Goodness-of-fit Tests
for Mixture of Two Weibull Distribution Function

Gennady Martynov

Institute for Information Transmission Problems
of the Russian Academy of Sciences (Kharkevich Institute)
Moscow

e-mail: martynov@iitp.ru
magevl@gmail.com

==

INTERNATIONAL CONFERENCE
MKCM-10, Divnomorsk, June 1, 2025

1. Introduction

The aim of this paper is to explore the feasibility of using parametric goodness-of-fit criteria, specifically the Cramér–von Mises–Smirnov criterion, in the context of a multiparameter hypothetical family of distributions. As an example, a five-parameter family of a mixture of two Weibull distributions was chosen. The family of such distributions is

$$F(x;\theta) = \{\rho W(x;\lambda_1,k_1) + (1-\rho)W(x;\lambda_2,k_2)\},$$

where $\theta = (\lambda_1, k_1, \lambda_2,\ k_2,\ \rho),\ 0 < x < \infty,\ \lambda_1 > 0,\ k_1 > 0,\ \sigma_2 > 0,\ k_2 > 0,\ \rho > 0,$ and $W(x;\lambda,\ k)$ is the standard Weibull distribution function $g(x,\lambda,\ k)$ is its density function. This distribution has many special cases where some parameters are known and/or related to each other. One such density is shown in the Figure 1.

We apply to our problem the general theory of the Cramér-von Mises test for the complex hypothesis that the observed random value belongs to the five-parametric family of the two Gaussian variables mixture. The limit distribution of such statistic depends of five unknown parameters. It was derived by the formula for the limit covariance function of the empirical process. Also, the new general effective method is proposed to compute the limit distribution of the considered statistics for mixed distributions. Let us now consider the general formulation of the problem of testing a parametric hypothesis in the absence of information about alternative distributions.

Let $X^n = \{X_1, X_2, ..., X_n\}$ be the sample from the r.v. with the unknown distribution function $F(x)$, $x \in R_1$. We will test the parametric hypothesis

$$H_0: \quad F(x) \in \mathcal{G} = \{G(x, \theta), \quad \theta = (\theta_1, \theta_2, ... \theta_k)^\top \in \Theta \subset R_k\}, \tag{1}$$

where θ is a vector of the parameters. We will consider the Cramér-von Mises statistic

$$\omega_n^2(\hat{\theta}_n) = n \int_{-\infty}^{\infty} (F_n(x) - G(x, \hat{\theta}_n))^2 \, dG(x, \hat{\theta}_n), \tag{2}$$

where $\hat{\theta}_n$ is the maximum likelihood estimator of θ, $F_n(x)$ is the empirical distribution function, based on X^n. The statistic ω_n^2 can be transformed to the interval $[0, 1]$ as follows

$$\omega_n^2(\hat{\theta}_n) = n \int_0^1 (\hat{F}_n(t) - t)^2 dt,$$

where $\hat{F}_n(t)$ is the estimated empirical distribution function based on the transformed sample $T^n = (t_1, \ldots, t_n)$ with $t_i = G(X_n, \hat{\theta}_n)$, $i = 1, ..., n$.

Under some regularity conditions, the following theorem holds (see [1], [4], [7])

Theorem 1. *The statistic* $\omega_n^2(\hat{\theta}_n)$ *converges in probability under* H_0 *to the functional*

$$\omega^2(\theta^0) = \int_0^1 \xi^2(t, \theta^0) dt \tag{3}$$

of the Gauss process $\xi(t, \theta^0)$ *with* $E\xi(t, \theta^0) = 0$ *and with the covariance function*

$$K(t, \tau, \theta^0) = E(\xi(t, \theta^0)\xi(\tau, \theta^0)) = (K_0(t, \tau) - q^\top(t, \theta^0) I^{-1}(\theta^0) q(\tau, \theta^0)), \tag{4}$$

where $K_0(t,\tau) = \min(t,\tau) - t\tau,\ \ t,\tau \in (0,\ 1),\ \theta^0$ *is an unknown value of the parameter* $\boldsymbol{\theta}$, *at which the sample is obtained,*

$$q(t,\theta^0) = (\partial G(x,\theta)/\partial\theta_1, ..., \partial G(x,\theta)/\partial\theta_k)|_{t=G(x,\theta^0),\theta=\theta^0}\,, \tag{5}$$

