The Comparative Analysis of Tests in the Problem of Testing the Hypothesis of Uniformity[1]

Pavel Yu. Blinov and Boris Yu. Lemeshko
Novosibirsk State Technical University, Novosibirsk, Russian Federation
e-mail: blindizer@yandex.ru,lemeshko@ami.nstu.ru

Abstract

In the paper some statistical tests intended for testing of uniformity have been considered. Distributions of test statistics, the power of tests under different competing hypotheses have been studied. Considered tests have been ranked by the test power. Advantages and disadvantages of individual tests have been shown. Also, it has been shown that the large part of the tests traditionally used for testing uniformity has the bias under some kind of competing hypotheses. It is underlines that special uniformity tests haven't clear advantage over nonparametric goodness-of-fit tests used for testing uniformity in general.

Keywords: uniform distribution, hypothesis testing, test statistic, test power.

Introduction

The uniform distribution is one of common distributions in applied mathematics statistics and probability theory. It is often used to describe the measurement error of some instruments or measuring systems. Simulation of pseudorandom values according to different parametric laws relies on sensors of uniform pseudorandom values. Parametric laws are urgently needed in the systems of statistical simulation. Testing the uniformity actually represents goodness-of-fit testing the hypothesis of uniform distribution of the observed sample $x_1, ..., x_n$. In some papers, the authors states that testing composite hypothesis can be reduced to test simple hypothesis of uniformity on the interval $[0, 1]$, because if $x_1, ..., x_n$ belong law with probability distribution function $F(x)$, then random variable $y_i = F(x_i)$ is uniformly distributed on unit interval. All of these factors explain the increasing interest in the choice of simple and computationally efficient procedures for testing hypotheses about the uniform law of analyzed samples.

The various statistical tests used for testing hypothesis of uniformity can be divided into two subsets. These are common goodness-of-fit tests applicable for testing of uniformity and special tests oriented on testing hypothesis that sample $x_1, ..., x_n$ is uniform distributed.

The presence of numerous tests put not simple problem of choosing for specialists, because available information in papers doesn't allows to give preference to certain test, while every specialist is interested not only in correctness of using of tests, but else in reliability of statistical inferences.

[1]This work is supported by the Russian Ministry of Education and Science (project 2.541.2014K).

In this paper, a lot of considered tests are studied by the method of statistical simulations. The number of experiments carried out for statistical modeling is assumed equal to 1 660 000 in the study of the distributions of test statistics. On the one hand, such number of experiments allows tracing the qualitative picture of test statistic distributions in depend on various factors. On the other hand, this number of experiments provides acceptable accuracy of the power estimates and unknown probabilities.

1 The statement of testing uniformity

In the most of uniformity tests, ordered statistics of quantity X are used ($x_{(1)} < x_{(2)} < ... < x_{(n)}$ are elements $x_{(i)}$ of variation series of the sample). Further designation $U_i = x_{(i)}, i = \overline{1, n}$ will be used in expressions of statistical tests.

As usually tests are oriented on testing of simple hypothesis H_0 on interval $[0, 1]$. However, if hypothesis of uniformity is tested on interval $[a, b]$ then elements $x_{(i)}$ of variation series $a < x_{(1)} < x_{(2)} < ... < x_{(n)} < b$ are modified to corresponding (required in the tests) ordered statistics as: $U_i = \frac{x_{(i)} - a}{b - a}, i = \overline{1, n}$.

To test composite hypothesis of uniformity H_0: $F(x) = (x - a)/(b - a)$, $x \in [a, b]$, where a and b are non-known, we proceed as follows. Using the variation series $x_{(1)} < x_{(2)} < ... < x_{(n)}$ of sample $X_1, X_2, ..., X_n$ the parameter estimates are obtained as follows:

$$\hat{a} = x_{(1)} - \frac{x_{(n)} - x_{(1)}}{n - 1}, \hat{b} = x_{(n)} + \frac{x_{(n)} - x_{(1)}}{n - 1}. \tag{1}$$

It is obviously that testing of composite hypothesis of uniformity for sample $X_1, X_2, ..., X_n$ on interval $[\hat{a}, \hat{b}]$ equal to testing of simple hypothesis of uniformity for sample with sample size $n - 2$ on interval $[x_{(1)}, x_{(n)}]$. The required values of order statistics for testing such hypothesis obtained by expressions: $U_{i-1} = \frac{x_{(i)} - x_{(1)}}{x_{(n)} - x_{(1)}}$, $i = \overline{2, (n - 1)}$.

