Измерительная техника. – 2010. – №5. – С. 11-18

УДК 519.24

О применении и мощности критериев проверки однородности дисперсий. Ч. II. Непараметрические критерии

Лемешко Б.Ю., Лемешко С.Б., Горбунова А.А.

Новосибирский государственный технический университет, Новосибирск, Россия,
e-mail: Lemeshko@fpm.ami.nstu.ru

Исследованы распределения и мощность непараметрических критериев однородности характеристик рассеяния (Ансари-Бредли, Муда, Сижела-Тьюки, Кейпена и Клотца). Проведен сравнительный анализ мощности с классическими критериями однородности дисперсий (Фишера, Бартлетта, Кокрена, Хартли, Левене). Приведены таблицы процентных точек для критерия Кокрена в случае законов, отличных от нормального.

Ключевые слова: непараметрический критерий, критерий Ансари-Бредли, критерий Муда, критерий Сижела-Тьюки, критерий Кейпена, критерий Клотца, мощность критерия.

Distributions and power of nonparametric test of variances homogeneity are researched (Ansari-Bradley test, Mood's test, Siegel-Tukey test, Capon's test, Klotz's test). The comparative analysis of power is made with classical tests of variance homogeneity (Fisher's, Bartlett's, Cochran's, Hartley's, Levene's tests). Tables of percentage points for Cochran's test in case of the laws which are distinct from normal are presented.

Key words: nonparametric test, Ansari-Bradley test, Mood's test, Siegel-Tukey test, Capon's test, Klotz's test, power of test.

Настоящая работа является продолжением [1]; здесь рассмотрены непараметрические (ранговые) критерии Ансари-Бредли [2], Муда [3], Сижела-Тьюки [4], Кейпена [5] и Клотца [6]. Перечисленные критерии предназначены для проверки гипотез об однородности параметров масштаба законов, соответствующих анализируемым выборкам. Как правило, характеристики рассеяния и, следовательно, стандартное отклонение σ пропорционально параметру масштаба закона. Поэтому рассматриваемые критерии можно считать непараметрическими аналогами критериев однородности дисперсий.

В рассматриваемых критериях сравниваются только две выборки. Поэтому проверяемая гипотеза об однородности дисперсий будет иметь вид

$$H_0 : \sigma_1^2 = \sigma_2^2, \tag{1}$$

а конкурирующая с ней –

$$H_1 : \sigma_1^2 \neq \sigma_2^2. \tag{2}$$

Для того, чтобы можно было сопоставить мощность непараметрических критериев с мощностью критериев Бартлетта, Кокрена, Фишера, Хартли и Левене в данном случае рассматриваются те же конкурирующие гипотезы (H_1: $\sigma_2 = 1{,}1\sigma_1$; H_2: $\sigma_2 = 1{,}2\sigma_1$; H_3: $\sigma_2 = 1{,}5\sigma_1$), что и в [1].

Отметим, что исследование распределений статистик, оценка мощности критериев относительно рассматриваемых конкурирующих гипотез проводилось с использованием методики статистического моделирования [7] и развиваемой на базе [8] программная система ISW (Интервальная статистика под Windows). При этом объем моделируемых выборок исследуемых статистик, как и прежде, составлял величину $N = 10^6$, чтобы обеспечить величину модуля разности между истинным законом распределения статистики и смоделированным эмпирическим не более 10^{-3}.

Исследования распределений статистик проводились при различных параметрических моделях законов, но для корректности сопоставления с результатами, полученными в [1], в данном случае при моделировании выборок также используется семейство распределений с плотностью

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

при различных значениях параметра формы θ_0. Распределение $De(\theta_0)$ включает в качестве частных случаев распределение Лапласа $(\theta_0 = 1)$ и нормальное $(\theta_0 = 2)$.

Критерий Ансари-Бредли (Ansari-Bradley test). Непараметрические аналоги критериев проверки однородности дисперсий предназначены для проверки гипотез о принадлежности двух выборок с объемами n_1 и n_2 общей генеральной совокупности с одинаковыми характеристиками рассеяния. При этом, как правило, предполагается равенство средних.

Пусть $n_1 < n_2$. Статистика критерия Ансари-Бредли [2] может быть вычислена следующим образом

$$S = \sum_{i=1}^{n_1} \left\{ \frac{n_1 + n_2 + 1}{2} - \left| R_i - \frac{n_1 + n_2 + 1}{2} \right| \right\}, \tag{4}$$

где R_i - ранги первой (меньшей по объему) выборки в общем вариационном ряду. Проверяемая гипотеза не отклоняется при $S_{\alpha/2} < S < S_{1-\alpha/2}$. Критические значения статистики для $n_1, n_2 \le 10$ доступны в таблице, приведенной, например, в [9], а для больших значений n_i соответствующая таблица может быть элементарно расширена методами статистического моделирования.

В случае принадлежности выборок случайных величин одному и тому же закону распределение $G(S | H_0)$ статистики (4) при справедливости проверяемой гипотезы H_0 не зависит от вида этого закона.

Математическое ожидание и дисперсия статистики (4) имеют вид:

$$E[S] = \begin{cases} \dfrac{n_1(n_1 + n_2 + 2)}{4} & \textit{при четном } (n_1 + n_2), \\ \dfrac{n_1(n_1 + n_2 + 1)^2}{4(n_1 + n_2)} & \textit{при нечетном } (n_1 + n_2); \end{cases}$$

$$D[S] = \begin{cases} \dfrac{n_1 n_2 (n_1 + n_2 - 2)(n_1 + n_2 + 2)}{48(n_1 + n_2 - 1)} & \textit{при четном} \ \ (n_1 + n_2), \\ \dfrac{n_1 n_2 (n_1 + n_2 + 1)\left[(n_1 + n_2)^2 + 3\right]}{48(n_1 + n_2)^2} & \textit{при нечетном} \ \ (n_1 + n_2). \end{cases}$$

При объемах выборок $n_1, n_1 > 10$ дискретное распределение нормированной статистики

$$S^* = (S - E[S]) / \sqrt{D[S]} \tag{5}$$

достаточно хорошо приближается стандартным нормальным законом. В этом случае проверяемая гипотеза не отклоняется при $N^*_{\alpha/2} < S^* < N^*_{1-\alpha/2}$, где N^*_α – соответствующая квантиль стандартного нормального закона. Дискретностью распределений статистик (4) и (5) практически можно пренебречь, начиная с $n_1, n_2 > 40$.

