APPLIED METHODS OF STATISTICAL ANALYSIS.

STATISTICAL COMPUTATION AND SIMULATION

PROCEEDINGS

OF THE INTERNATIONAL WORKSHOP

18-20 September 2019

Novosibirsk

2019

UDC 519.22(063)
A 67

E d i t o r s:

Prof. Boris Lemeshko, Prof. Mikhail Nikulin,
Prof. Narayanaswamy Balakrishnan

A 67 **Applied Methods of Statistical Analysis. Statistical Computation and Simulation** - AMSA'2019, Novosibirsk, Russia, 18-20 September, 2019: Proceedings of the International Workshop. - Novosibirsk: NSTU publisher, 2019. - 574 pp.
ISSN 2313-870X

ISSN 2313-870X **UDC 519.22(063)**

Applied Methods of Statistical Analysis. Statistical Computation and Simulation

Scientific Program Committee:

A. Vostretsov	Novosibirsk State Technical University, Russia
B. Lemeshko	Novosibirsk State Technical University, Russia
N. Balakrishnan	McMaster University, Canada
A. Medvedev	Siberian Federal University, Russia
G. Koshkin	Tomsk State University, Russia
A. Antonov,	Institute of Nuclear Power Engineering, Russia
E. Chimitova,	Novosibirsk State Technical University, Russia
Yu. Dmitriev,	Tomsk State University, Russia
M. Krnjajić,	National University of Ireland, Ireland
H. Liero,	University of Potsdam, Germany
N. Limnios,	Université de Technologie de Compiègne, France
G. Mikhailov,	Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
V. Melas,	St. Petersburg State University, Russia
V. Ogorodnikov,	Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
B. Ryabko,	Siberian State University of Telecommunications and Information Sciences, Russia
V. Rykov,	Institute of Computational Technologies SB RAS, Russia
V. Solev,	St.Petersburg Department of Steklov Mathematical Institute RAS, Russia
F. Tarasenko,	Tomsk State University, Russia
V. Timofeev	Novosibirsk State Technical University, Russia

Local Organizing Committee:

Ekaterina Chimitova, Evgenia Osintseva, Mariia Semenova

Preface

The Fifth International Workshop "Applied Methods of Statistical Analysis. Statistical Computation and Simulation" – AMSA'2019 is organized by Novosibirsk State Technical University.

The first two Workshops AMSA'2011 and AMSA'2013, as well as AMSA'2019, took place in Novosibirsk. AMSA'2015 was held in the resort Belokurikha located at the foothills of Altai. AMSA'2017, organized together with Siberian State University of Science and Technologies called after academician M.F. Reshetnev, took place in Krasnoyarsk.

The First Workshop "Applied Methods of Statistical Analysis" AMSA'2011 was focused on Simulations and Statistical Inference, AMSA'2013 – on Applications in Survival Analysis, Reliability and Quality Control, AMSA'2015 – on Nonparametric Approach and AMSA'2017 – on Nonparametric Methods in Cybernetics and System Analysis.

The Workshop AMSA'2019 was mainly oriented to the discussion of problems of Statistical Computation and Simulation, which are crucial for the development of methods of applied mathematical statistics and their effective application in practice.

The Workshop proceedings would certainly be interesting and useful for specialists, who use statistical methods for data analysis in various applied problems arising from engineering, biology, medicine, quality control, social sciences, economics and business. The Proceedings of International Workshop "Applied Methods of Statistical Analysis" are indexed in Scopus starting with 2017 materials.

The organization of the Fifth International Workshop "Applied Methods of Statistical Analysis. Statistical Computation and Simulation" – AMSA'2019 was supported by the Russian Ministry of Education and Science (project 1.1009.2017/4.6).

Prof. Boris Lemeshko

Contents

On Application of k-samples homogeneity tests

Boris Yu. Lemeshko and Irina V. Veretelnikova
Novosibirsk State Technical University, Novosibirsk, Russian Federation
e-mail: Lemeshko@ami.nstu.ru, ira-veterok@mail.ru

Abstract

New k-samples homogeneity tests based on the Smirnov, Lehmann-Rosenblatt and Anderson-Darling two-sample tests have been proposed. The maximum value of the statistics of the 2-sample test obtained during the analysis of combinations of pairs of samples is considered as a statistic of k-sample test. The constructed models for limit distributions of statistics of the proposed tests for $k = 3, \cdots, 11$ are given. Comparative analysis of the power of the set of k-samples tests, including the Zhang test, has been carried out. Power estimates of the studied tests are presented in relation to some competing hypotheses, which allows to order k-sample tests by preference with respect to different alternatives.

Keywords: k-samples tests, homogeneity tests, test statistic, distribution of statistics, power of test.

Introduction

The necessity of solving the task of checking the hypotheses of two (or more) samples of random values belonging to the same universe estimates (the homogeneity test) may arise in different areas. For example, this task may arise naturally when checking the measurement means and trying to be certain that the random measurement errors distribution law has not undergone any serious changes within some time period.

The task of testing the homogeneity of k-samples can be stated as follows. We have x_{ij}, where j is the observation in the set of order statistics of i-sample $j = \overline{1, n_i}, i = \overline{1, k}$. Let us assume that the i-sample correlates with the continuous distribution function of $F_i(x)$. It is required to test the hypothesis of $H_0 : F_1(x) = F_2(x) = \cdots = F_k(x)$ type without defining the common distribution law.