A number of considered tests can be divided into three groups. The first group has statistics based on interval between elements, in most of cases differences between neighbor elements denoted as:

$$D_i = U_i - U_{i-1}, \tag{2}$$

where $U_0 = 0, U_{n+1} = 1, n$ is the size of the sample. In the second group test statistics used difference between theoretical (expected) and empirical data. These tests also called as tests based on the empirical distribution function (EDF tests), and goodness-of-fit tests are contained in this group. The third group has statistics based on entropy estimator. The third group includes the tests based on the entropy estimator.

2 Alternative hypotheses

We compared the power of tests for relatively sample size $n = 10, 20, 30, 40, 50, 100, 150, 200, 300$. Empirical distributions of test statistics under either true null hypothesis or competing hypotheses were found based on 1 660 000 simulations also. The hypothesis under test H_0 was chosen as uniform law. Alternative hypothesis H_i was chosen as beta distribution with the density

$$f(x) = \frac{1}{\theta_2 B(\theta_0, \theta_1)} \left(\frac{x-\theta_3}{\theta_2}\right)^{\theta_0 - 1} \left(1 - \frac{x-\theta_3}{\theta_2}\right)^{\theta_1 - 1}, \tag{3}$$

where $B(\theta_0, \theta_1) = \Gamma(\theta_0)\Gamma(\theta_1)/\Gamma(\theta_0 + \theta_1)$ is beta-function, $\theta_0, \theta_1 \in (0, \infty)$ are parameters the of form, $\theta_2 \in (0, \infty)$ is shape parameter, $\theta_3 \in (-\infty, \infty)$ is bias parameter, $x \in [0, \infty]$. This distribution was chosen because the fact that the standard uniform distribution is a special case of the beta distribution with the parameters of form $\theta_0 = 1$ and $\theta_1 = 1$. We denote the function of beta distribution with values of parameters $B_I(\theta_0, \theta_1, \theta_2, \theta_3)$. So, three alternative hypotheses H_1, H_2, H_3, which are quite close to H_0, can be written by

$$H_1 : F(X) = B_I(1.5, 1.5, 1, 0), x \in [0, 1];$$
$$H_2 : F(X) = B_I(0.8, 1.0, 1, 0), x \in [0, 1];$$
$$H_3 : F(X) = B_I(1.1, 0.9, 1, 0), x \in [0, 1] \ .$$

The distribution functions and the density functions of these hypotheses are presented in Figure 1 and 2, respectively.

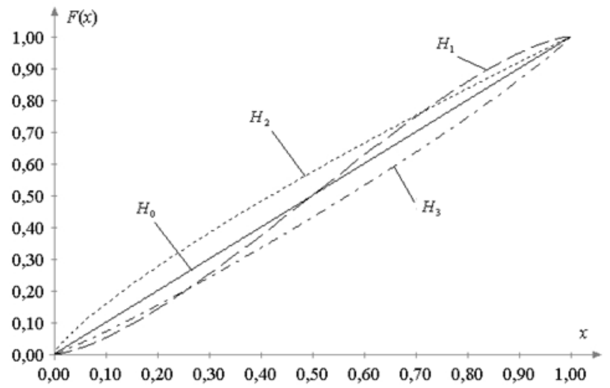

Figure 1: The distribution functions corresponding to the hypotheses

It is worth noting that the distribution function of alternative H_1 crossed the function of the uniform distribution, while the distribution functions of alternatives H_1 and H_3 are located above and below the function of uniform distribution, respectively. And abilities to distinguish hypothesis H_0 from H_1 and from H_2 and H_3 in tests are different. The comparative analysis shows that most of the considered tests have inability to distinguish hypothesis H_0 from H_1 under small sample size n , in other words these tests are biased in such cases.