$I(\theta^0)$ *is the Fisher information matrix,*

$$I(\theta^0)\,(E((\partial/\partial\theta_i)\log g(X,\theta)(\partial/\partial\theta_j)\log g(X,\theta)))_{\theta=\theta^0,\ 1\le i,j\le k}\,, \tag{6}$$

$g(x,\theta^0) = (\partial G(x,\theta)/\partial x)_{\theta=\theta^0}.$

The following condition must be fulfilled:

$$\int_0^1 K(t,t)dt < \infty.$$

Another regularity conditions can be found in [7], [9].
The limit distribution of ω_n^2 **under** H_0 **is the distribution of the functional**

$$\omega^2 = \int_0^1 \xi^2(t,\theta^0)dt,$$

represented as the quadratic form

$$\omega^2 = \sum_{k=1}^{\infty} \frac{x_k^2}{(\pi k)^2}, \tag{7}$$

where $\{x_i,\ i=1,2,...,\}$ is the sequence of the independent standard Gaussian random variables.

2. Family of the two mixing Weibull distribution

2.1. Introduction.

We will apply Theorem 1 to derive the formula for the limit covariance function of the empirical process of the Cramér-von Mises statistics applied to the five-parametric family of the mixture of two Gaussian random variables

$$F(x,\lambda_1,k_1,\lambda_2,k_2,\rho) = \rho W(x,\mu_1,\sigma_1) + (1-\rho)W(x,\mu_2,\sigma_2), \tag{8}$$

where

$$W(x,\lambda,k) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k};$$

ρ is the mixture parameter.The corresponding to densities are

$$f(x,\lambda_1,k_1,\lambda_2,k_2,\rho) = \rho\frac{k_1}{\lambda_1}\left(\frac{x}{\lambda 1}\right)^{k_1-1} e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}} + (1-\rho)\frac{k_2}{\lambda_2}\left(\frac{x}{\lambda 2}\right)^{k_2-1} e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}};$$

We will denote the parameters vector $(\mu_1,\sigma_1,\mu_2,\sigma_2,\rho)$ by θ.

$$\theta = (\theta_1,\theta_2,\theta_3,\theta_4,\theta_5) \equiv (\lambda_1,k_1,\lambda_2,k_2,\rho).$$

2.2. Vector Q.

$$\mathbf{Q}(x,\theta)=\{\partial_{\lambda_1}F(x,\theta),\partial_{k_1}F(x,\theta),\partial_{\lambda_2}F(x,\theta),\ \partial_{k_2}F(x,\theta),\partial_{\rho}F(x,\theta)\}$$

$$=\left(-\frac{e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho}{\lambda_1},e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\ln\left(\frac{x}{\lambda 1}\right),\frac{e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)}{\lambda_2},\right.$$

$$\left.-e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)\ln\left(\frac{x}{\lambda_2}\right),-e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}+e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\right\}$$

2.3. Derivatives on parameters of density logarithm .

$$f_{\lambda_1}(x,\theta)=\partial_{\lambda 1}\ln\left(\rho\frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1}e^{-\left(\frac{x}{\lambda 1}\right)^{k_1}}+(1-\rho)\frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}e^{-\left(\frac{x}{\lambda 2}\right)^{k_2}}\right)$$

$$=\left(e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1^2\left(-1+\left(\frac{x}{\lambda_1}\right)^{k_1}\right)\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)$$

$$/\left(\lambda_1\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right);$$

