Измерительная техника. – 2010. – №3. – С. 10-16

УДК 519.24

О применении и мощности критериев проверки однородности дисперсий. Ч. I . Параметрические критерии

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

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

Проведен сравнительный анализ мощности классических критериев однородности дисперсий (критериев Фишера, Бартлетта, Кокрена, Хартли, Левене). Исследованы распределения статистик критериев при нарушении предположений о принадлежности выборок нормальному закону.

Ключевые слова: критерий однородности дисперсий, критерий Фишера, критерий Бартлетта, критерий Кокрена, критерий Хартли, критерий Левене, мощность критерия.

Введение. Критерии проверки гипотез об однородности выборок представляют собой очень часто используемые в различных приложениях задачи статистического анализа. При этом речь может идти о проверке гипотез об однородности законов распределения, соответствующих анализируемым выборкам, или об однородности математических ожиданий, или об однородности дисперсий. Естественно, что наиболее полные выводы могут быть получены в первом случае, однако исследователя могут в большей степени интересовать вопросы о возможных отклонениях в средних значениях выборок или о различии в характеристиках рассеяния результатов измерений.

Особенности применения непараметрических критериев однородности Смирнова и Лемана-Розенблатта и анализ их мощность были рассмотрены в работе [1]. В [2] показана устойчивость классических критериев проверки гипотез об однородности средних к нарушениям предположения о принадлежности анализируемых выборок случайных величин нормальному закону и приведен сравнительный анализ мощности различных критериев, в том числе, непараметрических.

Применение классических критериев проверки однородности дисперсий всегда сопряжено с вопросом, насколько полученные выводы корректны в данной конкретной ситуации? Дело в том, что одним из основных предположений при построении этих критериев является принадлежность наблюдаемых случайных величин (погрешностей измерений) нормальному закону распределения. При этом давно известно, что параметрические критерии однородности дисперсий чрезвычайно чувствительны к малейшим отклонениям наблюдаемых случайных величин от нормального закона. При нарушении данного предположения условные распределения статистик критериев при справедливости проверяемой гипотезы, как правило, сильно изменяются. Так как погрешности измерительных приборов или наблюдаемые в различных приложениях величины далеко не всегда подчиняются нормальному закону, то применение классических результатов в таких условиях может приводить к неверным выводам.

В связи с этим вызывает практический интерес, каким же будет поведение критериев проверки однородности дисперсий (характеристик рассеяния) при определенных отклонениях закона распределения результатов измерений (контролируемого показателя) от нормального и будет ли корректным применение классического аппарата для проверки гипотез? Пока нет четкого ответа, как соотносится мощность различных параметрических критериев по отношению к конкретным конкурирующим гипотезам, насколько уступают им по мощности непараметрические критерии проверки гипотез о равенстве характеристик рассеяния (параметров масштаба)?

Данная работа продолжает исследования устойчивости критериев проверки гипотез о равенстве дисперсий [3, 4]. Сравниваются классические критерии Бартлетта [5], Кокрена [6], Фишера, Хартли [7], Левене [8], рассматриваются непараметрические (ранговые) критерии Ансари-Бредли [9], Муда [10], Сижела-Тьюки [11], Кейпена [12] и Клотца [13].

Цель настоящей работы заключалась: во-первых, в исследовании распределений статистик перечисленных критериев при законах распределения наблюдаемых случайных величин, отличных от нормального; во-вторых, в сравнительном анализе мощности критериев относительно конкретных конкурирующих гипотез; в-третьих, в реализации возможности применения классических критериев в условиях нарушения предположений о нормальности случайных величин; в-четвертых, в выработке рекомендаций по применению критериев в реальных условиях приложений.

Проверяемая гипотеза о постоянстве дисперсий m выборок имеет вид

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

а конкурирующая с ней гипотеза

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

где неравенство выполняется, по крайней мере, для одной пары индексов i_1, i_2.

При исследовании распределений статистик, построении для этих распределений процентных точек и оценке мощности критериев относительно различных конкурирующих гипотез использовалась методика статистического моделирования [14] и развиваемая на базе [15] программная система ISW (Интервальная статистика под Windows). При этом объем моделируемых выборок исследуемых статистик составлял величину $N = 10^6$. При таких N величина разности между истинным законом распределения статистики и смоделированным эмпирическим по модулю не превышает величины 10^{-3}.

Исследования распределений статистик проводились при различных наблюдаемых законах, в частности, в случае принадлежности моделируемых выборок семейству 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. Это семейство может быть хорошей моделью для законов распределения погрешностей различных измерительных систем. Распределение $De(\theta_0)$ включает в качестве частных случаев распределение Лапласа $(\theta_0 = 1)$ и нормальное $(\theta_0 = 2)$. Семейство (3) позволяет задавать различные симметричные законы распределения в той или иной мере отличающиеся от нормального: чем меньше значение параметра формы θ_0, тем «тяжелее» хвосты распределения $De(\theta_0)$, чем больше параметр, тем хвосты «легче».

