Modeling non-standard probability laws based on a mixture of truncated distributions

M S Tokmachev
Yaroslav-the-Wise Novgorod State University, 41, ul. B. St. Petersburgskaya, Veliky Novgorod, Russian Federation
E-mail: Mikhail.Tokmachov@novsu.ru

Abstract. A method for modeling probability distributions is presented using two operations: truncating multiple base distributions in various ways, and converting them to a mixture. The normal distribution is considered as the basic one. The use of truncations and varying the truncation intervals of the mixture components lead to distributions with a complex non-standard configuration, in particular, asymmetric and multi-vertex ones. Distributions simulated in this way are used as theoretical ones in computer programmed fitting to sample data with a non-standard structure. The conditions for the continuity of the final distribution are obtained. The construction algorithm allows implementing other distributions.

1. Introduction
Researchers in various fields of knowledge often encounter real statistical data that do not correspond to the classical laws of distribution. This is due to the heterogeneity of the studied general populations, as well as the complex structure of the studied trait itself. In this case, as a rule, the numerical values of a feature are formed when it is exposed to a multitude of multidirectional factors. As a result, the sample distributions are far from standards and have a variety of configurations. In such cases, it is very problematic to choose the appropriate theoretical distribution from the set of popular distributions.

According to the well-known remark of J. Tukey [1], “From the point of view of mathematics, the initial situation for the analysis would be much “cleaner” if
- we knew exactly the distribution of observables;
- this exact distribution could be simply expressed analytically.
In the real world, none of these requirements are met”.

Nevertheless, there remains the need to find the distribution for the study of the sample. Therefore, along with relatively simple (classical) distributions, researchers need other nonstandard distributions of various configurations.

Many methods of modeling probability distributions of varying complexity are known from the literature, for example, truncation of distributions in various ways [2], the use of mixtures of distributions in various combinations [3–7], the use of functional transformations of random variables [8–10], in particular, many classical distributions were obtained in such a way (distributions of Pearson, Student, Fisher, Weibull [7, 11]), the introduction of additional quantile characteristics [12, 13], etc. Let's supplement these methods.

In this work, to obtain new distributions based on previously known ones (in particular, the classical normal distribution), two successive techniques are used: truncating distributions, and transforming them into a mixture.
Using truncated distributions in a mixture simplifies models in many cases. If the components of the mixture are located at non-intersecting intervals, then instead of mixing, the final distribution is “glued” from separate pieces. Wherein, for arbitrary parameters, the continuity of the density curve is not guaranteed.

In the simplest case, two truncated distributions are used in the mixture. This condition is not essential for modeling. Similarly, it is possible to extend the results to three or more components of the mixture.

The distribution laws of the components in the mixture can also be different. But they affect the complexity of the computational formulas in the simulated distribution, therefore, in practical terms, it is easier to restrict ourselves to a few classical laws. In this article normally distributed random variables are considered.

2. A mixture of two distributions truncated at zero

Let two Gaussian random values \( X_1 \sim N \left( m_1, \sigma_1 \right) \) and \( X_2 \sim N \left( m_2, \sigma_2 \right) \) with the corresponding density functions be given as

\[
f_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}, i = 1,2,
\]

and distribution functions \( F_i(x) = \Phi \left( \frac{x-m_i}{\sigma_i} \right), i = 1,2 \), where the Laplace function is

\[
\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.
\]

For definiteness, we put \( m_1 \leq m_2 \). The values of \( \sigma_1 \) and \( \sigma_2 \), generally speaking, are different.

We introduce the density functions of the distributions truncated at zero:

\[
f_1^{-}(x) = \frac{f_1(x)}{F_1(0)} f_{1}^{-}(x) = \frac{f_1(x)}{1-F_1(0)} \quad \text{for } x < 0; f_1^{+}(x) = \frac{f_1(x)}{1-F_1(0)} = \frac{f_1(x)}{1-\Phi \left( \frac{m_1}{\sigma_1} \right)} \quad \text{for } x \geq 0;
\]

\[
f_2^{-}(x) = \frac{f_2(x)}{F_2(0)} = \frac{f_2(x)}{\Phi \left( \frac{m_2}{\sigma_2} \right)} \quad \text{for } x < 0; f_2^{+}(x) = \frac{f_2(x)}{1-F_2(0)} = \frac{f_2(x)}{1-\Phi \left( \frac{m_2}{\sigma_2} \right)} \quad \text{for } x \geq 0.
\]

