ALT`2008 – June 9-11 – Bordeaux

# On Testing Simple and Composite Goodness-of-Fit Hypotheses When Data are Censored

E.V. Chimitova[1], B.Yu. Lemeshko[1],

[1]Novosibirsk State Technical University, Russia

*Abstract*— **Problems of application of the nonparametric Kolmogorov, Cramer-von Mises-Smirnov, Anderson-Darling goodness-of-fit tests for censored data have been considered in this paper. The convergence of statistic distributions to the corresponding limiting distribution laws has been investigated under some true null hypothesis by means of statistical simulation methods as well as the test power against close competing hypotheses. The distributions of test statistics under study have been investigated for composite hypotheses.**

*Index Terms*—**goodness-of-fit tests, censored data, the Smirnov transformation, the Renyi test, the Kolmogorov-Smirnov test, the Cramer-von Mises-Smirnov test, the Anderson-Darling test.**

## I. Introduction

Les us consider a plan of experiment, when we observe the first $r$ of order statistics $X_{(1)}, X_{(2)}, ..., X_{(r)}$ from the sample $X_1, X_2, ..., X_n$ of independent identically distributed random variables of the size $n$ ($r \le n$).

The problem of testing simple and composite goodness-of-fit hypotheses with the nonparametric Kolmogorov-Smirnov, Renyi, Cramer-von Mises-Smirnov, Anderson-Darling criteria has been considered in this paper.

The limiting distributions of the above statistics were obtained [1]-[3] for testing simple hypotheses of the kind

$$H_0 : F(x) = F(x, \theta),$$

where $F(x,\theta)$ is the probability distribution function, with which an observed sample is tested for fit, and $\theta$ is a known (scalar or vector) parameter value. In case of composite hypotheses of the kind

$$H_0 : F(x) \in F(x, \theta), \theta \in \Omega \;,$$

where an estimate $\hat{\theta}$ is used instead of the unknown parameter $\theta$, the distribution of nonparametric statistics $G(S|H_0)$ essentially differs from the corresponding distribution when a simple hypothesis is tested. It depends on the form of the law $F(x,\theta)$ on which the null hypothesis is based on, estimation method for the parameter $\theta$ and a number of other factors [4]-[5]. Notice that the estimate $\hat{\theta}$ is calculated from the same sample that the goodness-of-fit hypothesis is tested by. If $\hat{\theta}$ is obtained from another sample, then the hypothesis under test is simple.

In spite of the fact that these criteria were obtained long ago, their properties for limited sample sizes haven't been properly studied. Probably, for this reason the usage of the tests for censored data hasn't been realized in any program system of statistical analysis we are familiar with. Hence they are not available for most specialists.

M. Nikulin in [6] attracted our attention to the possibility of effective application of nonparametric goodness-of-fit tests for the analysis of incomplete data by means of the Smirnov transformation and the "randomization". The advantages of such approach are evident as we move to the problem of testing goodness-of-fit of the empirical distribution obtained after transformations to the continuous (uniform) distribution law. For solving such a problem one can use the well studied technique of testing goodness-of-fit hypotheses from complete samples.

The paper is aimed at investigating practical aspects of nonparametric goodness-of-fit tests to analyze censored data on the right, using the Smirnov transformation and randomization among other things. In the paper the convergence of test statistic distributions to the corresponding limiting laws is investigated by means of the Monte-Carlo method. The power of above tests is also studied when testing close competing hypotheses.

**The Kolmogorov-Smirnov test**

The Kolmogorov-Smirnov statistics are defined as [1]:

$$D_n^+(0,1-a) = \sup_{F(x)\le 1-a} F_n(x) - F(x) \;,$$

$$D_n^-(0,1-a) = -\inf_{F(x)\le 1-a} F_n(x) - F(x) \;,$$

$$D_n(0,1-a) = \sup_{F(x)\le 1-a} |F_n(x) - F(x)|,$$

In practice the Kolmogorov statistic is more convenient to use with the Bolshev correction [7]:

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

where $D_n = \max\{D_n^+, D_n^-\}$, $D_n^+ = \max_{1\le i\le r}\left\{\frac{i}{n} - F(X_{(i)})\right\}$,

$$D_n^- = \max_{1\le i\le r}\left\{F(X_{(i)}) - \frac{i-1}{n}\right\}.$$

In [1] the limiting statement was obtained

$P \;\; S_K^c < S \;\; =$

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

$$= \sum_{i=-\infty}^{+\infty}(-1)^i \exp(-2i^2S^2)\cdot P\left\{\left|X-2iS\sqrt{\frac{a}{1-a}}\right|<\frac{S}{\sqrt{a-a^2}}\right\}=$$
$$=K_a^c(S),$$

where $X$ is the standard normal random variable. When the censoring degree $a=0$ the limiting distribution of the statistic $S_K^c$ coincides with the Kolmogorov distribution for complete samples:

$$K(S)=\sum_{i=-\infty}^{+\infty}(-1)^i\exp(-2i^2S^2).$$

**The Renyi test**

The Renyi test statistic to test a simple hypothesis $H_0$, are expressed as [2], [7]:

$$R_n^+(0,1-a)=\sup_{F(x)\le 1-a}\frac{F_n(x)-F(x)}{1-F(x)}=\max_{1\le i\le r}\frac{i/n-F(X_{(i)})}{1-F(X_{(i)})},$$

$$R_n^-(0,1-a)=\sup_{F(x)\le 1-a}\frac{F(x)-F_n(x)}{1-F(x)}=\max_{1\le i\le r}\frac{F(X_{(i)})-(i-1)/n}{1-F(X_{(i)})},$$

$$R_n(0,1-a)=\sup_{F(x)\le 1-a}\frac{|F_n(x)-F(x)|}{1-F(x)}=\max\left(R_n^+,R_n^-\right). \quad (2)$$

The random variables $R_n^+(0,1-a)$ and $R_n^-(0,1-a)$ are distributed identically and, as Renyi showed [2], the following limiting formulas are used:

$$\lim_{n\to\infty}P\left\{\sqrt{\frac{na}{1-a}}R_n^+(0,1-a)<S\right\}=2\Phi(S)-1,$$

$$\lim_{n\to\infty}P\left\{S_R^c=\sqrt{\frac{na}{1-a}}R_n(0,1-a)<S\right\}=L(S),\quad S>0,$$

where $\Phi(S)$ is the standard normal distribution, $L(S)$ is the Renyi distribution function:

$$L(S)=\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left\{-\frac{(2k+1)^2\pi^2}{8S^2}\right\}.$$

**The Cramer-von Mises-Smirnov test**

The Cramer-von Mises-Smirnov statistic is calculated by:

$$S_\omega=\frac{1}{12n}+\sum_{i=1}^{r}\left[F(X_{(i)})-\frac{2i-1}{2n}\right]^2. \quad (3)$$

In [3] the upper percentage points of the Cramer-von Mises-Smirnov statistic distribution are given for various censoring degrees $a$ in case of testing a simple hypothesis $H_0$. In case of a complete sample ($r=n$) this statistic has the distribution 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 $I_{-\frac{1}{4}}(\cdot)$, $I_{\frac{1}{4}}(\cdot)$ are the modified Bessel functions,

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

**The Anderson-Darling test**

The Anderson-Darling statistic from the censored sample on the right is calculated according to the formula:

$$S_\Omega=n\sum_{i=1}^{r-1}\left[F(X_{(i)})-F(X_{(i+1)})+\left(\frac{n-i}{n}\right)^2\ln\left(\frac{1-F(X_{(i)})}{1-F(X_{(i+1)})}\right)+\right.$$
$$\left.+\left(\frac{i}{n}\right)^2\ln\left(\frac{F(X_{(i+1)})}{F(X_{(i)})}\right)\right]-n\ln(1-F(X_{(1)}))-nF(X_{(1)}).$$

In [3] the upper percentage points of the Anderson-Darling statistic distribution are given for various censoring degrees $a$ when testing a simple hypothesis $H_0$. In case of a complete sample ($r=n$) this statistic has the distribution 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. THE INVESTIGATION OF STATISTIC DISTRIBUTIONS WHEN TESTING SIMPLE HYPOTHESES

We have studied the Renyi statistic distributions for various distributions laws with which an observed sample is tested for fit, various sample sizes $n$ and censoring degrees $a$ by statistical modeling methods. It has been shown that the Renyi statistic distributions essentially depend on the censoring degree. For example in figure 1 the empirical distributions of the Renyi statistic are represented when testing goodness-of-fit to the exponential law with the density function

$$f(x)=\frac{1}{\theta_0}e^{-x/\theta_0},\ \theta_0=1,$$

in case of the censoring degree $a=0.9$.