$$f_{k_1}(x,\theta)=\partial_{k_1}\ln\left(\rho\frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1}e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}+(1-\rho)\frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\right)$$

$$=-\left(e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\left(-1+k_1\left(-1+\left(\frac{x}{\lambda_1}\right)^{k_1}\right)\ln\left(\frac{x}{\lambda_1}\right)\right)\right)$$

$$/\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right);$$

$$f_{\lambda_2}(x,\theta)=\partial_{\lambda_2}\ln\left(\rho\frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1}e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}+(1-\rho)\frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\right)$$
$$=\frac{\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2^2\left(-1+\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)\right)}{\left(\lambda_2\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right)};$$

$$f_{k_2}(x,\theta)=\partial_{k_2}\ln\left(\rho\frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1}e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}+(1-\rho)\frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\right)$$
$$=-\frac{\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)\left(-1+k_2\left(-1+\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\ln\left(\frac{x}{\lambda_2}\right)\right)\right)}{\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)};$$

$$f_{\rho}(x,\theta)=\partial_{\rho}\ln\left(\rho\frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1}e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}+(1-\rho)\frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\right)$$
$$=\frac{(\left(-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}+e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}\right)}{\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)};$$

2.4. Subintegral expressions.

$$S_{ij}(\theta) = \int_0^{\infty} f_i(x,\theta) f_j(x,\theta) f(x,\theta) dx.$$

We note

$$s_{ij}(x,\theta) = f_i(x,\theta) f_j(x,\theta).$$

Further:

$$s_{\lambda_1\lambda_1}(x,\theta) = (f_{\lambda_1}(x,\theta)^2 f(x,\theta)$$

$$= \left(e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}} k_1^4 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right)^2 \left(\frac{x}{\lambda_1}\right)^{-1+2k_1} \rho^2 \right)$$

$$/ \left(\lambda_1^3 \left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1} - \left(\frac{x}{\lambda 2}\right)^{k_2}} k_2 \left(\frac{x}{\lambda_2}\right)^{k_2} (-1+\rho) + k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} \rho \right) \right);$$

$$s_{\lambda_1 k_1}(x,\theta) = f_{\lambda_1}(x,\theta) f_{k_1}(x,\theta) =$$

$$- \left(\left(e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}} k_1^2 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right) \left(\frac{x}{\lambda_1}\right)^{-1+2k_1} \rho^2 \left(-1 + k_1 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right) \ln\left(\frac{x}{\lambda_1}\right)\right)\right)\right)$$

$$/ \left(\lambda_1^2 \left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1} - \left(\frac{x}{\lambda 2}\right)^{k_2}} k_2 \left(\frac{x}{\lambda_2}\right)^{k_2} (-1+\rho) + k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} \rho \right) \right);$$

$$s_{\lambda_1\lambda_2}(x,\theta) = f_{\lambda_1}(x,\theta) f_{\lambda_2}(x,\theta) f(x,\theta)$$

$$= -\frac{\left(k_1^2 k_2^2 \left(-1+\left(\frac{x}{\lambda_1}\right)^{k_1}\right)\left(\frac{x}{\lambda_1}\right)^{-1+k_1}\left(-1+\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)\rho\right)}{\left(\lambda_1^2\lambda_2\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right)};$$

$$s_{\lambda_1 k_2}(x,\theta) = f_{\lambda_1}(x,\theta) f_{k_2}(x,\theta) f(x,\theta)$$
$$= -\left(k_1^2\left(-1+\left(\frac{x}{\lambda_1}\right)^{k_1}\right)\left(\frac{x}{\lambda_1}\right)^{-1+k_1}\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)\rho\left(-1+k_2\left(-1+\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\ln\left(\frac{x}{\lambda_2}\right)\right)\right)/$$
$$\left(\lambda_1^2\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right);$$