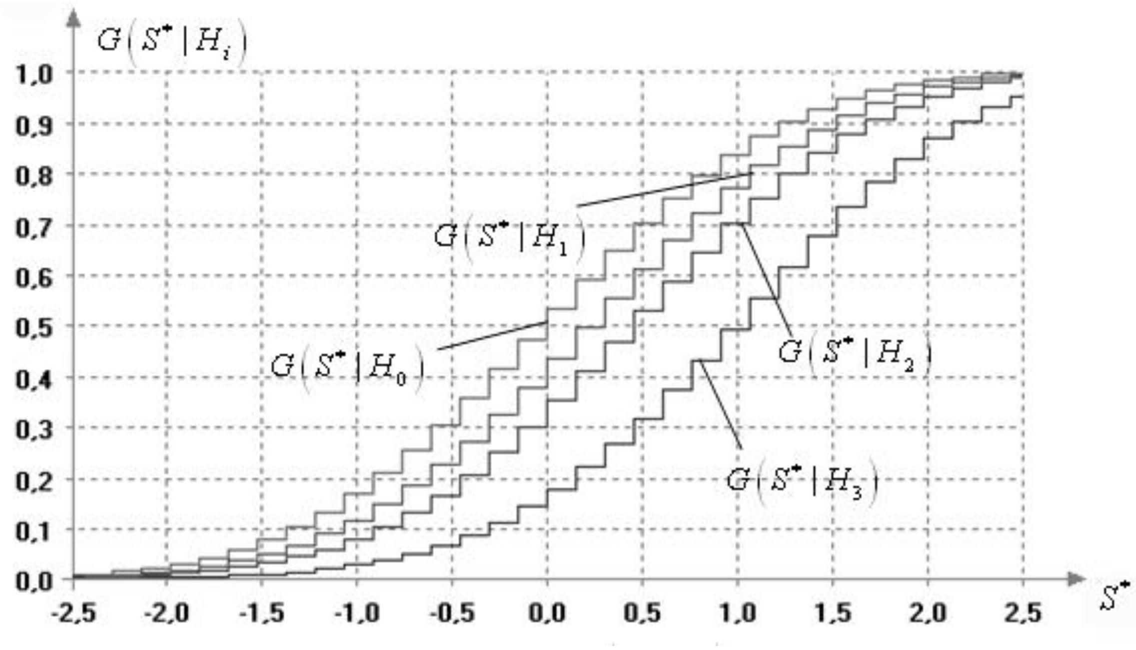

Рис. 1. Распределения $G(S^* \mid H_i)$ нормированной статистики критерия Ансари-Бредли (выборки объемом $n_1 = n_2 = 10$ принадлежат $De(3)$)

На рисунке 1 приведены графики распределения нормированной статистики S^* критерия Ансари-Бредли при справедливости различных конкурирующих гипотез H_0, H_1, H_2, H_3 при объемах выборок

$n_1 = n_2 = 10$, принадлежащих семейству распределений (3) с параметром формы $\theta_0 = 3$.

Критерий Муда (Mood's test). Статистика критерия имеет вид [3,10]

$$M = \sum_{i=1}^{n_1} \left(R_i - \frac{n_1 + n_2 + 1}{2} \right)^2, \tag{6}$$

где R_i - ранги меньшей по объему выборки в общем вариационном ряду двух выборок. Проверяемая гипотеза не отклоняется при $M_{\alpha/2} < M < M_{1-\alpha/2}$. Критические значения данной статистики для $n_1, n_2 \le 10$ доступны в [9], а для больших значений n_i таблица может быть легко расширена методами статистического моделирования.

При $n_1, n_2 > 10$ распределение нормированной статистики

$$M^* = \frac{\left(M - E[M] + \frac{1}{2} \right)}{\sqrt{D[M]}}, \tag{7}$$

где

$$E[M] = \frac{n_1(n_1 + n_2 + 1)(n_1 + n_2 - 1)}{12},$$

$$D[M] = \frac{n_1 n_2 \left(n_1 + n_2 + 1\right)(n_1 + n_2 + 2)(n_1 + n_2 - 2)}{180},$$

хорошо приближается стандартным нормальным законом [11], а при $n_1, n_2 > 20$, как показали исследования, дискретностью распределений статистик (6)-(7) вообще можно пренебречь. При использовании статистики (7) проверяемая гипотеза не отклоняется при $N^*_{\alpha/2} < M^* < N^*_{1-\alpha/2}$.

На рисунке 2 представлены распределения нормированной статистики (7) критерия Муда при справедливости различных конкурирующих гипотез H_0, H_1, H_2, H_3 при объемах выборок $n_1 = n_2 = 10$, принадлежащих семейству распределений (3) с параметром формы $\theta_0 = 3$. Как можно заметить, проблемы с дискретностью этой статистики существенно проще, чем со статистикой критерия Ансари-Бредли (см. рис. 1).

Критерий Сижела-Тьюки (Siegel-Tukey test) [4]. Вариационный ряд, построенный по объединенной выборке, $x_1 \le x_2 \le ... \le x_n$, где $n = n_1 + n_2$, преобразуется в последовательность вида

$$x_1, x_n, x_{n-1}, x_2, x_3, x_{n-2}, x_{n-3}, x_4, x_5, \ldots,$$

т.е. оставшийся ряд “переворачивается” каждый раз после приписывания рангов паре крайних значений. В качестве статистики критерия используется сумма рангов меньшей по объему выборки. При $n_1 < n_2$ статистика критерия имеет вид

