DOI 10.1007/s11182-021-02433-z
Russian Physics Journal, Vol. 64, No. 6, October, 2021 (Russian Original No. 6, June, 2021)

MATHEMATICAL PROCESSING OF PHYSICS EXPERIMENTAL DATA

ROBUST PARAMETRIC GENERATORS OF RANDOM VARIABLES

V. A. Simakhin,[1] L. G. Shamanaeva,[2,3] and A. V. Maer[1]

UDC 519.2; 53.082.4519.21; 551.596

A method of constructing consistent and effective algorithms for robust parametric generators of random variables intended for solving problems of statistical simulation and constructing bootstrap procedures is considered. The consistency and efficiency of the standard and robust generators are analyzed in the presence of asymmetric and symmetric outliers. It is shown on real examples that the standard parametric generators of random variables are inconsistent for heterogeneous samples, and their use can significantly and unpredictably distort simulation results and decision-making procedures. It is demonstrated that in the presence of outliers, the efficiency of the robust generators can considerably exceed that of the standard parametric random variable generators, especially in the presence of asymmetric outliers.

Keywords: statistical simulation, bootstrap, robust random variable generators.

INTRODUCTION

With the advent of highly efficient and accessible computer facilities, a possibility arose of simulating complex systems by the Monte Carlo methods and solving problems that cannot be solved analytically [1–7]. The major elements of mathematical system models are programs – generators of pseudo-random numbers (processes) [1]. Nowadays any program environment pretending to be popular necessary contains the developed random number generator software (RNGS). In the developed software system, the parametric generators are widely used provided that the form of the distribution $F(x,\theta)$ of random variable X is known to within a parameter θ. If θ is unknown, its value θ_N is estimated from a sample $\boldsymbol{X}_N = (x_1,..., x_N)$, and the generator is adjusted by substitution of the estimate θ_N for the unknown θ value. The need to know properties of statistical procedures for the sample of finite size N using the Monte Carlo methods gives rise to the bootstrap based on the parametric generators of random variables and for unknown distributions – on the nonparametric ones [4, 5]. Their creation is based on parametric and nonparametric methods of mathematical statistics widely used to find unbiased, well-grounded, and effective estimates of the parameters based on the available experimental data.

At the same time, the problem of occurrence of anomalous observations (outliers) is well known to researchers. The standard methods of processing inhomogeneous samples can cause a considerable bias and low efficiency of estimates of the parameters. This significantly distorts the results of decision-making procedures [6–9]. The centuries-old approach developed by experimenters is based on deleting outliers from a data sample, but for its objective application, nonparametric criteria of outlier detection are required the determination of which is a complicated problem [10].

[1]Kurgan State University, Kurgan, Russia, e-mail: sva_full@mail.ru; alex_povt@mail.ru; [2]V. E. Zuev Institute of Atmospheric Optics of the Siberian Branch of the Russian Academy of Sciences, Tomsk, Russia, e-mail: sima@iao.ru; [3]National Research Tomsk State University, Tomsk, Russia. Translated from Izvestiya Vysshikh Uchebnykh Zavedenii, Fizika, No. 6, pp. 145–156, June, 2021. Original article submitted February 19, 2021.

The development of robust statistics reoriented the problem of sample censoring to the problem of synthesis of procedures stable against the effect of the outliers [8, 9, 11]. Robust parametric and nonparametric PNGS has poorly been developed, though practical simulation of real problems suggests otherwise. Indeed, let the random variable X with distribution $F(x,\boldsymbol{\theta})=(1-\varepsilon)G(x,\boldsymbol{\theta})+\varepsilon H(x)$ is observed. Of interest is the generator for the aprioristic distribution $G(x,\boldsymbol{\theta})$ with unknown parameter θ, unknown outlier fraction ε, and distribution $H(x)$.

The maximum likelihood estimate (MLE) for θ based on the sample $\boldsymbol{X}_N=(x_1,..., x_N)$ from the distribution $G(x,\theta)$ yields a biased and inconsistent estimate of the parameter θ for the distribution $G(x,\boldsymbol{\theta})$. In a nonparametric case, the empirical distribution function (EDF) $F_N(x)$ is the unbiased and well-grounded estimate of $F(x,\theta)$, but the biased and inconsistent estimate of $G(x,\theta)$. The application of the RNGS system generators to obtain such estimates may yield unpredictable results and conclusions during simulation of complex systems characterized by the presence of a great number of generators and complex interrelations between the elements of the system being simulated. For example, even a simple visual analysis of numerous temperature and pressure distributions when modeling the reliability and durability of a gas pipeline system shows the presence of numerous and various outliers [3].

The influence of the outliers can be traced on the example of processing of experimental Doppler acoustic radar (sodar) data on the spatiotemporal dynamics of the wind velocity in the atmospheric boundary layer. Distributions of the wind vector components at different altitudes are characterized by the presence of various outliers [7]. Data processing using traditional and robust nonparametric methods demonstrated that the efficiency of classical nonparametric methods of data processing can be extremely low (in some cases, only ≈5 %) compared to the robust methods [7]. The examples presented above show the importance and urgency of introducing robust generators with distributions depending on various aprioristic statistical uncertainties of the problem under study.

In the present work, algorithms of constructing parametric random number generators for aprioristic distribution $G(x,\boldsymbol{\theta})$ in the presence of the outliers are considered. Considering the level of aprioristic statistical uncertainty, this class of problems belongs to semiparametric problems of mathematical statistics [10]. To construct the robust generators, estimates of the parameters by the weighed maximum likelihood method (WMLM) are used with application to inhomogeneous experimental data [12]. The robust estimates based on the WMLM allow unbiased, well-grounded, and effective algorithms for parametric random number generators of inhomogeneous experimental data to be constructed. Examples of robust random number generators are given. Of doubtless interest is an analysis of the efficiency of the standard parametric generator software for inhomogeneous samples. In this regards, the consistency and efficiency of standard and robust generators in the presence of asymmetric and symmetric outliers are estimated. On specific examples it is proved that for inhomogeneous samples, standard parametric generators of random variables are inconsistent and can unpredictably distort simulation results. It is shown that robust generators have much higher efficiency than standard parametric generators of random variables, especially in the presence of asymmetric outliers.