$$s_{\lambda_1\rho}(x,\theta) = f_{\lambda_1}(x,\theta) f_{\rho}(x,\theta) f(x,\theta)$$
$$= \left(e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}}k_1^2\left(-1+\left(\frac{x}{\lambda_1}\right)^{k_1}\right)\left(\frac{x}{\lambda_1}\right)^{-1+k_1}\left(e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda 1}\right)^{k_1}-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\rho\right)\Big/$$
$$\left(\lambda_1^2\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right);$$

$$s_{k_1 k_1}(x,\theta) = (f_{k_1}(x,\theta))^2 f(x,\theta)$$

$$= \left(e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}} \left(\frac{x}{\lambda_1}\right)^{2k_1} \rho^2 \left(-1 + k_1 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right) \ln\left(\frac{x}{\lambda_1}\right)\right)^2\right)$$
$$/ \left(x \left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1} - \left(\frac{x}{\lambda_2}\right)^{k_2}} k_2 \left(\frac{x}{\lambda 2}\right)^{k_2} (-1+\rho) + k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} \rho\right)\right);$$

$$s_{k_1\lambda_2}(x,\theta) = f_{k_1}(x,\theta) f_{\lambda_2}(x,\theta) f(x,\theta)$$
$$= \left(k_2^2 \left(\frac{x}{\lambda_1}\right)^{k_1} \left(-1 + \left(\frac{x}{\lambda_2}\right)^{k_2}\right) \left(\frac{x}{\lambda_2}\right)^{-1+k_2} (-1+\rho)\rho \left(-1 + k_1 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right) \ln\left(\frac{x}{\lambda_1}\right)\right)\right) \Big/$$
$$\left(\lambda_2^2 \left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}} k_2 \left(\frac{x}{\lambda_2}\right)^{k_2} (-1+\rho) + e^{\left(\frac{x}{\lambda_2}\right)^{k_2}} k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} \rho\right)\right);$$

$$s_{k_1k_2}(x,\theta) = f_{k_1}(x,\theta) f_{k_2}(x,\theta) f(x,\theta) =$$
$$-\left(\left(\left(\frac{x}{\lambda_1}\right)^{k_1} \left(\frac{x}{\lambda_2}\right)^{k_2} (-1+\rho)\rho \left(-1 + k_1 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right) \ln\left(\frac{x}{\lambda_1}\right)\right) \left(-1 + k_2 \left(-1 + \left(\frac{x}{\lambda_2}\right)^{k_2}\right) \ln\left(\frac{x}{\lambda_2}\right)\right)\right) \Big/\right.$$
$$\left.\left(x \left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}} k_2 \left(\frac{x}{\lambda_2}\right)^{k_2} (-1+\rho) + e^{\left(\frac{x}{\lambda_2}\right)^{k_2}} k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} \rho\right)\right)\right);$$

$$s_{k_1\rho}(x,\theta) = f_{k_1}(x,\theta) f_{\rho}(x,\theta) f(x,\theta) =$$
$$-\left(\left(e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}} \left(\frac{x}{\lambda_1}\right)^{k_1} \left(e^{\left(\frac{x}{\lambda 2}\right)^{k_2}} k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} - e^{\left(\frac{x}{\lambda_1}\right)^{k_1}} k_2 \left(\frac{x}{\lambda_2}\right)^{k_2}\right) \rho \left(-1 + k_1 \left(-1 + \left(\frac{x}{\lambda_1}\right)^{k_1}\right) \ln\left(\frac{x}{\lambda_1}\right)\right)\right) \Big/\right.$$
$$\left.\left(x \left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}} k_2 \left(\frac{x}{\lambda_2}\right)^{k_2} (-1+\rho) + e^{\left(\frac{x}{\lambda_2}\right)^{k_2}} k_1 \left(\frac{x}{\lambda_1}\right)^{k_1} \rho\right)\right)\right);$$