The general approach to the construction of k-sample homogeneity tests which are the counterparts of the two-sample Kolmogorov-Smirnov and Cramer-von Mises (Lehmann-Rosenblatt) tests, was considered in [1]. Under this approach, the statistics of the criterion is a measure of deviation of empirical distributions corresponding to specific samples from the empirical distribution based on the totality of the analyzed samples. The k-selective variant of the Kolmogorov–Smirnov test based on this principle is mentioned in [2, 3]. The k-selective version of the Anderson-Darling test is proposed in [4]. The homogeneity tests constructed by Zhang in [5, 6, 7] are the development of the homogeneity tests by Smirnov [8], Lehmann-Rosenblatt [9, 10] and Anderson–Darling [11] and allow us to analyze samples.

The application of k-samples tests in practice is constrained by the fact that, at best, only critical values of statistics for the relevant ones are known, as in the case of the Anderson-Darling test [4] or Kolmogorov-Smirnov tests [2, 12], and the possibility

of using Zhang's criteria rests on the need to look for the distribution of test statistics (or estimation of the achieved significance level p_{value}) using statistical modeling in order to form a conclusion about the results of the hypothesis test.

The only exception is the homogeneity test χ^2 for which the asymptotic distributions of statistics are known with the truth of H_0.

In the present work we illustrate the dependence of the distributions of statistics of the k-sample tests on the sample sizes and the number of k compared samples. For the k-sample Anderson–Darling test [4] we give models of limit distributions of statistics constructed by us [13, 14, 15]. Suggested variants of k-sample tests based on the use of 2-sample Smirnov test [8], Lehmann-Rosenblatt test [9, 10] and Anderson-Darling test [11], and present the constructed model for the limit distributions of the statistics of the proposed test for various k. The constructed models make it possible to carry out correct and informative conclusions with the calculation of p_{value} with the usage of the corresponding criteria. In addition, we present estimates of the power of the test considered with respect to some competing hypotheses, which allows us to organize the k-sample tests by preference with respect to various alternatives.

The studies were based on the intensive use of the Monte Carlo method in the simulation of distributions of tests statistics.

1 k-samples homogeneity tests

1.1 Anderson-Darling test

The Anderson-Darling k-sample test is proposed in [4]. Let us denote the empirical distribution function corresponding to the i^{th} sample $F_{in_i}(x)$, and the empirical distribution function corresponding to the combined sample volume $n = \sum_{i=1}^{k} n_i$ as $H_n(x)$. Statistics of the Anderson-Darling sample test (AD) is defined by the expression

$$A^2_{kn} = \sum_{i=1}^{k} n_i \int_{B_n} \frac{[F_{in_i}(x) - H_n(x)]^2}{(1 - H_n(x))H_n(x)} dH_n(x),$$

where $B_n = x \in R : H_n(x) < 1$. Under the assumption of continuity of $F_i(x)$ on the ordered combined sample $X_1 \leq X_2 \cdots = X_n$ in [4] this simple expression for the calculation of statistics is obtained:

$$A^2_{kn} = \frac{1}{n} \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{n-1} \frac{(nM_{ij} - jn_i)^2}{j(n-j)},$$

where M_{ij} is number of elements in i^{th} sample which are not larger than X_j. The hypothesis H_0 being tested is rejected for large values of statistics.

The statistics acquires the following final form in [4]:

$$T_{kn} = \frac{A^2_{kn} - (k-1)}{\sqrt{D[A^2_{kn}]}}. \tag{1}$$

where the dispersion is determined by the following expression [4]

$$D[A_{kn}^2] = \frac{an^3+bn^2+cn+d}{(n-1)(n-2)(n-3)}$$

with

$$a = (4g-6)(k-1) + (10-6g)H,$$
$$b = (2g-4)k^2 + 8hk + (2g-14h-4)H - 8h + 4g - 6,$$
$$c = (6h+2g-2)k^2 + (4h-4g+6)k + (2h-6)H + 4h,$$
$$d = (2h+6)k^2 - 4hk,$$

where

$$H = \sum_{i=1}^{k} \frac{1}{n_i}, h = \sum_{i=1}^{n-1} \frac{1}{i}, g = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \frac{1}{(n-i)j}.$$

Asymptotic (limiting) distributions of statistics (1) depend on the k-number of samples compared and do not depend on n_i. With the growth of k the distribution of statistics (1) slowly converges to the standard normal law.

In [4] for statistics (1) the table of critical values has been constructed for a number of k. Based on the results of statistical modeling, we built models of limiting distributions of statistics (1) for [13, 14, 15]. The laws of the family of beta-distributions of the III type with density turned out to be good models when having the density of

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 B(\theta_0,\theta_1)} \left[\frac{x-\theta_4}{\theta_3}\right]^{\theta_0-1} \left[1-\frac{x-\theta_4}{\theta_3}\right]^{\theta_1-1} / \left[1+(\theta_2-1)\frac{x-\theta_4}{\theta_3}\right]^{\theta_0+\theta_1}, \quad (2)$$

as shown in Table 1 as $B_{III}(\theta_0,\theta_1,\theta_2,\theta_3,\theta_4)$ having exact values for this law's parameters. These models are based on simulated samples of statistics with the number of simulation experiments $N = 10^6$ and $n_i = 10^3$.

1.2 Zhang test

The Zhang tests [5, 6, 7] allow comparing $k \geq 2$ samples.

