Научный вестник НГТУ. - 2006. - № 2(23)

УДК 62-83: 531.3

Исследование распределений статистик, используемых для проверки гипотез о равенстве дисперсий при законах ошибок наблюдений, отличных от нормального*

Б.Ю. ЛЕМЕШКО, В.М. ПОНОМАРЕНКО

Методами статистического моделирования исследуются распределения статистик критериев Хартли и критерия, предложенного Шеффе, при нарушении предположений о нормальном законе распределения ошибок наблюдений. Исследуется мощность критериев. Исследуется мощность критерия Шеффе в условиях нарушения предположения о нормальности. Приводятся рекомендации по использованию данных критериев.

1. ВВЕДЕНИЕ

Проверка гипотез в классическом дисперсионном анализе базируется на ряде предположений [1,2], и одно из основных – предположение о нормальном законе ошибок наблюдений. В рамках этих предположений оказалось возможным аналитически вывести предельные распределения ряда статистик.

Очевидно, что наблюдаемые данные далеко не всегда подчиняются нормальному закону. В такой ситуации правомерно возникает вопрос, насколько корректным оказывается применение классического аппарата проверки гипотез о средних, о дисперсиях или методов множественного сравнения? В каких случаях можно без боязни использовать классические критерии, а когда их применение чревато неверными выводами, и как следует поступать в таких ситуациях?

В ряде работ [1,3] приводятся результаты теоретических и численных исследований устойчивости различного рода критериев проверки гипотез по отношению к виду наблюдаемого закона. Данные источники содержат указания на существенную зависимость от вида закона критериев, касающихся проверки гипотез о дисперсиях, и на слабую зависимость в этом же случае критериев проверки гипотез о средних.

Применение компьютерных технологий моделирования [4,5] позволяет исследовать статистические свойства критериев в условиях нарушения классических предположений. Построение моделей распределений статистик при различных законах распределения ошибок расширяет аппарат и область применения как дисперсионного, так и многих других видов анализа.

Ранее нами проводились исследования статистик критериев дисперсионного анализа, используемых для проверки гипотез о средних в моделях с постоянными уровнями факторов [6,7], для проверки гипотез о дисперсиях в моделях с случайными уровнями факторов [8], проводились исследования, касающиеся проверки гипотез о равенстве дисперсий в задачах контроля качества [9].

Данная работа продолжает исследования устойчивости критериев проверки гипотез о равенстве дисперсий. Нами рассматривается два критерия. Во-первых, критерий Хартли [10], который во многом подобен традиционным критериям

* *Статья получена 20 марта 2006 г.*
Работа выполнена при поддержке Министерства образования и науки РФ (проект № 2006-РИ-19.0/001/119) и РФФИ (проект № 06-01-00059-а)

проверки однородности дисперсий, таким как критерий Кокрена и Бартлетта. Во-вторых, приближенный критерий проверки однородности дисперсий, предложенный Шеффе (далее критерий Шеффе), который по предположению автора [1] должен быть очень устойчив к нарушению предположений о нормальности.

Цель проводимых исследований заключалась в следующем. Во-первых, установлении того, что происходит с распределениями классических статистик, используемых в моделях с постоянными уровнями факторов для проверки гипотезы о равенстве дисперсий, если наблюдаемый закон в той или иной мере отличается от нормального. Во-вторых, в проверке, насколько будут корректны статистические выводы, базирующиеся на классических результатах, если нарушено предположение о нормальности. И, в-третьих, в создании необходимого математического аппарата, обеспечивающего исследователю корректность выводов при законах распределения, существенно отличающихся от нормального.

2. ПОСТАНОВКА ЗАДАЧИ

Критерии однородности дисперсий в дисперсионном анализе принято рассматривать для однофакторной модели вида:

$$y_{ij} = \mu_i + e_{ij}\ ,\ i = 1,\dots,I\ ,\ j = 1,\dots,n_i\ , \tag{1}$$

где $\mu_1,\dots,\mu_I$ – средние отклика y на I уровнях фактора, n_i – число наблюдений на i-м уровне, общее число наблюдений в модели составляет $n_1 + n_2 + \dots + n_I$. В этом случае совокупности наблюдений $\{y_{ij}\}$ при различных значениях i могут рассматриваться как элементы выборок из I генеральных совокупностей с математическим ожиданием равным μ_i, дисперсией σ_i^2 для i-той генеральной совокупности. В классической постановке предполагается, что все наблюдения распределены по нормальному закону.

Проверяемая гипотеза имеет вид

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

а конкурирующая с ней –

$$H_1 : \sigma_{i_1}^2 \neq \sigma_{i_2}^2\ , \tag{3}$$

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

Критерий Хартли был предложен для случая сбалансированного плана наблюдений, т.е. для случая, когда $n_1 = n_2 = \dots = n_I = n$. Статистика критерия имеет вид [10]

$$T_1 = \frac{s_{\max}^2}{s_{\min}^2}\ , \tag{4}$$

где

$$s_{\max}^2 = \max_i s_i^2\ ,\quad s_{\min}^2 = \min_i s_i^2\ ,\quad s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_{ij} - y_{i\bullet})^2\ ,\quad y_{i\bullet} = \frac{1}{n}\sum_{j=1}^{n} y_{ij}\ .$$

В [10] приводятся процентные точки условного распределения статистики (4) в случае справедливости проверяемой гипотезы и нормального закона ошибок наблюдения модели (1).