$$R = \sum_{i=1}^{n_1} R_i . \tag{8}$$

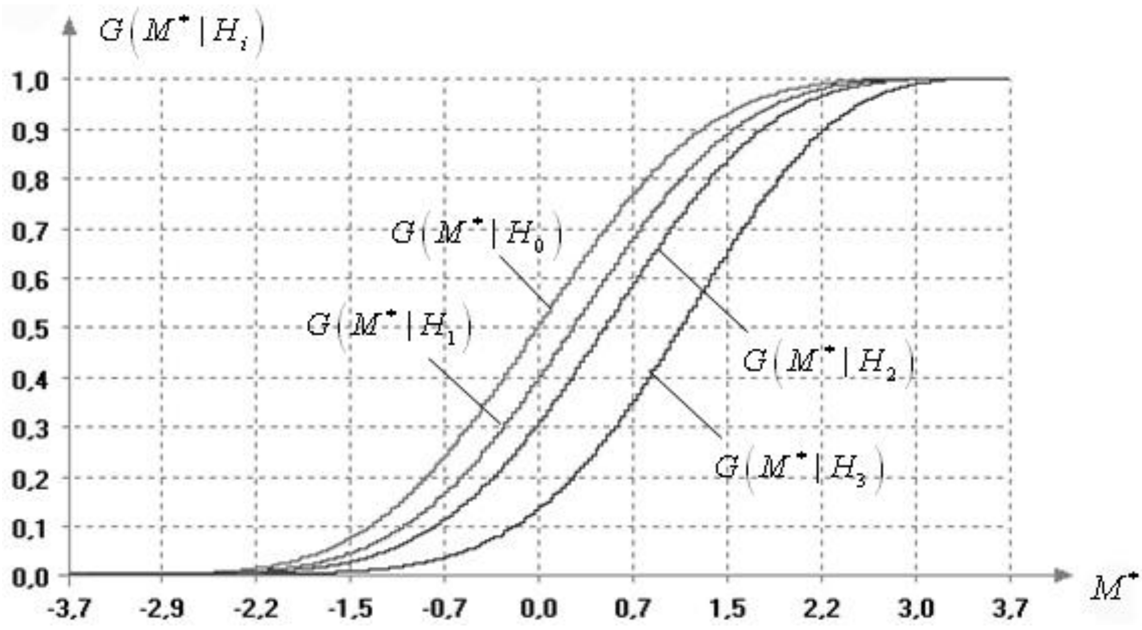

Рис. 2. Распределения $G(M^* | H_i)$ статистики нормированного критерия Муда (выборки объемом $n_1 = n_2 = 10$ принадлежат $De(3)$)

Проверяемая гипотеза не отклоняется при $R_{\alpha/2} < R < R_{1-\alpha/2}$. Статистика критерия Сижела-Тьюки является аналогом критерия Манна-Уитни, но предназначенным для проверки гипотез об однородности параметров масштаба. Поэтому при проверке гипотезы могут использоваться квантили распределения статистики Манна-Уитни [9]. При $n_1, n_2 > 10$ распределение нормированной статистики

$$R^* = (R - E[R]) / \sqrt{D[R]}, \tag{9}$$

где

$$E[R] = \frac{n_1(n_1 + n_2 + 1)}{2},$$

$$D[R] = \frac{n_1 n_2 \left(n_1 + n_2 + 1\right)}{12},$$

достаточно хорошо приближается стандартным нормальным законом. В этом случае проверяемая гипотеза не отклоняется при $N^*_{\alpha/2} < R^* < N^*_{1-\alpha/2}$. При этом дискретностью распределения статистики можно практически пренебречь с. $n_1, n_2 > 30$

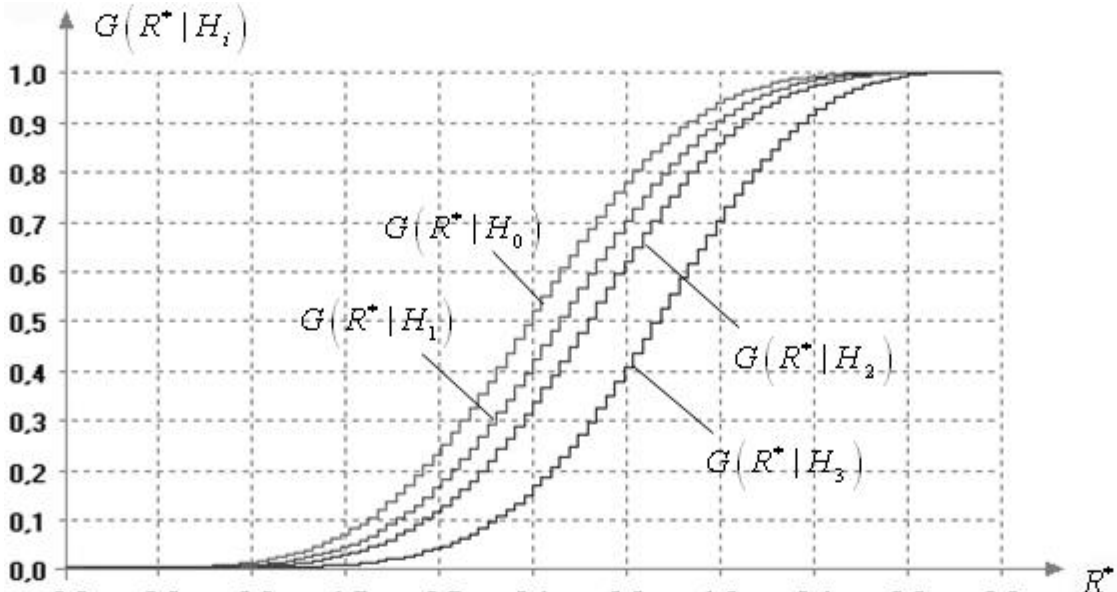

Рис. 3. Распределения $G(R^* | H_i)$ статистики нормированного критерия Сижела-Тьюки (выборки объемом $n_1 = n_2 = 10$ принадлежат $De(3)$)

На рисунке 3 приведены распределения нормированной статистики (9) критерия Сижела-Тьюки при справедливости различных конкурирующих гипотез H_0, H_1, H_2, H_3 при объемах выборок $n_1 = n_2 = 10$, принадлежащих семейству распределений (3) . Можно заметить, что данная статистика более дискретна по сравнению со статистикой Муда, но менее – по сравнению со статистикой критерия Ансари-Бредли (см. рис. 1 и 2).

Критерий Кейпена (Capon's test). Статистика критерия Кейпена [5] задается соотношением

$$K = \sum_{i=1}^{n_1} a_{n_1+n_2}(R_i), \tag{10}$$

где n_1 и n_2 – объемы сравниваемых выборок $(n_1 \le n_2)$, R_i – ранг i-го элемента меньшей по объему выборки в общем упорядоченном по возрастанию ряду (n_1+n_2) значений объединенной выборки, $a_{n_1+n_2}(i)$ - математическое ожидание квадрата i-й порядковой статистики в выборке объема (n_1+n_2) из стандартного нормального закона. Значения $a_N(i)$, называемые метками критерия, приведены, например, в [9].