Let $x_{i1}, x_{i2}, \cdots, x_{in_i}$ be ordered samples of continuous random variables with distribution functions $F_i(x)$, $(i = \overline{1,k})$ and, as previously, $X_1 < X_2 < \cdots < X_n$, where $n = \sum_{i=1}^{k} n_i$, is the unified ordered sample. Let us define the R_{ij} rank of the j^{th} ordered observation x_{ij} of the i^{th} sample in the unified sample. Let $X_0 = -\infty, X_{n+1} = +\infty$, and the ranks $R_{i,0} = 1, R_{i,n_i+1} = n+1$.

In the tests a modification of the empirical distribution function $\hat{F}(t)$ is used, having the values of $\hat{F}(X_m) = (m-0.5)/n$ at break points $X_m, m = \overline{1,n}$ [5].

The Z_K statistic of the Zhang homogeneity test is of the following form [5]:

$$Z_K = \max_{1\leq m\leq n} \left\{ \sum_{i=1}^{k} n_i \left[F_{i,m} \ln \frac{F_{i,m}}{F_m} + (1-F_{i,m}) \ln \frac{1-F_{i,m}}{1-F_m} \right] \right\}, \quad (3)$$

Table 1: Models of the limiting distributions of statistics (1)

k	Model
2	B_{III}(3.1575, 2.8730, 18.1238, 15.0000, -1.1600)
3	B_{III}(3.5907, 4.5984, 7.8040, 14.1310, -1.5000)
4	B_{III}(4.2657, 5.7035, 5.3533, 12.8243, -1.7500)
5	B_{III}(6.2992, 6.5558, 5.6833, 13.010, -2.0640)
6	B_{III}(6.7446, 7.1047, 5.0450, 12.8562, -2.2000)
7	B_{III}(6.7615, 7.4823, 4.0083, 11.800, -2.3150)
8	B_{III}(5.8057, 7.8755, 2.9244, 10.900, -2.3100)
9	B_{III}(9.0736, 7.4112, 4.1072, 10.800, -2.6310)
10	B_{III}(10.2571, 7.9758, 4.1383, 11.186, -2.7988)
11	B_{III}(10.6848, 7.5950, 4.2041, 10.734, -2.8400)
∞	N(0.0,1.0)

where $F_m = \hat{F}(X_m)$, so that $F_m = (m-0.5)/n$, and the calculation $F_{i,m} = \hat{F}_i(X_m)$ is done as follows. At the initial moment $j_i = 0, i = \overline{1,k}$. If $R_{i,j_i+1} = m$, then $j_i := j_i+1$ and $F_{i,m} = (j_i - 0.5)/n_i$, otherwise, with $R_{i,j_i} < m < R_{i,j_i+1}$, $F_{i,m} = j_i/n_i$.

This is a *right-hand* test: the hypothesis H_0 being tested is rejected at *high* statistical values (3).

Statistic Z_A of the homogeneity test of k samples is defined by the following expression [5]:

$$Z_A = -\sum_{m=1}^{n}\sum_{i=1}^{k} n_i \frac{F_{i,m}\ln F_{i,m} + (1-F_{i,m})\ln(1-F_{i,m})}{(m-0.5)(n-m+0.5)}, \tag{4}$$

where F_m and $F_{i,m}$ are calculated as shown above.

This is a *left-side* test: the hypothesis H_0 being tested is rejected for *small* values of statistics (4).

Distributions of the statistic (4) depend on the sample volume and the number of samples compared as well.

Statistic Z_C of the homogeneity test of k samples is defined by the following expression [5]:

$$Z_C = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \ln\left(\frac{n_i}{j-0.5} - 1\right)\ln\left(\frac{n}{R_{i,j}-0.5} - 1\right). \tag{5}$$

This is also a *left-hand* test: the tested hypothesis H_0 is rejected at *small* values of the statistic (5). The distributions $G(Z_C \mid H_0)$ of the statistic depend on the sample volume and the number of samples under analysis in the similar way.

The dependence of the distributions of statistics (3) - (5) of the volume of the samples complicates the use of the Zhang test since there are problems with the calculation of the evaluation of p_{value}.

At the same time, the lack of information on the laws of distribution of statistics and tables of critical values in modern conditions is not a serious disadvantage of the tests as it is easy to calculate the achieved levels of significance of p_{value} with the software that supports the application of the tests, merely using statistic simulating methods.

1.3 k-samples Tests Based on 2-sample Ones

In order to analyze the k-samples it is possible to apply a two-sample test with the S statistic to each pair (totaling $(k-1)k/2$ pairs), and the decision on accepting or rejecting the H_0 hypothesis will be made on the strength of all results. The following statistic can be taken as a statistic of this k-sample tests (when having a right-hand two-sample criterion):

$$S_{max} = \max_{\substack{1\le i\le k\\ i<j\le k}} \{S_{ij}\}, \tag{6}$$

where S_{ij} are the values of the statistics of the used two-sample criterion as calculated in the course of analysis of the i^{th} and the j^{th} samples.

The hypothesis H_0 to be tested will be rejected at **large** values of statistics S_{max}. The advantage of this kind of test is that as a result a pair of samples will be determined, the difference between them being the most significant from the standpoint of the two-sample test used.

Statistics of the two-sample Smirnov, Lehmann-Rosenblatt and Anderson-Darling tests can be used as S_{ij}. In this case the distributions of the relevant statistics S_{max} converge to some limiting ones, models of which can be found on the results of statistical modeling.

