Метрология. 2010. № 12. – С. 3-25.
УДК 519.24

Вопросы применения некоторых критериев проверки случайности и отсутствия тренда

Лемешко Б.Ю., Комиссарова А.С., Щеглов А.Е.

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

Исследованы распределения статистик ряда критериев, используемых при проверке гипотез об отсутствии тренда в математических ожиданиях и в дисперсиях наблюдаемых величин. Отмечены недостатки некоторых критериев. Проведен сравнительный анализ мощности критериев.

Ключевые слова: критерий наличия тренда, критерий автокорреляции, критерии Фостера-Стюарта, критерии Кокса-Стюарта, критерий Вальда-Вольфовитца, критерий Бартелса, критерий Хсу, ранговые критерии обнаружения сдвига, мощность критерия.

1. Введение

Для анализа выборок случайных величин (временных рядов) на предмет отсутствия тренда в характеристиках измеряемой величины в приложениях используется целый ряд параметрических и непараметрических критериев проверки гипотез. Совокупность этих критериев реализована в программных системах статистического анализа, описана в работах авторов, предложивших соответствующий критерий, в различных учебных пособиях, иногда изложение совокупности критериев сконцентрировано в одном источнике [1]. Однако касается ли это конкретного расчета или использования соответствующего программного обеспечения, общим остается одно: ответственность за корректное применение конкретного критерия возлагается на пользователя. И вот здесь, столкнувшись с необходимость проведения статистического анализа и оказавшись перед выбором того или иного критерия, специалист оказывается в тупике, а какой критерий предпочтительней?

Предпосылкой, обуславливающей применение параметрических критериев, как правило, является предположение о принадлежности анализируемых данных нормальному закону. А что произойдет с распределением статистики этого критерия в случае нарушения предположения о нормальности? Насколько будут оставаться корректными выводы, осуществляемые на основании классических результатов?

При ограниченных объемах выборок распределения статистик параметрических и непараметрических критериев могут существенно отличаться от асимптотических распределений этих статистик, используемых в процедуре проверки гипотезы. В случае применения непараметрических критериев эта проблема может усугубляться. Например, вследствие ярко выраженной дискретности значений статистики критерия использование асимптотического распределения вместо истинного распределения этой статистики оказывается не правомерным даже при больших объемах выборок.

Использование непараметрических критериев не требует выполнения предположения о принадлежности выборок анализируемых величин некоторому параметрическому закону (например, нормальному). Однако сведения о мощности непараметрических критериев, сведения о том, насколько они уступают параметрическим критериям, представляются достаточно туманными.

Исследователя может интересовать наличие или отсутствие тренда в математическом ожидании (в среднем) и в дисперсиях.

При *проверке отсутствия тренда* (случайности) *в математическом ожидании* задача формулируется следующим образом. Предполагается, что наблюдается временной ряд $x_1, ..., x_n$ значений взаимно независимых случайных величин с математическими ожиданиями $\mu_1,...,\mu_n$ и одинаковыми (но неизвестными) дисперсиями. Проверяется гипотеза $H_0 : \mu_i = \mu$, $i = 1, 2, ..., n$, о том, что все выборочные значения принадлежат к одной генеральной совокупности со средним μ, против конкурирующей гипотезы о наличии тренда $H_1 : |\mu_{i+1} - \mu_i| > 0$, $i = 1, 2, ..., n-1$.

Для проверки такого рода гипотез предназначен критерий Аббе. В [2, 3] была отмечена устойчивость распределения статистики критерия Аббе к отклонениям от нормального закона. В [4, 5] распределения статистики критерия Аббе были исследованы при справедливости H_0 и принадлежности ошибок измерений различным симметричным законам, была подтверждена устойчивость распределения статистики Аббе к нарушению предположения о нормальности $x_1, ..., x_n$ и исследована мощность критерия по отношению к различным конкурирующим гипотезам.

Критерий Аббе далеко не единственный критерий, ориентированный на проверку случайности и отсутствие тренда. Этим же целям служит критерий автокорреляции и целый ряд непараметрических критериев.

Проверяемая *гипотеза об отсутствии тренда в дисперсии* формулируется аналогичным образом (проверяется $H_0 : \sigma_i = \sigma$, $i = 1, 2, ..., n$, против $H_1 : |\sigma_{i+1} - \sigma_i| > 0$, $i = 1, 2, ..., n-1$).

При *проверке отсутствия сдвига в дисперсии* (в характеристиках рассеяния) предполагается, что наблюдаемая последовательность измерений x_1, ..., x_n имеет одно и то же среднее μ. Проверяется гипотеза H_0: $\sigma_1^2 = ... = \sigma_n^2 = \sigma_0^2$ (σ_0^2 неизвестно) против конкурирующей гипотезы

$$H_1 : \ \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2 = \sigma_0^2; \ \sigma_{k+1}^2 = \sigma_{k+2}^2 = ... = \sigma_n^2 = \sigma_0^2 + \delta \ (\delta > 0),$$

утверждающей, что значение дисперсии меняется в некоторой неизвестной точке, то есть k неизвестно ($1 \le k \le n-1$).

В настоящей работе продолжены исследования совокупности критериев, ориентированных на проверку гипотез отсутствия тренда в средних и дисперсиях [6]. Исследования опирались на развиваемые компьютерные технологии исследования статистических закономерностей, в основе которых лежит методика статистического моделирования и развиваемое программное обеспечение [7]. Объем выборок моделируемых распределений статистик, на которых строились выводы по работе, как правило, составлял величину от 10^4 до 10^6 в зависимости от требуемой точности.

2. Критерий автокорреляции

Если выборка значений   случайна, то значение каждого ее элемента не должно зависеть от величины предшествующего и последующего членов. Для проверки этой независимости используется статистика [8]

$$r_{1,n} = \frac{n\sum_{i=1}^{n-1} x_i x_{i+1} - \left(\sum_{i-1}^{n} x_i\right)^2 + n x_1 x_n}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i-1}^{n} x_i\right)^2}, \tag{1}$$

являющаяся коэффициентом корреляции первого порядка между элементами первичной выборки $(x_1,...,x_n)$ и элементами выборки, полученной из нее сдвигом на одну единицу $(x_2, x_3,..., x_n, x_1)$.

Величина $r_{1,n}$ распределена асимптотически нормально с математическим ожиданием и дисперсией

$$E[r_{1,n}] = -\frac{1}{n-1},$$

$$D[r_{1,n}] = \frac{n(n-3)}{(n+1)(n-1)^2}.$$

Поэтому в критерии используют нормализованную статистику

$$r_{1,n}^{*}=\frac{r_{1,n}-E[r_{1,n}]}{\sqrt{D[r_{1,n}]}}, \qquad (2)$$

которая подчиняется стандартному нормальному закону $N(0,1)$.

Асимптотическими результатами можно пользоваться и при достаточно малых объемах выборок. Исследование распределений статистики (2) методами статистического моделирования показало, что при $n>10$ распределение статистики достаточно хорошо согласуется со стандартным нормальным законом. В качестве примера в таблице 1 приведены результаты проверки согласия полученных в результате моделирования эмпирических распределений статистики (2) со стандартным нормальным законом для объемов выборок $n=10,15,20,25$.

Таблица 1. Проверка согласия эмпирического распределения статистики (2) со стандартным нормальным законом

n	Критерий согласия	S^*	$P\{S>S^*\}$
10	χ^2 Пирсона	98.325	9.3853e-18
	Колмогорова	2.3636	2.8081e-05
	ω^2 Крамера-Мизеса-Смирнова	1.2307	0.0007
	Ω^2 Андерсона-Дарлинга	8.2358	8.8646e-05
15	χ^2 Пирсона	60.920	3.0736e-10
	Колмогорова	1.6246	0.0102
	ω^2 Крамера-Мизеса-Смирнова	0.7182	0.0115
	Ω^2 Андерсона-Дарлинга	4.8892	0.0032
20	χ^2 Пирсона	20.862	0.0075
	Колмогорова	1.1880	0.1189
	ω^2 Крамера-Мизеса-Смирнова	0.3196	0.1192
	Ω^2 Андерсона-Дарлинга	2.0998	0.0810
25	χ^2 Пирсона	24.353	0.0020
	Колмогорова	1.1502	0.1418
	ω^2 Крамера-Мизеса-Смирнова	0.2994	0.1357
	Ω^2 Андерсона-Дарлинга	1.7711	0.1232

Проверка осуществлялась для выборок статистик объемом $N=10^4$ по критериям χ^2 Пирсона, Колмогорова, ω^2 Крамера-

Мизеса-Смирнова, Ω^2 Андерсона-Дарлинга. В случае критерия χ^2 Пирсона использовалось асимптотически оптимальное группирование с разбиением на 9 интервалов [9, 10]. По достигнутым уровням значимости, представляющим собой вероятность $P\{S > S^*\}$, где S – статистика соответствующего критерия согласия, а S^* – значение статистики критерия, вычисленное по анализируемой выборке, можно судить о сходимости распределений статистики (2) к стандартному нормальному закону.

