DOI 10.1007/s11018-017-1141-3
Measurement Techniques, Vol. 60, No. 1, April, 2017

GENERAL PROBLEMS OF METROLOGY AND MEASUREMENT TECHNIQUE

APPLICATION AND POWER OF PARAMETRIC CRITERIA FOR TESTING THE HOMOGENEITY OF VARIANCES. PART III

B. Yu. Lemeshko and T. S. Sataeva

UDC 519.24

The distributions of the statistics of parametric tests (Neyman–Pearson, O'Brien, Link, Newman, Bliss–Cochran–Tukey, Cadwell–Leslie–Brown, and the Overall–Woodward Z-variance and modified Z-variance tests) are studied, including the case in which the standard assumption of normality is violated. A comparative analysis of the power of the set of parametric tests is carried out.

Keywords: *test, statistic, homogeneity of variances, distribution of a statistic, power of tests.*

This article is a continuation of a cycle of studies of criteria for testing the homogeneity of variances and of estimates of the power of these tests for different distribution laws and relatively small sample sizes by computer simulation. Tests of the hypothesis of the homogeneity of variances are used for analyzing the results of measurements. In metrology, they are often used in comparisons of laboratory tests. The test hypothesis of uniformity of variances in the case of m samples has the form

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2. \tag{1}$$

The competing hypothesis is usually considered to be

$$H_1 : \sigma_{i_1}^2 \neq \sigma_{i_2}^2, \tag{2}$$

where the inequality is satisfied for at least one pair of subscripts i_1, i_2. Some of these tests can be used only for $m = 2$.

The quality of the statistical conclusions resulting from the test is ensured by the correctness of the tests with higher power that are applied. The standard assumption for applicability of the classical parametric tests is that the analyzed samples adhere to a normal law. Nonparametric tests are not subject to this kind of limit.

The power and properties of parametric (Bartlett, Cochran, Fisher, Hartley, Levene) and nonparametric (Ansari–Bradley, Mood, Siegel–Tukey, Capon, Klotz) tests have been compared in [1–4], including under conditions such that the standard assumption is violated. It was shown that, for $m = 2$, the parametric Bartlett, Cochran, Fisher, and Hartley tests are equivalent, and, for $m > 2$, the Cochran test is to be preferred. Here the power of the parametric tests is considerably higher than that of their nonparametric analogs, so it is necessary to reconsider the feasibility of using the parametric tests when the assumption of normality is violated [3, 5]. The sphere of applications for the nonparametric tests is limited by the assumption that the samples to be analyzed obey the same distribution law [2, 3].

In this paper, the conclusions of [1–4] are supplemented by studies of the following parametric tests of homogeneity of variances: Neyman–Pearson [6], O'Brien [7], Link (ratio of ranges) [8], Newman (studentized range) [9], Bliss–Cochran–Tukey [10], Cadwell–Leslie–Brown [11], Overall–Woodward Z-test [12], and the modified Overall–Woodward Z-test [13].

Novosibirsk State Technical University, Novosibirsk, Russia; e-mail: lemeshko@ami.nstu.ru. Translated from Izmeritel'naya Tekhnika, No. 1, pp. 8–13, January, 2017. Original article submitted July 27, 2016.

The distributions of the statistics have been studied and the powers of the tests relative to different competing hypotheses have been estimated using statistical modelling based on the ISW program package [14, 15]. $N = 10^6$ runs of the statistical simulations were carried out for modelling this set of statistics. As a rule, for these values of N the absolute value of the difference between the true distribution of the statistic and the empirically simulated distribution does not exceed 10^{-3}.

When the assumption of normality was violated, the distributions of the statistics of the tests were studied for the case in which the samples belong to a generalized normal law with density

$$f(x) = \frac{\theta_0}{2\theta_1 \Gamma(1/\theta_0)} \exp\left[-\left(|x-\theta_2|/\theta_1\right)^{\theta_0}\right], \tag{3}$$

where $\theta_0, \theta_1 \in (0, \infty)$, $\theta_2 \in (-\infty, \infty)$, and $x \in (-\infty, \infty)$.

