Методы статистического анализа в задачах принятия решений

Введение

Существующий аппарат прикладной математической статистики (методы построения различных вероятностных моделей и различные критерии проверки адекватности этих моделей) широко **востребован** и используется **в задачах, связанных с обоснованием принимаемых решений**.

Именно на эти методы опираются **при поиске** условных **вероятностей** результатов, необходимых **для принятия решений в условиях риска** или **в условиях неопределённости**. Оптимальность выбранной стратегии в соответствующих задачах напрямую зависит от адекватности построенных моделй законов распределения вероятностей и точности оценок вероятностей результатов и вероятностей состояний среды.

Построение моделей законов (идентификация закона) распределения вероятностей для наблюдаемых случайных величин, **проверка адекватности этих моделей** с использованием критериев проверки статистических гпотез и статистические выводы, формируемые в результате проверки этих гипотез, **представляют собой одну из основ**, на которых базируется принятие решений для выбора стратегии или стратегий.

Существующий аппарат статистического анализа данных включает широкое множество методов и критериев проверки гипотез. Это не означает, что он позволяет решить любую задачу. Аппарат и программное обеспечение развиваются.

В то же время приходится констатировать, что в задачах принятия решений, опирающихся на статистические и экспериментальные данные (результаты измерений, контроля и испытаний), **из обширного множества критериев**, предназначенных, например, для проверки различных статистических гипотез, **реально используется** лишь очень **ограниченный круг**.

Ряд перспективных критериев (и методов), которые “буквально напрашиваются” для применения в приложениях, по ряду причин практически не используются.

Зачастую критерии используются в условиях нарушения стандартных предположений, обуславливающих возможность их применения, что приводит к некорректности статистических выводов. (Нарушение предположения о виде закона. Выбор другого метода оценивания или вида оценки. Не учитывается, что проверяется сложная гипотеза, а не простая. Не учитывается влияние ошибок округления. И т.п.).

Причинами такого состояния в случае критериев проверки гипотез являются следующие:

1. Отсутствует доступная информация о существовании соответствующего математического аппарата (о соответствующих критериях).

2. Реальные свойства критериев при ограниченных объёмах анализируемых выборок могут существенно отличаться от асимптотических свойств.

3. Реальные свойства критериев могут существенно изменяться вследствие ошибок округления данных в анализируемых выборках.

4. Возможность применения многих критериев ограничена отсутствием информации о распределениях статистик этих критериев при справедливости проверяемой гипотезы. В связи с чем приходится опираться на ограниченные таблицы критических значений.

5. Отсутствует доступная информация о том, что происходит с критериями (с распределениями их статистик) в условиях нарушения стандартных предположений.

6. Отсутствует информация о достоинствах и недостатках отдельных критериев, о преимуществе в мощности тех или иных критериев из группы критериев, ориентированных на проверку одной и той же гипотезы, что позволило бы выбрать наиболее предпочтительный критерий.

7. Отсутствие программного обеспечения для использования соответствующих критериев, гарантируещего корректность выводов, как в условиях стандартных предположений, так и в условиях конкретных приложений.

При идентификации закона распределения вероятностей, соответствующего наблюдаемым данным, можно условно выделить 2 этапа.

1. Построение по экспериментальным данным некоторой модели закона распределения.

2. Проверка адекватности построенной модели экспериментальным данным.

На первом этапе выбирается некоторая модель закона, для которой по имеющейся выборке оцениваются параметры.

На втором, с использованием некоторого критерия проверяется гипотеза о соответствии модели экспериментальным данным (проверяется гипотеза о «согласии»).

Такая процедура может повторяться для различных видов параметрических моделей законов, в результат чего, выбирается наиболее предпочтительная модель.

Заметим, что при использовании различных методов оценивания получаются несколько отличающиеся модели. Более того, наличие аномальных измерений (и нарушение некоторых предположений) может приводить к существенным различиям в моделях.

Необходимо отметить, что выбранный метод оценивания отражается и на свойствах критериев, используемых при проверке адекватности моделей.

Методы оценивания параметров

1. Метод максимального правдоподобия

Оценки максимального правдоподобия вычисляются в результате максимизации по θ функции правдоподобия

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) \tag{1.1}$$

или ее логарифма

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i, \theta). \tag{1.2}$$

В случае скалярного параметра ОМП определяются как решение уравнения, а в случае векторного параметра – как решение системы уравнений правдоподобия вида

$$\frac{\partial \ln L(\theta)}{\partial \theta_l} = \sum_{i=1}^{n} \frac{\partial \ln f(x_i, \theta)}{\partial \theta_l} = 0, \quad l = \overline{1, m}, \tag{1.3}$$

где m – размерность вектора параметров θ. Чаще всего решение (1.3) может быть найдено только численными методами.

В отличие от других метод максимального правдоподобия позволяет определять оценки параметров по негруппированным, частично группированным, группированным и цензурированным данным.

Выборка называется *частично группированной*, если имеющаяся в нашем распоряжении информация связана с множеством непересекающихся интервалов, которые делят область определения случайной величины на k непересекающихся интервалов граничными точками

$$x_{(0)} < x_{(1)} < \ldots < x_{(k-1)} < x_{(k)},$$

где $x_{(0)}$ – нижняя грань области определения случайной величины X; $x_{(k)}$ – верхняя грань области определения случайной величины X, так что каждый интервал принадлежит к одному из двух типов:

а) i-й интервал принадлежит к первому типу, если число n_i известно, но индивидуальные значения x_{ij}, $j=\overline{1,n_i}$ неизвестны;

б) i-й интервал принадлежит ко второму типу, если известно не только число n_i, но и все индивидуальные значения x_{ij}, $j=\overline{1,n_i}$.

Суммирование по интервалам первого и второго типов (аналогично умножение) обозначим соответственно как $\left(\sum\limits_{(1)}\right)$ и $\left(\sum\limits_{(2)}\right)$.

Оценкой максимального правдоподобия неизвестного параметра по частично группированным наблюдениям называется такое значение параметра, при котором функция правдоподобия

$$L(\theta) = \prod_{(1)} P_i^{n_i}(\theta) \prod_{(2)} \prod_{j=1}^{n_i} f(x_{ij}, \theta), \tag{1.4}$$

где $f(x, \theta)$ – функция плотности случайной величины; $P_i(\theta) = = \int\limits_{x_{(i-1)}}^{x_{(i)}} f(x, \theta) dx$ – вероятность попадания наблюдения в i-й интервал значений, достигает максимума на множестве возможных значений параметра.

При вычислении ОМП максимизируют (1.4) или решают систему уравнений правдоподобия

$$\sum_{(1)} n_i \frac{\partial \ln P_i(\theta)}{\partial \theta_l} + \sum_{(2)} \sum_{j=1}^{n_i} \frac{\partial \ln f(x_{ij}, \theta)}{\partial \theta_l} = 0, \ l = \overline{1, m}, \tag{1.5}$$

где m – размерность вектора параметров $\theta = (\theta_1, \theta_2, \ldots, \theta_m)^{\mathrm{T}}$. В случае частично группированных данных система (1.5), за редким исключением, решается только численно.

2. Методы минимального расстояния

При вычислении *MD*-оценок (оценок минимального расстояния) по θ **минимизируется некоторая мера близости** (расстояние) $\rho(F(x,\theta), F_n(x))$ между теоретическим и эмпирическим распределениями. *MD*-оценки находятся в процессе решения задачи

$$\hat{\theta} = \arg\min_{\theta} \rho(F(x,\theta), F_n(x)) .$$

В качестве мер близости можно использовать, например, статистики непараметрических критериев согласия:

а) **статистику** D_n **Колмогорова**

$$D_n = \sup_x \left| F(x,\theta) - F_n(x) \right| , \tag{1.6}$$

которую можно вычислить на основании соотношений

$$D_n = \max\left(D_n^+, D_n^-\right),$$

$$D_n^+ = \max_{1 \le i \le n} \left\{ \frac{i}{n} - F(x_i,\theta) \right\},$$

$$D_n^- = \max_{1 \le i \le n} \left\{ F(x_i,\theta) - \frac{i-1}{n} \right\},$$

где n – объем выборки; $x_1, x_2, ..., x_n$ – упорядоченные по возрастанию выборочные значения;

б) **статистику ω^2 Крамера–Мизеса–Смирнова**

$$\omega_n^2\left[\psi(F)\right] = \int\limits_{-\infty}^{\infty} \left\{E\left[F_n(x)\right] - F(x)\right\}^2 \psi\left(F(x)\right) dF(x) =$$

$$= \frac{2}{n}\sum_{i=1}^{n}\left\{g\left[F(x_i)\right] - \frac{2i-1}{2n} f\left[F(x_i)\right]\right\} + \int\limits_0^1 (1-t)^2 \psi(t) dt\,, \qquad (1.7)$$

где $E\left[\cdot\right]$ – оператор математического ожидания; $\psi(t)$ – заданная на отрезке $0 \le t \le 1$ неотрицательная функция такая, что $\psi(t)$, $t\psi(t)$, $t^2\psi(t)$ интегрируемы на отрезке $0 \le t \le 1$, и

$$f(t) = \int\limits_0^t \psi(s) ds\,, \quad g(t) = \int\limits_0^t s\psi(s) ds\,.$$

При выборе $\psi(t) \equiv 1$ получается

$$\omega_n^2 = \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n}\left\{F(x_i,\theta) - \frac{2i-1}{2n}\right\}^2; \qquad (1.8)$$

в) **статистику Ω^2 Андерсона–Дарлинга**, которая получается при выборе в (1.7) $\psi(t) \equiv 1/t(1-t)$

$$\Omega_n^2 = -1 - \frac{2}{n}\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln F(x_i,\theta) + \left(1 - \frac{2i-1}{2n}\right)\ln(1 - F(x_i,\theta))\right\}. \qquad (1.9)$$

При вычислении *MD*-оценок могут использоваться статистики любых непараметрических критериев согласия, в т.ч.: Купера, Ватсона, Жанга Z_K, Z_A, Z_C.

3. Оценивание параметров по порядковым статистикам

Для нахождения оценок часто используются линейные комбинации порядковых статистик (элементов вариационного ряда) $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$, построенного по выборке, или выборочных квантилей. Такие оценки называются *L-оценками*. *L-оценки* обладают двумя важными для практического применения качествами: простотой вычислений и хорошими свойствами робастности.

При построении L-оценок по выборочным квантилям $\hat{z}_1 < \hat{z}_2 < \ldots. < \hat{z}_k$ рассматриваемого закона, где $P\{X \le z_i\} = F(z_i, \theta) = p_i$, а выборочная квантиль определяется, например, выражением $\hat{z}_i = \left(x_{([np_i])} + x_{([np_i]+1)} \right) / 2$, где $x_{(i)}$ – i-я порядковая статистика; n – объем выборки, соответствующую оценку находят в виде

$$\hat{\theta} = \sum_{i=1}^{k} \alpha_i \hat{z}_i ,$$

где α_i – набор коэффициентов, определяющий конкретную оценку.

4. Методы оценивания параметров по группированным данным

Группирование наблюдений используется как при оценке параметров распределений, так и в задачах проверки статистических гипотез. Чаще всего при группировании область определения случайной величины разбивается на интервалы равной длины или равной вероятности. Такой подход рассматривается в наиболее часто используемых для руководства источниках.

Оценки максимального правдоподобия параметров распределений по группированным данным являются асимптотически эффективными (если они существуют и единственны). И их асимптотическая дисперсия определяется соотношением

$$D(\theta) = n^{-1}\mathbf{J}_\Gamma^{-1}(\theta),$$

где $\mathbf{J}_\Gamma(\theta) = \left[\sum_{i=1}^{k}\left(\frac{\partial \ln P_i(\theta)}{\partial \theta_l}\frac{\partial \ln P_i(\theta)}{\partial \theta_r}\right)P_i(\theta)\right] = \sum_{i=1}^{k}\frac{\nabla P_i(\theta)\nabla P_i^{\mathrm{T}}(\theta)}{P_i(\theta)}$ – информационная матрица Фишера по группированным наблюдениям.

Это же справедливо для оценок, получаемых в результате **минимизации статистики** χ^2:

$$\chi^2 = n\sum_{i=1}^{k}\frac{\left(n_i / n - P_i(\theta)\right)^2}{P_i(\theta)},$$

в результате **минимизации модифицированной статистики** χ^2:

$$\operatorname{mod}\chi^2 = n\sum_{i=1}^{k}\frac{\left(n_i - nP_i(\theta)\right)^2}{n_i},$$

где n_i заменяется на 1, если $n_i = 0$.

Минимизации расстояния Хеллингера

$$H_D = \arccos \sum_{i=1}^{k} \sqrt{\left(n_i / n\right) P_i(\theta)}\ ,$$

дивергенции Кульбака–Лейблера

$$S_{KL} = \sum_{i=1}^{k} P_i(\theta) \ln\left[P_i(\theta) / \left(n_i / n\right)\right],$$

меры расхождения Холдейна

$$D_j = \frac{(n+j)!}{n!} \sum_{i=1}^{k} \frac{n_i!\, P_i^{j+1}(\theta)}{(n_i+j)!}, \quad j \neq -1,$$

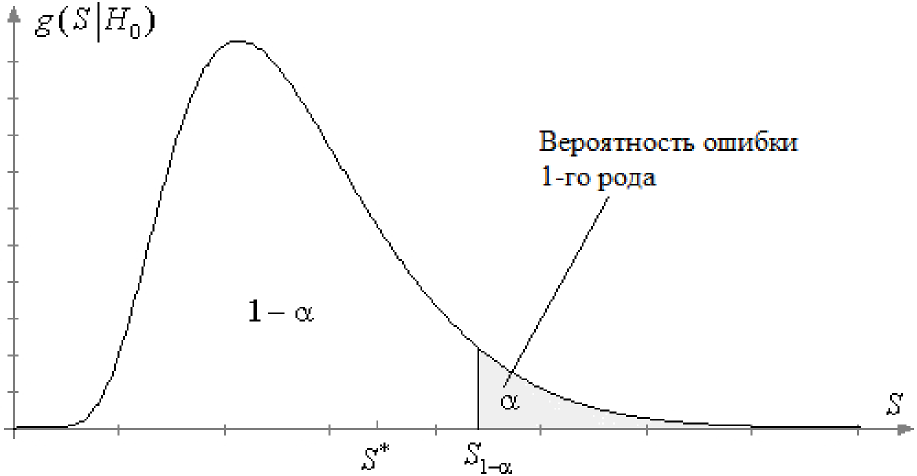
$$D_{-1} = \frac{1}{n} \sum_{i=1}^{k} n_i \ln P_i(\theta).$$

Все эти методы при соответствующих условиях регулярности дают состоятельные и асимптотически эффективные оценки.

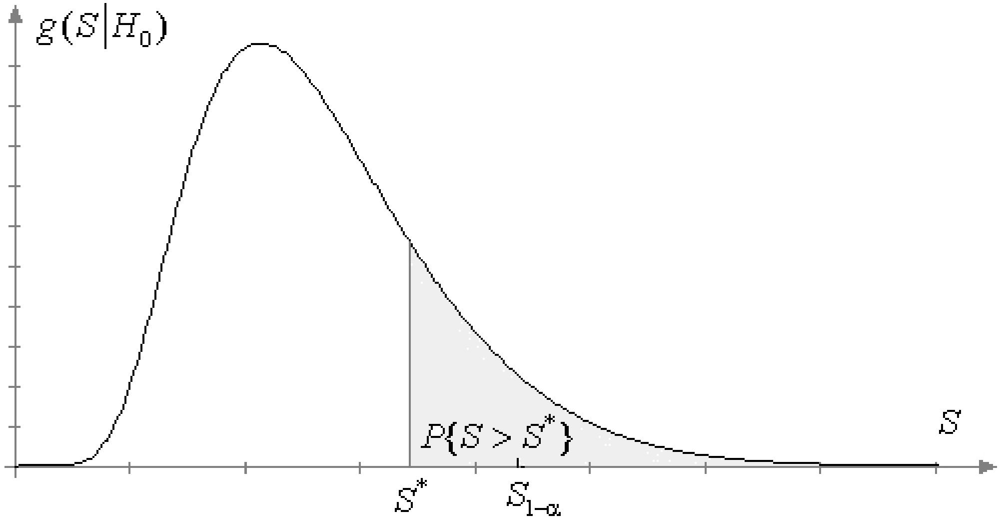
В то же вреямя следует учитывать, что их использование в случае проверки сложных гипотез с применением непараметрических критериев согласия соответствующим образом отражается на распределениях статистик этих критериев.

Общие сведения о проверке статистических гипотез

С каждым из используемых для проверки гипотезы H_0 критериев связана статистика S, которая представляет собой некоторую меру для измерения вероятности соответствия (несоответствия) анализируемых выборок проверяемой гипотезе H_0. Именно на основании значения статистики, вычисляемого по выборке или выборкам, делается вывод о принятии или отклонении проверяемой гипотезы. При справедливости проверяемой гипотезы H_0 статистика S подчиняется некоторому распределению $G(S|H_0)$.