Критерий автокорреляции относится к параметрическим критериям, в этой связи были проведены исследования распределений статистики критерия для случая принадлежности случайной величины различным законам, в том числе семейству с плотностью

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

с параметрами формы $\theta_2 = 0{,}2;\ 0{,}5;\ 1;\ 1{,}5;\ 2;\ 4;\ 8$. В этом случае вид закона менялся от близкого к распределению Коши, до близкого к равномерному закону. При $\theta_2 = 2$ выражение (3) дает плотность нормального закона распределения. Полученные в результате моделирования распределения статистики критерия автокорреляции в случае принадлежности случайной величины законам распределения семейства (3) при различных параметрах формы представлены на рис. 1.

Как следует из представленной на рисунке картины, в случае принадлежности выборок $x_1, \ldots, x_n$ достаточно широкому кругу законов распределение статистики критерия автокорреляции существенно не отличается от распределения, имеющего место в случае принадлежности $x_1, \ldots, x_n$ нормальному закону. Если закон, которому принадлежат случайные величины, симметричен и с не слишком тяжелыми хвостами, то распределение статистики не отличается значимо от “классического”.

При сильной асимметричности закона распределения случайных величин (например, в случае показательного закона) распределение статистики становится отличным от “классического”. В то же время, асимметричность закона влияет на распределение статистики менее значимо, чем “тяжесть” хвостов. В случае принадлежности выборок $x_1,...,x_n$ асимметричным законам экстремальных значений (минимального или максимального) распределения статистики практически не отличаются от “классического”.

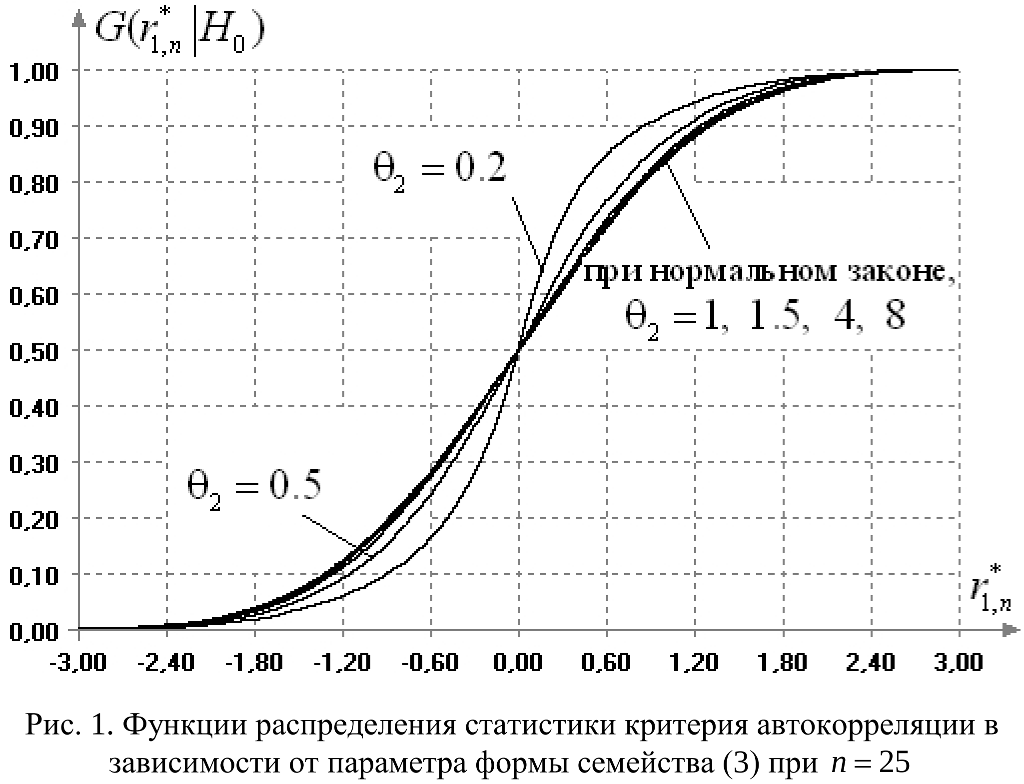

Рис. 1. Функции распределения статистики критерия автокорреляции в зависимости от параметра формы семейства (3) при $n = 25$

Отметим, что влияние закона распределения случайных величин на распределение статистики критерия автокорреляции такое же, как для критериев, связанных с проверкой гипотез о парной корреляции [11].

3. Критерий Фостера-Стюарта

Этот непараметрический критерий используется для проверки отсутствия тренда как в средних, так и в дисперсиях (в характеристиках рассеяния). Соответствующие статистики критерия имеют вид [12]:

$$S = \sum_{i=2}^{n} S_i \,, \tag{4}$$

$$d = \sum_{i=2}^{n} d_i \,, \tag{5}$$

где $d_i = u_i - l_i$, $S_i = u_i + l_i$; $u_i = 1$, если $x_i > x_{i-1}, x_{i-2}, ..., x_1$, иначе $u_i = 0$; $l_i = 1$, если $x_i < x_{i-1}, x_{i-2}, ..., x_1$, иначе $l_i = 0$.

Критерий со статистикой S используется для проверки наличия тренда в дисперсиях, а со статистикой d – для обнаружения тренда в средних. Очевидно, что $0 \le S \le n-1$ и $-(n-1) \le d \le n-1$.

При отсутствии тренда величины t и $\tilde{t}$:

$$t = \frac{d}{\hat{\sigma}_d} \,, \tag{6}$$

$$\tilde{t} = \frac{S - \mu}{\hat{\sigma}_S} \,, \tag{7}$$

где

$$\mu = 2\sum_{i=2}^{n} \frac{1}{i}, \;\; \hat{\sigma}_d = \sqrt{\mu} \approx \sqrt{2\ln n - 0{,}8456} \,,$$

$$\hat{\sigma}_S = \sqrt{\mu - 4\sum_{i=2}^{n} \frac{1}{i^2}} \approx \sqrt{2\ln n - 3{,}4253} \,,$$

приближенно описываются распределением Стьюдента с $\nu = n$ степенями свободы. Проверяемая гипотеза об отсутствии соответствующего тренда отклоняется при больших по модулю значениях статистик (6), (7).

На самом деле областью определения статистик t и $\tilde{t}$ является область дискретных значений. Исследование распределений статистик показало, что даже при достаточно больших объемах выборок ($n = 100, 200$) дискретные распределения статистик критерия *существенно отличаются от распределения Стьюдента* с n степенями свободы. На рис. 2 показаны функции распределения статистики t, а на рис. 3 – распределения статистики $\tilde{t}$, которые отличается еще большей дискретностью.

