Construction of statistic distribution models for nonparametric goodness-of-fit tests in testing composite hypotheses: the computer approach

B.Yu. Lemeshko and S.B. Lemeshko

Abstract— **In composite hypotheses testing, when the estimate of the scalar or vector parameter of the probabilities distribution laws is calculated by the same sample, the nonparametric goodness-of-fit Kolmogorov, Cramer-Mises-Smirnov, Anderson-Darling tests lose the free distribution property. In testing of composite hypotheses, the conditional distribution law of the statistic is affected by a number of factors: the form of the observed probabilities distribution law corresponding to the true testable hypothesis; the type of the parameter estimated and the number of parameters to be estimated; sometimes, it is a specific value of the parameter (e.g., in the case of gamma-distribution and beta-distribution families); the method of parameter estimation. In this paper we present more precise results (tables of percentage points and statistic distribution models) for the nonparametric goodness-of-fit tests in testing composite hypotheses using the maximum likelihood estimate (MLE) for some probabilities distribution laws.**
Statistic distributions of the nonparametric goodness-of-fit tests are investigated by the methods of statistical simulation. Constructed empirical statistic distributions are approximated with analytical law models.

***Index Terms*—Goodness-of-fit test, Composite hypotheses testing, Kolmogorov test, Cramer-Mises-Smirnov test, Anderson-Darling test.**

I. INTRODUCTION

IN composite hypotheses testing of the form H_0: $F(x) \in \{F(x,\theta), \theta \in \Theta\}$, when the estimate $\hat{\theta}$ of the scalar or vector distribution parameter $F(x,\theta)$ is calculated by the same sample, the nonparametric goodness-of-fit Kolmogorov, ω^2 Cramer-Mises-Smirnov, Ω^2 Anderson-Darling tests lose the free distribution property.

The value

Manuscript received April 30, 2008. This work was supported in part by the Russian Foundation for Basic Research, project no. 06-01-00059a.

B.Yu. Lemeshko, Novosibirsk State Technical University, Novosibirsk, Russia. (e-mail: Lemeshkor@fpm.ami.nstu.ru).

S.B. Lemeshko, Novosibirsk State Technical University, Novosibirsk, Russia. (e-mail: skyerr@mail.ru).

$$D_n = \sup_{|x|<\infty} \left| F_n(x) - F(x,\theta) \right|,$$

where $F_n(x)$ is the empirical distribution function, n is the sample size, is used in Kolmogorov test as a distance between the empirical and theoretical laws. In testing hypotheses, a statistic with Bolshev [1] correction of the form

$$S_K = \frac{6nD_n + 1}{6\sqrt{n}}, \tag{1}$$

where $D_n = \max(D_n^+, D_n^-)$,

$$D_n^+ = \max_{1 \le i \le n} \left\{ \frac{i}{n} - F(x_i,\theta) \right\},\; D_n^- = \max_{1 \le i \le n} \left\{ F(x_i,\theta) - \frac{i-1}{n} \right\},$$

n is the sample size, $x_1, x_2, \ldots, x_n$ are sample values in increasing order is usually used. The distribution of statistic (1) in testing simple hypotheses obeys the Kolmogorov distribution law $K(S) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 s^2}$.

In ω^2 Cramer-Mises-Smirnov test, one uses a statistic of the form

$$S_\omega = n\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ F(x_i,\theta) - \frac{2i-1}{2n} \right\}^2, \tag{2}$$

and in test of Ω^2 Anderson-Darling type, the statistic of the form

$$S_\Omega = -n - 2\sum_{i=1}^{n} \left\{ \frac{2i-1}{2n} \ln F(x_i,\theta) + \left(1 - \frac{2i-1}{2n}\right) \ln(1 - F(x_i,\theta)) \right\}. \tag{3}$$

In testing a simple hypothesis, statistic (2) obeys the distribution [1] of the form

$$a1(S) = \frac{1}{\sqrt{2s}} \sum_{j=0}^{\infty} \frac{\Gamma(j+1/2)\sqrt{4j+1}}{\Gamma(1/2)\Gamma(j+1)} \exp\left\{ -\frac{(4j+1)^2}{16S} \right\} \times$$
$$\times \left\{ I_{-\frac{1}{4}}\left[\frac{(4j+1)^2}{16S} \right] - I_{\frac{1}{4}}\left[\frac{(4j+1)^2}{16S} \right] \right\}$$

where $I_{-\frac{1}{4}}(\cdot)$, $I_{\frac{1}{4}}(\cdot)$ - modified Bessel function,

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}, \quad |z| < \infty, \quad |\arg z| < \pi ,$$

and statistic (3) obeys the distribution [1] of the form

$$a2(S)=\frac{\sqrt{2\pi}}{S}\sum_{j=0}^{\infty}(-1)^j\frac{\Gamma\left(j+\frac{1}{2}\right)(4j+1)}{\Gamma\left(\frac{1}{2}\right)\Gamma(j+1)}\exp\left\{-\frac{(4j+1)^2\pi^2}{8S}\right\}\times \times\int_0^{\infty}\exp\left\{\frac{S}{8(y^2+1)}-\frac{(4j+1)^2\pi^2y^2}{8S}\right\}dy \ .$$

II. STATISTIC DISTRIBUTIONS OF THE TESTS IN TESTING COMPOSITE HYPOTHESES

In composite hypotheses testing, the conditional distribution law of the statistic $G(S|H_0)$ is affected by a number of factors: the form of the observed law $F(x,\theta)$ corresponding to the true hypothesis H_0; the type of the parameter estimated and the number of parameters to be estimated; sometimes, it is a specific value of the parameter (e.g., in the case of gamma-distribution and beta-distribution families); the method of parameter estimation. The distinctions in the limiting distributions of the same statistics in testing simple and composite hypotheses are so significant that we cannot neglect them. For example, Figure 1 shows distributions of the Anderson-Darling statistic (3) while testing the composite hypotheses subject to different laws using maximum likelihood estimates (MLE) of two parameters.