При сравнительном анализе мощности критериев рассматривались следующие конкурирующие гипотезы: $H_1 : \sigma_m = 1.1\sigma_0$; $H_2 : \sigma_m = 1.2\sigma_0$; $H_3 : \sigma_m = 1.5\sigma_0$. То есть конкурирующей гипотезе соответствует ситуация, когда $m-1$ выборка принадлежат закону с некоторым $\sigma = \sigma_0$, в то время как одна из выборок, например, с номером m имеет некоторую отличную дисперсию. Проверяемой гипотезе соответствует ситуация $H_0:\ \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2 = \sigma_0^2$.

Критерий Бартлетта (Bartlett's test). Статистика критерия Бартлетта [16] вычисляется в соответствии с соотношением:

$$\chi^2 = M\left[1 + \frac{1}{3(m-1)}\left(\sum_{i=1}^{m}\frac{1}{\nu_i} - \frac{1}{N}\right)\right]^{-1}, \tag{4}$$

где

$$M = N\ln\left(\frac{1}{N}\sum_{i=1}^{m}\nu_i S_i^2\right) - \sum_{i=1}^{m}\nu_i \ln S_i^2,$$

m – количество выборок, n_i – объемы выборок, $\nu_i = n_i$, если математическое ожидание известно, и $\nu_i = n_i - 1$, если неизвестно, $N = \sum_{i=1}^{m} \nu_i$, S_i^2 – оценки выборочных дисперсий. При неизвестном математическом ожидании оценки $S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ji} - \overline{X}_i)^2$, где X_{ij} – j-е наблюдение в i-й выборке, $\overline{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ji}$.

Если гипотеза H_0 верна, все $\nu_i > 3$ и выборки извлекаются из нормальной генеральной совокупности, то статистика (4) приближенно подчиняется χ^2_{m-1}- распределению. Если вычисленное значение статистики $\chi^{2*} > \chi^2_{1-\alpha, m-1}$, то проверяемая гипотеза отклоняется на заданном уровне значимости α.

При нормально распределенных результатах измерений распределение статистики (4) практически не зависит от изменения объемов выборок [3]. То, что предельное распределение статистики критерия известно и его можно использовать при малых объемах выборок, являются серьезными преимуществами критерия Бартлетта.

При отклонении закона распределения наблюдаемого показателя от нормального закона распределение $G(\chi^2 | H_0)$ статистики (4) становится *зависящим от объема выборки* и отличным от χ^2_{m-1}. Этот факт отражает рисунок 1, на котором показаны распределения статистики при различных объемах 2-х сравниваемых выборок в случае принадлежности их закону распределения Лапласа (семейству (3) с параметром формы $\theta_0 = 1$), а также приведено предельное χ^2_1-распределение статистики (для классической ситуации с нормальным законом).

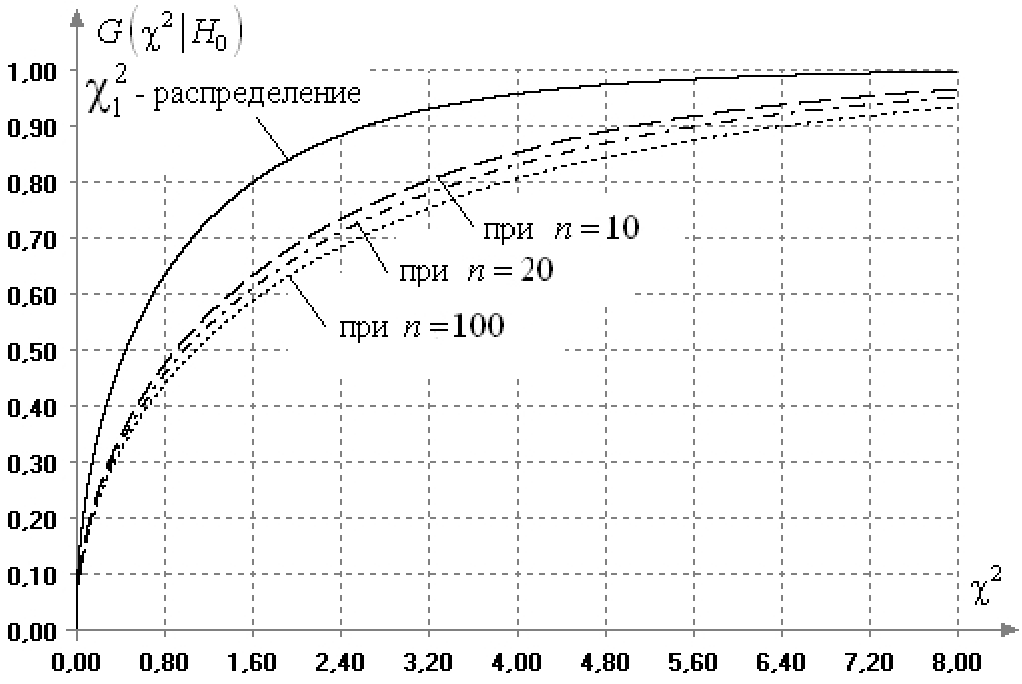

Рис. 1. Функции распределения статистики критерия Бартлетта при различных объемах выборок в случае распределения Лапласа при $m=2$