Критерий Шеффе может применяться при анализе моделей как со сбалансированным, так и с несбалансированным планом наблюдений. В критерии Шеффе статистика опирается не на собственно оценки дисперсий, как это обстоит в случае критериев Хартли, Кокрена и Бартлетта, а на средние значения логарифмов оценок дисперсий. При таком подходе задача сводится к сравнению средних, а критерии проверки гипотез “о средних” устойчивы по отношению к форме распределения ошибок наблюдений. Применяемое логарифмирование позволяет приблизить распределение к нормальному закону.

Чтобы перейти к сравнению средних, каждая i-я выборка наблюдений $\{y_{ij}\}$, $j = 1,\ldots,n_i$, разбивается на J_i групп объемом n_{ij}, так что $n_i = J_i n_{ij}$. Обозначим для удобства совокупность значений, полученную путем распределения значений $\{y_{ij}\}$ на подвыборки, через

$$\{x_{ijk}\},\ i = 1,\ldots,I,\ j = 1,\ldots,J_i,\ k = 1,\ldots,n_{ij}. \tag{5}$$

Тогда статистика критерия Шеффе [1] может быть записана в следующем виде:

$$T_2 = \frac{\nu}{I-1}\frac{\sum_{i=1}^{I} J_i \left(z_{i*} - \bar{z}\right)^2}{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left(z_{ij} - z_{i\bullet}\right)^2}, \tag{6}$$

где

$$z_{i\bullet} = \frac{1}{J_i}\sum_{j=1}^{J_i} z_{ij},\ \bar{z} = \frac{1}{\sum_{i=1}^{I} J_i}\sum_{i=1}^{I} J_i z_{i\bullet},\ \nu = \sum_{i=1}^{I}(J_i - 1). \tag{7}$$

Значения z_{ij}, выступающие в роли исходных наблюдаемых значений для критерия сравнения средних со статистикой (6) вычисляются как

$$z_{ij} = \ln s_{ij}^2, \tag{8}$$

где s_{ij}^2 – выборочная дисперсия подгруппы, определяемая по формуле

$$s_{ij}^2 = \frac{1}{n_{ij}-1}\sum_{j=1}^{n_{ij}}\left(x_{ijk} - x_{ij\bullet}\right)^2,\ x_{ij\bullet} = \frac{1}{n_{ij}}\sum_{j=1}^{n_{ij}} x_{ijk}. \tag{9}$$

По предположению Шеффе статистика T_2 должна подчиняться F-распределению Фишера со степенями свободы $I-1$ и ν. Причем распределение статистики не должно существенно зависеть от закона распределения ошибок $\{e_{ij}\}$, поскольку критерий строился как устойчивый к нарушению предположений о нормальности.

3. УСЛОВИЯ ПРОВЕДЕНИЯ ЭКСПЕРИМЕНТОВ

При верной нулевой гипотезе дисперсия ошибок наблюдения σ_i^2, $i = 1,\ldots,I$ (без потери общности) задавалась равной единице, значения средних отклика на уровнях μ_i (также без потери общности) задавались равными нулю. Выборочные значения наблюдений y_{ij} формировались в соответствии с видом модели (1).

Исследования распределений статистик проводились при различных законах ошибок наблюдений и случайного фактора модели (1). В данном случае приводятся результаты исследований, когда ошибки наблюдений подчинялись следующим законам: нормальному, распределению максимальных значений, семейству симметричных распределений с плотностью

$$De(\lambda) = f(x, \theta_1, \theta_2, \lambda) = \frac{\lambda}{2\sqrt{2}\theta_2 \Gamma(1/\lambda)} \exp\left(-\left(\frac{|x - \theta_1|}{\sqrt{2}\theta_2} \right)^{\lambda} \right) \tag{10}$$

при различных значениях параметра формы λ. Распределение $De(\lambda)$ включает в качестве частных случаев распределение Лапласа (λ=1) и нормальное (λ=2). В данной работе приводятся результаты исследований распределений статистик в случае принадлежности наблюдений семейству $De(\lambda)$, как правило, при значениях параметра формы λ=0.5, 1, 5, 10.

Использование распределений из семейства $De(\lambda)$ для моделирования ошибок наблюдения позволяет задавать симметричные законы распределения различной формы, изменяя значения параметра λ: чем меньше параметр λ, тем “тяжелее” хвосты распределения $De(\lambda)$, чем больше параметр λ, тем хвосты “легче”. Распределение максимальных значений использовалось для выявления влияния на распределения статистик асимметричности закона.

При исследовании поведения статистики критерия Хартли рассматривалось также логистическое распределение. В этом случае выяснялось поведение распределения статистики при “малых” отклонениях от нормальности.