The distribution of the statistics of the tests was estimated for different values of the parameter θ_0. Some special cases of this family of distributions include the normal law for $\theta_0 = 2$ and the Laplace distribution for $\theta_0 = 1$. When the shape factor θ_0 is smaller, the tails of distribution (3) are heavier and, vice versa, when this parameter is larger, the tails are lighter. This analysis of the distributions of the statistics of parametric tests of the homogeneity of dispersions shows that these distributions depend strongly on the form of the observed distribution law. Besides the Levene test [2], only the O'Brien test and the modified Z-test manifested a certain stability.

The Neyman–Pearson test (likelihood ratio test). The statistic for this test is given by the ratio of the arithmetic mean for all the estimates s_i^2 of the variances to their geometric mean[6]:

$$h = \frac{1}{m}\sum_{i=1}^{m} s_i^2 \Bigg/ \left(\prod_{i=1}^{m} s_i^2\right)^{1/m}, \tag{4}$$

where m is the number of samples; n_i are the volumes of the samples;

$$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

are the estimates of the sample variances;

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

are the sample means; and x_{ij} is the jth element of the ith sample. It is assumed that $n_1 = n_2 = \ldots = n_m = n$. The hypothesis to be tested, H_0 (see Eq. (1)), is rejected for large values of statistic (4) such that $h > h_{1-\alpha}$ (α is a specified level of significance).

Refined studies of the percentage points (assuming normality) and the dependence of distribution (4) on n and m are described in [16]. Tests with statistic (4), as in the case of any of the tests discussed below, can also be used for unequal n_i, but then the distributions of the statistics (when hypothesis (1) is true) will differ from the distributions for equal n_i. The tests are extremely sensitive to failure of the assumption of normality.

The O'Brien test. For calculating the statistic for the test [7] each jth element of the ith sample, x_{ij}, is transformed in accordance with the formula

$$V_{ij} = [(n_i - 1.5)n_i(x_{ij} - \bar{x}_i)^2 - 0.5 s_i^2 (n_i - 1)] \big/ [(n_i - 1)(n_i - 2)],$$

where $\bar{x}_i$ is the average value.

The test statistic has the form

$$V = \frac{1}{m-1}\sum_{i=1}^{m} n_i(\bar{V}_i - \bar{\bar{V}}_i)^2 \Bigg/ \frac{1}{N-m}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(V_{ij} - \bar{V}_i)^2, \tag{5}$$

where

$$\bar{V}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} V_{ij}; \quad \bar{\bar{V}}_i = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} V_{ij}; \quad N = \sum_{i=1}^{m} n_i .$$

The test is right-sided and the test hypothesis H_0 (see Eq. (1)) is rejected for large values of the (5) statistic. The limiting distribution of the statistic for the O'Brien test when H_0 holds is the Fisher $F_{m-1,N-m}$-distribution with $m-1$ and $N-m$ degrees of freedom [7]. Our studies showed that the distributions of statistic (5) converge fairly slowly to $F_{m-1,N-m}$-distributions. For example, in the case of $m = 2$, the difference between the real distribution $G(V|H_0)$ of the statistic (5) and the $F_{1,N-m}$-distribution can be neglected only for $n_1 = n_2 = n \geq 80$. For small sample volumes, a significant difference between the distribution $G(V|H_0)$ of the statistic and the $F_{m-1,N-m}$-distribution is observed for large values of V. Thus, using the percentage points of the $F_{m-1,N-m}$-distribution leads to an increase in the probability of type II β errors.

For $N - m \leq 80$, the distributions $G(V|H_0)$ for values of the statistic V such that $1 - G(V|H_0) < 0.1$ are closer to the $F_{m-1,\infty}$-distribution than to the $F_{m-1,N-m}$-distribution. For $N - m \leq 80$, the correctness of the conclusions can be increased if the $F_{m-1,\infty}$-distribution is used to estimate the attained level of significance $p_{\text{value}} = 1 - F_{m-1,\infty}(V)$, or if we rely on the upper critical values of the statistic (with different m for $n_1 = n_2 = n \leq 80$) [16].

