Материалы X международной конференции “Актуальные проблемы электронного приборостроения”
АПЭП-2010. Т.6, Новосибирск, 2010. – С.36-41.

Критерии проверки гипотез об однородности дисперсий при наблюдаемых законах, отличных от нормального

Алиса А. Горбунова, Борис Ю. Лемешко, Станислав Б. Лемешко
Новосибирский государственный технический университет, Новосибирск, Россия
gorbunova.alisa@gmail.com

***Аннотация* – 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. Исследованы распределения и мощность непараметрических критериев однородности характеристик рассеяния (Ансари-Бредли, Муда, Сижела-Тьюки, Кейпена и Клотца). Проведен сравнительный анализ мощности классических критериев однородности дисперсий с мощностью непараметрических критериев. Построены таблицы процентных точек для критерия Кокрена в случае некоторых законов, отличных от нормального.**

***Ключевые слова* – Однородность дисперсий, параметрические критерии, непараметрические критерии, мощность критерия.**

I. ВВЕДЕНИЕ

Проверка однородности выборок – часто используемая на практике задача статистического анализа. Естественно, что наиболее полные выводы могут быть получены в случае проверки гипотез об однородности законов распределения, однако исследователя могут в большей степени интересовать вопросы о возможных отклонениях в средних значениях выборок или о различии в характеристиках рассеяния результатов измерений.

Одним из основных предположений при построении классических критериев проверки однородности дисперсий является принадлежность наблюдаемых случайных величин (погрешностей измерений) нормальному закону. Поэтому применение классических критериев всегда сопряжено с вопросом, насколько корректны получаемые выводы в данной конкретной ситуации? При нарушении предположения о принадлежности анализируемых величин нормальному закону условные распределения статистик критериев при справедливости проверяемой гипотезы, как правило, сильно изменяются.

В имеющихся публикациях нет полной информации о мощности классических критериев проверки однородности дисперсий и о сравнительном анализе мощности классических критериев и непараметрических критериев проверки гипотез об однородности характеристик рассеяния (параметров масштаба).

Настоящая работа продолжает исследования устойчивости критериев проверки гипотез о равенстве дисперсий [1, 2]. Сравниваются классические критерии Бартлетта [3], Кокрена [4], Фишера, Хартли [5], Левене [6], рассматриваются непараметрические (ранговые) критерии Ансари-Бредли [7], Сижела-Тьюки [8], Муда [9], Кейпена [10] и Клотца [11].

Цель работы заключалась:

– в исследовании распределений статистик перечисленных критериев при законах распределения наблюдаемых случайных величин, отличных от нормального;
– в сравнительном анализе мощности критериев относительно конкретных конкурирующих гипотез;
– в реализации возможности применения классических критериев в условиях нарушения предположений о нормальности случайных величин.

II. ПОСТАНОВКА ЗАДАЧИ

Проверяемая гипотеза о постоянстве дисперсий m выборок имеет вид

$$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_m^2 , \quad (1)$$

а конкурирующая с ней гипотеза

$$H_1 : \sigma_{i_1}^2 \neq \sigma_{i_2}^2 , \quad (2)$$

где неравенство выполняется хотя бы для одной пары индексов i_1, i_2.

Исследование распределений статистик, построение для этих распределений процентных точек и оценка мощности критериев относительно различных конкурирующих гипотез опирались на развиваемые на кафедре прикладной математики Новосибирского государственного технического университета компьютерные технологии исследования вероятностных и статистических закономерностей, в основе которых лежит методика статистического моделирования и развиваемое программное обеспечение. 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

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

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 конкурирующие гипотезы вида H_1: $\sigma_m = d\sigma_0$. 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: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2 = \sigma_0^2$.

III. ИССЛЕДУЕМЫЕ КРИТЕРИИ

A. Критерий Бартлетта

Статистика критерия Бартлетта [3] вычисляется в соответствии с соотношением:

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

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. При нормально распределенных результатах измерений распределение статистики (4) практически не зависит от объемов выборок n_i [1]. Если законы распределения наблюдаемых величин отличаются от нормального закона, распределение $G(\chi^2 | H_0)$ статистики (4) *становится зависящим от* n_i и отличается от χ^2_{m-1}.

