DOI 10.1007/s10958-022-06101-6
Journal of Mathematical Sciences, Vol. 267, No. 1, October, 2022

GOODNESS-OF-FIT TEST BASED ON AN UNBIASED ESTIMATOR OF THE DISTRIBUTION FUNCTION IN THE CASE OF EXPONENTIAL DISTRIBUTION

V. V. Chichagov[1]

For exponential distribution, a modification of the Kolmogorov goodness-of-fit test is considered. This modification consists in the replacement of the empirical distribution function by the optimal unbiased estimator of the distribution function. Properties of the modified test statistics are described that make it possible to calculate its distribution. The power of the modified test is compared with that of the Kolmogorov test for various values of input parameters.

1. Introduction

Let $\overrightarrow{X_n} = (X_1, \ldots, X_n)$ be a repeated sample of size n from the population ξ. The sample elements are assumed to be independent random variables with one and the same known distribution function $F(x;\theta)$ with the unknown parameter $\theta \in \Theta$. Consider the problem of testing the simple hypothesis $H_0: \ F(x;\theta) = F_0(x)$, where $F_0(x) = F(x;\theta_0)$, θ_0 is a value of θ.

There are many goodness-of-fit tests for solving this problem, including the Kolmogorov test $\mathbf{K}$ based on the statistic

$$D_n = \sup_{x \in \mathbf{R}} |F_n(x) - F_0(x)|,$$

where $F_n(x)$ is the empirical distribution function constructed from the sample $\overrightarrow{X_n}$.

It is well known that the empirical distribution function is an unbiased estimator of the corresponding theoretical distribution function. In some cases (for example, if the distribution of ξ belongs to the exponential family), there exists an unbiased estimator of the distribution function that is a function of a sufficient statistic whose dimensionality is less than n. The variance of this unbiased estimator is less than that of the empirical distribution function. So, one can expect that the test $\widehat{\mathbf{K}}$ based on the statistic

$$\widehat{D_n} = \sup_{x \in \mathbf{R}} \left| \widehat{F}(x|S_n) - F_0(x) \right|$$

may be more powerful than the Kolmogorov test, where $\widehat{F}(x|S_n)$ is an unbiased estimator of the distribution function of the random variable ξ, S_n is a sufficient statistic for the parameter θ.

Define the test $\widehat{\mathbf{K}}$ in a way similar to the Kolmogorov test: if $\widehat{D}_n < c_{\alpha,n}$, then the null hypothesis is accepted. Here, as usual, $c_{\alpha,n}$ is the critical value of the test statistic correponding to the significance level α and sample size n.

As a result, the following problems need to be solved: the calculation of the critical value $c_{\alpha,n}$; the calculation of the power function of the test $\widehat{\mathbf{K}}$. In the present paper the solution of both of these problems is presented under the assumption that the random variable ξ has the exponential distribution with the distribution function

$$F(x,\theta) = 1 - e^{-x/\theta}, \quad \theta > 0, \ x \geqslant 0. \tag{1}$$

[1] Perm State University, Perm, Russia, e-mail: `chichagov@psu.ru`

Translated from *Statisticheskie Metody Otsenivaniya i Proverki Gipotez*, Vol. 22, pp. 43–56, 2010

2. Main theoretical results

Lemma 1. *If $\overrightarrow{X_n}$ is a repeated sample of size $n \geqslant 2$ from the population ξ having exponential distribution* (1)*, then*:

1) *the test statistic of $\widehat{\mathbf{K}}$ has the form*

$$\widehat{D_n} = \widehat{D_n}\,(S_n) = \sup_{t \geqslant 0} \left| e^{-t} - \left(1 - \frac{t\,\theta_0}{S_n}\right)_+^{n-1} \right|, \tag{2}$$

where $S_n = X_1 + \ldots + X_n$, $(x)_+ = \max\{0, x\}$;

2) *the value of the power function $W\left(k|\widehat{\mathbf{K}}\right)$ of the test $\widehat{\mathbf{K}}$ corresponding to the value $\theta = \theta_1 = k\theta_0$ of the parameter is determined by the relation*

$$W\left(k|\widehat{\mathbf{K}}\right) = \mathbf{P}\left(\sup_{t \geqslant 0} \left| e^{-t} - \left(1 - \frac{t}{k\,S_n^*}\right)_+^{n-1} \right| \geqslant c_{\alpha,n}\right), \quad k > 0, \tag{3}$$

where S_n^* is a random variable with the Erlang distribution with shape parameter n and scale parameter 1 defined by the density

$$f_{S_n^*}(s) = \frac{s^{n-1}}{\Gamma(n)}\, e^{-s}, \quad s > 0.$$

Moreover, $W\left(1|\widehat{\mathbf{K}}\right) = \alpha$.