$$s_{\lambda_2\lambda_2}(x,\theta)=(f_{\lambda_2}(x,\theta))^2 f(x,\theta)=$$
$$-\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}-\left(\frac{x}{\lambda_2}\right)^{k_2}}k_2^4\left(-1+\left(\frac{x}{\lambda 2}\right)^{k_2}\right)^2\left(\frac{x}{\lambda_2}\right)^{-1+2k_2}(-1+\rho)^2\right)$$
$$/\left(\lambda_2^3\left(e^{\left(\frac{x}{\lambda 1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right);$$

$$s_{\lambda_2 k_2}(x,\theta)=f_{\lambda_2}(x,\theta)f_{k_2}(x,\theta)f(x,\theta)=$$
$$\frac{e^{\left(\frac{x}{\lambda_1}\right)^{k_1}-\left(\frac{x}{\lambda_2}\right)^{k_2}}k_2^2\left(-1+\left(\frac{x}{\lambda 2}\right)^{k_2}\right)\left(\frac{x}{\lambda_2}\right)^{-1+2k_2}(-1+\rho)^2\left(-1+k_2\left(-1+\left(\frac{x}{\lambda 2}\right)^{k_2}\right)\ln\left(\frac{x}{\lambda_2}\right)\right)}{\lambda_2^2\left(e^{\left(\frac{x}{\lambda 1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)};$$

$$s_{\lambda_2\rho}(x,\theta)=f_{\lambda_2}(x,\theta)f_\rho(x,\theta)f(x,\theta)=$$
$$\frac{e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}k_2^2\left(-1+\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\left(-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}+e^{\left(\frac{x}{\lambda 1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\left(\frac{x}{\lambda_2}\right)^{-1+k_2}(-1+\rho)}{\lambda_2^2\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)};$$

$$s_{k_2\rho}(x,\theta)=f_{k_2}(x,\theta)f_\rho(x,\theta)f(x,\theta)=$$

$$\left(e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}}\left(-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}+e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda 2}\right)^{k_2}\right)\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)\left(-1+k_2\left(-1+\left(\frac{x}{\lambda 2}\right)^{k_2}\right)\ln\left(\frac{x}{\lambda_2}\right)\right)\right)/$$
$$\left(x\left(e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)-e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}\rho\right)\right);$$

$$s_{\rho\rho}(x,\theta)=(f_\rho(x,\theta))^2 f(x,\theta)=$$

$$\left(e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}-\left(\frac{x}{\lambda_2}\right)^{k_2}}\left(e^{\left(\frac{x}{\lambda 2}\right)^{k_2}}k_1\left(\frac{x}{\lambda_1}\right)^{k_1}-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}\right)^2\right)$$
$$/\left(x\left(-e^{\left(\frac{x}{\lambda_1}\right)^{k_1}}k_2\left(\frac{x}{\lambda_2}\right)^{k_2}(-1+\rho)+e^{\left(\frac{x}{\lambda_2}\right)^{k_2}}k_1\left(\frac{x}{\lambda 1}\right)^{k_1}\rho\right)\right);$$

2.5. Number example.

Let $theta^0=\{1,\ \ 2,\ \ 4,\ \ 5,\ \ 0.4\}$.

$$s^0_{\lambda_1\lambda_1}(x)=\frac{4e^{-x}(-1+x)^2x}{25\left(\frac{2x}{5}+\frac{3e^{x-\frac{x^5}{1024}}x^5}{1024}\right)};$$

$c_{11}=0.181950156997;$

$$s^0_{\lambda 1k1}(x)=-\frac{4096e^{-x}(-1+x)(-1+(-1+x)\ln(x))}{5\left(2048+15e^{x-\frac{x^5}{1024}}x^4\right)};$$