1. PROBLEM FORMULATION. ALGORITHM FOR THE RANDOM NUMBER GENERATOR

Let $\boldsymbol{X}_N=(x_1,..., x_N)$ be a sample of independent identically distributed (IID) random variables with distribution function (DF) $F(x,\boldsymbol{\theta})\in P$, where P is the class of distributions satisfying to the conditions of regularity of the maximum likelihood method (MLM) [13, 14]. Below we consider $F(x,\boldsymbol{\theta})\in P_\theta$, where $P_\theta\subset P$ and $P_\theta=\{F(x,\theta\}$ is the class of the Tukey distributions (the supermodel)

$$F(x,\boldsymbol{\theta})=(1-\varepsilon)G(x,\boldsymbol{\theta})+\varepsilon H(x), \tag{1.1}$$

$G(x,\boldsymbol{\theta}) \in P_0$ is the aprioristic model of the distribution function, $H(x) \in P_H$ is the outlier distribution, $\varepsilon \geq 0$ is the outlier fraction, $\{H(x), \varepsilon\}$ is the information on the outliers, $f(x,\boldsymbol{\theta}), g(x,\boldsymbol{\theta})$, and $h(x)$ are the corresponding distribution densities, and $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)^T$ is the vector of the unknown parameters of the distribution. The Tukey supermodel is used as a convenient model of real distributions $F(x,\boldsymbol{\theta})$ that can be considered approximately coinciding with the aprioristic distribution $G(x,\boldsymbol{\theta})$ [8–10].

Let us consider the problem of construction of the random number generator for the aprioristic distribution $G(x,\boldsymbol{\theta})$ on the inhomogeneous sample $\boldsymbol{X}_N = (x_1, ..., x_N)$ from the distribution $F(x,\boldsymbol{\theta}) \in P_0$. For this purpose, we take advantage of the classical method of generating random variables using the inverse transformation $X = G^{-1}(U)$ of the form $x_i = G^{-1}(u_i)$, $i = 1, 2, \ldots$, where u_i, $i = 1, 2, \ldots$, are realizations of random variables uniform in [0, 1] [1]. Hence, generation of a sample value x_i from the distribution $G(x)$ is reduced to finding the quantile at level u_i of the distribution $G(x)$. As a well-grounded estimate X_{pN} of the quantile X_p at level p ($0 < p < 1$) of the distribution function $G(x)$, we take the solution of the empirical equation $\widehat{G}_N(X_{pN}) = p$:

$$X_{pN} = \widehat{G}_N^{-1}(p), \tag{1.2}$$

where $\widehat{G}_N(x)$ is the well-grounded unbiased (asymptotically unbiased) estimate of $G(x)$. To obtain the estimate X_{pN} from Eq. (1.2), various recurrent stochastic approximation algorithms are used. Below we restrict ourselves to the sufficiently general class of asymptotically normal estimates $\widehat{G}_N(x)$, that is, assume that the random variable

$$\sqrt{N}\left[\widehat{G}_N(x) - G(x)\right] \Leftrightarrow \Phi\left(0, \sigma^2\right)$$

obeys the asymptotically normal distribution $\Phi(0, \sigma^2)$ with zero average and variance

$$\sigma^2 = D\left(\widehat{G}_N(x)\right) = \sigma^2\left(G(x)\right). \tag{1.3}$$

Using the Lagrange increment, we represent Eq. (1.2) in the form

$$\left[X_{pN} - X_p\right] = \frac{1}{g(X_p)}\left[\widehat{G}_N(X_{pN}) - G(X_p)\right]. \tag{1.4}$$

Theorem 1.1. If $\sqrt{N}\left[\widehat{G}_N - G(x)\right] \Leftrightarrow \Phi(0, \sigma^2)$ and σ^2 from Eq. (1.3) is a continuous function of $G(x)$, $\sqrt{N}\left[X_{pN} - X_p\right] \Leftrightarrow \Phi\left(0, D(X_{pN})\right)$, where

$$D(X_{pN}) = \frac{\sigma^2}{g^2(X_p)}. \tag{1.5}$$

The proof follows from representation (1.4) and the continuity theorems ([13], Subsection 1.5; [14], Subsection 3.5). No Section references are given below.

Let $\boldsymbol{\theta}_N^* = \left(\theta_{1N}^*, ..., \theta_{kN}^*\right)^T$ be the robust, unbiased, and well-grounded estimate of $\boldsymbol{\theta} = \left(\theta_1, ..., \theta_k\right)^T$, and the quantile X_p of the distribution $G\left(x, \boldsymbol{\theta}\right)$ has a unique solution. Using the substitution method [10, 13], we take $G\left(x, \boldsymbol{\theta}_N^*\right)$ for $\widehat{G}_N\left(x\right)$ and define the robust estimate X_{pN} of the quantile X_p of the distribution function $G\left(x, \boldsymbol{\theta}\right) \in P_0$ in the form $X_{pN} = G^{-1}\left(p, \boldsymbol{\theta}_N^*\right)$. As a result, we obtain the following robust generator algorithm:

$$x_i = G^{-1}(u_i, \boldsymbol{\theta}_N^*), \quad i = 1, 2, ... \,. \tag{1.6}$$

The problems and methods of finding the estimate $\boldsymbol{\theta}_N^* = \left(\theta_{1N}^*, ..., \theta_{kN}^*\right)^T$ and hence the quantile estimate X_{pN} depend on aprioristic information about the set $P_\theta = \left\{F\left(x, \boldsymbol{\theta}\right)\right\}$, each element $F\left(x, \boldsymbol{\theta}\right)$ of which is defined by aprioristic information on $G\left(x, \boldsymbol{\theta}\right) \in P_0$ and information on the outliers $\left\{H(x), \varepsilon\right\}$. Both parametric and nonparametric models of the distribution functions and their superpositions – semiparametric and semi-nonparametric models – can participate in mathematical model (1.1). In the parametric models, the form of the distribution function is known to within a finite number of the unknown parameters, whereas in the nonparametric models, it is unknown. Based on aprioristic information on the set $P_\theta = \left\{F\left(x, \boldsymbol{\theta}\right)\right\}$, we consider the following problem formulations for supermodel (1.1).

1. Parametric problem

$P_\theta = P_{\theta P}$ is a parametric class, that is, P_{0P} and P_{HP} are parametric classes; information on $\left\{H(x), \varepsilon \neq 0\right\}$ is known.

2. Semiparametric problem

$P_\theta = P_{\theta P}$ is a semiparametric class, P_{0P} is a parametric class, and P_{HN} is a nonparametric class; information on $\left\{H(x), \varepsilon\right\}$ is unknown.

3. Nonparametric problem

$P_\theta = P_{\theta N}$ is a nonparametric class, and P_{0N} is a nonparametric class with some additional information on $G\left(x, \boldsymbol{\theta}\right)$ $\left\{H(x), \varepsilon = 0\right\}$.