B. Критерий Кокрена

Когда все n_i одинаковы, $n_1 = n_2 = \ldots = n_m = n$, может использоваться более простой критерий Кокрена [4]. Статистика Q критерия Кокрена выражается формулой

$$Q = \frac{S^2_{\max}}{S_1^2 + S_2^2 + \cdots + S_m^2}, \quad (5)$$

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

Распределения статистики Кокрена сильно зависят от объема наблюдаемых выборок. В справочной литературе для ограниченного числа значений n приводятся только таблицы процентных точек, которые и используются при проверке гипотез.

C. Критерий Хартли

Критерий Хартли [5], так же как и критерий Кокрена, используется в случае выборок равного объема.

Статистика критерия Хартли проверки гипотезы однородности дисперсий имеет вид

$$F = \frac{s^2_{\max}}{s^2_{\min}}, \quad (6)$$

где $S^2_{\max} = \max(S_1^2, S_2^2, \ldots, S_m^2)$, $S^2_{\min} = \min(S_1^2, S_2^2, \ldots, S_m^2)$, m – число независимых оценок дисперсий (число выборок).

В литературе для распределений статистики (6) приводятся лишь таблицы процентных точек в зависимости от $\nu_1 = m$ и $\nu_2 = n - 1$.

D. Критерий Левене

Статистика критерия Левене [6] имеет вид:

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

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 некоторых описаниях критерия 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. 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).

Критерий Левене менее чувствителен к отклонениям анализируемых выборок от нормального закона. Однако он оказывается и менее мощным.

В оригинальном критерии Левене предусмотрено использование только выборочных средних. В работе [12] предложено в статистике вида (7) в качестве оценок среднего использовать выборочные медиану и усеченное среднее.

Однако наши исследования показали, что *использование в (7) выборочных медианы или усеченного среднего приводит к изменению распределения* $G(W|H_0)$ *статистики* (7).

E. *Критерий Фишера*

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 $\nu_2 = n_2 - 1$. 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.

F. *Критерий Ансари-Бредли*

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 значений.

Статистика критерия Ансари-Бредли [7] имеет вид

$$S = \sum_{i=1}^{n_1} \left\{ \frac{n_1 + n_2 + 1}{2} - \left| R_i - \frac{n_1 + n_2 + 1}{2} \right| \right\}, \tag{9}$$

где R_i - ранги первой выборки в общем вариационном ряду.

В случае принадлежности выборок одному и тому же закону распределение статистики (9) при справедливости проверяемой гипотезы H_0 не зависит от вида этого закона. *Дискретностью* распределения статистики (9) *можно пренебречь практически с* $n_1, n_2 > 40$.

G. *Критерий Сижела-Тьюки*

Вариационный ряд, построенный по объединенной выборке, $x_1 \leq x_2 \leq \ldots \leq x_n$, 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 [8] имеет вид

$$R = \sum_{i=1}^{n_1} R_i . \tag{10}$$

Дискретностью распределения статистики (10) *можно пренебречь практически с* $n_1, n_2 > 30$.

H. *Критерий Муда*

Статистика критерия [9] имеет вид

$$M = \sum_{i=1}^{n_1} \left(R_i - \frac{n_1 + n_2 + 1}{2} \right)^2, \tag{11}$$

где R_i - ранги меньшей по объему выборки в общем вариационном ряду двух выборок. *Дискретностью* распределения статистики (11) *можно вообще пренебречь с* $n_1, n_2 > 20$.

I. *Критерий Кейпена*

Если R_i - ранг i -го элемента меньшей по объему выборки в общем упорядоченном по возрастанию ряду, то статистика критерия [10] может быть записана в виде

$$K = \sum_{i=1}^{m} a_{m+n} \; R_i \;, \; m \leq n, \tag{12}$$

где $a_{m+n} \; R_i$ - математическое ожидание квадрата i -ой порядковой статистики в выборке объема $m + n$ из стандартного нормального распределения, называемые метками критерия. Значения $a_N \; i$ приведены в литературе, например в [13].

В отличие от предыдущих критериев распределение статистики (12) данного критерия является достаточно гладким.

