Measurement Techniques, Vol. 52, No. 6, 2009

GENERAL PROBLEMS OF METROLOGY AND MEASUREMENT TECHNIQUE

DISTRIBUTION MODELS FOR NONPARAMETRIC TESTS FOR FIT IN VERIFYING COMPLICATED HYPOTHESES AND MAXIMUM-LIKELIHOOD ESTIMATORS. PART 1

B. Yu. Lemeshko and S. B. Lemeshko UDC 519.24

Complicated-hypothesis testing is considered when the scalar or vector parameters of the probability distribution are calculated for a single sample. In that case, various factors influence the distributions of the statistics for the nonparametric tests for fit of Kolmogorov, Cramer–Mises–Smirnov, and Anderson–Darling. Revised results are given (tables of the percentage points and distribution models) for nonparametric tests of fit for verifying complicated hypotheses concerning a series of distributions when one uses maximum-likelihood estimators.

Key words: *fit tests, complicated hypothesis testing, Kolmogorov test, Cramer–Mises–Smirnov test, Anderson–Darling test.*

Practices in using nonparametric fit tests in statistical analysis give many examples of incorrect use of classical results, which apply on testing simple hypotheses, to a situation corresponding to testing complicated ones.

When one tests complicated hypotheses of the form H_0: $F(x) \in \{F(x, \theta), \theta \in \Theta\}$, where the estimator $\hat{\theta}$ for the scalar or vector parameters of the distribution $F(x, \theta)$ is calculated on the same sample, then the Kolmogorov fit test, the ω^2 Cramer–Mises–Smirnov test, and the Ω^2 Anderson–Darling test lose the property of freedom from the distribution.

In the Kolmogorov test, the distance between the empirical and theoretical distributions is defined by

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

in which $F_n(x)$ is the empirical distribution and n the sample volume. For $n \to \infty$, the distribution of the statistic $D_n\sqrt{n}$ if the test hypothesis is correct converges uniformly to the Kolmogorov distribution [1]:

$$K(S) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 s^2} .$$

When one tests hypotheses on the Kolmogorov test, it is recommended to use a statistic with the Bol'shev correction [2, 3] in the form [4]:

$$S_K = (6nD_n + 1)/6\sqrt{n}, \tag{1}$$

Novosibirsk State Technical University, Novosibirsk, Russia; e-mail: Lemeshko@fpm.ami.nstu.ru. Translated from Izmeritel'naya Tekhnika, No. 6, pp. 3–11, June, 2009. Original article submitted March 27, 2009.

in which

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

$x_1, x_2, ..., x_n$ are elements of the sample ranked in rising order. If the tested hypothesis is correct, the (1) statistic also fits a Kolmogorov distribution [4], and the convergence to it is substantially more rapid than for the statistic $D_n\sqrt{n}$.

It is very regrettable that the Bol'shev correction has not attracted attention of foreign experts. Papers on the Kolmogorov test up to now as a rule have used $D_n\sqrt{n}$, and therefore with restricted values of n it is necessary to incorporate the substantial dependence of the statistic distribution on this quantity.

In the ω^2 Cramer–Mises–Smirnov test, the statistic takes the form [4]

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

while a test of Anderson–Darling Ω^2 type [5, 6] takes the form

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

On testing simple hypotheses, the (2) statistic has the following distribution [4]:

$$a_1(S) = \frac{1}{\sqrt{2s}}\sum_{j=0}^{\infty}\frac{\Gamma(j+1/2)\sqrt{4j+1}}{\Gamma(1/2)\Gamma(j+1)}\exp\left\{-\frac{(4j+1)^2}{16S}\right\}\left\{I_{-1/4}\left[\frac{(4j+1)^2}{16S}\right] - I_{1/4}\left[\frac{(4j+1)^2}{16S}\right]\right\},$$

in which $I_{-1/4}(\cdot)$, $I_{1/4}(\cdot)$ are modified Bessel functions;

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

and the statistic of (3) has the distribution [4]

$$a_2(S) = \frac{\sqrt{2\pi}}{S}\sum_{j=0}^{\infty}(-1)^j\frac{\Gamma(j+1/2)(4j+1)}{\Gamma(1/2)\Gamma(j+1)}\exp\left\{-\frac{(4j+1)^2\pi^2}{8S}\right\}\int_0^\infty \exp\left\{\frac{S}{8(y^2+1)} - \frac{(4j+1)^2\pi^2y^2}{8S}\right\}dy.$$

Statistic Distributions on Testing Complicated Hypotheses. In that case, the conditional distributions of the statistic $G(S\,|\,H_0)$ are dependent on various factors: the form of the observed law $F(x, \theta)$ corresponding to correctness for the test hypothesis H_0; the types and numbers of the estimated parameters; the particular value of a parameter (in some cases for example for the families of the gamma and beta distributions); and the method of estimating the parameters. The differences between the limiting distributions for the same statistic on testing simple and complicated hypotheses are so substantial that one cannot neglect them in any case. For example, Fig. 1 shows the distribution of the Anderson–Darling statistic (3) on testing complicated hypotheses concerning various distributions with the use of maximum-likelihood estimators (MLE) for two parameters. Figure 2 illustrates the dependence of the distribution for the Kolmogorov test statistic (1) on the type and number of the estimated parameters by reference to the *Su*-Johnson distribution.