Proof. With the account of the fact that an unbiased estimator of distribution function (1) being a function of the sufficient statistic $S_n = X_1 + \ldots + X_n$ has the form (e.g., see [1])

$$F(x|S_n) = 1 - \left(1 - \frac{x}{S_n}\right)_+^{n-1},$$

make simple transformations

$$\widehat{D_n}\,(S_n) = \sup_{x \geqslant 0} \left| \left[1 - \left(1 - \frac{x}{S_n}\right)_+^{n-1}\right] - \left[1 - e^{-\dfrac{x}{\theta_0}}\right]_+ \right| =$$

$$= \sup_{t \geqslant 0} \left| e^{-\dfrac{t}{\theta_0}} - \left(1 - \frac{t}{S_n}\right)_+^{n-1} \right| = \sup_{t \geqslant 0} \left| e^{-t} - \left(1 - \frac{t\,\theta_0}{S_n}\right)_+^{n-1} \right|,$$

$$W\left(k|\widehat{\mathbf{K}}\right) = \mathbf{P}\left(\widehat{D_n}\,(S_n) \geqslant c_{\alpha,n}\right) = \mathbf{P}\left(\sup_{kS_n^+ \geqslant x \geqslant 0} \left| e^{-\dfrac{x}{\theta_0}} - \left(1 - \frac{x/\theta_0}{k\,S_n/\theta_1}\right)^{n-1} \right| \geqslant c_{\alpha,n}\right).$$

Hence, the statement of the lemma immediately follows.

Now note that for $\theta = \theta_1$ formulas (2)–(3) can be rewritten in another way

$$\widehat{D_n}\,(S_n) = \sup_{z \in [0,\, kS_n^*]} \left|\Upsilon_0(z; S_n^*, k)\right|, \quad S_n^* = \frac{S_n}{\theta_1}, \tag{4}$$

$$W\left(k|\widehat{\mathbf{K}}\right) = \mathbf{P}\left(\sup_{z \in [0,\, kS_n^*]} \left|\Upsilon_0(z; S_n^*, k)\right| \geqslant c_{\alpha,n}\right), \tag{5}$$

using the function

$$\Upsilon_0(z;s,k) \equiv \Upsilon(z;ks) = e^{-z} - \left(1 - \frac{z}{ks}\right)^{n-1}, \quad 0 \leqslant z \leqslant ks. \tag{6}$$

The following statement determines basic properties of the function $\Upsilon(z;x)$, where $x = ks$, making it possible to construct an algorithm of numeric solution of the problems posed above.

Theorem. I. *If* $n = 2$, *then for* $x \leqslant 1$, *the function* $\Upsilon(z;x)$ *increases in the argument* z, $z \leqslant x$, *from* 0 *to* e^{-x}; *for* $x > 1$, *on the interval* $(0, \ln x)$ *the function* $\Upsilon(z;x)$ *decreases from* 0 *to* $\frac{1+\ln x}{x} - 1$, *and on the interval* $(\ln x,\, x)$ *it increases from* $\frac{1+\ln x}{x} - 1$ *to* e^{-x}.

II. *If* $n \geqslant 3$, *then*:

1) *for* $x \leqslant n-1$, *the equation* $\Upsilon'_z(z;x) = 0$ *has one root* $z_1 = z_1(x)$ *on the interval* $(0,\, x)$, *and for* $x > n-1$ *it has two roots* $z_1(x)$ *and* $z_2(x)$ *such that* $z_1(x)$ *belongs to the interval* $(x-n+2,\, x)$ *and* $z_2(x)$ *belongs to the interval* $(0,\, x-n+2)$;

2) *the function* $\Upsilon_1(x) = \Upsilon(z_1(x);x)$ *decreases on the interval* $(0, +\infty)$ *from* 1 *to* 0;

3) *the function* $\Upsilon_2(x) = -\Upsilon(z_2(x);x)$ *increases on the interval* $(n-1, +\infty)$ *from* 0 *to* 1;

4) *there exists a unique value of* x *such that* $\Upsilon_1(x) = \Upsilon_2(x)$; *moreover, this value belongs to the interval* $(n-1,\, x)$, $x > n-1$.