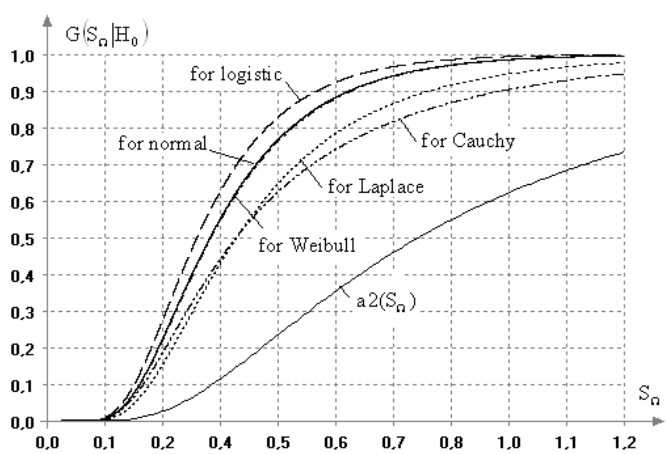

Fig. 1. The Anderson-Darling statistic (3) distributions for testing composite hypotheses with calculating MLE of two law parameters

Figure 2 illustrates the dependence of Kolmogorov test statistic (1) distribution upon the type and the number of estimated parameters by the example of *Su*-Jonson law.

The [2] was a pioneer in investigating statistic distributions of the nonparametric goodness-of-fit tests with composite hypotheses. Then, for the solution to this problem, various approaches where used [3-10].

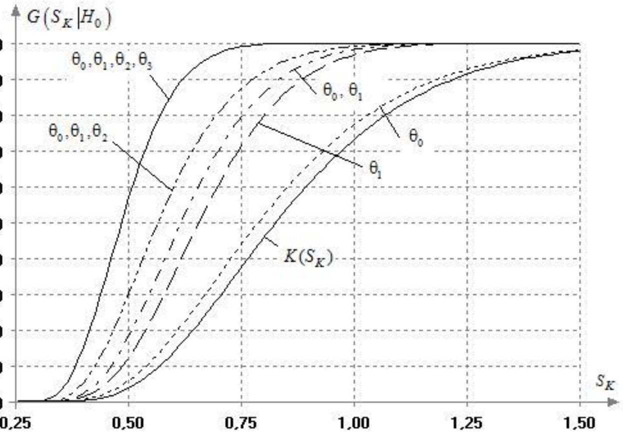

Fig. 2. The Cramer-Mises-Smirnov statistic (2) distributions for testing composite hypotheses with calculating MLE of *Su*-Jonson distribution law parameters

In our research [11-14] statistic distributions of the nonparametric goodness-of-fit tests are investigated by the methods of statistical simulating, and for constructed empirical distributions approximate models of law are found. The results obtained were used to develop recommendations for standardization R 50.1.037-2002 [15].

III. IMPROVEMENT OF STATISTIC DISTRIBUTION MODELS OF THE NONPARAMETRIC GOODNESS-OF-FIT TESTS

In this paper we present more precise results (tables of percentage points and statistic distribution models) for the nonparametric goodness-of-fit tests in testing composite hypotheses using the maximum likelihood estimate (MLE).

These investigations have been based on the developed software support. The software support allows of investigating probability regularities with the Monte-Carlo methods and constructing approximate analytical models for these regularities.

Table 1 contains a list of distributions relative to which we can test composite fit hypotheses using the constructed approximations of the limiting statistic distributions.

The tables of percentage points and statistic distributions models were constructed by modeled statistic samples with the size $N=10^6$. In this case, the samples of pseudorandom variables, belonging to $F(x,\theta)$, were generated with the size $n=10^3$.

Distributions $G(S|H_0)$ of the Kolmogorov statistic are best approximated by gamma-distributions family $\gamma(\theta_0,\theta_1,\theta_2)$ (see Table 1) or by the family of the III type beta-distributions with the density function

$$B_3(\theta_0,\theta_1,\theta_2,\theta_3,\theta_4)=\frac{\theta_2^{\theta_0}}{\theta_3\mathrm{B}(\theta_0,\theta_1)}\frac{\left(\frac{x-\theta_4}{\theta_3}\right)^{\theta_0-1}\left(1-\frac{x-\theta_4}{\theta_3}\right)^{\theta_1-1}}{\left[1+(\theta_2-1)\frac{x-\theta_4}{\theta_3}\right]^{\theta_0+\theta_1}} \ .$$

