Modeling of truncated probability distributions

M S Tokmachev[1]

[1]Novgorod State University, ul. B. St. Petersburgskaya, 41 173003 Veliky Novgorod, Russia

Abstract. A method for modeling truncated probability distributions based on the transformation of some basic distribution is developed. One of the parameters of the simulated distributions is a quantile of arbitrary order. The cases of the initial multiplication of the basic density function by linear and quadratic factors are considered. In the case of the basic normal law with linear multipliers, expressions for the density and the first moments are generally presented. Using different base distributions and varying parameters leads to a wide class of new distributions. The results of the work can be useful in the processing of statistics with a non-standard structure.

1. Introduction

In classical probability distributions, often, the theoretical values of random variables are distributed on infinite intervals. The corresponding sampling values are distributed over finite intervals. And then the so-called "tails of distributions" for the distribution of the general totality are required to be taken into account in some way. One such method is the use of truncated distributions. The truncation of the distribution over a smaller interval in one way or another leads to a rearrangement of the distribution. The variety of truncation methods serves as a source of modeling of new probability distributions.

The truncation essence is the translation of the probabilities, determined by the "distribution tails", to the probability concentration interval after the truncation operation. Geometrically for the density curve is the addition of the figure to the rest of the curvilinear trapezoid by the area equal to the area of the cut off part. Thus, for the density, the normalization condition is guaranteed. And what will be the shape of this added figure - a question at the discretion of the researcher. The shape of a given area can be represented in various ways.

Classical and most simple when truncating the distribution to some interval $[a;b]$ is (see, for example, [1, 2]) the method of multiplying the density function $f(x)$ on the correcting factor, which guarantees the validity of the normalization condition:

$$g(x)=\begin{cases}\dfrac{f(x)}{\int_a^b f(x)\mathrm{d}x} & to\ x\in[a;\ b]\\ 0\ \ to\ x<a\ \ or\ \ x>b & .\end{cases}$$

Modifications of the method are possible. In particular, B.F. Kiryanov [3] proposed in the classical truncation to additionally use the shift of the abscissa axis to the intersection with the density curve at one of the closest points to the axis, a or b (for symmetry, the intersection can be in both boundary points).

We note that in this way the distribution obtained on the interval, in general, does not repeat the configuration of the initial distribution. With a different truncation technique, the configuration can be saved.

The paper presents other possible ways of modeling distributions based on truncation, two-sided or one-sided. A more complicated way of truncating is density function transformation $f(x)$ in given intervals by multiplying it by linear or quadratic factors.

With different methods of transformation, the truncation operation has a different degree of computational complexity. For a number of simulated distributions, computer programs have been developed that not only form the models of truncated distributions numerically, but also test and fit the theoretical distribution of the species in question according to the sample data.

2. Simplest methods of modeling

2.1. Models with configuration saving

Let there be given a distribution with a density function $f(x)$. In particular, $f(x)$ can be the density of a normal law $N(m;\sigma)$. We reduce the distribution to the interval $[a;b]$ with the preservation of trends in the density function. In addition, we take a point x_0, $x_0 \in [a;b]$. The density of the truncated $[a;b]$ distribution is assumed equal to

$$g(x) = \begin{cases} f(x) + h_1 \ \ to\ x \in [a;\ x_0] \\ f(x) + h_2 \ \ to\ x \in [x_0;\ b] \\ 0 \ \ to\ x < a \ \ or \ \ x > b \end{cases} . \qquad (1)$$

The shape of the obtained curve $g(x)$ is shown in Figure 1.

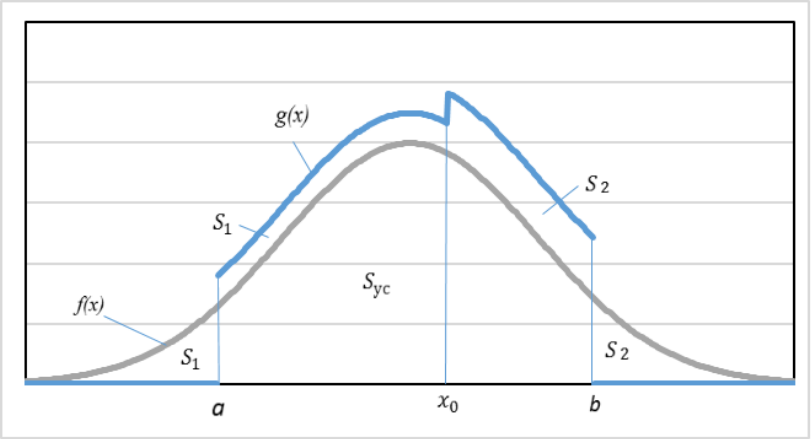

Figure 1. Graph of the density of a truncated distribution $g(x)$ kind (1).