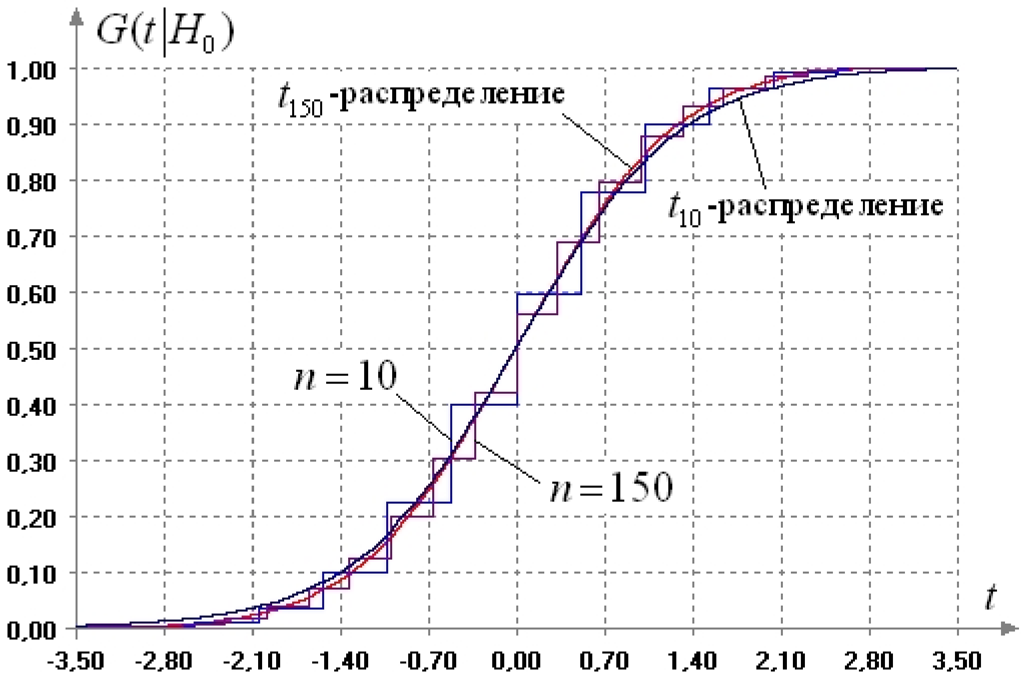

Рис. 2. Функции распределения статистики критерия Фостера-Стюарта для обнаружения тренда в средних при различных объемах выборок

Следовательно, использование распределений Стьюдента вместо истинных дискретных распределений статистик для определения достигнутого уровня значимости может приводить к ошибкам.

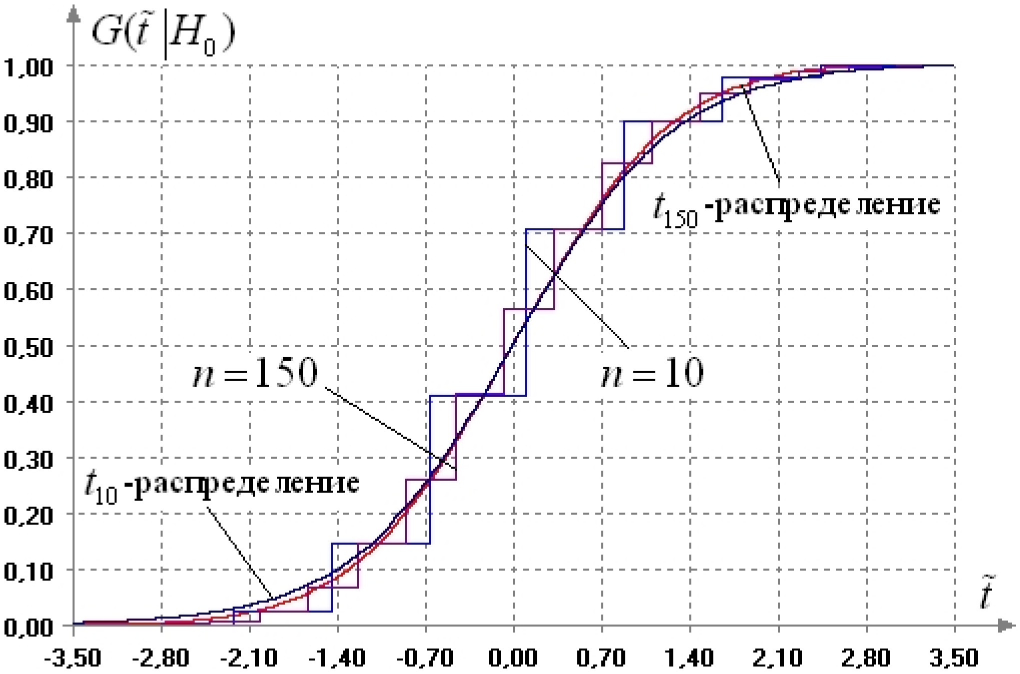

Рис. 3. Функции распределения статистики критерия Фостера-Стюарта для обнаружения тренда в дисперсиях при различных объемах выборок

4. Критерий Кокса-Стюарта

Как и предыдущий, данный непараметрический критерий используется для проверки последовательности измерений на предмет наличия тренда в среднем и в дисперсии.

В критерии проверки отсутствия тренда *в средних значениях* в выборке объема n используется статистика [13]

$$S_1 = \sum_{i=1}^{n/2} (n - 2i + 1) h_{i,n-i+1} , \tag{8}$$

где $h_{i,j} = 1$, если $x_i > x_j$, и $h_{i,j} = 0$, если $x_i \le x_j$ ($i < j$).

Нормализованная статистика

$$S_1^* = \frac{S_1 - E[S_1]}{\sqrt{D[S_1]}} , \tag{9}$$

где

$$E[S_1] = \frac{n^2}{8}, \quad D[S_1] = \frac{n(n^2 - 1)}{24} ,$$

при справедливости проверяемой гипотезы об отсутствии тренда приближенно подчиняется стандартному нормальному закону.

Распределение статистики S_1^* является дискретным и при малых n следует учитывать его отличие от стандартного нормального (см. рис. 4 при $n=10$). При объемах выборок порядка $n=100$ отличием распределения статистики от стандартного нормального закона при проведении анализа практически можно пренебречь и использовать для вычисления достигнутого уровня значимости функцию распределения стандартного нормального закона. Если реальные объемы выборок меньше, используя критерий и вычисляя достигнутый уровень значимости для полученного значения статистики S_1^* в соответствии со стандартным нормальным законом, надо иметь в виду введение возможной поправки на дискретность.

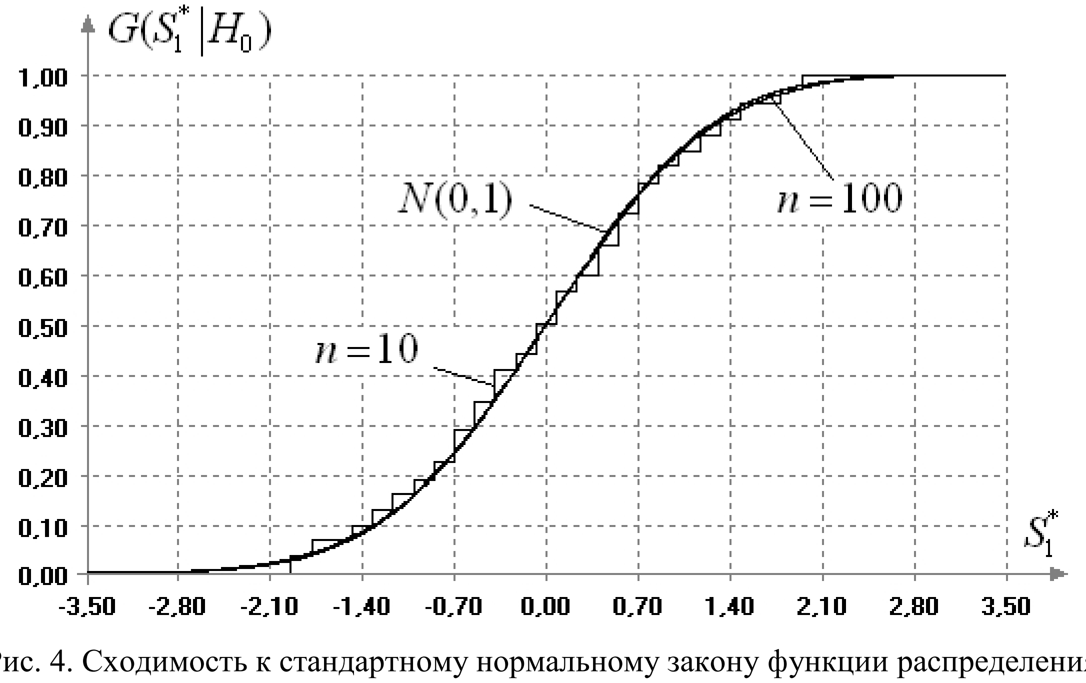

Рис. 4. Сходимость к стандартному нормальному закону функции распределения статистики (9) критерия Кокса-Стюарта для обнаружения тренда в средних

Критерий для проверки гипотезы отсутствия тренда *в дисперсии* (в характеристиках рассеяния) строится следующим образом. Выборка $x_1,...,x_n$ разбивается на $[n/k]$ подвыборок объемом k элементов $x_1,...,x_k$; $x_{k+1},...,x_{2k}$; $x_{2k+1},...,x_{3k}$; ...; $x_{n-k+1},...,x_n$ (если n не делится на k, отбрасывается необходимое число измерений в центре). Для каждой i-й подвыборки находится размах w_i $(1 \le i \le r)$ ($r = n/k$). Далее последовательность размахов w_i проверяется на наличие тренда критерием со статистикой S_1.