1.3.1 Smirnov Maximum Test

The D_{n_2,n_1} statistic used in the Smirnov test is calculated according to the following formulae [8]:

$$D^+_{n_2,n_1} = \max_{1\le r\le n_2} [\tfrac{r}{n_2} - F_{1,n_1}(x_{2r})] = \max_{1\le s\le n_1} [F_{2,n_2}(x_{2s}) - \tfrac{s-1}{n_1}],$$

$$D^-_{n_2,n_1} = \max_{1\le r\le n_2} [F_{1,n_1}(x_{2r}) - \tfrac{r-1}{n_2}] = \max_{1\le s\le n_1} [\tfrac{s}{n_1} - F_{2,n_2}(x_{1s})],$$

$$D_{n_2,n_1} = \max(D^+_{n_2,n_1}, D^-_{n_2,n_1}).$$

With the H_0 hypothesis being true and with unlimited increase of the number of samples the statistic

$$S_C = \sqrt{\frac{n_1 n_2}{n_1+n_2}} D_{n_2,n_1} \tag{7}$$

will in the limit fall with the Kolmogorov arrangement of $K(S)$ [8].

In case of using the k-samples variant of the Smirnov test as S_{ij} in (6) it seems more preferable to use a modification of the Smirnov statistic

$$S_{mod} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\left(D_{n_2,n_1} + \frac{n_1 + n_2}{4.6 n_1 n_2}\right), \tag{8}$$

its distribution being always closer to the limiting distribution of Kolmogorov $K(S)$ [16]. Statistic S_{max} will be defined as $S_{max}^{S_m}$ in this case.

With equal volumes of samples under comparison the statistic distributions $S_{max}^{S_m}$ will be of substantial discreteness (similar to the two-sample case, see Fig. 1) and be different from the asymptotic (limiting) distributions (see Fig. 2). If possible, it is preferable to use co-primes as n_i, then the distributions $G(S \mid H_0)$ of the $S_{max}^{S_m}$ statistic will not be actually different from the asymptotic ones.

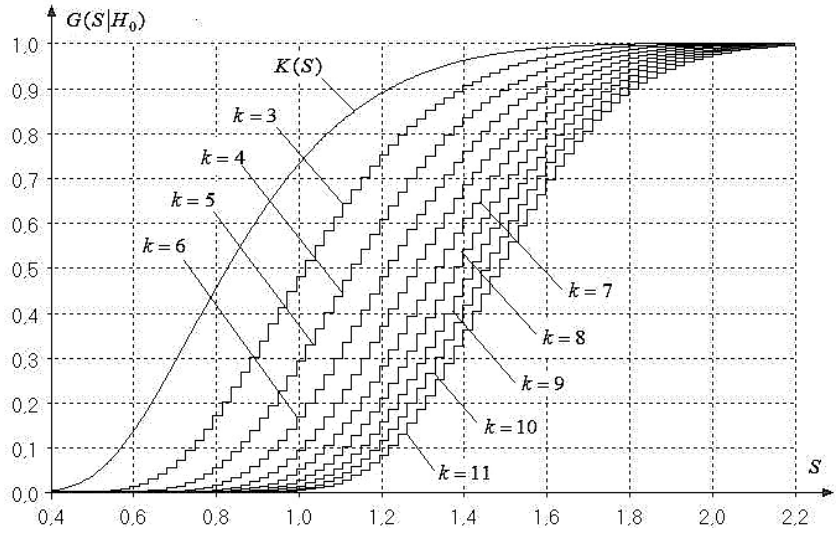

Figure 1: Statistic distributions with $n_i = 1000, i = \overline{1,k}$

Models of asymptotic $S_{max}^{S_m}$ statistic distributions with $k = 3 \div 11$ in the form of beta distributions of the III type (2) $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ having exact values of parameters and constructed in this paper based on the results of statistic modeling are shown in Table 2.

1.3.2 Lehman-Rosenblatt Maximum Test

Statistic of the two-sample Lehmann-Rosenblatt test as introduced in [9] is used in the following form [8]:

$$T = \frac{1}{n_1 n_2 (n_1 + n_2)}\left(n_2 \sum_{i=1}^{n_2} (r_i - i)^2 + n_1 \sum_{j=1}^{n_1} (s_j - j)^2\right) - \frac{4 n_1 n_2 - 1}{6(n_1 + n_2)}, \tag{9}$$

where r_i is the numerical order (rank) of x_{2i}; s_j is the numerical order (rank) of x_{1i} in the unified ordered series. In [10] it was shown that the statistic (9) at the limit is distributed as $a1(t)$ [8].