Curve $g(x)$ in Fig. 1, in contrast to the classical truncation, preserves the configuration of the original curve $f(x)$ on intervals $[a;x_0]$ and $(x_0;b]$. The total area of the figures above the curve $f(x)$ is necessarily equal to the sum of the areas of the curvilinear trapezoids of the "distribution tails" $S_1 + S_2$. This condition allows you to select values h_1 and h_2, see. (1), in different ways.

For example, $h_1 = \frac{S_1}{x_0 - a}$, $h_2 = \frac{S_2}{b - x_0}$, where S_1 and S_2 – the area of curvilinear trapezoids of "tails of distribution". Let us check the normalization condition:

$$\int_{-\infty}^{+\infty} g(x)\,dx = \int_{a}^{x_0} (f(x)+h_1)\,\mathrm{d}x + \int_{x_0}^{b} (f(x)+h_2)\,\mathrm{d}x$$

$$= \int_{a}^{x_0} f(x)\,\mathrm{d}x + \int_{x_0}^{b} f(x)\mathrm{d}x + h_1(x_0-a) + h_2(b-x_0) = S_{\text{yc}} + S_1 + S_2 = 1.$$

Note that the areas S_1 and S_2 are expressed in terms of the distribution function of a non-truncated random variable: $S_1 = F(a)$, $S_2 = 1 - F(b)$

Assuming in (1) $h = h_1 = h_2$, we come at a continuous $[a;b]$ function $g(x)$, wherein, $h = \frac{S_1+S_2}{b-a} = \frac{F(a)+1-F(b)}{b-a}$. Consequently,

$$g(x) = \begin{cases} f(x) + \frac{1-(F(b)-F(a))}{b-a} \; to\; x \in [a;\; b] \\ 0 \;\; to \;\; x \notin [a;\; b] \end{cases}. \qquad (2)$$

Thus, unlike the classical truncation in this case, the density function is transformed in an additive manner.

In a model with a probability density of the form (1), it is possible to use other relations. For example,

a) $h_1 = \frac{S_2}{x_0-a}$, $h_2 = \frac{S_1}{b-x_0}$; b) $h_1 = \frac{\alpha(S_1+S_2)}{x_0-a}$, $h_2 = \frac{(1-\alpha)(S_1+S_2)}{b-x_0}$, where $0 < \propto < 1$.

Distribution function $G(x)$ at a density of $g(x)$ kind (1) on interval $[a;b]$ differs from $F(x)$ on the linear summand:

$$G(x) = \begin{cases} F(x) - F(a) + h_1(x-a) \; to\; x \in [a;\; x_0] \\ F(x) - F(a) + h_1(x_0-a) + h_2(x-x_0) \; to\; x \in [x_0; b]. \end{cases}$$

Similarly, one-sided truncations can be made.

2.2. The use of parabolic components

Let $f(x)$ be the density function of the initial distribution. We perform its truncation on an interval using additions of parabolic type. A graph of the corresponding truncation of the distribution is shown in Figure 2.