Величину k в [13] рекомендуется выбирать из следующих соотношений:

$$n \ge 90 \to k = 5; \quad 64 \le n < 90 \to k = 4;$$
$$48 \le n < 64 \to k = 3; \quad n < 48 \to k = 2.$$

Дискретность распределения статистики S_1^* при обнаружении тренда в дисперсии заметно выше (см. рис. 5), это естественно, так как анализируемая выборка размахов содержит лишь $[n/k]$ элементов.

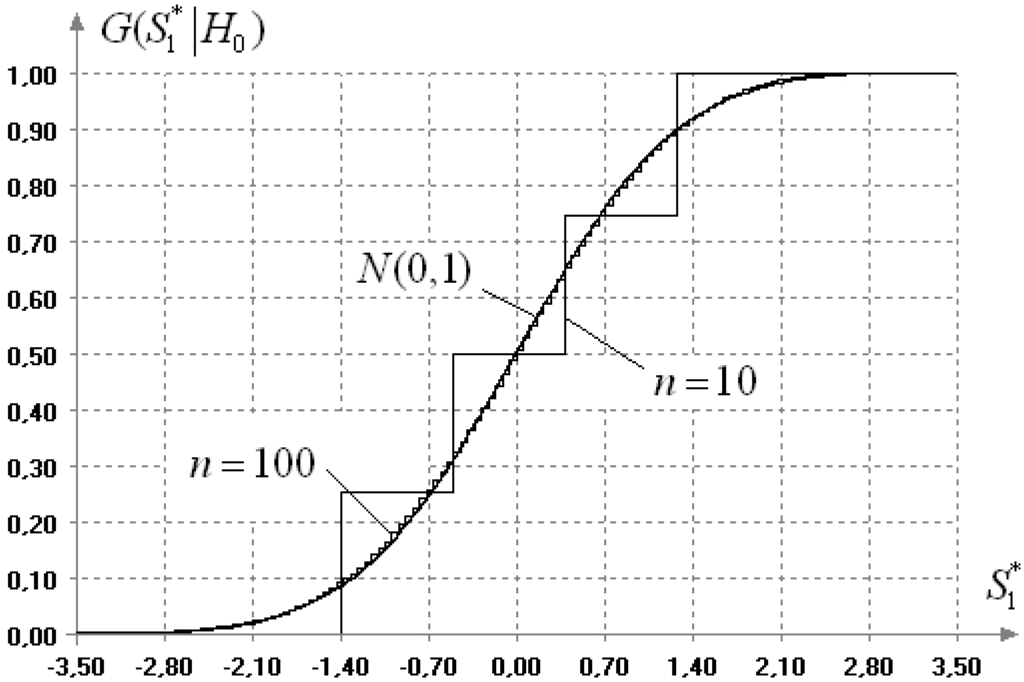

Рис. 5. Сходимость к стандартному нормальному закону функции распределения статистики критерия Кокса-Стюарта для обнаружения тренда в дисперсиях

5. Критерий Вальда-Вольфовитца

Статистика критерия Вальда-Вольфовитца, предложенная в [14], основана на коэффициенте сериальной корреляции и имеет вид:

$$R_1 = \sum_{i=1}^{n-1} x_i x_{i+1} + x_n x_1 \quad (10)$$

Величина R_1 распределена асимптотически нормально с математическим ожиданием и дисперсией

$$E[R_1] = (S_1^2 - S_2)/(n-1),$$

$$D[R_1] = \frac{S_2^2 - S_4}{n-1} + \frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(n-1)(n-2)} - \frac{(S_1^2 - S_2)^2}{(n-1)^2},$$

где $S_r = x_1^r + \ldots + x_n^r$.

Нормализованная статистика

$$R_1^* = \frac{R_1 - E[R_1]}{\sqrt{D[R_1]}} \quad (11)$$

подчиняется стандартному нормальному закону $N(0,1)$.

Исследование распределений статистики (11) в зависимости от объема выборки и для случая нарушения предположения о нормальности анализируемых выборок показало результаты, аналогичные для статистики (1): быструю сходимость распределения статистики (11) к стандартному нормальному закону и устойчивость этого распределения по отношению к отклонениям анализируемых данных от нормального закона.

В работе [14] была отмечена возможность построения непараметрического аналога сериального коэффициента корреляции. Такой аналог был предложен в работе [15]. Пусть R_i – ранг измерения x_i в упорядоченном по возрастанию ряду значений $x_1, x_2, \ldots, x_n$. Ранговый критерий сериальной корреляции Вальда-Вольфовитца имеет вид [15]:

$$R = \sum_{i=1}^{n-1} \left(R_i - \frac{n+1}{2} \right) \left(R_{i+1} - \frac{n+1}{2} \right). \quad (12)$$

Распределение этой статистики асимптотически нормально со средним и дисперсией

$$E[R]=0,$$

$$D[R]=\frac{n^2(n+1)(n-3)(5n+6)}{720}.$$

Гипотеза случайности (об отсутствии тренда) отклоняется при больших по модулю значениях статистики

$$R^*=\frac{R}{\sqrt{D[R]}}. \tag{13}$$

Результаты моделирования показали, что распределение статистики (13) критерия *смещено по отношению к предельному* и очень медленно сходится к стандартному нормальному закону (см. рис. 6). Дискретностью же распределения статистики практически можно пренебречь.

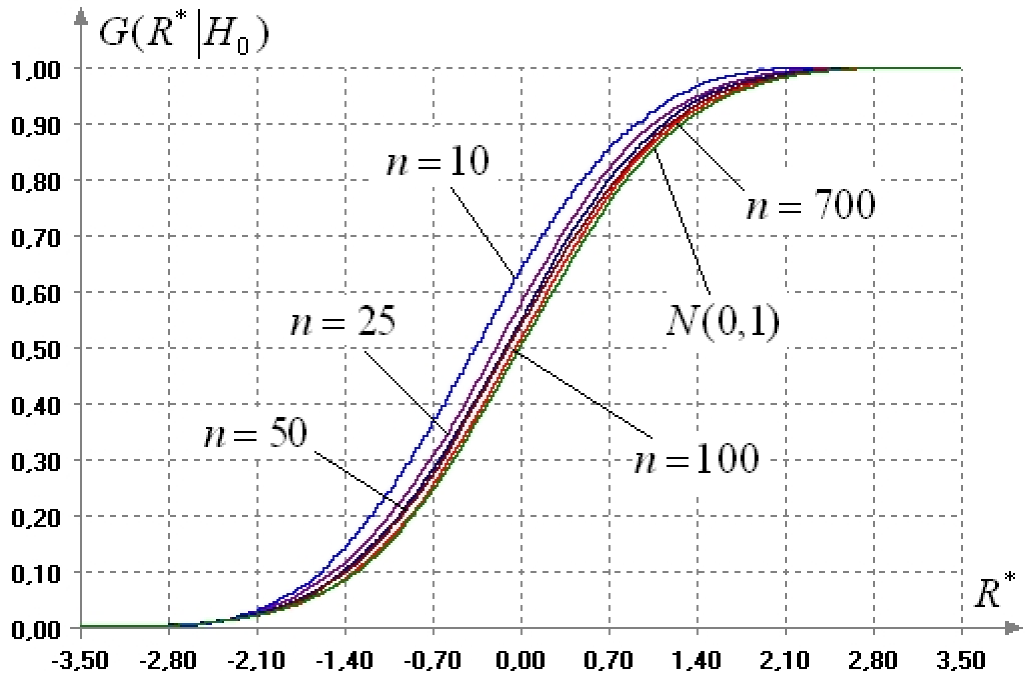

Рис. 6. Сходимость к стандартному нормальному закону распределения статистики (13) критерия Вальда-Вольфовитца

6. Критерий Бартелса

Пусть в последовательности n измерений R_i – ранг i-го наблюдения x_i. Бартелсом [16] был предложен ранговый критерий случайности ряда, основанный на статистике

$$B = \frac{\sum_{i=1}^{n-1}(R_i - R_{i+1})^2}{\sum_{i=1}^{n}(R_i - \bar{R})^2}. \tag{14}$$

Гипотеза о случайности (об отсутствии тренда) отклоняется при больших по модулю значениях статистики

$$B^* = \frac{B - E[B]}{\sqrt{D[B]}} = \frac{B-2}{2\sqrt{5/(5n+7)}}, \tag{15}$$

которая при отсутствии тренда приближенно подчиняется стандартному нормальному закону.