Table 1. Random variable distribution

Random variable distribution	Density function $f(x,\theta)$
Exponential	$\frac{1}{\theta_0}e^{-x/\theta_0}$
Seminormal	$\frac{2}{\theta_0\sqrt{2\pi}}e^{-x^2/2\theta_0^2}$
Rayleigh	$\frac{x}{\theta_0^2}e^{-x^2/2\theta_0^2}$
Maxwell	$\frac{2x^2}{\theta_0^3\sqrt{2\pi}}e^{-x^2/2\theta_0^2}$
Laplace	$\frac{1}{2\theta_0}e^{-\lvert x-\theta_1\rvert/\theta_0}$
Normal	$\frac{1}{\theta_0\sqrt{2\pi}}e^{-\frac{(x-\theta_1)^2}{2\theta_0^2}}$
Log-normal	$\frac{1}{x\theta_0\sqrt{2\pi}}e^{-(\ln x-\theta_1)^2/2\theta_0^2}$
Cauchy	$\frac{\theta_0}{\pi[\theta_0^2+(x-\theta_1)^2]}$
Logistic	$\frac{\pi}{\theta_0\sqrt{3}}\exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\Big/\left[1+\exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\right]^2$
Extreme-value (maximum)	$\frac{1}{\theta_0}\exp\left\{-\frac{x-\theta_1}{\theta_0}-\exp\left(-\frac{x-\theta_1}{\theta_0}\right)\right\}$
Extreme-value (minimum)	$\frac{1}{\theta_0}\exp\left\{\frac{x-\theta_1}{\theta_0}-\exp\left(\frac{x-\theta_1}{\theta_0}\right)\right\}$
Weibull	$\frac{\theta_0 x^{\theta_0-1}}{\theta_1^{\theta_0}}\exp\left\{-\left(\frac{x}{\theta_1}\right)^{\theta_0}\right\}$
Sb-Johnson $Sb(\theta_0,\theta_1,\theta_2,\theta_3)$	$\frac{\theta_1\theta_2}{(x-\theta_3)(\theta_2+\theta_3-x)}\times\exp\left\{-\frac{1}{2}\left[\theta_0-\theta_1\ln\frac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$
Sl-Johnson $Sl(\theta_0,\theta_1,\theta_2,\theta_3)$	$\frac{\theta_1}{(x-\theta_3)\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\theta_0+\theta_1\ln\frac{x-\theta_3}{\theta_2}\right]^2\right\}$
Su-Johnson $Su(\theta_0,\theta_1,\theta_2,\theta_3)$	$\frac{\theta_1}{\sqrt{2\pi}\sqrt{(x-\theta_3)^2+\theta_2^2}}\times\exp\left\{-\frac{1}{2}\left[\theta_0+\theta_1\ln\left\{\frac{x-\theta_3}{\theta_2}+\sqrt{\left(\frac{x-\theta_3}{\theta_2}\right)^2+1}\right\}\right]^2\right\}$
Gamma-distribution $\gamma(\theta_0,\theta_1,\theta_2)$	$\frac{1}{\theta_1^{\theta_0}\Gamma(\theta_0)}(x-\theta_2)^{\theta_0-1}e^{-(x-\theta_2)/\theta_1}$

And distributions of the Cramer-Mises-Smirnov and the Anderson-Darling statistics are well approximated by the family of the *Sb*-Johnson distributions $Sb(\theta_0,\theta_1,\theta_2,\theta_3)$ (see Table 1) or by the family of the III type beta-distributions $B_3(\theta_0,\theta_1,\theta_2,\theta_3,\theta_4,)$.

Upper percentage points and constructed models of the limiting statistic distributions of the Kolmogorov test when MLE are used are presented in Table 2 [for law: exponential, seminormal, Rayleigh, Maxwell, Laplace, normal, log-normal, Cauchy, logistic, extreme-value (maximum and minimum), Weibull].

Table 2. Upper percentage points and models of limiting statistic distributions of the Kolmogorov's test when MLE are used

Random variable distribution	Parameter estimated	Percentage points			Model
		0.1	0.05	0.01	
Exponential & Rayleigh	Scale	0.995	1.094	1.292	γ(5.1092; 0.0861; 0.2950)
Seminormal	Scale	1.051	1.160	1.381	γ(4.5462; 0.1001; 0.3100)
Maxwell	Scale	0.969	1.062	1.251	γ (5.4566; 0.0794; 0.2870)
Laplace	Scale	1.177	1.313	1.586	γ(3.3950; 0.1426; 0.3405)
	Shift	0.957	1.044	1.223	γ(5.1092; 0.0861; 0.2950)
	2 param-s	0.863	0.940	1.096	γ(6.2949; 0.0624; 0.2613)
Normal & Log-normal	Scale	1.191	1.327	1.600	γ(3.5609; 0.1401; 0.3375)
	Shift	0.888	0.963	1.114	γ(7.5304; 0.0580; 0.2400)
	2 param-s	0.835	0.909	1.057	γ(6.4721; 0.0580; 0.2620)
Cauchy	Scale	1.137	1.275	1.550	γ (3.0987; 0.1463; 0.3350)
	Shift	0.975	1.070	1.260	γ(5.9860; 0.0780; 0.2528)
	2 param-s	0.815	0.893	1.048	γ(5.3642; 0.0654; 0.2600)
Logistic	Scale	1.180	1.316	1.589	γ(3.4954; 0.1411; 0.3325)
	Shift	0.837	0.907	1.046	γ (7.6325; 0.0531; 0.2368)
	2 param-s	0.747	0.805	0.923	γ (7.5402; 0.0451; 0.2422)
Extreme-value & Weibull	Scale [1)]	1.182	1.316	1.583	γ(3.6805; 0.1355; 0.3350)
	Shift [2)]	0.995	1.093	1.292	γ(5.2194; 0.0848; 0.2920))
	2 param-s	0.824	0.895	1.037	γ(6.6012; 0.0563; .2598)

Note. [1)] - we estimated the Weibull distribution form parameter, [2)] - the Weibull distribution scale parameter.

For the same distribution laws upper percentage points and constructed models of the limiting statistic distributions of the Cramer-von Mises-Smirnov test are presented in Table 3, for Anderson-Darling test - in Table 4.

In Table 5 there are upper percentage points and models of limiting statistic distributions of nonparametric goodness-of-fit when MLE are used in the case of *Sb*-Johnson distribution, in Table 6 – in the case of *Sl*-Johnson distribution, in Table 7 – in the case of *Su*-Johnson distribution.