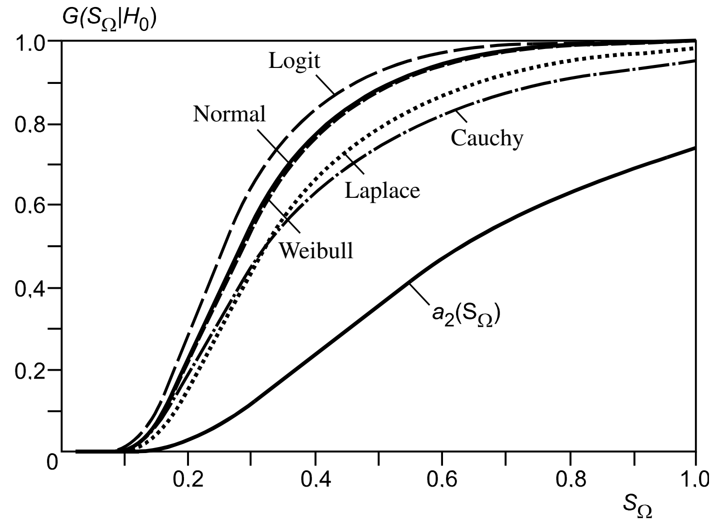

Fig. 1. Distributions for the Anderson–Darling statistic in testing complicated hypotheses based on maximum likelihood estimators for two parameters.

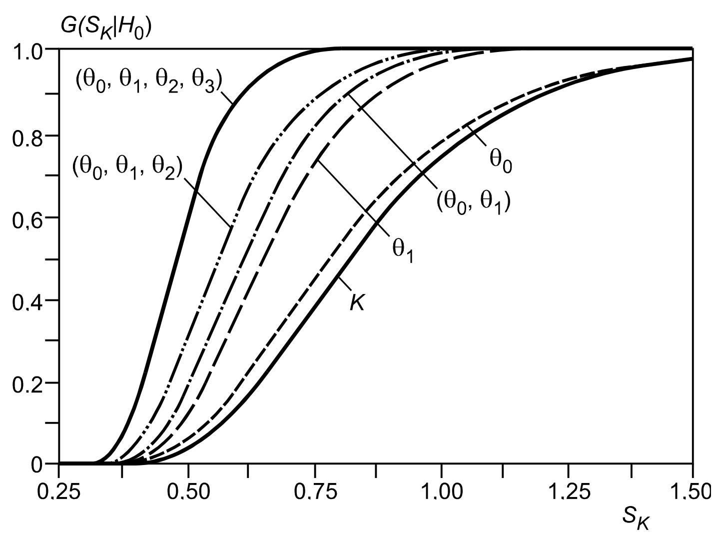

Fig. 2. Distributions of the Kolmogorov statistic on testing complicated hypotheses with the calculation of maximum-likelihood estimators for the parameters of the *Su*-Johnson distribution.

Paper [7] is the starting point for examining the limiting distributions for nonparametric tests of fit with complicated hypotheses. Subsequently, the task was handled by using various approaches: limiting distributions were examined by analytical methods [8–20]; the percentage points of the distributions were derived by Monte Carlo simulation [21–24]; and formulas were derived giving reasonably good approximations for small values of the corresponding probabilities [25–27].

Monte Carlo simulation has been used [28–35] to examine the distributions of statistics for nonparametric fitting tests; then the empirical distributions were used in constructing approximate analytic models for them. Some of the results were used in preparing the recommendations of [36].

Refinement of Models for the Statistic Distributions for Nonparametric Fit Tests. Here we present revised results (tables of the percentage points and models for their distributions) from nonparametric fitting tests in testing compli-