При проведении экспериментов по моделированию распределений статистик объем моделируемых выборок исследуемых статистик составлял, как правило, величину $N = 10000$.

Для контроля правильности результаты моделирования сравнивались с известными аналитическими или численными, полученными другими авторами для «классической» ситуации принадлежности наблюдений нормальному закону. В частности, для критерия Хартли при различных объемах выборок моделировались выборочные квантили, которые сравнивались с табличными [10] значениями процентных точек. А в случае критерия Шеффе эмпирическое распределение статистики (6) сравнивалось с теоретическим распределением Фишера, в соответствии с которым приближенно распределена статистика (6).

Количественной мерой близости эмпирических распределений статистик соответствующим теоретическим функциям распределения при проверке согласия служили достигнутые уровни значимости $P\{S > S^*\}$ для применяемых критериев согласия, где S^* – значение статистики S используемого критерия, вычисленное по конкретной выборке исследуемых статистик. Проверка согласия осуществлялась по критериям χ^2 Пирсона, Колмогорова, ω^2 Крамера-Мизеса-Смирнова, Ω^2 Андерсона-Дарлинга [11, 12].

В тех случаях, когда при исследовании критерия Шеффе не наблюдалось согласия эмпирического распределения с соответствующим распределением Фишера, и каждый раз при исследовании критерия Хартли, строилась приближенная модель для полученного эмпирического распределения статистики. Аппроксимация предельного закона распределения статистики формировалась как усредненное по параметрам распределение, полученное на основании десяти экспериментов.

Как при проверке согласия с известным распределением, так и при построении модели распределений, использовалась программная система ISW, в рамках которой реализованы результаты, полученные с помощью методики компьютерного моделирования, развиваемой на кафедре Прикладной математики НГТУ [13].

4. ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ T_1

Проверка корректности моделирования проводилась путем сравнения значений, взятых из таблиц верхних процентных точек распределения статистики (4) в нормальном случае, и выборочных квантилей, полученных при различных объемах смоделированных выборок статистик.

Значения табличных и выборочных квантилей, также как и точность полученных выборочных квантилей при N = 200000, представлены в таблице 1. Приводимые значения квантилей являются усредненными по 10 экспериментам. Из таблицы видно, что в пределах точности, достигаемой при используемых объемах выборок статистик, получаемые процентные точки весьма близки к табличным значениям, с которыми производится сравнение. В целом это позволяет говорить о корректности используемой методики моделирования.

Таблица 1

Табличные и выборочные (с оценкой точности при N = 200000) квантили распределения статистики (4) в случае нормального распределения ошибок наблюдения при различных значениях I и n, различных объемах N выборки статистик

Размерность задачи	Значение α	Табличное значение	Выборочные квантили при различных объемах выборок статистик		Точность выборочных квантилей, N = 200000
			N = 10000	N = 200000	
I =3, n =5	0.05	15.5	15.48	15.38	0.195
	0.01	37	36.41	36.19	0.732
I =3, n =61	0.05	1.85	1.84	1.84	0.005
	0.01	2.22	2.14	2.14	0.010
I =12, n =5	0.05	51.4	50.50	51.05	0.754
	0.01	120	118.17	118.77	3.496
I =12, n =61	0.05	2.36	2.35	2.35	0.005
	0.01	2.7	2.67	2.67	0.011

Рисунок 1 иллюстрирует поведение распределения статистики при нарушении предположений нормальности. Из рисунка видно, что на распределение статистики (4) сильно влияет закон распределения наблюдений $\{y_{ij}\}$. Чем “тяжелее” хвосты, тем ниже и правее от “нормального” случая находится распределение статистики.

Для ряда значений I и n построены модели распределения статистики (4) при различных законах ошибок наблюдения.

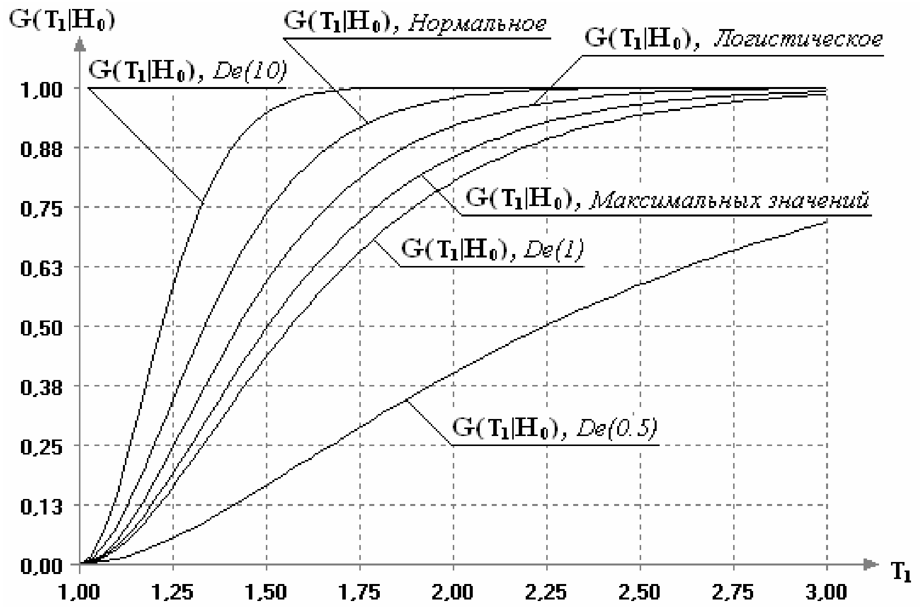

Рис. 1. Эмпирические функции распределения статистики (4) в случае $I=3$, $n=61$ при различных законах ошибок наблюдений

5. ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ T_2

При использовании критерия Шеффе для проверки гипотезы однородности дисперсий (2) имеющееся число наблюдений на уровне фактора n_i, как это видно из формул (5)-(9), может иметь различные разбиения на подгруппы (J_i - число подгрупп и n_{ij} - число наблюдений в подгруппе). При этом в [1] делается предположение о том, что статистика (6) приближенно будет подчиняться F-распределению Фишера с числом степеней свободы $I-1$ и $\nu=\sum_{i=1}^{I}(J_i-1)$. В случае сбалансированного плана наблюдений, который рассматривается в данной работе, имеем распределение $F_{I-1,I(J_i-1)}$, где J_i - одно и тоже для всех $i=1,\ldots,I$. Возникает вопрос: при каких значениях J_i и n_{ij} при нормальном законе распределения ошибок наблюдается достаточно хорошее согласие распределения статистики (6) и распределения $F_{I-1,I(J_i-1)}$, чтобы при проверке гипотезы (2) можно было пользоваться этим F-распределением для вычисления достигнутых уровней значимости.

В таблице 2 приведены достигнутые уровни значимости, полученные при проверке согласия эмпирического распределения статистики (6) при выполнении предположений о нормальности и соответствующего теоретического распределения Фишера. В таблице указаны достигнутые уровни значимости для минимальных значений n_{ij}, при которых уже наблюдается некоторое согласие сравниваемых распределений для ряда заданных значений I и J_i (одно и тоже для всех). Это согласие трудно назвать хорошим. Но, как иллюстрирует рисунок 2, уже при таких значениях n_{ij} использование соответствующего F-распределения в качестве предельного распределения статистики (6) не приведет к большим ошибкам при вычислении соответствующих вероятностей. Из рисунка также видно, что эмпирические распределения, полученные в случае размерностей, представленных в таблице 2, и в случае значений n_{ij}, больше приведенных (а значит и больших значений n_i), практически сливаются.

Как показывают исследования, согласие эмпирических и соответствующих теоретических распределений статистики (6) растет с ростом n_i. Это дает основания считать уровень согласия, отражаемый в таблице 2 (минимально) допустимым для того, чтобы уже при таких объемах выборок рекомендовать подходящее F-распределение к использованию в качестве предельного распределения статистики (6).

Таблица 2

Значения достигнутых уровней значимости, полученных в результате проверки согласия эмпирических распределений статистики (6) с соответствующими F-распределениями при справедливости гипотезы H_0 вида (2) и нормальном законе ошибок наблюдений

Критерий согласия	Размерность задачи; распределение, с которым проверяется согласие			
	$I=2, J_i=2, n_{ij}=5$; $F_{1,2}$	$I=2, J_i=5$ $n_{ij}=4$; $F_{1,8}$	$I=5, J_i=2, n_{ij}=6$; $F_{4,5}$	$I=5, J_i=5, n_{ij}=4$; $F_{4,20}$
Отн. правдоподобия	0.22	0.21	0.14	0.12
χ^2 Пирсона	0.22	0.21	0.14	0.12
Колмогорова	0.22	0.11	0.18	0.16
ω^2 Мизеса	0.24	0.09	0.16	0.16
Ω^2 Анд.-Дарлинга	0.24	0.09	0.12	0.14

Было исследовано, при каких значениях n_i (J_i и n_{ij}) при заданном значении I и нормальном законе ошибок наблюдений достаточно высока близость эмпирического распределения статистики (6) и соответствующего теоретического распределения Фишера. Рассматривались значения I в диапазоне от 2 до 5, J_i в диапазоне от 2 до 6.

По результатам исследований сделаны следующие заключения. Минимальное значение n_i, при котором допустимо использовать F-распределение в качестве предельного распределения статистики (6), составляет 10-12 наблюдений в группе.

Необходимый для приемлемого согласия объем n_i в наибольшей степени определяется выбором числа J_i. Например, при J_i=2 минимально допустимый объем n_i составляет 10-12, при J_i=4 – 16-20, а при J_i=6 – 20-25.

В целом следует отметить, что при нормальном законе ошибок и числе наблюдений n_i в группах около 30, можно без риска совершения больших ошибок использовать F-распределение в качестве предельного распределения статистики (6) при условии, что будет выбрано разбиение, в котором $n_{ij} \geq 4$.