$c_{12}=0.01569394979863;$

$$s^0_{\lambda 1\lambda 2}(x)=\frac{15(-1+x)x^4\left(-1024+x^5\right)}{2048\left(2048e^{\frac{x^5}{1024}}+15e^x x^4\right)};$$

$$c_{13}=-0.012805329547989;$$

$$s^0_{\lambda 1k2}(x)=\frac{15(-1+x)x^4\left(-1024+x^5\right)}{2048\left(2048e^{\frac{x^5}{1024}}+15e^x x^4\right)};$$

$$c_{14}=-0.0175592859005112;$$

$$s^0_{\lambda 1r}(x)=-\frac{2e^{-x}(-1+x)\left(-1024e^{\frac{x^5}{1024}}+5e^x x^4\right)}{2048e^{\frac{x^5}{1024}}+15e^x x^4};$$

$$c_{15}=\mathbf{-0.397235483675\,;}$$

$$s^0_{k1k1}(x)=\frac{4096e^{-x}(-1+(-1+x)\ln(x))^2}{5\left(2048+15e^{x-\frac{x^5}{1024}}x^4\right)};$$

$$c_{22}=\mathbf{0.5114664745443;}$$

$$s^0_{k1\lambda 2}(x)=-\frac{15x^4\left(-1024+x^5\right)(-1+(-1+x)\ln(x))}{2048\left(2048e^{\frac{x^5}{1024}}+15e^x x^4\right)};$$

$$c_{23}=\mathbf{-0.06665998553574;}$$

$$s^0_{k1k2}(x)=\frac{3x^4\left(-1024+5\left(-1024+x^5\right)\ln\left(\frac{x}{4}\right)\right)(-1+(-1+x)\ln(x))}{2560\left(2048e^{\frac{x^5}{1024}}+15e^x x^4\right)};$$

$$c_{24}=\mathbf{0.003695357449184;}$$

$$s^0_{k1\rho}(x)=\frac{2e^{-x}\left(-1024e^{\frac{x^5}{1024}}+5e^x x^4\right)(-1+(-1+x)\ln(x))}{2048e^{\frac{x^5}{1024}}+15e^x x^4};$$

$$c_{25}=\mathbf{0.2207103604933;}$$

$$s^0_{\lambda 2\lambda 2}(x) = \frac{1125 e^{x-\frac{x^5}{1024}} x^8 \left(-1024+x^5\right)^2}{17179869184\left(2048 e^{\frac{x^5}{1024}}+15 e^x x^4\right)};$$

$$c_{33} = 0.82511437808413;$$

$$s^0_{\lambda 2k2}(x) = \frac{3x^5\left(-1+5\left(-1+\frac{x^5}{1024}\right)\ln\left(\frac{x}{4}\right)\right)(-1+(-1+x)\ln(x))}{12800\left(\frac{2}{5} e^{\frac{x^5}{1024}} x+\frac{3e^x x^5}{1024}\right)};$$

$$c_{34} = \mathbf{0.003695357449184;}$$

$$s^0_{\lambda 2r}(x) = \frac{15 e^{-\frac{x^5}{1024}} x^4\left(-1+\frac{x^5}{1024}\right)\left(-e^{\frac{x^5}{1024}} x+\frac{5e^x x^5}{1024}\right)}{4096\left(\frac{2}{5} e^{\frac{x^5}{1024}} x+\frac{3e^x x^5}{1024}\right)};$$

$$c_{35} = \mathbf{-0.1609529386844\;;}$$

$$s^0_{k2k2}(x) = \frac{9 e^{x-\frac{x^5}{1024}} x^8\left(1024-5\left(-1024+x^5\right)\ln\left(\frac{x}{4}\right)\right)^2}{5368709120\left(2048 e^{\frac{x^5}{1024}}+15 e^x x^4\right)};$$

$$c_{44} = \mathbf{0.030396877016321;}$$

$$s^0_{k2\rho}(x) = \frac{3 e^{-\frac{x^5}{1024}} x^4\left(-1024 e^{\frac{x^5}{1024}}+5 e^x x^4\right)\left(-1024+5\left(-1024+x^5\right)\ln\left(\frac{x}{4}\right)\right)}{1048576\left(2048 e^{\frac{x^5}{1024}}+15 e^x x^4\right)};$$