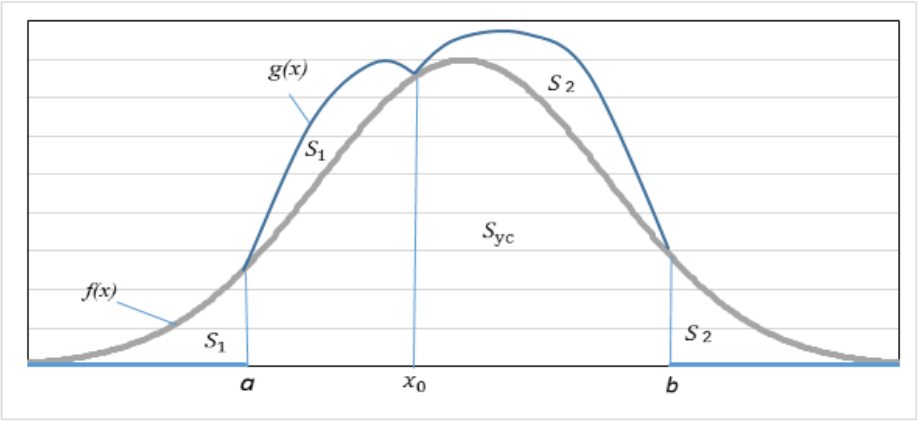

Figure 2. Graph of density $g(x)$ with superstructures of parabolic type.

Add-ons above the curve $f(x)$ figures should have given areas, for example, S_1 and S_2, and each of the curves of a parabolic type must pass through two given points, for example, $f(a)$, $f(x_0)$ and $f(x_0)$, $f(b)$. Wherein, x_0 is some value assigned by the researcher, $x_0 \in (a;b)$.
Then the density function of the truncated distribution $g(x)$ has the form

$$g(x) = \begin{cases} A_1x^2 + B_1x + C_1 \ \ to\ x \in [a; x_0] \\ A_2x^2 + B_2x + C_2 \ \ to\ x \in [x_0; b] \\ 0 \ \ \ to\ x \notin [a; b]\ . \end{cases} \quad (3)$$

Thus, the density curve in the truncation interval consists of two pieces of parabolas determined by the density $f(x)$ of basic distribution. Coefficients $A_1, B_1, C_1, A_2, B_2, C_2$ are calculated from the values of the function $f(x)$ in the boundary points, as well as the values of the areas of the superimposed figures. For the areas of the corresponding superstructures, in particular, S_1 and S_2 there derived two equations which are linear in regard to the mentioned coefficients:

$$\int_a^{x_0} \left(A_1x^2 + B_1x + C_1 - f(x)\right)\mathrm{d}x = S_1$$

$$\int_{x_0}^{b} \left(A_2x^2 + B_2x + C_2 - f(x)\right)\mathrm{d}x = S_2$$

Consequently, the coefficients in formula (3) can be found as a solution of two systems consisting of three linear algebraic equations each. In particular, in the case shown in Figure 2 (with areas of superstructures $S_1 = F(a)$ and $S_2 = 1 - F(b)$), we get

$$\begin{cases} A_1a^2 + B_1a + C_1 = f(a) \\ A_1x_0^2 + B_1x_0 + C_1 = f(x_0) \\ A_1\left(\frac{x_0^3 - a^3}{3}\right) + B_1\left(\frac{x_0^2 - a^2}{2}\right) + C_1(x_0 - a) = F(x_0) \end{cases} \quad (4)$$

$$\begin{cases} A_2x_0^2 + B_2x_0 + C_2 = f(x_0) \\ A_2b^2 + B_2b + C_2 = f(b) \\ A_2\left(\frac{b^3 - x_0^3}{3}\right) + B_2\left(\frac{b^2 - x_0^2}{2}\right) + C_2(b - x_0) = 1 - F(x_0) \end{cases} \quad (5)$$

The main determinants of each of the systems (4), (5) are respectively equal $\frac{(x_0-a)^4}{6}$ and $\frac{(b-x_0)^4}{6}$, and hence are different from zero for $x_0 \in (a;b)$. Therefore, the coefficients in the density function (3), as solutions of systems, are uniquely determined. In this case, as can be seen from the systems, they directly depend on the values a, x_0, b and the base distribution with the density function $f(x)$ and the distribution function $F(x)$.

Note that the area of the figures in parabolic superstructures when truncating the base distribution can be related not only as S_1 and S_2, but also differently, for example, S_2 and S_1 or $\alpha(S_1 + S_2)$, $\beta(S_1 + S_2)$, where $\alpha + \beta = 1$. The only restriction that guarantees the fulfillment of the normalization condition for the truncated distribution: the sum of the added areas is equal to $S_1 + S_2$, and $S_1 + S_2 = 1 - (F(b) - F(a))$.