Но, если объем выборок n_i меньше этого числа, что довольно часто встречается в дисперсионном анализе, то для того, чтобы распределение статистики (6) хорошо согласовалось с соответствующим F-распределением, следует выбирать такое разбиение на подгруппы, чтобы n_{ij} было наибольшим, а J_i - наименьшим из возможных.

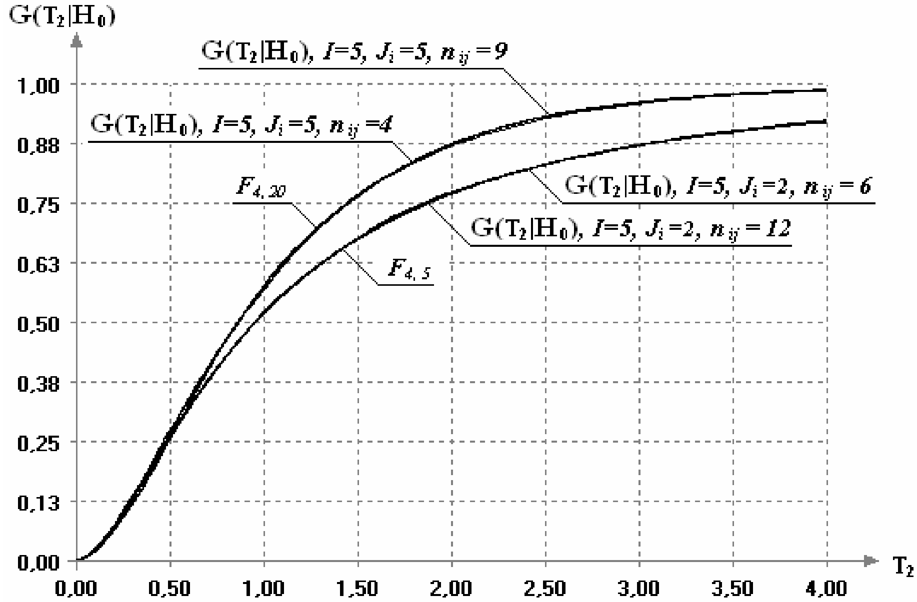

Рис. 2. Эмпирические и теоретические функции распределения статистики (6) в случае нормального закона ошибок наблюдения при I =5 и различных значениях J_i и n_{ij}

В таблице 3 для случая I =5, J_i=5, n_{ij}=4 представлены достигнутые уровни значимости при проверке согласия получаемых в результате моделирования эмпирических распределений статистики (6) и предполагаемого $F_{4,20}$-распределения Фишера при отличных от нормального законах наблюдений $\{x_{ijk}\}$.

Из таблицы следует, что на самом деле распределение $\{x_{ijk}\}$ оказывает влияние на степень согласия эмпирических распределений статистики (6) с распределениями Фишера. В случае распределения с "тяжелыми" хвостами (например, $De(0.5)$) наблюдается очень высокая степень согласия эмпирических распределений статистики (6) и $F_{4,20}$-распределения, по которому должна быть приблизительно распределена статистика (6) в “нормальном” случае. Из таблицы 3 видно, что в случае распределения наблюдений по закону $De(0.5)$ степень согласия распределения статистики (6) с $F_{4,20}$-распределением существенно выше, чем в нормальном случае. В случае принадлежности наблюдений распределению максимальных значений – несколько выше, чем в нормальном случае. А вот при ошибках наблюдений по законам с "легкими" хвостами, например, $De(5)$ и $De(10)$ согласие распределения статистики (6) с $F_{4,20}$-распределением при J_i =5, n_{ij} =4 уже практически не наблюдается. Причем согласие тем хуже, чем более "легкими" хвостами обладает закон распределения ошибок.

Таблица 3 иллюстрирует также влияние разбиения на подвыборки в случае принадлежности ошибок наблюдения законам, отличным от нормального. Из таблицы видно, что при увеличении n_{ij} (при постоянном $n_i = 20$) согласие распределения статистики (6) с $F_{4,20}$-распределением растет в случаях ошибок с несимметричными законами и законами с «легкими» хвостами. При этом в случае

$J_i=2$, $n_{ij}=10$ при законах распределения $De(5)$ и $De(10)$ согласие достигает приемлемого уровня.

В целом результаты исследований позволяют рекомендовать при разбиении наблюдений на подгруппы делать это так, чтобы объем подгруппы n_{ij} был максимален. Причем данная рекомендация справедлива как в случае нормального закона ошибок наблюдений, так и для ситуаций, когда ошибки наблюдений подчиняются законам распределения, отличным от нормального.