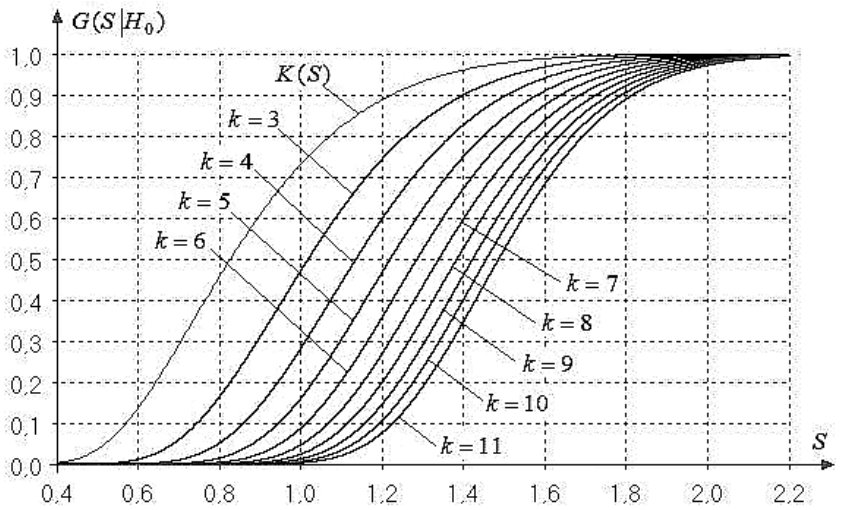

Figure 2: Asymptotic statistic distributions $S_{max}^{S_m}$

Table 2: Models of the limiting distributions of statistics $S_{max}^{S_m}$

k	Model
2	$K(S)$
3	B_{III}(6.3274, 6.6162, 2.8238, 2.4073, 0.4100)
4	B_{III}(7.2729, 7.2061, 2.6170, 2.3775, 0.4740)
5	B_{III}(7.1318, 7.3365, 2.4813, 2.3353, 0.5630)
6	B_{III}(7.0755, 8.0449, 2.3163, 2.3818, 0.6320)
7	B_{III}(7.7347, 8.6845, 2.3492, 2.4479, 0.6675)
8	B_{III}(7.8162, 8.9073, 2.2688, 2.4161, 0.7120)
9	B_{III}(7.8436, 8.8805, 2.1696, 2.3309, 0.7500)
10	B_{III}(7.8756, 8.9051, 2.1977, 2.3280, 0.7900)
11	B_{III}(7.9122, 9.0411, 2.1173, 2.2860, 0.8200)

In the case of using the k-samples variant of the Lehman-Rosenblatt test as S_{ij} in the statistic S_{max}^{LR} of form (6) statistic (9) is used. Dependence of distributions of statistic S_{max}^{LR} on the number of samples with H_0 being true is illustrated in Fig. 3.

The constructed models of asymptotic (limiting) distributions of statistic S_{max}^{LR} with the number of compared samples $k = 3 \div 11$ are shown in Table 3. In this case the Sb-Johnson distributions proved to be the best with the density of

$$f(x) = \frac{\theta_1\theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)} \exp\left\{-\frac{1}{2}\left[\theta_0 - \theta_1 \ln \frac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$$

with exact values of this law's parameters, the law being shown in Table 3 as $Sb(\theta_0, \theta_1, \theta_2, \theta_3)$. These represented models allow finding the estimates of p_{value} by the values of statistic S_{max}^{LR} with corresponding k number of samples under comparison.

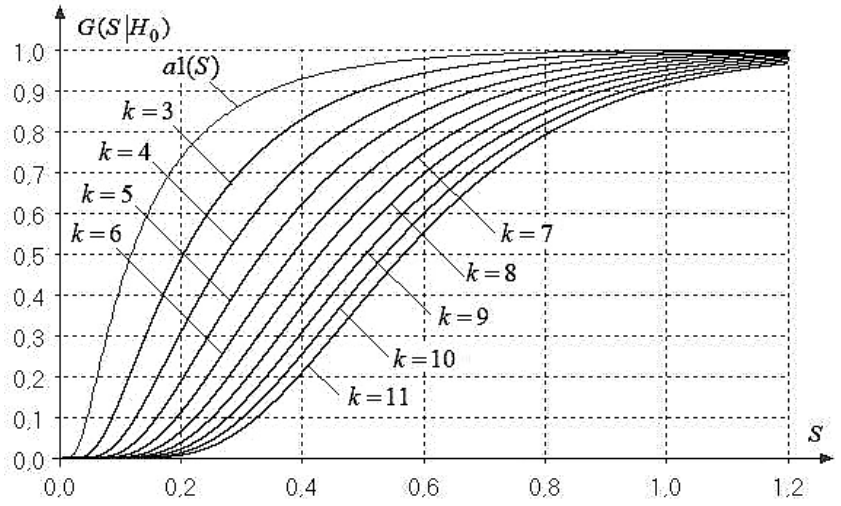

Figure 3: Distributions of statistic S_{max}^{LR}

Table 3: Models of the limiting distributions of statistics S_{max}^{LR}

k	Model
2	$a1(t)$
3	$Sb(3.2854, 1.2036, 3.0000, 0.0215)$
4	$Sb(2.5801, 1.2167, 2.2367, 0.0356)$
5	$Sb(3.1719, 1.4134, 3.1500, 0.0320)$
6	$Sb(2.9979, 1.4768, 2.9850, 0.0380)$
7	$Sb(3.2030, 1.5526, 3.4050, 0.0450)$
8	$Sb(3.2671, 1.6302, 3.5522, 0.0470)$
9	$Sb(3.4548, 1.7127, 3.8800, 0.0490)$
10	$Sb(3.4887, 1.7729, 3.9680, 0.0510)$
11	$Sb(3.4627, 1.8168, 3.9680, 0.0544)$

1.3.3 Anderson-Darling Maximum Test

The Anderson-Darling two-sample test was dealt with in [11]. This test's statistic is defined by the following expression:

$$A^2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1+n_2-1} \frac{(M_i(n_1+n_2) - n_1 i)^2}{i(n_1+n_2-i)}, \tag{10}$$

where M_i is the number of elements of the first sample, smaller or equal to the i^{th} element of the variation set of the unified sample. Distribution $a2(t)$ will be the limiting distribution (10) with the tested hypothesis H_0 being true [8].