Плотность распределения статистики при справедливости гипотезы H_0 и критическое значение для **правостороннего** критерия

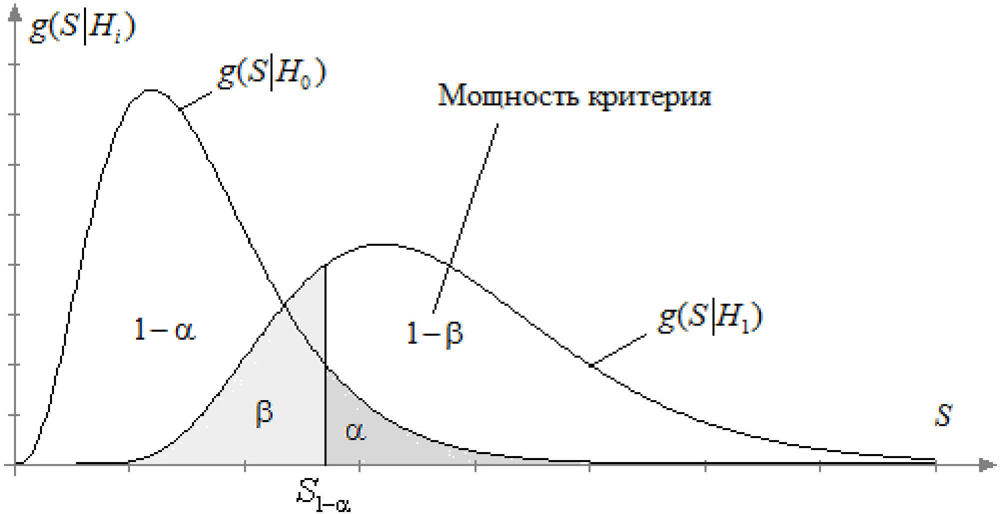


Плотность распределения статистики при справедливости гипотезы H_0 и достигнутый уровень значимости

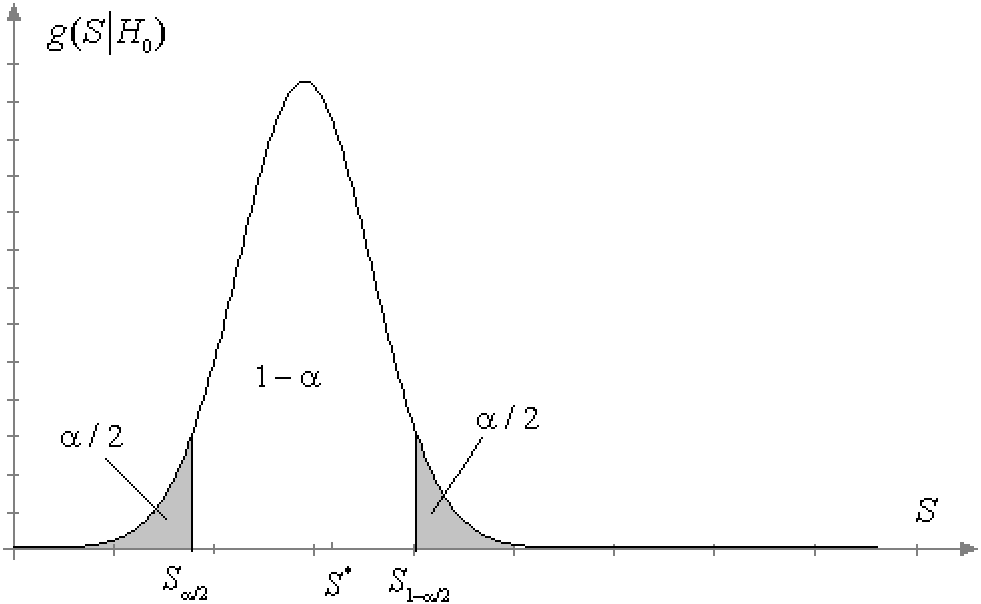
$$p_{value} = P\{S > S^*\} = \int_{S^*}^{\infty} g(s|H_0)\,ds = 1 - G(S^*|H_0)$$

Если задана конкурирующая гипотеза H_1, то можно говорить об ошибке 2-го рода и её вероятности

$$\beta = \int_{-\infty}^{S_{1-\alpha}} g(s|H_1)ds \,.$$

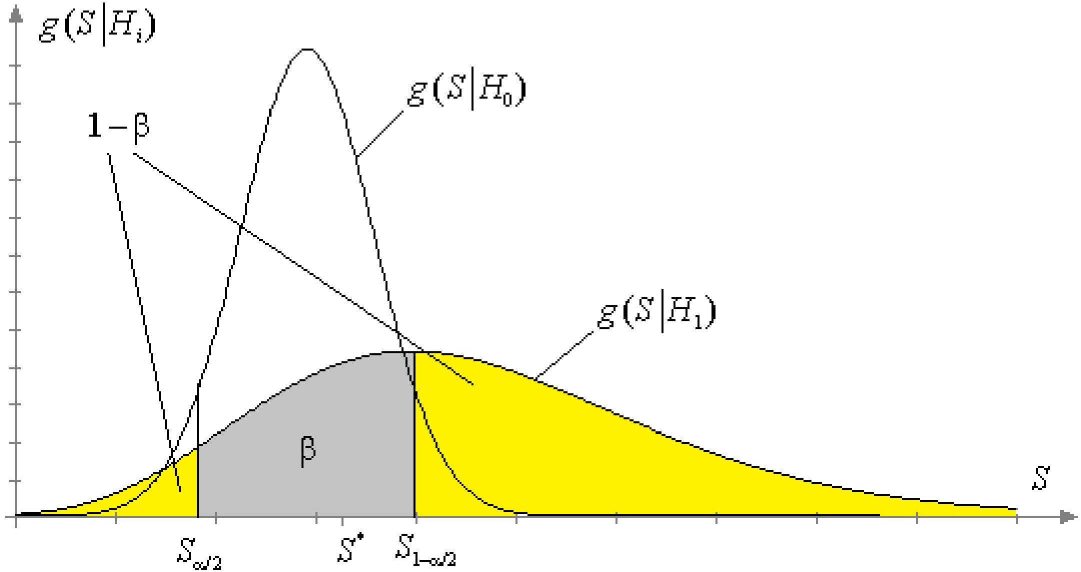


Плотности распределения статистик при справедливости соответственно гипотез H_0 и H_1 в случае **правостороннего** критерия



Плотность распределения статистики при справедливости гипотезы H_0
и критические значения для **двустороннего** критерия

$$p_{value} = 2\min\left\{G(S^*|H_0),\, 1-G(S^*|H_0)\right\}.$$



Плотности распределения статистик при справедливости соответственно гипотез H_0 и H_1 в случае двустороннего критерия

Чем критерий мощнее, тем он предпочтительней!

Из всего множества критериев проверки статистических гипотез, наиболее часто используемых в задачах принятия решений, мы рассмотрим критерии согласия, используемые при идентификации законов распределения случайных величин, в частности

Непараметрические критерии согласия.

А также множество критериев однородности, используемых, при сравнительном анализе закономерностей, связанных с анализируемыми выборками, в которых проверяется отсутствие различий в закономерностях. К ним относятся неапараметрические

Критерии однородности законов.

В задачах принятия решения нас могут интересовать различия в числовых характеристиках анализируемых выборок. В этом случае используются множества параметрических и непараметрических критериев:

Критерии однородности средних (о равенстве математических ожиданий)

Критерии однородности дисперсий (о равенстве дисперсий)

Свойства критериев согласия χ^2 Пирсона и различных модификаций критериев типа χ^2 (Никулина–Рао–Робсона, Никулина–Дапаридзе), а также влияние на распределения статистик и мощность этих критериев числа интервалов и способа разбиения на интервалы, достаточно полно представлены в различных источниках, в частности в рекомендациях ГОССТАНДАРТа РФ:

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.

0 Непараметрические критерии согласия

В данном случае будем говорить о применении критериев Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Купера, Ватсона, Жанга.

Практика применения такого рода критериев в приложениях богата большим числом примеров некорректного использования.

Наиболее распространенные ошибки применения связаны с использованием классических результатов, имеющих место при проверке простых гипотез, для ситуаций, соответствующих проверке сложных гипотез.

При проверке согласия различают простые и сложные гипотезы.

Простая проверяемая гипотеза имеет вид $H_0:\ F(x) = F(x, \theta)$, где $F(x, \theta)$ – функция распределения вероятностей, с которой проверяют согласие наблюдаемой выборки, а θ – известное значение параметра (скалярного или векторного).

Сложная проверяемая гипотеза имеет вид $H_0:\ F(x) \in \{F(x, \theta), \theta \in \Theta\}$, где Θ – область определения параметра θ.

Если процесс вычисления оценки $\hat{\theta}$ скалярного или векторного параметра закона не опирается на ту же самую выборку, по которой проверяют гипотезу о согласии, то алгоритм применения критерия согласия при проверке сложной гипотезы не отличается от проверки простой гипотезы.

0.1 Непараметрические критерии согласия при проверке простых гипотез

0.1.1 Критерий Колмогорова

Критерий Колмогорова [3] опирается на статистику

$$D_n = \sup_{|x|<\infty} \left| F_n(x) - F(x, \theta) \right|, \qquad (0.1)$$

где $F_n(x)$ – эмпирическая функция распределения; $F(x, \theta)$ – теоретическая функция распределения; n – объем выборки. При $n \to \infty$ функция распределения статистики $\sqrt{n} \cdot D_n$ сходится равномерно к функции распределения Колмогорова $K(s) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 s^2}$.

В критерии Колмогорова рекомендуется использовать статистику с поправкой Большева [4, 5] в форме [6]

$$S_K = \sqrt{n} D_n + \frac{1}{6\sqrt{n}} = \frac{6nD_n + 1}{6\sqrt{n}}, \qquad (0.3)$$

где $D_n = \max\left(D_n^+, D_n^-\right)$, $D_n^+ = \max_{1 \le i \le n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}$; $D_n^- = \max_{1 \le i \le n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}$,

n – объем выборки; $x_1, x_2, \ldots, x_n$ здесь и далее – упорядоченные по возрастанию выборочные значения; $F(x, \theta)$ – функция закона распределения, согласие с которым проверяют.

0.1.2 Критерий Крамера-Мизеса-Смирнова

Статистика критерия **Крамера–Мизеса–Смирнова** имеет вид

$$S_\omega = n\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ F(x_i, \theta) - \frac{2i-1}{2n} \right\}^2, \tag{0.13}$$

которая при простой гипотезе в пределе подчиняется закону с функцией распределения $a1(s)$, имеющей вид [6]

$$a1(s) = \frac{1}{\sqrt{2s}} \sum_{j=0}^{\infty} \frac{\Gamma(j+1/2)\sqrt{4j+1}}{\Gamma(1/2)\Gamma(j+1)} \exp\left\{ -\frac{(4j+1)^2}{16s} \right\} \times \left\{ I_{-\frac{1}{4}}\left[\frac{(4j+1)^2}{16s} \right] - I_{\frac{1}{4}}\left[\frac{(4j+1)^2}{16s} \right] \right\},$$

где $I_{-\frac{1}{4}}(\cdot), I_{\frac{1}{4}}(\cdot)$ – модифицированные функции Бесселя вида

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}, \ |z| < \infty, \ |\arg z| < \pi .$$

0.1.3 Критерий Андерсона-Дарлинга

Статистика критерия **Андерсона–Дарлинга** [7, 8] задается выражением

$$S_\Omega = n\Omega_n^2 = -n - 2\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln F(x_i,\theta) + \left(1-\frac{2i-1}{2n}\right)\ln(1-F(x_i,\theta))\right\}. \qquad (0.16)$$

В пределе при проверке простой гипотезы эта статистика подчиняется закону с функцией распределения $a2(s)$, имеющей вид [6]

$$a2(s) = \frac{\sqrt{2\pi}}{s}\sum_{j=0}^{\infty}(-1)^j\frac{\Gamma(j+1/2)(4j+1)}{\Gamma(1/2)\Gamma(j+1)}\exp\left\{-\frac{(4j+1)^2\pi^2}{8s}\right\}\times$$

$$\times\int_0^{\infty}\exp\left\{\frac{s}{8(y^2+1)} - \frac{(4j+1)^2\pi^2y^2}{8s}\right\}dy.$$

0.1.4 Критерий Купера

В критерии [9] Купера статистика V_n которого определяется соотношением

$$V_n = \sup_{-\infty<x<\infty} \left\{F_n(x) - F(x,\theta)\right\} - \inf_{-\infty<x<\infty} \left\{F_n(x) - F(x,\theta)\right\}$$

и используется в виде

$$V_n = D_n^+ + D_n^-, \tag{0.18}$$

где D_n^+, D_n^- определены выше), n – объем выборки, x_i – элементы вариационного ряда.

Статистика $\sqrt{n}V_n$ подчиняется распределению [10]:

$$G(s|H_0) = 1 - \sum_{m=1}^{\infty} 2(4m^2s^2 - 1)e^{-2m^2s^2} .$$

В критерии используется либо модификация статистики [10]

$$V = V_n\left(\sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}}\right), \tag{0.20}$$

либо [11]

$$V_n^{mod} = \sqrt{n}(D_n^+ + D_n^-) + \frac{1}{3\sqrt{n}}, \tag{0.21}$$

где идея использования поправки вытекает из выражения для статистики критерия согласия Смирнова [6, с. 81].

0.1.5 Критерий Ватсона

Статистика критерия Ватсона [12, 13] имеет вид

$$U_n^2 = n \int_{-\infty}^{\infty} \left\{ F_n(x) - F(x,\theta) - \int_{-\infty}^{\infty} \left(F_n(y) - F(y,\theta) \right) dF(y,\theta) \right\}^2 dF(x,\theta)$$

и используется в следующей удобной для расчетов форме

$$U_n^2 = \sum_{i=1}^{n} \left(F(x_i,\theta) - \frac{i - \frac{1}{2}}{n} \right)^2 - n \left(\frac{1}{n} \sum_{i=1}^{n} F(x_i,\theta) - \frac{1}{2} \right)^2 + \frac{1}{12n} . \qquad (0.23)$$

Предельное распределение $G\left(U_n^2 \middle| H_0\right)$ статистики U_n^2 приведено в [12, 13] в виде

$$G\left(s \middle| H_0\right) = 1 - 2 \sum_{m=1}^{\infty} (-1)^{m-1} e^{-2m^2\pi^2 s} .$$

0.1.6 Критерии Жанга

В диссертации Жанга [14] и в последующих работах [15, 16] предложены непараметрические критерии согласия, статистики которых имеют вид:

$$Z_K = \max_{1 \le i \le n} \left(\left(i - \frac{1}{2} \right) \log \left\{ \frac{i - \frac{1}{2}}{nF(x_i, \theta)} \right\} + \left(n - i + \frac{1}{2} \right) \log \left[\frac{n - i + \frac{1}{2}}{n\{1 - F(x_i, \theta)\}} \right] \right), \tag{0.27}$$

$$Z_A = -\sum_{i=1}^{n} \left[\frac{\log\{F(x_i, \theta)\}}{n - i + \frac{1}{2}} + \frac{\log\{1 - F(x_i, \theta)\}}{i - \frac{1}{2}} \right], \tag{0.28}$$

$$Z_C = \sum_{i=1}^{n} \left[\log \left\{ \frac{[F(x_i, \theta)]^{-1} - 1}{(n - \frac{1}{2}) / (i - \frac{3}{4}) - 1} \right\} \right]^2 \tag{0.29}$$

Использование критериев со статистиками (0.27) – (0.29) осложняет сильная зависимость распределений статистик от объема выборки n.

Естественно, зависимость от n сохраняется и в случае проверки сложных гипотез.

0.2 Непараметрические критерии согласия при проверке сложных гипотез

0.2.1 Факторы, влияющие на распределения статистик критериев при проверке сложных гипотез

При проверке сложных гипотез, когда по той же самой выборке оценивают параметры наблюдаемого закона распределения вероятностей, все рассматриваемые непараметрические критерии согласия теряют свойство «свободы от распределения» [17]. Более того, предельные распределения статистик непараметрических критериев согласия зависят от целого ряда факторов, определяющих «сложность» гипотезы.

На закон распределения статистики $G(S|H_0)$ влияют следующие факторы:

– вид наблюдаемого закона распределения $F(x, \theta)$, соответствующего истинной гипотезе H_0;

– тип оцениваемого параметра и число оцениваемых параметров;

– в некоторых ситуациях конкретное значение параметра (например, в случае гамма-распределения);

– используемый метод оценивания параметров.

Игнорирование того, что проверяют сложную гипотезу, и того, что сложные гипотезы могут быть различными, приводит к некорректному применению непараметрических критериев согласия и, как следствие, к неверным статистическим выводам.

Различия в предельных распределениях тех же самых статистик при проверке простых и сложных гипотез настолько существенны, что пренебрегать этим абсолютно недопустимо.

0.2.2 Зависимость распределений статистик критериев от метода оценивания параметров

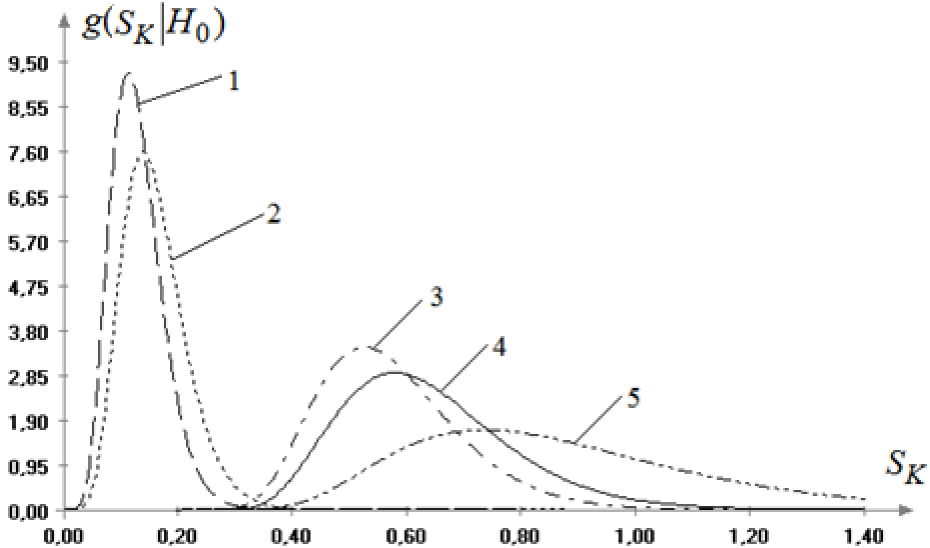


Рис. 0.1. Плотности распределения $g(S_K|H_0)$ статистики S_K критерия Колмогорова при проверке сложной гипотезы (H_0 – нормальный закон, оцениваются оба параметра: *1* – с использованием *MD*-оценок S_K; *2* – *MD*-оценок S_ω; *3* – *MD*-оценок S_Ω; *4* – ОМП; 5 – плотность распределения Колмогорова $k(s)$)

0.2.3 Зависимость распределений статистик критериев от вида закона

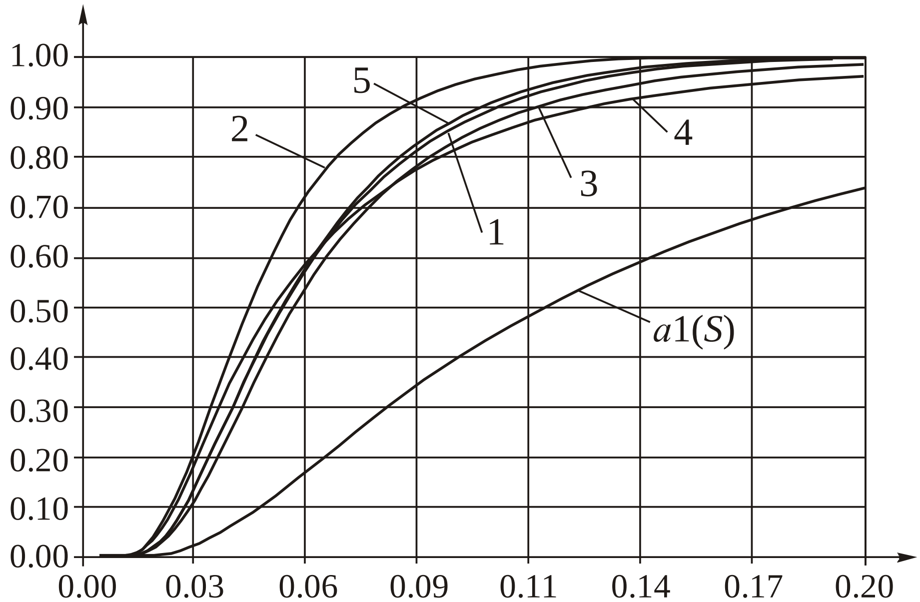


Рис. 0.2. Распределения $G(S_\omega | H_0)$ статистики S_ω Крамера–Мизеса–Смирнова при оценивании двух параметров закона, соответствующего гипотезе H_0 (*1* – нормального, *2* – логистического, *3* – Лапласа, *4* – наименьшего значения, *5* – Коши), при использовании ОМП. $a1(s)$ – функция распределения, предельная при простой гипотезе