4. Semi-nonparametric problem

$P_\theta = \overline{P}_{\theta N}$ is a semi-nonparametric class of the distribution functions, $\overline{P}_{0P}$ is a nonparametric class of the distribution functions with some additional information on $G\left(x, \boldsymbol{\theta}\right)$, and P_{HN} is a nonparametric class; information on $\left\{H(x), \varepsilon \neq 0\right\}$ is unknown.

2. ROBUST PARAMETRIC GENERATORS

Let us consider the problem of constructing a robust parametric generator for $G(x,\boldsymbol{\theta})$ on supermodel (1.1) $(\varepsilon \neq 0)$ at a parametric level of aprioristic uncertainty. Let $\boldsymbol{X}_N = (x_1,..., x_N)$ be a sample of independent identically distributed random variables with distribution functions $F(x,\boldsymbol{\theta}) \in P_{\theta P}$, $G(x,\boldsymbol{\theta}) \in P_{0\theta P}$, and $H(x) \in P_{HP}$: P_θ, $P_{\theta P}$, P_{0P}, and P_{HP} are the parametric classes, and information on $\{H(x), \varepsilon \neq 0\}$ is known. The parametric problems for a homogeneous sample ($\varepsilon = 0$) have been sufficiently studied. Exactly these generators are mainly present in the RNGS. As a robust estimate of the quantile X_p of the distribution $G(x,\boldsymbol{\theta})$, we take a solution of the empirical equation of the form

$$G\left(X_p,\boldsymbol{\theta}_N\right) = p \quad \text{or } X_{pN} = G^{-1}\left(p,\boldsymbol{\theta}_N\right), \tag{2.1}$$

where $\boldsymbol{\theta}_N$ is the unbiased (asymptotically unbiased), well-grounded estimate of $\boldsymbol{\theta}$ on the distribution $G(x,\boldsymbol{\theta}) \in P_{0P}$ and the sample $\boldsymbol{X}_N = (x_1,..., x_N)$ from the distribution $F(x,\boldsymbol{\theta}) \in P_{\theta P}$.

Theorem 2.1. Let $G(x,\boldsymbol{\theta}) \in P_{0P}$ be a continuous function of $\boldsymbol{\theta}$, and $\sqrt{N}\left(\boldsymbol{\theta}_N - \boldsymbol{\theta}\right)$ has the asymptotically normal distribution $\sqrt{N}\left(\boldsymbol{\theta}_N - \boldsymbol{\theta}\right) \Leftrightarrow \Phi\left(0, \boldsymbol{B}\right)$, where $\boldsymbol{B}$ is the covariance matrix in $\boldsymbol{\theta}_N = \left(\theta_{1N},..., \theta_{kN}\right)^T$; then

$$\sqrt{N}\left[G\left(x, \boldsymbol{\theta}_N\right) - G\left(x, \boldsymbol{\theta}\right)\right] \Leftrightarrow \Phi\left(0, \sigma^2\left[G\left(x, \boldsymbol{\theta}_N\right)\right]\right), \tag{2.2}$$

where

$$\sigma^2\left[G\left(x, \boldsymbol{\theta}_N\right)\right] = C^T \boldsymbol{B} C, \tag{2.3}$$

$$C^T = \left(\frac{\partial G(x, \theta)}{\partial \theta_1}, ..., \frac{\partial G(x, \theta)}{\partial \theta_k}\right). \tag{2.4}$$

Proof. Applying the finite increment theorem, we represent $\left[G\left(x, \boldsymbol{\theta}_N\right) - G\left(x, \boldsymbol{\theta}\right)\right]$ in the form

$$\left[G\left(x, \boldsymbol{\theta}_N\right) - G\left(x, \boldsymbol{\theta}\right)\right] = C^T\left(\boldsymbol{\theta}_N - \boldsymbol{\theta}\right). \tag{2.5}$$

The proof follows from the continuity theorems [13, 14]. Let us designate by E_L the operator of averaging over the distribution L.

Theorem 2.2. If $F(x,\boldsymbol{\theta}) \in P_{\theta P} \in P$, $G(x,\boldsymbol{\theta}) \in P_{0P}$, and $\boldsymbol{\theta}_N$ is the maximum likelihood estimate (MLE) for $F(x,\boldsymbol{\theta}) \in P_{\theta P} \in P$, then

1) $E_G\left(\boldsymbol{\theta}_N - \boldsymbol{\theta}\right) = 0$,

2) $\boldsymbol{\theta}_N$ is the well-grounded estimate of $\boldsymbol{\theta}$ for the distribution G,

3) $\sqrt{N}\left(\boldsymbol{\theta}_N - \boldsymbol{\theta}\right) \Leftrightarrow \Phi\left(0, \boldsymbol{B}\right)$ has the asymptotically normal distribution with zero vector of the averages and the covariance matrix $\boldsymbol{B} = \left\|b_{ij}\right\|$, where

$$b_{ij}=\int U_i(x,\theta)U_j(x,\theta)W(x,\theta)dG(x,\theta), \tag{2.6}$$

$$U_j(x,\boldsymbol{\theta}_N)=\frac{\partial}{\partial\theta_j}\ln g(x,\boldsymbol{\theta})\bigg|_{\theta=\boldsymbol{\theta}_N},\ W(x,\boldsymbol{\theta})=\frac{(1-\varepsilon)g(x,\boldsymbol{\theta})}{f(x,\boldsymbol{\theta})}.$$

Proof. Applying the multidimensional analog of the maximum likelihood method (MLM) to the sample $\boldsymbol{X}_N=(x_1,..., x_N)$ from $F(x,\boldsymbol{\theta})$, we find the MLE $\boldsymbol{\theta}_N$ for $F(x,\boldsymbol{\theta})$ from the system of evaluation equations [13, 14]

$$\int\frac{\partial}{\partial\theta_j}\ln f(x,\boldsymbol{\theta})\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_N}dF_N(x)=\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial\theta_j}\ln f(x,\boldsymbol{\theta})\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_N}=0,\ \ j=\overline{1,k}. \tag{2.7}$$

In this case, $\sqrt{N}\left(\boldsymbol{\theta}_N-\boldsymbol{\theta}\right)\Leftrightarrow\Phi\left(0,\ \boldsymbol{B}\right)$ obeys the asymptotically normal distribution with zero vector of averages $E_F\left(\boldsymbol{\theta}_{jN}-\boldsymbol{\theta}\right)=0$ and the covariance matrix $\boldsymbol{B}=\left\|b_{ij}\right\|$, where

$$b_{ij}=E_F\left[\frac{1}{f^2(x,\theta)}\frac{\partial f(x,\theta)}{\partial\theta_i}\frac{\partial f(x,\theta)}{\partial\theta_j}\right]. \tag{2.8}$$