Table 3. Upper percentage points and models of limiting statistic distributions of the Cramer-Mises-Smirnov's test when MLE are used

Random variable distribution	Parameter estimated	Percentage points			Model
		0.1	0.05	0.01	
Exponential & Rayleigh	Scale	0.174	0.221	0.337	*Sb*(3.3738;1.2145;1.0792;0.011)
Seminormal	Scale	0.205	0.266	0.415	*Sb*(3.527;1.1515;1.5527;0.012)
Maxwell	Scale	0.162	0.204	0.306	*Sb*(3.353;1.220;0.9786;0.0118)
Laplace	Scale	0.323	0.438	0.719	*Sb*(3.2262;0.9416;2.703;0.015)
	Shift	0.151	0.187	0.267	*Sb*(2.9669;1.2534;0.6936;0.01)
	2 param-s	0.115	0.144	0.214	*Sb*(3.768;1.2865;0.8336;0.0113)
Normal & Log-normal	Scale	0.327	0.443	0.727	*Sb*(3.153;0.9448;2.5477;0.016)
	Shift	0.134	0.165	0.238	*Sb*(3.243;1.315;0.6826;0.0095)
	2 param-s	0.103	0.126	0.178	*Sb*(4.3950;1.4428;0.915;0.009)
Cauchy	Scale	0.316	0.430	0.711	*Sb*(3.1895;0.9134;2.690;0.013)
	Shift	0.172	0.216	0.319	*Sb*((2.359;1.0732;0.595;0.0129)
	2 param-s	0.129	0.170	0.271	*Sb*(3.4364;1.0678;1.000;0.011)
Logistic	Scale	0.323	0.438	0.719	*Sb*(3.264;0.9581;2.7046;0.014)
	Shift	0.119	0.148	0.216	*Sb*(4.0026;1.2853;1.00;0.0122)
	2 param-s	0.081	0.098	0.135	*Sb*(3.2137;1.3612;0.36;0.0105)
Extreme-value & Weibull	Scale [1)]	0.320	0.431	0.704	*Sb*(3.343;0.9817;2.753;0.015)[1)]
	Shift [2)]	0.174	0.221	0.336	*Sb*(3.498;1.2236;1.1632;0.01)[2)]
	2 param-s	0.102	0.124	0.174	*Sb*(3.3854;1.4453;0.4986;0.007)

Note. [1)] - we estimated the Weibull distribution form parameter, [2)] - the Weibull distribution scale parameter.

Table 4. Upper percentage points and models of limiting statistic distributions of the Anderson-Darling's test when MLE are used

Random variable distribution	Parameter estimated	Percentage points			Model
		0.1	0.05	0.01	
Exponential & Rayleigh	Scale	1.060	1.319	1.954	*Sb*(3.8386;1.3429;7.500;0.090)
Seminormal	Scale	1.188	1.499	2.267	*Sb*(4.2019;1.2918;11.500;0.100)
Maxwell	Scale	1.010	1.247	1.832	*Sb*(3.9591;1.3296;7.800;0.101)
Laplace	Scale	1.726	2.286	3.684	*Sb*(4.3260;1.0982;27.00;0.110)
	Shift	1.070	1.301	1.832	*Sb*(3.1506;1.3352;4.9573;0.096)
	2 param-s	0.797	0.982	1.440	*Sb*(3,8071;1,3531;5,1809;0,10)
Normal & Log-normal	Scale	1.745	2.309	3.706	*Sb*(4.3271;1.0895;28.000;0.12)
	Shift	0.892	1.087	1.551	*Sb*(3.3085;1.4043;4.2537;0.08)
	2 param-s	0.629	0.750	1.030	*Sb*(3.5601;1.4846;3.0987;0.08)
Cauchy	Scale	1.716	2.277	3.673	*Sb*(3.7830;1.0678;18.0;0.11)
	Shift	1.215	1.512	2.211	*Sb*(3.4814;1.2375;7.810;0.1)
	2 param-s	0.948	1.226	1.913	*Sb*(3.290;1.129;5.837;0.099)
Logistic	Scale	1.724	2.285	3.682	*Sb*(3.516;1.054;14.748;0.117)
	Shift	0.856	1.043	1.495	*Sb*(5.1316;1.5681;10.0;0.065)
	2 param-s	0.562	0.665	0.903	*Sb*(3.409;1.434;2.448;0.095)
Extreme-value & Weibull	Scale [1)]	1.723	2.273	3.634	*Sb*(3.512;1.064;14.496;0.125)[1)]
	Shift [2)]	1.059	1.318	1.952	*Sb*(4.799;1.402;13.0;0.085)[2)]
	2 param-s	0.634	0.755	1.040	*Sb*(3.4830;1.5138;3.00;0.07)

Note. [1)] - we estimated the Weibull distribution form parameter, [2)] - the Weibull distribution scale parameter.

Table 5. Upper percentage points and models of limiting statistic distributions of nonparametric goodness-of-fit when MLE are used in the case of *Sb*-Johnson distribution

Parameter estimated	Percentage points			Model
	0.9	0.95	0.99	
for Kolmogorov's test				
θ_0	0.888	0.963	1.115	B_3(6.3484, 7.4913, 2.3663, 1.4790, 0.27)
θ_1	1.189	1.326	1.600	B_3(6.8242, 4.7737, 5.2621, 2.3878, 0.27)
θ_0 & θ_1	0.836	0.909	1.058	B_3(6.6559,8.1766, 2.9405, 1.6143, 0.27)
for Cramer-Mises-Smirnov's test				
θ_0	0.134	0.165	0.238	B_3(4.2304,3.806,13.1934,0.6908,0.0086)
θ_1	0.327	0.442	0.724	B_3(2.9153,2.0048,33.416,2.0782,0.0114)
θ_0 & θ_1	0.104	0.126	0.179	B_3(4.3897,4.0574,12.101,0.5119,0.0086)
for Anderson-Darling's test				
θ_0	0.893	1.087	1.553	B_3(4.2657,4.3788,11.4946,4.655,0.084)
θ_1	1.741	2.309	3.702	B_3(4.1703,2.3363,42.083,12.602,0.088)
θ_0 & θ_1	0.631	0.751	1.034	B_3(4.0891,5.9708,9.6497,4.000,0.082)

Table 6. Upper percentage points and models of limiting statistic distributions of nonparametric goodness-of-fit when MLE are used in the case of *Sl*-Johnson distribution