Исследования показали, что распределение статистики достаточно быстро сходится к стандартному нормальному (см. рис. 7).

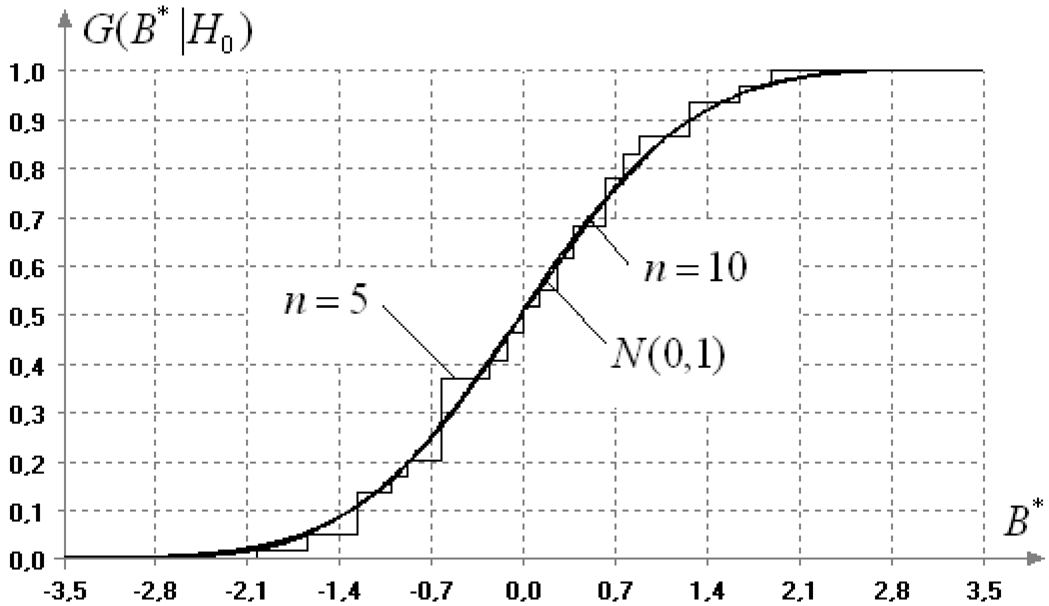

Рис. 7. Сходимость к стандартному нормальному закону распределения статистики критерия Бартелса

7. Критерий Хсу

Критерий обнаружения сдвига дисперсий основан на статистике [17]

$$H = \frac{\sum_{i=1}^{n}(i-1)(x_i - m_x)^2}{(n-1)\sum_{i=1}^{n}(x_i - m_x)^2}, \quad 0 \le H \le 1. \tag{16}$$

Обычно критерий используется в стандартизированной форме

$$H^* = \frac{H - 0,5}{\sqrt{D[H]}}, \tag{17}$$

где $D[H] = \dfrac{n+1}{6(n-1)(n+2)}$. Статистика (17) при справедливости гипотезы об отсутствии сдвига дисперсии подчиняется стандартному нормальному закону.

Результаты моделирования показали, что при $n > 30$ распределение статистики достаточно хорошо согласуется со стандартным нормальным законом (см. рис. 8).

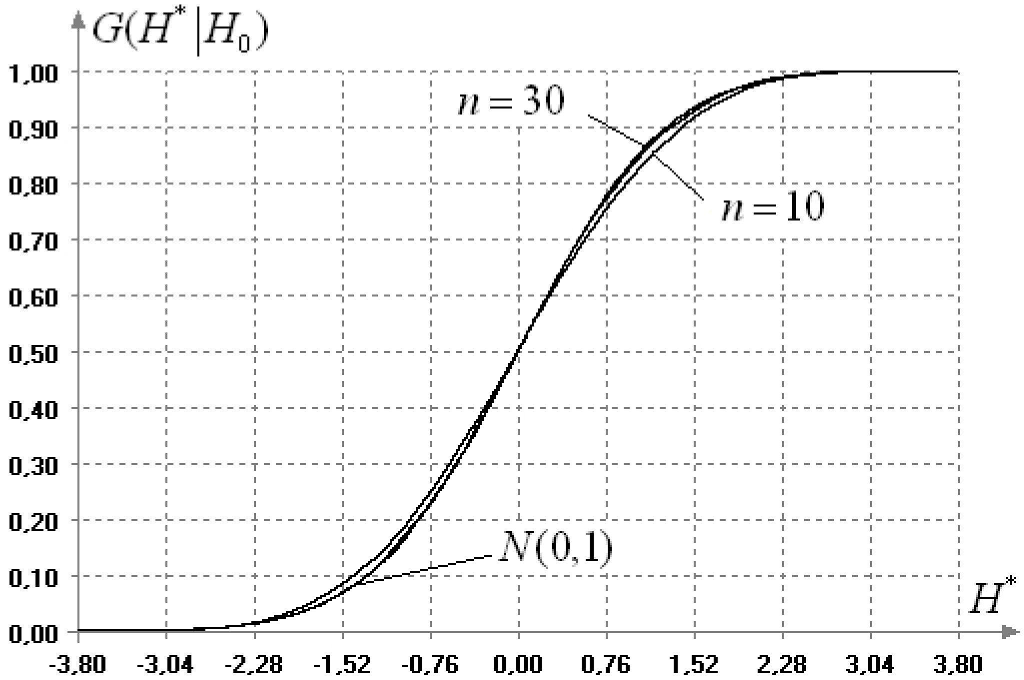

Рис. 8. Сходимость распределения статистики (17) критерия Хсу к стандартному нормальному закону

Критерий Хсу относится к параметрическим критериям. Поэтому, как и в случае любого параметрического критерия, связанного с проверкой гипотез о дисперсиях, распределения его статистики зависят от закона, которому принадлежат анализируемые случайные величины. Полученные в результате моделирования распределения статистики критерия (17) в случае принадлежности случайных величин (ошибок измерений) законам распределения семейства (3) при различных параметрах формы представлены на рис. 9.

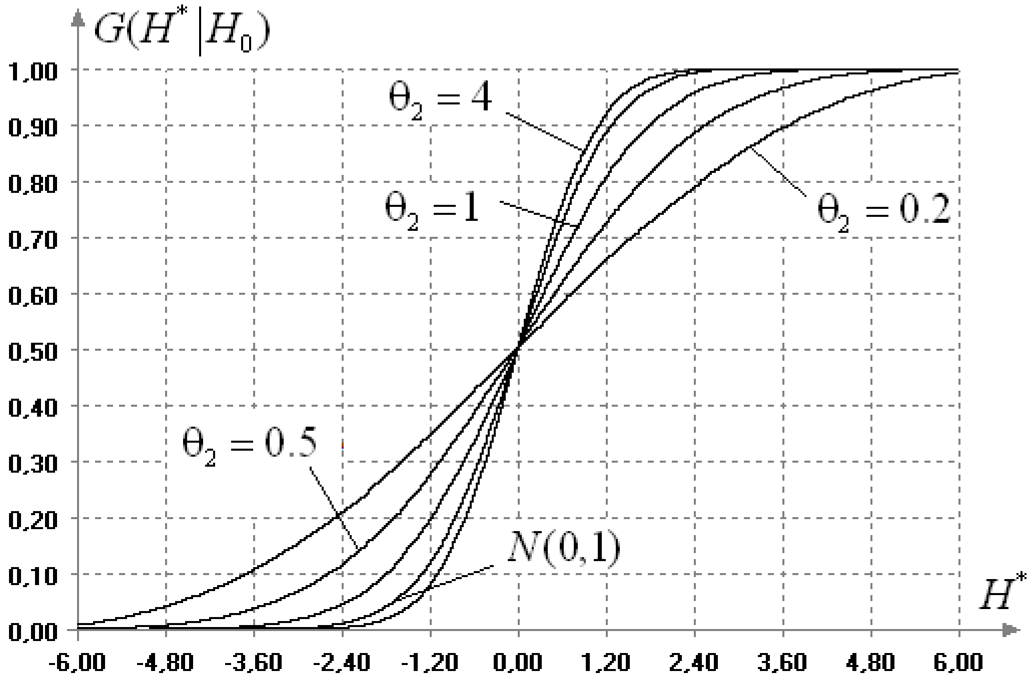

Рис. 9. Распределения статистики (17) в случае принадлежности случайных величин законам семейства (3) при различных значениях параметра формы $n = 30$

Как можно видеть, распределения статистики (17) сильно зависит от закона распределения, которому принадлежат случайные величины. При этом наибольшее отклонение от стандартного нормального закона оказывается в случае принадлежности случайных величин законам с тяжелыми хвостами. Существенно влияет на распределение статистики и асимметричность закона.

