Измерительная техника, 2008, №9. – С.23-28.

УДК 519.24

ОБ УСТОЙЧИВОСТИ И МОЩНОСТИ КРИТЕРИЕВ ПРОВЕРКИ ОДНОРОДНОСТИ СРЕДНИХ

Лемешко Б.Ю., Лемешко С.Б.

Показана устойчивость параметрических критериев проверки однородности средних к нарушению предположения нормальности наблюдений случайных величин. Исследована мощность параметрических и непараметрических критериев проверки однородности средних.

Ключевые слова: мощность критерия, устойчивость критерия, критерий Стьюдента, критерий Манна-Уитни, критерий Краскела-Уаллиса, *F*-критерий

Введение

К критериям проверки гипотез об однородности математических ожиданий (об однородности средних) прибегают при контроле средств измерений, при статистическом анализе результатов экспериментов, при статистическом управлении качеством для проверки наличия возмущения в ходе процесса.

В общем случае проверяемая гипотеза о равенстве математических ожиданий, соответствующих k выборкам, имеет вид

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

при конкурирующей гипотезе

$$H_1 : \mu_{i_1} \neq \mu_{i_2}$$

хотя бы для некоторой пары индексов i_1, i_2.

Для проверки гипотезы H_0 может использоваться ряд параметрических критериев: сравнения двух выборочных средних при известных дисперсиях; сравнения двух выборочных средних при неизвестных, но равных дисперсиях (критерий Стьюдента); сравнения двух выборочных средних при неизвестных и неравных дисперсиях; *F*-критерий. В этих же целях применяется целая совокупность непараметрических критериев: критерий Уилкоксона, критерий Манна–Уитни, критерий Краскела-Уаллиса.

Основным предположением, обуславливающим применение параметрических критериев, является принадлежность анализируемых выборок нормальному закону. Непараметрические критерии свободны от этого требования.

Не смотря на, казалось бы, полную ясность для многих авторов всех нюансов, связанных с применением данных критериев, можно выделить, по крайней мере, два вопроса, недостаточно четко освещенных в литературе и послуживших поводом для проведения настоящих исследований.

Во-первых, насколько важно убедиться в принадлежности анализируемых выборок нормальному закону при использовании параметрических критериев проверки однородности средних? В настоящее время, исследователи в силу различных объективных и субъективных причин зачастую пренебрегают проверкой нормальности наблюдений, вследствие чего и подвергаются справедливой критике за возможную некорректность выводов. Особенно типична такая ситуация при анализе биомедицинских измерений, где встретить выборки, хорошо согласующиеся с нормальным законом весьма проблематично. С другой стороны, можно сослаться на авторитетные суждения о предпочтительности применения непараметрических критериев или об отсутствии необходимости в проверке нормальности, например, при использовании критерия Стьюдента в случае больших объемов выборок.

Второй повод обусловлен достаточно туманными сведениями о мощности упомянутых критериев.

Цель настоящей работы заключалась в исследовании влияния нарушений предположения о нормальности на распределения статистик параметрических критериев и в сравнительном анализе мощности наиболее популярных критериев проверки однородности средних.

1. Критерий сравнения двух выборочных средних при известных дисперсиях

Применение критерия сравнения двух выборочных средних (по двум выборкам) при известных и равных дисперсиях предусматривает вычисление статистики

$$z = \frac{\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}, \qquad (1)$$

где $\bar{x}_i = \dfrac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$, n_i - объем i-й выборки, $i = 1, 2$. В случае принадлежности наблюдений (ошибок измерений) нормальным законам статистика z подчиняется стандартному нормальному закону.

2. Критерий Стьюдента сравнения двух выборочных средних при неизвестных, но равных дисперсиях

Применение критерия сравнения двух выборочных средних при неизвестных, но равных дисперсиях предусматривает вычисление статистики t [1]:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2}{\sqrt{\left[\dfrac{n_1 + n_2}{n_1 n_2}\right]\left[\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]}}, \qquad (2)$$

где

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

В случае принадлежности выборок нормальному закону при справедливости гипотезы H_0 эта статистика подчиняется t_ν-распределению Стьюдента с числом степеней свободы $\nu = n_1 + n_2 - 2$.

3. Критерий сравнения двух выборочных средних при неизвестных и неравных дисперсиях

При неравных объемах выборок $n_1 \neq n_2$ статистика критерия имеет вид [2,3]:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2}{\sqrt{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]}} . \qquad (3)$$

В случае нормального закона и справедливости гипотезы H_0 статистика (3) подчиняется распределению t_ν-Стьюдента с числом степеней свободы

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} .$$

В случае равенства неизвестных дисперсий статистика (3) эквивалентна статистике (2). При неравенстве дисперсий всегда число степеней свободы $\nu < n_1 + n_2 - 2$. Чем больше разница в дисперсиях, соответствующих выборкам, тем сильнее распределение статистики (3) отличается от распределения статистики (2).