Parameter estimated	Percentage points			Model
	0.9	0.95	0.99	
for Kolmogorov's test				
θ_0	0.888	0.963	1.115	B_3(6.3484,7.4913,2.3663,1.4790,0.27)
θ_1	1.189	1.326	1.600	B_3(6.8242,4.7737,5.2621,2.3878,0.27)
θ_2	0.888	0.963	1.115	B_3(6.3484,7.4913,2.3663,1.4790,0.27)
θ_0,θ_1	0.836	0.909	1.058	B_3(6.6559,8.1766,2.9405,1.6143,0.27)
θ_0,θ_2	0.888	0.963	1.115	B_3(6.3484,7.4913,2.3663,1.4790,0.27)
θ_1,θ_2	0.836	0.909	1.058	B_3(6.6559,8.1766,2.9405,1.6143,0.27)
$\theta_0,\theta_1,\theta_2$	0.836	0.909	1.058	B_3(6.6559,8.1766,2.9405,1.6143,0.27)
for Cramer-Mises-Smirnov's test				
θ_0	0.134	0.165	0.238	B_3(4.2304,3.8058,13.193,0.6908,0.0086)
θ_1	0.327	0.442	0.724	B_3(2.9153,2.0048,33.414,2.0782,0.0114)
θ_2	0.134	0.165	0.238	B_3(4.2304,3.8058,13.193,0.6908,0.0086)
θ_0,θ_1	0.104	0.126	0.179	B_3(4.3897,4.0574,12.101,0.5119,0.0086)
θ_0,θ_2	0.134	0.165	0.238	B_3(4.2304,3.8058,13.193,0.6908,0.0086)
θ_1,θ_2	0.104	0.126	0.179	B_3(4.3897,4.0574,12.101,0.5119,0.0086)
$\theta_0,\theta_1,\theta_2$	0.104	0.126	0.179	B_3(4.3897,4.0574,12.101,0.5119,0.0086)
for Anderson-Darling's test				
θ_0	0.893	1.087	1.553	B_3(4.2657,4.3788,11.4946,4.6551,0.084)
θ_1	1.741	2.309	3.702	B_3(4.1703,2.3363,42.083,12.6019,0.088)
θ_2	0.893	1.087	1.553	B_3(4.2657,4.3788,11.4946,4.6551,0.084)
θ_0,θ_1	0.631	0.751	1.034	B_3(4.0891,5.9708,9.6497,4.0000,0.082)
θ_0,θ_2	0.893	1.087	1.553	B_3(4.2657,4.3788,11.4946,4.6551,0.084)
θ_1,θ_2	0.631	0.751	1.034	B_3(4.0891,5.9708,9.6497,4.0000,0.082)
$\theta_0,\theta_1,\theta_2$	0.631	0.751	1.034	B_3(4.0891,5.9708,9.6497,4.0000,0.082)

Table 7. Upper percentage points and models of limiting statistic distributions of nonparametric goodness-of-fit when MLE are used in the case of *Su*-Johnson distribution

Parameter estimated	Percentage points			Model
	0.9	0.95	0.99	
θ_0	0.888	0.963	1.115	B3(6.3484,7.4913,2.3663,1.4790,0.27)
θ_1	1.189	1.326	1.600	B3(6.8242,4.7737,5.2621,2.3878,0.27)
θ_2	1.161	1.300	1.576	B3(5.3417,4.6440,4.7448,2.3802,0.29)
θ_3	0.880	0.960	1.122	B3(6.6252,7.4025,3.0590,1.6516,0.27)
θ_0, θ_1	0.836	0.909	1.058	B3(6.4792,7.0243,2.8437,1.4260,0.27)
θ_0, θ_2	0.798	0.872	1.024	B3(6.4496,6.7714,3.3119,1.4226,0.27)
θ_0, θ_3	0.802	0.875	1.023	B3(6.3069,6.1065,3.2916,1.3317,0.27)
θ_1, θ_3	1.142	1.282	1.561	B3(5.9751,4.4559,5.6810,2.4123,0.27)
θ_1, θ_3	0.792	0.858	0.994	B3(6.4839,7.0152,2.7376,1.2838,0.27)
θ_2, θ_3	0.733	0.791	0.910	B3((6.2438,6.9161,2.5011,1.0904,0.27)
$\theta_0, \theta_1, \theta_2$	0.776	0.851	1.007	B3(6.2414,6.4027,3.7458,1.4361,0.27)
$\theta_0, \theta_1, \theta_3$	0.720	0.780	0.901	B3(6.4262,6.9732,2.7325,1.1317,0.26)
$\theta_0, \theta_2, \theta_3$	0.658	0.706	0.806	B3(6.1239,7.9516,2.24033,0.9839,0.26)
$\theta_1, \theta_2, \theta_3$	0.704	0.760	0.878	B3(7.1354,8.0363,2.7466,1.1766,0.25)
$\theta_0, \theta_1, \theta_2, \theta_3$	0.622	0.666	0.755	B3(6.6889,8.1712,2.3857,0.9291,0.25)
θ_0	0.134	0.165	0.238	B3(3.6736,3.9355,11.2146,0.6908,0.01)
θ_1	0.327	0.442	0.724	B3(2.9153,2.0048,33.414,2.078,0.0114)
θ_2	0.318	0.433	0.716	B3(2.2077,1.7250,28.4959,1.75,0.015)
θ_3	0.125	0.154	0.225	B3(3.6990,3.8775,11.9942,0.6601,0.01)
θ_0, θ_1	0.104	0.126	0.179	B3(4.3897,4.0574,12.101,0.512,0.0086)
θ_0, θ_2	0.090	0.110	0.161	B3(5.203,3.9325,15.697,0.4659,0.0075)
θ_0, θ_3	0.104	0.133	0.203	B3(5.954,3.1023,30.6943,0.638,0.0071)
θ_1, θ_2	0.314	0.428	0.711	B3(2.4905,1.6985,45.967,2.3084,0.012)
θ_1, θ_3	0.094	0.113	0.158	B3(4.6011,5.7370,19.1580,1.0,0.0075)
θ_2, θ_3	0.080	0.096	0.137	B3(4.7686,4.6085,11.142,0.393,0.0075)
$\theta_0, \theta_1, \theta_2$	0.083	0.104	0.155	B3(5.2574,3.644,19.921,0.4707,0.0075)
$\theta_0, \theta_1, \theta_3$	0.071	0.086	0.122	B3(5.775,4.7935,18.118,0.4777,0.0065)
$\theta_0, \theta_2, \theta_3$	0.056	0.066	0.089	B3(7.350,5.4726,13.745,0.2883,0.0052)
$\theta_1, \theta_2, \theta_3$	0.073	0.089	0.130	B3(5.6379,4.0985,18.5518,0.421,0.007)
$\theta_0, \theta_1, \theta_2, \theta_3$	0.048	0.056	0.075	B3(6.9739,6.6406,13.743,0.315,0.0052)
θ_0	0.893	1.087	1.553	B3(4.2329,4.5369,10.8807,4.655,0.082)
θ_1	1.741	2.309	3.702	B3(4.1703,2.3363,42.083,12.603,0.088)
θ_2	1.707	2.275	3.667	B3(2.6348,1.9774,21.3842,7.75,0.125)
θ_3	0.952	1.161	1.649	B3(3.5597,4.9656,11.4180,6.520,0.092)
θ_0, θ_1	0.631	0.751	1.034	B3(4.0891,5.9708,9.6497,4.0000,0.082)
θ_0, θ_2	0.577	0.689	0.961	B3(5.5368,4.9114,13.1278,3.0625,0.07)
θ_0, θ_3	0.737	0.920	1.386	B3(5.6629,3.4912,25.1600,4.5052,0.07)
θ_1, θ_2	1.666	2.232	3.627	B3(3.8896,1.6253,31.1820,5.80,0.09)
θ_1, θ_3	0.694	0.842	1.200	B3(4.6199,5.2874,19.2708,6.561,0.074)
θ_2, θ_3	0.642	0.935	1.140	B3(4.4276,4.30288,14.669,3.7865,0.08)
$\theta_0, \theta_1, \theta_2$	0.518	0.627	0.898	B3(5.5158,4.3512,14.770,2.6199,0.067)
$\theta_0, \theta_1, \theta_3$	0.454	0.536	0.733	B3(5.3306,5.8858,10.751,2.5087,0.065)
$\theta_0, \theta_2, \theta_3$	0.396	0.459	0.606	B3(5.7098,6.8325,7.9837,1.8803,0.06)
$\theta_1, \theta_2, \theta_3$	0.585	0.729	1.087	B3(5.1840,3.2993,19.364,2.7865,0.073)
$\theta_0, \theta_1, \theta_2, \theta_3$	0.329	0.378	0.489	B3(7.1015,5.8708,7.1323,1.0517,0.05)