В [17] описан критерий, позволяющий определить точку изменения дисперсии в случае принадлежности ошибок измерений нормальному закону. Его статистика строится следующим образом. Пусть для $k = 1, 2, ..., n-1$

$$w_k = \sum_{i=1}^{k} (x_i - m_x)^2 ,$$

$$W_k = \frac{w_n - w_k}{w_k} \frac{k}{n-k}.$$

где k соответствует *искомой точке* изменения дисперсии.

Далее из уравнения $F_{\gamma_k}(n-k,k) = W_k$, где $F_\gamma(f_1, f_2)$ – γ-квантиль F - распределения Фишера с f_1 и f_2 степенями свободы, находятся оценки γ_k, которые должны подчиняться равномерному закону. Статистика G-критерия имеет вид

$$G = \frac{1}{n-1} \sum_{k=1}^{n-1} \gamma_k, \ 0 \le G \le 1. \tag{18}$$

Гипотеза об отсутствии изменения дисперсии отклоняется с уровнем значимости α, если $G > G_{1-\alpha}$. В этом случае значение k, которому соответствует максимальная величина $|\gamma_k - 1/2|$, дает оценку искомой точке изменения значения дисперсии в наблюдаемом ряду. Критические значения $G_{1-\alpha}$ можно найти в [1].

Как и критерий со статистиками (16)–(17) данный критерий применим только при извлечении выборок из нормальной генеральной совокупности.

8. Ранговый критерий обнаружения сдвига дисперсии

Ранговый критерий обнаружения сдвига дисперсии (характеристики рассеяния) в неизвестной точке основан на использовании семейства ранговых статистик вида [18]

$$S_R = \sum_{i=1}^{n} i a_n(R_i), \tag{19}$$

где R_i - ранги выборочных значений в упорядоченном ряду измерений.

Метки критерия a_n могут быть различными, например:

– метки Клотца $a_{1n}(i)=U^2_{i/(n+1)}$, где U_γ – γ-квантиль стандартного нормального закона;

– метки Сэвиджа $a_{2n}(i)=\sum_{j=1}^{i}\frac{1}{n-j+1}$.

Соответствующие (19) статистики обозначим $S_{R,j}=\sum_{i=1}^{n} i a_{jn}(R_i)$, $j=1,2$. При отсутствии сдвига дисперсии в ряду измерений $S_{R,j}$-статистики свободны от распределения и симметричны относительно $E[S_{R,j}]=\frac{n+1}{2}\sum_{i=1}^{n} a_{jn}(i)$.

При $n>15$ и справедливости гипотезы об отсутствии сдвига в характеристике рассеяния случайной величины статистики

$$S^*_{R,j}=\frac{S_{R,j}-E[S_{R,j}]}{\sqrt{D[S_{R,j}]}}, \tag{20}$$

где

$$E[S_{R,1}]=\frac{n+1}{2}\sum_{i=1}^{n}U^2_{i/(n+1)}, \quad E[S_{R,2}]=\frac{n(n+1)}{2};$$

$$D[S_{R,1}]=\frac{n(n+1)}{12}\sum_{i=1}^{n}U^4_{i/(n+1)}-\frac{1}{3n+3}\left[E[S_{R,1}]\right]^2;$$

$$D[S_{R,2}]=\frac{n(n+1)}{12}\left(n-\sum_{j=1}^{n}\frac{1}{j}\right),$$

приближенно подчиняются стандартному нормальному закону (см. рис. 10 и 11).

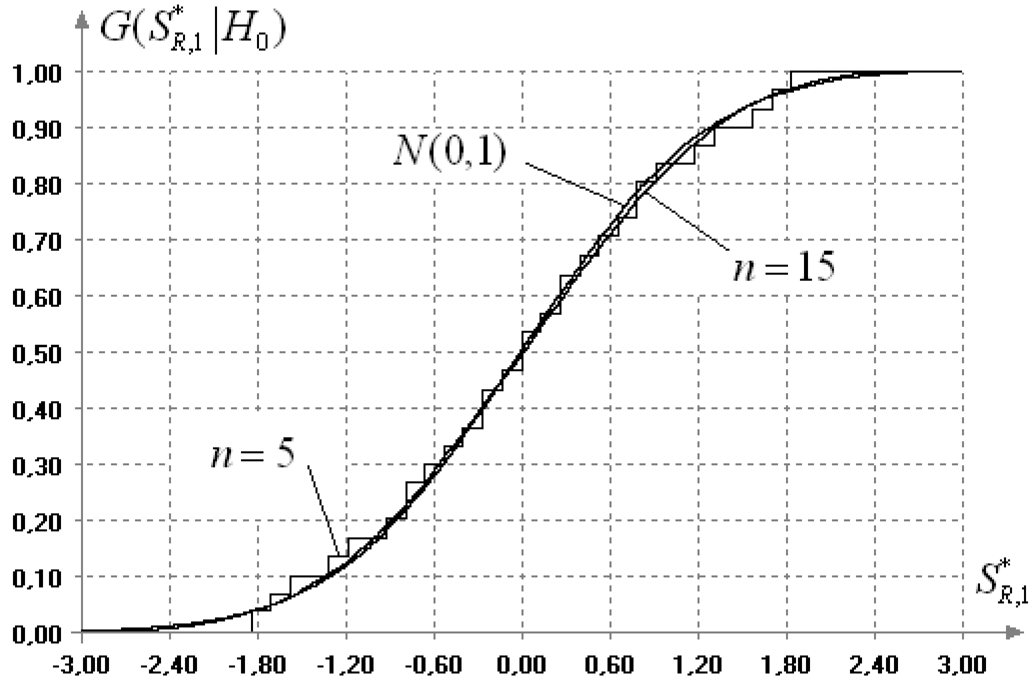

Рис. 10. Сходимость к стандартному нормальному закону распределения статистики $S_{R,1}^{*}$ рангового критерия (с метками Клотца) обнаружения сдвига дисперсии в неизвестной точке

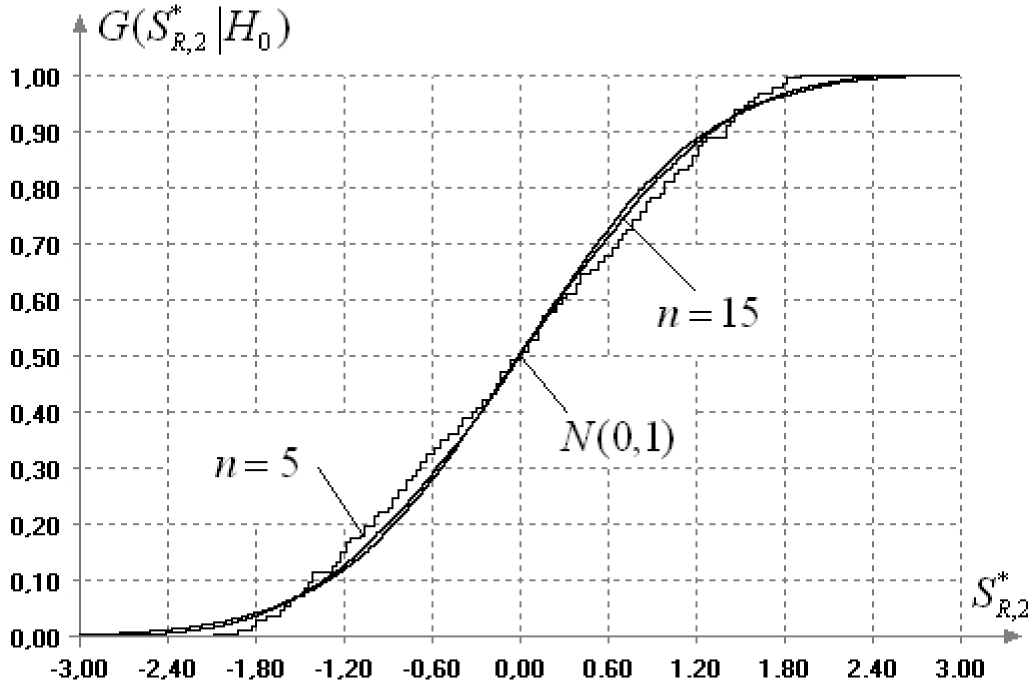

Рис. 11. Сходимость к стандартному нормальному закону распределения статистики $S_{R,2}^{*}$ рангового критерия (с метками Сэвиджа) обнаружения сдвига дисперсии в неизвестной точке

9. Сравнительный анализ мощности критериев

Анализ мощности критериев проводился для ситуации принадлежности наблюдаемых случайных величин нормальному закону. Проверяемой гипотезе H_0 соответствует выполнение предположения о независимости наблюдаемых случайных величин (отсутствие тренда). В качестве конкурирующих гипотез рассматривались различные ситуации при наличии тренда в средних или в дисперсии.