By combining pairs of different truncated distributions, we obtain different configurations of the density curve for the mixture. In particular, for positive \( \alpha \) and \( \beta, \alpha+\beta = 1 \), we denote this

\[
f_{12}(x) = \alpha f_1^{-}(x) + \beta f_2^{+}(x) = \begin{cases} \alpha \frac{f_1(x)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} & \text{for } x < 0 \\ \beta \frac{f_2(x)}{1-\Phi \left( \frac{m_2}{\sigma_2} \right)} & \text{for } x \geq 0 \end{cases} = \begin{cases} \alpha \frac{f_1(x)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} & \text{for } x < 0 \\ \beta \frac{f_2(x)}{\Phi \left( \frac{m_2}{\sigma_2} \right)} & \text{for } x \geq 0 \end{cases}
\]

(2)

For the continuity of the density function \( f_{12} (x) \), it is necessary to eliminate a possible gap at zero. We introduce the condition:

\[
\alpha \frac{f_1(0)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} = \beta \frac{f_2(0)}{\Phi \left( \frac{m_2}{\sigma_2} \right)}, \quad \text{i.e.} \quad \alpha \left[ \frac{f_1(0)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} + \frac{f_2(0)}{\Phi \left( \frac{m_2}{\sigma_2} \right)} \right] = \frac{f_2(0)}{\Phi \left( \frac{m_2}{\sigma_2} \right)}
\]

Therefore,

\[
\alpha = \frac{f_2(0)}{\Phi \left( \frac{m_2}{\sigma_2} \right)} \left[ \frac{f_1(0)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} + \frac{f_2(0)}{\Phi \left( \frac{m_2}{\sigma_2} \right)} \right]^{-1}, \quad \beta = 1 - \alpha = \frac{f_1(0)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} \left[ \frac{f_1(0)}{\Phi \left( \frac{m_1}{\sigma_1} \right)} + \frac{f_2(0)}{\Phi \left( \frac{m_2}{\sigma_2} \right)} \right]^{-1}
\]

(3)

Thereby, we have proved the theorem.

*Theorem 1.* Under conditions (3) for the coefficients \( \alpha \) and \( \beta \), the density function \( f_{12} (x) \) of the form (2) is continuous for all \( x \).
For different $m_1, m_2, \sigma_1, \sigma_2$ the corresponding graphs of the function $f_{12}(x)$ are shown in Figures 1–2.

![Graph of the density function $f_{12}(x)$ at $X_1 \sim N(-1; 1)$ and $X_2 \sim N(2; 1)$.](image1)

**Figure 1.** Graph of the density function $f_{12}(x)$ at $X_1 \sim N(-1; 1)$ and $X_2 \sim N(2; 1)$.

![Graph of the density function $f_{12}(x)$ at $X_1 \sim N(-2; 2)$ and $X_2 \sim N(-0.5; 0.5)$.](image2)

**Figure 2.** Graph of the density function $f_{12}(x)$ at $X_1 \sim N(-2; 2)$ and $X_2 \sim N(-0.5; 0.5)$.

For the mixture $f_{21}(x)$ for positive $\alpha$ and $\beta$, $\alpha + \beta = 1$, we get

$$f_{21}(x) = \alpha f_2^-(x) + \beta f_1^+(x) = \begin{cases} \alpha \frac{f_2(x)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} & \text{for } x < 0 \\ \beta \frac{f_1(x)}{1 - \Phi\left(-\frac{m_1}{\sigma_1}\right)} & \text{for } x \geq 0 \end{cases} = \begin{cases} \alpha \frac{f_2(x)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} & \text{for } x < 0 \\ \beta \frac{f_1(x)}{\Phi\left(-\frac{m_1}{\sigma_1}\right)} & \text{for } x \geq 0 \end{cases}$$ (4)

Similarly to the above, for the density function $f_{21}(x)$ to be continuous, the condition is necessary

$$\alpha \frac{f_2(0)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} = \beta \frac{f_1(0)}{\Phi\left(-\frac{m_1}{\sigma_1}\right)} \quad \text{i.e.} \quad \alpha \left[ \frac{f_2(0)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} + \frac{f_1(0)}{\Phi\left(-\frac{m_1}{\sigma_1}\right)} \right] = f_2(0) \frac{\Phi\left(-\frac{m_2}{\sigma_2}\right)}{\Phi\left(-\frac{m_1}{\sigma_1}\right)}$$

Therefore,

$$\alpha = \frac{f_1(0)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} \left[ \frac{f_2(0)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} + \frac{f_1(0)}{\Phi\left(-\frac{m_1}{\sigma_1}\right)} \right]^{-1}, \quad \beta = 1 - \alpha = \frac{f_2(0)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} \left[ \frac{f_2(0)}{\Phi\left(-\frac{m_2}{\sigma_2}\right)} + \frac{f_1(0)}{\Phi\left(-\frac{m_1}{\sigma_1}\right)} \right]^{-1}$$ (5)

Thereby, we have proved the theorem.