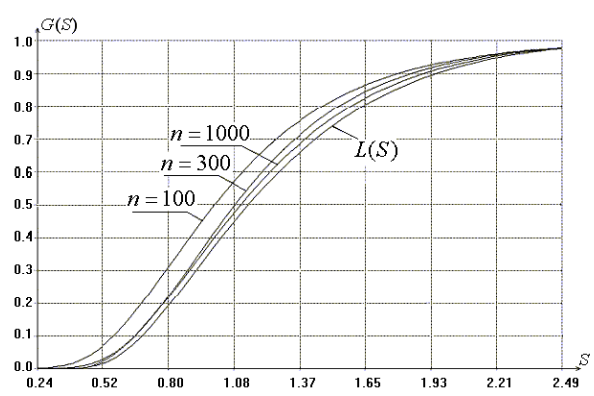


Fig. 1. The distributions of the Renyi statistic $S_R^c$ when the sample size $n=100$, 300, 1000 and $a=0.9$

As it is seen from the figure 1 for the sample size $n < 1000$ and $a = 0.9$ the Renyi statistic distributions essentially differ from the limiting law $L(S)$. The investigation of the Renyi statistic distributions depending on the censoring degree has shown that the best goodness-of-fit to the limiting law $L(S)$ is reached for $a = 0.5$ and for small or on the contrary high censoring degrees the empirical statistic distributions essentially differ from $L(S)$.

In contrast to the Renyi test the investigations of the Kolmogorov statistic distributions have shown their good convergence to the corresponding limiting distribution functions. For example, in figure 2 the Kolmogorov empirical distributions $G(S_K^c \mid H_0)$ and the corresponding limiting laws are represented while testing simple hypothesis of goodness-of-fit to the exponential law.

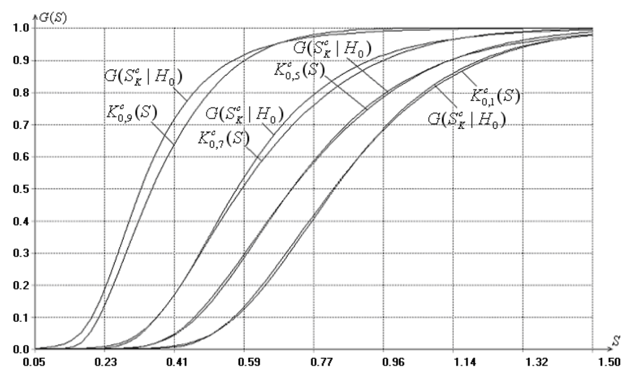


Fig. 2. The distributions of the Kolmogorov statistic $S_K^c$ for various censoring degrees, $n = 50$

As is it seen from the figure 2 the empirical distributions of the Kolmogorov statistic are very close the corresponding limiting laws when the sample size $n = 50$ and the censoring degree $a < 0.5$, but for $a = 0.7$ and $a = 0.9$ the empirical distributions of the statistic (1) considerably differ from their limiting laws.

Thus, as a result of investigations with computer simulation methods a sufficient goodness-of-fit of the empirical distributions $G(S_K^c \mid H_0)$ to the limiting law $K_c^a(S)$ has been shown for the sample size beginning from $n = 30$ when the censoring degree less than 0.5. While increasing censoring degree up to 0.95 a sufficient goodness-of-fit of $G(S_K^c \mid H_0)$ to $K_c^a(S)$ has been observed only for $n \geq 500$.

Similar regularities for the statistic distributions have been observed for the Cramer-von Mises-Smirnov test and the Anderson-Darling test. For these criteria we have compared sample quantiles, obtained from the empirical distributions of the test statistics, with the upper percentage points given in [3] depending on the sample size and the censoring degree.

## III. The Smirnov Transformation and Randomization for Censored Data

Smirnov transformation is used rather often in statistical analysis. It is possible to move from a censored sample to the sample of random variables $U_1, U_2, ..., U_n$, uniformly distributed on [0,1]. In case of right censoring we have $U_1 = F(X_{(1)})$, $U_2 = F(X_{(2)})$, …, $U_r = F(X_{(r)})$, and the values $U_{r+1}, U_{r+2}, ..., U_n$ are simulated uniformly on the interval $\left. F(X_{(r)}), 1 \right]$. The randomization as a technique of conversion censored data to a sample of "complete" observations is really applicable only in computer analysis.