При исследовании мощности критериев об отсутствии тренда в средних рассматривались модели задания линейного, периодического и смешанного тренда.

В случае наличия линейного тренда случайные величины моделировались в соответствии с

$$x_i = a \cdot t + \xi_i, \tag{21}$$

где ξ_i представляют собой независимые случайные величины, распределённые в соответствии с заданным законом (например, по стандартному нормальному закону), $t \in [0,1]$. Справедливой проверяемой гипотезе H_0 соответствует значение параметра $a = 0$.

Величины x_i (21) вычислялись в соответствии с выражением $x_i = a \cdot (i-1)\Delta t + \xi_i$, где шаг Δt определялся как $\Delta t = 1/n$ в зависимости от объема выборки n. Псевдослучайные величины ξ_i генерировались в соответствии со стандартным нормальным законом (возможно и по любому другому). Исследовалась мощность критерия относительно конкурирующих гипотез с линейным трендом, задаваемым параметром $a = 0.5$ и $a = 4$. Соответствующие конкурирующие гипотезы обозначены в дальнейшем как H_1, H_2. Примеры временных рядов при тренде с параметром $a = 0.5$ и $a = 4$ при объеме выборки $n = 100$ приведены на рис.12а и 12б.

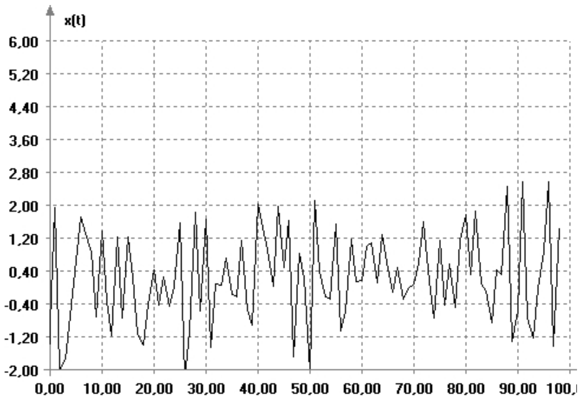

Рис. 12а. Линейный тренд при $a = 0.5$

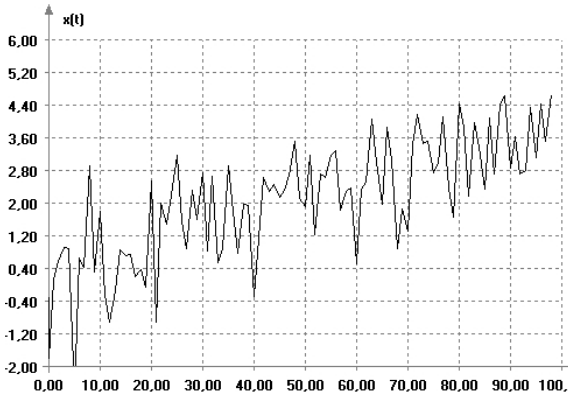

Рис. 12б. Линейный тренд при $a = 4$

В случае периодического тренда случайные величины моделировались в соответствии с соотношением

$$x_i = a \cdot \sin(2\pi t) + \xi_i, \tag{22}$$

а в случае смешанного – в соответствии с

$$x_i = a \cdot t + a \cdot \sin(2\pi t) + \xi_i. \tag{23}$$

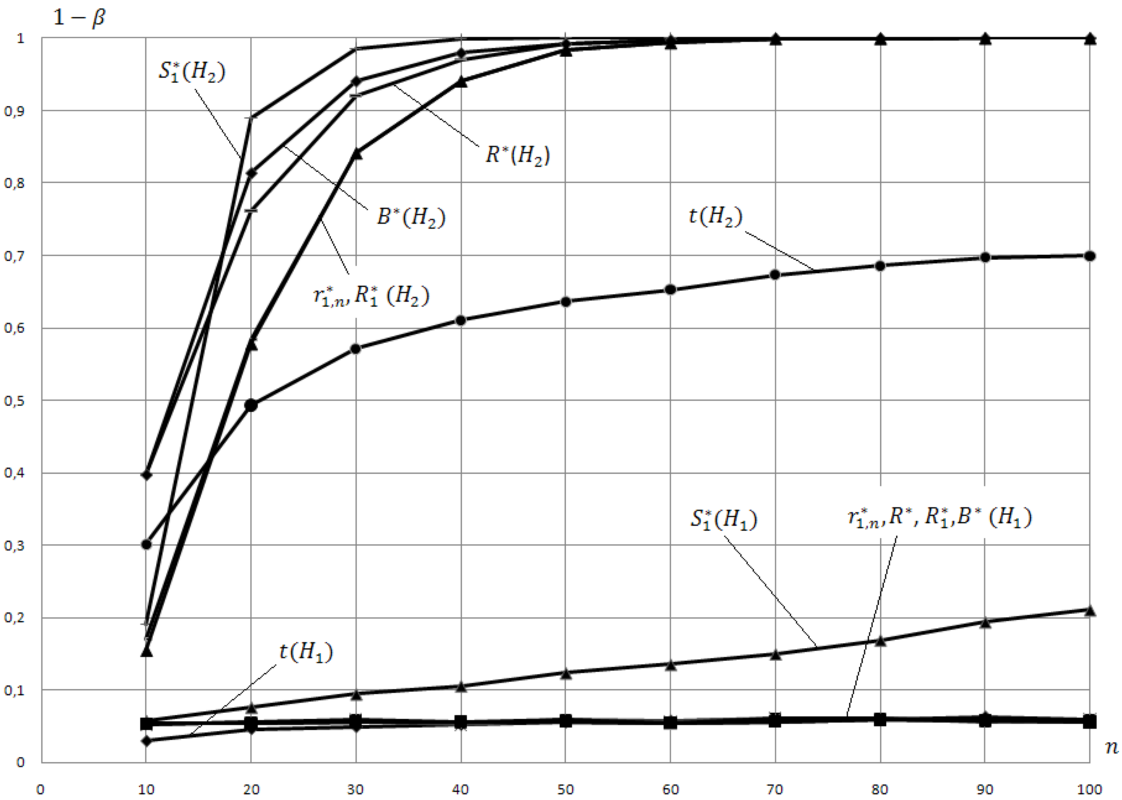

Рис.13. Мощность критериев тренда и случайности относительно конкурирующих гипотез с линейным трендом

На рис. 13 для уровня значимости $\alpha = 0.05$ приведены значения мощности $1-\beta$ рассматриваемых в данной работе критериев тренда и случайности относительно конкурирующих гипотез H_1 (при $a = 0.5$) и H_2 (при $a = 4$) с линейным трендом (21) в зависимости от объема выборки n. Анализ мощности критериев показал, что критерий Фостера-Стюарта (t) значительно уступает критериям Кокса-Стюарта (S_1^*), критерию автокорреляции ($r_{1,n}^*$), Вальда-Вольфовитца (R^*) и Бартелса (B^*).

Рассматриваемые критерии проверки гипотез об отсутствии тренда в средних, включая критерий Аббе, можно упорядочить по мощности (относительно линейного тренда) следующим образом: Аббе (S_A) $\succ$ Кокса-Стюарта (S_1^*) $\succ$ Бартелса (B^*) $\succ$ ранговой сериальной корреляции Вальда-Вольфовитца (R^*) $\succ$ автокорреляции ($r_{1,n}^*$), сериальной корреляции Вальда-Вольфовитца (R_1^*) $\succ$ Фостера-Стюарта (t). Если рассматривать в качестве конкурирующих гипотез наличие произвольного тренда, то целесообразно рекомендовать использование критериев Аббе, Бартелса, Вальда-Вольфовитца (ранговый и неранговый вариант) и автокорреляции.