The distributions of the statistic for the O'Brien test (like the Levene test [2]) are fairly stable with respect to failure of the standard assumption of normality. Deviations toward distributions with lighter tails than the normal law essentially have no effect on the distribution of the statistic. For distributions with heavier tails, the deviations from the distributions corresponding to a normal law are not so large as in the case of other parametric tests.

The Link test (ratio of ranges). This test is an analog of the Fisher test. It is used only for analyzing two samples. The statistic for the test is defined as [8]

$$F^* = \omega_{n_1} / \omega_{n_2}, \tag{6}$$

where $\omega_{n_1} = x_{1\max} - x_{1\min}, \omega_{n_2} = x_{2\max} - x_{2\min}$ are the spreads; and $x_{1\max}, x_{2\max}, x_{1\min}, x_{2\min}$ are, respectively, the maximum and minimum elements in the samples being compared.

The test hypothesis is rejected with a significance level of α if $F^* > F^*_{1-\alpha/2}$ or $F^* < F^*_{\alpha/2}$, where $F^*_{1-\alpha/2}, F^*_{\alpha/2}$ are the upper and lower critical values of the statistic. The distribution of the test statistic depends substantially on the volumes of the samples being compared. Refined lower and upper percentage points for statistic (6) in the case where the samples obey a normal law for $n_1, n_2 \leq 20$ are given in [16]. This test is extremely sensitive to any deviations from normality.

The Newman test (studentized range). The test statistic is [9]

$$q = \omega_{n_1} / s_{n_2}, \tag{7}$$

where $\omega_{n_1} = x_{1\max} - x_{1\min}; s_{n_2} = \left[\frac{1}{n_2 - 1}\sum_{i=1}^{n_2}\left(x_{2i} - \bar{x}_2\right)^2\right]^{1/2}$.

The (1) tested hypothesis of equality of the variances is rejected if $q < q_{\alpha/2}$ or $q > q_{1-\alpha/2}$, where $q_{\alpha/2}, q_{1-\alpha/2}$ are, respectively, the lower and upper critical values of the statistic for a given level of significance α. Refined lower and upper critical values of statistic (7) are given in Ref. 16. This test has the same shortcomings as the Link test.

The Bliss–Cochran–Tukey test. The statistic for this test [10] proposed as an analog of the Cochran test is given by

$$c = \max_{1 \leq i \leq m} \omega_i \Big/ \sum_{i=1}^{m} \omega_i ,$$

where $\omega_i = \max_{1 \leq j \leq n_i} x_{ij} - \min_{1 \leq j \leq n_i} x_{ij}$ is the spread in the ith sample.

If the statistic $c > c_{1-\alpha}$, where $c_{1-\alpha}$ is the upper critical value for a specified level of significance α, then the test hypothesis H_0 of equal dispersions is rejected. The upper critical values of the statistic for volumes $n_1 = n_2 = n \leq 20$ and $m \leq 10$ when the normality assumption is satisfied are given in [16].

The distributions of the test statistic depend strongly on the volume and quality of the samples being compared and change when the normality assumption fails.

TABLE 1. Power of Tests with Respect to the Hypothesis H_1: $\sigma_2 = 1.1\sigma_0$