Таблица 3

Значения достигнутых уровней значимости, полученных в результате проверки согласия эмпирического распределения статистики (6) с теоретическим $F_{4,20}$-распределением при справедливости гипотезы H_0 вида (2) при $I=5$, $n_i=20$, при различных разбиениях на подвыборки и различных законах ошибок наблюдений

J_i, n_{ij}	Критерий согласия	Распределение ошибок наблюдений			
		Max	$De(0.5)$	$De(5)$	$De(10)$
$J_i=5$, $n_{ij}=4$	Отн. правдоподобия	0.2162	0.5199	0.0034	0.0014
	χ^2 Пирсона	0.2176	0.5195	0.0037	0.0014
	Колмогорова	0.1526	0.5431	0.0395	0.0399
	ω^2 Мизеса	0.1896	0.5291	0.0323	0.0225
	Ω^2 Анд.-Дарлинга	0.1751	0.5158	0.0112	0.0078
$J_i=2$, $n_{ij}=10$	Отн. правдоподобия	0.434	0.374	0.078	0.109
	χ^2 Пирсона	0.433	0.374	0.078	0.109
	Колмогорова	0.525	0.388	0.142	0.171
	ω^2 Мизеса	0.526	0.358	0.132	0.146
	Ω^2 Анд.-Дарлинга	0.523	0.312	0.093	0.125

Исследования показали, что наблюдаемые различия в степени близости распределения статистики (6) и соответствующего распределения Фишера при различных законах наблюдений $\{x_{ijk}\}$, объясняется тем, как распределены $z_{ij}=\ln s_{ij}^2$, непосредственно входящие в статистику T_2 (6). Распределение z_{ij} становится тем симметричнее, чем "тяжелее" хвосты распределения наблюдений $\{x_{ijk}\}$. Так, в случае принадлежности $\{x_{ijk}\}$ закону $De(0.5)$ распределение z_{ij} наиболее симметрично, в случае принадлежности $\{x_{ijk}\}$ закону $De(10)$ распределение z_{ij} наименее симметрично.

6. ИССЛЕДОВАНИЕ МОЩНОСТИ КРИТЕРИЕВ ХАРТЛИ И ШЕФФЕ

В таблице 4 приведены значения мощности для 4-х критериев проверки однородности дисперсий: Шеффе, Хартли, Бартлетта и Кокрена при достаточно больших объемах выборок $n_i=200$ и $n_i=500$. Данные по критериям Бартлетта и Кокрена взяты из работы [9]. При этом для критерия Шеффе для каждого набора значений $I=5$ и n_i приведены значения мощности, полученные при различных способах разбиения на подгруппы. Рассматривается альтернатива $\sigma_I^2 = 1.44\sigma_1^2$, достаточно близкая к проверяемой гипотезе H_0 вида (2).

Анализируя данные таблицы 4, можно сделать несколько выводов. Во-первых, если мощность критерия Хартли в целом сравнима с мощностью критерия Бартлетта, то мощность критерия Шеффе в большинстве случаев ниже мощности остальных рассматриваемых критериев. Во-вторых, мощность критерия Шеффе очень сильно зависит от того, как именно разбить n_i наблюдений на J_i число подгрупп по n_{ij} наблюдений в каждой. В большинстве случаев наблюдается следующая тенденция: мощность возрастает при увеличении числа подгрупп J_i, но может резко упасть при слишком большом значении J_i, как в случае J_i=50, n_{ij}=10.

Таблица 4

Мощность критериев Кокрена, Бартлетта, Шеффе и Хартли относительно альтернативы вида $H_1: \sigma_I^2 = 1.44\sigma_1^2$ при I=5 и различных значениях n_i в случае нормального закона ошибок наблюдений

α	n_i	Разбиение на подгруппы в случае критерия Шеффе	Критерий Шеффе	Критерий Хартли	Критерий Бартлетта	Критерий Кокрена
0.1	200	J_i=2, n_{ij}=100	0.542	0.846	0.835	0.920
		J_i=10, n_{ij}=20	0.762			
		J_i=20, n_{ij}=10	0.735			
	500	J_i=5, n_{ij}=100	0.987	0.997	0.997	0.999
		J_i=10, n_{ij}=50	0.992			
		J_i=50, n_{ij}=10	0.266			
0.05	200	J_i=2, n_{ij}=100	0.362	0.760	0.757	0.837
		J_i=10, n_{ij}=20	0.647			
		J_i=20, n_{ij}=10	0.619			
	500	J_i=5, n_{ij}=100	0.967	0.993	0.993	0.997
		J_i=10, n_{ij}=50	0.981			
		J_i=50, n_{ij}=10	0.172			
0.01	200	J_i=2, n_{ij}=100	0.111	0.535	0.556	0.671
		J_i=10, n_{ij}=20	0.390			
		J_i=20, n_{ij}=10	0.373			
	500	J_i=5, n_{ij}=100	0.853	0.965	0.970	0.986
		J_i=10, n_{ij}=50	0.920			
		J_i=50, n_{ij}=10	0.062			

В таблице 5 представлены значения мощности критериев Шеффе и Хартли, полученные для альтернатив вида $\sigma_I^2 = C\sigma_1^2$ при различных значениях C в случае различных объемов выборок n_i. Как и ожидалось, при близких альтернативах (C =1.2, C =1.44) и относительно небольшом числе наблюдений в группах: n_i =50 и n_i =20, мощность критериев Шеффе и Хартли уже крайне низка. Приемлемым уровень мощности становится в случае альтернатив, при которых дисперсии отличаются в «разы». Причем у критерия Шеффе при n_i =20 мощность все еще очень низка.