Отметим, что при $n_1 + n_2 > 200$ различие между критериями со статистиками (1)-(3) практически исчезает, так как с ростом числа степеней свободы распределение Стьюдента сходится к стандартному нормальному. И при $n_1 + n_2 > 200$ соответствующие распределения Стьюдента практически не отличаются от стандартного нормального. При использовании в таких ситуациях для вычисления достигаемых уровней значимости стандартного нормального закона вместо соответствующего распределения Стьюдента погрешности в определении вероятностей не превышают 0,001.

4. *F*-критерий

В случае справедливости гипотезы о постоянстве (о равенстве) дисперсий с помощью F-критерия можно проверять гипотезу об однородности математических ожиданий по k выборкам [4].

Пусть у нас имеется k выборок объема n. Общая сумма квадратов отклонений по всем выборкам

$$Q_{kn} = \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{kn})^2 ,$$

где

$$\bar{x}_{kn} = \frac{1}{kn} \sum_{i=1}^{k} \sum_{j=1}^{n} x_{ij} = \bar{\bar{x}}_k ,$$

раскладывается на два компонента

$$Q_{kn} = Q_1 + Q_2 ,$$

$$Q_1 = n \sum_{i=1}^{k} (\bar{x}_{in} - \bar{\bar{x}}_k)^2 = n \sum_{i=1}^{k} (\bar{x}_{in}^2 - k\bar{\bar{x}}_k)^2 ,$$

$$Q_2 = \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{in})^2 = (n-1) \sum_{i=1}^{k} s_{in}^2 .$$

Компонент Q_1 является мерой различия в уровнях настройки между k выборками, в то время как Q_2 определяет различие в уровнях настройки внутри этих k выборок.

Для проверки гипотезы используется критерий со статистикой

$$F = \frac{Q_1 / (k-1)}{Q_2 / [k(n-1)]} . \tag{4}$$

Если все выборки извлекаются из нормальной генеральной совокупности, то при справедливости гипотезы H_0 статистка (4) подчиняется F_{ν_1,ν_2}-распределению Фишера со степенями свободы $\nu_1 = k-1$ и $\nu_2 = k(n-1)$ [4].

Принадлежность выборок нормальному закону явилось основным предположением для перечисленных выше критериев, использованным при построении распределений статистик в случае справедливости проверяемой гипотезы H_0.

5. Критерий Манна и Уитни

Ранговый критерий Манна и Уитни [5-8] основан на критерии Уилкоксона [9] для независимых выборок. Он является непараметрическим аналогом *t*-критерия для сравнения двух средних значений непрерывных распределений. Для вычисления статистики упорядочивают $n_1 + n_2$ значений объединенной выборки, определяют сумму рангов R_1, соответствующую элементам первой выборки, и сумму рангов второй R_2. Вычисляются

$$U_1 = n_1 n_2 + \frac{n_1(n_1 - 1)}{2} - R_1,$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 - 1)}{2} - R_2.$$

Статистика критерия имеет вид

$$U = \min\{U_1, U_2\}.$$

Вместо *U*-статистики удобнее использовать статистику

$$\tilde{z} = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}, \tag{5}$$

дискретное распределение которой в случае справедливости проверяемой гипотезы H_0 при $n_1 + n_2 > 60$ хорошо приближается стандартным нормальным законом, когда объем каждой из выборок не слишком мал $n_1 \geq 8$, $n_2 \geq 8$. При меньших объемах выборок следует учитывать, что достигаемый уровень значимости (*p*-значение), вычисляемый по значению статистики в соответствии с функцией распределения стандартного нормального закона, может заметно отличаться от истинного.

6. Критерий Краскела–Уаллиса

H –критерий Краскела–Уаллиса [10,11] является развитием U – критерия для проверки гипотезы о равенстве средних по k выборкам. Объединенную выборку $n = \sum_{i=1}^{k} n_i$ упорядочивают и вычисляют суммы рангов R_i для i-й выборки, $i = \overline{1,k}$. Статистика для проверки гипотезы H_0 имеет вид

$$H = \left[\frac{12}{n(n+1)} \right] \left[\sum_{i=1}^{k} \frac{R_i^2}{n_i} \right] - 3(n+1). \qquad (6)$$

H представляет собой дисперсию ранговых сумм. При больших n_i и k в случае справедливости проверяемой гипотезы H_0 статистика подчиняется χ^2_{k-1}-распределению [11]. В описаниях критерия говорится, что χ^2_{k-1}-распределением практически можно пользоваться при $n_i \geq 5,\ k \geq 4$.