Test	α	Sample volume n				
		10	20	40	60	100
Bartlett, Cochran, Hartley, Fisher, Neyman–Pearson, Z-test	0.1	0.112	0.128	0.157	0.188	0.246
	0.05	0.058	0.068	0.090	0.111	0.156
	0.01	0.012	0.016	0.023	0.032	0.051
Modified Z-test, O'Brien	0.1	0.109	0.125	0.154	0.184	0.243
	0.05	0.056	0.066	0.087	0.108	0.153
	0.01	0.012	0.015	0.022	0.030	0.049
Link	0.1	0.110	0.123	0.150	0.176	0.228
	0.05	0.056	0.065	0.084	0.103	0.141
	0.01	0.012	0.014	0.021	0.028	0.044
Newman	0.1	0.111	0.123	0.143	0.159	0.186
	0.05	0.57	0.066	0.080	0.091	0.112
	0.01	0.012	0.015	0.020	0.025	0.033
Bliss–Cochran–Tukey, Cadwell–Leslie–Brown, Link	0.1	0.111	0.119	0.133	0.141	0.154
	0.05	0.057	0.063	0.072	0.078	0.087
	0.01	0.012	0.014	0.018	0.019	0.023
Mood	0.1	0.111	0.120	0.43	0.166	0.211
	0.05	0.057	0.064	0.080	0.096	0.128
	0.01	0.012	0.014	0.020	0.026	0.039
Ansari–Bradley	0.1	0.101	0.125	0.135	0.154	0.190
	0.05	0.052	0.064	0.074	0.087	0.113
	0.01	0.011	0.014	0.019	0.023	0.033
Siegel–Tukey	0.1	0.106	0.121	0.135	0.154	0.190
	0.05	0.055	0.062	0.075	0.087	0.113
	0.01	0.011	0.010	0.018	0.023	0.033

The Cadwell–Leslie–Brown test. The statistic for this test, proposed as an analog of the Hartley test, is [11]

$$K = \max_{1 \le i \le m} \omega_i \Big/ \min_{1 \le i \le m} \omega_i . \tag{8}$$

The test hypothesis H_0 is rejected for $K > K_{1-\alpha}$, where $K_{1-\alpha}$ is the upper critical value of the statistic. Refined critical values $K_{1-\alpha}$ of statistic (8) for $m \le 10$ samples with equal sample volumes $n_1 = n_2 = n \le 20$, $i = 1, ..., m$, are given in [16]. The properties of the test are similar to those of the preceding test.

The Overall–Woodward Z-test. The statistic for this test is given by [12]

$$Z = \frac{1}{m-1} \sum_{i=1}^{m} Z_i^2 , \tag{9}$$

where

$$Z_i = [c_i(n_i - 1)s_i^2 / MSE]^{1/2} - [c_i(n_i - 1) - c_i / 2]^{1/2};$$

$$MSE = \frac{1}{N-m}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2; \quad c_i = 2 + 1/n_i; \quad N = \sum_{i=1}^{m} n_i .$$

When the test hypothesis in (1) is true and the analyzed samples are normally distributed, the limiting distribution of statistic in (9) is the Fisher $F_{m-1,\infty}$-distribution. For small n_i, however, distribution (9) differs substantially from the $F_{m-1,\infty}$-distribution. The difference between the real distribution of the statistic and the $F_{m-1,\infty}$-distribution can be neglected for $n_i \geq 50$. When $n_i \leq 50$, tables of the upper critical values of $Z_{1-\alpha}$ [16] can be used. The distribution of the statistic in (9) for the Z-test is very sensitive to failure of the normality assumption.

The modified *Z*-test. As a more stable test, a modified statistic has been proposed [13] with

$$c_i = 2.0\left[K_i^{-1}(2.9 + 0.2 / n_i)\right]^{1.6(n_i - 1.8K_i + 14.7)/n_i}, \tag{10}$$

where

$$K_i = \frac{1}{n_i - 2}\sum_{j=1}^{n_i} G_{ij}^4$$

is the estimated excess coefficient for the ith sample and

$$G_{ij} = (x_{ij} - \bar{x}_i) / [s_i^2(n_i - 1) / n_i]^{1/2}.$$

Our studies showed that the distribution of the modified statistic converges to the $F_{m-1,\infty}$-distribution very slowly as n_i increases. Even for large sample volumes, the distribution of the modified statistics is not consistent with the $F_{m-1,\infty}$-distribution, although in the region of large values the difference between it and the $F_{m-1,\infty}$-distribution is insignificant. For a correct application of the test with small sample volumes, a table [16] of the critical values can be used.