Proof. Let $n = 2$. Then, $\Upsilon(z;x) = e^{-z} - \left(1 - \frac{z}{x}\right)$ and $\Upsilon'_z(z;x) = -e^{-z} + \frac{1}{x}$, if $z \leqslant x$. Therefore, for $0 \leqslant x \leqslant 1$ the function $\Upsilon(z;x)$ is increasing in $z \in [0,\, x)$, since $\Upsilon'_z(z;x) > 0$ for any $z \geqslant 0$. Moreover, $\Upsilon(0;x) = 0$, $\Upsilon(x;x) = e^{-x}$.

If $x > 1$, then, since the equation $-e^{-z} + \frac{1}{x} = 0$ has a solution $z = \ln x$, on the interval $(0, \ln x)$ the function decreases in z to the value $\Upsilon(\ln x; x) = \frac{1+\ln x}{x} - 1$, whereas on the interval $(\ln x,\, x)$ it increases to the value e^{-x}. This reasoning implies statement I of the theorem.

Now consider the case $n \geqslant 3$. Making account of the relation

$$\Upsilon'_z(z;x) = -e^{-z} + \frac{n-1}{x}\left(1 - \frac{z}{x}\right)^{n-2}, \quad 0 < z < x,$$

define the function

$$\Psi(z) = \ln \frac{\frac{n-1}{x}\left(1 - \frac{z}{x}\right)^{n-2}}{e^{-z}} = \ln \frac{n-1}{x} + (n-2)\ln\left(1 - \frac{z}{x}\right) + z, \quad 0 \leqslant z < x,$$

that does not change signs on the same intervals with respect to z as the function $\Upsilon'_z(z;x)$.

The following properties of $\Psi(z)$ and $\Psi'(z)$ are obvious in their domains:

$$\Psi(0) = \ln \frac{n-1}{x}, \quad \Psi(x-0) = -\infty, \quad \Psi'(z) \text{ decreases},$$
$$\Psi'(0) > 0 \text{ for } x > n-2, \quad \Psi'(0) < 0 \text{ for } x < n-2,$$
$$\Psi'(0) = 0 \text{ for } x = n-2, \quad \Psi''(z) = -\frac{n-2}{(x-z)^2} < 0 \text{ for } 0 \leqslant z < x.$$

Let $0 \leqslant x \leqslant n-1$. Then, since in this case $\Psi(0) \geqslant 0$ and $\Psi'(z)$ decreases, the equation $\Psi(z) = 0$ has one root $z_1 = z_1(x)$ on the interval $[0,\, x)$. Moreover, on the interval $[0,\, z_1]$ the function $\Upsilon(z;x)$ increases from 0 to $\Upsilon_1(x)$, and on the interval $[z_1,\, x]$ it decreases from $\Upsilon_1(x)$ to e^{-x}.

Let $n-1 < x$. Then, since in this case $\Psi(0) < 0$ and $\Psi'(x-n+2) = 0$, the function $\Psi(z)$ increases on the interval $[0,\, x-n+2]$ and decreases on the interval $[x-n+2,\, x)$. Moreover, $\Psi(x-n+2) > 0$, since otherwise $\Psi(z) \leqslant 0$ for $0 \leqslant z < x$, which is impossible because $\Upsilon(0;x) = 0$ and $\Upsilon(x;x) = e^{-x} > 0$.

This means that for $x > n-1$ the equation $\Psi(z) = 0$ has two roots on the interval $[0,\, x)$, one belonging to the interval $(0,\, x-n+2)$, the other belonging to the interval $(x-n+2,\, x)$. Denote the second root $z_1 = z_1(x)$ (which corresponds to the notation introduced in the case $x \leqslant n-1$), since it corresponds to the maximum value of the function $\Upsilon(z;x)$. The first root will be denoted $z_2 = z_2(x)$.

Note that to the value $z_1(x)$ there corresponds the maximum value of the function $\Upsilon(z;x)$ for $x \geqslant 0$, which is equal to $\Upsilon_1(x)$, whereas to the value $z_2(x)$ there corresponds the minimum value of the function $\Upsilon(z;x)$ for $x \geqslant n-1$, which is equal to $-\Upsilon_2(x)$.