**Theorem 2.** Under conditions (5) for the coefficients $\alpha$ and $\beta$, the density function $f_{21}(x)$ of the form (4) is continuous for all $x$.

The configuration of the distribution can be judged by the graph of the function $f_{21}(x)$ in Figure 3 for specific parameters $m_1, m_2, \sigma_1, \sigma_2$. 

![Graph of the density function $f_{21}(x)$ for specific parameters $m_1, m_2, \sigma_1, \sigma_2$.](image3)
For relations $\alpha$ and $\beta$ different from (3), (5), $\alpha + \beta = 1$, the density functions $f_{i2}(x)$ and $f_{i1}(x)$ of the forms (2), (4) are discontinuous at zero.

3. One-sided truncation at the point of mathematical expectation of one of the distributions

Consider again two Gaussian random variables $X_1 \sim N(m_1; \sigma_1)$ and $X_2 \sim N(m_2; \sigma_2)$ with density functions $f_i(x)$ and distribution functions $F_i(x) = \Phi\left(\frac{x-m_i}{\sigma_i}\right)$, $i = 1, 2$. We put $m_1 \leq m_2$.

Consider the truncation of distributions at a point coinciding with one of the mathematical expectations: for definiteness, we take the point $m_2$ as the truncation boundary.

Similarly to (1), we introduce the density functions of the distributions truncated at the point $m_2$:

$$f^+_i(x) = \frac{f_i(x)}{F_i(m_2)} = -1 - \Phi\left(\frac{m_2-m_i}{\sigma_i}\right) \text{ for } x \geq m_2;$$

$$f^-_i(x) = \frac{f_i(x)}{1 - F_i(m_2)} = \Phi\left(\frac{m_2-m_i}{\sigma_i}\right) \text{ for } x < m_2;$$

$$f^+_2(x) = \frac{f_2(x)}{1 - F_2(m_2)} = \frac{f_2(x)}{1 - \Phi\left(\frac{m_2-m_2}{\sigma_2}\right)} = 2f_2(x) \text{ for } x \geq m_2;$$

For pairs of different distributions, mixtures are obtained. In particular, for positive $\alpha$ and $\beta$, $\alpha + \beta = 1$, we denote this

$$f_{12}(x) = \alpha f^-_1(x) + \beta f^+_2(x) = \begin{cases} \alpha \frac{f_1(x)}{\Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} & \text{for } x < m_2 \\ 2\beta f_2(x) & \text{for } x \geq m_2 \end{cases}. \quad (7)$$

For the function $f_{12}(x)$ to be continuous, the following condition is necessary:

$$\alpha \frac{f_1(m_2)}{\Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} = 2\beta f_2(m_2), \quad \text{i.e.} \quad \alpha \left[ f_1(m_2) \Phi\left(\frac{m_2-m_1}{\sigma_1}\right) + 2f_2(m_2) \right] = 2f_2(m_2).$$

Therefore,

$$\alpha = 2f_2(m_2) \left[ \frac{f_1(m_2)}{\Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} + 2f_2(m_2) \right]^{-1}, \quad \beta = \frac{f_1(m_2)}{\Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} \left[ \frac{f_1(m_2)}{\Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} + 2f_2(m_2) \right]^{-1}.$$

Replacing the density functions $f_1(m_2)$ and $f_2(m_2)$ with specific expressions, we get

![Graph of the density function $f_{21}(x)$ at $X_1 \sim N(-0.5; 2)$ and $X_2 \sim N(4; 1)$.](image-url)
\[
\alpha = \left[ 1 + \frac{\sigma_2}{2\sigma_1\Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} e^{-\frac{(m_2-m_1)^2}{2\sigma_1^2}} \right]^{-1}, \quad \beta = 1 - \alpha.
\]

Thereby, the following statement has been proved.

**Theorem 3.** Under conditions (8) for the coefficients \(\alpha\) and \(\beta\), the density function \(f_{12}(x)\) of the form (7) is continuous for all \(x\).