Гипотеза о равенстве параметров масштаба не отклоняется с достоверностью α, если $K_{\alpha/2} < K < K_{1-\alpha/2}$, где критические значения $K_{\alpha/2}$, $K_{1-\alpha/2}$ также можно найти в [9].

При $n_1, n_2 > 10$ справедлива нормальная аппроксимация

$$K^* = \frac{K - E[K]}{\sqrt{D[K]}}, \tag{11}$$

где $E[K] = n_1$, $D[K] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} a_{n_1+n_2}^2(i) - \frac{n_1 n_2}{n_1+n_2-1}$. Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $N_{\alpha/2}^* < K^* < N_{1-\alpha/2}^*$.

В отличие от предыдущих критериев распределения статистик (10)-(11) данного критерия являются достаточно гладкими.

Критерий Клотца (Klotz's test). Статистика критерия имеет вид [6]

$$L = \sum_{i=1}^{n_1} u^2_{\frac{R_i}{n_1+n_2+1}}, \tag{12}$$

где n_1 и n_2 – объемы сравниваемых выборок ($n_1 \le n_2$), R_i – ранг i-го элемента меньшей по объему выборки в общем упорядоченном по возрастанию ряду (n_1+n_2) значений объединенной выборки, u_γ - γ-квантиль стандартного нормального распределения Значения $u^2_{\frac{i}{N+1}}$, называемые метками критерия, приведены, например, в [9].

Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $L_{\alpha/2} < L < L_{1-\alpha/2}$, где критические значения $L_{\alpha/2}$, $L_{1-\alpha/2}$ можно найти в [9].

При $n_1, n_2 > 10$ справедлива нормальная аппроксимация

$$L^* = \frac{L - E[L]}{\sqrt{D[L]}}, \tag{13}$$

где

$$E[L] = \frac{n_1}{n_1+n_2}\sum_{i=1}^{n_1+n_2} u^2_{\frac{i}{n_1+n_2+1}},$$

$$D[L] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)}\sum_{i=1}^{n_1+n_2} u^4_{\frac{i}{n_1+n_2+1}} - \frac{n_2}{n_1(n_1+n_2-1)}\left[\frac{n_1}{n_1+n_2}\sum_{i=1}^{m+n} u^2_{\frac{i}{n_1+n_2+1}}\right].$$

Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $N^*_{\alpha/2} < L^* < N^*_{1-\alpha/2}$. Распределения статистик (12)-(13) также достаточно гладкие.

Сравнительный анализ мощности. Ниже в таблицах 1-3 для уровней значимости $\alpha = 0{,}1;\ 0{,}05;\ 0{,}01$ приведены полученные в результате исследований значения мощности критериев Муда, Ансари-Бредли и Сижела-Тьюки относительно конкурирующих гипотез H_1: $\sigma_2 = 1{,}1\sigma_1$, H_2: $\sigma_2 = 1{,}2\sigma_1$, H_3: $\sigma_2 = 1{,}5\sigma_1$ (в случае принадлежности выборок нормальному закону).

Результаты анализа показывают заметное преимущество в мощности критерия Муда и практическую эквивалентность критериев Ансари-

Бредли и Сижела-Тьюки. Некоторый “разнобой” в данных таблиц при объемах выборок $n=10$ и $n=20$ объясняется различной степенью дискретности распределений статистик этих критериев.

Таблица 1. Мощность критериев относительно конкурирующей гипотезы $H_1:\ \sigma_2=1{,}1\sigma_1$

Критерий	α	Объемы выборок				
		n =10	n =20	n =40	n =60	n =100
Муда	0,1	0,111	0,120	0,143	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
Ансари-Бредли	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
Сижела-Тьюки	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

Таблица 2. Мощность критериев относительно конкурирующей гипотезы $H_2:\ \sigma_2=1{,}2\sigma_1$

Критерий	α	Объемы выборок				
		n =10	n =20	n =40	n =60	n =100
Муда	0,1	0,135	0,173	0,254	0,3300	0,468
	0,05	0,073	0,100	0,161	0,2225	0,344
	0,01	0,016	0,027	0,052	0,082	0,152
Ансари-Бредли	0,1	0,128	0,172	0,226	0,289	0,406
	0,05	0,070	0,097	0,141	0,189	0,287
	0,01	0,015	0,025	0,045	0,066	0,119
Сижела-Тьюки	0,1	0,124	0,165	0,226	0,289	0,405
	0,05	0,066	0,092	0,141	0,190	0,287
	0,01	0,014	0,013	0,044	0,066	0,119

Применение критериев Кейпена и Клотца для произвольных объемов выборок крайне затруднительно, в связи с необходимостью использования соответствующих “меток”, которые приводятся только для малых объемов выборок. Исследование мощности этих критериев при объемах выборок до 10 показало, что в этом случае они имеют мощность, которая не превосходит мощности критерия Муда. Поэтому дальнейшие исследования не имели смысла.

Таблица 3. Мощность критериев относительно конкурирующей гипотезы $H_3: \sigma_2 = 1,5\sigma_1$

Критерий	α	Объемы выборок				
		n =10	n =20	n =40	n =60	n =100
Муда	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
Ансари-Бредли	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
Сижела-Тьюки	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

Мощность критериев рассматривалась для различных законов распределения, отличных от нормального. Следует заметить, что, как и в случае параметрических критериев [1], мощность всех рассмотренных в данной части работы критериев относительно тех же конкурирующих гипотез повышается вместе с “облегчением” хвостов распределений (по сравнению с нормальным законом), которым подчиняются анализируемые выборки.

Естественно, что непараметрические критерии уступают по мощности рассмотренным в [1] критериям Бартлетта, Кокрена, Хартли и Фишера. На рисунке 4 приведены графики мощности критериев относительно конкурирующих гипотез H_1 и H_3 в зависимости от объема выборки n при $\alpha = 0,1$ в случае нормального закона. Как видим, преимущество в мощности критерия Бартлетта по сравнению с наиболее мощным из непараметрических критерием Муда достаточно существенное. Напомним, что в случае 2-х выборок мощности критериев Бартлетта, Кокрена, Хартли и Фишера совпадают [1]. Но в случае более 2-х выборок, когда применение непараметрических критериев не предусмотрено, различие в мощности критериев Бартлетта, Кокрена и Хартли ощутимо.