Now study the behavior of the functions $\Upsilon_i(x)$, $i=1,2$, in their domains. For this purpose calculate the derivatives

$$\begin{aligned}\Upsilon_i'(x) =&(-1)^{i-1}\left[-e^{-z_i(x)}z_i'(x)+(n-1)\left(1-\frac{z_i(x)}{x}\right)^{n-2}\frac{z_i'(x)\cdot x - z_i(x)}{x^2}\right] = \\ =&(-1)^{i}\frac{(n-1)z_i(x)}{x^2}\left(1-\frac{z_i(x)}{x}\right)^{n-2}, \quad i=1,2,\end{aligned}$$

with the account of that $z_i'(x)=0$ in any internal point of their domain. This immediately implies that $\Upsilon_1'(x)<0$ for $x>0$ and $\Upsilon_2'(x)>0$ for $x>n-1$. Therefore, the function $\Upsilon_1(x)$ decreases on $(0,+\infty)$ and the function $\Upsilon_2(x)$ increases on $(n-1,+\infty)$.

To complete the proof it suffices to find the limit values of the functions $\Upsilon_i(x)$, $i=1,2$, on the boundaries of their domains.

First make sure that the behavior of the roots $z_1(x)$ and $z_2(x)$ on the boundaries of their domains is determined by the relations

$$z_1(x) \sim x\left[1-\left(\frac{x}{n-1}\right)^{1/(n-2)}\right] \quad \text{for} \quad x \to +0, \tag{7}$$

$$z_1(x) \sim x\left[1-\exp\left\{-\frac{1}{n-2}\left(x+\ln\frac{n-1}{x}\right)\right\}\right] \quad \text{for} \quad x\to+\infty, \tag{8}$$

$$z_2(x) \sim x-n+1 \quad \text{for} \quad x \to n-1+0, \tag{9}$$

$$z_2(x) \sim \ln\frac{x}{n-1} \quad \text{for} \quad x\to+\infty. \tag{10}$$

We write $u_1(x) \sim u_2(x)$ for the asymptotic equivalence of $u_1(x)$ and $u_2(x)$.

It is easy to see that the values (7)–(10) so defined satisfy the inequalities determining the intervals in which the roots $z_1(x)$ and $z_2(x)$ fall:

$$0<z_1(x)<x \quad \text{for} \quad 0<x\leqslant n-1, \qquad x-n+2<z_1(x)<x \quad \text{for} \quad x>n-1; \tag{11}$$

$$0<z_2(x)<x-n+2 \quad \text{for} \quad x>n-1. \tag{12}$$

Make sure that the limit values (7)–(10) of the roots satisfy the equation $\Psi(z)=0$:

$$\lim_{x\to+0}\Psi\left(z_1(x)\right) = \lim_{x\to+0}\left[\ln\frac{n-1}{x}+\ln\left(\frac{x}{n-1}\right)+x\left[1-\frac{x}{n-1}\right]^{1/(n-2)}\right]=0, \tag{13}$$

$$\begin{aligned}\lim_{x\to+\infty}\Psi\left(z_1(x)\right) =& \lim_{x\to+\infty}\Bigg[\ln\frac{n-1}{x}-\left(x+\ln\frac{n-1}{x}\right)+ \\ &+x\left[1-\exp\left\{-\frac{1}{n-2}\left(x+\ln\frac{n-1}{x}\right)\right\}\right]\Bigg] = \\ &=-\lim_{x\to+\infty}x\exp\left\{-\frac{1}{n-2}\left(x+\ln\frac{n-1}{x}\right)\right\}=0,\end{aligned} \tag{14}$$

$$\lim_{x\to n-1+0}\Psi\left(z_2(x)\right) = \lim_{x\to n-1+0}\left[\ln\frac{n-1}{x}+(n-2)\ln\left(1-\frac{x-n+1}{x}\right)+x-n+1\right]=0, \tag{15}$$

$$\lim_{x\to+\infty}\Psi\left(z_2(x)\right)=\lim_{x\to+\infty}\left[\ln\frac{n-1}{x}+(n-2)\ln\left(1-\ln\frac{\ln\frac{x}{n-1}}{x}\right)+\ln\frac{x}{n-1}\right]=0. \quad (16)$$