Критерий Фостера-Стюарта показал наименьшую мощность. В определенной степени данный результат объясняется сильной дискретностью распределения статистики критерия Фостера-Стюарта. Вследствие дискретности действительный уровень значимости существенно отличается от задаваемого $\alpha = 0.05$ и получаемые оценки мощности оказываются заниженными.

При исследовании мощности критериев обнаружения тренда в характеристиках рассеяния в качестве конкурирующих гипотез рассматривалась ситуация с линейным трендом в дисперсии. Исследовалась мощность критериев Фостера-Стюарта ($\tilde{t}$), Кокса-Стюарта (S_1^*), автокорреляции ($r_{1,n}^*$), ранговой сериальной корреляции

Вальда-Волфовитца (R^*), сериальной корреляции Вальда-Вольфовитца (R_1^*) и Бартелса (B^*). Оказалось, что критерии Фостера-Стюарта и Кокса-Стюарта, специально построенные для выявления тренда в дисперсии, весьма значительно превосходят по мощности остальные критерии. Для обнаружения тренда в дисперсии можно рекомендовать к применению критерии Фостера-Стюарта и Кокса-Стюарта, расположив их по предпочтению: Фостера-Стюарта ($\tilde{t}$) $\succ$ Кокса-Стюарта (S_2^*).

При исследовании мощности критериев сдвига дисперсии в неизвестной точке в качестве конкурирующих гипотез рассматривалось наличие скачкообразного сдвига в величине дисперсии. Исследовалась мощность ранговых критериев обнаружения сдвига дисперсии в неизвестной точке с метками Клотца ($S_{R,1}^*$), с метками Сэвиджа ($S_{R,2}^*$), критерия Хсу (H^*) и G-критерия. В случае принадлежности случайных величин нормальному закону критерии сдвига дисперсии в неизвестной точке можно упорядочить следующим образом: критерий Хсу (H^*) $\succ$ критерий сдвига дисперсии с метками Клотца ($S_{R,1}^*$) $\succ$ G-критерий (G) $\succ$ Критерий сдвига дисперсии с метками Сэвиджа ($S_{R,2}^*$).

10. Заключение

Таким образом, на основании проведенных исследований, можно констатировать, что применение параметрических критериев обнаружения тренда в средних (критерий автокорреляции, Аббе) будет корректным и в тех случаях, когда мы имеем дело с законом, существенно отличающимся от нормального (но симметричным и без “тяжелых” хвостов). Это общая тенденция устойчивости распределений параметрических критериев так или иначе связанных с проверкой гипотез о математических ожиданиях [19, 20, 21] или проверкой гипотез о равенстве нулю коэффициентов парной, частной и множественной корреляции [11]. Однако параметрические критерии обнаружения тренда в средних лишь не многим превосходят по мощности непараметрические.

Параметрические критерии обнаружения сдвига в дисперсии (критерий Хсу) мощнее непараметрических, но очень чувствительны к нарушению предположений о нормальности случайных величин (как и любые критерии, связанные с проверкой гипотез о дисперсиях [19, 22, 23, 24, 25, 26, 27]).

Использование критериев Фостера-Стюарта затруднено дискретностью распределений статистик и плохой сходимостью к соответствующим *t*-распределениям Стьюдента.

Распределение статистики рангового критерия сериальной корреляции Вальда-Вольфовитца смещено относительно стандартного нормального закона, к которому очень медленно сходится. При использовании нерангового критерия такой проблемы не возникает.

Полученные оценки мощности критериев позволяют судить о способности критериев обнаруживать наличие линейного и нелинейного тренда в среднем или в характеристиках рассеяния.

Работа выполнена при поддержке РФФИ (проект № 09-01-00056а) и Федеральной целевой программы Минобрнауки РФ “Научные и научно-педагогические кадры инновационной России”.

Список литературы

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

2. Струнов В.И.. О применении критерия Аббе для анализа независимости рядов измерений, характеризующихся отличными от

нормального законами распределения // Измерительная техника. 2006. № 8. – С. 13-17.

3. Strunov V.I. Applying the Abbé test to the independence of measurement series with distributions deviating from normal // Measurement Technique. Vol. 49, No. 8, 2008. – P.962-969.

4. Лемешко С.Б. Критерий независимости Аббе при нарушении предположений нормальности // Измерительная техника. 2006. № 10. – С.9-14.

5. Lemeshko S.B. The Abbé independence test with deviations from normality // Measurement Technique. Vol. 49, No. 10, 2006. – P.962-969.

6. Беркович А.С., Лемешко Б.Ю., Щеглов А.Е. Исследование распределений статистик критериев тренда и случайности // Материалы X международной конференции “Актуальные проблемы электронного приборостроения” АПЭП-2010. Новосибирск, 2010. – Т.6. – С.13-17.

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

8. Knoke J.D. Testing for randomness against autocorrelation: The parametric case // Biometrica. 1975. – V.62. – P.571-575.

9. Денисов В.И., Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть I. Критерии типа χ^2. – Новосибирск: Изд-во НГТУ, 1998. – С. 126.

10. 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.

11. Лемешко Б.Ю., Помадин С.С. Исследование распределений статистик корреляционного анализа при отклонении многомерного закона от нормального // Тр. V международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2000. Новосибирск, 2000. - Т. 7. - С. 184-187.

12. Foster F.G., Stuart A. Distribution-free tests in time series dated on the breaking of records // JRSS. 1954. – V. B16, №1. – P.1-22.

13. Cox D.R., Stuart A. Quick sign tests for trend in location and dispersion // Biometrica. 1955. – V.42. – P.80-95.

14. Wald A., Wolfowitz J. An exact test for randomness in the non-parametric case based on serial correlation // AMS. 1943. V. 14. P. 378-388.

15. Dufour J.-M., Roy R. Some robust exact results on sample autocorrelations and tests of randomness// J. of Econometrics. 1985. V. 29, P. 257-273.

16. Bartels R. The rank version of von Neumann's ratio test for randomness // JASA. 1982. V. 77, №377. P. 40-46.

17. Hsu D.A. Test for variance shift at an unknown time point // Appl. Statist., 1977. – V.26, № 3. – P.279-284.

18. Hsieh H.K. Nonparametric tests for scale shift at a unknown time point // Commun. Stat. – Theor. Meth., 1984. – V.13. № 11. – P.1335-1355.

19. Лемешко Б.Ю., Помадин С.С. Проверка гипотез о математических ожиданиях и дисперсиях в задачах метрологии и контроля качества при вероятностных законах, отличающихся от нормального // Метрология. 2004. – № 3.- С.3-15.

20. Лемешко Б.Ю., Лемешко С.Б. Об устойчивости и мощности критериев проверки однородности средних // Измерительная техника. 2008. № 9. – С.23-28.

21. 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.

22. Лемешко Б.Ю., Лемешко С.Б., Горбунова А.А. О применении и мощности критериев проверки однородности дисперсий. Ч. I. Параметрические критерии // Измерительная техника. 2010. № 3. – С.10-16.

23. Lemeshko B.Yu., Lemeshko S.B., and A. A. Gorbunova. Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria // Measurement Techniques, Vol. 53, No. 3, 2010. – P.237-246.

24. Лемешко Б.Ю., Лемешко С.Б., Горбунова А.А. О применении и мощности критериев проверки однородности дисперсий. Ч. II. Непраметрические критерии // Измерительная техника. 2010. № 5. – С.11-18.

25. Lemeshko B.Yu., Lemeshko S.B., and A. A. Gorbunova. Application and power of criteria for testing the homogeneity of variances. Part II. Nonparametric criteria // Measurement Techniques, Vol. 53, No. 5, 2010. – P.476-486.

26. 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.

27. Lemeshko B., Mirkin E. Bartlett and Cochran tests in measurements with probability laws different from normal // Measurement Techniques, 2004, Vol. 47, № 10. – P. 960-968.

UDC 519.24

Application of tests for trend detection and checking for randomness

Lemeshko B.Yu., Komissarova A.S., Tsheglov A.E.
Novosibirsk state technical university, Novosibirsk, Russia,
e-mail: Lemeshko@fpm.ami.nstu.ru

Abstract. An analysis of asymptotic distributions of various tests for trend detection in mean and variance is conducted. Disadvantages of applying several tests under analysis are registered. The results of a comparative test power analysis are presented.

Keywords: test for trend detection, autocorrelation test, Foster-Stuart test, Cox-Stuart test, Wald-Wolfowitz test, Bartels test, Hsu test, rank tests for scale shift, test power