In the case of using the k-samples variant of the Anderson-Darling test as S_{ij} in the S_{max}^{AD} statistic (6) statistic (10) will be used. Dependence of distributions of

statistic S^{AD}_{max} on the number of samples with H_0 being true is shown in Fig. 4.

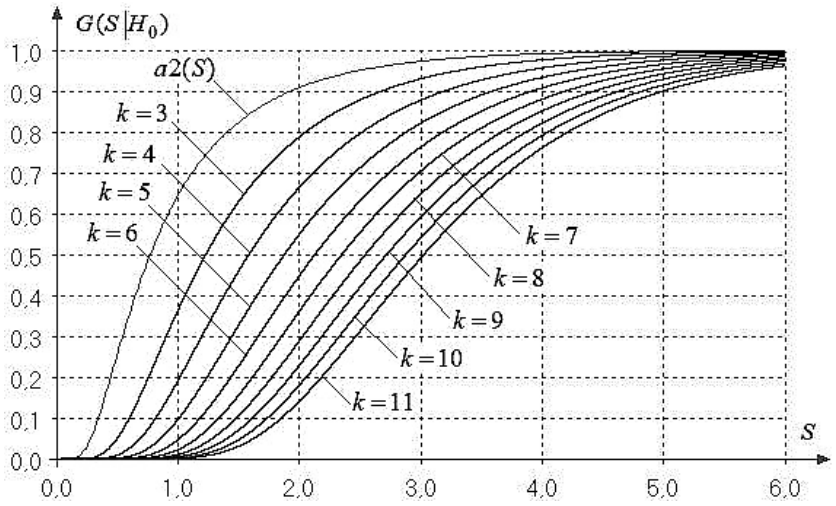

Figure 4: Distributions of statistic S^{AD}_{max}

Models of asymptotic (limiting) distributions of statistic S^{AD}_{max} for the k number of samples under comparison $k = 3 \div 11$ have been constructed for distributions $G(S^{AD}_{max} \mid H_0)$ and shown in Table 4. In this case the beta distributions of the III type proved to be the best (2) as shown as $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ with exact values of parameters shown in Table 4; these can be used for estimating p_{value} with the k number of compared samples.

Table 4: Models of the limiting distributions of statistics S^{AD}_{max}

k	Model
2	$a2(t)$
3	$B_{III}(4.4325, 2.7425, 12.1134, 8.500, 0.1850)$
4	$B_{III}(5.2036, 3.2160, 10.7792, 10.000, 0.2320)$
5	$B_{III}(5.7527, 3.3017, 9.7365, 10.000, 0.3000)$
6	$B_{III}(5.5739, 3.4939, 7.7710, 10.000, 0.3750)$
7	$B_{III}(6.4892, 3.6656, 8.0529, 10.500, 0.3920)$
8	$B_{III}(6.3877, 3.8143, 7.3602, 10.800, 0.4800)$
9	$B_{III}(6.7910, 3.9858, 7.1280, 11.100, 0.5150)$
10	$B_{III}(6.7533, 4.2779, 6.6457, 11.700, 0.5800)$
11	$B_{III}(7.1745, 4.3469, 6.6161, 11.800, 0.6100)$

1.4 Homogeneity Test χ^2

The homogeneity test χ^2 can successfully be used to analyze $k \geq 2$ samples. In this case the common area of the samples is split into r intervals (groups). Let η_{ij} be the

number of elements of the i^{th} sample of the j^{th} interval, then $n_i = \sum_{j=1}^{r} \eta_{ij}$.

The χ^2 homogeneity test statistic will be of the following form:

$$\chi^2 = n \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{(\eta_{ij} - \nu_j n_i/n)^2}{\nu_j n_i} = n\left(\sum_{i=1}^{k} \sum_{j=1}^{r} \frac{\eta_{ij}^2}{\nu_j n_i} - 1 \right), \tag{11}$$

where $\nu_j = \sum_{l=1}^{k} \eta_{lj}$ is the total number of elements of all samples falling into the j^{th} interval. The χ^2-distribution with the number of degrees of freedom $(k-1)(r-1)$ shall be the asymptotic distribution of statistic [17].

2 Comparative analysis of powers

One of the main characteristics of the statistical test is its power relative to a given competing hypothesis H_1. The power is the remainder of $1-\beta$, where β is the possibility of type II error (accept hypothesis H_0 with H_1 being true) at specified probability α of type I error (reject H_0 when true).

The power of k-samples tests was investigated for various k and situations when the tested hypothesis H_0 was whether all samples belonged to the standard normal law, the competing hypothesis H_1 being if all samples but the last one belonged to the standard normal law and the last sample belonged to the normal law with the shift parameter $\theta_0 = 0.1$ and the scale parameter $\theta_1 = 1$; hypothesis H_2 being that the last sample belonged to the normal law with the shift parameter $\theta_0 = 0$ and the scale parameter $\theta_1 = 1.1$, the competing hypothesis H_3 being the last sample belonged to the logistic law with the density of

$$f(x) = \tfrac{1}{\theta_1\sqrt{3}} \exp\{-\tfrac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\}/[1+\exp\{-\tfrac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\}]^2$$

and parameters $\theta_0 = 0$ and $\theta_1 = 1$.