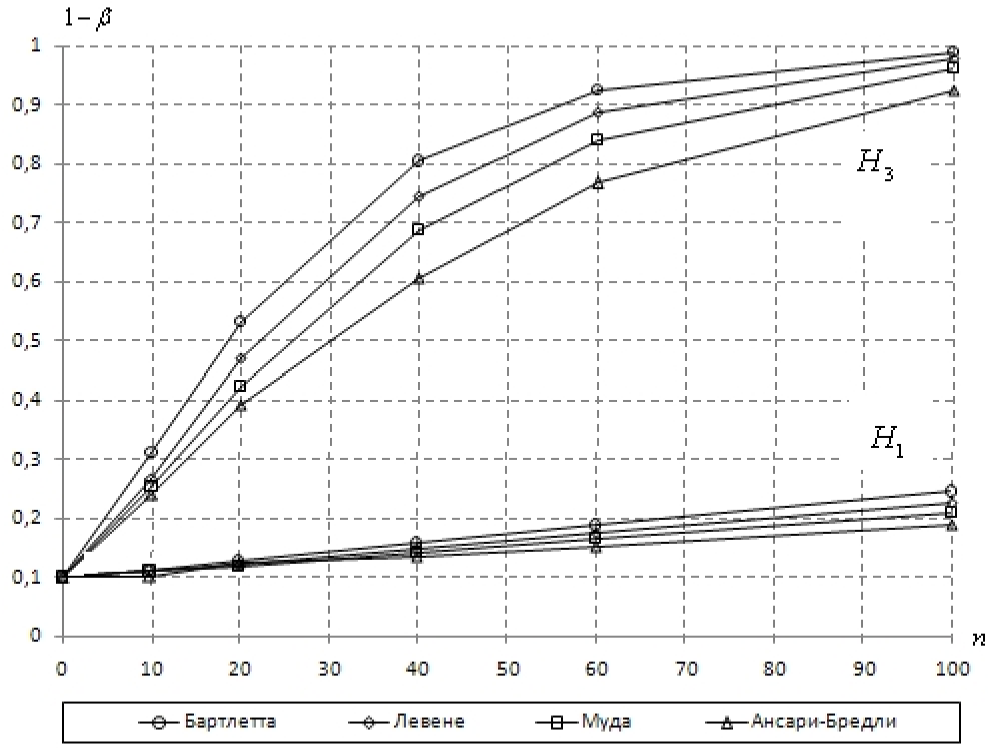

Рис. 4. Мощность критериев относительно конкурирующих гипотез H_1 и H_3 в зависимости от объема выборки n при $\alpha = 0.1$ в случае нормального закона

На рисунке 5 представлены графики мощности многовыборочных критериев относительно тех же конкурирующих гипотез H_1 и H_3 в зависимости от объема выборки n при $\alpha = 0{,}1$ в случае нормального закона для числа анализируемых выборок $m = 5$.

В процессе дискуссий, связанных с обсуждением достоинств и недостатков параметрических и непараметрических критериев, можно услышать радикальные мнения, суть которых заключается в рекомендации применения только непараметрических критериев. При этом основной и вполне справедливый довод связывают с тем, что на практике закон распределения, которому подчиняются анализируемые выборки, не известен и, как правило, отличается от нормального. Это действительно так. Если обе выборки принадлежат одной и той же генеральной совокупности, распределения $G(S|H_0)$ статистик непараметрических критериев не зависят от вида закона.

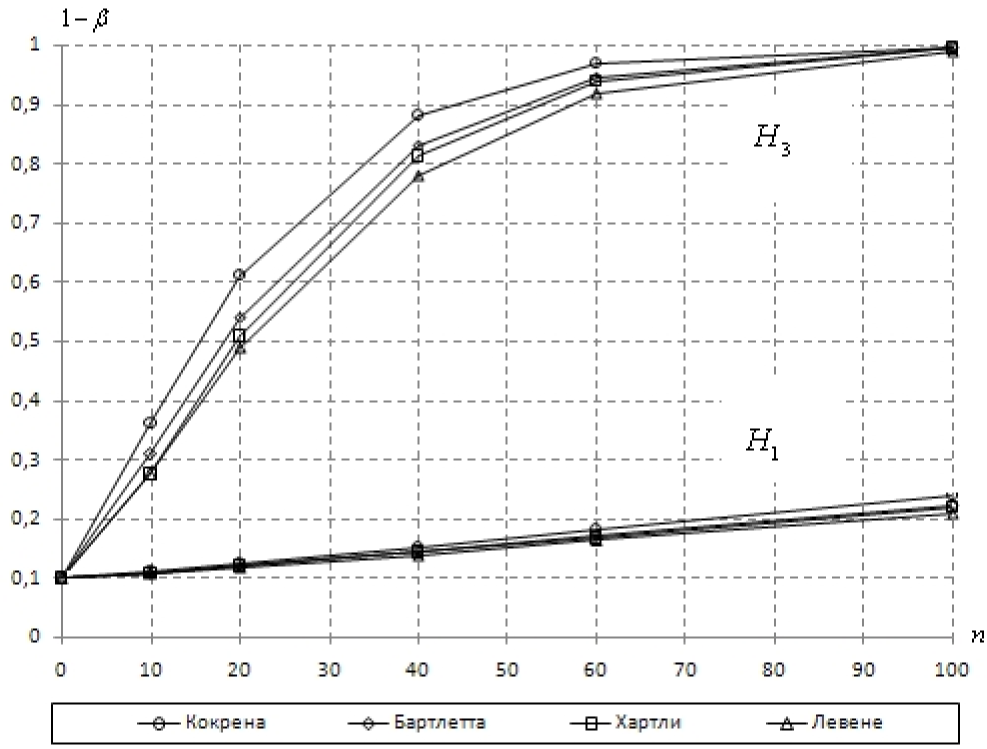

Рис. 5. Мощность многовыборочных критериев относительно конкурирующих гипотез H_1 и H_3 в зависимости от объема выборки n при $\alpha = 0.1$ в случае нормального закона при числе анализируемых выборок $m = 5$

Но, если при справедливости гипотезы H_0 о равенстве дисперсий выборки принадлежат разным законам, мы *наблюдаем зависимость* $G(S|H_0)$ *от этих законов*. На рисунке 6, в качестве примера, подтверждающего этот факт, показаны распределения нормированной статистики (7) критерия Муда, когда две выборки одинакового объема $n = 10$ подчиняются различным парам законов семейства (3) при равенстве дисперсий. Как видим, распределение зависит и от того, какая по порядку выборка, какому закону принадлежит.