The distribution of the statistic for the modified Z-test actually does have greater stability with respect to failure of the standard assumption of normality. Equation (10), however, does not have the required accuracy, since, in the case of the modified Z-test, the monotonic dependence of the distributions of the statistics on the degree of deviation of the observed distribution from normal common to all parametric tests breaks down.

Analysis of the power of the tests. The power of the individual tests has been examined in [1–3, 17–19]. In this article, for a comparative analysis, as competition to the test hypothesis H_0, we consider the situation where $m - 1$ of the samples belong to a distribution law with some $\sigma = \sigma_0$ and one of the samples, say number m, belongs to a law with another σ (H_1: $\sigma_m = 1.1\sigma_0$; H_2: $\sigma_m = 1.5\sigma_0$). Besides the previously examined tests, we compare the Bartlett, Cochran, Hartley, Fisher, and Levene distributions, as well as the nonparametric tests of Mood, Ansari–Bradley, and Siegel–Tukey, the powers of which are taken from Refs. 2 and 3. The estimated powers of the entire set of tests with adherence to normality (for $\alpha = 0.1, 0.05, 0.01$ and $m = 2$) are listed in Tables 1 and 2, where the tests are ordered in terms of decreasing power.

The Neyman–Pearson and Overall–Woodward Z-test were equivalent in power to the Bartlett, Cochran, Hartley, and Fisher tests. A difference in the powers of the modified Z-test and the O'Brien test is noticeable only for a rather distant competing hypothesis H_2 (see Eq. (2)). There they have an advantage in power over the Levene test.

The Newman test falls behind the Levene test in terms of power more noticeably as the sample volumes increase. At the same time, it has an obvious advantage (except for the size of volume $n = 10$) over the Bliss–Cochran–Tukey, Cadwell–Leslie–Brown, and Link tests, which are equivalent in power.

The group of "stable" tests (modified Z-test, O'Brien test, and Levene test) for small sample sizes (see for $n = 10$) has less power than the Newman, Link, Bliss–Cochran–Tukey, and Cadwell–Leslie–Brown tests, but with increasing n it has an obvious advantage over these and over the nonparametric tests. Within this group, the O'Brien test has some remaining advantage.

The Newman, Bliss–Cochran–Tukey, Cadwell–Leslie–Brown, and Link tests have a slight advantage in power over the nonparametric tests only for small sample sizes ($n = 10$–20), but they fall significantly behind the nonparametric tests as the sizes increase.

TABLE 2. Power of Tests with Respect to the Hypothesis H_2: $\sigma_2 = 1.5\sigma_0$

Test	α	Sample volume n				
		10	20	40	60	100
Bartlett, Cochran, Hartley, Fisher, Neyman–Pearson, Z-test	0.1	0.312	0.532	0.806	0.926	0.991
	0.05	0.201	0.402	0.705	0.871	0.980
	0.01	0.064	0.182	0.463	0.692	0.924
O'Brien	0.1	0.266	0.490	0.783	0.917	0.990
	0.05	0.155	0.344	0.664	0.849	0.976
	0.01	0.039	0.127	0.379	0.628	0.903
Modified Z-test	0.1	0.265	0.489	0.781	0.916	0.990
	0.05	0.158	0.348	0.666	0.849	0.976
	0.01	0.043	0.138	0.397	0.639	0.906
Link	0.1	0.269	0.471	0.746	0.888	0.981
	0.05	0.163	0.338	0.628	0.812	0.960
	0.01	0.045	0.131	0.364	0.590	0.866
Newman	0.1	0.296	0.473	0.682	0.796	0.901
	0.05	0.190	0.348	0.566	0.699	0.840
	0.01	0.060	0.153	0.326	0.473	0.667
Bliss–Cochran–Tukey, Cadwell–Leslie–Brown, Link	0.1	0.285	0.425	0.584	0.674	0.776
	0.05	0.181	0.305	0.458	0.554	0.671
	0.01	0.057	0.127	0.237	0.314	0.430
Mood	0.1	0.255	0.425	0.688	0.841	0.964
	0.05	0.158	0.302	0.565	0.751	0.931
	0.01	0.045	0.121	0.319	0.518	0.802
Ansari–Bradley	0.1	0.242	0.393	0.608	0.768	0.926
	0.05	0.150	0.270	0.484	0.659	0.869
	0.01	0.041	0.104	0.254	0.413	0.693
Siegel–Tukey	0.1	0.246	0.383	0.609	0.768	0.926
	0.05	0.155	0.261	0.484	0.659	0.869
	0.01	0.043	0.056	0.251	0.414	0.693