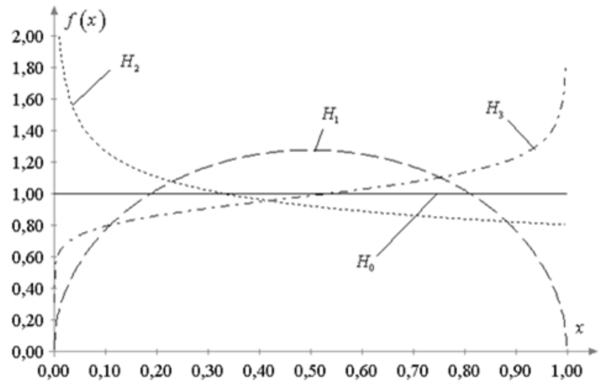

Figure 2: The density functions corresponding to the hypotheses

3 Simulation result

The expressions for statistics of special uniformity tests are presented in Table 1. The Table 2 contains considered tests ordered by decreasing of power (quantity $1-\beta$) under alternatives H_1, H_2 and H_3 ($n = 100$ and $\alpha = 0.1$). The dark mark means that the test is biased under small sample size n, in other words that quantity α larger than $1-\beta$. This bias take a place to a lesser extent in Neyman-Barton tests N_2 and N_3 [14]. This advantage isn't observed only for some tests: Kuper test [9], Watson test [19, 20], Dudewicz-Van Der Mulllen test [5], Cheng-Spiring test [3], Swartz test [18], second Cressie [4] test and chi-squared Pearson test.

Entropy procedure used different entropy estimator gives high power under alternative hypothesis H_1. Whereas their power is relatively worst under alternatives H_2 and H_3. It should be noted that only modifications of entropy test have bias under alternative H_2 for small sample size n. It is recognized that power of these tests and also Cressie tests and Pardo test [15] depends from choosing of parameter m called as window size also.

The Neyman-Barton test N_2 shows good power under H_1 and relatively good power under H_2 and H_3. The Hegazy-Green tests [7] and Frosini test demonstrate consistently good ability to distinguish alternative hypotheses from uniformity distribution. The low powers are shown by tests, the statistics of which use the differences (2) of successive values of order sample $U_i - U_{i-1}$ (Sherman test [17], Kimball test [8], Moran tests [12, 13], Greenwood test [6], Greenwood-Quesenberry-Miller test [16]). The Cheng-Spiring test, demonstrated quite high power under H_1, shows low power under H_2 and H_3. The lowest power is demonstrated by Yang test [22], under all considered alternative hypotheses. Among the non-parametric goodness-of-fit tests, the good powers are obtained by Zhang tests Z_A and Z_C [24], and Anderson-Darling tests [1].

Conclusions

Unfortunately, the distributions of most special uniformity tests depend on the sample size, therefore the researchers must rely on the tables of percent points. The similar issue occurs in using nonparametric goodness-of-fit Zhang tests.

It is found from comparative analysis of tests, which can be used for testing the hypothesis of uniformity, that using of single certain test can be incorrect in forming the reliable statistical inference. The applying more than one test based on different measure of deviation of empirical distribution from theoretical distribution improves the quality of statistical inference. It is better to use some series of tests, which have certain advantages for more objective inferences.

References

[1] Anderson T.W., Darling D.A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*. Vol. **49**, pp. 765-769.

[2] Blinov P.Yu., Lemeshko B.Yu. (2014). Entropy-based test of uniformity. *12th International Conference on Actual Problems of Electronics Instrument Engineering (APEIE)*. Vol. **1**, pp. 540-548.

[3] Cheng S.W., Spiring F.A. (1987). A test to Identify the uniform distribution with applications. *IEEE Transactions on Reliability*. Vol. **36**, pp. 98-105.

[4] Cressie N. (1979). An optimal statistic based on higher order gaps. *Biometrika*. Vol. **66**, pp. 619-627.