Using (7)–(10), we can describe the behavior of the functions $\Upsilon_1(x)$ and $\Upsilon_2(x)$ on the boundaries of their domains:

$$\lim_{x\to+0}\Upsilon_1(x)=\lim_{x\to+0}\left[\exp\left\{-x\left[1-\left(\frac{x}{n-1}\right)^{\dfrac{1}{n-2}}\right]\right\}-\left(\frac{x}{n-1}\right)^{\dfrac{n-1}{n-2}}\right]=1,$$

$$\lim_{x\to+\infty}\Upsilon_1(x)=\lim_{x\to+\infty}\left[\exp\left\{-x\left[1-\exp\left(-\frac{x+\ln\frac{n-1}{x}}{n-2}\right)\right]\right\}--\exp\left\{-\frac{n-1}{n-2}\left(x+\ln\frac{n-1}{x}\right)\right\}\right]=0,$$

$$\lim_{x\to n-1+0}[-\Upsilon_2(x)]=\lim_{x\to n-1+0}\left[e^{-(x-n+1)}-\left(1-\frac{x-n+1}{x}\right)^{n-1}\right]=0,$$

$$\lim_{x\to+\infty}[-\Upsilon_2(x)]=\lim_{x\to+\infty}\left[e^{-\ln\dfrac{x}{n-1}}-\left(1-\frac{\ln\dfrac{x}{n-1}}{x}\right)^{n-1}\right]=-1,$$

which completes the proof of the theorem.

Corollary 1. *Under the conditions of the theorem, the statistic $\widehat{D_n}(S_n)$ has the following properties.*

I. $\widehat{D_2}(x)=e^{-x}$ *for* $0\leqslant x\leqslant w_1$, $\widehat{D_2}(x)=1-\frac{1+\ln x}{x}$ *for* $x\geqslant w_1$, *where* w_1 *is the solution of the equation*

$$1-\frac{1+\ln w}{w}=e^{-w},\quad w\geqslant 1. \quad (17)$$

II. *If* $n\geqslant 3$, *then* $\widehat{D_n}(x)=\Upsilon_1(x)$ *for* $0\leqslant x\leqslant w_2$, $\widehat{D_n}(x)=\Upsilon_2(x)$ *for* $x>w_2$, *where* w_2 *is the solution of the equation*

$$\Upsilon_1(w)=\Upsilon_2(w),\quad w>n-1. \quad (18)$$

3. Comparison of the Kolmogorov test and its modification

The Kolmogorov test was compared with its modification in the case of exponential distribution (1). The computation of the modified test $\widehat{\mathbf{K}}$ in the problem of testing the simple hypothesis $H_0:\ F(x;\theta)= = 1-e^{-x/\theta_0}$, $x\geqslant 0$, was conducted with the use of the theorem and corollary presented in Section 1. Table 1 presents the critical values of the modified test for significance levels $\alpha=0.9$, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.01 and sample sizes $n=3$, 5, 10, 50, 100, 200, 500, 1000.

Table 2 contains normalized critical values $\sqrt{n}\,c_{\alpha,n}$. As is seen from this table, for each given significance level, as the sample size increases, the convergence of the sequence of normalized critical values to some limit value is observed. This effect agrees with the asymptotic normality of the sequence of random variables $\sqrt{n}\,\widehat{D_n}$ established in [2]. Tables 3–6 contain the values of power functions of the tests $\widehat{\mathbf{K}}$ and $\mathbf{K}$ depending on significance level α with different sample sizes n for the alternatives of the form $H_1:\ F(x;\theta)=1-e^{-x/(k\,\theta_0)}$, $k=1$, 1.2, 1.4, …, 5.

Figure 1 demonstrates the results of the comparison of the test power functions for the significance level $\alpha=0.05$ and sample size $n=20$ with k varying from 0.2 to 3.7.

Table 1. Critical values $c_{\alpha,n}$ depending on confidence level $1-\alpha$ and sample size n

$1-\alpha$	Sample size n								
	3	5	10	20	50	100	200	500	1000
0.1	0.0901	0.0478	0.0245	0.0141	0.0076	0.0050	0.0034	0.0021	0.0015
0.2	0.1129	0.0650	0.0370	0.0234	0.0138	0.0095	0.0067	0.0042	0.0030
0.3	0.1383	0.0846	0.0514	0.0339	0.0206	0.0144	0.0101	0.0064	0.0045
0.4	0.1667	0.1066	0.0672	0.0452	0.0278	0.0195	0.0137	0.0086	0.0061
0.5	0.1984	0.1312	0.0846	0.0575	0.0356	0.0250	0.0176	0.0111	0.0079
0.6	0.2345	0.1590	0.1042	0.0713	0.0443	0.0311	0.0220	0.0139	0.0098
0.7	0.2772	0.1916	0.1271	0.0874	0.0545	0.0383	0.0270	0.0171	0.0121
0.8	0.3312	0.2328	0.1559	0.1077	0.0672	0.0473	0.0334	0.0211	0.0149
0.9	0.4137	0.2941	0.1986	0.1377	0.0862	0.0607	0.0429	0.0271	0.0191
0.95	0.4920	0.3494	0.2363	0.1640	0.1026	0.0723	0.0511	0.0323	0.0228
0.99	0.6687	0.4729	0.3148	0.2169	0.1352	0.0952	0.0672	0.0424	0.0300