The power was evaluated on the results of modeling statistic distributions with the tested $G(S \mid H_0)$ being true, and competing hypotheses $G(S \mid H_1)$, $G(S \mid H_2)$ and $G(S \mid H_3)$ having equal volumes of n_i compared samples. As an example, Tables 5 and 6 show evaluation of the power of tests with $\alpha = 0.1$ for $k = 3$ and $k = 4$ correspondingly. In the case of the homogeneity test χ^2 the unified sample was split into $r = 10$ equifrequent intervals.

Thus-conducted power analysis of k-samples tests allows making some conclusions.

The tests can be organized power-wise with respect to changes in the shift parameter in the following way:

$$S_{max}^{AD} \succ AD \succ S_{max}^{LR} \succ S_{max}^{Sm} \succ Z_C \succ Z_A \succ Z_K \succ \chi^2.$$

With respect to changes in the scale parameter:

$$Z_C \succ Z_A \succ Z_K \succ AD \succ \chi^2 \succ S_{max}^{AD} \succ S_{max}^{Sm} \succ S_{max}^{LR}.$$

Table 5: Assessment of the power of test against alternatives H_1, H_2 and H_3, $k = 3$, $n_i = n$

Test	$n_i = 20$	$n_i = 50$	$n_i = 100$	$n_i = 300$	$n_i = 500$	$n_i = 10^3$
Against alternative hypothesis H_1						
S^{AD}_{max}	0.113	0.134	0.171	0.314	0.450	0.712
AD	0.113	0.134	0.171	0.313	0.449	0.711
S^{LR}_{max}	0.114	0.134	0.168	0.306	0.437	0.694
S^{Sm}_{max}	0.110	0.128	0.155	0.272	0.383	0.622
Z_C	0.113	0.131	0.160	0.273	0.380	0.612
Z_A	0.112	0.130	0.158	0.268	0.371	0.599
Z_K	0.110	0.125	0.144	0.231	0.321	0.525
χ^2	0.100	0.108	0.120	0.173	0.226	0.385
Against alternative hypothesis H_2						
Z_C	0.107	0.125	0.160	0.319	0.475	0.771
Z_A	0.107	0.126	0.162	0.319	0.470	0.767
Z_K	0.107	0.123	0.147	0.263	0.376	0.621
AD	0.104	0.111	0.124	0.191	0.273	0.509
χ^2	0.105	0.114	0.129	0.202	0.277	0.495
S^{AD}_{max}	0.102	0.107	0.114	0.165	0.231	0.446
S^{Sm}_{max}	0.103	0.104	0.114	0.136	0.164	0.253
S^{LR}_{max}	0.103	0.104	0.108	0.127	0.152	0.241
Against alternative hypothesis H_3						
Z_A	0.103	0.108	0.116	0.181	0.279	0.580
Z_C	0.103	0.108	0.116	0.176	0.270	0.568
Z_K	0.104	0.110	0.117	0.170	0.233	0.423
χ^2	0.100	0.113	0.121	0.173	0.226	0.382
AD	0.103	0.107	0.114	0.148	0.189	0.315
S^{Sm}_{max}	0.102	0.105	0.111	0.148	0.183	0.288
S^{AD}_{max}	0.102	0.104	0.110	0.134	0.166	0.272
S^{LR}_{max}	0.103	0.104	0.107	0.124	0.145	0.218

At that, the Zhang tests of Z_A and Z_C statistics are almost equivalent power-wise, and the Anderson-Darling test is noticeably inferior to the Zhang tests.

The tests can be organized power-wise with respect to situations when all but one sample belongs to the normal law and the last one belongs to the logistic law, in the following way:

$$Z_A \succ Z_C \succ Z_K \succ \chi^2 \succ AD \succ S^{Sm}_{max} \succ S^{AD}_{max} \succ S^{LR}_{max}.$$

It can be noted that with the increase in the number of compared samples of the same volumes the power of the criterion relative to similar competing hypotheses

decreases as a rule, which is absolutely natural. It is more difficult to single out a situation and to give preference to a competing hypothesis, when only one of the analyzed samples belongs to some other law. We can't but mention that the Zhang tests with statistics of Z_K, Z_A, Z_C possess quite substantial advantage in power with respect to some alternatives.

Table 6: Assessment of the power of test against alternatives H_1, H_2 and H_3, $k = 4$, $n_i = n$

Test	$n_i = 20$	$n_i = 50$	$n_i = 100$	$n_i = 300$	$n_i = 500$	$n_i = 10^3$
Against alternative hypothesis H_1						
S_{max}^{AD}	0.112	0.131	0.165	0.302	0.438	0.706
AD	0.112	0.131	0.164	0.301	0.433	0.701
S_{max}^{LR}	0.113	0.130	0.162	0.293	0.425	0.686
S_{max}^{Sm}	0.111	0.125	0.151	0.261	0.366	0.605
Z_C	0.111	0.126	0.155	0.260	0.368	0.595
Z_A	0.111	0.127	0.153	0.255	0.360	0.579
Z_K	0.109	0.121	0.141	0.219	0.300	0.502
χ^2	0.102	0.109	0.118	0.167	0.221	0.358
Against alternative hypothesis H_2						
Z_C	0.106	0.122	0.158	0.306	0.468	0.761
Z_A	0.107	0.124	0.158	0.305	0.463	0.745
Z_K	0.106	0.120	0.145	0.249	0.367	0.606
AD	0.104	0.110	0.123	0.180	0.254	0.474
χ^2	0.107	0.113	0.127	0.189	0.271	0.458
S_{max}^{AD}	0.101	0.104	0.111	0.145	0.195	0.381
S_{max}^{Sm}	0.102	0.105	0.108	0.128	0.153	0.221
S_{max}^{LR}	0.102	0.103	0.105	0.118	0.135	0.197
Against alternative hypothesis H_3						
Z_A	0.103	0.107	0.116	0.179	0.274	0.566
Z_C	0.103	0.107	0.115	0.173	0.257	0.555
Z_K	0.103	0.107	0.114	0.161	0.222	0.410
χ^2	0.102	0.110	0.116	0.164	0.218	0.357
AD	0.102	0.106	0.113	0.143	0.179	0.291
S_{max}^{Sm}	0.103	0.104	0.112	0.138	0.166	0.257
S_{max}^{AD}	0.101	0.103	0.107	0.124	0.147	0.229
S_{max}^{LR}	0.102	0.102	0.105	0.116	0.130	0.183