0.2.4 Зависимость распределений статистик от числа и типа оцениваемых параметров

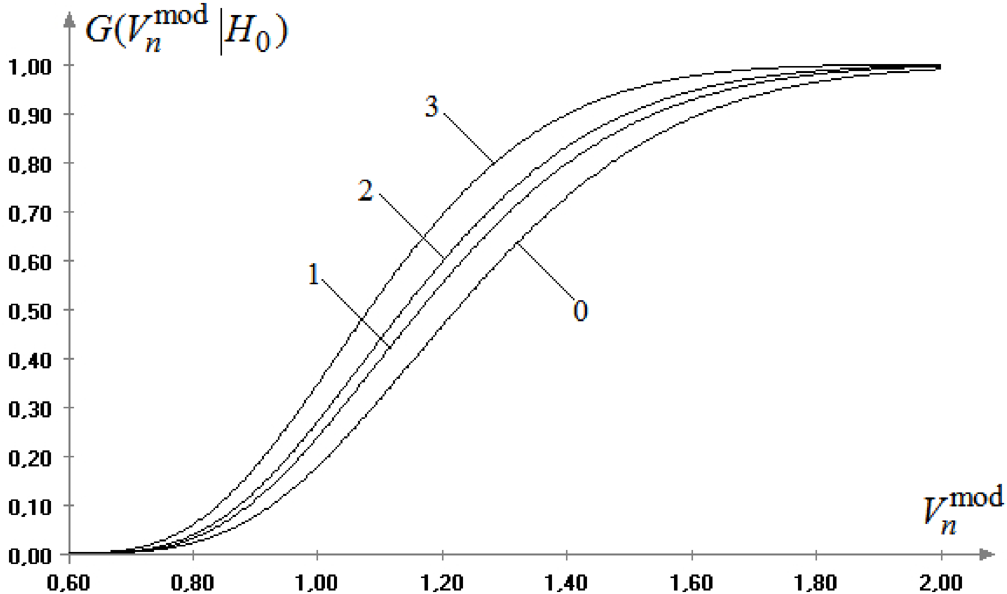
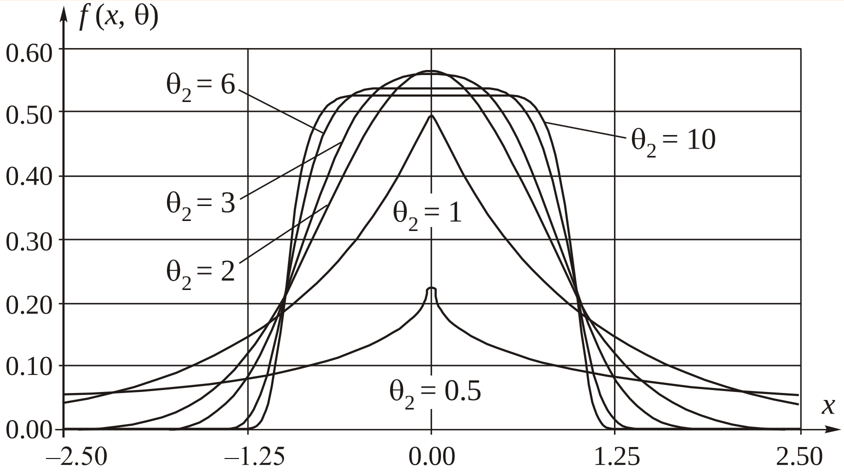


Рис. 0.3. Распределения $G(V_n^{\text{mod}}|H_0)$ статистики V_n^{mod} Купера при использовании ОМП для оценивания параметров нормального закона (*0* – без оценивания (распределение Купера), *1* – при оценивании параметра сдвига, *2* – параметра масштаба, 3 – при оценивании двух параметров)

0.2.5 Зависимость распределений статистик непараметрических критериев от конкретных значений параметра или параметров

Предельные распределения $G(S|H_0)$ рассматриваемых статистик при проверке сложных гипотез могут зависеть от конкретных значений параметров закона, с которым проверяют согласие. Это касается таких законов, как семейства гамма-распределений, семейства бета-распределений I, II и III рода, обобщенного нормального распределения, обобщенного распределения Вейбулла, обратного гауссовского распределения и других. Например, для **обобщенного нормального распределения** с плотностью:

$$f(x) = \frac{\theta_2}{2\theta_1\Gamma(1/\theta_2)} \exp\left\{-\left(\frac{|x-\theta_0|}{\theta_1}\right)^{\theta_2}\right\}. \qquad (0.7)$$



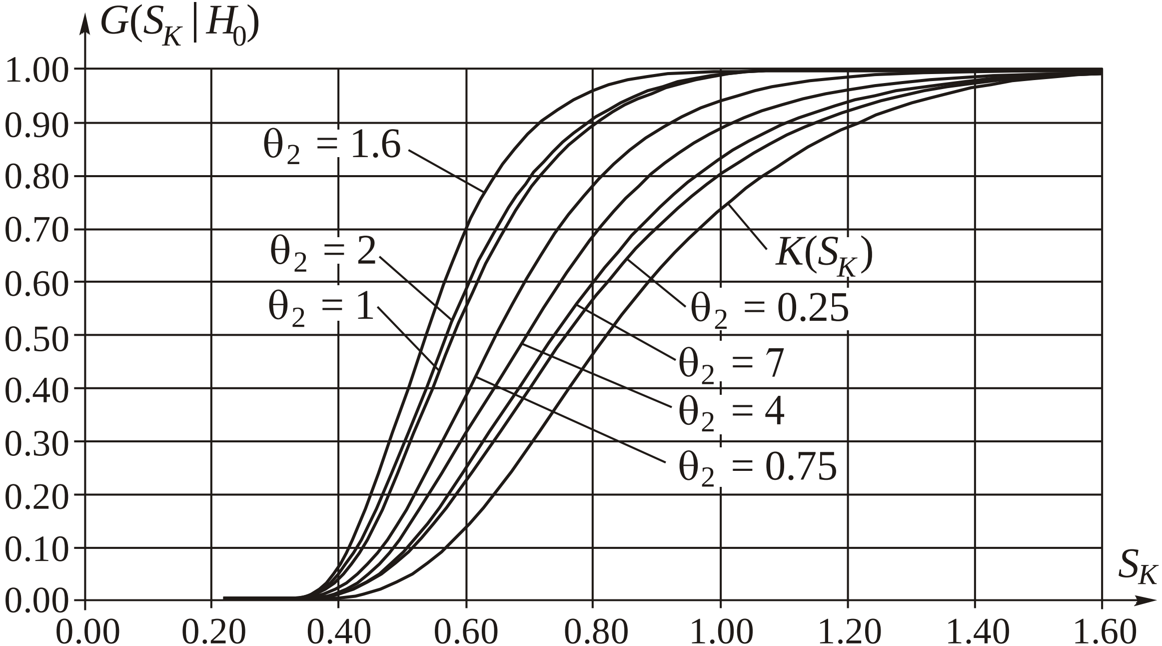


Рис. 0.5. Зависимость распределения статистики критерия типа Колмогорова от параметра формы θ_2 при оценивании всех трех параметров распределения (0.7)

0.2.6 Методика компьютерного моделирования распределений статистик

Если для описания выборки используется закон распределения вероятностей $F(x,\theta)$ и найдена оценка его параметра $\hat{\theta}$, а для проверки сложной гипотезы H_0: $F(x)\in\{F(x,\theta),\theta\in\Theta\}$, то для нахождения неизвестного распределения $G(S|H_0)$ статистики соответствующего критерия согласия целесообразно воспользоваться методикой компьютерного анализа статистических закономерностей.

Для этого следует в соответствии с законом $F(x,\hat{\theta})$ смоделировать N выборок того же объема n, что и выборка, для которой необходимо проверить гипотезу H_0: $F(x)\in\{F(x,\theta),\theta\in\Theta\}$. Далее для каждой из N выборок вычислить оценки тех же параметров закона, а затем – значение статистики S соответствующего критерия согласия.

В результате будет получена выборка значений статистики $S_1, S_2, ..., S_N$ с законом распределения $G(S_n|H_0)$ для проверяемой гипотезы H_0. По этой выборке при достаточно большом N можно построить достаточно гладкую эмпирическую функцию распределения $G_N(S_n|H_0)$, которой можно непосредственно воспользоваться для вывода о том, есть ли основания для отклонения гипотезы H_0.

При необходимости можно по $G_N(S_n|H_0)$ **построить приближенную аналитическую модель**, аппроксимирующую $G_N(S_n|H_0)$, и тогда, опираясь уже на эту модель, принимать решение относительно проверяемой гипотезы H_0.

0.2.7 Перечень законов, для которых построены модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез

№ п/п	Распределение случайной величины, область определения	Функция плотности
1	Экспоненциальное, $x \ge 0$	$f(x) = \frac{1}{\theta_0} e^{-x/\theta_0}$
2	Полунормальное, $x \ge 0$	$f(x) = \frac{2}{\theta_0 \sqrt{2\pi}} e^{-x^2/2\theta_0^2}$
3	Рэлея, $x \ge 0$	$f(x) = \frac{x}{\theta_0^2} e^{-x^2/2\theta_0^2}$
4	Максвелла, $x \ge 0$	$f(x) = \frac{2x^2}{\theta_0^3 \sqrt{2\pi}} e^{-x^2 / 2\theta_0^2}$
5	Лапласа, $x \in (-\infty, \infty)$	$f(x) = \frac{1}{2\theta_0} e^{-(x-\theta_1)/\theta_0}$
6	Нормальное, $x \in (-\infty, \infty)$	$f(x) = \frac{1}{\theta_0 \sqrt{2\pi}} e^{-\frac{(x-\theta_1)^2}{2\theta_0^2}}$

7	Логнормальное, $x \in (0, \infty)$	$f(x) = \dfrac{1}{x\theta_0\sqrt{2\pi}} e^{-(\ln x - \theta_1)^2/2\theta_0^2}$
8	Коши, $x \in (-\infty, \infty)$	$f(x) = \dfrac{\theta_0}{\pi\left[\theta_0^2 + (x-\theta_1)^2\right]}$
9	Логистическое, $x \in (-\infty, \infty)$	$f(x) = \dfrac{1}{\theta_0} \exp\left\{-\dfrac{(x-\theta_1)}{\theta_0}\right\} \bigg/ \left[1 + \exp\left\{-\dfrac{(x-\theta_1)}{\theta_0}\right\}\right]^2$
10	Наибольшего значения, $x \in (-\infty, \infty)$	$f(x) = \dfrac{1}{\theta_0} \exp\left\{-\dfrac{x-\theta_1}{\theta_0} - \exp\left(-\dfrac{x-\theta_1}{\theta_0}\right)\right\}$
11	Наименьшего значения, $x \in (-\infty, \infty)$	$f(x) = \dfrac{1}{\theta_0} \exp\left\{\dfrac{x-\theta_1}{\theta_0} - \exp\left(\dfrac{x-\theta_1}{\theta_0}\right)\right\}$
12	Вейбулла, $x \in (0, \infty)$	$f(x) = \dfrac{\theta_0 x^{\theta_0 - 1}}{\theta_1^{\theta_0}} \exp\left\{-\left(\dfrac{x}{\theta_1}\right)^{\theta_0}\right\}$
13	Sb-Джонсона, $x \in [\theta_3, \theta_2 + \theta_3]$	$f(x) = \dfrac{\theta_1\theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)} \times \exp\left\{-\dfrac{1}{2}\left[\theta_0 - \theta_1 \ln\dfrac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$

14	Sl-Джонсона, $x \in [\theta_3, \infty)$	$$f(x) = \frac{\theta_1}{\sqrt{2\pi}\left(x-\theta_3\right)} \exp\left\{-\frac{1}{2}\left[\theta_0 + \theta_1 \ln\frac{x-\theta_3}{\theta_2}\right]^2\right\}$$
15	Su-Джонсона, $x \in (-\infty, \infty)$	$$f(x) =$$ $$\frac{\theta_1}{\sqrt{2\pi}\sqrt{\left(x-\theta_3\right)^2 + \theta_2^2}} \times \exp\left\{-\frac{1}{2}\left[\theta_0 + \theta_1 \ln\left\{\frac{x-\theta_3}{\theta_2} + \sqrt{\left(\frac{x-\theta_3}{\theta_2}\right)^2 + 1}\right\}\right]^2\right\}$$

Построенные модели (*в виде параметрических законов распределения с соответствующими значениями параметров*), аппроксимирующие предельные распределения статистик критериев **Колмогорова, Смирнова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга, Купера и Ватсона,** предусматривающие, в основном, применение оценок максимального правдоподобия и, в меньшей степени, применение MD-оценок, представлены в таблицах приложения **А** руководства:

Лемешко Б. Ю. **Непараметрические критерии согласия.** Руководство по применению: Монография. – М.: ИНФРА-М, 2014. – 163 с.

Текущая версия программной системы ISW.

Доступна по адресу: https://ami.nstu.ru/~headrd/

0.2.8 Интерактивный подход к проверке гипотез в нестандартных условиях

№ п/п	Распределение случайной величины, область определения	Функция плотности
16	Гамма-распределение, $x \in (\theta_2, \infty)$	$f(x) = \dfrac{1}{\theta_1^{\theta_0}\Gamma(\theta_0)}(x-\theta_2)^{\theta_0-1} e^{-(x-\theta_2)/\theta_1}$
17	Обобщенное нормальное, $x \in (-\infty, \infty)$	$f(x) = \dfrac{\theta_2}{2\theta_1\Gamma(1/\theta_2)} \exp\left\{-\left(\dfrac{\lvert x-\theta_0\rvert}{\theta_1}\right)^{\theta_2}\right\}$
18	Обратное гауссовское, $x \in (0, \infty)$	$f(x) = \left(\dfrac{\theta_1}{2\pi x^3}\right)^{1/2} \exp\left(-\dfrac{\theta_1(x-\theta_0)^2}{2\theta_0^2 x}\right)$
19	Обобщенное Вейбулла, $x \in (0, \infty)$	$f(x) = \dfrac{\theta_0}{\theta_1}\theta_2^{\theta_0} x^{\theta_0-1}\left(1+\left(\dfrac{x}{\theta_2}\right)^{\theta_0}\right)^{\frac{1}{\theta_1}-1} e^{1-\left(1+\left(\frac{x}{\theta_2}\right)^{\theta_0}\right)^{\frac{1}{\theta_1}}}$

20	Бета-распределение I рода	$f(x)=\frac{1}{\theta_2 B(\theta_0,\theta_1)}\left(\frac{x}{\theta_2}\right)^{\theta_0-1}\left(1-\frac{x}{\theta_2}\right)^{\theta_1-1}$
21	Бета-распределение II рода	$f(x)=\frac{1}{\theta_2 B(\theta_0,\theta_1)}\frac{[x/\theta_2]^{\theta_0-1}}{[1+x/\theta_2]^{\theta_0+\theta_1}}$
22	Бета-распределение III рода	$f(x)=\frac{\theta_2^{\theta_0}}{\theta_3 B(\theta_0,\theta_1)}\frac{\left(\frac{x-\theta_4}{\theta_3}\right)^{\theta_0-1}\left(1-\frac{x-\theta_4}{\theta_3}\right)^{\theta_1-1}}{\left[1+(\theta_2-1)\frac{x-\theta_4}{\theta_3}\right]^{\theta_0+\theta_1}}$

В случае этих законов **распределения статистик критериев зависят от конкретных значений параметра или параметров формы**.

В ситуациях, когда распределения статистик критериев зависят от конкретных значений параметров формы законов, заранее найти модель распределения статистики применяемого критерия нельзя, так как значения оценок этих параметров находятся в ходе самой проверки.

В случае критериев со статистиками Жанга проблема усугубляется зависимостью распределений статистик от объемов выборок.

Следовательно, искать их (исследовать распределения статистик) можно только после вычисления оценок параметров, используя **интерактивный режим** для моделирования неизвестного распределения статистики или для вычисления p_{value}, на основании которых будет формироваться вывод по итогам проверки сложной гипотезы.

Такой интерактивный режим реализован в программной системе ISW.

В то же время следует заметить, что это не снимает всех проблем, так как моделирование неизвестного распределения статистики используемого критерия далеко не тривиальная задача, особенно, в плане оценивания параметров законов распределения по моделируемым выборкам.

Эта задача существенно усложняется при использовании моделей в виде смесей законов.

0.3 О мощности критериев

Результаты сравнительного анализа мощности непараметрических критериев относительно различных пар *близких* конкурирующих законов показали, что критерии Колмогорова (K), Крамера-Мизеса-Смирнова (KMS), Андерсона-Дарлинга (AD), Купера (V_n), Ватсона (U_n^2), Жанга (Z_C , Z_A и Z_K), **как правило**, можно упорядочить по мощности следующим образом:

– при проверке простых гипотез –

$$Z_C \succ Z_A \succ Z_K \succ U_n^2 \succ V_n \succ AD \succ KMS \succ \approx K \ ;$$

– при проверке сложных гипотез –

$$Z_A \succ \approx Z_C \succ Z_K \approx AD \succ KMS \succ U_n^2 \succ V_n \succ K \ .$$

Мощность непараметрических критериев при проверке сложных гипотез при тех же объемах выборок n всегда существенно выше, чем при проверке простых.

Если критерий χ^2 Пирсона использовать асимптотически оптимальное группирование [119] и выбирать число интервалов, при котором он будет иметь максимальную мощность , то при проверке простых гипотез **критерий χ^2 окажется на третьей позиции**.

При проверке сложных гипотез – преимущество за непараметрическими критериями согласия, а критерии типа χ^2 Никулина-Рао-Робсона [18, 19, 20] и χ^2 Пирсона оказываются соответственно на **7-й – 8-й позициях** в общем ряду критериев по убыванию мощности.

0.4 Применение критериев согласия для анализа больших выборок

Вопросы применения статистических методов к анализу больших массивов данных (Big Data) в последние годы вызывают большой интерес. В приложениях всё чаще приходится сталкиваться с необходимостью анализа гигантских объёмов накапливаемых данных. Возникают потребности извлечения и использования закономерностей, в том числе вероятностных, скрытых в этих данных.

При попытках применения для анализа больших данных классического аппарата прикладной математической статистики, как правило, встречаются со специфическими проблемами, ограничивающими возможности их корректного применения. Например, сталкиваются с тем, что хорошо зарекомендовавшие себя методы и алгоритмы становятся неэффективными из-за “проклятия размерности”. Одни популярные критерии проверки гипотез оказываются не приспособленными для анализа выборок даже порядка тысячи наблюдений. Другие, которые формально можно использовать при объёмах выборок $n \to \infty$, всегда приводят к отклонению даже справедливой проверяемой гипотезы H_0. Такие проблемы характерны для многих критериев, в том числе для непараметрических критериев согласия. И связаны они не только с ростом вычислительных затрат.

То, что очень часто информация о законе распределения $G(S|H_0)$ статистики критерия ограничена лишь узкими рамками таблицы критических значений, совсем не ограничивает возможность корректного применения критерия при объемах выборок за рамками этой таблицы. Для этого достаточно лишь воспользоваться интерактивным режимом для моделирования и последующего использования $G_N(S_n|H_0)$.

Основная причина, препятствующая корректному применению множества классических критериев проверки статистических гипотез, заключается в следующем. Как правило, объёмы выборок в Big Data (принадлежащие некоторому непрерывному закону распределения) практически неограничены, но сами данные представлены с ограниченной точностью (округлены с некоторым Δ). По сути, “нарушается предположение” о том, что наблюдается непрерывная случайная величина.

Допустим, для критерия существует предельное распределение статистики $G(S|H_0)$. Эмпирическое распределение $F_n(x)$, соответствующее выборке непрерывных случайных величин (без округления), при $n \to \infty$ сходится к функции распределения $F(x)$ этой случайной величины. Эмпирическое распределение $G_N(S_n|H_0)$ статистики, строящейся по выборке непрерывной случайной величины при $n \to \infty$ (и $N \to \infty$) сходится к предельному $G(S|H_0)$.

Пусть теперь наблюдаемые данные округляются с некоторым Δ.

Тогда, начиная с некоторого n, зависящего от вида $F(x)$, от области определения случайной величины и от Δ, $\max|F_n(x) - F(x)|$ перестанет уменьшаться, а распределение $G_N(S_n|H_0)$ – станет с ростом n отклоняться от предельного $G(S|H_0)$ (чем больше Δ, тем при меньшем n).

Рассмотрим поведение распределений $G(S_n|H_0)$ статистик критериев на примере проверки согласия со стандартным нормальным законом.

При округлении с точностью до 1 в выборках, принадлежащих N(0,1), может появляться **9** уникальных значений, при округлении с точностью $\Delta = 0.1$ – порядка **86** уникальных значений, с точностью $\Delta = 0.01$ – порядка **956**, с точностью $\Delta = 0.001$ – порядка **9830**.