The superstructure in the truncation of the distribution can also consist of one parabolic curve. In this case, only three unknown coefficients. They are calculated as the solution of the corresponding system of three linear algebraic equations constructed according to the above algorithm. We represent the expressions for the density function for such truncations.

1). $g(x) = \begin{cases} Ax^2 + Bx + C \ \ to\ x \in [a; b] \\ 0\ to\ x \notin [a; b] \end{cases}$

The resulting truncated distribution is a modification of the symmetric classical parabolic distribution. In this case, for arbitrary and different basic functions $f(x)$ symmetry, generally speaking, is absent.

$$2).\ g(x) = \begin{cases} Ax^2 + Bx + C\ to\ x \in [a; x_0] \\ f(x)\ \ to\ x \in (x_0; b] \\ 0\ \ to\ x \notin [a; b], \end{cases}$$

where $x_0 \in (a; b)$.

$$3).\ g(x) = \begin{cases} f(x)\ to\ x \in [a; x_0] \\ Ax^2 + Bx + C\ \ to\ x \in (x_0; b] \\ 0\ \ to\ x \notin [a; b], \end{cases}$$

where $x_0 \in (a; b)$.

$$4).\ g(x) = \begin{cases} f(x)\ to\ x \in [a; x_0]\ \cup [x_1; b] \\ Ax^2 + Bx + C\ \ to\ x \in (x_0; x_1) \\ 0\ \ to\ x \notin [a; b], \end{cases}$$

where $a < x_0 < x_1 < b$.

$$5).\ g(x) = \begin{cases} A_1x^2 + B_1x + C_1\ \ to\ x \in [a; x_0] \\ f(x)\ \ to\ \ x \in (x_0; x_1) \\ A_2x^2 + B_2x + C_2\ \ to\ x \in [x_1; b] \\ 0\ \ to\ x \notin [a; b], \end{cases}$$

where $a < x_0 < x_1 < b$.

The computer implementation of this method of truncating different distributions makes it easy to calculate the coefficients, the distribution function, and the numerical characteristics associated with the moments.

Varying by position parameters a, b, x_0, x_1 and the ratio of areas in superstructures above the density function graph $f(x)$. The basic distribution leads to a large variety of simulated truncated distributions.

Similarly, one-sided truncations can be considered, using a larger number of parabolic-type components, as well as other more complex configurations and their combinations.

3. Simulation of truncated distributions with a given quantile and additional factors

Let $f(x)$ is density of the base distribution of a random variable X. We transform this distribution by truncating to a given interval $[a; b]$. We introduce the restriction in the form of quantiles of order α with the value of this quantile x_α. At the same time, the interval $[a; b]$ is divided into two intervals $[a; x_\alpha]$, $(x_\alpha; b]$. On each of the intervals obtained for the density function we introduce additional factors.

For linear factors, the density of the modeled distribution $g(x)$ has the form:

$$g(x) = \begin{cases} (B_1x + C_1)f(x)\ \ to\ x \in [a; x_\alpha] \\ (B_2x + C_2)f(x)\ \ to\ x \in (x_\alpha; b] \\ 0\ \ \ to\ x \notin [a; b]. \end{cases} \quad (6)$$

In this case, $g(x)$ is the density of a certain distribution, it is sufficient that it is nonnegative and that the normalization condition $\int_a^b g(x)\mathrm{d}x = 1$.

Coefficients B_1, C_1, B_2, C_2 are found as solutions of the system of four linear equations that guarantee and the fulfillment of the normalization condition:

$$\begin{cases} \int_a^{x_\alpha}(B_1x + C_1)f(x)dx = \alpha \\ \int_{x_\alpha}^{b}(B_2x + C_2)f(x)dx = \beta \\ (B_1x_\alpha + C_1)f(x_\alpha) = g(x_\alpha) \\ (B_2x_\alpha + C_2)f(x_\alpha) = g(x_\alpha) \end{cases} \quad (7)$$

Numerical values α, $g(x_\alpha)$ are chosen by the researcher, and the probabilities α and β are related by $\alpha + \beta = 1$. Consequently, $a, b, \alpha, x_\alpha, g(x_\alpha)$ are parameters of the simulated truncated distribution. The non-negativity of the function $g(x)$ is limited by $g(a) \geq 0,\ g(b) \geq 0,\ g(x_\alpha) \geq 0$

We note that the solution of the system (7) essentially depends on the type of the base distribution with the density function $f(x)$, which is clearly present in each of the equations of the system. We give the result under the basic normal distribution.