[5] Dudewics E.J., van der Meulen E.C. (1981). A review of the properties of tests for uniformity. *Journal of the American Statistical Association*. Vol. **76**, pp. 967-974.

[6] Greenwood V. (1946). The statistical study of Infection disease. *Journal of the Royal Statistical Society: Series A*. Vol. **109**, pp. 257-261.

[7] Hegazy Y.A., Green J.R.S. (1975). Some new goodness-of-fit tests using order statistics. *Applied Statistics.*. Vol. **24**, pp. 299-308.

[8] Kimball B.F. (1947). Some basic theorems for developing tests of fit for the case of the non-parametric probability distribution function. *The Annals of Mathematical Statistics*. Vol. **18**, pp. 540-541.

[9] Kuiper N.H. (1960). Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen. Ser. A.*. Vol. **63**, pp. 38-47.

[10] Lemeshko B.Yu., Gorbunova A.A. (2013). Application and Power of the Nonparametric Kuiper, Watson, and Zhang Tests of Goodness-of-Fit. *Measurement Techniques*. Vol. **56**, pp. 465-475.

[11] Lemeshko B.Yu., Gorbunova A.A, Lemeshko S.B., Rogozhnikov A.P. (2013). Solving problems of using some nonparametric goodness-of-fit tests. *Optoelectronics, Instrumentation and Data Processing.*. Vol. **50**, pp. 21-35.

[12] Moran P.A.P. (1947). The random division of an intervals. *Journal of the Royal Statistical Society: Series B*. Vol. **9**, pp. 92-98.

[13] Moran P.A.P. (1951). The random division of an intervals, II. *Journal of the Royal Statistical Society: Series B*. Vol. **13**, pp. 147-150.

[14] Neyman J. (1937). “Smooth” tests for goodness-of-fit. *Scandinavisk Aktuarietidskrift*. Vol. **20**, pp. 149-199.

[15] Pardo M.C. (2003). A test for uniformity based on informational energy. *Statistical Papers*. Vol. **44**, pp. 521-534.

[16] Quesenberry C.P, Miller F.L. (1977). Power studies of some tests for uniformity. *Journal of Statistical Computation and Simulation*. Vol. **5**, pp. 169-191.

[17] Sherman B.A. (1950). A random variable related to the spacing of sample values. *The Annals of Mathematical Statistics*. Vol. **21**, pp. 339-361.

[18] Swartz T. (1992). Goodness-of-fit tests using Kullback–Leibler information. *Communications in Statistics – Theory and Methods*. Vol. **21**, pp. 711-729.

[19] Watson G.S. (1961). Goodness-of-fit tests on a circle. *Biometrika*. Vol. **48**, pp. 109-114.

[20] Watson G.S. (1962). Goodness-of-fit tests on a circle, II. *Biometrika*. Vol. **49**, pp. 57-63.

[21] Yousefzadeh A., Arghami N.R. (2008). Testing Exponentiality Based on Type II Censored Data and a New cdf Estimator. *Communications in Statistics - Simulation and Computation*. Vol. **37**, pp. 1479-1499.

[22] Young D.L. (1982). The linear nearest neighbour statistic. *Biometrika*. Vol. **69**, pp. 477-480.

[23] Zamanzade E. (2014). Testing uniformity based on new entropy estimators. *Journal of Statistical Computation and Simulation*. pp. 1-15, DOI: 10.1080/00949655.2014.958085.

[24] Zhang J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B*. Vol. **64**, pp. 281-294.