IV. IMPROVEMENT OF STATISTIC DISTRIBUTION MODELS OF THE NONPARAMETRIC GOODNESS-OF-FIT TESTS IN THE CASE OF GAMMA-DISTRIBUTION

In composite hypotheses testing subject to gamma-distribution with the density function

$$f(x,\theta) = \frac{x^{\theta_0 - 1}}{\theta_1^{\theta_0}\Gamma\ \theta_0}\exp\left(-\frac{x}{\theta_1}\right)$$

limiting statistics distributions of the nonparametric goodness-of-fit tests depend on values of the form parameter θ_0. For example, Figure 3 illustrates dependence of the Kolmogorov statistic distribution upon the value θ_0 in testing a composite hypothesis in the case of calculating MLE for the scale parameter of gamma-distribution only.

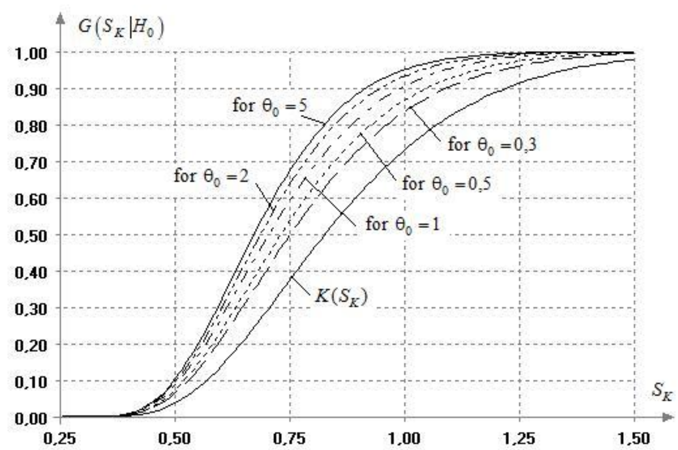

Fig. 3. The Kolmogorov statistic (1) distributions for testing composite hypotheses with calculating MLE of scale parameter depend on the form parameter value of gamma-distribution

Upper percentage points constructed as a result of statistical modeling and constructed models of limiting statistic distributions of the Kolmogorov test when MLE are used in the case of gamma-distribution are given in Table 8, for Cramer-von Mises-Smirnov test – in Table 9, for Anderson-Darling test - in Table 10. In this case statistic distributions are well approximated by the family of the III type beta-distributions $B_3\ \theta_0,\theta_1,\theta_2,\theta_3,\theta_4,\ .$

The upper percentage points and the models of statistic distributions, presented at Tables 8-10, improve the results given in recommendations [15].