Theorem, [4]. Let $f(x)$ be the density of a Gaussian random variable $X: X \sim N(m; \sigma)$. The following parameters have been fulfilled: $x_\alpha \in (a; b)$, $g(x_\alpha) \geq 0$, $0 < \alpha < 1$, $\beta = 1 - \alpha$. Then the coefficients of the density function $g(x)$ of a simulated distribution of the form (6) satisfy the relations

$$B_1 = \frac{\alpha - \frac{g(x_\alpha)}{f(x_\alpha)}\left[\Phi\left(\frac{x_\alpha - m}{\sigma}\right) - \Phi\left(\frac{a - m}{\sigma}\right)\right]}{\sigma\left[\varphi\left(\frac{a - m}{\sigma}\right) - \varphi\left(\frac{x_\alpha - m}{\sigma}\right)\right] + (m - x_\alpha)\left[\Phi\left(\frac{x_\alpha - m}{\sigma}\right) - \Phi\left(\frac{a - m}{\sigma}\right)\right]};$$

$$B_2 = \frac{\beta - \frac{g(x_\alpha)}{f(x_\alpha)}\left[\Phi\left(\frac{b - m}{\sigma}\right) - \Phi\left(\frac{x_\alpha - m}{\sigma}\right)\right]}{\sigma\left[\varphi\left(\frac{x_\alpha - m}{\sigma}\right) - \varphi\left(\frac{b - m}{\sigma}\right)\right] + (m - x_\alpha)\left[\Phi\left(\frac{b - m}{\sigma}\right) - \Phi\left(\frac{x_\alpha - m}{\sigma}\right)\right]};$$

$$C_1 = \frac{g(x_\alpha)}{f(x_\alpha)} - B_1x_\alpha;\ C_2 = \frac{g(x_\alpha)}{f(x_\alpha)} - B_2x_\alpha,$$

$$B_1a + C_1 \geq 0,\ B_2b + C_2 \geq 0,$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ is the probability density of the standard normal law, $\mathrm{F}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$ is Laplace function.

Theorem, [4]. Let $f(x)$ be the density of a Gaussian random variable $X: X \sim N(0; 1)$. The density of a truncated distribution $g(x)$ satisfies the relation (6). Then for the mathematical expectation and variance of the truncated distribution, the following relations hold:

$$M(X) = B_1[a\varphi(a) - x_\alpha\varphi(x_\alpha)] + B_1[\Phi(x_\alpha) - \Phi(a)] - C_1[\varphi(x_\alpha) - \varphi(a)]$$

$$+B_2[x_\alpha\varphi(x_\alpha) - b\varphi(b)] + B_2[\Phi(b) - \Phi(x_\alpha)] - C_2[\varphi(b) - \varphi(x_\alpha)];$$

$$D(X) = B_1[a^2\varphi(a) - {x_\alpha}^2\varphi(x_\alpha)] + C_1[\Phi(x_\alpha) - \Phi(a)] + C_1[\varphi(a) - \varphi(x_\alpha)]$$

$$+B_2[{x_\alpha}^2\varphi(x_\alpha) - b^2\varphi(b)] + C_2[\Phi(b) - \Phi(x_\alpha)] + C_2[\varphi(x_\alpha) - \varphi(b)] - M^2(X).$$

Similarly, using functions $\varphi(x)$ and $\Phi(x)$, in distributions modeled on the basis of the basic normal law, we obtain expressions for the distribution function, for the characteristic function, and for higher-order moments.