The Bartlett, Cochran, Hartley, Levene, Neyman–Pearson, O'Brien, Bliss–Cochran–Tukey, and Cadwell–Leslie–Brown tests, as well as the Overall–Woodward Z-test and modified Z-test, can be used for $m > 2$. Here, the Bartlett, Cochran, Hartley, Neyman–Pearson, and Overall–Woodward Z-tests no longer form a group of equivalent tests with the same power. The Bartlett and Neyman–Pearson tests, which are still essentially equivalent in terms of power, are the only exceptions.

Tables 3 and 4 list the estimated powers of the multi-sample tests (for $m = 3$, $m = 5$ and for $n_i = 100$, $i = 1, ..., m$) with respect to the competing hypotheses H_1 and H_2 (with a distinct variance for one of the samples) where the analyzed samples are normally distributed. These tests are listed in order of decreasing power, so that the preferability of one test over another can be established.

TABLE 3. Power of m-Sample Tests with Respect to Hypothesis H_1

Test	m = 3 for different α			m = 5 for different α		
	0.1	0.05	0.01	0.1	0.05	0.01
Cochran	0.250	0.161	0.056	0.241	0.156	0.056
O'Brien	0.243	0.153	0.051	0.230	0.144	0.048
Z-test	0.243	0.153	0.051	0.227	0.141	0.046
Neyman–Pearson, Bartlett	0.242	0.152	0.049	0.224	0.138	0.044
Modified *Z*-test	0.240	0.150	0.048	0.223	0.137	0.044
Hartley	0.239	0.148	0.046	0.219	0.133	0.040
Levene	0.225	0.139	0.043	0.209	0.127	0.039
Cadwell–Leslie–Brown	0.149	0.083	0.021	0.139	0.075	0.018
Bliss–Cochran–Tukey	0.147	0.082	0.021	0.136	0.075	0.019

TABLE 4. Power of m-Sample Tests with Respect to Hypothesis H_2

Test	m = 3 for different α			m = 5 for different α		
	0.1	0.05	0.01	0.1	0.05	0.01
Cochran	0.997	0.994	0.974	0.998	0.997	0.987
O'Brien	0.996	0.990	0.961	0.997	0.994	0.976
Z-test	0.996	0.991	0.964	0.997	0.993	0.974
Neyman–Pearson, Bartlett	0.996	0.990	0.962	0.996	0.992	0.970
Modified *Z*-test	0.995	0.989	0.955	0.996	0.991	0.967
Hartley	0.995	0.988	0.947	0.995	0.989	0.955
Levene	0.990	0.979	0.926	0.991	0.982	0.944
Cadwell–Leslie–Brown	0.820	0.728	0.501	0.829	0.742	0.524
Bliss–Cochran–Tukey	0.795	0.691	0.444	0.783	0.675	0.432

The first ranking test is the Cochran test, with a clear advantage in terms of power, as in [2]. The second ranked test is the O'Brien test, but it does not have significant advantages over the Overall–Woodward *Z*-tests or the Neyman–Pearson and Bartlett tests for analyzing three samples and close competing hypotheses. At the same time, the O'Brien test is more powerful than the modified *Z*-test and the Levene test, which are also stable with respect to violation of the standard assumption of normality. When the competing hypotheses are eliminated, the power of the Bliss–Cochran–Tukey test is higher than that of the Cadwell–Leslie–Brown.

Conclusion. The need to test hypotheses of homogeneity of variances arises during processing of groups of measurement results, including the analysis of results from interlaboratory comparisons. The data shown here can be used to choose the most appropriate test.