Table 8. Upper percentage points and models of limiting statistic distributions of the Kolmogorov's test when MLE are used in the case of gamma-distribution

Value of the form parameter	Parameter estimated	Percentage points			Model
		0.1	0.05	0.01	
0.3	Scale	1.096	1.211	1.444	B_3(6.687;4.837;4.4047;1.944;0.281)
	form	0.976	1.070	1.262	B_3(6.4536;5.7519;3.310;1.650;0.28)
	2 param-s	0.905	0.990	1.162	B_3(6.9705;5.678;3.6297;1.507;0.27)
0.5	Scale	1.051	1.160	1.379	B_3(6.9356;5.008;4.3582;1.847;0.28)
	form	0.961	1.052	1.236	B_3(6.386;5.9685;3.123;1.6154;0.28)
	2 param-s	0.884	0.965	1.131	B_3(6.408;5.934;3.206;1.448;0.2774)
1.0	Scale	0.994	1.095	1.299	B_3(6.719;5.374;3.7755;1.688;0.282)
	form	0.936	1.022	1.191	B_3(6.1176;6.4704;2.6933;1.55;0.28)
	2 param-s	0.862	0.940	1.097	B_3(5.603;6.129;2.7065;1.361;0.29)
2.0	Scale	0.952	1.044	1.228	B_3(5.8359;22.6032;2.192;4.0;0.282)
	form	0.915	0.995	1.155	B_3(6.1387;6.5644;2.602;1.484;0.28)
	2 param-s	0.849	0.924	1.077	B_3(5.832;6.145;2.755;1.328;0.2862)
3.0	Scale	0.933	1.020	1.200	B_3(5.9055;24.431;2.0996;4.0;0.282)
	form	0.906	0.985	1.140	B_3(6.1221;6.613;2.5536;1.459;0.28)
	2 param-s	0.845	0.919	1.070	B_3(6.039;6.1276;2.831;1.32;0.2837)
4.0	Scale	0.923	1.008	1.181	B_3(5.9419;27.1264;1.915;4.000.282)
	form	0.901	0.980	1.132	B_3(6.083;6.7095;2.496;1.4494;0.28)
	2 param-s	0.843	0.916	1.066	B_3(6.1584;6.119;2.875;1.317;0.281)
5.0	Scale	0.917	1.000	1.170	B_3(5.8774;30.0692;1.720;4.0;0.282)
	form	0.899	0.977	1.127	B_3(6.089;6.7265;2.4894;1.443;0.28)
	2 param-s	0.842	0.915	1.063	B_3(6.196;6.1114;2.8894;1.314;0.28)

Table 9. Upper percentage points and models of limiting statistic distributions of the Cramer-Mises-Smirnov's test when MLE are used in the case of gamma-distribution

Value of the form parameter	Parameter estimated	Percentage points			Model
		0.1	0.05	0.01	
0.3	Scale	0.233	0.305	0.482	B_3(3.272;1.9595;16.177;0.75;0.013)
	form	0.166	0.209	0.316	B_3(3.025;3.226;11.11;0.776;0.0125)
	2 param-s	0.127	0.158	0.233	B_3(2.361;4.084;7.061;0.619;0.0145)
0.5	Scale	0.205	0.264	0.413	B_3(3.230;2.1984;14.315;0.70;0.013)
	form	0.159	0.199	0.298	B_3(3.014;3.35;10.095;0.721;0.0125)
	2 param-s	0.119	0.146	0.212	B_3(2.7216;3.984;7.499;0.537;0.013)
1.0	Scale	0.175	0.220	0.336	B_3(3.1201;2.546;11.120;0.60;0.013)
	form	0.149	0.186	0.273	B_3(2.993;3.472;8.828;0.635;0.0125)
	2 param-s	0.111	0.136	0.194	B_3(3.00;3.896;7.3247;0.4508;0.012)
2.0	Scale	0.155	0.193	0.288	B_3(2.9463;3.1124;9.116;0.60;0.013)
	form	0.142	0.176	0.256	B_3(2.991;3.533;8.20;0.5786;0.0125)
	2 param-s	0.107	0.131	0.185	B_3(3.053;3.940;7.117;0.425;0.0118)
3.0	Scale	0.148	0.184	0.272	B_3(2.884;3.3796;8.4342;0.60;0.013)
	form	0.139	0.172	0.251	B_3(2.974;3.553;7.884;0.555;0.0125)
	2 param-s	0.106	0.129	0.182	B_3(3.070;3.962;7.034;0.416;0.0117)
4.0	Scale	0.145	0.179	0.264	B_3(2.852;3.5285;8.1044;0.60;0.013)
	form	0.138	0.171	0.248	B_3(2.968;3.543;7.763;0.542;0.0125)
	2 param-s	0.105	0.128	0.180	B_3(3.097;3.954;7.064;0.412;0.0116)
5.0	Scale	0.142	0.176	0.259	B_3(2.8249;3.628;7.8756;0.60;0.013)
	form	0.137	0.169	0.246	B_3(2.964;3.547;7.656;0.533;0.0125)
	2 param-s	0.105	0.128	0.179	B_3(4.433;3.626;10.55;0.410;0.0084)

Table 10. Upper percentage points and models of limiting statistic distributions of the Anderson- Darling's test when MLE are used in the case of gamma-distribution

Value of the form parameter	Parameter estimated	Percentage points			Model
		0.1	0.05	0.01	
0.3	Scale	1.300	1.655	2.543	B_3(3.385;2.883;14.684;6.042;0.109)
	form	1.021	1.258	1.865	B_3(3.1073;3.704;8.672;4.344;0.112)
	2 param-s	0.718	0.870	1.233	B_3(4.532;4.060;10.072;2.921;0.078)
0.5	Scale	1.183	1.490	2.260	B_3(5.005;2.936;18.852;5.244;0.077)
	form	0.993	1.221	1.791	B_3(3.1104;3.729;8.068;4.013;0.112)
	2 param-s	0.684	0.824	1.145	B_3(5.008;4.056;10.029;2.587;0.073)
1.0	Scale	1.058	1.313	1.955	B_3(5.031;3.185;15.463;4.380;0.077)
	form	0.952	1.166	1.696	B_3(3.115;3.792;7.4813;3.677;0.112)
	2 param-s	0.657	0.785	1.084	B_3(5.0034;4.109;9.161;2.343;0.073)
2.0	Scale	0.980	1.203	1.771	B_3(4.945;3.375;13.043;3.830;0.077)
	form	0.922	1.125	1.625	B_3(3.0434;4.162;7.1516;3.85;0.112)
	2 param-s	0.643	0.766	1.051	B_3(4.924;4.209;8.664;2.2754;0.073)
3.0	Scale	0.952	1.163	1.702	B_3(5.037;3.413;12.901;3.687;0.077)
	form	0.912	1.110	1.601	B_3(3.056;3.909;6.7844;3.397;0.112)
	2 param-s	0.639	0.761	1.043	B_3(4.9475;4.207;8.669;2.251;0.073)
4.0	Scale	0.937	1.141	1.662	B_3(4.943;3.504;12.224;3.630;0.077)
	form	0.906	1.103	1.590	B_3(3.053;3.944;6.762;3.3993;0.112)
	2 param-s	0.637	0.758	1.039	B_3(4.9274;4.228;8.557;2.239;0.073)
5.0	Scale	0.927	1.130	1.640	B_3(4.881;3.576;11.789;3.605;0.077)
	form	0.902	1.099	1.586	B_3(3.050;3.964;6.751;3.4024;0.112)
	2 param-s	0.636	0.757	1.037	B_3(4.921;4.243;8.488;2.2314;0.073)