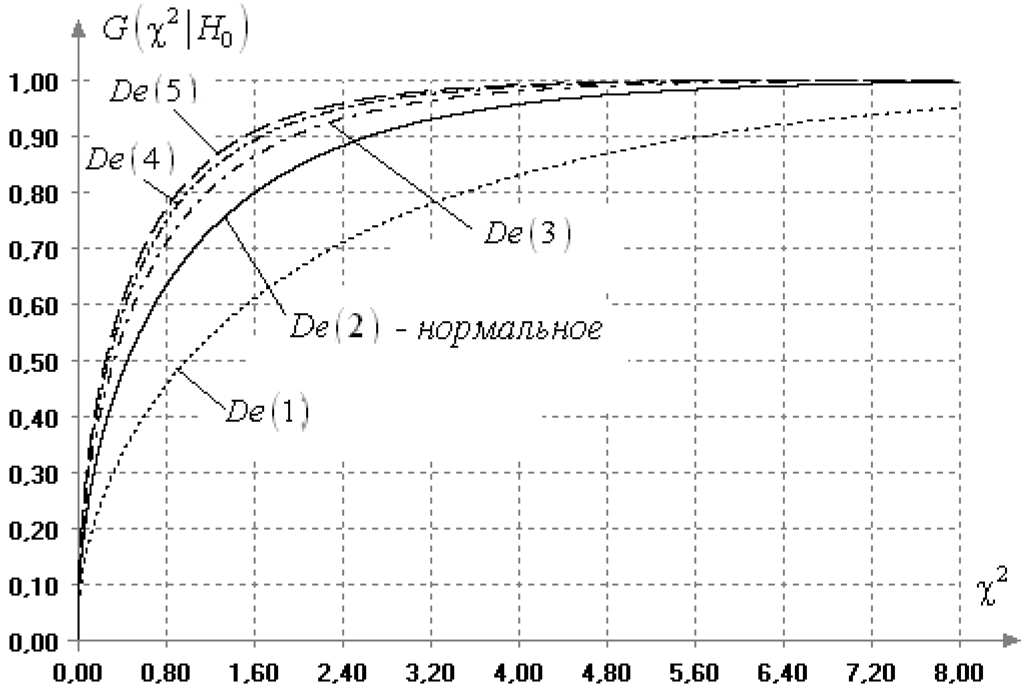

Рис. 2. Функции распределения статистики критерия Бартлетта в случае распределений семейства (3) с различными значениями параметра формы $(n=20,\ \ m=2)$

Распределения статистики (4) очень чувствительны к отклонениям наблюдаемого закона от нормального. На рисунке 2 показано, как меняется распределение статистики критерия Бартлетта, если результаты измерений подчиняются семейству распределений (3) с различными значениями параметра формы.

На рисунке 3 приведены графики распределения статистики (4) при справедливости конкурирующих гипотез $H_1 : \sigma_m = 1.1\sigma_0$, $H_2 : \sigma_m = 1.2\sigma_0$, $H_3 : \sigma_m = 1.5\sigma_0$ в ситуации, когда выборки принадлежат закону распределения семейства (3) с параметром формы $\theta_0 = 3$.

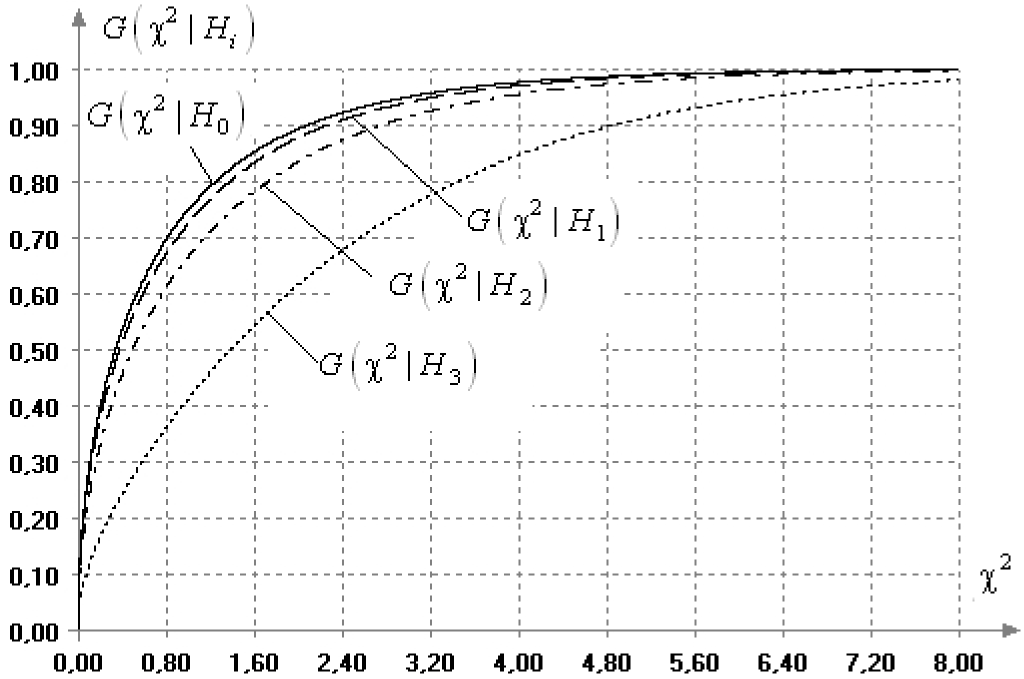

Рис. 3. Условные распределения $G(\chi^2 | H_i)$ статистики критерия Бартлетта, $m = 2$, $n = 10$, выборки принадлежат семейству (3) с параметром формы $\theta_0 = 3$.