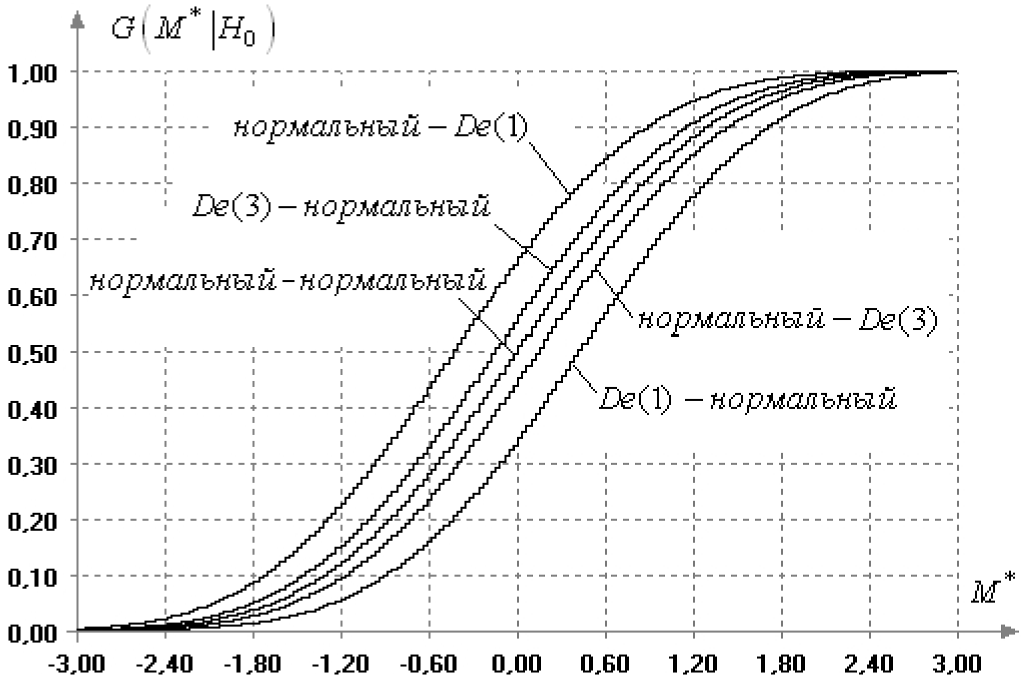

Рис. 6. Распределения нормированной статистики критерия Муда при справедливости H_0 в случае принадлежности пары выборок различным законам семейства (3)

Естественно, что для параметрических критериев отмеченная зависимость также характерна, но изменения в такой ситуации, например, распределений статистики критерия Кокрена, на наш взгляд, оказываются несколько меньшими.

Применение критерия Кокрена при законах, отличных от нормального. Сравнивая значения мощности критериев в таблицах 1-3 [1] для параметрических критериев со значениями в таблицах 1-3 настоящей работы для непараметрических, следует констатировать, что параметрические критерии имеют значительное преимущество в мощности. Причем это преимущество сохраняется и в ситуациях, когда анализируемые выборки принадлежат законам, существенно отличающимся от нормального. Ясно, что в этих ситуациях мы не можем использовать ставшие классическими результаты, связанные с распределениями (или процентными точками) статистик критериев Бартлетта, Кокрена, Хартли, Левене.

Проблема осложняется тем, что в условиях нарушения классических предположений о нормальности, распределения статистик упомянутых критериев при справедливости проверяемой гипотезы зависят от законов распределения, которым подчиняются анализируемые выборки, и от объемов выборок. В принципе, то же самое мы имеем, например, для критериев Кокрена, Хартли, Левене при нормальном законе. В этой связи понятно, что построить (найти) модели распределения статистики критерия для любых законов, для любых объемов выборок – задача нереальная. Однако для конкретных параметрических моделей законов распределения, зарекомендовавших себя в различных приложениях в качестве хороших моделей наблюдаемых случайных величин, такая задача (как и для нормального закона) может быть (относительно) легко решена с использованием компьютерных технологий, как, например в [12,13].

Результаты исследований в [1] показали предпочтительность критерия Кокрена, который в 2-х выборочном варианте не уступает любому другому, а в многовыборочном оказывается наиболее мощным (за исключением законов с тяжелыми хвостами, где хорошо в этом плане зарекомендовал себя критерий Левене).

В случае принадлежности наблюдаемых величин распределениям семейства (3) при значениях параметра формы $\theta_0 = 1, 2, 3, 4, 5$ и ряда значений n на основании результатов статистического моделирования построена таблица верхних процентных точек (1%, 5%, 10%) статистики критерия Кокрена (см. статистику (5) в [1]) для 2-х выборок. Полученные результаты представлены в таблицах 4-7 и могут использоваться в случае, когда есть основания считать, что распределение (3) с соответствующим параметром θ_0 представляет собой хорошую модель для наблюдаемых случайных величин. Построенные процентные точки уточняют некоторые результаты, представленные в [14], и расширяют возможности применения критерия Кокрена.

Таблица 4 - Верхние процентные точки для статистики критерия Кокрена в случае 2-х выборок равного объема n

n	$De(1)$			$De(2)$			$De(3)$			$De(4)$			$De(5)$		
	α			α			α			α			α		
	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01
5	0,917	0,947	0,980	0,865	0,906	0,959	0,845	0,890	0,950	0,836	0,883	0,947	0,831	0,879	0,945
8	0,862	0,900	0,949	0,791	0,833	0,899	0,764	0,807	0,877	0,751	0,794	0,866	0,744	0,787	0,861
10	0,836	0,875	0,930	0,761	0,801	0,868	0,733	0,773	0,842	0,720	0,759	0,829	0,713	0,751	0,822
15	0,789	0,829	0,890	0,713	0,748	0,811	0,686	0,719	0,780	0,674	0,706	0,765	0,667	0,698	0,757
20	0,759	0,797	0,858	0,684	0,716	0,774	0,660	0,689	0,743	0,648	0,676	0,728	0,642	0,669	0,720
25	0,736	0,772	0,834	0,665	0,694	0,748	0,642	0,668	0,717	0,632	0,656	0,703	0,626	0,649	0,695
30	0,718	0,753	0,814	0,650	0,677	0,727	0,629	0,653	0,699	0,619	0,642	0,685	0,614	0,635	0,677
40	0,693	0,725	0,782	0,630	0,654	0,699	0,611	0,632	0,672	0,603	0,622	0,660	0,598	0,616	0,653
50	0,674	0,704	0,758	0,617	0,638	0,679	0,599	0,618	0,654	0,591	0,609	0,642	0,587	0,604	0,636
60	0,660	0,689	0,740	0,606	0,626	0,664	0,591	0,608	0,640	0,583	0,599	0,630	0,579	0,594	0,624
70	0,649	0,676	0,724	0,598	0,617	0,652	0,584	0,599	0,630	0,577	0,591	0,620	0,573	0,587	0,614
80	0,640	0,665	0,712	0,592	0,609	0,642	0,578	0,593	0,621	0,572	0,585	0,612	0,568	0,581	0,607
90	0,632	0,657	0,701	0,587	0,603	0,634	0,573	0,587	0,614	0,567	0,580	0,605	0,564	0,576	0,600
100	0,626	0,649	0,692	0,582	0,598	0,628	0,570	0,583	0,609	0,564	0,576	0,600	0,561	0,572	0,595