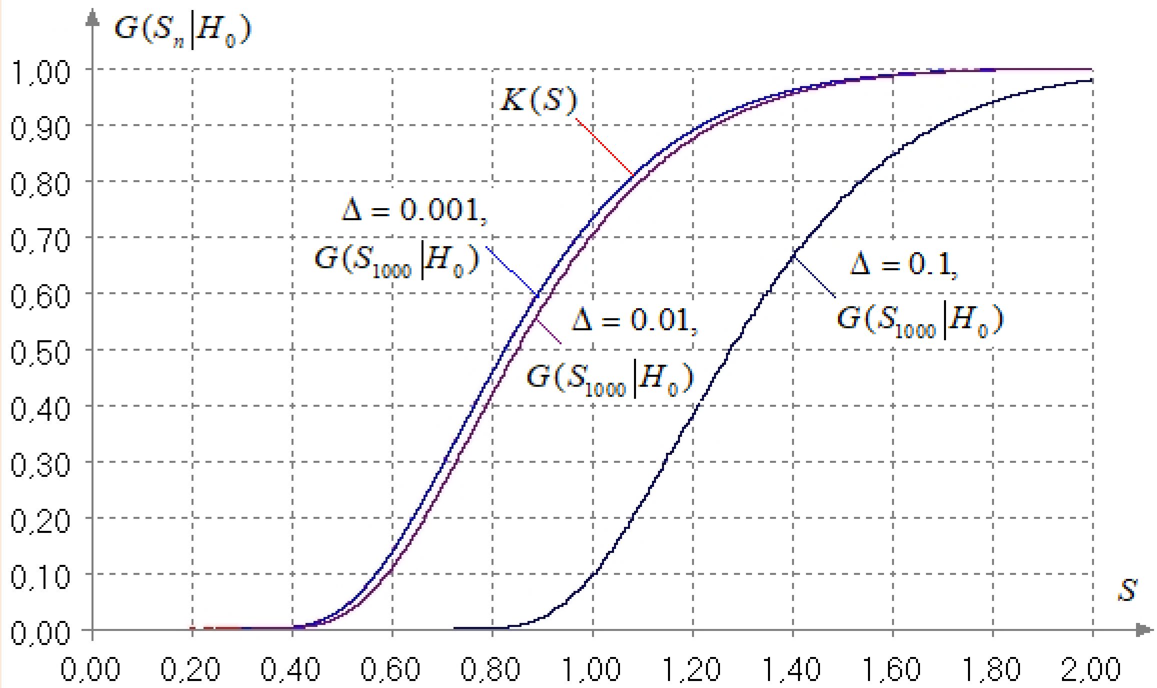


Рис. 0.6. Эмпирические распределения $G(S_{1000}|H_0)$ статистики критерия Колмогорова при проверке согласия со стандартным нормальным законом в зависимости от Δ

В результате исследований, проверденных в работе[1] были сформированы следующие рекомендции.

При анализе очень больших выборок для того, чтобы при использовании соответствующих критериев для вычисления **p**-value можно было использовать предельное распределение статистики $G(S|H_0)$ (имеющее место при проверке простой или сложной гипотезы), рекомендуется применять критерий не ко всему массиву Big Data, а **извлекать для анализа из этого массива выборки ограниченного объема**.

То есть, применять критерий следует при таких n, при которых для данной степени округления Δ реальное распределение статистики $G(S_n|H_0)$ ещё практически не отличается от $G(S|H_0)$.

[1] *Лемешко Б. Ю. Лемешко С.Б., Семёнова М.А.* К вопросу статистического анализа больших данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. – № 44. – С. 40-49. DOI: 10.17223/19988605/44/5

Влияние ошибок округления на распределения статистик

Любые измерения фиксируются с некоторой погрешностью округления.

Очевидность возможного влияния ошибок округления на статистические выводы признавалась многими авторами [137-141].

Вопрос оставался лишь в том, как ошибки округления сказываются на распределениях статистик критериев.

В ситуации, **когда ошибки округления** $\Delta << \sigma$**,** где σ – среднеквадратичное отклонение ошибки измерения, и n меньше некоторого $n_{\max}$, зависящего от n и σ, **влиянием** Δ **на** $G(S_n | H_0)$ **можно пренебречь** [142]. И можно спокойно пользоваться всеми асимптотическими результатами, имеющими место при проверке простых и сложных гипотез.

Всё меняется при соизмеримости Δ **и** σ (когда $\Delta \sim \sigma$).

Влияние ошибок округления на распределения статистик критериев проверки различных гипотез рассматривалось нами в [143-146]

В данном случае в условиях соизмеримости Δ и σ (и **без потери общности**) проиллюстрируем влияние Δ на распределения $G(S | H_0)$ статистик (1) и (2) критериев Колмогорова и Крамера–Мизеса–Смирнова на примере проверки гипотезы о принадлежности выборок нормальному закону распределения.

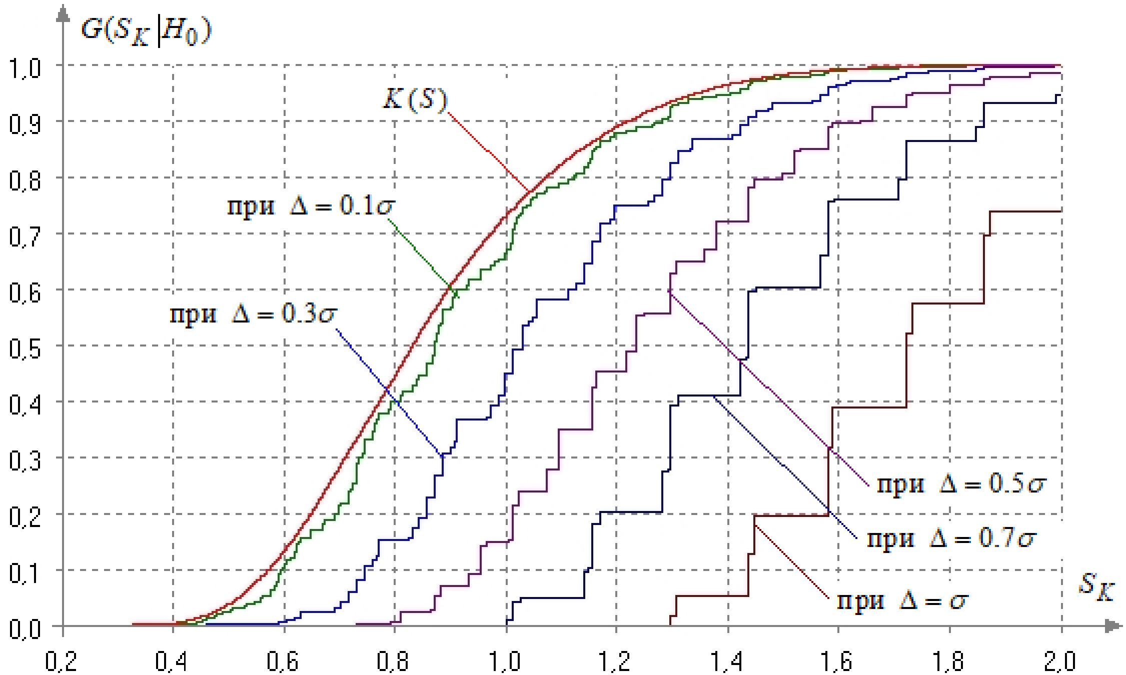


Рис. 4. Зависимость распределения статистики (1) критерия Колмогорова **от ошибки округления** Δ при справедливости **простой гипотезы** H_0 о принадлежности выборки нормальному закону при $n = 50$

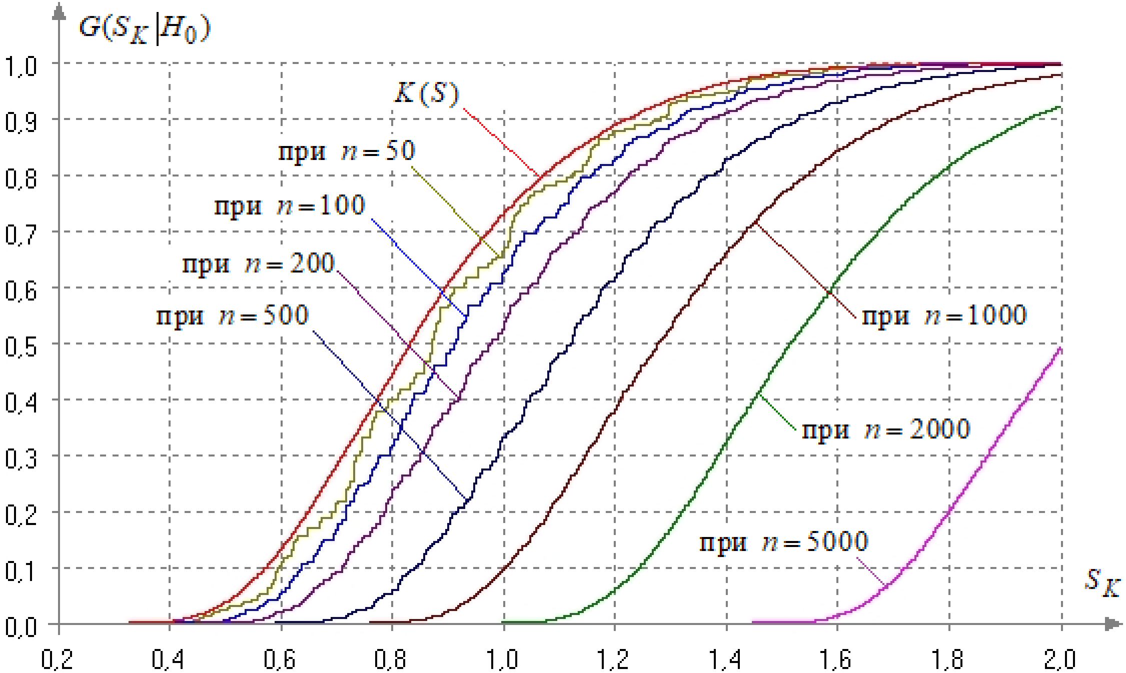


Рис. 5. Зависимость распределения статистики (1) критерия Колмогорова **от объёма выборки** n при справедливости **простой гипотезы** H_0 о принадлежности выборки нормальному закону и ошибке округления $\Delta = 0.1\sigma$

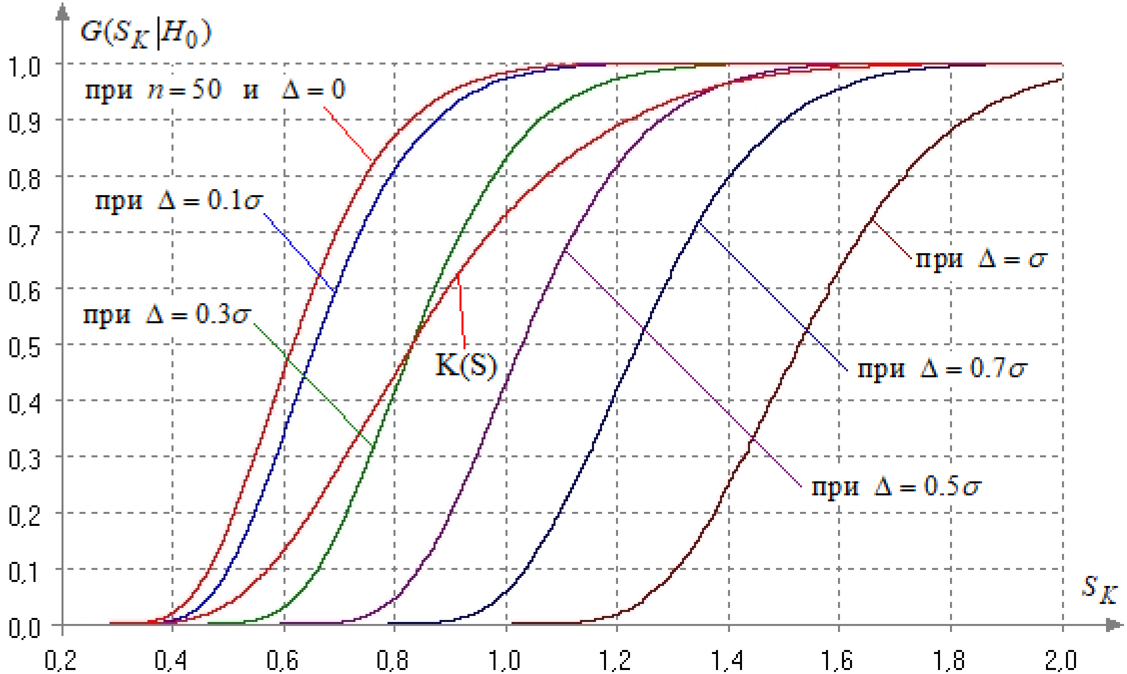


Рис. 6. Зависимость распределения статистики (1) критерия Колмогорова **от ошибки округления** Δ при справедливости **сложной гипотезы** H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $n = 50$

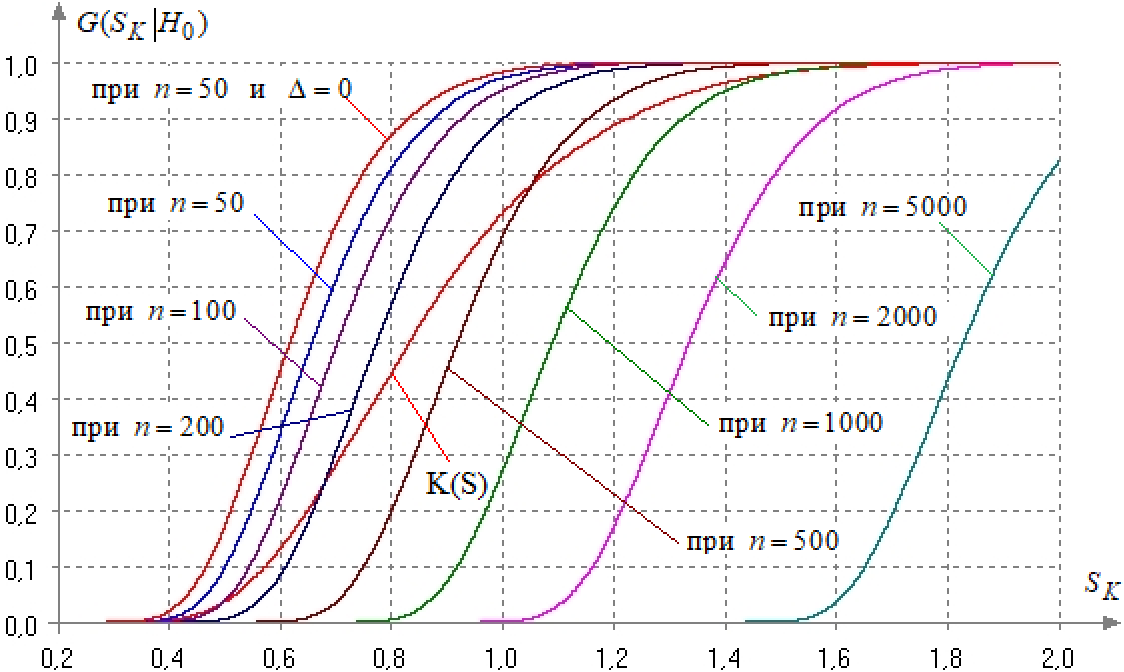


Рис. 7. Зависимость распределения статистики (1) критерия Колмогорова **от объёма выборки** n при справедливости **сложной гипотезы** H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $\Delta = 0.1\sigma$

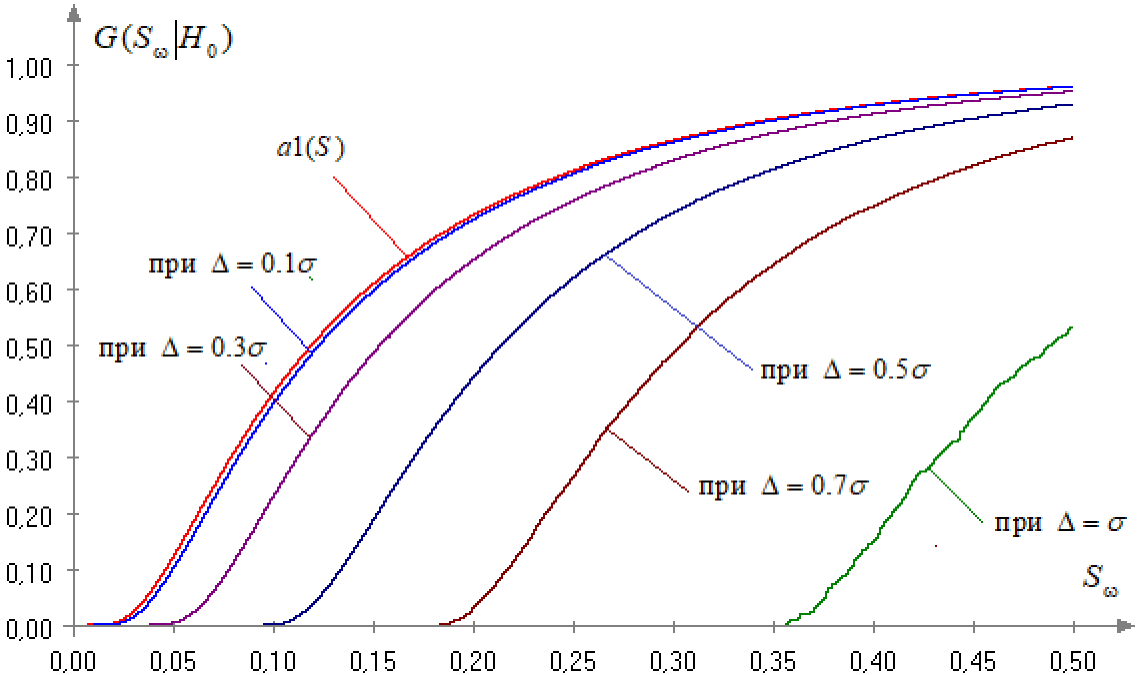


Рис. 8. Зависимость распределения статистики (2) критерия Крамера–Мизеса–Смирнова **от ошибки округления** Δ при справедливости **простой гипотезы** H_0 о принадлежности выборки нормальному закону при $n = 50$

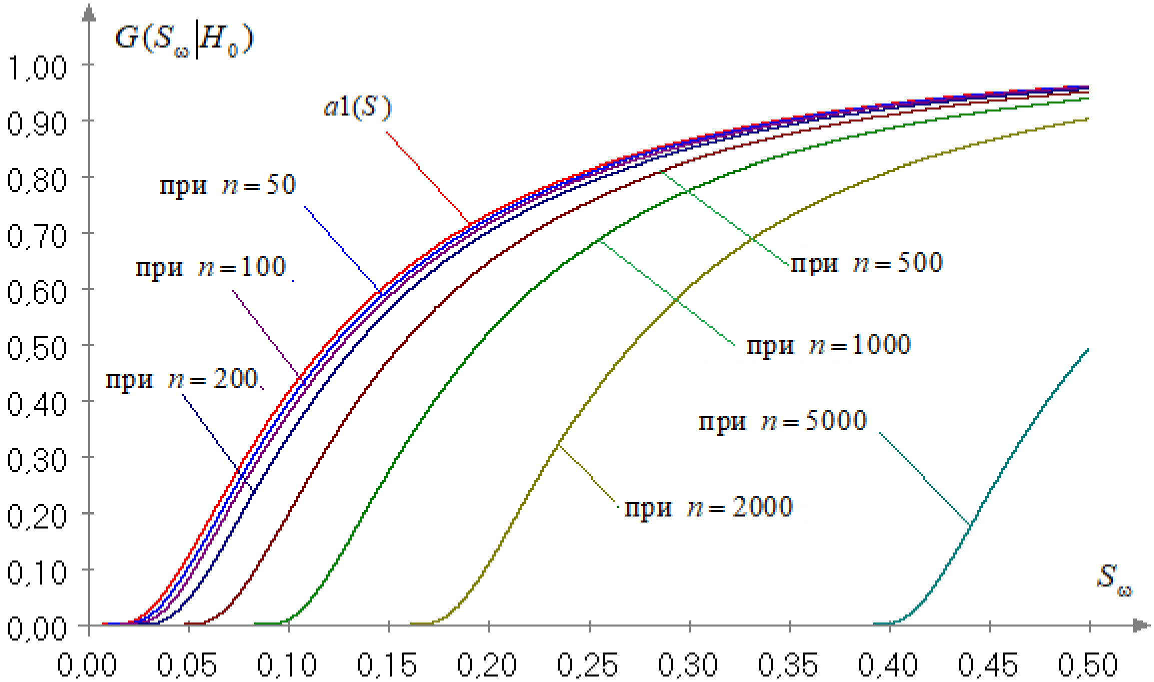


Рис. 9. Зависимость распределения статистики (2) критерия Крамера–Мизеса–Смирнова **от объёма выборки** n при справедливости **простой гипотезы** H_0 о принадлежности выборки нормальному закону и ошибке округления $\Delta = 0.1\sigma$