Table 2. Critical values $\sqrt{n}\,c_{\alpha,n}$ depending on confidence level $1-\alpha$ and sample size n

$1-\alpha$	Sample size n								
	3	5	10	20	50	100	200	500	1000
0.1	0.156	0.107	0.077	0.063	0.053	0.050	0.048	0.047	0.047
0.2	0.196	0.145	0.117	0.105	0.098	0.095	0.094	0.094	0.093
0.3	0.240	0.189	0.162	0.151	0.145	0.144	0.143	0.142	0.142
0.4	0.289	0.238	0.212	0.202	0.196	0.195	0.194	0.193	0.193
0.5	0.344	0.293	0.268	0.257	0.252	0.250	0.249	0.248	0.248
0.6	0.406	0.356	0.330	0.319	0.313	0.311	0.311	0.310	0.310
0.7	0.480	0.429	0.402	0.391	0.385	0.383	0.382	0.382	0.382
0.8	0.574	0.521	0.493	0.482	0.475	0.473	0.473	0.472	0.472
0.9	0.716	0.658	0.628	0.616	0.609	0.607	0.606	0.605	0.605
0.95	0.852	0.781	0.747	0.733	0.726	0.723	0.722	0.721	0.721
0.99	1.158	1.058	0.996	0.970	0.956	0.952	0.950	0.948	0.948

Table 3. The values of power functions of the tests $\widehat{\mathbf{K}}$ and $\mathbf{K}$ depending on the significance level α for the sample size $n = 3$

k	$\alpha = 0.2$		$\alpha = 0.1$		$\alpha = 0.05$		$\alpha = 0.01$	
	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$
1	0.200	0.200	0.100	0.100	0.050	0.050	0.010	0.010
1.2	0.224	0.226	0.105	0.117	0.044	0.062	0.006	0.015
1.4	0.276	0.268	0.136	0.147	0.054	0.082	0.004	0.025
1.6	0.339	0.316	0.184	0.184	0.078	0.107	0.003	0.038
1.8	0.405	0.365	0.239	0.224	0.113	0.134	0.004	0.054
2	0.468	0.413	0.298	0.265	0.155	0.162	0.006	0.071
2.2	0.526	0.457	0.356	0.305	0.200	0.190	0.010	0.091
2.4	0.578	0.498	0.412	0.345	0.248	0.217	0.017	0.110
2.6	0.624	0.536	0.464	0.382	0.296	0.244	0.027	0.131
2.8	0.665	0.571	0.512	0.418	0.344	0.269	0.039	0.151
3	0.701	0.602	0.555	0.451	0.389	0.294	0.053	0.171
3.2	0.732	0.630	0.595	0.482	0.432	0.317	0.070	0.191
3.4	0.760	0.656	0.631	0.511	0.472	0.339	0.089	0.211
3.6	0.784	0.679	0.663	0.538	0.510	0.360	0.110	0.230
3.8	0.805	0.701	0.692	0.564	0.545	0.379	0.132	0.248
4	0.823	0.720	0.718	0.587	0.577	0.398	0.156	0.266
4.2	0.840	0.737	0.741	0.609	0.607	0.416	0.180	0.283
4.4	0.854	0.753	0.762	0.629	0.634	0.433	0.205	0.300
4.6	0.867	0.768	0.781	0.648	0.659	0.449	0.230	0.316
4.8	0.879	0.782	0.798	0.666	0.682	0.464	0.255	0.332
5	0.889	0.794	0.814	0.682	0.704	0.478	0.280	0.347