В действительности же при $k = 2$ дискретностью можно практически пренебречь при $n_i \geq 30$. С ростом числа выборок влияние дискретности быстро убывает. При $k = 3$ распределение статистики достаточно хорошо приближается χ^2_{k-1}-распределением, начиная с $n_i \geq 20$, а при $n_i \geq 30$ согласие распределения статистики с χ^2_{k-1}-распределением не отклоняется по всем применяемым критериям согласия [12,13]. При $k \geq 5$ согласие распределения статистики с χ^2_{k-1}-распределением не отклоняется при $n_i \geq 20$.

7. Исследование устойчивости параметрических критериев к нарушению предположений нормальности

При проведении данных исследований использована та же методика компьютерного моделирования и анализа статистических закономерностей, что и в других наших работах, например, [12, 13].

Распределения статистик (1)-(3) при справедливой проверяемой гипотезе H_0 исследовались при различных законах распределения, в частности, в случае принадлежности наблюдений семейству с плотностью

$$f(x)=\frac{\lambda}{2\theta_1\Gamma(1/\lambda)}\exp\left\{-\left(\frac{|x-\theta_0|}{\theta_1}\right)^{\lambda}\right\} \quad (7)$$

с различными значениями параметра формы λ. При $\lambda=2$ выражение (7) дает плотность нормального закона распределения. При больших значениях λ распределение (7) стремится к равномерному, при малых λ получаем симметричные законы с “тяжелыми хвостами”.

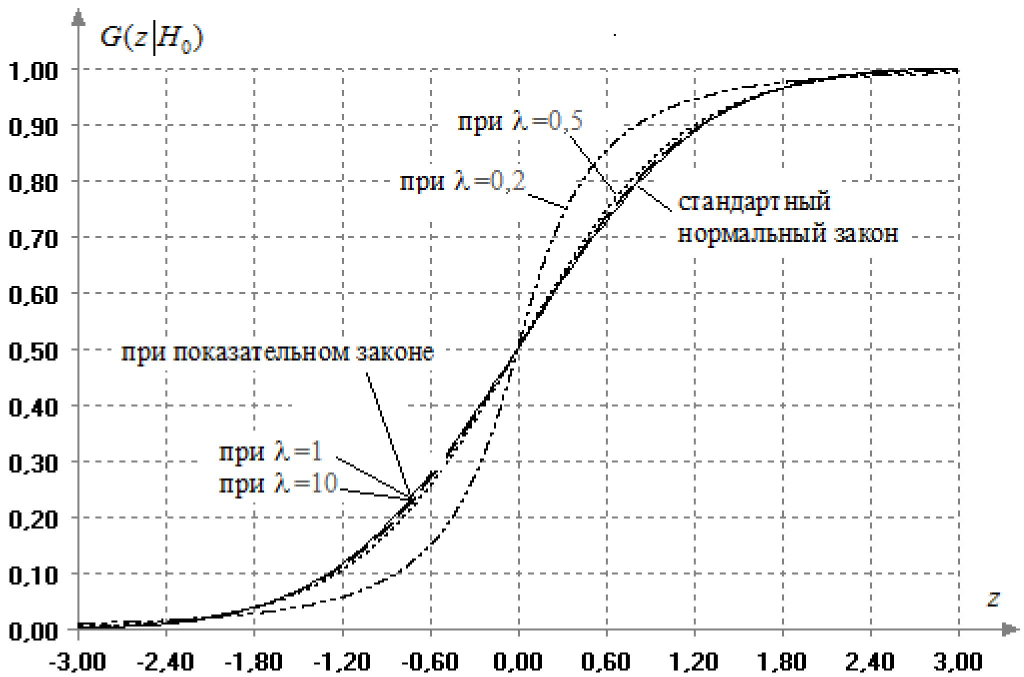

Рис. 1. Эмпирические распределения статистики (1) при различных законах распределения наблюдаемых величин и справедливости гипотезы H_0 при объемах выборок $n_1=n_2=10$

На рис. 1 представлены полученные в результате моделирования распределения статистики (1) в случае принадлежности наблюдаемых величин законам распределения семейства (7) при различных значениях параметра формы и в случае показательного закона с плотностью

$$f(x) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right).$$

Результаты исследований позволяют сделать следующие выводы. Конечно, распределение статистики (1) зависит от законов, которым принадлежат анализируемые выборки. Асимметрия наблюдаемых законов приводит к отличию распределения статистики от стандартного нормального, однако это отличие не настолько велико, чтобы приводить к серьезным ошибкам при использовании критерия. В случае симметричных законов наблюдается устойчивость распределения статистики к значительным отклонениям наблюдаемых законов от нормального (вплоть до равномерного): распределения статистики существенно отличаются от стандартного нормального только в случае законов с “тяжелыми хвостами” (например, при распределении Коши или см., при $\lambda = 0,5$ и $\lambda = 0,2$).