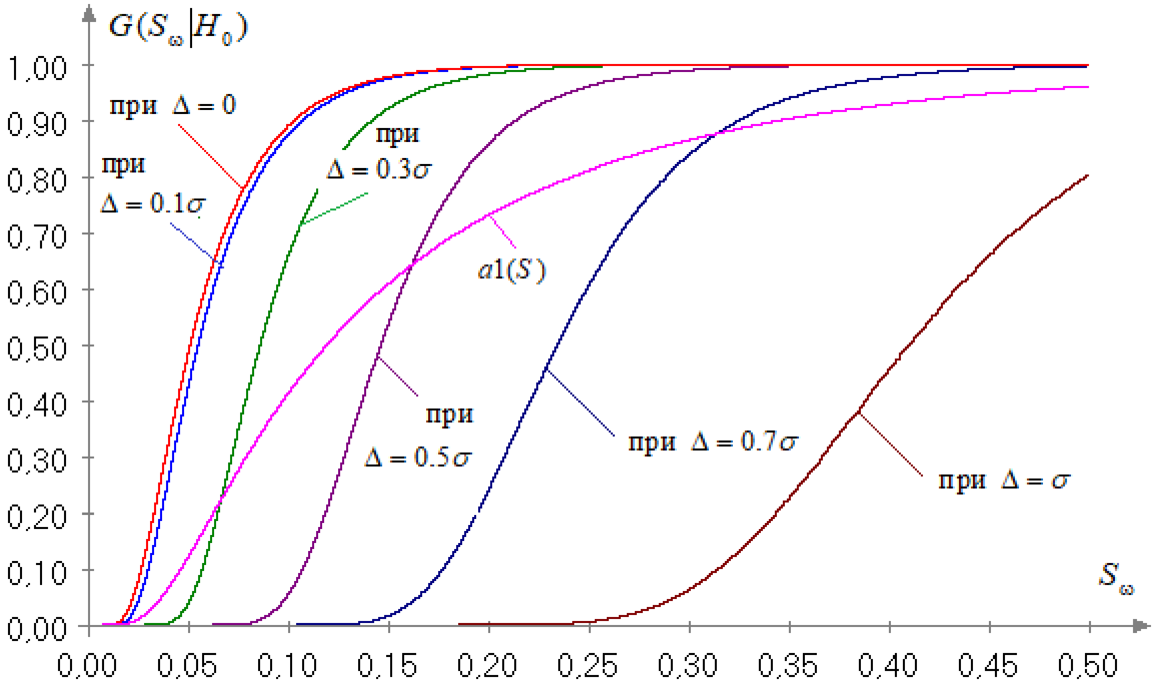


Рис. 10. Зависимость распределения статистики (2) критерия Крамера–Мизеса–Смирнова **от ошибки округления** Δ при справедливости **сложной гипотезы** H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $n = 50$

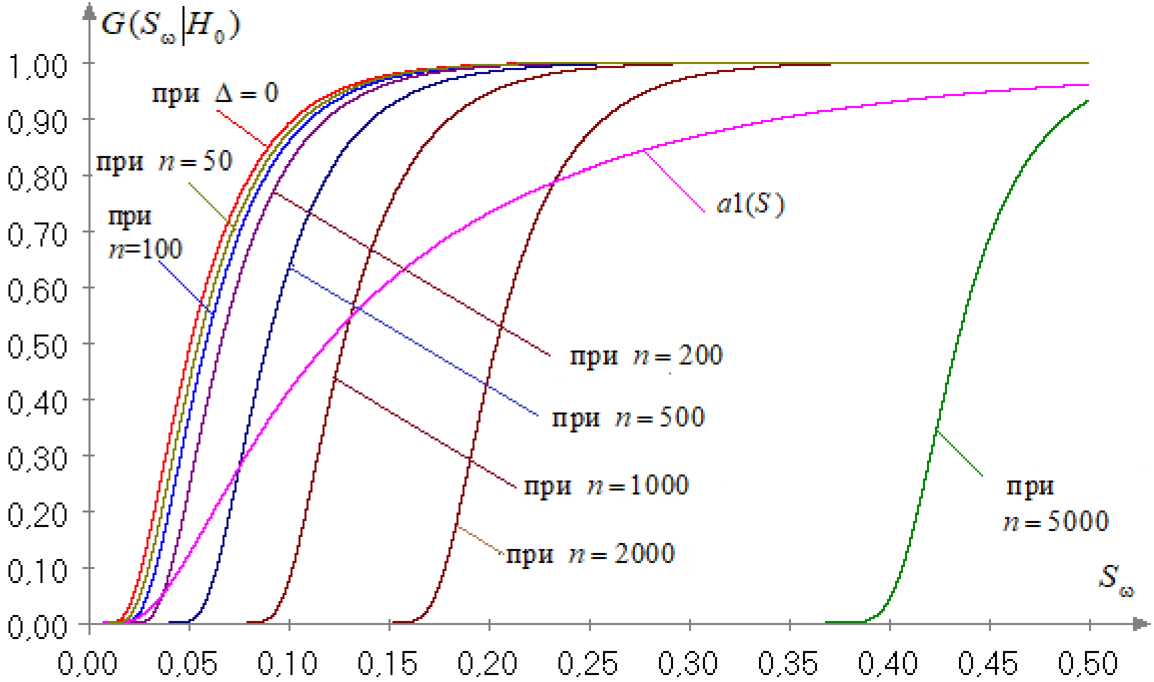


Рис. 11. Зависимость распределения статистики (2) критерия Крамера–Мизеса–Смирнова **от объёма выборки** n при справедливости **сложной гипотезы** H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $\Delta = 0.1\sigma$

Какие выводы следуют из представленных результатов исследований?

Во-первых, можно видеть, что наличие ошибок округления приводит к появлению зависимости $G(S|H_0)$ от n.

Во-вторых, признание самого факта наличия округлений в данных **исключает** возможность использования предельных распределений в качестве распределений статистик **в условиях больших выборок** [142].

В-третьих, в условиях соизмеримости Δ и σ и при относительно небольших объёмах выборок распределения $G(S_n|H_0)$ статистик непараметрических критериев могут значительно отличаться от предельных, что полностью исключает возможность при проверке гипотезы использовать классические результаты.

В-четвёртых, проведенные исследования показали, что вследствие округлений потеря свойства “свободы от распределения” происходит и в условиях проверки простых гипотез. В частности, при одном и том же соотношении $\Delta = w\sigma$ между Δ и σ наблюдаемого симметричного закона степень отклонения $G(S_n|H_0)$ от предельного увеличивается в случае законов с более тяжёлыми “хвостами” (по сравнению с нормальным законом).

В ситуации проверки сложной гипотезы при использовании ОМП для оценки параметров закона мы имеем аналогичную картину влияния Δ на распределения статистик непараметрических критериев согласия.

Влияние ошибок округления на распределения статистик критериев на примере анализа характеристик ирисов

На выборке измерений ширины лепестка **ириса разноцветного** объёмом $n = 50$, заимствованной в работе [147], проверим сложную гипотезу о её принадлежности нормальному закону.

1.4	1.5	1.5	1.3	1.5	1.3	1.6	1.0	1.3	1.4
1.0	1.5	1.0	1.4	1.3	1.4	1.5	1.0	1.5	1.1
1.8	1.3	1.5	1.2	1.3	1.4	1.4	1.7	1.5	1.0
1.1	1.0	1.2	1.6	1.5	1.6	1.5	1.3	1.3	1.3
1.2	1.4	1.2	1.0	1.3	1.2	1.3	1.3	1.1	1.3

С подобными результатами измерений часто сталкиваются в различных приложениях.

Измерения представлены в см и зафиксированы с ошибкой округления $\Delta = 0.1$. ОМП параметров сдвига и масштаба нормального закона, вычисленные по выборке, $\hat{\mu} = 1.3260$, $\hat{\sigma} = 0.1958$.

В таблице 1 приведены значения статистик применяемых критериев согласия и соответствующие оценки p_{value}, вычисленные по распределениям статистик, имеющим место в отсутствие округлений (при $\Delta = 0$) и при ошибке округления $\Delta = 0.1$.

Таблица1. Результаты проверки гипотезы о нормальности

№ п/п	Критерий	Статистика	p_{value}	
			$\Delta = 0$	$\Delta = 0.1$
1	Колмогорова	1.065	0.010	0.453
2	Крамера–Мизеса–Смирнова	0.154	0.023	0.467
3	Андерсона–Дарлинга	0.975	0.015	0.319
4	Купера	1.886	0.001	0.432
5	Ватсона	0.153	0.014	0.451
6	Жанга Z_A	3.387	0.050	0.214
7	Жанга Z_C	12.31	0.055	0.226
8	Жанга Z_K	2.563	0.008	0.283

По разности значений p_{value} можно судить о различии соответствующих распределений статистик.

И можно видеть, что при учёте округления $\Delta = 0.1$ гипотеза о нормальности не отклоняется.

Результаты проверки принадлежности нормальным законам всех выборок из [147] приведены далее.

Таблица 1. Результаты измерений характеристик ирисов [147]

№ п/п	Iris setosa				Iris versicolor				Iris virginica			
	Sepal length	Sepal width	Petal length	Petal width	Sepal length	Sepal width	Petal length	Petal width	Sepal length	Sepal width	Petal length	Petal width
1	5.1	3.5	1.4	0.2	7.0	3.2	4.7	1.4	6.3	2.2	6.0	2.5
2	4.9	3.0	1.4	0.2	6.4	3.2	4.5	1.5	5.8	2.5	5.1	1.9
3	4.7	3.2	1.3	0.2	6.9	3.1	4.9	1.5	7.1	2.5	5.9	2.1
4	4.6	3.1	1.5	0.2	5.5	2.3	4.0	1.3	6.3	2.5	5.6	1.8
5	5.0	3.6	1.4	0.2	6.5	2.8	4.6	1.5	6.5	2.5	5.8	2.2
6	5.4	3.9	1.7	0.4	5.7	2.8	4.5	1.3	7.6	2.6	6.6	2.1
7	4.6	3.4	1.4	0.3	6.3	3.3	4.7	1.6	4.9	2.6	4.5	1.7
8	5.0	3.4	1.5	0.2	4.9	2.4	3.3	1.0	7.3	2.7	6.3	1.8
9	4.4	2.9	1.4	0.2	6.6	2.9	4.6	1.3	6.7	2.7	5.8	1.8
10	4.9	3.1	1.5	0.1	5.2	2.7	3.9	1.4	7.2	2.7	6.1	2.5
11	5.4	3.7	1.5	0.2	5.0	2.0	3.5	1.0	6.5	2.7	5.1	2.0
12	4.8	3.4	1.6	0.2	5.9	3.0	4.2	1.5	6.4	2.8	5.3	1.9
13	4.8	3.0	1.4	0.1	6.0	2.2	4.0	1.0	6.8	2.8	5.5	2.1
14	4.3	3.0	1.1	0.1	6.1	2.9	4.7	1.4	5.7	2.8	5.0	2.0
15	5.8	4.0	1.2	0.2	5.6	2.9	3.6	1.3	5.8	2.8	5.1	2.4
16	5.7	4.4	1.5	0.4	6.7	3.1	4.4	1.4	6.4	2.8	5.3	2.3
17	5.4	3.9	1.3	0.4	5.6	3.0	4.5	1.5	6.5	2.8	5.5	1.8
18	5.1	3.5	1.4	0.3	5.8	2.7	4.1	1.0	7.7	2.8	6.7	2.2
19	5.7	3.8	1.7	0.3	6.2	2.2	4.5	1.5	7.7	2.8	6.9	2.3
20	5.1	3.8	1.5	0.3	5.6	2.5	3.9	1.1	6.0	2.9	5.0	1.5
21	5.4	3.4	1.7	0.2	5.9	3.2	4.8	1.8	6.9	2.9	5.7	2.3
22	5.1	3.7	1.5	0.4	6.1	2.8	4.0	1.3	5.6	3.0	4.9	2.0
23	4.6	3.6	1.0	0.2	6.3	2.5	4.9	1.5	7.7	3.0	6.7	2.0
24	5.1	3.3	1.7	0.5	6.1	2.8	4.7	1.2	6.3	3.0	4.9	1.8
25	4.8	3.4	1.9	0.2	6.4	2.9	4.3	1.3	6.7	3.0	5.7	2.1
26	5.0	3.0	1.6	0.2	6.6	3.0	4.4	1.4	7.2	3.0	6.0	1.8
27	5.0	3.4	1.6	0.4	6.8	2.8	4.8	1.4	6.2	3.0	4.8	1.8
28	5.2	3.5	1.5	0.2	6.7	3.0	5.0	1.7	6.1	3.0	4.9	1.8
29	5.2	3.4	1.4	0.2	6.0	2.9	4.5	1.5	6.4	3.0	5.6	2.1
30	4.7	3.2	1.6	0.2	5.7	2.6	3.5	1.0	7.2	3.0	5.8	1.6
31	4.8	3.1	1.6	0.2	5.5	2.4	3.8	1.1	7.4	3.0	6.1	1.9
32	5.4	3.4	1.5	0.4	5.5	2.4	3.7	1.0	7.9	3.0	6.4	2.0
33	5.2	4.1	1.5	0.1	5.8	2.7	3.9	1.2	6.4	3.0	5.6	2.2
34	5.5	4.2	1.4	0.2	6.0	2.7	5.1	1.6	6.3	3.1	5.1	1.5
35	4.9	3.1	1.5	0.2	5.4	3.0	4.5	1.5	6.1	3.1	5.6	1.4
36	5.0	3.2	1.2	0.2	6.0	3.4	4.5	1.6	7.7	3.1	6.1	2.3
37	5.5	3.5	1.3	0.2	6.7	3.1	4.7	1.5	6.3	3.1	5.6	2.4
38	4.9	3.6	1.4	0.1	6.3	2.3	4.4	1.3	6.4	3.2	5.5	1.8
39	4.4	3.0	1.3	0.2	5.6	3.0	4.1	1.3	6.0	3.2	4.8	1.8
40	5.1	3.4	1.5	0.2	5.5	2.5	4.0	1.3	6.9	3.2	5.4	2.1
41	5.0	3.5	1.3	0.3	5.5	2.6	4.4	1.2	6.7	3.2	5.6	2.4
42	4.5	2.3	1.3	0.3	6.1	3.0	4.6	1.4	6.9	3.2	5.1	2.3
43	4.4	3.2	1.3	0.2	5.8	2.6	4.0	1.2	5.8	3.3	5.1	1.9
44	5.0	3.5	1.6	0.6	5.0	2.3	3.3	1.0	6.8	3.3	5.9	2.3
45	5.1	3.8	1.9	0.4	5.6	2.7	4.2	1.3	6.7	3.3	5.7	2.5
46	4.8	3.0	1.4	0.3	5.7	3.0	4.2	1.2	6.7	3.4	5.2	2.3
47	5.1	3.8	1.6	0.2	5.7	2.9	4.2	1.3	6.3	3.4	5.0	1.9
48	4.6	3.2	1.4	0.2	6.2	2.9	4.3	1.3	6.5	3.6	5.2	2.0
49	5.3	3.7	1.5	0.2	5.1	2.5	3.0	1.1	6.2	3.8	5.4	2.3
50	5.0	3.3	1.4	0.2	5.7	2.8	4.1	1.3	5.9	3.8	5.1	1.8

Таблица 2. Результаты проверки гипотез о принадлежности нормальному закону характеристик для Iris setosa

№ п/п	Критерий	Sepal length			Sepal width			Petal length			Petal width		
		$\mu = 5.006,\ \sigma = 0.3489$			$\mu = 3.428,\ \sigma = 0.3753$			$\mu = 1.462,\ \sigma = 0.1719$			$\mu = 0.246,\ \sigma = 0.1043$		
		S	p_{value}		S	p_{value}		S	p_{value}		S	p_{value}	
			$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$
1	K	0.828	0.106	0.479	0.758	0.192	0.614	1.102	0.006	0.521	2.501	0.000	0.000
2	CMS	0.071	0.269	0.659	0.074	0.248	0.558	0.187	0.009	0.429	0.982	0.000	0.0001
3	AD	0.414	0.339	0.731	0.484	0.231	0.496	0.999	0.013	0.522	4.747	0.000	0.0001
4	Ku	1.511	0.047	0.388	1.420	0.086	0.485	2.110	0.000	0.343	4.148	0.000	0.0004
5	W	0.070	0.229	0.627	0.072	0.217	0.540	0.188	0.004	0.402	0.940	0.000	0.0001
6	Z_A	3.312	0.636	0.828	3.356	0.148	0.220	3.365	0.107	0.520	3.706	0.000	0.001
7	Z_C	5.643	0.518	0.689	6.690	0.376	0.515	8.390	0.216	0.756	41.661	0.0002	0.002
8	Z_K	1.175	0.262	0.630	1.168	0.078	0.249	0.206	0.019	0.610	12.067	0.000	0.000

Таблица 3. Результаты проверки гипотез о принадлежности нормальному закону характеристик для Iris versicolor

№ п/п	Критерий	Sepal length			Sepal width			Petal length			Petal width		
		$\mu = 5.936,\ \sigma = 0.5110$			$\mu = 2.770,\ \sigma = 0.3106$			$\mu = 4.260,\ \sigma = 0.4652$			$\mu = 1.3260,\ \sigma = 0.1958$		
		S	p_{value}		S	p_{value}		S	p_{value}		S	p_{value}	
			$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$
1	K	0.716	0.265	0.557	0.888	0.061	0.415	0.860	0.079	0.259	1.065	0.010	0.453
2	CMS	0.059	0.395	0.602	0.105	0.094	0.345	0.091	0.147	0.262	0.154	0.023	0.467
3	AD	0.374	0.421	0.613	0.573	0.139	0.450	0.562	0.149	0.254	0.975	0.015	0.319
4	Ku	1.282	0.199	0.519	1.403	0.097	0.678	1.252	0.232	0.623	1.886	0.001	0.432
5	W	0.057	0.362	0.576	0.098	0.093	0.373	0.076	0.191	0.359	0.153	0.014	0.451
6	Z_A	3.313	0.624	0.717	3.320	0.500	0.742	3.354	0.155	0.201	3.387	0.050	0.214
7	Z_C	6.214	0.441	0.515	5.900	0.483	0.691	8.593	0.203	0.256	12.31	0.055	0.226
8	Z_K	0.849	0.520	0.742	1.348	0.176	0.568	1.285	0.204	0.399	2.563	0.008	0.283

Таблица 4. Результаты проверки гипотез о принадлежности нормальному закону характеристик для Iris virginica

№ п/п	Критерий	Sepal length			Sepal width			Petal length			Petal width		
		$\mu=6.5880,\ \sigma=0.6295$			$\mu=2.9740,\ \sigma=0.3193$			$\mu=5.552,\ \sigma=0.5463$			$\mu=2.0260,\ \sigma=0.2719$		
		S	p_{value}		S	p_{value}		S	p_{value}		S	p_{value}	
			$\Delta=0$	$\Delta=0.1$		$\Delta=0$	$\Delta=0.1$		$\Delta=0$	$\Delta=0.1$		$\Delta=0$	$\Delta=0.1$
1	K	0.841	0.095	0.202	0.925	0.042	0.314	0.843	0.092	0.238	0.895	0.057	0.504
2	CMS	0.089	0.155	0.210	0.106	0.093	0.313	0.087	0.165	0.250	0.121	0.059	0.308
3	AD	0.557	0.153	0.204	0.611	0.112	0.355	0.619	0.108	0.159	0.760	0.048	0.245
4	Ku	1.331	0.151	0.336	1.744	0.008	0.170	1.322	0.159	0.411	1.746	0.008	0.259
5	W	0.085	0.140	0.198	0.102	0.081	0.308	0.074	0.201	0.322	0.121	0.044	0.273
6	Z_A	3.332	0.334	0.377	3.342	0.238	0.389	3.356	0.146	0.175	3.356	0.146	0.296
7	Z_C	6.711	0.373	0.420	7.263	0.314	0.486	9.466	0.148	0.178	9.737	0.133	0.275
8	Z_K	1.306	0.194	0.296	1.496	0.124	0.439	1.554	0.107	0.201	2.076	0.028	0.230

Обратите внимание на замечания!