J. *Критерий Клотца*

Статистика критерия [11] имеет вид

$$L = \sum_{i=1}^{m} u^2_{\frac{R_i}{m+n+1}}, \tag{13}$$

где u_γ - γ -квантиль стандартного нормального распределения, R_i - ранги, полученные элементами меньшей по объему выборки (объема m) в общем упорядоченном по возрастанию ряду.

Распределение статистики (13) также достаточно гладкое.

При объемах выборок $n_1, n_2 > 10$ дискретные распределения статистик (9)- (13) достаточно хорошо приближается нормальным законом. Поэтому вместо статистик (9)-(13) чаще используются их нормированные аналоги, которые приближенно подчиняются стандартному нормальному закону.

IV. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

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 [14] однозначно говорится о более низкой мощности критерия Кокрена по сравнению с критерием Бартлетта. В [1] на примере проверки гипотез об однородности дисперсий пяти выборок было показано, что выше мощность у критерия Кокрена.

Исследование мощности критериев Бартлетта, Кокрена, Хартли, Фишера и Левене относительно конкурирующих гипотез вида H_1: $\sigma_2 = d\sigma_1$, $d \neq 1$ (при числе выборок $m = 2$ в случае принадлежности случайных величин нормальному закону) показало, что ***критерии Бартлетта, Кокрена, Хартли и Фишера в случае проверки однородности 2-х выборок имеют одинаковую мощность***. Критерий Левене заметно им уступает.

При законах 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 с более “тяжелыми хвостами” (например, в случае принадлежности выборок распределению Лапласа) критерий Левене имеет преимущество в мощности.

Критерии Бартлетта, Кокрена, Хартли и Левене могут применяться при числе выборок $m > 2$. В таких ситуациях мощность этих критериев оказывается различной. При $m > 2$ в случае выполнения предположений о нормальном законе данные критерии можно упорядочить по убыванию мощности следующим образом:

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

Порядок предпочтения сохраняется и в случае нарушений предположений о нормальном законе. Исключение касается ситуаций, когда выборки принадлежат законам с более “тяжелыми хвостами” по сравнению с нормальным законом. Например, в случае принадлежности выборок закону Лапласа критерий Левене оказывается несколько мощнее трех других**.**

Результаты исследований мощности непараметрических критериев показали заметное преимущество критерия Муда и практическую эквивалентность критериев Ансари-Бредли, Сижела-Тьюки, Кейпена и Клотца. Естественно, что непараметрические критерии уступают по мощности параметрическим. На Рис. 1 приведены графики мощности критериев относительно конкурирующих гипотез H_1^1: $\sigma_2 = 1.1\sigma_1$ и H_1^2: $\sigma_2 = 1.5\sigma_1$ в зависимости от объема выборок n_i при $\alpha = 0.1$ в случае нормального закона. Как видим, преимущество в мощности критерия Кокрена 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.

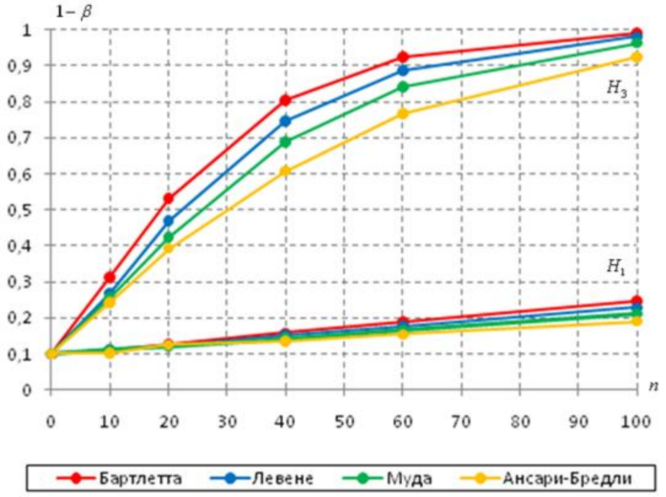


Рис. 1. Мощность критериев относительно конкурирующих гипотез H_1 и H_3 в зависимости от объема выборки n при $\alpha = 0.1$ в случае нормального закона.