TABLE 1. Random-Quantity Distributions

Name	Density $f(x, \theta)$
Exponential	$\theta_0^{-1}\exp(-x/\theta_0)$
Seminormal	$\frac{2}{\theta_0\sqrt{2\pi}}\exp\left(-x^2/2\theta_0^2\right)$
Rayleigh	$\frac{x}{\theta_0^2}\exp\left(-x^2/2\theta_0^2\right)$
Maxwell	$\frac{2x^2}{\theta_0^3\sqrt{2\pi}}\exp\left(-x^2/2\theta_0^2\right)$
Laplace	$(2\theta_0)^{-1}\exp\left(-\lvert x-\theta_1\rvert/\theta_0\right)$
Normal (Gauss)	$\frac{1}{\theta_0\sqrt{2\pi}}\exp\left(-\frac{(x-\theta_1)^2}{2\theta_0^2}\right)$
Log-normal	$\frac{1}{x\theta_0\sqrt{2\pi}}\exp\left(-\frac{(\ln x-\theta_1)^2}{2\theta_0^2}\right)$
Cauchy	$\theta_0\left\{\pi\left[\theta_0^2+(x-\theta_1)^2\right]\right\}^{-1}$
Logit	$\frac{\pi}{\theta_0\sqrt{3}}\exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\Big/\left[1+\exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\right]^2$
Extreme value (max)	$\frac{1}{\theta_0}\exp\left\{-\frac{x-\theta_1}{\theta_0}-\exp\left(-\frac{x-\theta_1}{\theta_0}\right)\right\}$
Extreme value (min)	$\frac{1}{\theta_0}\exp\left\{\frac{x-\theta_1}{\theta_0}-\exp\left(\frac{x-\theta_1}{\theta_0}\right)\right\}$
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 $Sb(\theta_0, \theta_1, \theta_2, \theta_3)$	$\frac{\theta_1\theta_2}{(x-\theta_3)(\theta_2+\theta_3-x)}\exp\left\{-\frac{1}{2}\left[\theta_0-\theta_1\ln\frac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$
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 $Su(\theta_0, \theta_1, \theta_2, \theta_3)$	$\frac{\theta_1}{\sqrt{2\pi}\sqrt{(x-\theta_3)^2+\theta_2^2}}\exp\left\{-\frac{1}{2}\left[\theta_0+\theta_1\ln\left\{\frac{x-\theta_3}{\theta_2}+\sqrt{\left(\frac{x-\theta_3}{\theta_2}\right)^2+1}\right\}\right]^2\right\}$
Gamma distribution $\gamma(\theta_0, \theta_1, \theta_2)$	$\frac{1}{\theta_1^{\theta_0}\Gamma(\theta_0)}(x-\theta_2)^{\theta_0-1}\exp\left(-\frac{x-\theta_2}{\theta_1}\right)$
Two-sided exponential	$\frac{\theta_0}{2\theta_1\Gamma(1/\theta_0)}\exp\left\{-\left(\frac{\lvert x-\theta_2\rvert}{\theta_1}\right)^{\theta_0}\right\}$

TABLE 2. Upper Percentage Points and Limiting Distribution Models for Kolmogorov Test Statistics

Law	Estimated parameter	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	Ditto	1.051	1.160	1.381	γ(4.5462; 0.1001; 0.3100)
Maxwell	»	0.969	1.062	1.251	γ(5.4566; 0.0794; 0.2870)
Laplace	Scale	1.177	1.313	1.587	B_3(4.4680; 4.8450; 3.9105; 2.3784; 0.324)
	Shift	0.957	1.045	1.223	B_3(5.3541; 7.2519; 2.5630; 1.7652; 0.302)
	Both parameters	0.863	0.940	1.095	γ(6.2949; 0.0624; 0.2613)
Normal, log-normal	Ditto	1.190	1.327	1.600	B_3(4.8849; 5.2341; 3.6279; 2.3872; 0.303)
		0.888	0.963	1.114	B_3(5.2604; 7.4327; 2.1872; 1.4774; 0.30)
		0.835	0.909	1.057	γ(6.4721; 0.0580; 0.2620)
Cauchy	»	1.137	1.275	1.550	γ(3.0987; 0.1463; 0.3350)
		0.975	1.070	1.260	γ(5.9860; 0.0780; 0.2528)
		0.815	0.893	1.048	γ(5.3642; 0.0654; 0.2600)
Logit	»	1.180	1.316	1.589	γ(3.4954; 0.1411; 0.3325)
		0.837	0.907	1.046	γ(7.6325; 0.0531; 0.2368)
		0.746	0.805	0.923	γ(7.5402; 0.0451; 0.2422)
Extremal values, Weibull	»	1.182	1.316	1.583	γ(3.6805; 0.1355; 0.3350)*
		0.995	1.093	1.292	γ(5.2194; 0.0848; 0.2920)**
		0.824	0.895	1.037	γ(6.6012; 0.0563; 0.2598)

*, ** on estimating shape and scale parameters of a Weibull distribution, respectively.

cated hypotheses when MLE are used. The data are based on software enabling one to examine probability regularities by Monte Carlo simulation and construct approximate analytic models for them.

Table 1 contains a list of distributions that can be used in testing complicated hypotheses on the basis of the approximations for the limiting distributions for nonparametric fitting tests, together with their densities.

The tables for the percentage points and distribution models have been constructed by means of sample sets of volume $N = 10^6$ for the statistics. With such N, the difference between the true distribution $G(S|H_0)$ and the simulated empirical one $G_N(S|H_0)$ does not exceed 10^{-3} in modulus. With that value for the tests, the calculations were performed with sets of random numbers generated in accordance with the observed law $F(x, \theta)$ of volume $n = 10^3$. In that case the distribution $G(S_n|H_0)$ virtually coincides with the limiting distribution $G(S|H_0)$; in statistical analysis, one can use these models beginning with sample volumes $n > 25$.