From the proof of the MLM and representations for the MLE [13, 14] it follows that

$$\left(\theta_{jN}-\theta_j\right)=\left[\int\frac{\partial^2}{\partial\theta_j^2}\ln f(x,\boldsymbol{\theta})dF(x,\boldsymbol{\theta})\right]^{-1}\left[\int\frac{\partial}{\partial\theta_j}\ln f(x,\boldsymbol{\theta})dF_N(x,\boldsymbol{\theta})\right],$$

and the condition of the unbiased MLE $E_F\left(\boldsymbol{\theta}_{jN}-\boldsymbol{\theta}\right)=0$ is reduced to

$$E_F\left[\int\frac{\partial}{\partial\theta_j}\ln f(x,\boldsymbol{\theta})dF_N(x,\boldsymbol{\theta})\right]=\int\frac{\partial}{\partial\theta_j}\ln f(x,\boldsymbol{\theta})dF(x,\boldsymbol{\theta})=0,\ j=\overline{1,k}. \tag{2.9}$$

Let us consider Eq. (2.7) and expressions (2.8) and (2.9) for $F(x,\boldsymbol{\theta})$ defined by supermodel (1.1), $F(x,\boldsymbol{\theta})\in P_\theta\subset P$, and $G(x,\boldsymbol{\theta})\in P_{0P}\subset P$. In this case, system of evaluation equations (2.7) is transformed to the form

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial\theta_j}\ln f(x_i,\boldsymbol{\theta}_N)=\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial\theta_j}U(x_i,\boldsymbol{\theta}_N)W(x_i,\boldsymbol{\theta}_N)=0,\ \ j=\overline{1,k}, \tag{2.10}$$

where $U_j(x,\boldsymbol{\theta}_N)=\frac{\partial}{\partial\theta_j}\ln g(x,\boldsymbol{\theta})\Big|_{\theta=\boldsymbol{\theta}_N}$ is the contribution function for $G(x,\boldsymbol{\theta})$ and $W(x,\boldsymbol{\theta})$ is the weight function:

$$W(x,\boldsymbol{\theta})=\frac{(1-\varepsilon)g(x,\boldsymbol{\theta})}{f(x,\boldsymbol{\theta})}=\left[1+\frac{\varepsilon\cdot h(x)}{(1-\varepsilon)g(x,\boldsymbol{\theta})}\right]^{-1}. \tag{2.11}$$

Since the conditions of MLM regularity are fulfilled, condition (2.9) for the unbiased estimate can be transformed to the form

$$\int \frac{\partial}{\partial\theta_j} \ln f(x,\boldsymbol{\theta}) dF(x,\boldsymbol{\theta}) = \int \frac{\partial}{\partial\theta_j} \ln g(x,\boldsymbol{\theta}) \frac{g(x,\boldsymbol{\theta})}{f(x,\boldsymbol{\theta})} dF(x,\boldsymbol{\theta})$$

$$= \int \frac{\partial}{\partial\theta_j} \ln g(x,\boldsymbol{\theta}) dG(x,\boldsymbol{\theta}) = \frac{\partial}{\partial\theta_j} \int g(x,\boldsymbol{\theta}) dx = 0.$$

As a result, we obtain

$$E_F\left[\int \frac{\partial}{\partial\theta_j} \ln f(x,\boldsymbol{\theta}) dF_N(x,\boldsymbol{\theta})\right] = E_G\left[\int \frac{\partial}{\partial\theta_j} \ln g(x,\boldsymbol{\theta}) dF_N(x,\boldsymbol{\theta})\right] = 0 .$$

Hence, the MLE θ_N is the unbiased estimate for the distribution $G(x,\theta) \in P_{0P} \subset P$.

Let us consider expressions (2.8):

$$b_{ij} = E_F\left[\frac{1}{f^2(x,\theta)} \frac{\partial f(x,\theta)}{\partial\theta_i} \frac{\partial f(x,\theta)}{\partial\theta_j}\right] = \int \frac{1}{f^2(x,\theta)} \frac{\partial g(x,\theta)}{\partial\theta_i} \frac{\partial g(x,\theta)}{\partial\theta_j} dF(x,\theta)$$

$$= \int \frac{\partial \ln g(x,\theta)}{\partial\theta_i} \frac{\partial \ln g(x,\theta)}{\partial\theta_j} W(x,\theta) dG(x,\theta) = \int U_i(x,\theta) U_j(x,\theta) W(x,\theta) dG(x,\theta) .$$

As a result, we obtain

$$b_{ij} = \int U_i(x,\theta) U_j(x,\theta) W(x,\theta) dG(x,\theta) .$$

According to formulas (2.6), (2.10), and (2.11), we have the weighed maximum likelihood method for the aprioristic distribution $G(x,\theta)$. The theorem has been proved.

Computationally, the weight functions of the form [14]

$$W(x,\boldsymbol{\theta}) = \frac{(1-\varepsilon) g(x,\boldsymbol{\theta}) C(g(x,\boldsymbol{\theta}) - \delta)}{f(x,\boldsymbol{\theta})} \tag{2.12}$$

are more convenient, especially for the distribution functions with unbounded carrier and *U*-shaped distributions, where $\delta \approx 0$ is the constant defined by the user and $C(x) = \begin{cases} 1, x \ge 0, \\ 0, x < 0. \end{cases}$ Algorithms (2.1) and (2.10)–(2.12) of constructing generators based on the WMLM allow robust effective estimates of the quantiles of distributions $G(x,\theta)$ to be obtained for the Tukey supermodel given by Eq. (1.1) in the presence of outliers with known parametric distribution. As a result, we obtain the following algorithm for the effective parametric random number generator with the distribution $G(x,\theta)$ for the inhomogeneous sample from the distribution $F(x,\theta)$:

$$x_i = G^{-1}(u_i, \theta_N^*), \quad i = 1, 2, \ldots , \tag{2.13}$$

where u_i, $i = 1, 2, \ldots$, are random variables uniform in [0, 1], and the robust effective estimates $\theta_N^* = \left(\theta_{1N}^*, \ldots, \theta_{kN}^*\right)^T$ are determined from the system of evaluation equations

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial\theta_j}U(x_i,\theta_N^*)W(x_i,\theta_N^*) = 0, \quad j = \overline{1,k}\,, \tag{2.14}$$

$$U_j(x,\theta_N^*) = \frac{\partial}{\partial\theta_j}\ln g(x,\theta_N^*)\Big|_{\theta=\theta_N^*}\,, \tag{2.15}$$

$$W(x,\theta_N^*) = \frac{(1-\varepsilon)g(x,\theta_N^*)}{f(x,\theta_N^*)}\,. \tag{2.16}$$