Таблица 5

Мощность критериев Шеффе и Хартли относительно альтернативы вида H_1:

$\sigma_I^2 = C\sigma_1^2$ при I =5 в случае нормального закона распределения

C	n_i	J_i, n_{ij} для критерия Шеффе	α					
			0.1		0.05		0.01	
			Шеф-фе	Харт-ли	Шеф-фе	Харт-ли	Шеф-фе	Харт-ли
1.2	500	J_i =10, n_{ij} =50	0.590	0.655	0.455	0.532	0.221	0.296
	200	J_i =10, n_{ij} =20	0.290	0.343	0.185	0.232	0.060	0.087
	50	J_i =5, n_{ij} =10	0.136	0.163	0.072	0.092	0.016	0.024
	20	J_i =4, n_{ij} =5	0.107	0.130	0.054	0.069	0.011	0.015
1.44	500	J_i =10, n_{ij} =50	0.992	0.997	0.981	0.993	0.920	0.965
	200	J_i =10, n_{ij} =20	0.762	0.846	0.647	0.760	0.390	0.535
	50	J_i =5, n_{ij} =10	0.255	0.324	0.155	0.215	0.045	0.077
	20	J_i =4, n_{ij} =5	0.142	0.178	0.075	0.101	0.017	0.026
3	50	J_i =5, n_{ij} =10	0.946	0.992	0.894	0.983	0.700	0.937
	20	J_i =4, n_{ij} =5	0.519	0.802	0.375	0.701	0.149	0.445

Исследования показали (см. таблицу 6), что на мощность критерия Шеффе также влияет закон распределения наблюдений $\{x_{ijk}\}$. Чем “легче” хвосты распределений, тем выше мощность, и наоборот: чем хвосты “тяжелее”, тем ниже мощность. Влияние на распределение статистики (6) при верной альтернативе несимметричности распределения схоже с влиянием распределения с более “тяжелыми”, чем у нормального закона хвостами. Рисунок 3 иллюстрирует влияние закона распределения наблюдений на распределение статистики при справедливости конкурирующей гипотезы и большом объеме выборки n_i =500.

Но при малых объемах n_i, как видно из таблицы 6, мощность уже настолько мала, что различия, обусловленные видом закона ошибок наблюдений $\{x_{ijk}\}$, незначительны.

Таблица 6

Мощность критерия Шеффе при $I=5$ при различных объемах n_i относительно альтернативы вида H_1: $\sigma_I^2 = 1.2\sigma_1^2$ при различных законах ошибок наблюдений

J_i, n_{ij}	α	Распределение ошибок наблюдения				
		De(0.5)	*Max*	*Norm*	*De*(5)	*De*(10)
$J_i=10$, $n_{ij}=50$	0.1	0.163	0.160	0.590	0.837	0.897
	0.05	0.091	0.090	0.455	0.740	0.822
	0.01	0.023	0.022	0.221	0.494	0.605
$J_i=5$, $n_{ij}=10$	0.1	0.107	0.105	0.136	0.159	0.169
	0.05	0.054	0.053	0.072	0.087	0.093
	0.01	0.011	0.011	0.016	0.021	0.0226

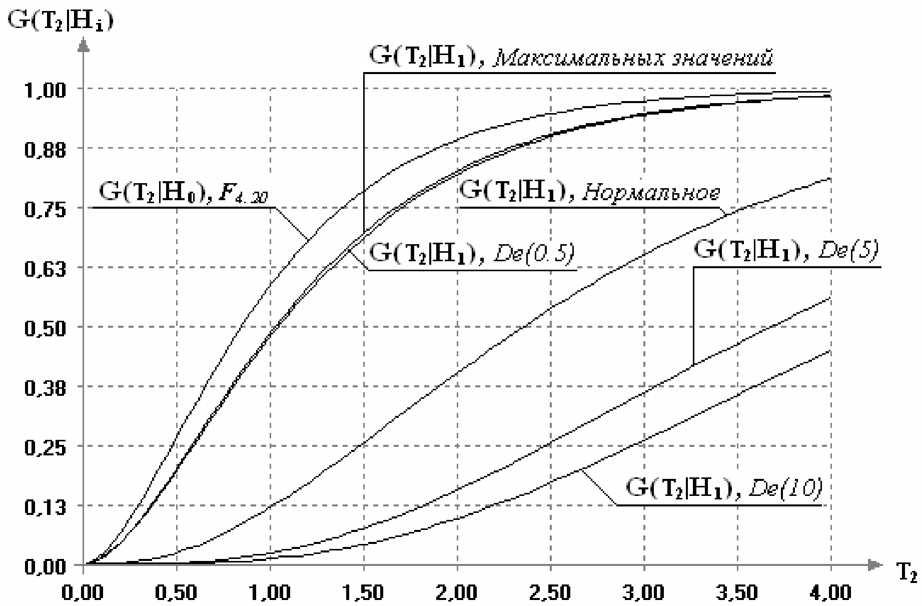

Рис. 3. Функции распределения статистики (6) в случаях справедливости проверяемой гипотезы H_0 вида (2) и справедливости конкурирующей гипотезы H_1 вида $\sigma_I^2 = 1.2\sigma_1^2$ при различных законах наблюдений, при $I=5$, $J_i=10$, $n_{ij}=50$

ЗАКЛЮЧЕНИЕ

Как показали исследования, критерий Хартли крайне чувствителен к закону распределения наблюдений.