Замечание 1: В данном случае каждой проверке по каждому критерию при $\Delta=0.1$ и $n=50$ **соответствует своё распределение статистики** $G(S|H_0)$**,** зависящее от σ нормального закона.

Замечание 2. В общем случае произвольного закона при проверке **сложной** гипотезы к факторам, влияющим на распределения статистик $G(S|H_0)$ (при сложной гипотезе), добавляется зависимость от n, от Δ и от **значений оценок** параметров **формы** и **масштаба** закона.

Замечание 3. В общем случае произвольного закона при проверке **простой** гипотезы распределения статистик $G(S|H_0)$ становятся зависящими от n, от Δ, от вида закона $F(x,\theta)$ и от **значений** параметров **формы** и **масштаба** закона.

Выводы по влиянию ошибок округления

Таким образом, при идентификации закона распределения, соответствующего анализируемой выборке, с использованием непараметрических критериев согласия наряду со сложностью проверяемой гипотезы следует учитывать возможное влияние ошибок округления.

Ошибки округления всегда сопровождают процесс измерений. В ситуации, когда $\Delta << \sigma$ и n меньше некоторого $n_{\max}$, зависящего от n и σ, влиянием Δ на $G(S_n | H_0)$ можно пренебречь.

Но при $n > n_{\max}$ реальное распределение статистики отклоняется от асимптотического. В таком случае использование асимптотического распределения приводит к увеличению вероятностей ошибок 1-го рода, т.е. к отклонению справедливой гипотезы H_0.

В условиях соизмеримости Δ и σ такая ситуация (т.е. отклонение реальных распределений статистик от асимптотических или от имеющих место при $\Delta = 0$ в случае зависимости распределений статистик от n) может проявляться и при малых объёмах выборок.

А с ростом n она будет только усугубляться.

С такого рода выборками сталкиваются, например, в биологии, зоологии, медицине и т.п., где в связи со спецификой измерения традиционно обладают не высокой точностью (с большими Δ).

С такими же выборками сталкиваются при высокоточных измерениях в технических приложениях, когда измерения осуществляются на пределе точности измерительных систем.

Общим в том и другом случаях являются измерения на пределе точности.

Единственным выходом, способным обеспечивать корректность выводов по применяемым критериям в нестандартных условиях (при проверке сложных гипотез и влиянии Δ на $G(S_n|H_0)$), является использование реальных распределений статистик этих критериев (имеющих место в этих нестандартных условиях).

Эта задача может (должна!) решаться в интерактивном режиме (в процессе проверки) и опираться на компьютерные технологии исследования (**на метод Монте-Карло!**) и аппарат математической статистики.

В различных приложениях **важную роль** играют определённые модели законов распределения, в частности, **нормальный** закон распределения, **равномерный** закон распределения, **экспоненциальный** закон распределения. В этой связи появилось множество специальных критериев, ориентированных только на проверку гипотез о принадлежности выборок этим конкретным законам.

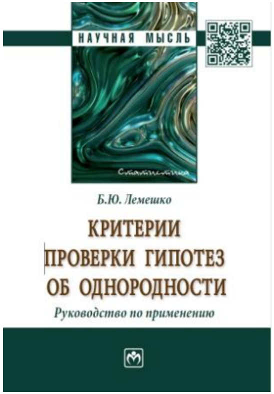
И таких критериев достаточно много, и часто они имеют преимущество в мощности по сравнению с общими критериями согласия. Но и имеют определённые недостатки.

Результатам исследований таких групп критериев посвящены следующие руководства по применению:

1. Лемешко Б.Ю. **Критерии проверки отклонения распределения от нормального закона**. Руководство по применению: Монография. – М.: ИНФРА-М, 2015. – 160 с. – (Научная мысль). DOI: 10.12737/6086

2. Лемешко Б.Ю., Блинов П.Ю. **Критерии проверки отклонения распределения от равномерного закона**. Руководство по применению: Монография. – М.: ИНФРА-М, 2015. – 183 с. – (Научная мысль). DOI: 10.12737/11304

3. Лемешко Б.Ю., Блинов П.Ю. **Критерии проверки отклонения от экспоненциального закона**. Руководство по применению : Монография. – Москва : ИНФРА-М, 2021. – 352 с. – (Научная мысль). – DOI 10.12737/1097477



Результаты исследований критериев проверки однородности законов, критериев проверки гипотез об однородности средних (о равенстве математических ожиданий), критериев проверки гипотез об однородности дисперсий (о равенстве дисперсий), соответствующих анализируемым выборкам, а также вопросы применения этого множества критериев рассмотрены в руководстве по применению:

Лемешко Б.Ю. **Критерии проверки гипотез об однородности**. Руководство по применению : монография. – 2-е изд., перераб. и доп. – Москва : ИНФРА-М, 2021. – 248 с. – (Научная мысль). – DOI 10.12737/986695

Исследовались распределения статистик и мощность множества критериев, применяемых для проверки гипотез об отсутствии тренда различного характера в наблюдаемых данных.

Лемешко Б.Ю., Веретельникова И.В. **Критерии проверки гипотез о случайности и отсутствии тренда.** Руководство по применению: Монография / Б.Ю. Лемешко, И.В. Веретельникова. – Москва : ИНФРА-М. 2021. – 221 с. – (Научная мысль).

k-выборочные критерии проверки однородности законов

С необходимостью решения задач проверки гипотез о принадлежности двух (или более) выборок случайных величин одной и той же генеральной совокупности (проверки однородности) постоянно сталкиваются в различных приложениях. При этом возникают проблемы корректности применения и выбора наиболее предпочтительного критерия.

Задача проверки однородности k выборок формулируется следующим образом. Пусть x_{ij} j-е наблюдение i-й выборки $j=\overline{1,n_i}$, $i=\overline{1,k}$. Предположим, что i-й выборке соответствует непрерывная функция распределения $F_i(x)$. Необходимо проверить гипотезу вида $H_0: F_1(x)=F_2(x)=\ldots=F_k(x)$ при любом x без указания общего для них закона распределения. Эмпирическую функцию распределения, соответствующую i-й выборке обозначим как $F_{in_i}(x)$.

На практике чаще всего применяются двухвыборочные критерии Смирнова [97] и Лемана–Розенблатта [52, 79]. Значительно реже упоминается об использовании критерия Андерсона–Дарлинга [78] (Андерсона–Дарлинга–Петита) или его k-выборочного варианта [83] и ещё реже о применении k-выборочных вариантов критериев Смирнова или Лемана–Розенблатта [45, 29, 30]. Практически не говорится об использовании критериев однородности Жанга [14, 95].

1. Рассматриваемые критерии

1.1 Критерий Смирнова

Критерий однородности Смирнова предложен в работе [135]. Предполагается, что функции распределения $F_1(x)$ и $F_2(x)$ являются непрерывными. Статистика критерия Смирнова измеряет расстояние между эмпирическими функциями распределения, построенными по выборкам

$$D_{n_1,n_2} = \sup_x \left| F_{1,n_i}(x) - F_{2,n_2}(x) \right| .$$

При практическом использовании критерия статистика D_{n_1,n_2} вычисляется в соответствии с соотношениями [6]:

$$D^+_{n_1,n_2} = \max_{1 \le r \le n_1} \left[\frac{r}{n_1} - F_{2,n_2}(x_{1r}) \right] = \max_{1 \le s \le n_2} \left[F_{1,n_1}(x_{2s}) - \frac{s-1}{n_2} \right],$$

$$D^-_{n_1,n_2} = \max_{1 \le r \le n_1} \left[F_{2,n_2}(x_{1r}) - \frac{r-1}{n_1} \right] = \max_{1 \le s \le n_2} \left[\frac{s}{n_2} - F_{1,n_1}(x_{2s}) \right],$$

$$D_{n_1,n_2} = \max\left(D^+_{n_1,n_2}, D^-_{n_1,n_2} \right).$$

При справедливости гипотезы H_0 статистика критерия Смирнова

$$S_{\text{C}} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_{n_1,n_2} \tag{1}$$

в пределе подчиняется распределению Колмогорова $K(S)$ [6].

Однако при ограниченных значениях n_1 и n_2 случайная величина D_{n_1,n_2} является дискретной, а число её возможных значений представляет собой наименьшее общее кратное n_1 и n_2 [6]. Ступенчатость условного распределения $G(S_{\text{C}} | H_0)$ статистики S_{C} при равных n_1 и n_2 сохраняется даже при $n_i = 1000$. Поэтому предпочтительнее применять критерий, когда объемы выборок n_1 и n_2 не равны и представляют собой взаимно простые числа.

Другим недостатком критерия со статистикой (1) является то, что распределения $G(S_{\text{C}} | H_0)$ с ростом n_1 и n_2 медленно приближаются к предельному распределению слева и при ограниченных n_1 и n_2 существенно отличаются от $K(s)$ (см. рис.1). В этой связи в [112] предложена простая модификация статистики (1):

$$S_{\text{CM}} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(D_{m,n} + \frac{n_1 + n_2}{4.6 n_1 n_2} \right),$$

у которой практически отсутствует последний недостаток.

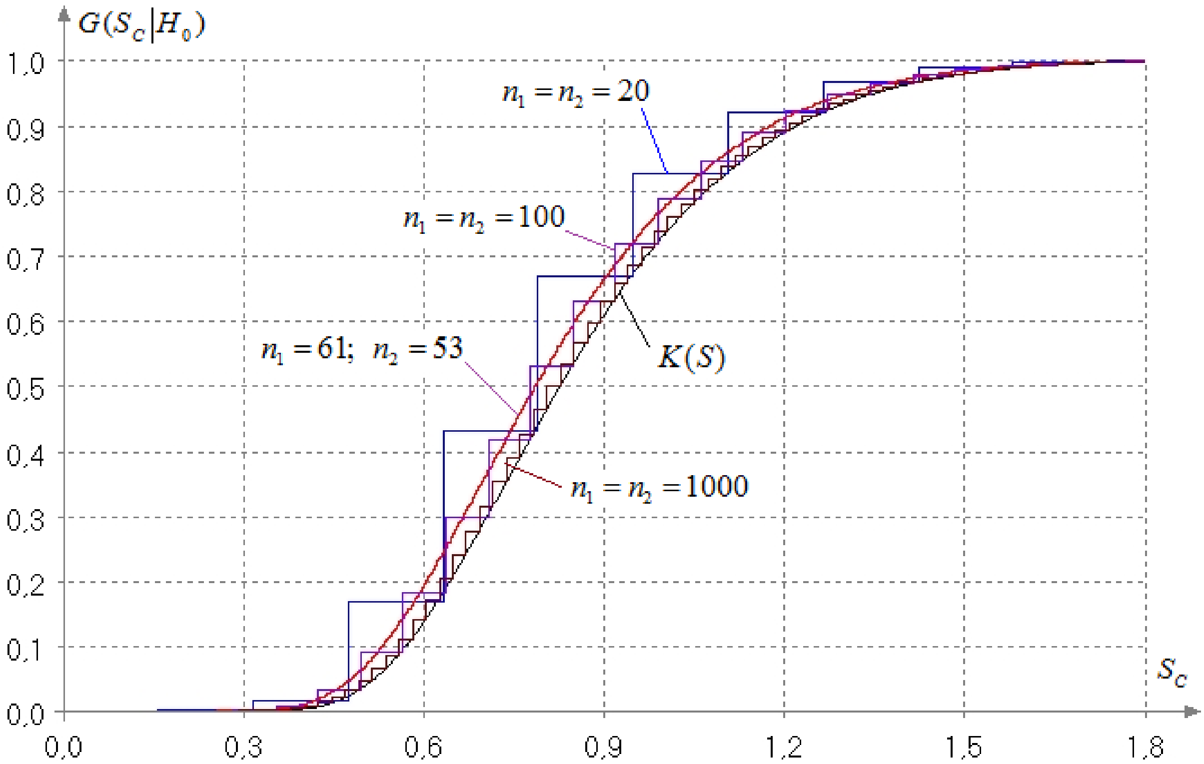


Рис. 1. Распределения статистики (1) при справедливости H_0 в зависимости от n_1 и n_2

1.2 Критерий Лемана–Розенблатта

Критерий однородности Лемана–Розенблатта представляет собой критерий типа ω^2. Критерий предложен в работе [53] и исследован в [79].

Статистика критерия используется в форме [53]

$$T = \frac{1}{n_1 n_2 (n_1 + n_2)} \left[n_2 \sum_{i=1}^{n_2} \left(r_i - i \right)^2 + n_1 \sum_{j=1}^{n_1} \left(s_j - j \right)^2 \right] - \frac{4 n_1 n_2 - 1}{6(n_1 + n_2)}, \qquad (2)$$

где r_i – порядковый номер (ранг) x_{2i}; s_j – порядковый номер (ранг) x_{1j} в объединенном вариационном ряде. В [79] было показано, что статистика (4) в пределе распределена как $a1(t)$ [97].

В отличие от критерия Смирнова распределение статистики T быстро сходится к предельному $a1(T)$. При $n_1 = n_2 = 100$ распределение $G\left(T \middle| H_0\right)$ визуально совпадает с $a1(T)$, а при $n_1, n_2 \geq 45$ отклонением $G\left(T \middle| H_0\right)$ от $a1(T)$ на практике можно пренебречь.

1.3 Критерий Андерсона–Дарлинга

Двухвыборочный критерий Андерсона–Дарлинга (критерий однородности) рассмотрен в работе [78]. Статистика применяемого критерии определяется выражением

$$A^2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1+n_2-1} \frac{\left(M_i(n_1+n_2) - n_1 i\right)^2}{i(n_1+n_2-i)}, \tag{3}$$

где M_i – число элементов первой выборки, меньших или равных i-му элементу вариационного ряда объединенной выборки.

Предельным распределением статистики (3) при справедливости проверяемой гипотезы H_0 является то же самое распределение $a2(t)$ [97], которое является предельным для статистики критерия согласия Андерсона–Дарлинга.

Сходимость распределения $G\left(A^2 \middle| H_0\right)$ статистики (3) к $a2(A^2)$ при ограниченных объёмах выборок была исследована в [130], где было показано, что при $n_1, n_2 \geq 45$ отклонение функции распределения $G\left(A^2 \middle| H_0\right)$ от $a2(A^2)$ не превышает 0.01.

1.4 Многовыборочный критерий Андерсона–Дарлинга

Многовыборочный вариант критерия согласия Андерсона–Дарлинга предложен в [83]. В предположении о непрерывности $F_i(x)$ по анализируемым выборкам строится объединённая общим объёмом $n=\sum_{i=1}^{k} n_i$ и упорядочивается $X_1 \le X_2 \le \ldots \le X_n$. Статистика критерия имеет вид [83]:

$$A_{kn}^2 = \frac{1}{n}\sum_{i=1}^{k}\frac{1}{n_i}\sum_{j=1}^{n-1}\frac{\left(nM_{ij}-jn_i\right)^2}{j(n-j)}, \tag{4}$$

где M_{ij} – число элементов в i-й выборке, которые не больше чем X_j. Проверяемая гипотеза H_0 отклоняется при больших значениях статистики (4).

В [83] таблица верхних процентных точек представлена не для статистики (4), а для статистики вида:

$$T_{kn} = \frac{A_{kn}^2-(k-1)}{\sqrt{D[A_{kn}^2]}}. \tag{5}$$

Дисперсия статистики A_{kn}^2 определяется выражением [83]

$$D\left[A_{kn}^2\right] = \frac{an^3+bn^2+cn+d}{(n-1)(n-2)(n-3)}$$

при

$$a=(4g-6)(k-1)+(10-6g)H,$$

$$b=(2g-4)k^2+8hk+(2g-14h-4)H-8h+4g-6,$$

$$c=(6h+2g-2)k^2+(4h-4g+6)k+(2h-6)H+4h,$$

$$d = (2h+6)k^2 - 4hk,$$

где

$$H = \sum_{i=1}^{k} \frac{1}{n_i}, \quad h = \sum_{i=1}^{n-1} \frac{1}{i}, \quad g = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \frac{1}{(n-i)\,j}.$$

Зависимость предельных распределений статистики (5) от числа сравниваемых выборок k иллюстрирует рис. 2. С ростом числа сравниваемых выборок это распределение медленно сходится к стандартному нормальному закону.

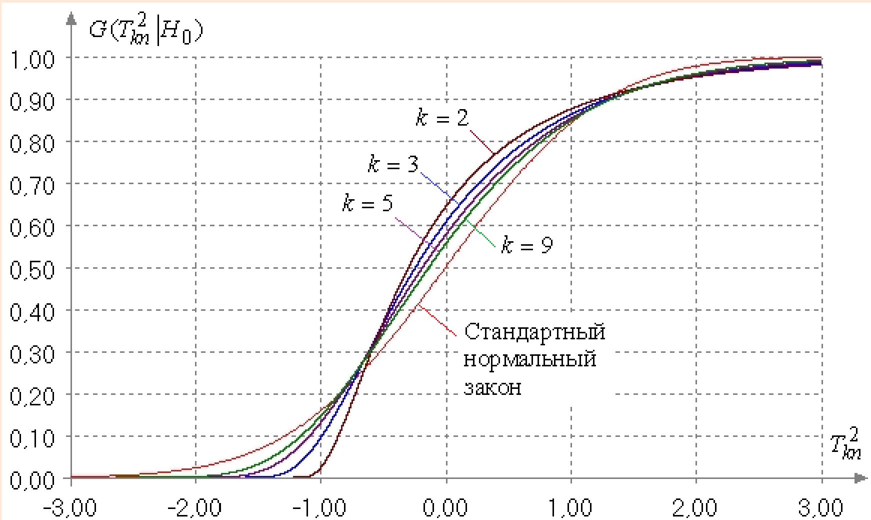


Рис. 2. Зависимость предельных распределений статистики (5) от числа сравниваемых выборок

Исследование распределений статистик методами статистического моделирования показало, что при использовании критериев отличие распределений статистик от соответствующих предельных не имеет практического значения при $n_i \geq 30$.

Таблица верхних процентных точек предельных распределений для статистики (5) представлена в [83].

Таблица 1

Уточненные верхние критические значения $T_{kn}^2(\alpha)$ статистики (5)

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	0.325	1.228	1.966	2.731	3.784
3	0.439	1.300	1.944	2.592	3.429
4	0.491	1.321	1.925	2.511	3.277
5	0.523	1.331	1.900	2.453	3.153
6	0.543	1.333	1.885	2.410	3.078
7	0.557	1.337	1.870	2.372	3.017
8	0.567	1.335	1.853	2.344	2.970
9	0.577	1.334	1.847	2.323	2.927
10	0.582	1.3345	1.838	2.306	2.899
11	0.589	1.332	1.827	2.290	2.867
∞	0.674	1.282	1.645	1.960	2.326

Мы несколько уточнили и расширили таблицу критических значений.

Одновременно для предельных распределений статистики (5) были построены приближенные модели законов (для $k = 2 \div 11$). Хорошими моделями оказались законы семейства бета-распределений III рода с плотностью

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 \mathrm{B}(\theta_0, \theta_1)} \frac{\left(\dfrac{x-\theta_4}{\theta_3}\right)^{\theta_0 - 1} \left(1 - \dfrac{x-\theta_4}{\theta_3}\right)^{\theta_1 - 1}}{\left[1 + (\theta_2 - 1)\dfrac{x-\theta_4}{\theta_3}\right]^{\theta_0 + \theta_1}}$$

при конкретных значениях параметров этого закона $\mathrm{B_{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$, найденным по полученным в результате моделирования выборкам статистик объёмом $N = 10^6$.