3. EXAMPLES OF THE ROBUST PARAMETRIC GENERATORS

Let us consider the algorithms of robust parametric generators for a number of typical distributions $G(x,\theta)$. Of doubtless interest is the efficiency of the standard parametric generators from the RNGS system for an inhomogeneous sample from the distribution $F(x,\theta) \in P_\theta$. Investigation of the generators is reduced to consideration of various estimates of distribution quantiles since the algorithms of generators are defined through the quantile estimates. We now designate by $X_p^{(0)}$ and $X_p^{(f)}$ the quantiles at level p of the distribution function $G(x,\theta^{(0)})$ and by $\theta^{(0)}$ the true parameters. Let there are two quantile estimates: the standard estimate $X_{pN}^{(1)}$ synthesized at $\varepsilon = 0$ and the robust estimate $X_{pN}^{(2)}$. The estimates are compared by the criterion of the asymptotic relative efficiency [9, 13]:

$$Eff\,(X_{pN}^{(2)}, X_{pN}^{(1)}) = \frac{V^{(2)}}{V^{(1)}}\,, \tag{3.1}$$

where $V^{(i)} = D^{(i)} + \left(b^{(i)}\right)^2$ is the root mean square error (RMSE) of the quantile estimate $X_{pN}^{(i)}$, $b^{(i)} = (E_L X_{pN}^{(i)} - X_p^{(0)})$ is the bias, and $D^{(i)} = N \cdot E_L(X_{pN}^{(i)} - E_L X_{pN}^{(i)})^2$ is the variance of the estimate $X_{pN}^{(i)}$ on the distribution L. For $\varepsilon = 0$, the standard estimate $X_{pN}^{(1)}$ and the robust estimate $X_{pN}^{(2)}$ coincide, and $Eff\,(X_{pN}^{(2)}, X_{pN}^{(1)}) = 1$. The variant with $\varepsilon \neq 0$ is considered below.

Example 3.1. It is required to construct the random number generator for the exponential random variable X with distribution

$$G(x,\theta) = 1 - \exp\left(-\frac{x}{\theta}\right), \quad x \ge 0; \quad g(x,\theta) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right), \quad x \ge 0\,, \tag{3.2}$$

where θ is the unknown parameter. The standard algorithm ($\varepsilon = 0$) for the generator of exponential random variable [1] is

$$x_i = G^{-1}\left(u_i, \theta_N^{(1)}\right) = -\theta_N^{(1)}\ln\left(1 - u_i\right), \quad i = 1, \ldots, N\,, \tag{3.3}$$

where the MLE $\theta_N^{(1)}$ is

$$\theta_N^{(1)} = \frac{1}{N}\sum_{i=1}^{N} x_i \, . \tag{3.4}$$

For the robust generator, according to formulas (2.13)–(2.16), we obtain the algorithm

$$x_i = -\theta_N^{(2)} \cdot \ln\left(1-u_i\right), \quad i = 1, \ldots, N \, , \tag{3.5}$$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \theta_N^{(2)})\frac{g(x_i,\theta_N^{(2)})}{f(x_i,\theta_N^{(2)})} = 0 \, . \tag{3.6}$$

Let us designate $0 < u_i = p < 1$. For distribution (3.2) at $\varepsilon = 0$ we subsequently obtain

$$R = -\ln(1-p)\,,\ E_G X_{pN}^{(1)} = \theta^{(0)} R\,,\ X_p^{(0)} = R\theta^{(0)}\,,\ b^{(1)} = 0\,,\ D^{(0)} = R^2\left(\theta^{(0)}\right)^2,\ V^{(1)} = D^{(0)} = R^2\left(\theta^{(0)}\right)^2 .$$

For the distribution $F\left(x,\theta\right)$ at $\varepsilon \neq 0$, we subsequently obtain

$$E_F\theta_N^{(1)} = (1-\varepsilon)\theta^{(0)} + \varepsilon E_H X\,,\ E_F X_{pN}^{(1)} = R\left[(1-\varepsilon)\theta^{(0)} + \varepsilon E_H X\right],\ b^{(1)} = R\varepsilon\left[E_H X - \theta^{(0)}\right],$$

$$D^{(1)} = E_F(X_{pN}^{(1)} - E_F X_{pN}^{(1)})^2 = R^2\left\{(1-\varepsilon)[(\theta^{(0)})^2 + (b_{10})^2] + \varepsilon\left[(b_{11})^2 + E_H(X - E_H X)^2\right]\right\},$$

$$b_{10} = E_F X - \theta^{(0)}, \quad b_{11} = E_F X - E_H X\,,\ V^{(1)} = D^{(1)} + (b^{(1)})^2 .$$

Let us consider the robust estimate of the quantile $X_{pN}^{(2)}$. From Eq. (3.6) we determine the robust estimate of the parameter $\theta_N^{(2)}$. From theorem 2.2 it follows that $E_F\theta_N^{(2)} = E_G\theta_N^{(2)} = \theta^{(0)}$, $S = E_F[\theta_N^{(2)} - E_F\theta_N^{(2)}]^2$ $= E_F[\theta_N^{(2)} - \theta^{(0)}]^2 = \int [x-\theta^{(0)}]^2 \frac{g^2(x,\theta^{(0)})}{f(x,\theta^{(0)})}dx$. Consequently, $E_F X_{pN}^{(2)} = R\theta^{(0)}$, $b^{(2)} = 0$, $D^{(2)} = R^2 S$, and $V^{(2)} = R^2 S$. Let us consider an example of the supermodel for $\theta^{(0)} = 1$, $\varepsilon = 0.1$, and $\sigma = 0.2$:

$$f\left(x,\,\theta\right) = (1-\varepsilon)\exp\left(-x\right) + \frac{\varepsilon}{\sqrt{2\pi}\cdot\sigma} exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right).$$

Results of investigations are given in Tables 3.1–3.4.