The classical Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling tests (for complete samples) can be applied to analyze transformed sample. We have investigated the statistic distributions of these criteria by statistical simulation for various distributions on which the null goodness-of-fit hypothesis is based on and various sample sizes. As an example let us consider the problem of testing simple hypothesis of goodness-of-fit to the Weibull distribution with the density function

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

by a censored sample of size $n = 20$ and $r = 4$. The empirical distributions of the Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling statistics obtained when testing goodness-of-fit of the samples $U_1, U_2, ..., U_n$ obtained after Smirnov transformation and randomization with the uniform on [0,1] distribution are represented in the fig. 3.

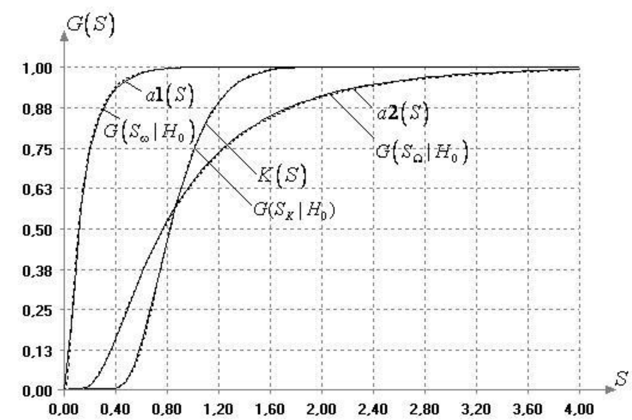


Fig. 3. The distributions of the Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling statistics when $n = 20$, $r = 4$

As it is seen from the figure 3 the empirical distributions of the nonparametric Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling statistics calculated by a transformed sample $U_1, U_2, ..., U_n$ are in accord with the corresponding limiting laws already when $n = 20$ *independently on the censoring degree*. This conclusion is confirmed by the high significance levels achieved while testing goodness-of-fit of the empirical statistic distributions to the corresponding limiting laws.

### IV. THE INVESTIGATION OF STATISTIC DISTRIBUTIONS WHEN TESTING COMPOSITE HYPOTHESES

When testing composite hypotheses the estimate $\hat{\theta}$, which is calculated from the same sample that the goodness-of-fit hypothesis is tested by, is used as the unknown parameter $\theta$. Estimates of the distribution parameters can be obtained by the maximum likelihood method which is universal concerning the form of data registration. When calculating maximum likelihood estimates (MLE) from censored sample on the right the following system of likelihood equations is solved

$$\sum_{j=1}^{r} \frac{\partial \ln f(X_{(j)},\theta)}{\partial \theta_i} + (n-r)\frac{\partial \ln P_c(\theta)}{\partial \theta_i} = 0, \quad i = \overline{1,m},$$

where $m$ is the dimension of the parameter vector $\theta = (\theta_1, \theta_2, \ldots, \theta_m)^{\mathrm{T}}$, $f(x,\theta)$ is the density function of the random variable, $P_c(\theta) = \int_{X_{(r)}}^{\infty} f(x,\theta)dx$.

B. Lemeshko in his papers obtained the distribution models approximating the limiting distribution laws of the nonparametric statistics for a number distribution laws with which an observed sample is tested for fit using MLE. The approximations for various distributions are represented in [8].

Let us investigate the nonparametric statistic distributions when using the Smirnov transformation and randomization. MLE calculated from the censored samples are used as the unknown parameter. Then we compare the empirical statistic distributions obtained with the corresponding approximations given in [8].

It has been shown that the distributions of the nonparametric statistics under consideration when testing composite hypotheses essentially depend on the censoring degree.

For example, the empirical distributions of the Kolmogorov statistic (1) obtained when testing composite hypothesis of goodness-of-fit to the Weibull distribution are represented in the figure 4 for the sample size $n = 100$ and various censoring degrees. The approximation of the limiting distribution law $\gamma(4.9738, 0.0660, 0.3049)$ for the Kolmogorov statistic when testing goodness-of-fit to the Weibull distribution and estimating the scale and form parameters by maximum likelihood method is also represented in the figure.