For quadratic multipliers, the density of the model distribution $g(x)$ has the form:

$$g(x)=\begin{cases}(A_1x^2+B_1x+C_1)f(x) \ \ to\ x\in[a;x_\alpha]\\(A_2x^2+B_2x+C_2)f(x) \ \ to\ x\in(x_\alpha;b]\\0 \ \ to\ x\notin[a;b]\,.\end{cases} \tag{8}$$

Coefficients A_i, B_i, C_i, $(i=1,2)$ in the relation (8) are both solutions of systems of linear algebraic equations with respect to these coefficients

$$\begin{cases}\int_a^{x_\alpha}(A_1x^2+B_1x+C_1)f(x)dx=\alpha\\(A_1x_\alpha{}^2+B_1x_\alpha+C_1)f(x_\alpha)=g(x_\alpha)\\(A_1a^2+B_1a+C_1)f(a)=g(a)\end{cases}\qquad\begin{cases}\int_{x_\alpha}^{b}(A_2x^2+B_2x+C_2)f(x)dx=\beta\\(A_2x_\alpha{}^2+B_2x_\alpha+C_2)f(x_\alpha)=g(x_\alpha)\\(A_2b^2+B_2b+C_2)f(b)=g(b)\end{cases} \tag{9}$$

When $\alpha+\beta=1$, the normalization condition is satisfied. We note that after integration of each of the systems (9) in the first equation, we arrive at precisely the linear algebraic equation with respect to the coefficients A_i, B_i, C_i. When $x_\alpha\in(a;b)$, the solution of systems is unique. Again $a, b, \alpha, x_\alpha, g(x_\alpha)$ are the parameters of a simulated truncated distribution based on the base distribution with density $f(x)$.

Under the conditions formulated, an arbitrarily given set of parameters does not guarantee a probabilistic solution of systems (9). For the density function to be nonnegative $g(x)$ the nonnegativity of its quadratic multipliers is necessary:

$$\begin{cases}(A_1x^2+B_1x+C_1)\geq 0\ to\ x\in[a;x_\alpha]\\(A_2x^2+B_2x+C_2)\geq 0\ \ to\ x\in(x_\alpha;b]\end{cases}.$$

For verification, we can use two lemmas.

Lemma 1. If the square trinomial $Q(x)=Ax^2+Bx+C$ on interval $[r;s]$ satisfies the conditions $Q(r)=0, Q(s)=0, A<0$ $A<0$, then $Q(x)$ is nonnegative on $[r;s]$.

Lemma 2. If the square trinomial $Q(x)=Ax^2+Bx+C$ has no roots on the interval $(r;s)$ and $Q(r)\geq 0,\ Q(s)>0$ (or $Q(r)>0, Q(s)\geq 0$), then $Q(x)$ is nonnegative on $[r;s]$.

For the basic function of the standard normal law, the corresponding formulas for computing the coefficients, the distribution function, and the moments of even and odd orders are presented in [5]. In addition, along with the functions $\varphi(x)$ and $\Phi(x)$, In the relations obtained, we use double factorials and polynomials of a certain class. In [6], specific relationships were found for calculating the coefficients, the distribution function, and the moments for the basic exponential distribution.

By the same algorithm for constructing a distribution with density $g(x)$ The uniform distribution on the segment, the distribution with a sinusoidal density and the distribution [7, 8] of the hyperbolic cosine type are investigated as a basis. In all cases, depending on the parameters, non-trivial distributions with a complex structure are obtained. Example of density function graph $g(x)$ of kind (8) the simulated distribution is shown in Figure 3. Note that the configuration of the graph is very sensitive to the distribution parameters, and also to the choice of the basic function.