Example 3.2. Let *X* be the random variable with the Weibull distribution:

$$G\left(x,\,\theta\right) = 1 - \exp\left(-\frac{x}{\theta}\right)^{\alpha}, \quad x \geq 0, \quad \theta > 0, \quad \alpha > 0\,, \tag{3.7}$$

TABLE 3.1. Values of Quantiles $X_p^{(0)}$ and $X_p^{(f)}$ and Estimates $E_f X_{pN}^{(2)}$ and $E_f X_{pN}^{(1)}$ Depending on p for $\theta^{(0)} = 1, \varepsilon = 0.1, \sigma = 0.2$, and $\mu = 5$

p / X_p	$X_p^{(0)}$	$E_f X_{pN}^{(2)}$	$E_f X_{pN}^{(1)}$	$X_p^{(f)}$
p = 0.55	0.80	0.80	1.28	0.94
p = 0.75	1.39	1.39	2.23	1.79
p = 0.95	3.00	3.00	4.81	4.92

TABLE 3.2. Efficiency $Eff(X_{pN}^{(2)}, X_{pN}^{(1)})$ Depending on p and μ for $\theta^{(0)} = 1$, $\varepsilon = 0.1$, and $\sigma = 0.2$

X_p / μ	$X_p^{(0)}$	$\mu = 3$	$\mu = 5$	$\mu = 7$	$\mu = 10$
$X_{0.55}$	0.90	0.66	0.31	0.16	0.08
$X_{0.75}$	1.39	0.66	0.31	0.16	0.09
$X_{0.95}$	3.00	0.66	0.31	0.16	0.09

TABLE 3.3. Values of $E_f X_{0.75N}^{(1)}$, Bias $b^{(1)}$, and RMSE $X_{0.75N}^{(1)}$ of the Quantile $X_{0.75}^{(0)} = 1.39$ Depending on μ for $\theta^{(0)} = 1, \varepsilon = 0.1$, and $\sigma = 0.2$

μ	$\mu = 3$	$\mu = 5$	$\mu = 7$	$\mu = 10$
$E_f X_{0.75N}^{(1)}$	1.84	2.23	2.62	3.21
$b^{(1)}$	0.45	0.84	1.23	1.82
RMSE	2.86	2.90	11.74	24.11

TABLE 3.4. Values of $E_f X_{0.75N}^{(1)}$, Bias $b^{(1)}$, and RMSE $X_{0.75N}^{(1)}$ of the Quantile $X_{0.75}^{(0)} = 1.39$ Depending on σ for $\theta^{(0)} = 1$ and $\mu = 10$

ε	$\varepsilon = 0$	$\varepsilon = 0.05$	$\varepsilon = 0.1$	$\varepsilon = 0.2$
$E_f X_{0.75N}^{(1)}$	1.39	2. 30	3.21	5.03
$b^{(1)}$	0	0.91	1.82	3.64
RMSE	1.92	12.91	24.11	45.54

$$g(x, \theta) = \frac{\alpha x^{\alpha-1}}{\theta^{\alpha}} \exp\left(-\frac{x}{\theta}\right)^{\alpha}, \quad x \geq 0, \quad \theta > 0, \quad \alpha > 0,$$

where θ is unknown parameter. The standard algorithm ($\varepsilon = 0$) for the generator with the Weibull distribution [1, 2] has the form

$$x_i = G^{-1}\left(u_i,\, \theta_N^{(1)}\right) = \theta_N^{(1)} R^{1/\alpha}, \quad i = 1, \ldots, N, \tag{3.8}$$

where $R = -\ln(1-p)$ and the MLE $\theta_N^{(1)}$ is determined as

$$\theta_N^{(1)} = \left[\frac{1}{N}\sum_{i=1}^{N} x_i^{\alpha}\right]^{1/\alpha}. \tag{3.9}$$

According to formulas (2.13)–(2.16), for the robust generator ($\varepsilon \geq 0$) we have

$$x_i = -\theta_N^{(2)} R^{1/\alpha}, \quad i = 1, \ldots, N, \tag{3.10}$$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i^{\alpha} - \theta_N^{(*)})\frac{g(x_i, \theta_N^{(*)})}{f(x_i, \theta_N^{(*)})} = 0, \tag{3.11}$$

$$\theta_N^{(2)} = \left[\theta_N^{(*)}\right]^{1/\alpha}. \tag{3.12}$$

Let us designate by $T_N = \frac{1}{N}\sum_{i=1}^{N} x_i^{\alpha}$, $0 < u_i = p < 1$. The statistics $\sqrt{N}(T_N - E_L(T_N)) \Rightarrow \Phi(0, D_L(T_N))$ for the distribution L is the well-grounded asymptotically normal nonparametric estimate of the initial α-order moment with zero average and variance [14]

$$D_L(T_N) = E_L x^{2\alpha} - \left(E_L x^{\alpha}\right)^2.$$

Proceeding from the continuity theorem [13, 14], we obtain for the statistic $W_N = \varphi\left(T_N\right)$

$$\sqrt{N}(W_N - E_L(W_N)) \Rightarrow \Phi(0, D_L(W_N)), \tag{3.13}$$

where

$$E_L(W_N) = \varphi(E_L(T_N)),\; D_L(W_N) = \left[\frac{d\varphi(t)}{dt}\Big|_{t=E_L(T_N)}\right]^2 D_L(T_N).$$

For $\varepsilon = 0$, $L = F = G$, and $\varphi(t) = (t)^{1/\alpha}$, taking into account formula (3.13) and the fact that for the Weibull distribution, $E_g x^k = \theta^k \Gamma\left(1 + \frac{k}{\alpha}\right)$, where $\Gamma(x)$ is the gamma function, we obtain

$$E_G \theta_N^{(1)} = E\left\{\left[T_N\right]^{1/\alpha}\right\} = \left[E_G x^{\alpha}\right]^{1/\alpha} = \theta,\; X_p^{(0)} = R^{1/\alpha}\theta^{(0)},\; E_G X_{pN}^{(1)} = R^{1/\alpha}\theta,$$

$$b^{(1)} = 0,\; N[E_G(T_N - E_G(T_N))^2] = \left[E_G x^{2\alpha} - \left(E_G x^{\alpha}\right)^2\right] = \theta^{2\alpha},$$

$$\frac{d\varphi(t)}{dt}\Big|_{t=E_L(T_N)} = \frac{1}{\alpha}(\theta^{\alpha})^{\frac{1-\alpha}{\alpha}} = \frac{1}{\alpha}\theta^{1-\alpha},\ D^{(0)} = [E_G(\theta_N^{(1)} - \theta)^2](R)^{2/\alpha} = \frac{1}{\alpha^2}\theta^2(R)^{2/\alpha},$$

$$V^{(0)} = D^{(0)} = \frac{1}{\alpha^2}(\theta^{(0)})^2(R)^{2/\alpha}.$$

Taking into account formula (3.13) for the distribution $F(x,\theta)$, for $\varepsilon \neq 0$ we obtain

$$E_F(x^{\alpha}) = (1-\varepsilon)(\theta^{(0)})^{\alpha} + \varepsilon E_H(X)^{\alpha},\ E_F\theta_N^{(1)} = E_F\left\{[T_N]^{1/\alpha}\right\} = \left[E_F(x^{\alpha})\right]^{1/\alpha},$$

$$E_F X_{pN}^{(1)} = R^{1/\alpha}E_F(\theta_{pN}^{(1)}) = R^{1/\alpha}\left[E_F(x^{\alpha})\right]^{1/\alpha},\ b^{(1)} = E_F X_{pN}^{(1)} - X_p^{(0)},$$

$$D^{(1)} = E_F\left[X_{pN}^{(1)} - E_F X_{pN}^{(1)}\right]^2 = \frac{R^{2/\alpha}}{\alpha^2}\left[E_F(x^{\alpha})\right]^{\frac{2(1-\alpha)}{\alpha}}\left[E_F x^{2\alpha} - \left(E_F x^{\alpha}\right)^2\right],\ V^{(1)} = D^{(1)} + (b^{(1)})^2.$$