Рис. 2 отражает аналогичную картину зависимости распределений статистики (2) от законов распределения наблюдаемых величин, которая позволяет сделать идентичные выводы об устойчивости критерия Стьюдента. Таким же образом от наблюдаемых законов зависят распределения статистики (3), используемой в критерии при неравных неизвестных дисперсиях. С ростом объемов выборок критерий становится еще более устойчивым к отклонениям наблюдаемых законов от нормального. На рис. 3 показаны распределения статистики (2) при объемах выборок $n_1 = n_2 = 100$ в случае принадлежности выборок нормальному закону и распределениям семейства (7) при $\lambda = 0,5$ и $\lambda = 0,2$.

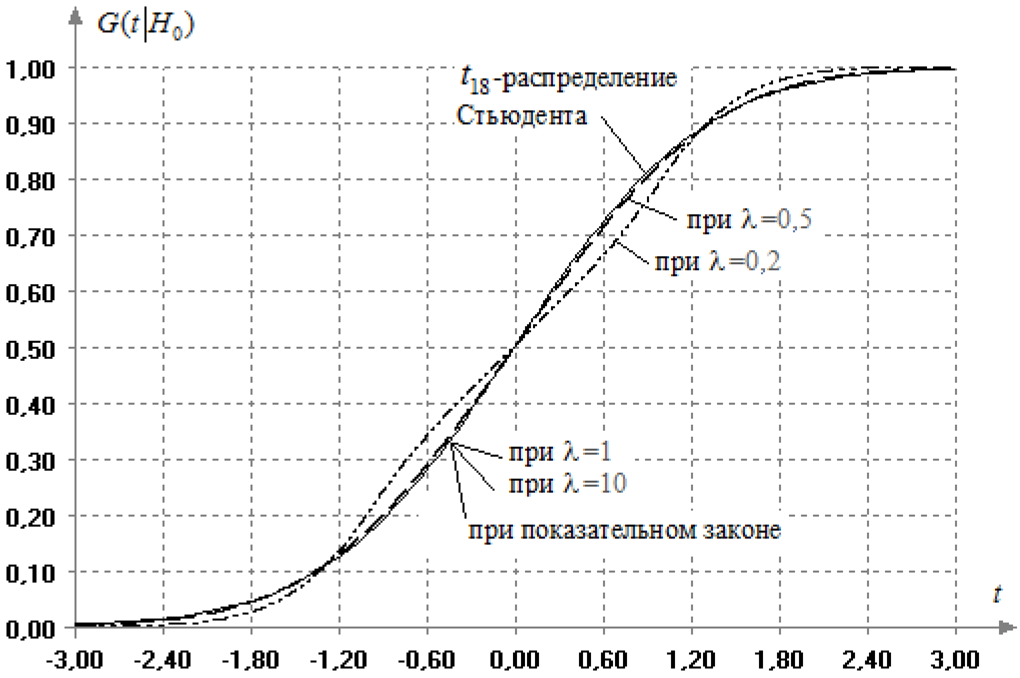

Рис. 2. Эмпирические распределения статистики (2) при различных законах распределения наблюдаемых величин и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 10$