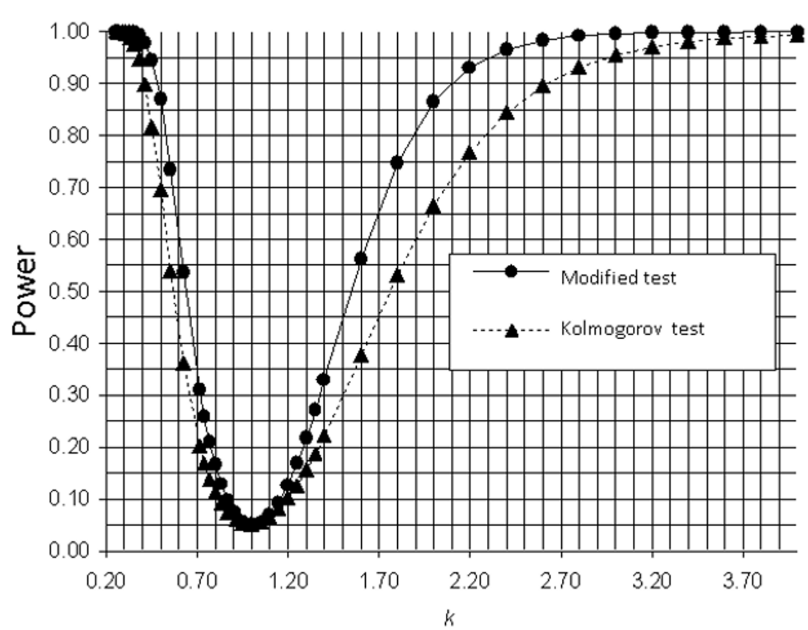

Fig. 1. The test power functions depending on k

Table 4. The values of power functions of the tests $\widehat{\mathbf{K}}$ and $\mathbf{K}$ depending on the significance level α for the sample size $n = 10$

	$\alpha = 0.2$		$\alpha = 0.1$		$\alpha = 0.05$		$\alpha = 0.01$	
k	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$
1	0.200	0.200	0.100	0.100	0.050	0.050	0.010	0.010
1.2	0.286	0.253	0.158	0.140	0.082	0.077	0.013	0.020
1.4	0.441	0.355	0.291	0.223	0.181	0.139	0.042	0.045
1.6	0.596	0.468	0.447	0.324	0.317	0.220	0.106	0.085
1.8	0.721	0.573	0.590	0.427	0.461	0.311	0.201	0.137
2	0.812	0.663	0.706	0.524	0.590	0.402	0.313	0.198
2.2	0.875	0.737	0.793	0.609	0.696	0.487	0.429	0.264
2.4	0.917	0.795	0.856	0.681	0.778	0.564	0.537	0.332
2.6	0.945	0.841	0.900	0.741	0.839	0.632	0.632	0.397
2.8	0.963	0.876	0.931	0.790	0.884	0.690	0.711	0.460
3	0.975	0.903	0.952	0.830	0.917	0.739	0.776	0.518
3.2	0.983	0.924	0.966	0.862	0.940	0.780	0.827	0.572
3.4	0.988	0.940	0.976	0.888	0.957	0.815	0.867	0.620
3.6	0.992	0.953	0.983	0.909	0.968	0.844	0.897	0.663
3.8	0.994	0.962	0.988	0.925	0.977	0.869	0.921	0.702
4	0.996	0.970	0.991	0.939	0.983	0.889	0.939	0.736
4.2	0.997	0.976	0.994	0.950	0.987	0.906	0.953	0.767
4.4	0.998	0.980	0.995	0.959	0.991	0.920	0.964	0.793
4.6	0.999	0.984	0.997	0.966	0.993	0.931	0.972	0.817
4.8	0.999	0.987	0.997	0.971	0.995	0.941	0.978	0.838
5	0.999	0.989	0.998	0.976	0.996	0.950	0.983	0.856

Table 5. The values of power functions of the tests $\widehat{\mathbf{K}}$ and $\mathbf{K}$ depending on the significance level α for the sample size $n = 20$