Критерий Кокрена (Cochran's test). Когда все n_i одинаковы, $n_1 = n_2 = ... = n_m = n$, возможно использование более простого критерия Кокрена [6]. Статистика Q критерия Кокрена выражается формулой

$$Q = \frac{S_{\max}^2}{S_1^2 + S_2^2 + \cdots + S_m^2}, \tag{5}$$

где $S_{\max}^2 = \max(S_1^2, S_2^2, ..., S_m^2)$, m – число независимых оценок дисперсий (число выборок), S_i^2 – оценки выборочных дисперсий.

Распределения статистики Кокрена сильно зависят от объема наблюдаемых выборок. Распределения статистики неизвестны. Поэтому в справочной литературе для ограниченного числа значений n приводятся только таблицы процентных точек, которые и используются при проверке гипотез. Нулевая гипотеза отклоняется при больших значениях статистики.

Характер зависимости распределений $G(Q|H_0)$ статистики (5) критерия Кокрена при справедливости проверяемой гипотезы H_0 от закона распределения наблюдаемых выборок аналогичен зависимости для критерия Бартлетта [3].

Критерий Хартли (Hartley's test). Критерий Хартли [7], так же как и критерий Кокрена, используется в случае выборок равного объема.

Статистика критерия Хартли проверки гипотезы однородности дисперсий имеет вид

$$F = \frac{S_{\max}^2}{S_{\min}^2}, \tag{6}$$

где $S_{\max}^2 = \max(S_1^2, S_2^2, ..., S_m^2)$, $S_{\min}^2 = \min(S_1^2, S_2^2, ..., S_m^2)$, m – число независимых оценок дисперсий (число выборок), S_i^2 – оценки выборочных дисперсий.

Степенями свободы для распределения статистики являются числа $\nu_1 = m$ и $\nu_2 = n-1$, откуда понятно, что распределения статистики существенно зависят от объема выборок. Проверяемая гипотеза отклоняется при больших значениях статистики. В литературе для статистики (6) приводятся лишь таблицы процентных точек.

Характер зависимости распределений $G(F|H_0)$ статистики (6) критерия Хартли при справедливости проверяемой гипотезы H_0 от закона распределения наблюдаемых выборок аналогичен зависимостям для критериев Бартлетта и Кокрена.

Критерий Левене (Levene's test). Статистика критерия Левене [8] имеет вид:

$$W = \frac{N-m}{m-1} \frac{\sum_{i=1}^{m} n_i (\overline{Z}_{i\bullet} - \overline{Z}_{\bullet\bullet})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i} (Z_{ij} - \overline{Z}_{i\bullet})^2}, \tag{7}$$

где m – количество выборок, n_i – объем i-й выборки, $N = \sum_{i=1}^{m} n_i$, X_{ij} – j-е наблюдение в i-й выборке, $Z_{ij} = \left| X_{ij} - \overline{X}_{i\bullet} \right|$, в котором $\overline{X}_{i\bullet}$ – среднее в i-й выборке. $\overline{Z}_{i\bullet}$ – среднее Z_{ij} по i-й выборке, $\overline{Z}_{\bullet\bullet}$ – среднее Z_{ij} по всем выборкам.

В многочисленных описаниях критерия, например [17], говорится, что в случае принадлежности выборок нормальному закону и справедливости H_0 эта статистика подчинятся F_{ν_1,ν_2}-распределению Фишера с числом степеней свободы $\nu_1 = m-1$ и $\nu_2 = N-m$. На самом деле при объемах выборок в 10÷20 элементов *распределения статистики существенно отличаются от* F_{ν_1,ν_2}*-распределения*, и это надо иметь ввиду при использовании критерия. То, что распределение статистики (7) не является F_{ν_1,ν_2}- распределением Фишера очевидно из определения величин Z_{ij}, которые в любом случае не принадлежат нормальному закону, а, следовательно, (7) не может подчиняться F_{ν_1,ν_2}- распределению. В этой связи процентные точки распределения исследовались методами статистического моделирования [18]. В то же время, как показали наши исследования, уже при объемах выборок $n_i \geq 40$ максимальное отклонение распределения статистики от F_{ν_1,ν_2}- распределения Фишера не превышает величины 0.005 (в случае принадлежности анализируемых выборок нормальному закону).

Проверяемая гипотеза отклоняется при больших значениях статистики.