Similarly, for the mixture \(f_{12}(x)\) at positive \(\alpha\) and \(\beta\), \(\alpha + \beta = 1\), we get

\[
f_{12}(x) = \alpha f_{12}^-(x) + \beta f_{12}^+(x) = \begin{cases} 
2 \alpha f_1(x) & \text{for } x < m_2 \\
\beta \frac{f_1(x)}{1 - \Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} & \text{for } x \geq m_2,
\end{cases}
\]

For the density function \(f_{12}(x)\) to be continuous at the point of possible discontinuity \(m_2\), the following condition must be performed

\[
2 \alpha f_2(m_2) = \beta \frac{f_1(m_2)}{1 - \Phi\left(\frac{m_2-m_1}{\sigma_1}\right)}.
\]

For \(\beta = 1 - \alpha\) it follows that

\[
2 \alpha f_2(m_2) + \alpha \frac{f_1(m_2)}{1 - \Phi\left(\frac{m_2-m_1}{\sigma_1}\right)} = f_1(m_2).
\]

Substituting the values of the functions \(f_i(m_2)\) and \(f_j(m_2)\), we get

\[
\alpha = \left[ 1 + \frac{2\sigma_1}{\sigma_2} \left( 1 - \Phi\left(\frac{m_2-m_1}{\sigma_1}\right) \right) e^{-\frac{(m_2-m_1)^2}{2\sigma_1^2}} \right]^{-1}, \quad \beta = 1 - \alpha.
\]

Let us formulate the proved statement.

**Theorem 4.** Under conditions (10) for the coefficients \(\alpha\) and \(\beta\), the density function \(f_{12}(x)\) of the form (9) is continuous for all \(x\).

### 4. One-sided truncation of distributions at one arbitrary point

Let, as before, the random variables \(X_1 \sim N(m_1; \sigma_1)\) and \(X_2 \sim N(m_2; \sigma_2)\) are given. For definiteness, we put \(m_1 \leq m_2\).

Truncate the distributions at an arbitrary point \(x_0\). The density functions for the distributions truncated at \(x_0\) have the form:

\[
f_{1-}(x) = \frac{f_1(x)}{\phi\left(\frac{x_0-m_1}{\sigma_1}\right)} \text{ for } x < x_0; \quad f_{1+}(x) = \frac{f_1(x)}{1 - \Phi\left(\frac{x_0-m_1}{\sigma_1}\right)} \text{ for } x \geq x_0; \quad f_{2-}(x) = \frac{f_2(x)}{\phi\left(\frac{x_0-m_2}{\sigma_2}\right)} \text{ for } x < x_0; \quad f_{2+}(x) = \frac{f_2(x)}{1 - \Phi\left(\frac{x_0-m_2}{\sigma_2}\right)} \text{ for } x \geq x_0.
\]

For the density functions of truncated distributions located on opposite sides of the point \(x_0\), mixtures of the gluing type are obtained. In particular,

\[
f_{12}(x) = \alpha f_{1-}(x) + \beta f_{1+}(x) = \begin{cases} 
\alpha \frac{f_1(x)}{\phi\left(\frac{x_0-m_1}{\sigma_1}\right)} & \text{for } x < x_0 \\
\beta \frac{f_2(x)}{1 - \Phi\left(\frac{x_0-m_2}{\sigma_2}\right)} & \text{for } x \geq x_0
\end{cases},
\]

with continuity condition

\[
\alpha = \frac{f_2(x_0)}{1 - \Phi\left(\frac{x_0-m_2}{\sigma_2}\right)} \frac{f_1(x_0)}{\phi\left(\frac{x_0-m_1}{\sigma_1}\right)} + \frac{f_2(x_0)}{1 - \Phi\left(\frac{x_0-m_2}{\sigma_2}\right)} \right]^{-1}, \quad \beta = 1 - \alpha.
\]
\[ f_{21}(x) = \alpha f_2^-(x) + \beta f_1^+(x) = \begin{cases} \alpha \frac{f_2(x)}{\Phi\left(\frac{x-m_2}{\sigma_2}\right)} & \text{for } x < x_0 \\ \beta \frac{f_1(x)}{1 - \Phi\left(\frac{x-m_1}{\sigma_1}\right)} & \text{for } x \geq x_0 \end{cases} \]

with continuity condition

\[ \alpha = \frac{f_1(x_0)}{1 - \Phi\left(\frac{x_0-m_1}{\sigma_1}\right)} \left[ \frac{f_2(x_0)}{\Phi\left(\frac{x_0-m_2}{\sigma_2}\right)} + \frac{f_1(x_0)}{1 - \Phi\left(\frac{x_0-m_1}{\sigma_1}\right)} \right]^{-1}, \quad \beta = 1 - \alpha. \]

For distributions truncated at one point and on one side of this point, the mixture results in asymmetric, continuous distributions with density functions

\[ af_2^-(x) + \beta f_1^+(x) \] or \[ af_2^+(x) + \beta f_1^-(x). \]

The configuration of these distributions (“double humped”, “single humped” and without humps) depends on the location of the point \( x_0 \) in relation to \( m_1 \) and \( m_2 \) and the orientation of the truncations (both left-sided or both right-sided).