В отличие от него, критерий Шеффе действительно достаточно устойчив к нарушению предположений о нормальности. В то же время вид закона ошибок наблюдений влияет на объем выборки, при котором достигается достаточно хорошее согласие распределения статистики с распределением Фишера. Чем "легче" хвосты распределений, тем большие объемы выборок требуются для достаточно хорошего согласия.

Мощность критерия Хартли в нормальном случае сравнима с мощностью критерия Бартлетта, ниже мощности критерия Кокрена и выше мощности критерия Шеффе.

Мощность критерия Шеффе ниже мощности других критериев и зависит от того, каким именно образом наблюдения разбиваются на подгруппы.

Сложен выбор оптимальных значений n_{ij} и J_i при заданном n_i, так как он зависит от вида закона распределения ошибок наблюдения. Если ошибки подчиняются закону с «легкими» хвостами, например, законам $De(10)$ или $De(5)$, то следует выбирать разбиение на подгруппы с минимальным значением J_i. Это позволит с большей уверенностью использовать распределение Фишера в качестве предельного распределения статистики (6). Если же у закона распределения ошибок «тяжелые» хвосты, то следует выбирать разбиение с достаточно большим значением J_i. Это позволит увеличить, малую в этих случаях, мощность критерия. Таким образом, наиболее удачное разбиение зависит от вида закона распределения ошибок.

В целом по критерию Шеффе можно сделать следующий вывод: свойства критерия зависят от закона ошибок наблюдения и от того, насколько удачно выбрано разбиение на подвыборки.

СПИСОК ЛИТЕРАТУРЫ

[1] **Шеффе Г.** Дисперсионный анализ. - М.: Физматгиз, 1980. - 628 с.

[2] **Маркова Е.В. и др.** Дисперсионный анализ и синтез планов на ЭВМ. - М.: Наука, 1982.-195 с.

[3] **Кендалл М., Стьюарт А.** Статистические выводы и связи. - М.: Наука,1973. - 900 с.

[4] **Лемешко Б.Ю.** Компьютерные методы исследования статистических закономерностей // Информационные системы и технологии: ИСТ`2000: Сб. научн. ст. - Новосибирск. 2001. - С.26-41.

[5] **Лемешко Б.Ю., Постовалов С.Н.** Компьютерное моделирование как способ познания статистических закономерностей в технике, экономике, естествознании // Материалы региональной конференции “Вероятностные идеи в науке и философии”. - Новосибирск: Ин-т философии и права СО РАН / НГУ. 2003. - С. 110-113.

[6] **Лемешко Б.Ю. Пономаренко В.М.** Проблемы применения классического аппарата дисперсионного анализа в приложениях технического, экономического и естественно-научного характера // Материалы региональной конференции (с участием иностранных ученых) “Вероятностные идеи в науке и философии”. - Новосибирск: Ин-т философии и права СО РАН / НГУ. 2003. - С. 106-109.

[7] **Lemeshko B.Yu., Ponomarenko V.M.** Statistical Hypotheses Testing In Variance Analysis In Case Of Classical Assumptions Failure // Proceedings of the Seventh International Conference “Computer Data Analysis and Modeling: Robustness and Computer Intensive Methods”, September 6-10, 2004, Minsk. Vol. 1. - P. 110-113.

[8] **Лемешко Б.Ю. Пономаренко В.М.** Проверка гипотез в моделях дисперсионного анализа со случайными факторами при нарушении предположений о нормальности / Доклады АН ВШ РФ, № 2(5).- С. 26-39

[9] **Лемешко Б.Ю., Миркин Е.П.** Критерии Бартлетта и Кокрена в измерительных задачах при вероятностных законах, отличающихся от нормального // Измерительная техника. 2004. № 3. - С. 10-16.

[10] **Закс Л.** Статистическое оценивание. // Пер. с нем. В.Н. Варыгина / Под ред. **Адлера Ю.П., Горского В.Г.**. - М.: “Статистика”, 1976. - 598 с.

[11] Р 50.1.033-2001. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа хи-квадрат. — М.: Изд-во стандартов, 2002. — 87 с.

[12] Р 50.1.037-2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии. — М.: Изд-во стандартов,2002. — 64 с.

[13] **Лемешко Б.Ю.** Статистический анализ одномерных наблюдений случайных величин: Программная система. - Новосибирск: Издательство НГТУ, 1995. - 125 с.