Table 1: Statistics of considered tests for uniformity

Number	Test	Test statistic
1	Sherman	$\omega_n = \frac{1}{2}\sum\limits_{i=1}^{n+1} \lvert D_i - \frac{1}{n+1} \rvert$
2	Kimball	$A = \sum\limits_{i=1}^{n+1} (D_i - \frac{1}{n+1})^2$
3	Moran 1	$B = \sum\limits_{i=1}^{n+1} (D_i)^2$
4	Moran 2	$M_n = -\sum\limits_{i=1}^{n+1} \ln\left[(n+1)D_i\right]$
5	Yang	$M = \frac{1}{l}\sum\limits_{i=1}^{n} \min(D_i, D_{i+1});\ l = b - a$
6	Greenwood	$G = (n+1)\sum\limits_{i=1}^{n+1} (D_i)^2$
7	Greenwood-Qesenberry-Miller	$Q = \sum\limits_{i=1}^{n+1} (D_i)^2 + \sum\limits_{i=1}^{n} (D_{i+1} D_i)$
8	Swartz	$A_n^* = \frac{n}{2}\sum\limits_{i=1}^{n} \left(\frac{U_{i+1} - U_{i-1}}{2} - \frac{1}{n}\right)^2$, where $U_0 = -U_1$, $U_{n+1} = 2 - U_n$
9	Cressie 1	$S_n^{(m)} = \sum\limits_{i=0}^{n+1-m} \left(U_{i+m} - U_i - \frac{m}{n+1}\right)^2,\ m < \frac{n}{2}$
10	Cressie 2	$L_n^{(m)} = \sum\limits_{i=0}^{n+1-m} \ln[\frac{n+1}{m}(U_{i+m} - U_i)],\ m < \frac{n}{2}$
11	Cheng-Spiring	$W_p = \left[(U_n - U_1)\frac{n+1}{n-1}\right]^2 / \sum\limits_{i=1}^{n} \left(U_i - \bar{U}\right)^2$
12	Hegazy-Green T_1	$T_1 = \frac{1}{n}\sum\limits_{i=1}^{n} \lvert U_i - \frac{i}{n+1} \rvert$
13	Hegazy-Green T_1^*	$T_1^* = \frac{1}{n}\sum\limits_{i=1}^{n} \lvert U_i - \frac{i-1}{n-1} \rvert$
14	Hegazy-Green T_2	$T_2 = \frac{1}{n}\sum\limits_{i=1}^{n} \left(U_i - \frac{i}{n+1}\right)^2$
15	Hegazy-Green T_2^*	$T_2^* = \frac{1}{n}\sum\limits_{i=1}^{n} \left(U_i - \frac{i-1}{n-1}\right)^2$
16	Frosini	$B_n = \frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n} \lvert U_i - \frac{i-0.5}{n} \rvert$
17	Neyman-Barton N_k; $k = 2, 3, 4$	$N_k = \sum\limits_{j=1}^{k} V_j^2$, where $V_j = \frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n} \pi_j(U_i - 0.5)$, $\pi_1(y) = 2\sqrt{3}y$; $\pi_2(y) = \sqrt{5}(6y^2 - 0.5)$; $\pi_3(y) = \sqrt{7}(20y^3 - 3y)$; $\pi_4(y) = 3(70y^4 - 15y^2 + 0.375)$

Table 1 (continued)

Number	Test	Test statistic
18	Dudewicz-Van Der Mulen	$H(m,n) = -\frac{1}{n}\sum\limits_{i=1}^{n} \ln\left[\frac{n}{2m}(U_{i+m} - U_{i-m})\right]$, where $m < \frac{n}{2}$; if $i+m \geq n$, then $U_{i+m} = U_n$, and if $i-m \leq 1$, then $U_{i-m} = U_1$
19	Pardo	$E_{m,n} = \frac{1}{n}\sum\limits_{i=1}^{n} \frac{2m}{n(U_{i+m} - U_{i-m})}$
20	The first modification of entropy test [23],	$HY_1 = -\frac{1}{n}\sum\limits_{i=1}^{n} \ln\left(\frac{U_{i+m} - U_{i-m}}{\hat{F}(U_{i+m}) - \hat{F}(U_{i-m})}\right)$, where $\hat{F}(U_i) = \frac{n-1}{n(n+1)}\left(i + \frac{1}{n-1} + \frac{U_i - U_{i-1}}{U_{i+1} - U_{i-1}}\right)$, $i = \overline{2,(n-1)}$, $\hat{F}(U_1) = 1 - \hat{F}(U_n) = \frac{1}{n+1}$
21	The first modification of entropy test [21],	$HY_2 = -\sum\limits_{i=1}^{n} \ln\left(\frac{U_{i+m} - U_{i-m}}{\hat{F}(U_{i+m}) - \hat{F}(U_{i-m})}\right) * \left(\frac{\hat{F}(U_{i+m}) - \hat{F}(U_{i-m})}{\sum\limits_{j=1}^{n}\left(\hat{F}(U_{j+m}) - \hat{F}(U_{j-m})\right)}\right)$