**5. One-sided truncation of two distributions at two points**

Consider random variables \( X_1 \sim N(m_1; \sigma_1) \) and \( X_2 \sim N(m_2; \sigma_2) \) and put \( m_1 < m_2 \). Let us make one-sided truncations of these distributions, and the truncation points, generally speaking, being different. Exclusively for simplicity of calculations, we assume truncation \( X_1 \) at point \( m_1 \) and truncation \( X_2 \) at point \( m_2 \).

The density functions of the distributions truncated at the points \( m_1 \) and \( m_2 \), respectively, have the form:

\[ f_1^-(x) = \frac{f_1(x)}{F_1(m_1)} = \frac{f_1(x)}{\Phi\left(\frac{m_1-m_1}{\sigma_1}\right)} = 2f_1(x) \] for \( x < m_1; \)

\[ f_1^+(x) = \frac{f_1(x)}{1 - F_1(m_1)} = \frac{f_1(x)}{1 - \Phi\left(\frac{m_1-m_1}{\sigma_1}\right)} = 2f_1(x) \] for \( x \geq m_1; \)

\[ f_2^-(x) = \frac{f_2(x)}{F_2(m_2)} = \frac{f_2(x)}{\Phi\left(\frac{m_2-m_2}{\sigma_2}\right)} = 2f_2(x) \] for \( x < m_2; \)

\[ f_2^+(x) = \frac{f_2(x)}{1 - F_2(m_2)} = \frac{f_2(x)}{1 - \Phi\left(\frac{m_2-m_2}{\sigma_2}\right)} = 2f_2(x) \] for \( x \geq m_2. \)

Consider pairs of these functions included in the mixture, which lead to different configurations of the simulated distribution.

**5.1. Components \( f_1^-(x) \) and \( f_2^+(x) \). Non-overlapping versatile truncations**

The density function of the mixture \( af_1^-(x) + \beta f_2^+(x) \) for positive \( \alpha \) and \( \beta \), \( \alpha + \beta \leq 1 \) turns out to be zero on the interval \([m_1; m_2]\). Let us supplement the distribution on the indicated interval with some function \( \phi(x) \). A non-negative function \( \phi(x) \) is chosen in a given class of functions taking into account the fulfillment of the normalization condition. Then

\[ f_{12}(x) = \begin{cases} 2\alpha f_1(x) & \text{for } x < m_1 \\ 2\beta f_2(x) & \text{for } x \geq m_2 \\ \phi(x) & \text{for } x \in [m_1; m_2]. \end{cases} \] (11)

Assuming the complementary function to be linear, \( \phi(x) = ax + b \), and the density function \( f_{12}(x) \) to be continuous, we find the equation of the straight line passing through two given points...
(m_1, 2\alpha f_1(m_1)) and (m_2, 2\beta f_2(m_2)), i.e. \( m_1, \frac{2\alpha}{\sqrt{2\pi}\sigma_1} \) and \( m_2, \frac{2\beta}{\sqrt{2\pi}\sigma_2} \). The equation of this straight line looks like this

\[
y = \frac{2}{\sqrt{2\pi}(m_2 - m_1)} \left( \frac{\beta}{\sigma_2} - \frac{\alpha}{\sigma_1} \right) x + \frac{2\alpha}{\sqrt{2\pi}\sigma_1} - \frac{2m_1}{\sqrt{2\pi}(m_2 - m_1)} \left( \frac{\beta}{\sigma_2} - \frac{\alpha}{\sigma_1} \right).
\]

To provide the condition for normalizing the distribution with density \( f_{12}(x) \), it is necessary that the area \( S \) of the corresponding trapezoid on the interval \([m_1; m_2] \) was equal to \( 1 - \alpha - \beta \). Thereby, we arrive at the relation

\[
S = \frac{1}{\sqrt{2\pi}} \left( \frac{\beta}{\sigma_2} + \frac{\alpha}{\sigma_1} \right) (m_2 - m_1) = 1 - \alpha - \beta.
\]

This implies the condition for the positive parameters \( \alpha \) and \( \beta \):

\[
\left( \frac{m_2 - m_1}{\sqrt{2\pi}\sigma_1} + 1 \right) \alpha + \left( \frac{m_2 - m_1}{\sqrt{2\pi}\sigma_2} + 1 \right) \beta = 1. \tag{12}
\]

Let us formulate the proved statement.