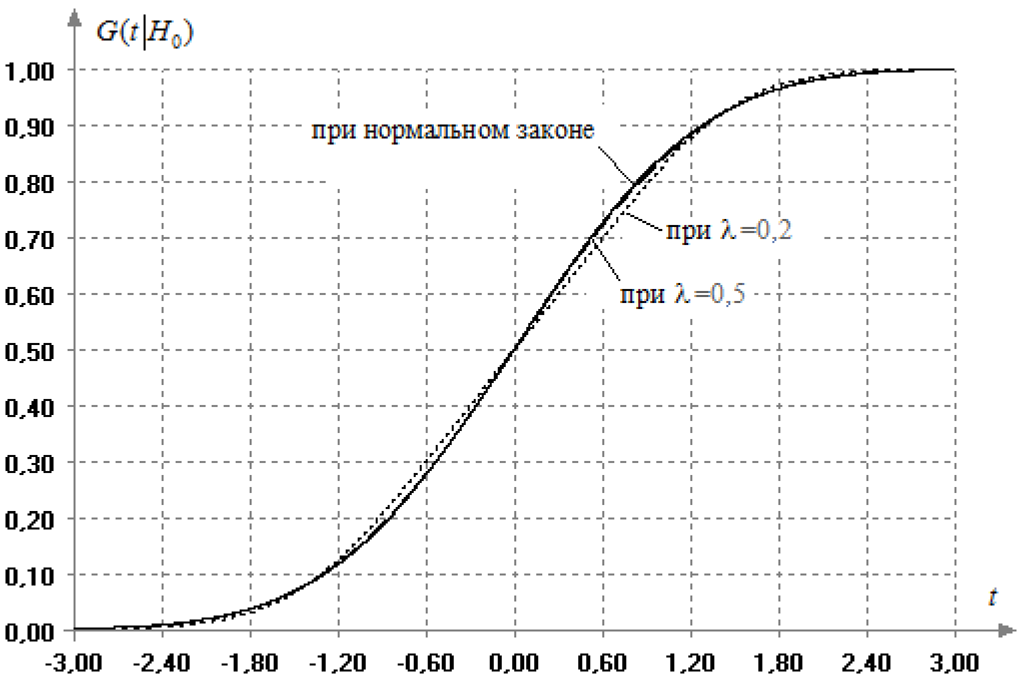

Рис. 3. Эмпирические распределения статистики (2) при различных законах распределения наблюдаемых величин и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 100$

Данные результаты подтверждают общую закономерность: параметрические критерии, связанные с проверкой гипотез о математических ожиданиях, весьма устойчивы к отклонениям наблюдаемых законов от нормального. Это справедливо даже в случае многомерных случайных величин [14].

Распределения статистики (4) F-критерия также устойчивы к отклонениям законов, соответствующих анализируемым выборкам, от нормального. Однако следует подчеркнуть, что применение данного критерия для проверки однородности средних предполагает примерное равенство дисперсий анализируемых выборок. При невыполнении данного условия распределения статистики $G(F|H_0)$ становятся отличными от соответствующего F_{ν_1,ν_2}-распределения. Если отношение максимальной дисперсии к минимальной соответствующих анализируемых выборок не превышает 4, то отклонение распределения статистики $G(F|H_0)$ от F_{ν_1,ν_2}-распределения Фишера не превосходит величины 0,01.

8. Исследование мощности критериев

Рассматривалась проверка однородности средних 2-х выборок. Для исследуемых в работе критериев мощность проанализирована относительно следующих альтернатив $H_1^1: \mu_2 = \mu_1 + 0{,}1\sigma$; $H_1^2: \mu_2 = \mu_1 + 0{,}2\sigma$; $H_1^3: \mu_2 = \mu_1 + 0{,}3\sigma$; $H_1^4: \mu_2 = \mu_1 + 0{,}4\sigma$; H_1^5: $\mu_2 = \mu_1 + 0{,}5\sigma$; H_1^6: $\mu_2 = \mu_1 + \sigma$ при одинаковых дисперсиях выборок. На рис. 4 приведены распределения $\tilde{z}$-статистики (5) критерия Манна-Уитни при справедливости проверяемой $G(\tilde{z}|H_0)$ и конкурирующих гипотез $G(\tilde{z}|H_1^i)$ при объемах выборок $n_1 = n_2 = 10$. Рисунок, с одной стороны, позволяет судить о мощности критерия относительно рассматриваемых альтернатив, а с другой – демонстрирует дискретность распределений статистики, которую следует учитывать, сравнивая мощности критериев.

Вычисленные на основании результатов моделирования оценки мощности $1-\beta$ критериев, где β - вероятность ошибки второго рода, для различных значений уровня значимости α (вероятностей ошибок первого рода) и различных объемов выборок представлены в таблицах 1-4 (для H_1^1, H_1^2, H_1^5, H_1^6). Критерии в таблицах упорядочены по мощности. Выборки распределений статистик моделировались объемом $N=10^6$, что позволило оценивать значения мощности с погрешностью в пределах $\pm 10^{-3}$.