For $\varepsilon = 0$, we obtain $b^{(1)} = b^{(0)} = 0$, $D^{(1)} = D^{(0)}$, and $V^{(1)} = V^{(0)}$. To obtain the robust estimate of the quantile $X_{pN}^{(2)}$, the robust estimate of the parameter $\theta_N^{(*)}$ is found from Eq. (3.11), and from theorem 2.2 with allowance for formula (3.13) it follows that

$$E_F\theta_N^{(*)} = E_G T_N = (\theta)^{\alpha},\ S = E_F[\theta_N^{(*)} - E_F\theta_N^{(*)}]^2 = \int[x^{\alpha} - \theta^{\alpha}]^2\frac{g^2(x,\theta)}{f(x,\theta)}dx,\ E_F X_{pN}^{(2)} = (R)^{1/\alpha}\theta^{(0)},$$

$$b^{(2)} = 0,\ D^{(2)} = \frac{R^{2/\alpha}}{\alpha^2}\left[E_F(x^{\alpha})\right]^{\frac{2(1-\alpha)}{\alpha}}S = \frac{R^{2/\alpha}}{\alpha^2}[\theta]^{2(1-\alpha)}S,\ V^{(2)} = \frac{R^{2/\alpha}}{\alpha^2}[\theta]^{2(1-\alpha)}S.$$

Let us consider an example of the supermodel for $\alpha = 2$, $\theta^{(0)} = 1$, and $\varepsilon = 0.1$:

$$f(x,\theta) = (1-\varepsilon)2x\cdot\exp\left(-x^2\right) + \frac{\varepsilon}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right).$$

The results obtained are given in Tables 3.5–3.7.

Example 3.3. Let X be the random variable with normal distribution $\Phi(\theta_1,\theta_2)$ and

$$G(x,\theta) = \frac{1}{\sqrt{2\pi}\cdot\theta_2}\int_{-\infty}^{x}\exp\left(-\frac{(t-\theta_1)^2}{2\theta_2^2}\right)dt. \tag{3.14}$$

The standard algorithm ($\varepsilon = 0$) for the generator of random variables with the normal distribution $\Phi(\theta_1,\theta_2)$ is

$$x_i = G^{-1}\left(u_{pi},\theta_{1N}^{(1)},\theta_{2N}^{(1)}\right) = \theta_{1N}^{(1)} + u_{pi}\theta_{2N}^{(1)},\quad i = 1, ..., N, \tag{3.15}$$

TABLE 3.5. Values of the Efficiency $Eff(X_{pN}^{(2)}, X_{pN}^{(1)})$ Depending on p and μ for $\alpha = 2$, $\theta^{(0)} = 1$, and $\varepsilon = 0.1$

X_p/μ	$X_p^{(0)}$	$\mu = 3$	$\mu = 5$	$\mu = 7$	$\mu = 10$
$X_{0.1}$	0.33	0.44	0.16	0.07	0.03
$X_{0.3}$	0.60	0.44	0.16	0.07	0.03
$X_{0.5}$	0.83	0.44	0.16	0.07	0.03
$X_{0.75}$	1.18	0.44	0.16	0.07	0.03

TABLE 3.6. Values of $E_f X_{0.3N}^{(1)}$, Bias $b^{(1)}$, and RMSE of the Estimate $X_{0.3N}^{(1)}$ of the Quantile $X_{0.3}^{(0)} = 0.60$ Depending on μ for $\alpha = 2, \theta^{(0)} = 1$, and $\varepsilon = 0.1$

μ	$\mu = 3$	$\mu = 5$	$\mu = 7$	$\mu = 10$
$E_f X_{0.3N}^{(1)}$	0.80	1.10	1.44	1.97
$b^{(1)}$	0.21	0.51	0.84	1.38
RMSE	0.20	0.17	1.23	2.86

TABLE 3.7. Values of $E_f X_{0.3N}^{(1)}$, Bias $b^{(1)}$, and RMSE of the estimate $X_{0.3N}^{(1)}$ of the quantile $X_{0.3}^{(0)} = 0.60$ Depending on ε for $\theta^{(0)} = 1, \mu = 10$, and $\sigma = 0.2$

ε	$\varepsilon = 0$	$\varepsilon = 0.05$	$\varepsilon = 0.1$	$\varepsilon = 0.2$
$E_f X_{0.3N}^{(1)}$	0.60	1. 46	1.97	2.72
$b^{(1)}$	0	0.86	1.38	2.13
RMSE	0.09	1.27	2.86	6.38

where u_{pi}, $i = 1, 2, \ldots$, are random numbers from the standard normal distribution $\Phi(0, 1)$ and

$$\theta_{1N}^{(1)} = \frac{1}{N}\sum_{i=1}^{N} x_i \, , \; \theta_{2N}^{(1)} = \sqrt{S_N^2} \, , \; S_N^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \theta_{1N}^{(1)}\right)^2 . \tag{3.16}$$

For the robust generator ($\varepsilon \geq 0$), according to formulas (2.13)–(2.16), we obtain the algorithm

$$x_i = G^{-1}\left(u_{pi}, \theta_{1N}^{(2)}, \theta_{2N}^{(2)}\right) = \theta_N^{(1)} + u_{pi}\theta_{2N}^{(2)}, \quad i = 1, \ldots, N \, , \tag{3.17}$$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \theta_{1N}^{(2)})\frac{g(x_i, \theta_{1N}^{(2)}, \theta_{2N}^{(2)})}{f(x_i, \theta_{1N}^{(2)}, \theta_{2N}^{(2)})} = 0 \, , \tag{3.18}$$

TABLE 3.8. Efficiency $Eff(X_{pN}^{(2)}, X_{pN}^{(1)})$ Depending on p and μ in the Presence of Asymmetric Outliers for $\theta^{(0)} = 0, \theta_2 = 1$, and $\varepsilon = 0.1$