The $G(S|H_0)$ distributions for the Kolmogorov statistic are best approximated by the families of gamma distributions $\gamma(\theta_0, \theta_1, \theta_2)$ (Table 1) or beta distributions of the third kind with density

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

TABLE 3. Upper Percentage Points and Limiting Distribution Models for Cramer–Mises–Smirnov Test

Law	Estimated parameter	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	Ditto	0.205	0.266	0.415	*Sb*(3.527; 1.1515; 1.5527; 0.012)
Maxwell	»	0.162	0.204	0.306	*Sb*(3.353; 1.220; 0.9786; 0.0118)
Laplace	Scale	0.323	0.439	0.719	B_3(3.9800; 1.4667; 38.0035; 1.13; 0.0111)
	Shift	0.152	0.187	0.268	B_3(3.3130; 3.8338; 10.097; 0.7517; 0.011)
	Both parameters	0.115	0.144	0.213	B_3(4.489; 3.7706; 17.577; 0.7065; 0.0085)
Normal, log-normal	Ditto	0.327	0.442	0.725	*Sb*(3.153; 0.9448; 2.5477; 0.016)
		0.134	0.165	0.238	B_3(4.433; 3.6365; 13.920; 0.6632; 0.0084)
		0.103	0.126	0.178	B_3(4.1153; 4.1748; 11.035; 0.5116; 0.009)
Cauchy	»	0.315	0.430	0.711	*Sb*(3.1895; 0.9134; 2.690; 0.013)
		0.172	0.216	0.319	*Sb*(2.359; 1.0732; 0.595; 0.0129)
		0.129	0.170	0.271	*Sb*(3.4364; 1.0678; 1.000; 0.011)
Logit	»	0.323	0.438	0.719	*Sb*(3.264; 0.9581; 2.7046; 0.014)
		0.119	0.148	0.216	*Sb*(4.0026; 1.2853; 1.00; 0.0122)
		0.081	0.098	0.135	*Sb*(3.2137; 1.3612; 0.36; 0.0105)
Extremal values, Weibull	»	0.320	0.431	0.704	*Sb*(3.343; 0.9817; 2.753; 0.015)*
		0.174	0.221	0.336	*Sb*(3.498; 1.2236; 1.1632; 0.01)**
		0.102	0.124	0.174	*Sb*(3.3854; 1.4453; 0.4986; 0.007)

*, ** the same as in Table 2.

TABLE 4. Upper Percentage Points and Limiting Distribution Models for Anderson–Darling Test

Law	Estimated parameter	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	Ditto	1.188	1.499	2.267	*Sb*(4.2019; 1.2918; 11.500; 0.100)
Maxwell	»	1.009	1.247	1.832	*Sb*(3.9591; 1.3296; 7.800; 0.1010)
Laplace	Scale	1.725	2.290	3.685	B_3(4.0842; 1.7532; 28.1434; 6.00; 0.105)
	Shift	1.071	1.302	1.837	B_3(4.0842; 1.7532; 28.1434; 6.00; 0.105)
	Both parameters	0.798	0.982	1.439	B_3(5.3576; 3.8690; 17.2148; 4.2386; 0.073)
Normal, log-normal	Ditto	1.743	2.309	3.704	B_3(3.4638; 2.330; 35.7115; 12.603; 0.105)
		0.892	1.087	1.552	B_3(4.1081; 5.0598; 16.9721; 7.9065; 0.09)
		0.630	0.750	1.032	B_3(4.7262; 4.6575; 9.4958; 2.717; 0.0775)
Cauchy	»	1.716	2.277	3.673	*Sb*(3.7830; 1.0678; 18.0; 0.11)
		1.215	1.512	2.211	*Sb*(3.4814; 1.2375; 7.810; 0.1)
		0.948	1.226	1.913	*Sb*(3.290; 1.129; 5.837; 0.099)
Logit	»	1.724	2.285	3.682	*Sb*(3.516;1.054; 14.748; 0.117)
		0.856	1.043	1.495	*Sb*(5.1316; 1.5681; 10.0; 0.065)
		0.562	0.665	0.903	*Sb*(3.409; 1.434; 2.448; 0.095)
Extremal values, Weibull	»	1.723	2.273	3.634	*Sb*(3.512; 1.064; 14.496; 0.125)*
		1.059	1.318	1.952	*Sb*(4.799; 1.402; 13.0; 0.085)**
		0.634	0.755	1.040	*Sb*(3.4830; 1.5138; 3.00; 0.07)

*, ** the same as in Tables 2 and 3.