$$c_{45} = \mathbf{-0.06632321061149;}$$

$$s^0_{\rho\rho}(x) = \frac{5 e^{-x-\frac{x^5}{1024}}\left(1024 e^{\frac{x^5}{1024}}-5 e^x x^4\right)^2}{1024\left(2048 e^{\frac{x^5}{1024}}+15 e^x x^4\right)};$$

$$c_{55} = \mathbf{3.201293530454;}$$

3. Covariance function

Now, the following theorem derived from Theorem 1 can be formulated using the obtained above formulas for vector Q and for the elements of the Fisher information matrix:

Theorem 2. *The limit covariation function for empirical covariance function corresponding to (**??**) can be represented as*

$$\begin{aligned} C(t,\tau,\theta) &= \min(t,\tau) - t\tau \\ -Q^\top(t,\theta) & \\ &\times \left(\int_0^\infty \begin{pmatrix} s_{\mu_1\mu_1}(x,\theta) & s_{\mu_1\sigma_1}(x,\theta) & s_{\mu_1\mu_2}(x,\theta) & s_{\mu_1\sigma_2}(x,\theta) & s_{\mu_1\rho}(x,\theta) \\ & s_{\sigma_1\sigma_1}(\theta) & s_{\sigma_1\mu_2}(x,\theta) & s_{\sigma_1\sigma_2}(x,\theta) & s_{\sigma_2\rho}(x,\theta) \\ & & s_{\mu_2\mu_2}(x,\theta) & s_{\mu_2\sigma_2}(x,\theta) & s_{\mu_2\rho}(x,\theta) \\ & & & s_{\sigma_2\sigma_2}(x,\theta) & s_{\sigma_2\rho}(x,\theta) \\ & & & & s_{\rho\rho}(x,\theta) \end{pmatrix} dx \right)^{-1} \\ \times Q(\tau,\theta). & \end{aligned} \quad (9)$$

We see that the obtained limit covariance function depends on an unknown parameter vector θ. We have at our disposal only the estimated vector of parameters.

Finding the maximum likelihood estimate for the parameters of a mixture of distributions is a difficult problem. (See, for example,[5], [8]).

One example of the covariance function is shown in Figure 2.

4. Formula for calculation of the Cramér-von Mises statistic

There exist simple formulas for computing the value of the Cramér-von Mises statistic for traditional Cramér-von Mises type tests, there is a very simple formula for calculating the values of their statistic. This is the formula:

$$\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left(t_{(i)} - \frac{i - 1/2}{n} \right)^2 \tag{10}$$

where

$$t_{(i)} = G(X_{(i)}),\ X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)},$$

designed for the simple hypothesis, and

$$t_{(i)} = G(X_{(i)}, \hat{\theta}_n),\ X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)},$$

designed for the parametric hypothesis. Here $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ is an ordered sample X.

5. Computing the distribution function of the squared integral of Gaussian process

In this section we will consider a new method for calculating the limit distributions of any Cramér-von Mises type statistic if the limit covariance function of its empirical function is known.

The asymptotic distribution of the Cramér-von Mises type statistics can be represented in general by the integral

$$\omega^2 = \int_0^1 \xi^2(t)dt, \tag{11}$$

where $\xi(t)$ **is the Gaussian process having under hypothesis mathematical expectation** $E\xi(t) = 0$**. Here, we will consider the approximation of (11) by the sum**

$$\omega^{2,m} = \frac{1}{m}\sum_{i=1}^{m} \xi^2(t_i), \tag{12}$$

where $t_i = \frac{i-1/2}{m}$**. This approximation was proposed in [2]. This numerical integration formula corresponds to the Darboux method. The distribution of** $\omega^{2,m}$ **approximates with high precision the distribution of** ω^2**. Let us find the eigenvalues** $\mu_1, \mu_2, ..., \mu_m$ **of the covariance matrix**

$$K^m = \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1m} \\ K_{21} & K_{22} & \cdots & K_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ K_{m1} & K_{m2} & \cdots & K_{mm} \end{pmatrix}. \tag{13}$$