Поведение условного распределения $G(W \mid H_0)$ статистики критерия Левене при справедливости проверяемой гипотезы H_0 в зависимости от вида закона распределения, соответствующего наблюдаемым выборкам, иллюстрирует рисунок 4. Очевидно, что критерий Левене менее чувствителен к отклонениям анализируемых выборок от нормального закона (сравните рисунки 2 и 4).

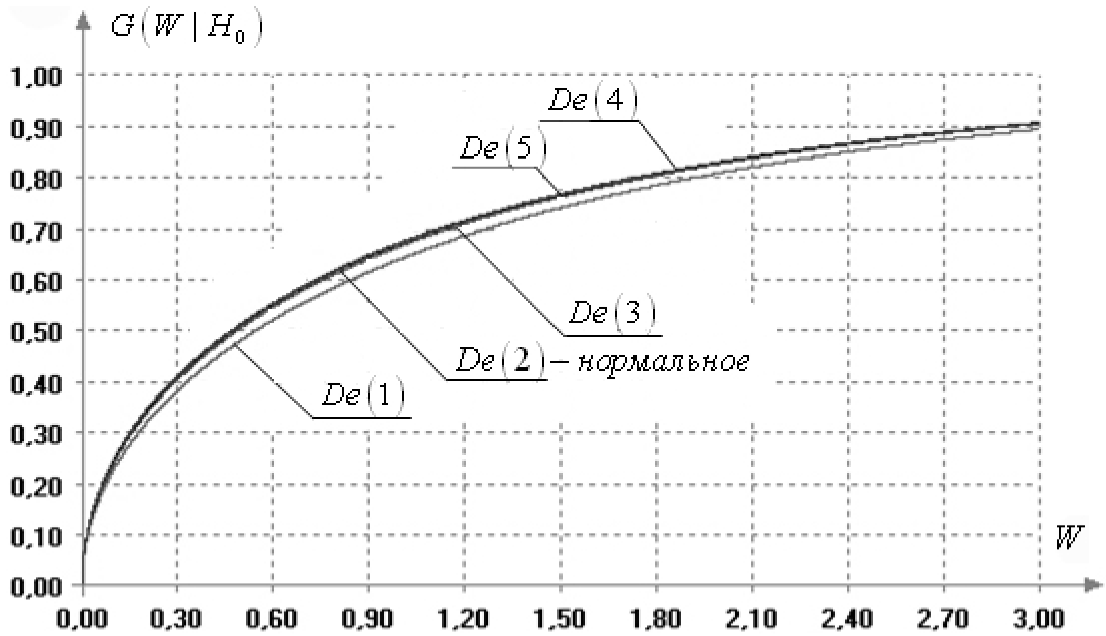

Рис. 4. Функции распределения статистики критерия Левене в случае распределений семейства (3) с различными значениями параметра формы $(n = 20, m = 2)$

В оригинальном критерии Левене предусмотрено использование только выборочных средних. В работе [19] предложено в статистике вида (7) в качестве оценок среднего использовать выборочные медиану и усеченное среднее ($Z_{ij} = \left| X_{ij} - \tilde{X}_{i\bullet} \right|$, где $\tilde{X}_{i\bullet}$ − медиана в i-й выборке; $Z_{ij} = \left| X_{ij} - \bar{X}'_{i\bullet} \right|$, где $\bar{X}'_{i\bullet}$ − усеченное среднее в i-й выборке). Считается, что в этих случаях критерий становится ещё устойчивей к нарушению предположений о нормальности.

Однако наши исследования этого явно не подтвердили: различия между распределениями $G(W|H_0)$ в случае принадлежности анализируемых выборок нормальному закону и в случае принадлежности закону Лапласа имеют тот же порядок как при использовании в статистике (7) обычного среднего значения, так и при использовании, например, медианы. При этом следует учитывать, что в случае использования медианы распределение статистики (7) при выполнении предположений о нормальности существенно отличается от распределения статистики при использовании обычного среднего и в большей степени отличается от F_{ν_1,ν_2}-распределения Фишера. То, что предшествующие исследователи [19] подчеркивали как повышенную устойчивость (при использовании медианы) критерия со статистикой (7) к нарушению предположений о нормальности по отношению к законам с более тяжелыми по сравнению с нормальным законом хвостами, на самом деле представляет собой изменение $G(W|H_0)$ в связи с *изменением метода оценивания* математического ожидания соответствующих выборок. Рисунок 5 показывает, как в случае $m=2$ выборок объемом $n=10$, принадлежащих нормальному закону, изменяется распределение статистики (7) в зависимости от того, используется ли оценка среднего или оценка медианы. Там же приведено соответствующее $F_{1,18}$-распределение Фишера. Картина, представленная на рисунке 5, подчеркивает, что при ограниченных объемах выборок, необходимо учитывать различие в распределениях статистики, связанное с различием используемых оценок.