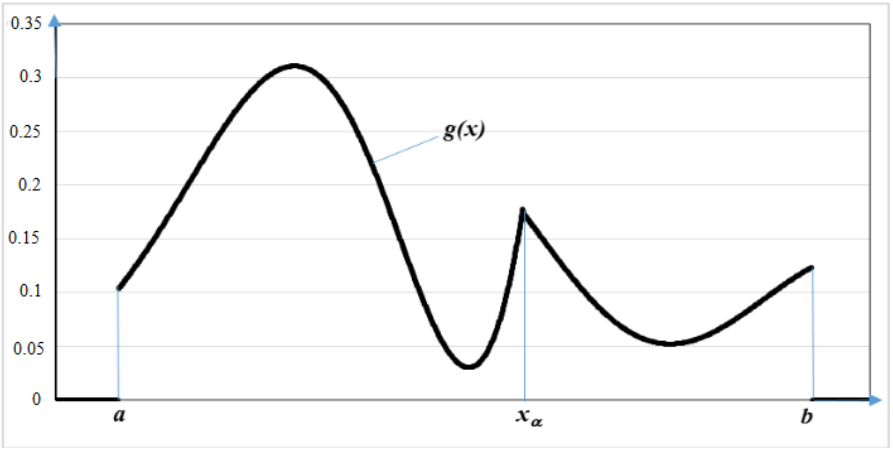

Figure 3. Graph of the density function $g(x)$ at the basic normal distribution.

As a result of the transformations, the final distribution turns out to be a mixture of different distributions, which, as a rule, corresponds to the multifactority and inhomogeneity of the impact in the formation of the indicator under study, based on real data. The input of quantiles makes it possible to take into account possible qualitative transitions corresponding to the data structure.

According to the form of the function (8) for any base function $f(x)$ the mixture turns out to consist of six truncated distributions: three for each of the intervals $[a; x_\alpha]$, $(x_\alpha; b]$. In particular, with the basic normal distribution, the components of the mixture are: initial density $f(x)$, displaced Rayleigh distribution density $\left(\frac{x-m}{\sigma^2}e^{-\frac{(x-m)^2}{2\sigma^2}},\ x > m\right)$ and the displaced Maxwell distribution density $\left(\frac{2(x-m)^2}{\sigma^3\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}},\ x > m\right)$. At the base density of the exponential distribution, in addition to the initial density, the components of the mixture are the Erlang distribution densities of the second and third orders.

When modeling a particular type of function $g(x)$ in accordance with real data, the computer program checks compliance with Pearson's consent criterion. Distribution parameters are adjusted to minimize the value χ^2_{view}.

4. Conclusion

Statistical indicators are formed as a result of the interaction of many factors, heterogeneous and multidirectional. Therefore, for many real statistical data, as a rule, not only normal, but also other classical distribution laws do not work. Naturally, in statistical data possessing a complex nonstandard structure, the theoretical probability distribution must also be structurally complex. It is necessary to model non-standard distributions. The obtained distributions allow us to consider and analyze the estimated inhomogeneous general populations, and also help in justifying their decomposition into homogeneous components.

The considered method of modeling based on truncation with further transformations defines an extensive class of probability distributions with non-standard structure. Let's formulate the algorithm of actions:

1) basic distribution (selection of the basic function $f(x)$;
2) truncation, two-sided or one-sided (border selection, a and b);
3) truncation, two-sided or one-sided (border selection α and the corresponding quantile x_α);
4) density boundary values (choice of values $g(a), g(x_\alpha), g(b)$);

5) calculation of coefficients $g(x)$ with the control of the non-negativity of the density function.

The results of modeling new probability distributions can be very useful when processing statistical data with nonstandard probability ratios.

References

[1] Cramer G 1975 *Mathematical Methods of Statistics* (Moscow: Mir) p 631
[2] Vadzinsky R N 2001 *Handbook of Probability Distributions* (SanktPetersburg: Science) p 295
[3] Kiryanov B F, Tokmachev M S 2009 *Mathematical Models in Public Health* (Veliky Novgorod: Publishing House of Novgorod State University) p 279
[4] Tokmachev M S, Ryazantsev P P 2010 Simulation of truncated distributions *Bulletin of NovSU*. No. **55** pp 34–36
[5] Tokmachev M S 2011 Simulation of truncated distributions with a given quantile *Bulletin of KSTU A N Tupolev*. No. **2** pp 157–162
[6] Tokmachev M S 2011 Simulation of truncated distributions with a given quantile based on the exponential distribution *Bulletin of NovSU. S.: Techn. science*. No. **65** pp 97–100
[7] Tokmachev M S 1995 Constant regression of quadratic statistics on linear statistics *Bulletin of NovSU. S: Natural. and techn. Science.* No. **1** pp 139–141
[8] Tokmachev M S 2017 Investigation of the probability distribution of hyperbolic cosine type *Bulletin SUSU. S: Mathematics. Mechanics. Physics.* T. 9, No. **3** pp 18–26