Таблица 5 - Верхние процентные точки для статистики критерия Кокрена в случае 3-х выборок равного объема n

n	$De(1)$			$De(2)$			$De(3)$			$De(4)$			$De(5)$		
	α			α			α			α			α		
	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01
5	0,794	0,847	0,918	0,700	0,752	0,839	0,665	0,717	0,806	0,649	0,700	0,790	0,641	0,690	0,781
8	0,716	0,768	0,852	0,614	0,658	0,741	0,579	0,620	0,698	0,563	0,602	0,677	0,554	0,591	0,665
10	0,681	0,732	0,817	0,581	0,622	0,698	0,548	0,584	0,654	0,533	0,567	0,634	0,524	0,557	0,622
15	0,623	0,669	0,751	0,531	0,564	0,628	0,503	0,531	0,588	0,489	0,516	0,569	0,482	0,508	0,558
20	0,587	0,629	0,707	0,502	0,531	0,588	0,477	0,501	0,550	0,466	0,488	0,533	0,459	0,480	0,524
25	0,562	0,600	0,673	0,484	0,509	0,560	0,461	0,482	0,526	0,450	0,470	0,510	0,444	0,463	0,501
30	0,543	0,578	0,647	0,470	0,493	0,539	0,449	0,468	0,507	0,439	0,457	0,493	0,434	0,451	0,485
40	0,515	0,547	0,608	0,450	0,470	0,510	0,432	0,449	0,482	0,424	0,439	0,470	0,419	0,434	0,463
50	0,496	0,525	0,581	0,437	0,455	0,490	0,421	0,436	0,465	0,414	0,427	0,454	0,410	0,422	0,448
60	0,482	0,508	0,560	0,428	0,444	0,476	0,413	0,426	0,453	0,406	0,418	0,443	0,402	0,414	0,437
70	0,471	0,495	0,543	0,421	0,435	0,465	0,407	0,419	0,444	0,401	0,412	0,434	0,397	0,408	0,429
80	0,462	0,485	0,530	0,415	0,429	0,456	0,402	0,413	0,436	0,396	0,406	0,427	0,393	0,403	0,422
90	0,455	0,476	0,518	0,410	0,423	0,449	0,398	0,408	0,430	0,392	0,402	0,422	0,389	0,398	0,417
100	0,449	0,469	0,509	0,406	0,418	0,443	0,394	0,405	0,425	0,389	0,398	0,417	0,386	0,395	0,413

Таблица 6 - Верхние процентные точки для статистики критерия Кокрена в случае 4-х выборок равного объема n

n	$De(1)$			$De(2)$			$De(3)$			$De(4)$			$De(5)$		
	α			α			α			α			α		
	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01
5	0,696	0,755	0,848	0,584	0,634	0,727	0,545	0,591	0,679	0,527	0,571	0,656	0,517	0,560	0,643
8	0,611	0,666	0,761	0,501	0,541	0,619	0,466	0,500	0,569	0,450	0,482	0,546	0,441	0,471	0,533
10	0,575	0,626	0,720	0,470	0,506	0,576	0,438	0,468	0,529	0,423	0,451	0,507	0,415	0,441	0,495
15	0,517	0,561	0,646	0,424	0,453	0,510	0,397	0,421	0,468	0,385	0,406	0,450	0,378	0,398	0,439
20	0,482	0,521	0,598	0,399	0,422	0,471	0,375	0,395	0,435	0,364	0,382	0,419	0,358	0,375	0,410
25	0,457	0,493	0,563	0,382	0,403	0,445	0,360	0,378	0,413	0,351	0,366	0,398	0,346	0,360	0,390

30	0,439	0,471	0,536	0,369	0,388	0,427	0,350	0,365	0,397	0,341	0,355	0,384	0,336	0,349	0,377
40	0,413	0,441	0,498	0,352	0,368	0,401	0,335	0,348	0,348	0,328	0,340	0,364	0,324	0,335	0,358
50	0,395	0,420	0,470	0,340	0,355	0,384	0,326	0,337	0,361	0,319	0,329	0,351	0,315	0,325	0,345
60	0,382	0,404	0,451	0,332	0,345	0,371	0,319	0,329	0,350	0,313	0,322	0,341	0,309	0,318	0,336
70	0,372	0,392	0,435	0,326	0,337	0,361	0,313	0,323	0,342	0,308	0,316	0,334	0,305	0,313	0,329
80	0,364	0,383	0,422	0,320	0,331	0,354	0,309	0,318	0,336	0,304	0,312	0,328	0,301	0,309	0,324
90	0,357	0,375	0,412	0,316	0,326	0,348	0,305	0,314	0,331	0,300	0,308	0,324	0,298	0,305	0,320
100	0,352	0,368	0,403	0,313	0,322	0,342	0,302	0,310	0,327	0,298	0,305	0,320	0,295	0,302	0,316