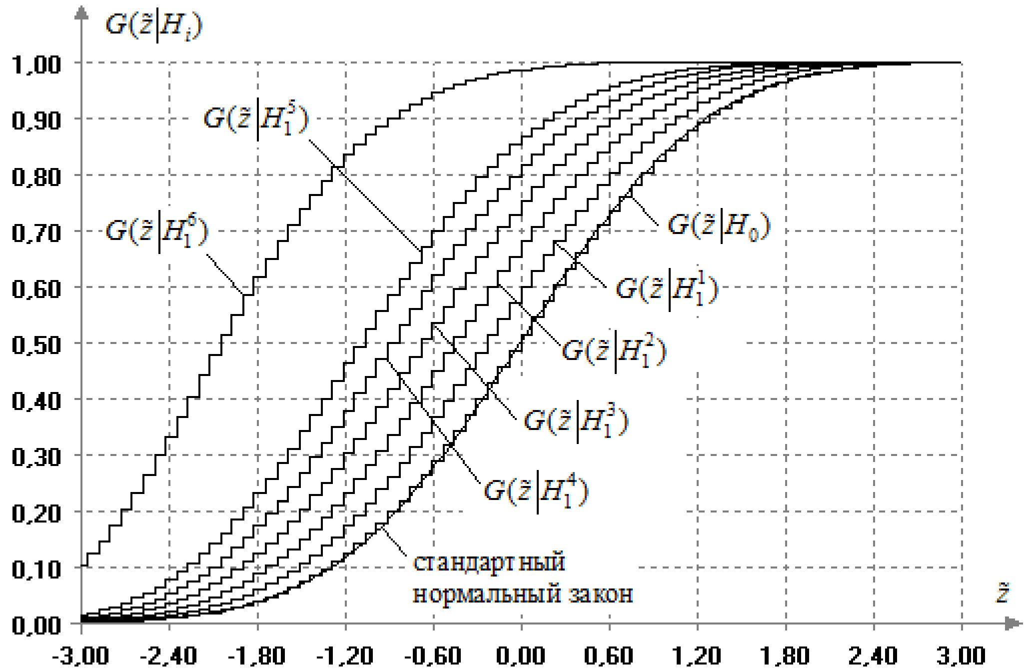

Рис. 4. Эмпирические распределения статистики (5) Манна-Уитни при справедливости различных конкурирующих гипотез при объемах выборок $n_1 = n_2 = 10$

В рассматриваемой ситуации критерий со статистикой (3) эквивалентен критерию со статистикой (2) и имеет ту же мощность. Поэтому он опущен в приводимых таблицах.

Значения мощности для многовыборочных критериев в таблицах 1-4 существенно ниже, чем для критериев со статистиками (1) – (3) и критерия Манна-Уитни. Дело в том, что *F*-критерий и критерий Краскела-Уаллиса в силу вида своих статистик в отличие от критериев со статистиками (1) – (3) и (5) не могут различать альтернативы $\mu_2 > \mu_1 + \Delta\mu$ и $\mu_2 < \mu_1 - \Delta\mu$. Если в выражениях для статистик (1) – (3) и (5) числители взять по модулю, то получим аналогичную ситуацию. Тогда значения мощностей этих критериев для сравнения с много-выборочными надо брать при уровнях значимости $\alpha / 2$.

Какие выводы можно сделать на основе результатов, представленных в таблицах. Первый, очевидный, параметрические критерии обладают большей мощностью по сравнению с непараметрическими. Во-вторых, можно констатировать, что непараметрические критерии совсем немного уступают по мощности параметрическим: критерий Манна-Уитни критерию Стьюдента, а Краскела-Уаллиса – *F*-критерию, соответственно. Кажущееся преимущество $\tilde{z}$-критерия Манна-Уитни при $n = 10$, отраженное в таблицах, объясняется тем, что вследствие дискретности распределения статистики $\tilde{z}$-критерия Манна-Уитни действительные уровни значимости отличаются от заданных в таблице значений α и несколько превышают их. Этим же объясняется "преимущество" в некоторых случаях H-критерия Краскела-Уаллиса по отношению к *F*-критерию.

И, в-третьих. На практике, применяя рассматриваемые критерии для проверки гипотезы об однородности математических ожиданий, как правило, задаются только вероятностью α ошибки 1-го рода. Процедуры контроля предусматривают, чаще всего, небольшие объемы выборок. Как правило, не заходит речи о вероятности β ошибки 2-го рода: не отклонить проверяемую гипотезу при справедливости конкурирующей. В то же время в процедуре контроля при задании α желательно гарантировать величину $\beta \le \alpha$. В данном же случае мы видим, что при конкурирующей гипотезе H_1^1 для $\alpha = 0{,}1$ и объемах выборок $n = 100$

вероятность ошибки 2-го рода составит величину $\beta = 1 - 0,283 = 0,717$ для z-критерия со статистикой (1). При $\alpha = 0,1$ и объемах выборок $n = 100$ данный критерий обеспечит величину $\beta = 0,061 < 0,1$ только для более далекой альтернативы H_1^4. А чтобы с заданным качеством различать гипотезы H_0 и H_1^1 необходимо иметь выборки объемом порядка 1350 наблюдений!

А какие альтернативы можно с таким же качеством (при $\alpha \le 0,1$ и $\beta \le 0,1$) различать по выборкам объемом $n = 10$? Альтернативы, когда μ_2 отличается от μ_1 на величину не менее, чем $1,15\sigma$! При $n = 20$, когда μ_2 отличается от μ_1 на величину порядка $0,82\sigma$, при $n = 30$ - на величину порядка $0,67\sigma$, при $n = 50$ - на величину порядка $0,51\sigma$, при $n = 100$ - на величину порядка $0,364\sigma$.