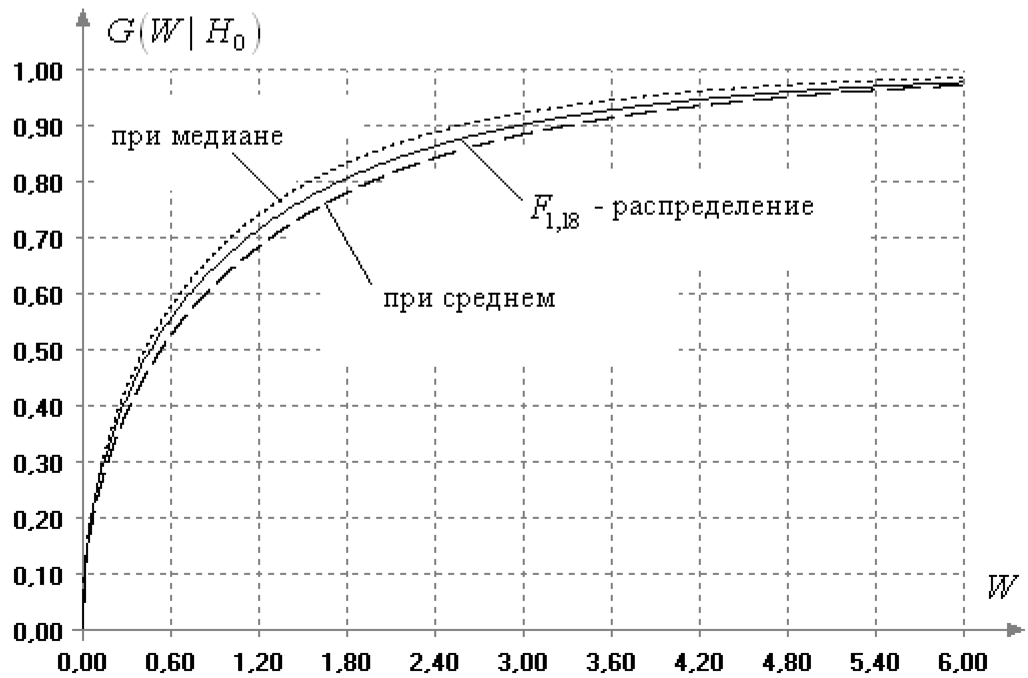

Рис. 5. Функции распределения статистики критерия Левене в случае принадлежности выборок нормальному закону при $n=10$, $m=2$ и $F_{1,18}$ - распределение Фишера

Критерий Фишера (Fisher's test). Критерий Фишера используется для проверки гипотезы о равенстве дисперсий *двух* выборок случайных величин. Статистика критерия имеет простой вид

$$F = \frac{s_1^2}{s_2^2}, \tag{8}$$

где s_1^2 и s_2^2 – несмещенные оценки дисперсий, вычисленные по выборкам.

В случае принадлежности выборок нормальному закону и справедливости $H_0 : \sigma_1^2 = \sigma_2^2$ эта статистика подчинятся F_{ν_1,ν_2} -распределению Фишера с числом степеней свободы $\nu_1 = n_1 - 1$ и $\nu_2 = n_2 - 1$, где n_1 и n_2 - объемы сравниваемых выборок. Проверяемая гипотеза отклоняется при малых $F^* < F_{\alpha/2,\nu_1,\nu_2}$ или больших $F^* > F_{1-\alpha/2,\nu_1,\nu_2}$ значениях статистики. Как и остальные, критерий Фишера очень чувствителен к отклонениям от нормальности.

Сравнительный анализ мощности. О преимуществах того или иного критерия при заданной вероятности ошибки 1-го рода α (отклонить верную гипотезу H_0) можно судить по величине мощности $1-\beta$, где β - вероятность ошибки 2-го рода (не отклонить гипотезу H_0 при справедливости конкурирующей гипотезы H_1). В [16] однозначно говорится о более низкой мощности критерия Кокрена по сравнению с мощностью критерия Бартлетта. В нашей работе [3] на примере проверки гипотез об однородности дисперсий пяти выборок было показано, что выше оказывается мощность критерия Кокрена.

Ниже в таблицах 1-3 для уровней значимости (для вероятностей ошибок 1-го рода) $\alpha = 0{,}1;\ 0{,}05;\ 0{,}01$ приведены значения мощности рассматриваемых в данной работе критериев однородности дисперсий для случая 2-х выборок ($H_0:\ \sigma_2 = \sigma_1$) 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). Полученные в результате исследований значения мощности критериев Бартлетта, Кокрена, Хартли и Фишера в этом случае оказались *одинаковыми*. При вычислении мощности по смоделированным распределениям статистик (при справедливости H_0 и H_i, $i = 1, 2, 3$) совпали все 3 значащие цифры, приведенные в таблицах. Следовательно, в этих условиях (при двух выборках, извлекаемых из нормальных совокупностей) данные четыре критерия являются эквивалентными. В то же время применение критерия Бартлетта оказывается более предпочтительным, так как известно его асимптотическое распределение, которое не зависит от объемов выборок, которые могут быть не равными. Критерий Левене заметно уступает по мощности критериям Бартлетта, Кокрена, Хартли и Фишера.