TABLE 5. Upper Percentage Points and Limiting Distribution Models for the Nonparametric Fitting Test Statistics in the Case of Testing Hypotheses with Respect to *Sb*-Johnson Distributions

Estimated quantity	Percentage points			Model
	0.9	0.95	0.99	
Kolmogorov 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)
Cramer–Mises–Smirnov test				
θ_0	0.134	0.165	0.238	B_3(4.2304; 3.8058; 13.1934; 0.6908; 0.0086)
θ_1	0.327	0.442	0.724	B_3(2.9153; 2.0048; 33.4135; 2.07821; 0.0114)
θ_0, θ_1	0.104	0.126	0.179	B_3(4.3897; 4.0574; 12.1009; 0.5119; 0.0086)
Anderson–Darling test				
θ_0	0.893	1.086	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; 42.0833; 12.6019; 0.088)
θ_0, θ_1	0.631	0.751	1.034	B_3(4.0891; 5.9708; 9.6497; 4.0000; 0.082)

TABLE 6. Upper Percentage Points and Limiting Distribution Models for the Nonparametric Fitting Test Statistics in the Case of Testing Hypotheses with Respect to *Sl*-Johnson Distributions

Estimated quantity	Percentage points			Model
	0.9	0.95	0.99	
Kolmogorov 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)
θ_0, θ_1, θ_2	0.836	0.909	1.058	B_3(6.6559; 8.1766; 2.9405; 1.6143; 0.27)
Cramer–Mises–Smirnov test				
θ_0	0.134	0.165	0.238	B_3(4.2304; 3.8058; 13.1934; 0.6908; 0.0086)
θ_1	0.327	0.442	0.724	B_3(2.9153; 2.0048; 33.4135; 2.07821; 0.0114)
θ_2	0.134	0.165	0.238	B_3(4.2304; 3.8058; 13.1934; 0.6908; 0.0086)
θ_0, θ_1	0.104	0.126	0.179	B_3(4.3897; 4.0574; 12.1009; 0.5119; 0.0086)
θ_0, θ_2	0.134	0.165	0.238	B_3(4.2304; 3.8058; 13.1934; 0.6908; 0.0086)
θ_1, θ_2	0.104	0.126	0.179	B_3(4.3897; 4.0574; 12.1009; 0.5119; 0.0086)
θ_0, θ_1, θ_2	0.104	0.126	0.179	B_3(4.3897; 4.0574; 12.1009; 0.5119; 0.0086)
Anderson–Darling test				
θ_0	0.893	1.086	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; 42.0833; 12.6019; 0.088)
θ_2	0.893	1.086	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	1.086	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)
θ_0, θ_1, θ_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 Limiting Distribution Models for the Nonparametric Fitting Test Statistics in the Case of Testing Hypotheses with Respect to *Su*-Johnson Distributions

Estimated quantity	Percentage points			Model
	0.9	0.95	0.99	
Kolmogorov 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	1.161	1.300	1.576	B_3(5.3417; 4.6440; 4.7448; 2.3802; 0.29)
θ_3	0.880	0.960	1.122	B_3(6.6252; 7.4025; 3.0590; 1.6516; 0.27)
θ_0, θ_1	0.836	0.909	1.058	B_3(6.4792; 7.0243; 2.8437; 1.4260; 0.27)
θ_0, θ_2	0.798	0.872	1.024	B_3(6.4496; 6.7714; 3.3119; 1.4226; 0.27)
θ_0, θ_3	0.802	0.875	1.023	B_3(6.3069; 6.1065; 3.2916; 1.3317; 0.27)
θ_1, θ_2	1.142	1.282	1.561	B_3(5.9751; 4.4559; 5.6810; 2.4123; 0.27)
θ_1, θ_3	0.792	0.858	0.994	B_3(6.4839; 7.0152; 2.7376; 1.2838; 0.27)
θ_2, θ_3	0.733	0.791	0.910	B_3(6.2438; 6.9161; 2.5011; 1.0904; 0.27)
θ_0, θ_1, θ_2	0.776	0.851	1.007	B_3(6.2414; 6.4027; 3.7458; 1.4361; 0.27)
θ_0, θ_1, θ_3	0.720	0.780	0.901	B_3(6.4262; 6.9732; 2.7325; 1.1317; 0.26)
θ_0, θ_2, θ_3	0.658	0.706	0.806	B_3(6.1239; 7.9516; 2.24033; 0.9839; 0.26)
θ_1, θ_2, θ_3	0.704	0.760	0.878	B_3(7.1354; 8.0363; 2.7466; 1.1766; 0.25)
θ_0, θ_1, θ_2, θ_3	0.622	0.666	0.755	B_3(6.6889; 8.1712; 2.3857; 0.9291; 0.25)
Cramer–Mises–Smirnov test				
θ_0	0.134	0.165	0.238	B_3(3.6736; 3.9355; 11.2146; 0.6908; 0.01)
θ_1	0.327	0.442	0.724	B_3(2.9153; 2.0048; 33.4135; 2.07821; 0.0114)
θ_2	0.318	0.433	0.716	B_3(2.2077; 1.7250; 28.4959; 1.75; 0.015)
θ_3	0.125	0.154	0.225	B_3(3.6990; 3.8775; 11.9942; 0.6601; 0.01)
θ_0, θ_1	0.103	0.126	0.178	B_3(4.3897; 4.0574; 12.1009; 0.5119; 0.0086)
θ_0, θ_2	0.090	0.110	0.161	B_3(5.2030; 3.9325; 15.6968; 0.4659; 0.0075)
θ_0, θ_3	0.104	0.133	0.203	B_3(5.9540; 3.1023; 30.6943; 0.6380; 0.0071)
θ_1, θ_2	0.314	0.428	0.711	B_3(2.4905; 1.6985; 45.9674; 2.3084; 0.012)
θ_1, θ_3	0.094	0.113	0.158	B_3(4.6011; 5.7370; 19.1580; 1.0; 0.0075)
θ_2, θ_3	0.080	0.096	0.137	B_3(4.7686; 4.6085; 11.1421; 0.3929; 0.0075)
θ_0, θ_1, θ_2	0.083	0.104	0.155	B_3(5.2574; 3.6440; 19.9213; 0.4707; 0.0075)
θ_0, θ_1, θ_3	0.071	0.086	0.122	B_3(5.7750; 4.7935; 18.1182; 0.4777; 0.0065)
θ_0, θ_2, θ_3	0.056	0.066	0.089	B_3(7.3500; 5.4726; 13.7452; 0.2883; 0.0052)
θ_1, θ_2, θ_3	0.073	0.089	0.130	B_3(5.6379; 4.0985; 18.5518; 0.42100; 0.007)
θ_0, θ_1, θ_2, θ_3	0.048	0.056	0.075	B_3(6.9739; 6.6406; 13.7433; 0.3151; 0.0052)