X_p/μ	$X_p^{(0)}$	$\mu = 3$	$\mu = 5$	$\mu = 7$	$\mu = 10$
$X_{0.2}$	–0.84	0.58	0.32	0.18	0.09
$X_{0.5}$	0.00	0.55	0.29	0.17	0.09
$X_{0.9}$	1.28	0.51	0.23	0.12	0.09

TABLE 3.9. Efficiency $Eff(X_{pN}^{(2)}, X_{pN}^{(1)})$ Depending on p and σ in the Presence of Symmetric Outliers for $\mu = 0, \theta_1 = 0, \theta_2 = 1$, and $\varepsilon = 0.1$

X_p/σ	$X_p^{(0)}$	$\sigma = 3$	$\sigma = 6$	$\sigma = 9$
$X_{0.2}$	–0.84	0.54	0.24	0.17
$X_{0.5}$	0.00	0.56	0.27	0.19
$X_{0.9}$	1,28	0,53	0.22	0.14

TABLE 3.10. Values of $E_f X_{0,2N}^{(1)}$, Bias $b^{(1)}$, and RMSE of the Estimate $X_{0.2N}^{(1)}$ of the Quantile $X_{0.2}^{(0)} = -0.84$ Depending on μ for $\theta_1 = 0, \theta_2 = 1$, and $\varepsilon = 0.1$

μ	$\mu = 3$	$\mu = 5$	$\mu = 7$	$\mu = 10$
$E_f X_{0.2N}^{(1)}$	–0.80	–1.00	–1.24	–1.65
$b^{(1)}$	0.04	–0.15	–0.40	–0.81
RMSE	2.32	4.29	7.40	14.06

$$\frac{1}{N}\sum_{i=1}^{N}\left[(x_i - \theta_{1N}^{(2)})^2 - (\theta_{2N}^{(2)})^2\right]\frac{g(x_i, \theta_{1N}^{(2)}, \theta_{2N}^{(2)})}{f(x_i, \theta_{1N}^{(2)}, \theta_{2N}^{(2)})} = 0\,. \tag{3.19}$$

Let $X_{pN} = \theta_{1N}^{(1)} + U_p\theta_{2N}^{(1)}$ be the estimate of the quantile X_p for the distribution $\Phi(\theta_1, \theta_2)$ and U_p be the quantile at level p of the standard normal distribution $\Phi(0, 1)$. Let us consider an example of the supermodel for $\theta_1 = 0, \theta_2 = 1$, and $\varepsilon = 0.1$:

$$f(x, \theta) = \frac{(1-\varepsilon)}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) + \frac{\varepsilon}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right).$$

The results obtained are given in Tables 3.8–3.10.

CONCLUSIONS

1. In the present work, the robust effective parametric generators of random variables have been synthesized (algorithm (2.13)–(2.16)).

2. The efficiencies of the robust and standard generators of the RNGS system in the absence of the outliers coincide.

3. The robust generators are efficient, well-grounded, and asymptotically unbiased for the aprioristic distribution $G(x, \theta)$ in the presence of the outliers.

4. In the presence of the outliers, **it is impossible to use algorithms of standard generators** since they generate random numbers $X_{pN}^{(1)}$ from **uncertain distribution**. The random numbers $X_{pN}^{(1)}$ coincide neither with $X_{\delta}^{(0)}$ from the aprioristic distribution $G(x, \theta)$, nor with $X_{p}^{(f)}$ from the real distribution $F(x, \theta)$ for any arbitrary *p* (see Table 3.1).

5. In the presence of the outliers, the efficiency of the standard generators from the RNGS system compared to the robust ones $Eff(X_{pN}^{(2)}, X_{pN}^{(1)})$ is independent of *p*, but significantly depends on the bias parameter μ for the asymmetric outliers (see Tables 3.2, 3.5, and 3.8) and on the scale parameter σ for the symmetric outliers (Table 3.9).

6. In the presence of the asymmetric and symmetric outliers, the efficiency of the standard generators compared to robust ones $Eff(X_{pN}^{(2)}, X_{pN}^{(1)})$ is independent of *p* and tends to zero with increase in the bias parameter μ (Tables 3.2, 3.5, and 3.8) and the scale parameter σ (Table 3.9).

7. In the presence of the asymmetric and symmetric outliers, the standard generators become inconsistent with the aprioristic distribution $G(x, \theta)$ for any arbitrary *p*, the bias and the RMSE (Tables 3.3, 3.4, 3.6, 3.7, and 3.10) increase with the bias parameters μ (Tables 3.3, 3.4, 3.6, and 3.7) and the scale parameters σ (Table 3.10).

REFERENCES

1. S. Ermakov and G. A. Mikhailov, Statistical Simulation [in Russian], Nauka, Moscow (1982).
2. I. V. Pavlov, Statistical Methods of Estimating the Reliability of Complex Systems Using Test Results [in Russian], Radio i Svyaz’, Moscow (1982).
3. V. N. Syzrantsev, Ya. P. Nevelev, and S. L. Golofast, Calculation of Strength Reliability of Products Based on Methods of Nonparametric Statistics [in Russian], Nauka, Novosibirsk (2008).
4. B. Èfron, Nonconventional Methods of Multidimensional Statistical Analysis [in Russian], Finansy i Statistika, Moscow (1998).
5. A. S. Davison and D. V. Hinkley, Bootstrap Methods and Their Application, Cambridge University Press, Cambridge (1997).
6. V. N. Syzrantsev, S. L. Golofast, and A. V. Maer, Oil and Gas Studies, No. 5, 87–92 (2012).
7. V. A. Simakhin, O. S. Cherepanov, and L. G. Shamanaeva, Russ. Phys. J., **58**, No. 12, 1868–1874 (2016).
8. F. Hampel, E. Ronchetti, P. Rausseu, and V. Shtael, Robustness in Statistics [Russian translation], Mir, Moscow (1989).
9. V. P. Shulenin, Robust Methods of Mathematical Statistics [in Russian], Publishing House of Scientific and Technological Literature, Tomsk (2016).
10. A. I. Orlov, Zavodsk. Lab., No. 7, 40–42 (1992).
11. V. A. Simakhin, Adaptive Estimates [in Russian], Publishing House of Kurgan State University, Kurgan (2019).
12. V. A. Simakhin, L. G. Shamanaeva, and A. E. Avdyushina, Russ. Phys. J., **63**, No. 9, 1510–1518 (2020).
13. A. A. Borovkov, Mathematical Statistics [in Russian], Nauka, Novosibirsk (1997).
14. G. I. Ivchenko and Yu. I. Medvedev, Introduction to Mathematical Statistics [in Russian], LKI Publishing House, Moscow (2010).