	$\alpha = 0.2$		$\alpha = 0.1$		$\alpha = 0.05$		$\alpha = 0.01$	
k	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$
1	0,200	0.200	0.100	0.100	0.050	0.050	0.010	0.010
1.2	0.352	0.294	0.215	0.173	0.126	0.101	0.030	0.028
1.4	0.602	0.471	0.455	0.327	0.329	0.221	0.127	0.084
1.6	0.793	0.643	0.680	0.500	0.561	0.376	0.305	0.178
1.8	0.903	0.775	0.832	0.653	0.746	0.532	0.509	0.298
2	0.957	0.865	0.918	0.770	0.864	0.665	0.685	0.428
2.2	0.981	0.920	0.961	0.852	0.931	0.768	0.812	0.550
2.4	0.992	0.954	0.982	0.907	0.966	0.843	0.893	0.656
2.6	0.996	0.973	0.992	0.942	0.983	0.896	0.941	0.742
2.8	0.998	0.985	0.996	0.964	0.992	0.931	0.968	0.810
3	0.999	0.991	0.998	0.978	0.996	0.955	0.983	0.861
3.2	1.000	0.995	0.999	0.986	0.998	0.970	0.991	0.900
3.4	1.000	0.997	1.000	0.991	0.999	0.981	0.995	0.927
3.6	1.000	0.998	1.000	0.995	1.000	0.987	0.997	0.948
3.8	1.000	0.999	1.000	0.997	1.000	0.991	0.999	0.962
4	1.000	0.999	1.000	0.998	1.000	0.994	0.999	0.973
4.2	1.000	1.000	1.000	0.999	1.000	0.996	1.000	0.980
4.4	1.000	1.000	1.000	0.999	1.000	0.997	1.000	0.986
4.6	1.000	1.000	1.000	0.999	1.000	0.998	1.000	0.990
4.8	1.000	1.000	1.000	1.000	1.000	0.999	1.000	0.992
5	1.000	1.000	1.000	1.000	1.000	0.999	1.000	0.994

Table 6. The values of power functions of the tests $\widehat{\mathbf{K}}$ and $\mathbf{K}$ depending on the significance level α for the sample size $n = 50$

	$\alpha = 0.2$		$\alpha = 0.1$		$\alpha = 0.05$		$\alpha = 0.01$	
k	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$	$\widehat{\mathbf{K}}$	$\mathbf{K}$
1	0.200	0.200	0.100	0.100	0.050	0.050	0.010	0.010
1.2	0.518	0.408	0.370	0.268	0.254	0.173	0.093	0.059
1.4	0.861	0.722	0.770	0.586	0.666	0.459	0.421	0.235
1.6	0.975	0.906	0.948	0.827	0.908	0.733	0.768	0.499
1.8	0.997	0.974	0.991	0.942	0.982	0.895	0.936	0.734
2	1.000	0.994	0.999	0.983	0.997	0.965	0.986	0.880
2.2	1.000	0.999	1.000	0.996	1.000	0.989	0.997	0.952
2.4	1.000	1.000	1.000	0.999	1.000	0.997	1.000	0.982
2.6	1.000	1.000	1.000	1.000	1.000	0.999	1.000	0.994
2.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
3.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 7. The sample size required for the test **K** to attain the same power as the test $\widehat{\mathbf{K}}$

k	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8
$n = 20$	31	34	34	34	34	33	33	33	33
	0.126	0.330	0.563	0.750	0.870	0.937	0.971	0.987	0.994
$n = 50$	84	83	81	80	78	-	-	-	-
	0.254	0.670	0.911	0.986	0.998	-	-	-	-

The presented results of computation demonstrate that, as a rule, the modified test $\widehat{\mathbf{K}}$ is more powerful than **K** uniformly over the whole range of k. The only exceptions are separate results related to small sample sizes and k close to 1. The results presented in Table 7 indicate even more vividly the substantial advantage of the test $\widehat{\mathbf{K}}$ over the test **K**. Table 7 presents the values of the sample size and power of the test **K**, corresponding to the given sample size n of the test $\widehat{\mathbf{K}}$ at the significance level $\alpha = 0.05$ and some values of the coefficient k. As is seen from this table, within the range of input parameters described above, the Kolmogorov test requires a 1.6–1.7 times greater number of observations.

In conclusion, it should be noted that the computation of the power of the Kolmogorov test was conducted by the computer program written by E. A. Kosyanova on the base of [3], which can be treated as an alternative to the approach proposed in [4].

REFERENCES

1. V. G. Voinov and M. S. Nikulin, *Unbiased Estimators and their Applications. Vol. 1: Univariate Case*, Springer, Berlin (1996).
2. V. V. Chichagov, "Finding distribution of Kolmogorov statistic modification based on unbiased estimator of distribution function from exponential family," in: *Transactions of the XXVI International Seminar on Stability Problems for Stochastic Models*, Ort Braude College, Karmiel (2007), pp. 60–63.
3. V. V. Chichagov, "On some applications of Poisson random walk plans," *J. Math. Sci.*, **81**, No. 4, 2874–2879 (1996).
4. V. A. Epanechnikov, "The significance level and power of the two-sided Kolmogorov test in the case of small sample sizes," *Theor. Probab. Appl.*, **13**, No. 4, 686–690 (1968).