Распределения статистик непараметрических критериев не зависят от вида закона, если обе выборки принадлежат одной и той же генеральной совокупности. Но, если выборки принадлежат разным законам, то при справедливости проверяемой гипотезы H_0 о равенстве дисперсий *распределения статистик непараметрических критериев изменяются: они зависят от вида этих законов.*

На Рис. 2, в качестве примера, подтверждающего этот факт, показаны распределения нормированной статистики 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.

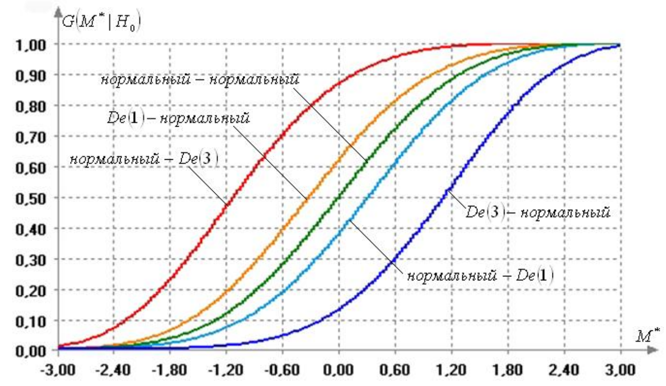


Рис. 2. Распределения нормированной статистики критерия Муда при справедливости H_0 в случае принадлежности пары выборок различным законам семейства (3).

V. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Классические критерии имеют значительное преимущество в мощности по сравнению с непараметрическими. Это преимущество сохраняется и тогда, когда анализируемые выборки принадлежат законам, существенно отличающимся от нормального. Поэтому есть основания для исследования распределений статистик классических критериев проверки однородности дисперсий (построение моделей распределений или таблиц процентных точек) при законах, наиболее часто используемых на практике (отличающихся от нормального закона). Из рассмотренных критериев в наибольшей степени на эту роль подходит критерий Кокрена.

В случае принадлежности наблюдаемых величин семейству распределений (3) при значениях параметра формы $\theta_0 = 1,\ 2,\ 3,\ 4,\ 5$ и ряда значений n на основании результатов статистического моделирования построены таблицы верхних процентных точек (1%, 5%, 10%) статистики критерия Кокрена (для $2 - 5$ выборок). Полученные результаты могут использоваться в ситуации, 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 [1], и расширяют возможности применения критерия Кокрена.

VI. ВЫВОДЫ И ЗАКЛЮЧЕНИЕ

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 развитого программного обеспечения решение таких задач не вызывает принципиальных трудностей, а наличие программного обеспечения позволяет решать эти задачи по мере возникновения потребности.

Разработано программное обеспечение, позволяющее корректно применять критерии однородности дисперсий в случае принадлежности анализируемых выборок произвольным законам.

Настоящие исследования выполнены при поддержке Российского фонда фундаментальных исследований (проект № 09-01-00056-а) и ФЦП "Развитие научного потенциала высшей школы" на 2009-2013 годы.

СПИСОК ЛИТЕРАТУРЫ

[1] 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].

[2] 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.

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

[4] 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.

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

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

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

[8] 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.

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

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

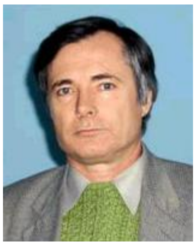
[11] Klotz J. Nonparametric tests for scale // AMS. – 1962. – V.33. – P. 498-512.

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

[13] Кобзарь, А.И. Прикладная математическая статистика: Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006. – 816 с.

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

Горбунова Алиса Александровна получила степень бакалавра прикладной математики в Новосибирском государственном техническом университете в 2009 году. В настоящее время продолжает обучение в магистратуре факультета прикладной математики и информатики. Имеет 5 публикаций.



Лемешко Борис Юрьевич, д.т.н., профессор кафедры прикладной математики Новосибирского государственного технического университета, декан факультета прикладной математики и информатики. Область научных интересов – прикладная математическая статистика, компьютерные методы анализа данных и исследования статистических закономерностей. Имеет около 300 публикаций.

Лемешко Станислав Борисович, к.т.н., с.н.с. кафедры прикладной математики Новосибирского государственного технического университета. Область научных интересов – компьютерные методы анализа данных и исследования статистических закономерностей. Имеет 35 публикаций.