Estimated quantity	Percentage points			Model
	0.9	0.95	0.99	
Anderson–Darling test				
θ_0	0.892	1.087	1.552	B_3(4.2329; 4.5369; 10.8807; 4.6551; 0.082)
θ_1	1.743	2.309	3.704	B_3(4.1703; 2.3363; 42.0833; 12.6019; 0.088)
θ_2	1.707	2.275	3.667	B_3(2.6348; 1.9774; 21.3842; 7.75; 0.125)
θ_3	0.952	1.161	1.648	B_3(3.5597; 4.9656; 11.4180; 6.5202; 0.092)
θ_0, θ_1	0.630	0.750	1.032	B_3(4.0891; 5.9708; 9.6497; 4.0000; 0.082)
θ_0, θ_2	0.576	0.689	0.961	B_3(5.5368; 4.9114; 13.1278; 3.0625; 0.07)
θ_0, θ_3	0.737	0.920	1.386	B_3(5.6629; 3.4912; 25.1600; 4.5052; 0.07)
θ_1, θ_2	1.666	2.232	3.627	B_3(3.8896; 1.6253; 31.1820; 5.80; 0.09)
θ_1, θ_3	0.694	0.842	1.200	B_3(4.6199; 5.2874; 19.2708; 6.5610; 0.074)
θ_2, θ_3	0.642	0.935	1.140	B_3(4.4276; 4.30288; 14.6688; 3.7865; 0.08)
θ_0, θ_1, θ_2	0.518	0.627	0.898	B_3(5.5158; 4.3512; 14.7750; 2.6199; 0.067)
θ_0, θ_1, θ_3	0.454	0.536	0.733	B_3(5.3306; 5.8858; 10.7581; 2.5087; 0.065)
θ_0, θ_2, θ_3	0.395	0.459	0.606	B_3(5.7098; 6.8325; 7.9837; 1.8803; 0.06)
θ_1, θ_2, θ_3	0.585	0.729	1.087	B_3(5.1840; 3.2993; 19.3614; 2.7865; 0.073)
θ_0, θ_1, θ_2, θ_3	0.329	0.378	0.488	B_3(7.1015; 5.8708; 7.1323; 1.0517; 0.05)

The distributions for the Cramer–Mises–Smirnov and Anderson–Darling statistics are closely approximated by families of *Sb*-Johnson distributions $Sb(\theta_0, \theta_1, \theta_2, \theta_3)$ (see Table 1) or beta distributions of the third kind $B_3(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$, as for the Kolmogorov test.

Table 2 gives the upper percentage points and models for the limiting distributions of the Kolmogorov test statistic when MLE are used for the parameters of the various laws; and Tables 3 and 4 are correspondingly for the Cramer–Mises–Smirnov test and the Anderson–Darling one. Analogous data are given in Tables 5–7 for the distributions *Sb*-Johnson (on using MLE), *Sl*-Johnson, and *Su*-Johnson, respectively. In all these cases, the distributions of $G(S \mid H_0)$ for the fitting test statistics are not dependent on the particular values of the unknown parameters of $F(x, \theta)$.