Таблица 1. Мощность критериев относительно альтернативы H_1^1: $\mu_2 = \mu_1 + 0,1\sigma$

z -критерий при известных дисперсиях					
α	*n=10*	*n=20*	*n=30*	*n=50*	*n=100*
0,1	0,145	0,167	0,186	0,217	0,283
0,05	0,078	0,091	0,105	0,126	0,175
0,01	0,018	0,022	0,026	0,034	0,053
t -критерий Стьюдента при неизвестных и равных дисперсиях					
0,1	0,144	0,166	0,185	0,216	0,283
0,05	0,077	0,091	0,104	0,125	0,174
0,01	0,017	0,021	0,026	0,034	0,053
$\tilde{z}$ -критерий Манна-Уитни					
0,1	*0,153*	0,165	0,184	0,214	0,277
0,05	*0,079*	0,091	0,101	0,123	0,170
0,01	*0,016*	0,021	0,024	0,032	0,051
F-критерий					
0,1	0,109	0,116	0,125	0,141	0,183
0,05	0,055	0,061	0,067	0,078	0,108
0,01	0,012	0,013	0,015	0,020	0,031
H -критерий Краскела-Уаллиса					
0,1	0,113	0,118	0,123	0,141	0,178
0,05	0,057	0,059	0,066	0,078	0,104
0,01	0,008	0,013	0,015	0,019	0,030

Таблица 2. Мощность критериев относительно альтернативы $H_1^2: \mu_2 = \mu_1 + 0,2\sigma$

z-критерий при известных дисперсиях					
α	*n=10*	*n=20*	*n=30*	*n=50*	*n=100*
0,1	0,202	0,258	0,306	0,389	0,552
0,05	0,115	0,155	0,192	0,259	0,409
0,01	0,030	0,045	0,060	0,092	0,182
t-критерий Стьюдента при неизвестных и равных дисперсиях					
0,1	0,199	0,256	0,304	0,387	0,551
0,05	0,112	0,153	0,190	0,257	0,407
0,01	0,028	0,043	0,059	0,090	0,179
$\tilde{z}$-критерий Манна-Уитни					
0,1	*0,209*	0,251	0,299	0,379	0,538
0,05	*0,115*	0,151	0,184	0,250	0,395
0,01	*0,026*	0,041	0,054	0,085	0,170
F-критерий					
0,1	0,131	0,165	0,198	0,261	0,408
0,05	0,071	0,094	0,119	0,168	0,290
0,01	0,017	0,025	0,034	0,056	0,121
H-критерий Краскела-Уаллиса					
0,1	0,136	0,164	0,192	0,261	0,394
0,05	0,073	0,091	0,115	0,168	0,278
0,01	0,011	0,023	0,032	0,054	0,113

Таблица 3. Мощность критериев относительно альтернативы $H_1^5: \mu_2 = \mu_1 + 0,5\sigma$

z-критерий при известных дисперсиях					
α	*n=10*	*n=20*	*n=30*	*n=50*	*n=100*
0,1	0,434	0,618	0,743	0,888	0,988
0,05	0,299	0,475	0,614	0,804	0,971
0,01	0,113	0,228	0,348	0,568	0,887
t-критерий Стьюдента при неизвестных и равных дисперсиях					
0,1	0,424	0,611	0,738	0,886	0,988
0,05	0,285	0,463	0,607	0,799	0,970
0,01	0,099	0,211	0,335	0,556	0,882
$\tilde{z}$-критерий Манна-Уитни					
0,1	*0,430*	0,596	0,725	0,875	0,985
0,05	*0,283*	0,451	0,586	0,781	0,964
0,01	*0,090*	0,199	0,310	0,529	0,865
F-критерий					
0,1	0,288	0,464	0,606	0,799	0,963
0,05	0,185	0,338	0,478	0,697	0,929
0,01	0,060	0,144	0,244	0,453	0,801
H-критерий Краскела-Уаллиса					
0,1	0,286	0,452	0,586	0,799	0,970
0,05	0,184	0,321	0,457	0,697	0,940
0,01	0,043	0,132	0,227	0,434	0,824