Таблица 1. Мощность критериев однородности дисперсий 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
Бартлетта, Кокрена, Хартли, Фишера	0,1	0,112	0,127	0,158	0,188	0,246
	0,05	0,058	0,068	0,090	0,112	0,157
	0,01	0,012	0,016	0,024	0,032	0,051
Левене	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

Таблица 2. Мощность критериев однородности дисперсий 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
Бартлетта, Кокрена, Хартли, Фишера	0,1	0,144	0,199	0,304	0,401	0,564
	0,05	0,079	0,119	0,201	0,283	0,438
	0,01	0,018	0,033	0,071	0,114	0,218
Левене	0,1	0,135	0,184	0,276	0,363	0,515
	0,05	0,072	0,107	0,177	0,250	0,388
	0,01	0,016	0,028	0,058	0,095	0,180

Таблица 3. Мощность критериев однородности дисперсий 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
Бартлетта, Кокрена, Хартли, Фишера	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
Левене	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

В случае законов распределения, отличных от нормального, например, в случае принадлежности двух анализируемых выборок семейству законов распределения (3) критерии Бартлетта, Кокрена, Хартли и Фишера *остаются эквивалентными* по мощности, а критерий Левене заметно им уступает. Однако в случае законов распределения с более тяжелыми хвостами (например, в случае принадлежности выборок распределению Лапласа) критерий Левене имеет преимущество в мощности.

Следует обратить внимание, при каких объемах выборок с заданными вероятностями ошибок 1-го и 2-го рода, например, $\alpha \le 0{,}1$ и $\beta \le 0{,}1$ критерии Бартлетта, Кокрена, Хартли и Фишера могут различать конкурирующие гипотезы H_0 и H_i? В случае относительно далекой конкурирующей гипотезы H_3: $\sigma_2 = 1{,}5\sigma_1$ при $n_1 = n_2$ объемы выборок n_i должны быть не менее 53, а в случае более близкой гипотезы H_1: $\sigma_2 = 1{,}1\sigma_1$ требуется уже порядка 950 наблюдений!

Критерии Бартлетта, Кокрена, Хартли и Левене могут применяться при числе выборок больше 2-х. И в таких ситуациях мощность этих критериев оказывается различной. В таблицах 4-6 приведены оценки мощности (в случае принадлежности выборок нормальному закону) при числе выборок 3 и 5, для случая, когда $(m-1)$-й выборкам соответствует значение σ_1, а последней – σ_m, отличное от σ_1.

Таблица 4. Мощность многовыборочных критериев однородности дисперсий относительно конкурирующей гипотезы H_1: $\sigma_m = 1{,}1\sigma_1$

Количество выборок $m=3$, $n=100$				
α	Кокрена	Бартлетта	Хартли	Левене
0,1	0,250	0,242	0,239	0,225
0,05	0,161	0,152	0,148	0,139
0,01	0,056	0,049	0,046	0,043
Количество выборок $m=5$, $n=100$				
0,1	0,241	0,224	0,219	0,209
0,05	0,156	0,138	0,133	0,127
0,01	0,056	0,044	0,040	0,039

Как видим, в многовыборочном варианте в случае выполнения предположений о нормальном законе данные критерии можно упорядочить по убыванию мощности следующим образом:

$$\text{Кокрена} \succ \text{Бартлетта} \succ \text{Хартли} \succ \text{Левене}.$$

Это порядок предпочтения сохраняется и в случае нарушений предположений о нормальном законе. Исключение касается ситуаций, когда выборки принадлежат законам с более тяжелыми, чем у нормального закона, хвостами, например, закону распределения Лапласа: в этом случае критерий Левене оказывается несколько мощнее трех других.

Таблица 5. Мощность многовыборочных критериев однородности дисперсий относительно конкурирующей гипотезы $H_2: \sigma_m = 1,2\sigma_1$

Количество выборок $m=3$, $n=100$				
α	Кокрена	Бартлетта	Хартли	Левене
0,1	0,609	0,577	0,568	0,530
0,05	0,494	0,459	0,443	0,409
0,01	0,286	0,237	0,217	0,200
Количество выборок $m=5$, $n=100$				
0,1	0,624	0,557	0,545	0,513
0,05	0,515	0,434	0,418	0,390
0,01	0,316	0,227	0,204	0,197

Таблица 6. Мощность многовыборочных критериев однородности дисперсий относительно конкурирующей гипотезы $H_3: \sigma_m = 1,5\sigma_1$

Количество выборок $m=3$, $n=100$				
α	Кокрена	Бартлетта	Хартли	Левене
0,1	0,997	0,996	0,995	0,990
0,05	0,994	0,990	0,988	0,979
0,01	0,974	0,962	0,947	0,926
Количество выборок $m=5$, $n=100$				
0,1	0,998	0,996	0,995	0,991
0,05	0,997	0,991	0,989	0,982
0,01	0,987	0,969	0,955	0,944

Выводы. Таким образом, исследования распределений статистик и мощности параметрических критериев проверки гипотез об однородности дисперсий показали следующее. Во-первых, в случае анализа 2-х выборок критерии Бартлетта, Кокрена, Хартли и Фишера обладают идентичной мощностью как при выполнении предположений о нормальности, так и при их нарушении. Очевидно, что при выполнении предположений о нормальности выборок применение критерия Бартлетта предпочтительней, так как распределения его статистики практически не зависят от объемов выборок.