Table 2: The tests ranked by power ($n = 100$, $\alpha = 0.1$)

	hypothesis H_1	$1-\beta$	hypothesis H_2	$1-\beta$	hypothesis H_3	$1-\beta$
1	The second modification of entropy test	0.883	Anderson–Darling	0.648	Anderson–Darling	0.526
2	Zhang Z_A	0.850	Hegazy-Green T_1	0.610	Hegazy-Green T_1	0.522
3	Neyman-Barton N_2	0.837	Zhang Z_C	0.606	Frosini	0.522
4	Cressie 2	0.820	Frosini	0.603	Hegazy-Green T_1^*	0.520
5	Zhang Z_C	0.819	Hegazy-Green T_2	0.602	Hegazy-Green T_2	0.508
6	Dudewicz-Van Der Mulen	0.790	Neyman-Barton N_2	0.597	Kramer-von-Misses-Smirnov	0.507
7	The first modification of entropy test	0.789	Kramer-von-Misses-Smirnov	0.595	Hegazy-Green T_2^*	0.506
8	Watson	0.779	Hegazy-Green T_1^*	0.595	Zhang Z_C	0.463
9	Neyman-Barton N_3	0.766	Zhang Z_K	0.590	Zhang Z_A	0.459
10	Neyman-Barton N_4	0.739	Hegazy-Green T_2^*	0.585	Kolmogorov	0.450

Table 2 (continued)

	hypothesis H_1	$1-\beta$	hypothesis H_2	$1-\beta$	hypothesis H_3	$1-\beta$
11	Kuper	0.736	Neyman-Barton N_3	0.577	Neyman-Barton N_2	0.447
12	Cheng-Spiring	0.722	Zhang Z_A	0.574	Zhang Z_K	0.438
13	Zhang Z_K	0.617	Neyman-Barton N_4	0.557	Neyman-Barton N_3	0.416
14	Pearson χ^2	0.593	Kolmogorov	0.542	Neyman-Barton N_4	0.381
15	Swartz	0.583	Pardo	0.463	Pearson χ^2	0.374
16	Anderson–Darling	0.505	Pearson χ^2	0.448	Pardo	0.291
17	Hegazy-Green T_1^*	0.443	Kuper	0.364	Dudewicz-Van Der Mulen	0.275
18	Hegazy-Green T_2^*	0.409	Watson	0.356	The first modification of entropy test	0.275
19	Pardo	0.408	The first modification of entropy test	0.328	The second modification of entropy test	0.267
20	Frosini	0.384	Dudewicz-Van Der Mulen	0.327	Watson	0.257
21	Kramer-von-Misses-Smirnov	0.358	Cressie 1	0.314	Kuper	0.254
22	Hegazy-Green T_1	0.322	The second modification of entropy test	0.266	Cressie 2	0.226
23	Kolmogorov	0.322	Greenwood-Qesenberry-Miller	0.244	Cressie 1	0.218
24	Hegazy-Green T_2	0.308	Swartz	0.226	Swartz	0.206
25	Greenwood-Qesenberry-Miller	0.290	Cressie 2	0.217	Greenwood-Qesenberry-Miller	0.186
26	Kimball	0.279	Sherman	0.204	Kimball	0.165
27	Moran 1	0.279	Kimball	0.201	Moran 1	0.165
28	Greenwood	0.279	Moran 1	0.201	Greenwood	0.165
29	Sherman	0.215	Greenwood	0.201	Sherman	0.154
30	Cressie 1	0.187	Moran 2	0.193	Moran 2	0.143
31	Moran 2	0.187	Cheng-Spiring	0.168	Cheng-Spiring	0.106
32	Yang	0.115	Yang	0.108	Yang	0.104