V. CONCLUSION

In this paper we present more precise models of statistic distributions of the nonparametric goodness-of-fit tests in testing composite hypotheses subject to some laws considered in recommendations for standardization R 50.1.037-2002 [15]. The models of statistic distributions of the nonparametric goodness-of-fit tests for testing composite hypotheses of the Exponential family [14] were made more precise earlier and haven't been improved now.

In the case of the I, II, III type beta-distribution families' statistic distributions depend on a specific value of two form parameter of these distributions. Statistic distributions models and tables of percentage points for various combinations of values of two form parameters (more than 1500 models) were constructed in the thesis of Lemeshko S.B. [16] and partly were published in the paper [17].

The results of comparative analysis of goodness-of-fit tests power (nonparametric and χ^2 type) subject to some sufficiently close pair of alternative are presented in [18].

ACKNOWLEDGMENT

This research was supported by the Russian Foundation for Basic Research, project no. 06-01-00059a.

REFERENCES

[1] Bolshev L.N., Smirnov N.V. Tables of Mathematical Statistics. Moscow: Science, 1983. (in Russian)

[2] Kac M., Kiefer J., Wolfowitz J. On Tests of Normality and Other Tests of Goodness of Fit Based on Distance Methods. *Ann. Math. Stat.*, vol. 26, 1955, pp. 189-211.

[3] Durbin J. Kolmogorov–Smirnov Test when Parameters are Estimated. *Lect. Notes Math*. No. 566, 1976, pp. 33–44.

[4] Martinov G.V. Omega-Quadrate Tests. Moscow: Science, 1978. (in Russian)

[5] Pearson E.S., Hartley H.O. Biometrica Tables for Statistics, vol 2. Cambridge: University Press, 1972.

[6] Stephens M.A. Use of Kolmogorov–Smirnov, Cramer – von Mises and Related Statistics – Without Extensive Table. *J. R. Stat. Soc*., vol. 32, 1970, pp. 115-122.

[7] Stephens M.A. EDF Statistics for Goodness of Fit and Some Comparisons. *J. Am. Statist. Assoc*., vol.69, 1974. pp. 730-737.

[8] Chandra M., Singpurwalla N.D., Stephens M.A. Statistics for Test of Fit for the Extrem–Value and Weibull Distribution. *J. Am. Statist. Assoc*. 76(375), 1981, pp. 729-731.

[9] Tyurin Yu.N. On the Limiting Kolmogorov-Smirnov Statistic Distribution for Composite Hypothesis. *News of the AS USSR. Ser. Math*., vol. 48(6), 1984, pp. 1314-1343. (in Russian)

[10] Tyurin Yu.N., Savvushkina N.E. Goodness-of-Fit Test for Weibull-Gnedenko Distribution. *News of the AS USSR. Ser. Techn. Cybernetics*, no. 3, 1984, pp. 109-112. (in Russian)

[11] Lemeshko B.Yu., Postovalov S.N. Statistical Distributions of Non-parametric Goodness-of-Fit Tests as Estimated by the Sample Parameters of Experimentally Observed Laws. *Industrial laboratory*, vol. 64, no. 3, 1998, pp. 197-208.

[12] Lemeshko B.Yu., Postovalov S.N. Application of the Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses. *Optoelectronics, Instrumentation and Data Processing*, vol. 37, no. 2, 2001, pp. 76-88.

[13] Lemeshko B.Yu., Postovalov S.N. The nonparametric goodness-of-fit tests about fit with Johnson distributions in testing composite hypotheses // News of the SB AS HS, no. 1(5), 2002, pp. 65-74. (in Russian)

[14] Lemeshko B.Yu., Maklakov A.A. Nonparametric Test in Testing Composite Hypotheses on Goodness of Fit Exponential Family Distributions. *Optoelectronics, Instrumentation and Data Processing*, vol. 40, no. 3, pp. 3-18.

[15] R 50.1.037-2002. Recommendations for Standardization. Applied statistics. Rules of check of experimental and theoretical distribution of the consent. Part II. Nonparametric goodness–of–fit test. Moscow: Publishing house of the standards, 2002. (in Russian)

[16] Lemeshko S.B. Expansion of applied opportunities of some classical methods of mathematical statistics. The dissertation on competition of a scientific degree of Cand. Tech. Sci. Novosibirsk State Technical University. Novosibirsk, 2007. (in Russian)

[17] Lemeshko B.Yu., Lemeshko S.B. Statistic distributions of the nonparametric goodness-of-fit tests in testing hypotheses relative to beta-distributions. // News of the SB AS HS, no.2(9), 2007, pp. 6-16. (in Russian)

[18] Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N. (2007). The power of goodness-of-fit tests for close alternatives. *Measurement Techniques* , vol. 50, no. 2, 2007, pp. 132-141.