В случае анализа более 2-х выборок предпочтение следует отдать критерию Кокрена. Определенным затруднением для следования этой рекомендации является то, что в литературных источниках содержится достаточно ограниченные данные с таблицами процентных точек для статистики этого критерия (для различного числа и объемов выборок). В современных условиях этот недостаток может быть быстро ликвидирован с использованием технологий компьютерного моделирования [14].

В случае нарушения предположений о нормальности анализируемых выборок распределения всех рассмотренных критериев становятся зависящими от объемов выборок (критерий Бартлетта теряет свое преимущество). Естественно, что в этом случае распределения статистик критериев зависят от вида законов, которым подчиняются выборки. Следует отметить, что мощность всех рассмотренных в данной части работы критериев относительно тех же конкурирующих гипотез тем выше, чем “легче” по сравнению с нормальным законом хвосты симметричного распределения, которому подчиняются анализируемые выборки. При этом, как правило, наиболее предпочтительным является критерий Кокрена. Однако в случае распределений с более “тяжелыми хвостами”, более высокую мощность имеет критерий Левене.

В следующей части работы мы рассмотрим особенности применения и мощность непараметрических критериев, используемых для проверки гипотез об однородности характеристик рассеяния, а также остановимся на реализации возможности применения классических критериев в условиях нарушения предположений о нормальности анализируемых величин.

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. О сходимости распределений статистик и мощности критериев однородности Смирнова и Лемана-Розенблатта // Измерительная техника. 2005. № 12. – С.9-14. [Lemeshko B.Yu., Lemeshko S.B. Statistical distribution convergence and homogeneity test power for Smirnov and Lehmann–Rosenblatt tests // Measurement Techniques, 2005. V. 48, № 12. – P.1159-1166]
2. Лемешко Б.Ю., Лемешко С.Б. Об устойчивости и мощности критериев проверки однородности средних // Измерительная техника. 2008. № 9. – С.23-28. [Lemeshko B.Yu., Lemeshko S.B. Power and robustness of criteria used to verify the homogeneity of means // Measurement Techniques. 2008. Vol. 51, № 9. – P.950-959].
3. 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].
4. Лемешко Б.Ю., Пономаренко В.М. Исследование распределений статистик, используемых для проверки гипотез о равенстве дисперсий при законах ошибок, отличных от нормального // Научный вестник НГТУ. – 2006. – №2(23). – С. 21-33.

5. Bartlett M.S. Properties of sufficiency of statistical tests // Proc. Roy. Soc.. – 1937. – A 160. – P. 268-287.

6. Cochran W.G. The distribution of the largest of a set of estimated variances as a fraction of their total // Annals of Eugenics. – 1941. – V.11. – P. 47-52.

7. Hartley H.O. The maximum F-ratio as a short-cut test of heterogeneity of variance // Biometrika. – 1950. – V.37. – P. 308-312.

8. Levene H. Robust tests for equality of variances // Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. – 1960. – P. 278-292.

9. Ansari A.R., Bradley R.A. Rank-tests for dispersions // AMS.1960. V.31. №4. – P.1174-1189.

10. Mood A. On the asymptotic efficiency of certain nonparametric tests // AMS. – 1954. – V.25. – P. 514-522.

11. 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.

12. Capon J. Asymptotic efficiency of certain locally most powerful rank tests // AMS. – 1961. – V.32, №1. – P. 88-100.

13. Klotz J. Nonparametric tests for scale // AMS. – 1962. – V.33. – P. 498-512.

14. Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2004. – 120 с.

15. Лемешко Б.Ю. Статистический анализ одномерных наблюдений случайных величин: Программная система. - Новосибирск: Изд-во НГТУ, 1995. - 125 с.

16. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.

17. Levene Test for Equality of Variances. In: e-Handbook of Statistical Methods (http://www.itl.nist.gov/div898/handbook/eda/section3/eda35a.htm).

18. Neel J.H., Stallings W.M. A Monte Carlo Study of Levene`s Test of Homogeneity of Variance: Empirical Frequencies of Type I Error in Normal Distributions. (Paper presented at the Annual Meeting of the American Educational Reserch Association Convention (Chicago, Illinois, April 1974).

19. Brown M.B., Forsythe A.B. Robust Tests for Equality of Variances // J. Amer. Statist. Assoc. 1974. V.69. – P.364-367.

Lemeshko B.Yu., Lemeshko S.B., Gorbunova A.A.

About using and power of variance homogeneity tests. P. I

The comparative analysis of power is made for classical tests of variance homogeneity (Fisher's, Bartlett's, Cochran', Hartley' and Levene's tests). Statistics distributions of tests are investigated at infringement of assumptions that samples belongs to the normal law.

Key words: test of homogeneity of variances, Fisher's test, Bartlett's test, Cochran's test, Hartley's test, Levene's test, power of test.