As it seen from the figure 4, the empirical statistic distribution is close to the approximation of the limiting law for complete samples when the censoring degree is small ($n = 100$, $r = 95$). But with decreasing the number of observations $r$ from 60 and lower the deviation to the right from the limiting law considerably increases.

Similar regularities have been observed when testing composite hypotheses of goodness-of-fit to other distribution laws. The deviation to the right of the empirical distributions from the limiting law for high censoring degrees can be explained with the fact that for limited sample sizes and high censoring degree the parameter MLE turns out to be biased [9].

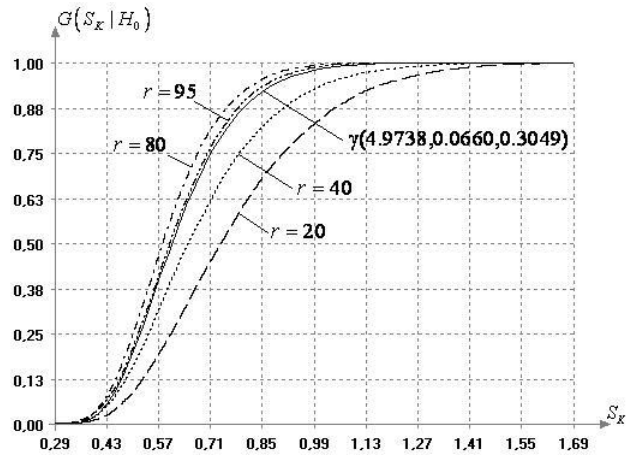


Fig. 4. The Kolmogorov statistics distributions for $n = 100$

### V. THE TEST POWER INVESTIGATION

The power of the nonparametric tests has been investigated for various pairs of close competing hypotheses depending on the sample size $n$ and the censoring degree in the paper. The test power increases with the sample size growth for any pair of competing hypotheses. The test power behavior relative to the censoring degree growth essentially depends on the kind of hypotheses $H_0$ and $H_1$ under test.

Let us consider as an example the problem of testing simple hypothesis $H_0$: exponential distribution with the scale parameter equal to 1 against $H_1$: the Weibull distribution with the form parameter 1.2 and the scale parameter 1. The power's estimates for the Kolmogorov-Smirnov test are given in case of the sample size $n = 100$ and the significance level $\alpha = 0.1$ in the table 1.

Table 1

| *When calculating the statistic from the original (without transformation) censored sample* | | | | | |
|---|---|---|---|---|---|
| $r = 95$ | $r = 80$ | $r = 65$ | $r = 50$ | $r = 40$ | $r = 10$ |
| 0.24 | 0.23 | 0.22 | 0.26 | 0.31 | 0.46 |
| *When calculating the statistic using the Smirnov transformation and randomization* | | | | | |
| $r = 95$ | $r = 80$ | $r = 65$ | $r = 50$ | $r = 40$ | $r = 10$ |
| 0.22 | 0.21 | 0.21 | 0.20 | 0.20 | 0.18 |

As it is seen from the table 1 the test power in the first case (without transformation) changes not steadily for this pair of competing hypotheses.

The power's estimates for the Kolmogorov-Smirnov test in case of testing simple hypothesis $H_0$: exponential distribution with the scale parameter 0.5 against $H_1$: exponential distribution with the scale parameter 0.7 are given in the table 2. As in the previous example here the sample size $n = 100$ and the significant level $\alpha = 0.1$.

Table 2

| *When calculating the statistic from the original (without transformation) censored sample* | | | | | |
|---|---|---|---|---|---|
| $r=95$ | $r=80$ | $r=65$ | $r=50$ | $r=40$ | $r=10$ |
| 0.85 | 0.84 | 0.83 | 0.76 | 0.71 | 0.42 |
| *When calculating the statistic using the Smirnov transformation and randomization* | | | | | |
| $r=95$ | $r=80$ | $r=65$ | $r=50$ | $r=40$ | $r=10$ |
| 0.85 | 0.84 | 0.82 | 0.74 | 0.65 | 0.19 |

Here the test power steadily decreases with the censoring degree growth in both cases.

We also have studied the test power for a number of other pairs of competing hypotheses. Analyzing the results given in the tables 1 and 2 as well as the results of other simulations the following regularities have been observed: when testing simple hypotheses for small censoring degrees the power of the Kolmogorov-Smirnov test insignificantly higher when calculating statistic from the original censored samples $X_{(1)},X_{(2)},...,X_{(r)}$ than in case of using the Smirnov transformation and randomization. But while increasing the censoring degree the advantage of the test by the original censored sample becomes more considerable comparing the test in which the Kolmogorov statistic is calculated from the transformed sample $U_1,U_2,...,U_n$.