**Theorem 5.** Under conditions (12) for the coefficients \( \alpha \) and \( \beta \), the density function of the form (11)

\[
f_{12}(x) = \begin{cases} 2\alpha f_1(x) & \text{for } x < m_1 \\ 2\beta f_2(x) & \text{for } x \geq m_2 \\ ax + b & \text{for } x \in [m_1; m_2), \end{cases}
\]

is continuous for all \( x \) for

\[
a = \frac{2}{\sqrt{2\pi}(m_2 - m_1)} \left( \frac{\beta}{\sigma_2} - \frac{\alpha}{\sigma_1} \right), \quad b = \frac{2\alpha}{\sqrt{2\pi}\sigma_1} - \frac{2m_1}{\sqrt{2\pi}(m_2 - m_1)} \left( \frac{\beta}{\sigma_2} - \frac{\alpha}{\sigma_1} \right).
\]

A characteristic graph of a function \( f_{12}(x) \) of the form (11) with a linear insertion \( \phi(x) \) at \( x \in [m_1; m_2] \) is shown in figure 4.

![Figure 4](image_url)

**Figure 4.** Graph of the density function \( f_{12}(x) \) at \( X_1 \sim N(1; 1) \) and \( X_2 \sim N(3; 2) \) for \( \alpha = 0.25, \beta = 0.39353 \).

We give two corollaries of the theorem.

**Corollary 5.1.** For \( \left( \frac{\beta}{\sigma_2} = \frac{\alpha}{\sigma_1} \right) \), the straight line segment in formula (11) is parallel to the abscissa axis, namely: \( \phi(x) = \frac{2\alpha}{\sqrt{2\pi}\sigma_1} \) on the interval \([m_1; m_2] \).

**Corollary 5.2.** For \( \sigma_1 = \sigma_2 = \sigma \) the coefficients \( \alpha \) and \( \beta \) satisfy the relation below

\[
\alpha + \beta = \left( \frac{m_2 - m_1}{\sqrt{2\pi}\sigma} + 1 \right)^{-1}.
\]

Next, consider the density function \( f_{12}(x) \) of the form (11) with an additional quadratic function \( \phi(x) \)

\[
f_{12}(x) = \begin{cases} 2\alpha f_1(x) & \text{for } x < m_1 \\ 2\beta f_2(x) & \text{for } x \geq m_2 \\ Ax^2 + Bx + C & \text{for } x \in [m_1; m_2), \end{cases} \tag{13}
\]

for \( Ax^2 + Bx + C \geq 0 \) on the interval \([m_1; m_2] \).
In the indicated relation (13), the positive parameters $\propto$ and $\beta$, $\alpha + \beta < 1$, are chosen by the researcher, the values of $A, B, C$ are calculated.

To find $A, B, C$, we compose the system of equations

$$\begin{aligned}
\int_{m_1}^{m_2} Ax^2 + Bx + C \, dx &= 1 - \alpha - \beta \\
Am_1^2 + Bm_1 + C &= 2a f_1(m_1) \\
Am_2^2 + Bm_2 + C &= 2\beta f_2(m_2).
\end{aligned}$$

The first of the equations follows from the normalization condition for the density function $f_{12}(x)$. The other two equations guarantee the continuity of the function $f_{12}(x)$ at points $m_1$ and $m_2$, respectively.

After performing the integration, we arrive at a system of three linear algebraic equations with three unknowns:

$$\begin{aligned}
A \left( \frac{m_2^3 - m_1^3}{3} \right) + B \left( \frac{m_2^2 - m_1^2}{2} \right) + C(m_2 - m_1) &= 1 - \alpha - \beta \\
Am_1^2 + Bm_1 + C &= \frac{2\alpha}{\sqrt{2\pi}\sigma_1} \\
Am_2^2 + Bm_2 + C &= \frac{2\beta}{\sqrt{2\pi}\sigma_2}.
\end{aligned}$$

(14)

Determinant of the system is $\Delta = \frac{(m_2-m_1)^4}{6}$, i.e. for $m_2 \neq m_1$ we obtain $\Delta \neq 0$. Consequently, the system of equations (13) has a unique solution with respect to $A, B, C$, which, obviously, can be found by standard methods, for example, the Cramer method.

Note that the solution has a probabilistic meaning. Therefore, the condition $Ax^2 + Bx + C \geq 0$ for $x \in [m_1; m_2)$ is obligatory. For $A < 0$, the inequality holds due to $f_{12}(m_1) > 0, f_{12}(m_2) > 0$. For $A > 0$, the trinomial should not have different roots on $[m_1; m_2)$. In particular, the inequality $B^2 - 4AC \leq 0$ is a sufficient condition for the trinomial to be non-negative.

Theorem 6. If $A, B, C$ satisfy system (14) and the condition $Ax^2 + Bx + C \geq 0$ is satisfied for $x \in [m_1; m_2)$, then the density function $f_{12}(x)$ of the form (13) with given positive coefficients $\propto$ and $\beta$, where $\alpha + \beta < 1$, is continuous for all $x$.