Example. We test a complicated hypothesis on the assignment of a sample to a normal law. The ranked sample of volume 200 observations takes the form

–1.270	–1.196	–1.043	–1.018	–1.011	–0.929	–0.916	–0.892	–0.886	–0.877	–0.827	–0.801
–0.791	–0.782	–0.776	–0.766	–0.757	–0.736	–0.723	–0.714	–0.701	–0.700	–0.677	–0.673
–0.667	–0.658	–0.627	–0.615	–0.615	–0.615	–0.604	–0.602	–0.572	–0.567	–0.565	–0.556
–0.543	–0.542	–0.482	–0.475	–0.468	–0.162	–0.458	–0.450	–0.448	–0.437	–0.433	–0.416
–0.410	–0.108	–0.403	–0.382	–0.370	–0.367	–0.365	–0.340	–0.337	–0.331	–0.326	–0.322
–0.309	–0.304	–0.288	–0.282	–0.271	–0.252	–0.245	–0.234	–0.224	–0.224	–0.211	–0.205
–0.178	–0.143	–0.131	–0.129	–0.127	–0.117	–0.115	–0.108	–0.095	–0.093	–0.078	–0.051
–0.044	–0.035	–0.035	–0.034	–0.031	–0.012	–0.012	–0.010	–0.004	0.006	0.008	0.019

0.023	0.037	0.045	0.055	0.056	0.083	0.086	0.092	0.101	0.104	0.114	0.121
0.122	0.124	0.136	0.141	0.141	0.142	0.147	0.150	0.151	0.188	0.194	0.204
0.206	0.211	0.211	0.228	0.231	0.232	0.242	0.252	0.252	0.257	0.270	0.278
0.283	0.294	0.319	0.321	0.336	0.342	0.343	0.361	0.382	0.384	0.404	0.404
0.404	0.405	0.408	0.435	0.468	0.474	0.480	0.482	0.510	0.514	0.533	0.541
0.543	0.552	0.580	0.582	0.582	0.587	0.589	0.600	0.611	0.617	0.327	0.641
0.643	0.643	0.645	0.651	0.668	0.677	0.678	0.679	0.693	0.714	0.737	0.737
0.749	0.750	0.756	0.762	0.780	0.784	0.808	0.809	0.827	0.853	0.853	0.866
0.877	0.883	0.902	0.960	0.994	1.048	1.075	1.114				

For the sample with the use of MLE, we calculated the parameters of the normal distribution for scale $\hat{\theta}_0 = 0.5383$ and shift $\hat{\theta}_1 = 0.0294$.

The value of the (1) Kolmogorov statistic is $S_K^* = 0.8152$ and is less than the critical value $S_{K;0.1} = 0.835$ (see Table 2), i.e., the tested hypothesis is not declined for significance level $\alpha = 0.1$. The values calculated in accordance with the gamma distribution $\gamma(6.4721; 0.0580; 0.2620)$ (Table 2) represent a level of significance $P\{S_K > S_K^*\} \approx 0.119$.

The value of the (2) Kramer–Mises–Smirnov statistic $S_\omega^* = 0.1451$ is greater than 0.126 but less than 0.178 (see Table 3); the tested hypothesis is not declined for significance level $\alpha = 0.01$. The values calculated in accordance with the beta distribution of the third type $B_3(4.1153; 4.1748; 11.035; 0.5116; 0.009)$ (Table 3) represent a level of significance $P\{S_\omega > S_\omega^*\} \approx 0.029$.

The value of the (3) Anderson–Darling statistic $S_\Omega^* = 1.0157$ is greater than 0.750 but less than 1.032 (see Table 4), and one calculates in accordance with a beta distribution of the third type $B_3(4.7262; 4.6575; 9.4958; 2.7174; 0.0775)$ that the significance level attained is $P\{S_\Omega > S_\Omega^*\} \approx 0.0117$. Then the tested hypothesis is not declined at the significance level $\alpha = 0.01$.

This sample was generated on a symmetrical law substantially differing from normal.

In conclusion, we note as follows. We have used computer methods of examining statistical regularities, which are based on Monte Carlo simulation of empirical statistic distributions followed by analysis of them and construction of more precise models for the distribution of statistics for nonparametric fitting tests of Kolmogorov type, ω^2 Cramer–Mises–Smirnov type, and Ω^2 Anderson–Darling type in testing complicated hypotheses on the number of probability distributions. These results revise and extend the recommendations of [36].

In the next part of the paper, we present distribution models for the test statistics and tables of the percentage points for testing complicated hypotheses with respect to families of gamma distributions and two-sided exponential ones (Table 1).

The researches have been performed with partial support from the Russian Foundation for Basic Research (Grant No. 06-01-00059-a) and from the Federal Agency for Education of the Ministry of Education and Science of the Russian Federation within the framework of the analytical domestic target program “Development of the scientific potential of higher schools” (Project No. 2.1.2/3970).