It is correct to use the set of parametric tests examined here when the analyzed samples adhere to a normal distribution. When this assumption fails, the method recommended in Ref. 16 can be employed to ensure that they are used correctly.

These studies were supported by the Ministry of Education and Science of the Russian Federation as part of the State Program on Support for Scientific Research and the project part of the State Assignment (Project No. 2.541.2014.K).

REFERENCES

1. B. Yu. Lemeshko and E. P. Mirkin, “Bartlett and Cochran tests in measurements with probability laws different from normal,” *Izmer. Tekhn.*, No. 10, 10–16 (2004).
2. B. Yu. Lemeshko, S. B. Lemeshko, and A. A. Gorbunova, “Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria,” *Izmer. Tekhn.*, No. 3, 10–16 (2010).
3. B. Yu. Lemeshko, S. B. Lemeshko, and A. A. Gorbunova, “Application and power of criteria for testing the homogeneity of variances. Part II. Nonparametric criteria,” *Izmer. Tekhn.*, No. 5, 11–18 (2010).
4. B. Yu. Lemeshko, S. B. Lemeshko, S. N. Postovalov, and E. V. Chimitova, *Statistical Analysis of Data, Modelling and Study of Probability Behavior. A Computer Approach: Monograph*, Izd. NGTU, Novosibirsk (2011).
5. A. A. Gorbunova and B. Yu. Lemeshko, “Application of parametric inhomogeneity of variances tests under violation of classical assumption,” in: *Proc. 2nd Int. Conf. on Stochastic Modeling Techniques and Data Analysis*, June 5–8, 2012, Chania, Greece, pp. 253–260.
6. A. I. Kobzar’, *Applied Mathematical Statistics for Engineers and Scientific Workers*, FIZMATLIT, Moscow (2006).
7. R. G. O’Brien, “Robust techniques for testing heterogeneity of variance effects in factorial designs,” *Psychometrika*, **43**, No. 3, 327–342 (1978).
8. R. F. Link, “The sampling distribution of the ratio of two ranges from independent samples,” *Ann. Math. Stat.*, **21**, No. 1, 112–116 (1950).
9. D. Newman, “The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation,” *Biometrika*, **32**, 20–30 (1939).
10. C. I. Bliss, W. G. Cochran, and J. W. Tukey, “A rejection criterion based upon the range,” *Biometrika*, **43**, 418–422 (1956).
11. R. T. Leslie and B. M. Brown, “Use of range in testing heterogeneity of variance,” *Biometrika*, **53**, 221–227 (1966).
12. J. E. Overall and J. A. Woodward, “A simple test for heterogeneity of variance in complex factorial design,” *Psychometrika*, **39**, 311–318 (1974).
13. J. E. Overall and J. A. Woodward, *A Robust and Powerful Test for Heterogeneity of Variance*, Univ. of Texas Med. Branch Psychometric Lab. (1976).
14. B. Yu. Lemeshko and S. N. Postovalov, *Computer Technology for Data Analysis and Study of Statistical Behavior: Textbook*, NGTU, Novosibirsk (2004).
15. *ISW-Program System for Statistical Analysis of Univariate Random Quantities*, http://ami.nstu.ru/~headrd/ISW.htm, acces. Jan. 23, 2016.
16. B. Yu. Lemeshko, *Tests for Hypotheses of Homogeneity: Applic. Handbook*, INFRA-M, Moscow (2017).
17. H. B. Lee, G. S. Katz, and A. F. Restori, “A Monte Carlo study of seven homogeneity of variance tests,” *J. Math. Stat.*, **6**, No. 3, 359–366 (2010).
18. W. J. Conover, M. E. Johnson, and M. M. Johnson, “A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data,” *Technometrics*, **23**, No. 4, 351–361 (1981).
19. T. S. Lim and W. Y. Loh, “A comparison of tests of equality of variances,” *Comp. Stat. Data Anal.*, **22**, No. 5, 287–301 (1996).