Typical examples of graphs of the function $f_{12}(x)$ supplemented by the quadratic function $\varphi(x)$ at $x \in [m_1; m_2)$ are shown in figures 5, 6.

![Figure 5. Graph of the density function $f_{12}(x)$ at $X_1 \sim N(1; 1)$ and $X_2 \sim N(3; 2)$ for $\alpha = 0.3, \beta = 0.4$.](image-url)
Let three Gaussian random variables \( X_1 \sim N (1; 1) \) and \( X_2 \sim N (3; 2) \) for \( \alpha = 0.2 \), \( \beta = 0.2 \).

Graph view \( y = Ax^2 + Bx + C \) at \( x \in [m_1; m_2] \), convex downward (figure 5) or convex upward (figure 6), is determined by the area of the corresponding curved trapezoid with the base \( m_2-m_1 \), which is equal to \( 1 - \alpha - \beta \).

5.2. Components \( f_i^{-}(x) \) and \( f_i^{+}(x) \). Intersecting versatile truncations

The simulated density function \( f_{21}(x) \) has the form:

\[
f_{21}(x) = \alpha f_2^{-}(x) + \beta f_2^{+}(x) = \begin{cases} 
2\alpha f_2(x) & \text{for } x < m_1 \\
2\alpha f_2(x) + 2\beta f_2(x) & \text{for } x \in [m_1; m_2] \\
2\beta f_2(x) & \text{for } x \geq m_2.
\end{cases}
\]

Note that, for nonzero \( \alpha \) and \( \beta \), the function \( f_{21}(x) \) is discontinuous at the points \( m_1 \) and \( m_2 \).

6. A mixture of three truncated normal distributions

Let three Gaussian random variables \( X_1 \sim N (m_1; \sigma_1) \), \( X_2 \sim N (m_2; \sigma_2) \) and \( X_3 \sim N (m_3; \sigma_3) \) be given. For definiteness, we assume that \( m_1 < m_2 < m_3 \). We simulate a new distribution using the above method: truncation of distributions, and mixture (in particular, gluing).

Consider the simplest case when the distributions are truncated at the points of the expectation: \( m_1 \) and \( m_3 \).

Let us present the density functions of the truncated distributions.

Right-sided truncation of the distribution of the random variable \( X_1 \):

\[
f_1^{-}(x) = 2f_1(x) \quad \text{for } x < m_1.
\]

Left-sided truncation of the distribution of the random variable \( X_1 \):

\[
f_1^{+}(x) = 2f_1(x) \quad \text{for } x \geq m_1.
\]

Two-sided truncation of the distribution of the random variable \( X_2 \):

\[
f_2^{\pm}(x) = \frac{f_2(x)}{p \left( \frac{m_3-m_2}{\sigma_2} \right) - p \left( \frac{m_1-m_2}{\sigma_2} \right)} \quad \text{for } x \in [m_1; m_2].
\]

Right-sided truncation of the distribution of the random variable \( X_3 \):

\[
f_3^{-}(x) = 2f_3(x) \quad \text{for } x < m_3.
\]

Left-sided truncation of the distribution of the random variable \( X_3 \):

\[
f_3^{+}(x) = 2f_3(x) \quad \text{for } x \geq m_3.
\]

Note that when modeling distributions, it is possible to use other truncations, for example, setting \( m_2 \) as the truncation boundary.

Combining the components of the mixture of three truncated distributions, we simulate new distributions of various configurations. In particular, we write the density function of the mixture of three non-intersecting truncations:

\[
f_{123}(x) = \alpha f_1^{-}(x) + \beta f_2^{+}(x) + (1 - \alpha - \beta)f_3^{+}(x)
\]
For the density function \( f_{123}(x) \) to be continuous, the following condition is necessary:
\[
\begin{align*}
2\alpha f_1(m_1) &= \beta \frac{f_2(m_1)}{\Phi \left( \frac{m_3-m_2}{\sigma_2} \right) - \Phi \left( \frac{m_1-m_2}{\sigma_2} \right)} \\
\beta \frac{f_2(m_3)}{\Phi \left( \frac{m_3-m_2}{\sigma_2} \right) - \Phi \left( \frac{m_1-m_2}{\sigma_2} \right)} &= 2(1 - \alpha - \beta) f_3(m_3).
\end{align*}
\]
This system of two linear algebraic equations has a solution
\[
\begin{align*}
\beta &= \frac{2f_1(m_1)f_3(m_3)}{2f_1(m_1)f_3(m_3) + f_1(m_2)f_2(m_3) + f_2(m_1)f_3(m_3)} \\
\alpha &= \frac{f_2(m_1)}{2f_1(m_1)\Phi \left( \frac{m_3-m_2}{\sigma_2} \right) - \Phi \left( \frac{m_1-m_2}{\sigma_2} \right)}.
\end{align*}
\]
Similarly, using the same principles, other distributions can be obtained, in which there will be not only “gluing” components, but also classical mixtures.