Here,

$$K_{ij} = K(t_i, \tau_j) = E\xi(t_i)\xi(t_j), i = 1, ..., m,\ j = 1, ..., m.$$

The decomposition of (12) on eigenvectors of K^m leads to the sum

$$Q = \sum_{k=1}^{m} \frac{x_k^2}{\mu_k},$$

where $x_1, x_2, ..., x_m$ are mutually independent standard normal variables. With a high degree of certainty, all eigenvalues will be different, otherwise it should slightly move the integration nodes in the formula (12). Let m be an even number. The distribution of this quadratic form can be obtained in the case of single values of $\mu_i,\ i = 1, ..., \mu$ from the Smirnov's formula

$$F(x) = 1 - \frac{1}{\pi} \sum_{k=1}^{m/2} (-1)^k \int_{\mu_{2k-1}}^{2k} \frac{e^{-x\mu/2}}{-\left(\prod_{i=1}^{\infty}(1 - \mu/\mu_i)\right)^{1/2}} \frac{d\mu}{\mu}.$$

The Smirnov's formula [10] has been studied, for example, in [6] and [7]

6. Conclusion

The results of this study indicate that applying the theory of goodness-of-fit tests to the cluster analysis problem considered in this work is entirely feasible. Furthermore, the described theory can be extended to the cases involving more than two clusters. It is also noted that in practice the requirements for the quality

of the estimation of the unknown parameters can be significantly relaxed, as the critical values of the test change relatively slowly.

Литература

[1] Anderson, T. W., Darling, D. A.: Asymptotic theory of certain "Goodness of Fit"criteria based on stochastic processes. *The Ann. of Math. Statist.*,**- 1952.- V. 23.- No. 2.- P. 193–212.**

[2] Deheuvels, P.: Martynov G.: Cramer-von-Mises-type tests with applications to test of independence for multivariate extreme-value distributions. *Communications in Statistics-Theory and Methods*,**- 1996.- Vol. 25.- No. 4.- P. 871–908.**

[3] Gikhman I. I.: On one concept from the theory of the omega-square criterion. *Mathematical journal of Kiev state university*,**- 1954.- V. 5.- P. 51-59 (in Ukrainian)**

[4] Kac, M., Kiefer, J., Wolfowitz, J.: On tests of Gaussianity and other tests of goodness-of-fit based on distance methods. *Ann. Math. Statist.*,**- 1955.- Vol. 30.- P. 420–447.**

[5] **Lemeshko B.Yu., Postovalov S.N.: Statistical analysis of mixtures of distributions for partially grouped data.** *Collection of Scientific Works of Novosibirsk State Technical University*,**- 1995.- Vol. 1, No. 1,- P 25-31. (in Russian)**

[6] **Martynov G. V.: Computation of distribution functions of quadratic forms of Gaussianly distributed random variables.** *Theory Probab. Appl.*,**- 1976.- Vol. 20.- No. 4.- P. 782–793.**

[7] **Martynov G. V.:** *Omega square test*, **Moccow: Nauka. 1978. (in Russian)**

[8] **Mezbahur Rahman and SabihaMahzaben: Maximum likelihood parameter estimation for the mixtures of normal distributions.** *Advances and applications in statistics.*,**- 2018.-Vol. 53.- No. 5.- P. 501-518.**

[9] **Neuhaus G.: Asemptotic properties of the Cramér-von-Mises statisticswhen parameters are estimated.** *Pric. Prague Symp. Asymptotic Stat. Prgue Charles Univ., 1973*,**- 1974.- Vol. 2.- P. 257–297.**

[10] **Smirnov N. V.: Sur la distribution de** ω^2**. C. R.** *Acad. Sci.* **Paris,- 1936.- V. 202.- P. 449–452.**