Представленные в таблице 2 модели $\mathrm{B_{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ с приведенными значениями параметров позволяют по значениям статистики, вычисленным по соотношению (5), находить оценки p_{value} при соответствующем числе k сравниваемых выборок.

Т а б л и ц а 2

Модели предельных распределений статистики (11)

k	Модель
2	B_{III} (3.1575, 2.8730, 18.1238, 15.0000, –1.1600)
3	B_{III} (3.5907, 4.5984, 7.8040, 14.1310, –1.5000)
4	B_{III} (4.2657, 5.7035, 5.3533, 12.8243, –1.7500)
5	B_{III} (6.2992, 6.5558, 5.6833, 13.010, –2.0640)
6	B_{III} (6.7446, 7.1047, 5.0450, 12.8562, –2.2000)
7	B_{III} (6.7615, 7.4823, 4.0083, 11.800, –2.3150)
8	B_{III} (5.8057, 7.8755, 2.9244, 10.900, –2.3100)
9	B_{III} (9.0736, 7.4112, 4.1072, 10.800, –2.6310)
10	B_{III} (10.2571, 7.9758, 4.1383, 11.186, –2.7988)
11	B_{III} (10.6848, 7.5950, 4.2041, 10.734, –2.8400)
∞	N(0.0, 1.0)

1.5 Критерии однородности Жанга

Предложенные Жангом критерии однородности [14, 95] являются развитием критериев Смирнова, Лемана–Розенблатта и Андерсона–Дарлинга и дают возможность сравнивать $k \geq 2$ выборок.

Критерии согласия Жанга [106] показывают некоторое преимущество в мощности по сравнению с критериями согласия Крамера–Мизеса–Смирнова и Андерсона–Дарлинга, но недостатком, ограничивающим применение критериев согласия Жанга, является зависимость распределений статистик от объёмов выборок.

Этим же недостатком обладают варианты критериев Жанга для проверки однородности законов.

Для преодоления этого недостатка автор [14] предлагает для оценивания p_{value} использовать метод Монте–Карло.

Задача моделирования распределений статистик критериев однородности Жанга, по сравнению с аналогичной задачей для критериев согласия, оказывается много проще, так как приходится моделировать распределения статистик $G(S|H_0)$ критериев в случае принадлежности анализируемых выборок равномерному закону.

Пусть $x_{i1}, x_{i2}, \ldots, x_{in_i}$ упорядоченные выборки непрерывных случайных величин с функциями распределения $F_i(x)$, $(i = \overline{1,k})$ и пусть $X_1 < X_2 < \ldots < X_n$, где $n = \sum_{i=1}^{k} n_i$, объединённая упорядоченная выборка.

Обозначим R_{ij} ранг j-го упорядоченного наблюдения x_{ij} i-й выборки в объединённой выборке. Пусть $X_0 = -\infty$, $X_{n+1} = +\infty$, а ранги $R_{i,0} = 1$, $R_{i,n_i+1} = n+1$.

В критериях используется модификация эмпирической функции распределения $\hat{F}(t)$, принимающая в точках разрыва X_m, $m = \overline{1,n}$, значения $\hat{F}(X_m) = (m-0.5)/n$ [14].

Статистика Z_K критерия однородности Жанга имеет вид [14]:

$$Z_K = \max_{1 \le m \le n} \left\{ \sum_{i=1}^{k} n_i \left[F_{i,m} \ln \frac{F_{i,m}}{F_m} + \left(1 - F_{i,m}\right) \ln \frac{1 - F_{i,m}}{1 - F_m} \right] \right\}, \tag{6}$$

где $F_m = \hat{F}(X_m)$, так что $F_m = (m-0.5)/n$, а вычисление $F_{i,m} = \hat{F}_i(X_m)$ осуществляется следующим образом. В начальный момент значения $j_i = 0$, $i = \overline{1,k}$. Если $R_{i,j_i+1} = m$, то $j_i := j_i + 1$ и $F_{i,m} = (j_i - 0.5)/n_i$, в противном случае если $R_{i,j_i} < m < R_{i,j_i+1}$, то $F_{i,m} = j_i/n_i$.

Критерий правосторонний: проверяемая гипотеза H_0 отклоняется при **больших** значениях статистики (6). Распределения статистики зависят от n_i и от k. На принятие решения влияет дискретность статистики, которая с ростом k становится менее выраженной (см. рис.3).

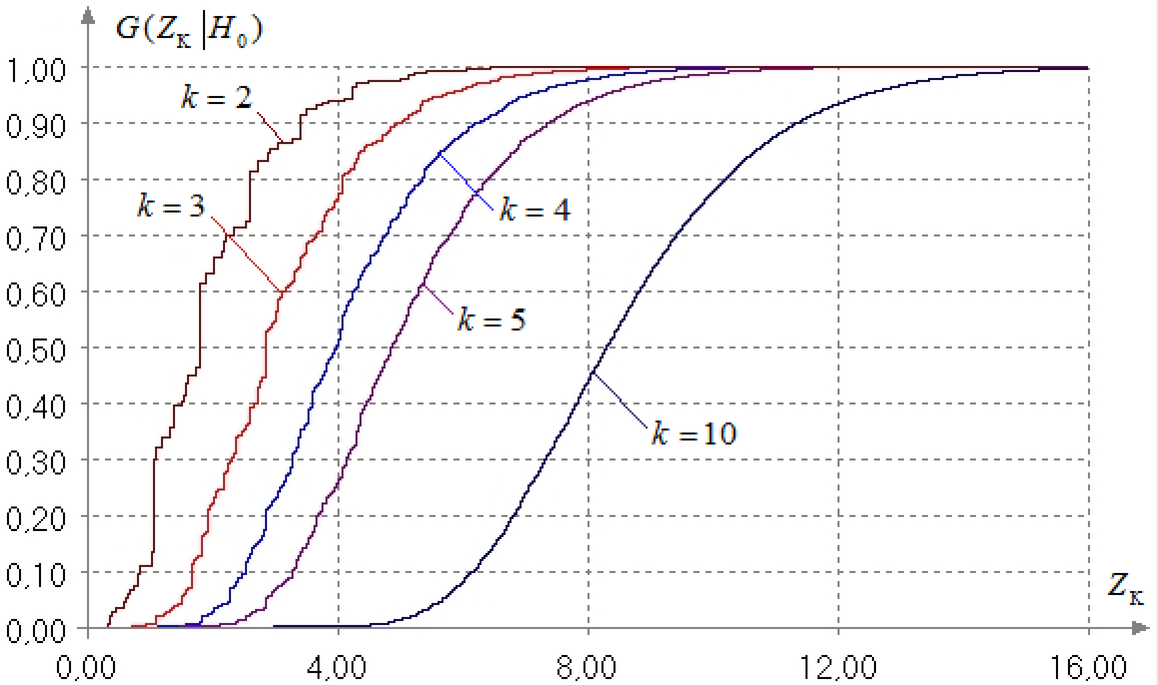


Рис. 3. Зависимость распределений статистики (6) от k при $n_i = 20$

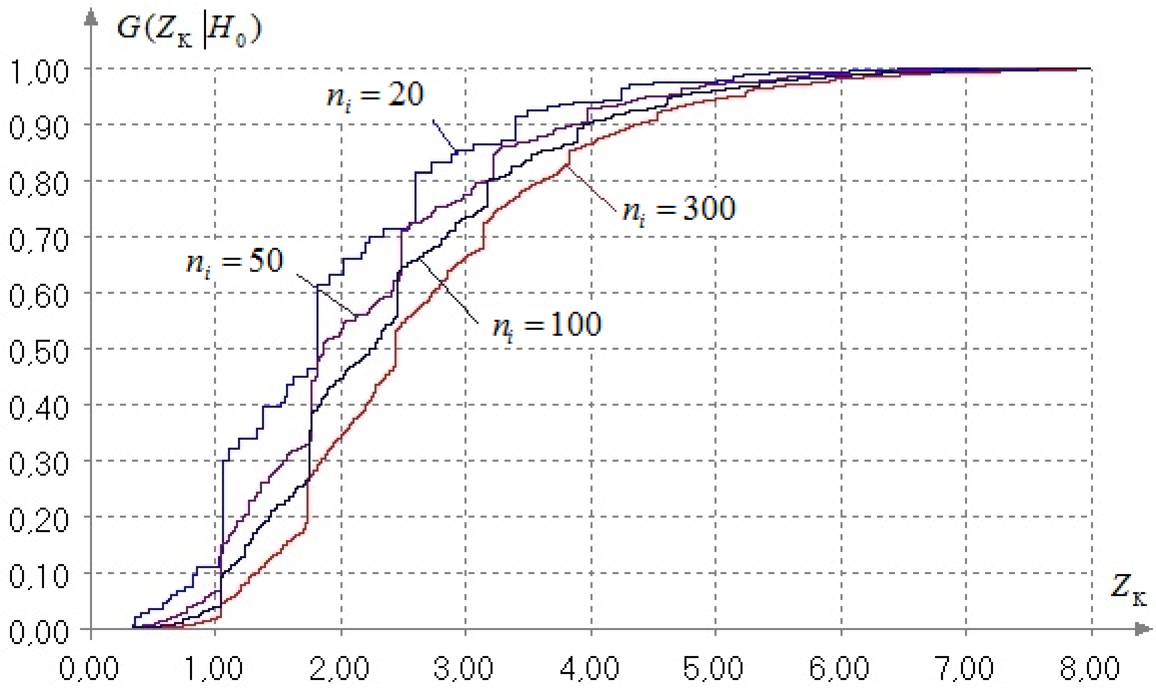


Рис. 3а. Зависимость распределений статистики (6) от объёмов выборок

Статистика Z_{A} **критерия однородности Жанга** определяется выражением [14]:

$$Z_{\text{A}} = -\sum_{m=1}^{n}\sum_{i=1}^{k} n_i \frac{F_{i,m}\ln F_{i,m} + \left(1-F_{i,m}\right)\ln\left(1-F_{i,m}\right)}{(m-0.5)(n-m+0.5)}, \quad (7)$$

где F_m и $F_{i,m}$ вычисляются, как определено выше.

Критерий левосторонний: проверяемая гипотеза H_0 отклоняется при **малых** значениях статистики (7). Распределения статистики зависят от n_i и от k.

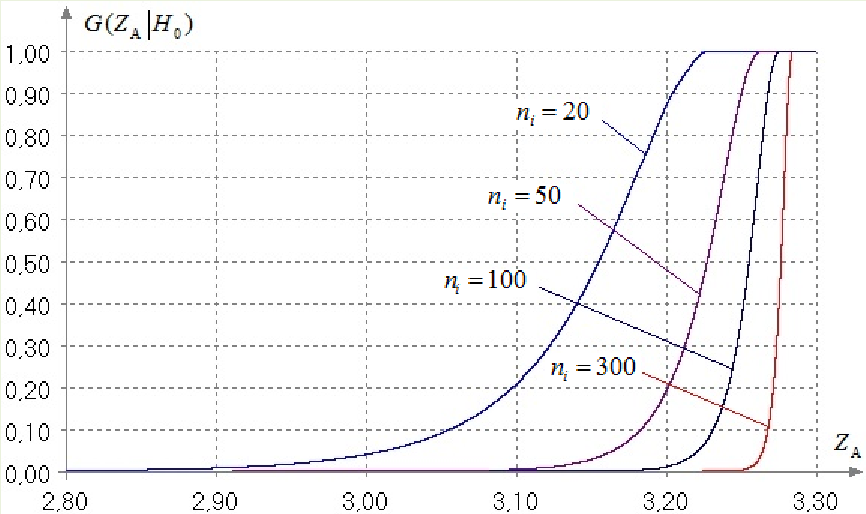


Рис. Зависимость распределений статистики (7) от объёмов выборок

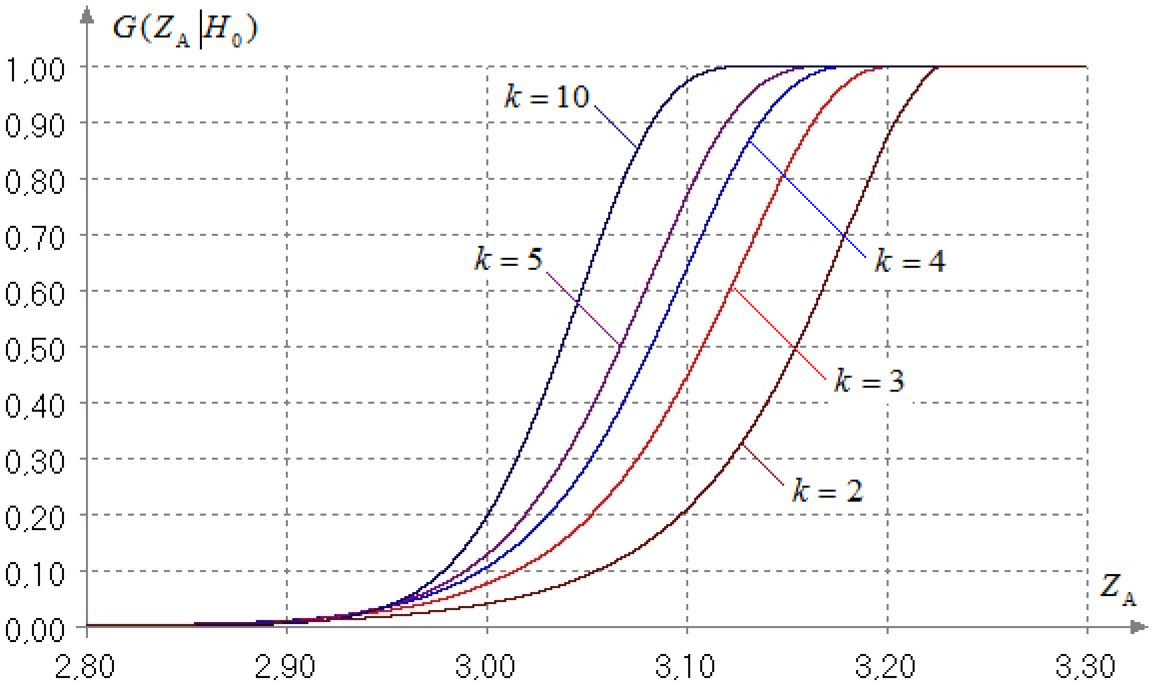


Рис. Зависимость распределений статистики (7) от k при $n_i = 20$

Статистика Z_C **критерия однородности Жанга** k выборок вычисляется в соответствии с выражением [14]:

$$Z_C = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left(\frac{n_i}{j-0.5}-1\right)\ln\left(\frac{n}{R_{i,j}-0.5}-1\right). \qquad (8)$$

Критерий также левосторонний: проверяемая гипотеза H_0 отклоняется при **малых** значениях статистики (8). Распределения статистики зависят от n_i и от k, зависимость от k показана на рис. 4.

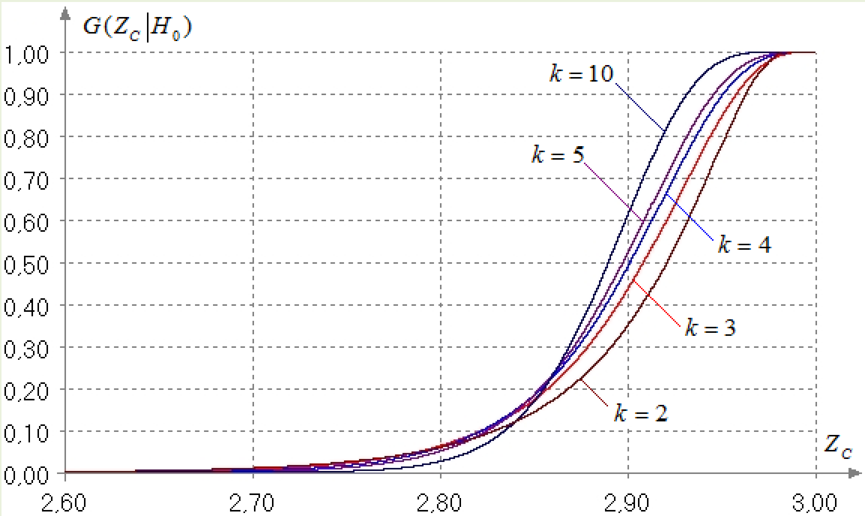


Рис. 4. Зависимость распределений статистики (8) от k при $n_i = 20$

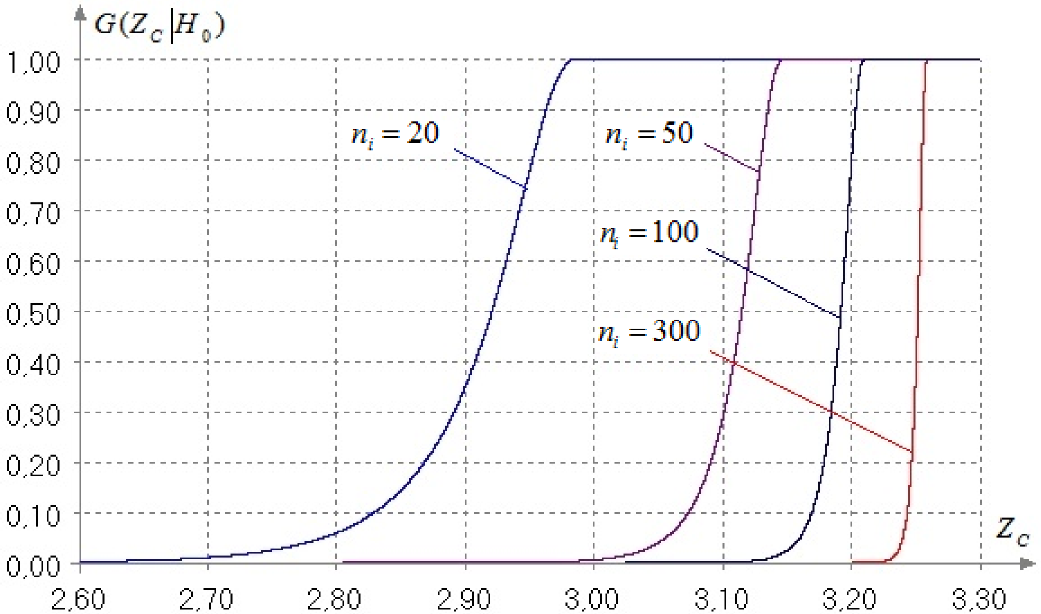


Рис. Зависимость распределений статистики (8) от объёмов выборок

Отсутствие информации о законах распределения статистик и таблиц критических значений в современных условиях не является серьёзным недостатком критериев Жанга, так как в программном обеспечении, осуществляющем поддержку применения критериев, несложно организовать вычисление достигнутых уровней значимостей p_{value}, используя методы статистического моделирования.

2. Двухвыборочные критерии при анализе k выборок

Различные подходы к построению k -выборочных аналогов критериев Смирнова, Лемана–Розенблатта и Андерсона–Дарлинга рассматривались в работе [45].

k -выборочный вариант критерия Колмогорова–Смирнова, основанный на таком подходе, был построен в [29] и рассматривается в последовательных изданиях книги [30].

На использовании такого же подхода построен k -выборочный критерий Андерсона–Дарлинга [83], рассмотренный выше, для которого нами были построены модели предельных распределений.

В этих критериях, так же как и в критериях однородности Жанга, строится объединённая выборка, а статистики измеряют отклонение эмпирических распределений отдельных выборок от эмпирического распределения, построенного по совокупности анализируемых выборок.

Возможен другой путь. Для анализа k выборок можно к каждой паре выборок применить двухвыборочный критерий со статистикой S (всего $(k-1)k/2$ вариантов), а решение принимать по совокупности результатов.