REFERENCES

1. A. N. Kolmogoroff, *Giornale dell' Istituto Italiano degli Attuari*, **4**, No. 1, 83 (1933).
2. L. N. Bol'shev, “Asymptotic Pearson transformations,” *Teor. Veroyatn. Ee Primen.*, **8**, No. 2, 129 (1963).
3. L. N. Bol'shev, “Probability theory and mathematical statistics,” in: *Selected Works* (edited by Yu. V. Prokhorov) [in Russian], Nauka, Moscow (1987).
4. L. N. Bol'shev and N. V. Smirnov, *Tables for Mathematical Statistics* [in Russian], Nauka, Moscow (1983).
5. T. W. Anderson and D. A. Darling, *Ann. Math. Statist.*, **23**, 193 (1952).
6. T. W. Anderson and D. A. Darling, *J. Amer. Statist. Assoc.*, **29**, 765 (1954).
7. M. Kac, J. Kiefer, and J. Wolfowitz, *Ann. Math. Statist.*, **26**, 189 (1955).

8. D. A. Darling, *Ann. Math. Statist.*, 26, 1 (1955).
9. D. A. Darling, *Ann. Math. Statist.*, 28, 823 (1957).
10. H. W. Lilliefors, *J. Amer. Statist. Assoc.*, **62**, 399 (1967).
11. H. W. Lilliefors, *J. Amer. Statist. Assoc.*, **64**, 387 (1969).
12. J. Durbin, *Ann. Math. Statist.*, **1**, 279 (1973).
13. J. Durbin, *Biometrika*, **62**, 5 (1975).
14. J. Durbin, *Lect. Notes Math.*, **566**, 33 (1976).
15. I. I. Gikhman, *Dokl. AN SSSR*, **91** (4), 715 (1953).
16. I. I. Gihman, *Selected Translation in Mathematical Statistics and Probability: Abstr. Amer. Math. Soc.*, Providence, RI, Vol. 1 (1961), p. 77.
17. G. V. Martynov, *The Omega Squared Test* [in Russian], Nauka, Moscow (1978).
18. K. O. Dzhaparidze and M. S. Nikulin, *J. Soviet. Math.*, **20**, 2147 (1982).
19. M. S. Nikulin, *Math. Rep. of the Academy of Science of the Royal Society of Canada*, **14** (4), 151 (1992).
20. M. S. Nikulin, *Zapiski Nauchnykh Seminarov LOMI*, **177**, 108 (1989); *J. Soviet Math.*, **61** (4), 1896 (1992).
21. E. S. Pearson and H. O. Hartley, *Biometric Tables for Statistics*, Vol. 2, University Press, Cambridge (1972).
22. M. A. Stephens, *J. R. Stat. Soc.*, **32**, 115 (1970).
23. M. A. Stephens, *J. Am. Statist. Assoc.*, **69**, 730 (1974).
24. M. Chandra, N. D. Singpurwalla, and M. A. Stephens, *J. Am. Statist. Assoc.*, **76** (375), 729 (1981).
25. Yu. N. Tyurin, *Izv. AN SSSR, Ser. Matematika*, **48**, No. 6, 1314 (1984).
26. Yu. N. Tyurin and N. E. Savvushkina, *Izv. AN SSSR, Ser. Tekhnich. Kibernet.*, No. 3, 109 (1984).
27. N. E. Savvushkina, *Trudy VNII Sistem. Issled.*, No. 8, 50 (1990).
28. B. Yu. Lemeshko and S. N. Postovalov, *Nadezhn. Kontrol Kach.*, No. 11, 3 (1997).
29. B. Yu. Lemeshko and S. N. Postovalov, *Zav. Lab.*, **64**, No. 3, 61 (1998); *Ind. Lab.*, **64**, No. 3, 197 (1998).
30. B. Yu. Lemeshko and S. N. Postovalov, *Methods of Quality Management: Reliability and Quality Control*, No. 11 [in Russian] (1999), p. 34.
31. B. Yu. Lemeshko and S. N. Postovalov, *Applied Statistics: Rules for Checking Fit of an Experimental Distribution to a Theoretical One. Recommendations on Methods. Part 2: Nonparametric Tests* [in Russian], Izd. NGTU, Novosibirsk (1999).
32. B. Yu. Lemeshko and S. N. Postovalov, *Avtometriya*, No. 2, 88 (2001); *Optoelectr., Instrum., and Data Proces.*, No. 2, 76 (2001).
33. B. Yu. Lemeshko and S. N. Postovalov, *Dokl. SO AN VSh*, No. 1 (5), 65 (2002).
34. B. Yu. Lemeshko and A. A. Maklakov, *Avtometriya*, No. 3, 3 (2004); *Optoelectr., Instrum., and Data Proces.*, **40**, No. 3, 3 (2004).
35. V. I. Denisov et al., *Design of Experiments and Statistical Analysis for Grouped Observations*, NSTU Publishing House, Novosibirsk (2004).
36. R 50.1.037-2002, *Recommendations on Standardizing Applied Statistics. Rules for Fit Testing of Experimental Distributions to Theoretical Ones. Part 2. Nonparametric Tests* [in Russian].