7. Conclusion
Probability distributions, presented as mixtures of truncated normal distributions, have a diverse structure. This is evidenced by various configurations of the graphs of the density function of mixtures. Modeling of these distributions is easily programmable according to the specified parameters \( m, \sigma \) and truncation boundaries.

The mixture coefficients \( \alpha \) and \( \beta \) can either be specified initially or calculated using the formulas given. The coefficients \( A, B, C \) of insertions between the truncation boundaries are obtained as solutions of the system of linear algebraic equations.

Note that the simulated density functions are not always continuous: at the boundaries of the truncations, there can be a discontinuity of the “jump” type. In the cases considered, the ratios of the mixture coefficients are found that ensure the continuity.

The indicated modeling method can be extended to other distributions of mixture components, for example, asymmetric and with nonzero kurtosis. For this, it is sufficient to know the density functions and the distribution functions of the components. Other types of distributions in the mixture increase the structural diversity of the simulated distributions. In particular, such components of a mixture can be random variables with a lognormal distribution, Weibull distribution, and distribution of the hyperbolic cosine type [14]. However, as noted above, with more complex distribution laws, the complexity of the computational formulas for the parameters also increases.

The obtained typical models with various parameters are used as theoretical distributions in the study of real sample data that do not correspond to the classical distribution laws. The selection of the parameters of the assumed distribution can be made by the method of computer fitting in order to minimize the observed value of the goodness-of-fit criterion of the sample and theoretical distributions, for example [15], the criterion \( \chi^2 \).

Also, there are no fundamental restrictions on the choice of truncation points of distributions: they can be arbitrary, not coinciding with the mathematical expectations of distributions. Truncation can be omitted for some components in the mixture, which is more consistent with the classical form of a mixture of standard distributions.
The freedom to choose the type of distributions, truncation points, and the combination of components in the mixture, up to the use of inserts, expands the possibilities of modeling an adequate theoretical distribution when examining real data.

References
[1] Tukey J 1981 Analysis of observation results (M.: Mir) p 693 [in Russ]
[2] Tokmachev M S 2018 Modeling of truncated probability distributions IOP Conf. Series: Materials Science and Engineering electronic edition p 012056
[3] Feller V 1984 Introduction to probability theory and its applications (M.: Mir) Vol. 2 p 752 [in Russ]
[4] Lukach E 1917 Characteristic functions (M.: Nauka) p 424 [in Russ]
[5] Medgyessy P 1977 Decomposition of superpositions of density functions and discrete distributions (Budapest: Akademia Kiado)
[6] Ayvazyan S A and Mkhitaryan V S 1998 Applied statistics and fundamentals of econometrics (M.: YuNITI) p 1022 [in Russ]
[7] Dzhonson N L, Kots S and Balakrishnan N 2012 One-dimensional continuous distributions. Part 1. Tutorial (M.: Binom. Laboratoriya znaniy) p 703 [in Russ]
[8] 2008 Collection of problems in probability theory, mathematical statistics and the theory of random functions ed. A A Sveshnikov (SPb.: Lan) p 448 [in Russ]
[9] Tokmachev M S 2004 Functions of random variables Teaching aid (Veliky Novgorod: NovGU) p 74 [in Russ]
[10] Rao S P 1968 Linear statistical methods and their applications (Moscow: Nauka) p 548 [in Russ]
[11] Vadzinsky R N 2001 Probability distributions handbook (SPb.: Nauka) p 295 [in Russ]
[12] Tokmachev M S 2011 Modeling truncated distributions with a given quantile Bulletin of KSTU (Kazan State Technical University) named after A. N. Tupolev) 2 157–162 [in Russ]
[13] Tokmachev M S 2019 Simulation of Probability Distributions Using two quantile Characteristics Journal of Physics: Conference Series electronic edition p 012057
[14] Tokmachev M S 2017 Investigation of probabilistic distribution of hyperbolic cosine type Bulletin of SUSU (South Ural State University) Ser.: “Mathematics. Mechanics. Physics” Vol. 9 3 18–26 [in Russ]
[15] Kramer G 1984 Mathematical methods of statistics (M.: Inostrannaya literatura) p 631 [in Russ]