In case of composite hypothesis testing the test power is higher than for simple hypotheses for the same kind of competing hypotheses, what is confirmed by statistical modeling. As an example we consider testing composite hypothesis $H_0$: exponential distribution against $H_1$: the Weibull distribution with the form parameter 1.2 and the scale parameter 1 by samples of size $n=100$. The power's estimates are given in table 3 for the Kolmogorov-Smirnov test, the significance level $\alpha=0.1$.

Table 3

| *When calculating the statistic from the original (without transformation) censored sample* | | | | | |
|---|---|---|---|---|---|
| $r=95$ | $r=80$ | $r=65$ | $r=50$ | $r=40$ | $r=30$ |
| 0.54 | 0.45 | 0.37 | 0.31 | 0.25 | 0.23 |
| *When calculating the statistic using the Smirnov transformation and randomization* | | | | | |
| $r=95$ | $r=80$ | $r=65$ | $r=50$ | $r=40$ | $r=30$ |
| 0.53 | 0.44 | 0.34 | 0.22 | 0.15 | 0.12 |

As it is seen from the table 3, the test power steadily decreases with censoring degree growth. Similarly to the case of simple hypotheses the test power is higher when the statistic is calculated from the original censored samples $X_{(1)},X_{(2)},...,X_{(r)}$ than in case of using the Smirnov transformation and randomization and difference in the power increases with the censoring degree growth.

## VI. Conclusion

The Renyi statistic distributions have been investigated for various sample sizes and censoring degrees. It has been shown that the statistic distributions converge to the limiting law $L(S)$ very slowly, especially for high or on the contrary very low censoring degrees. This result doesn't allow recommending using the Renyi test in practice.

The results of investigating the distributions the Kolmogorov statistic, Cramer-von Mises-Smirnov statistic and the Anderson-Darling statistic enable to conclude a good possibility to use the approach considered (Smirnov transformation with randomization) for correct application of the classical nonparametric goodness-of-fit tests for censored data. In case of simple hypothesis testing, the distributions of the above statistics converge to the corresponding limiting distributions very quickly. For the sample size $n\geq 20$ one can use the limiting laws without risk of making a great mistake for any censoring degree.

The application of Smirnov transformation with randomization is quite efficient for realization in software systems of statistical analysis.

The nonparametric tests using the Smirnov transformation and randomization are at a disadvantage by power comparing with the corresponding tests without data transformation, especially for high censoring degrees.

When testing composite hypotheses from censored data and estimating unknown parameters by maximum likelihood method the distributions of nonparametric test statistics considerably depend on the censoring degree.

## References


[1] Barr D.M., Davidson T. A Kolmogorov-Smirnov test for censored samples. Technometrics, 1973. Vol. 15. N. 4.

[2] Renyi A. On the theory of order statistics // Acta Mathem. Acad. Sci. Hung. 1953. Vol. 4. – P. 191-232.

[3] Pettitt A.N., Stephens M.A. Modified Cramer von Mises statistics for censored data // Biometrika, 1976. Vol. 63. N. 2.

[4] Lemeshko B. Yu. Errors when using nonparametric fitting criteria // Measurement Techniques, 2004.Vol. 47. N. 2, – P.134-142.

[5] Lemeshko B.Yu., Postovalov S.N., Frantsuzov A.V. Application of the nonparametric goodness-of-fit tests to testing nonparametric model adequacy // Optoelectronics, Instrumentation and Data Processing, 2002. N. 2. – P. 3-12.

[6] Greenwood P.E., Nikulin M.S. A Guide to Chi-Squared Testing. – John Wiley & Sons, Inc. 1996. – 280 p.

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

[8] 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. - 64 p. (in Russian).

[9] Lemeshko B.Yu., Chimitova E.V. Investigation of the estimates properties and goodness-of-fit test statistics from censored samples with computer modeling technique // Proceedings of the Seventh International Conference "Computer Data Analysis and Modeling: Robustness and Computer Intensive Methods", September 6-10, 2004, Minsk. Vol. 1. – P. 143-146.