Таблица 4. Мощность критериев относительно альтернативы $H_1^6: \mu_2 = \mu_1 + \sigma$

z *-критерий при известных дисперсиях*					
α	*n=10*	*n=20*	*n=30*	*n=50*	*n=100*
0,1	0,830	0,970	0,995	1,000	1,000
0,05	0,723	0,935	0,987	1,000	1,000
0,01	0,463	0,798	0,939	0,996	1,000
t *-критерий Стьюдента при неизвестных и равных дисперсиях*					
0,1	0,816	0,967	0,995	1,000	1,000
0,05	0,693	0,927	0,985	1,000	1,000
0,01	0,398	0,764	0,928	0,995	1,000
$\tilde{z}$ *-критерий Манна-Уитни*					
0,1	*0,811*	0,961	0,993	1,000	1,000
0,05	*0,681*	0,917	0,981	0,999	1,000
0,01	*0,368*	0,739	0,911	0,993	1,000
F-критерий					
0,1	0,693	0,927	0,985	1,000	1,000
0,05	0,562	0,868	0,967	0,999	1,000
0,01	0,294	0,673	0,882	0,990	1,000
H *-критерий Краскела-Уаллиса*					
0,1	0,680	0,917	0,981	1,000	1,000
0,05	0,548	0,849	0,981	0,999	1,000
0,01	0,231	0,640	0,861	0,987	1,000

Выводы

Таким образом, исследования подтвердили устойчивость параметрических критериев проверки однородности математических ожиданий. Это означает, что если закон (законы) распределения анализируемых выборок отличается от нормального, но нет оснований полагать, что наблюдаемые величины принадлежат законам с “тяжелыми хвостами”, применение параметрических критериев со статистиками (1) – (3) остается корректным, по крайней мере, не приводит к существенным погрешностям.

Если дисперсии анализируемых выборок неизвестны и, возможно, различны, лучше воспользоваться критерием со статистикой (3), так как при малых объемах выборок распределение статистики (2) будет существенно отличаться от $t_{n_1+n_2-2}$-распределения Стьюдента.

При $n_1 + n_2 > 200$ для всех критериев со статистиками (1) – (3) в качестве распределений статистик можно использовать стандартный нормальный закон.

$\tilde{z}$-критерий Манна-Уитни, являющийся непараметрическим аналогом критериев со статистиками (1) – (3), совсем немного уступает им по мощности.

Применение *F*-критерия проверки однородности математических ожиданий серии выборок целесообразно, если есть основание полагать, что дисперсии, соответствующие выборкам, примерно одинаковы. В противном случае от него следует отказаться и воспользоваться критерием Краскела-Уаллиса, который немного уступает *F*-критерию по мощности.

Следует помнить, что кроме ошибок 1-го рода есть ошибки 2-го рода. Если проверяемая гипотеза при заданном α не была отклонена, это еще не означает, что она справедлива. Организуя процедуру проверки и предполагая, какие альтернативы должны различаться, необходимо выбирать такие объемы выборок, чтобы вероятность ошибки 2-го рода β оказалась не меньше α.

Работа выполнена при поддержке РФФИ (проект № 06-01-00059).

1. Закс Л. Статистическое оценивание. – М.: Статистика, 1976. – 598 с.
2. Welch B.L. The generalization of “Student`s” problem when several different population variances are involved // Biometrika, 1947. V.34. – P. 29-35.
3. Мардиа К., Земроч П. Таблицы F-распределений и распределений, связанных с ними. М.: Наука, 1984. – 255 с.
4. Миттаг Х.-Й., Ринне Х. Статистические методы обеспечения качества. – М.: Машиностроение. 1995. – 600 с.

5. Mann H.B., Whitney D.R. On a test of whether one of two random variables is stochastically larger than the other // Ann. Math. Statist. 1947. V.18. – P.50-60.
6. Milton R.C. An extended table of critical values for the Mann –Whitney (Wilcoxon) two-sample statistic // J. Amer. Statist. Ass. 1964. V.59. – P.925-934.
7. Hollander M., Wolfe D.A. Non-parametric Statistical Methods (2nd edition). New York: Wiley, 1999.
8. Conover W.J. Practical Nonparametric Statistics (3rd edition). Wiley 1999. 584 p.
9. Wilcoxon F. Individual comparisons by ranking methods // Biometrics Bulletin, 1945. № 1. - P.80-83.
10. Kruskal W.H., Wallis W.A. Use of ranks in one-criterion variance analysis // J. Amer. Statist. Assoc. 1952. V.47. – P.583-621.
11. Kruskal W.H., Wallis W.A. Use of ranks in one-criterion variance analysis // J. Amer. Statist. Assoc. 1953. V.48. – P.907-911.
12. 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.
13. 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.
14. Лемешко Б.Ю., Помадин С.С. Корреляционный анализ наблюдений многомерных случайных величин при нарушении предположений о нормальности // Сибирский журнал индустриальной математики. 2002. – Т.5. – № 3. – С.115-130.