Conclusions

The constructed models of statistic limiting distributions for k-samples homogeneity tests (the Anderson-Darling ones and those proposed in this paper) allows obtaining correct and informational conclusions on and calculating the tests significance p_{value}. Software can is available for this purpose [18].

Funding

The studies were carried out with the support of the Ministry of Education and Science of the Russian Federation in the framework of the state work 'Ensuring the conduct of scientific research'(No. 1.4574.2017 / 6.7) and the design part of the state task (No. 1.1009.2017 / 4.6).

References

[1] Kiefer J. (1959). K-samples Analogues of the Kolmogorov–Smirnov and Cramer-v. Mises Tests. *Annals of Mathematical Statistics*. Vol. **30. No. 2.**, – pp. 420-447. URL: http://www.jstor.org/ stable/2237091

[2] Conover W.J.(1965). Several k-samples Kolmogorov–Smirnov tests. *The Annals of Mathematical Statistics*. Vol. **36, No. 3**, pp. 1019-1026. URL: http://www.jstor.org/ stable/2238210

[3] Conover W.J. (1999). *Practical Nonparametric Statistics – 3d ed.*. Wiley.

[4] Scholz F.W., Stephens M.A. (1987). K-samples Anderson–Darling Tests. *Journal of the American Statistical Association*. Vol. **82. No. 399**, pp. 918-924. DOI: 10.1080/ 01621459.1987.10478517

[5] Zhang J. (2001). *Powerful goodness-of-fit and multi-sample tests*. PhD Thesis. York University, Toronto. URL: http://www.collectionscanada.gc.ca/ obj/s4/f2/dsk3/ ftp05/ NQ66371.pdf (accessed 28.01.2013).

[6] Zhang J. (2006). Powerful Two-Sample Tests Based on the Likelihood Ratio. *Technometrics*. Vol. **48. No. 1**, pp. 95-103. DOI: 10.1198/ 004017005000000328

[7] Zhang J., Wu Y. (2007). k-samples tests based on the likelihood ratio. *Computational Statistics & Data Analysis*. Vol. **51. No. 9**, pp. 4682-4691. DOI: 10.1016/ j.csda.2006.08.029

[8] Bolshev L.N., Smirnov N.V. (1983). *Tables of mathematical statistic*. M.: Nauka.

[9] Lehmann E.L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statist*. Vol. **22, № 1**,pp. 165–179. URL: http://www.jstor.org/stable/2236420

[10] Rosenblatt M. (1952). Limit theorems associated with variants of the von Mises statistic. *Ann. Math. Statist*. Vol. **23**, pp. 617–623. URL: https://projecteuclid.org/euclid.aoms/1177729341

[11] Pettitt A.N. (1976). A two-sample Anderson-Darling rank statistic.*Biometrika*. Vol. **63. No.1**, pp. 161-168. DOI: 10.1093/biomet/63.1.161

[12] Wolf E.H., Naus J.I. (1973). Tables of critical values for a k-sample Kolmogorov–Smirnov test statistic.*J. Amer. Statist. Assoc*. Vol. **68**, pp. 994–997. DOI: 10.1080/01621459.1973.10481462

[13] Lemeshko B.Y. (2017).*Tests for homogeneity. Guide on the application*. M: INFRA–M. DOI: 10.12737/22368

[14] Lemeshko B.Y., Lemeshko S.B., Veretelnikova I.V. (2017). On application of distribution laws homogeneity tests.*Tomsk State University Journal of Control and Computer Science*. No. **41**, pp. 24-31. DOI: 10.17223/19988605/41/3

[15] Lemeshko B.Y., Veretelnikova I.V., Lemeshko S.B., Novikova A.Y. (2017). Application of Homogeneity Tests: Problems and Solution.*In: Rykov V., Singpurwalla N., Zubkov A. (eds) Analytical and Computational Methods in Probability Theory. ACMPT 2017. Lecture Notes in Computer Science*. Vol. **10684**.

[16] Lemeshko B.Y., Lemeshko S.B. (2005). Statistical distribution convergence and homogeneity test power for Smirnov and Lehmann–Rosenblatt tests.*Measurement Techniques*. Vol. **48,No.12**, pp. 1159-1166. DOI: 10.1007/s11018-006-0038-3

[17] Cramer H. (1975).*Mathematical methods of statistics*. M.: Mir.

[18] ISW-Software for statistic analysis of one-dimensional observations. https://ami.nstu.ru/ headrd/ISW.htm. (accessed date 06.06.2019)