Таблица 7 - Верхние процентные точки для статистики критерия Кокрена в случае 5-ти выборок равного объема n

n	$De(1)$			$De(2)$			$De(3)$			$De(4)$			$De(5)$		
	α			α			α			α			α		
	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01	0,1	0,05	0,01
5	0,623	0,684	0,787	0,504	0,551	0,642	0,464	0,505	0,588	0,446	0,484	0,562	0,436	0,472	0,548
8	0,537	0,591	0,690	0,426	0,461	0,533	0,392	0,421	0,482	0,376	0,403	0,458	0,367	0,393	0,446
10	0,501	0,550	0,645	0,397	0,428	0,491	0,366	0,392	0,444	0,352	0,375	0,422	0,344	0,366	0,411
15	0,445	0,485	0,567	0,355	0,379	0,429	0,330	0,349	0,390	0,318	0,336	0,372	0,312	0,329	0,363
20	0,412	0,447	0,520	0,332	0,352	0,394	0,310	0,326	0,360	0,300	0,315	0,345	0,295	0,308	0,337
25	0,388	0,420	0,485	0,316	0,334	0,371	0,297	0,311	0,341	0,288	0,301	0,328	0,283	0,295	0,320
30	0,370	0,399	0,459	0,305	0,321	0,354	0,287	0,300	0,327	0,279	0,291	0,315	0,275	0,286	0,308
40	0,347	0,371	0,371	0,290	0,303	0,331	0,275	0,285	0,308	0,268	0,278	0,298	0,264	0,273	0,292
50	0,330	0,352	0,397	0,280	0,291	0,316	0,266	0,276	0,295	0,260	0,269	0,286	0,257	0,265	0,281
60	0,,318	0,337	0,378	0,272	0,283	0,304	0,220	0,227	0,242	0,254	0,262	0,278	0,252	0,259	0,274
70	0,309	0,326	0,363	0,266	0,276	0,296	0,255	0,263	0,279	0,250	0,257	0,272	0,247	0,254	0,268
80	0,301	0,318	0,352	0,262	0,271	0,289	0,251	0,259	0,274	0,247	0,253	0,267	0,244	0,250	0,263
90	0,295	0,310	0,342	0,258	0,266	0,284	0,248	0,255	0,269	0,244	0,250	0,263	0,242	0,247	0,259
100	0,290	0,304	0,334	0,255	0,263	0,279	0,246	0,252	0,265	0,242	0,247	0,259	0,239	0,245	0,256

Выводы. Таким образом, подводя окончательные итоги по результатам исследования критериев проверки гипотез об однородности дисперсий, следует обратить внимание на следующие основные факты.

В случае анализа 2-х выборок критерии Фишера, Бартлетта, Кокрена, Хартли обладают одинаковой мощностью относительно тех же конкурирующих гипотез. При анализе более 2-х выборок преимущество в мощности имеет критерий Кокрена.

Среди рассмотренных непараметрических критериев наибольшей мощностью обладает критерий Муда, который заметно уступает критериям Фишера, Бартлетта, Кокрена, Хартли и Левене.

Действие параметрических критериев при необходимости можно распространить на ситуации, когда выборки описываются законами, отличающимися от нормального, воспользовавшись, как в нашем случае, методикой компьютерного моделирования для исследования распределений статистик и построения для этих распределений моделей или таблиц процентных точек. Из рассмотренных критериев в наибольшей степени на эту роль подходит критерий Кокрена. Однако необходимо учитывать, что распределение статистики критерия будет зависеть от вида закона, объема выборок и, во многих случаях, от конкретных значений некоторых параметров законов. При использовании соответствующего программного обеспечения [7] решение таких задач не вызывает принципиальных трудностей, а наличие программного обеспечения позволяет решать эти задачи по мере возникновения потребности [15]. Не лишне заметить, что специфика задач по исследованию методами компьютерного моделирования вероятностных закономерностей позволяет эффективно распараллеливать вычислительные процессы, используя ресурсы многоядерных и многопроцессорных компьютеров и компьютерных сетей, и получать искомое решение практически в реальном масштабе времени.

Настоящие исследования выполнены при частичной поддержке Российского фонда фундаментальных исследований (проект № 09-01-00056-а), Федерального агентства по образованию в рамках Аналитической ведомственной целевой программы "Развитие научного потенциала высшей школы" (проект № 2.1.2/3970) и федеральной целевой программы Минобрнауки РФ “Научные и научно-педагогические кадры инновационной России” в рамках мероприятия 1.2.1 (проект № НК-15П/15).

Литература

1. Лемешко Б.Ю., Лемешко С.Б., Горбунова А.А. О применении и мощности критериев проверки однородности дисперсий. Ч. I // Измерительная техника. 2010. № 3. – С.10-16.
2. Ansari A.R., Bradley R.A. Rank-tests for dispersions // AMS.1960. V.31. №4. – P.1174-1189.
3. Mood A. On the asymptotic efficiency of certain nonparametric tests // AMS. – 1954. – V.25. – P. 514-522.
4. Siegel S., Tukey J.W. A nonparametric sum of rank procedure for relative spread in unpaired samples // JASA. – 1960. – V.55, №291. – P. 429-445.
5. Capon J. Asymptotic efficiency of certain locally most powerful rank tests // AMS. – 1961. – V.32, №1. – P. 88-100.
6. Klotz J. Nonparametric tests for scale // AMS. – 1962. – V.33. – P. 498-512.
7. Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2004. – 120 с.
8. Лемешко Б.Ю. Статистический анализ одномерных наблюдений случайных величин: Программная система. - Новосибирск: Изд-во НГТУ, 1995. - 125 с.
9. Кобзарь, А.И. Прикладная математическая статистика: Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006. – 816 с.
10. Sukhatme B.V. On certain Two-sample nonparametric tests for variances //AMS. 1957. V.28. №1. – P.188-194.
11. Laubsher N.F., Steffens F.E., De Lange E.M. Exact critical Values for Mood`s distribution-free test statistic for dispersion and its normal approximation // Technometrics. 1968. V.10. №3. – P.497-508.
12. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с

использованием оценок максимального правдоподобия. Ч. I // Измерительная техника. 2009. № 6. – С.6-11.

13. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. II // Измерительная техника. 2009. № 8. – С.17-26.

14. Лемешко Б.Ю., Миркин Е.П. Критерии Бартлетта и Кокрена в измерительных задачах при вероятностных законах, отличающихся от нормального // Измерительная техника. – 2004. – №10. – С. 10-16. [Lemeshko B.Yu., Mirkin E.P. Bartlett and Cochran tests in measurements with probability laws different from normal // Measurement Techniques. 2004. Vol. 47, № 10. – P. 960-968]

15. Лемешко Б.Ю., Постовалов С.Н., Лемешко С.Б. Компьютерные технологии анализа данных и исследования статистических закономерностей: Методические указания к выполнению лабораторных работ. – Новосибирск: Изд-во НГТУ, 2007. – 71 с.