В качестве статистики такого k-выборочного критерия (в случае использования правостороннего двухвыборочного критерия) можно рассмотреть, например, статистику вида

$$S_{\max} = \max_{\substack{1 \le i \le k \\ i < j \le k}} \{S_{i,j}\}, \tag{2.17}$$

где $S_{i,j}$ – значения статистик используемого двухвыборочного критерия, полученные при анализе i-й и j-й выборок.

Проверяемая гипотеза H_0 будет отклоняться при **больших** значениях статистики $S_{\max}$.

Преимуществом такого рода критерия является и то, что в результате будет определена пара выборок, различие между которыми оказывается наиболее значимым с позиций используемого двухвыборочного критерия.

В качестве $S_{i,j}$ можно использовать статистики двухвыборочных критериев Смирнова (лучше в модифицированном виде), Лемана–Розенблата, Андерсона–Дарлинга. В этом случае распределения соответствующих статистик $S_{\max}$ сходятся к некоторым предельным, модели которых могут быть найдены по результатам статистического моделирования.

2.1. k –выборочный критерий Смирнова (max)

В качестве $S_{i,j}$ в (2.17) в этом случае будет рассматриваться модификация статистики Смирнова (2.3), распределение которой всегда ближе к предельному распределению Колмогорова $K(S)$. Статистику $S_{\max}$ в этом случае будем обозначать как $S_{\max}^{Sm}$.

При равных объёмах сравниваемых выборок распределения статистики $S_{\max}^{Sm}$ (как и в двухвыборочном варианте) обладают существенной дискретностью (см. рис. 2.28) и отличаются от асимптотических (предельных) распределений (см. рис. 2.29). Если есть такая возможность, то предпочтительней в качестве n_i выбирать взаимно простые числа, тогда распределения $G(S|H_0)$ статистики $S_{\max}^{Sm}$ практически не будут отличаться от асимптотических.

Верхние критические значения для статистики $S_{\max}^{Sm}$, построенные по эмпирическим распределениям статистик, полученным методом Монте–Карло при количестве имитационных экспериментов $N=10^6$, представлены в таблице 2.20, а в таблице 2.21 приведены построенные модели асимптотических (предельных) распределений статистики $S_{\max}^{Sm}$ при числе сравниваемых выборок $k=3\div11$.

Представленные в таблице 2.21 модели $\mathrm{B_{III}}(\theta_0,\theta_1,\theta_2,\theta_3,\theta_4)$ бета-распределения 3-го рода (2.13) с приведенными значениями параметров, позволяют по значениям статистики, вычисленным в соответствии с соотношением (2.17) с использованием в качестве $S_{i,j}$ статистики Смирнова (2.1) или её модификации (2.3), находить оценки p_{value} при соответствующем числе k сравниваемых выборок.

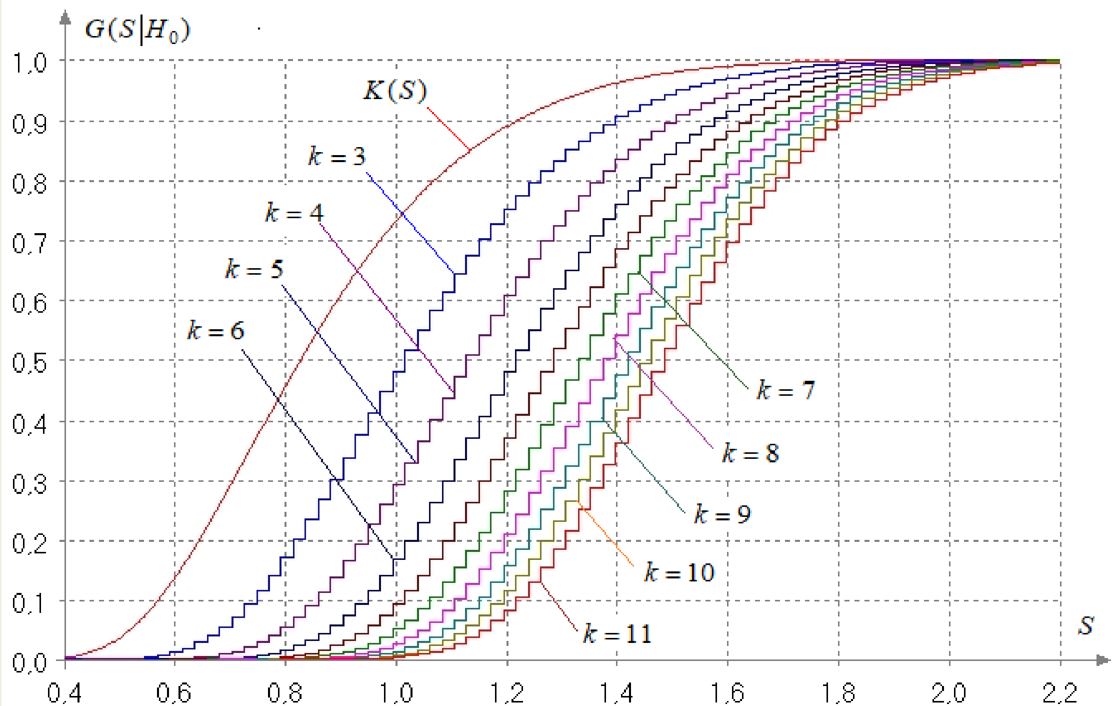


Рис. 2.28. Распределения статистики $S_{\max}^{Sm}$ k -выборочного критерия Смирнова при $n_i = 1000$, $i = \overline{1, k}$

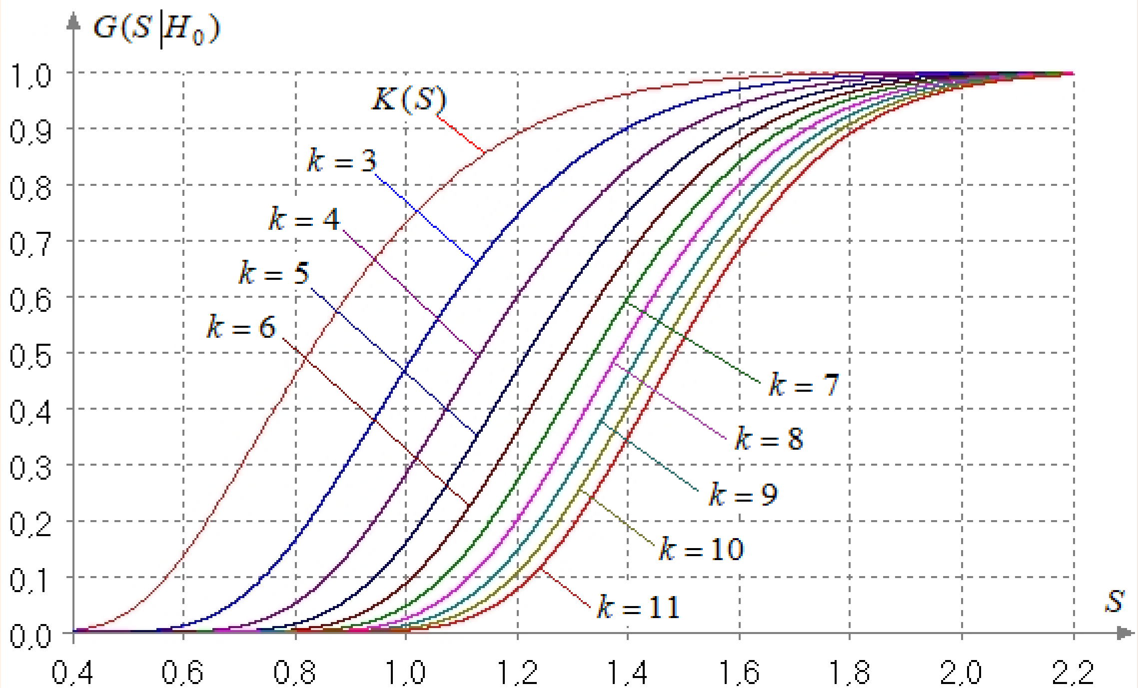


Рис. 2.29. Асимптотические распределения статистики $S_{\max}^{Sm}$ k -выборочного критерия Смирнова

Т а б л и ц а 2.20

Верхние критические значения статистики $S_{\max}^{Sm}$ Смирнова

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	1.0192	1.2238	1.3581	1.4802	1.6276
3	1.2059	1.4006	1.5266	1.6396	1.7782
4	1.3194	1.5070	1.6278	1.7372	1.8701
5	1.4002	1.5822	1.6997	1.8057	1.9329
6	1.4620	1.6396	1.7545	1.8573	1.9833
7	1.5127	1.6867	1.7989	1.9009	2.0245
8	1.5546	1.7264	1.8362	1.9362	2.0596
9	1.5911	1.7598	1.8682	1.9672	2.0879
10	1.6224	1.7893	1.8969	1.9940	2.1148
11	1.6506	1.8156	1.9217	2.0182	2.1371

Таблица 2.21

Модели предельных распределений статистики $S_{\max}^{Sm}$

k	Модель
2	$K(S)$
3	B_{III} (6.3274, 6.6162, 2.8238, 2.4073, 0.4100)
4	B_{III} (7.2729, 7.2061, 2.6170, 2.3775, 0.4740)
5	B_{III} (7.1318, 7.3365, 2.4813, 2.3353, 0.5630)
6	B_{III} (7.0755, 8.0449, 2.3163, 2.3818, 0.6320)
7	B_{III} (7.7347, 8.6845, 2.3492, 2.4479, 0.6675)
8	B_{III} (7.8162, 8.9073, 2.2688, 2.4161, 0.7120)
9	B_{III} (7.8436, 8.8805, 2.1696, 2.3309, 0.7500)
10	B_{III} (7.8756, 8.9051, 2.1977, 2.3280, 0.7900)
11	B_{III} (7.9122, 9.0411, 2.1173, 2.2860, 0.8200)

2.6.2. k –выборочный критерий Лемана–Розенблатта (max)

В качестве $S_{i,j}$ в статистике $S_{\max}^{LR}$ вида (2.17) в этом случае используется статистика (2.4) Лемана–Розенблатта. Зависимость распределений статистики при справедливости H_0 от числа выборок иллюстрирует рис. 2.30.

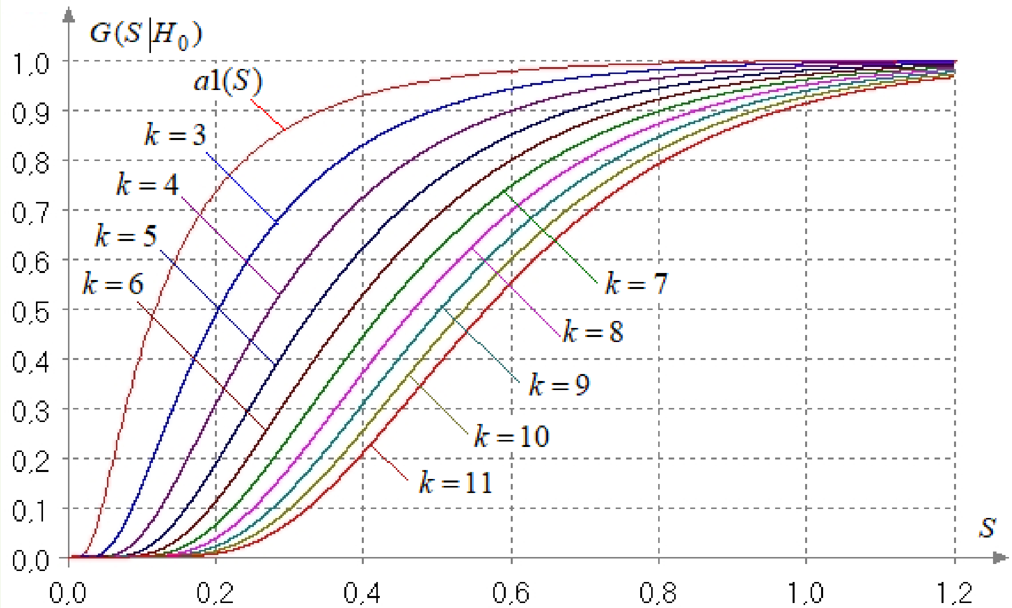


Рис. 2.30. Распределения статистики $S_{\max}^{LR}$ k -выборочного критерия Лемана–Розенблатта

Верхние критические значения для статистики $S_{\max}^{LR}$, построенные по результатам статистического моделирования при количестве имитационных экспериментов $N = 10^6$, представлены в таблице 2.24.

Построенные модели асимптотических (предельных) распределений статистики $S_{\max}^{LR}$ при числе сравниваемых выборок $k = 3 \div 11$ представлены в таблице 2.25.

Таблица 2.24

Верхние критические значения статистики $S_{\max}^{LR}$ Лемана–Розенблатта

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	0.2094	0.3473	0.4614	0.5806	0.7435
3	0.3306	0.4995	0.6283	0.7581	0.9308
4	0.4206	0.6050	0.7413	0.8770	1.0550
5	0.4924	0.6856	0.8267	0.9653	1.1429
6	0.5521	0.7512	0.8959	1.0365	1.2175
7	0.6037	0.8072	0.9524	1.0947	1.2774
8	0.6481	0.8550	1.0015	1.1444	1.3298
9	0.6876	0.8976	1.0457	1.1902	1.3781
10	0.7237	0.9355	1.0858	1.2303	1.4171
11	0.7563	0.9703	1.1214	1.2655	1.4535

В данном случае наилучшими моделями оказались распределения Sb–Джонсона с плотностью

$$f(x) = \frac{\theta_1\theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)} \exp\left\{ -\frac{1}{2}\left[\theta_0 - \theta_1 \ln \frac{x-\theta_3}{\theta_2+\theta_3-x} \right]^2 \right\}$$

при конкретных значениях параметров этого закона, обозначенного в таблице 2.25 как $Sb(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$. Представленные модели позволяют по значениям статистики $S_{\max}^{LR}$ при соответствующем числе k сравниваемых выборок находить оценки p_{value}.

Таблица 2.25

Модели предельных распределений статистики $S_{\max}^{LR}$

k	Модель
2	$a1(t)$
3	Sb(3.2854, 1.2036, 3.0000, 0.0215)
4	Sb(2.5801, 1.2167, 2.2367, 0.0356)
5	Sb(3.1719, 1.4134, 3.1500, 0.0320)
6	Sb(2.9979, 1.4768, 2.9850, 0.0380)
7	Sb(3.2030, 1.5526, 3.4050, 0.0450)
8	Sb(3.2671, 1.6302, 3.5522, 0.0470)
9	Sb(3.4548, 1.7127, 3.8800, 0.0490)
10	Sb(3.4887, 1.7729, 3.9680, 0.0510)
11	Sb(3.4627, 1.8168, 3.9680, 0.0544)

2.6.3. k –выборочный критерий Андерсона–Дарлинга (max)

В статистике $S_{\max}^{AD}$ вида (2.17) качестве $S_{i,j}$ используется статистика (2.7) Андерсона–Дарлинга. Зависимость распределений статистики $S_{\max}^{AD}$ при справедливости H_0 от числа выборок иллюстрирует рис. 2.31.

Верхние критические значения для статистики $S_{\max}^{AD}$, построенные по результатам статистического моделирования при количестве имитационных экспериментов $N = 10^6$, представлены в таблице 2.28.

Для распределений $G(S_{\max}^{AD} | H_0)$ также построены модели асимптотических (предельных) распределений статистики $S_{\max}^{AD}$ для числа сравниваемых выборок $k = 3 \div 11$, которые представлены в таблице 2.29. В этом случае лучшими моделями оказались бета-распределения 3-го рода (2.13), которые в виде $\mathrm{B_{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ с конкретными значениями параметров приведены в таблице 2.29 и могут использоваться для оценки p_{value} при k сравниваемых выборках.

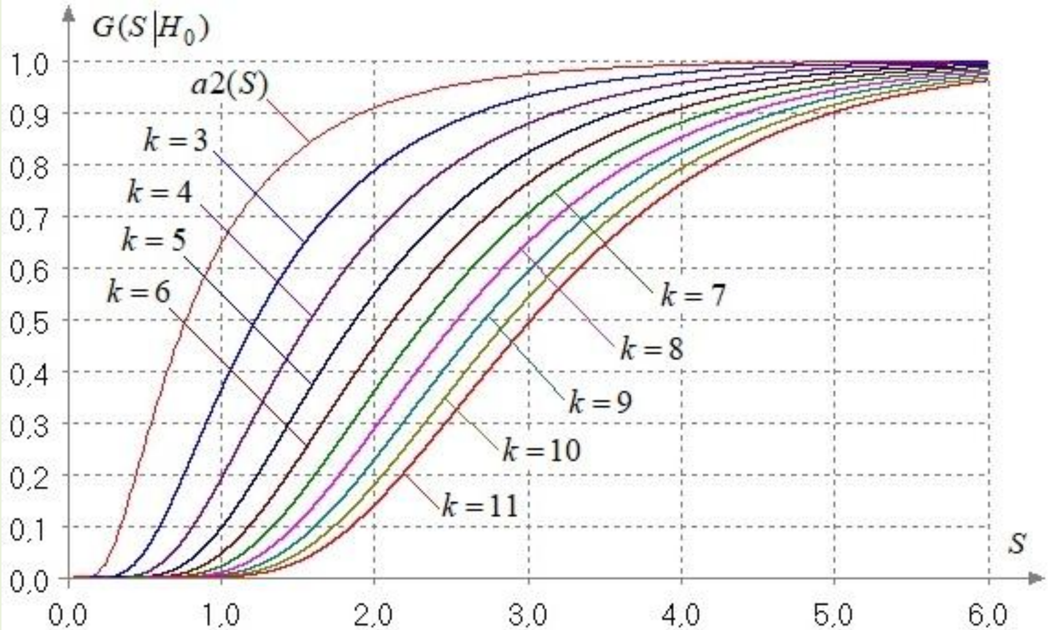


Рис. 2.31. Распределения статистики $S_{\max}^{AD}$ k -выборочного критерия Андерсона–Дарлинга

Таблица 2.28

Верхние критические значения статистики $S_{\max}^{AD}$ Андерсона–Дарлинга

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	1.2479	1.9330	2.4924	3.0775	3.8781
3	1.8535	2.6796	3.31215	3.95176	4.7924
4	2.2990	3.1966	3.8682	4.5368	5.4076
5	2.6514	3.5948	4.2877	4.9686	5.8472
6	2.9431	3.9187	4.6292	5.3175	6.2089
7	3.1943	4.1950	4.9097	5.6063	6.5118
8	3.4135	4.4292	5.1501	5.8531	6.7710
9	3.6094	4.6395	5.3660	6.0733	7.0076
10	3.7867	4.8270	5.5616	6.2720	7.2042
11	3.9466	4.9974	5.7384	6.4512	7.3837

Таблица 2.29

Модели предельных распределений статистики $S_{\max}^{AD}$

k	Модель
2	$a2(t)$
3	B_{III} (4.4325, 2.7425, 12.1134, 8.500, 0.1850)
4	B_{III} (5.2036, 3.2160, 10.7792, 10.000, 0.2320)
5	B_{III} (5.7527, 3.3017, 9.7365, 10.000, 0.3000)
6	B_{III} (5.5739, 3.4939, 7.7710, 10.000, 0.3750)
7	B_{III} (6.4892, 3.6656, 8.0529, 10.500, 0.3920)
8	B_{III} (6.3877, 3.8143, 7.3602, 10.800, 0.4800)
9	B_{III} (6.7910, 3.9858, 7.1280, 11.100, 0.5150)
10	B_{III} (6.7533, 4.2779, 6.6457, 11.700, 0.5800)
11	B_{III} (7.1745, 4.3469, 6.6161, 11.800, 0.6100)

2.7. Критерий однородности χ^2

Пусть имеется k выборок $x_{i1}, x_{i2}, \ldots, x_{in_i}$, $i = \overline{1,k}$, непрерывных случайных величин и $n = \sum_{i=1}^{k} n_i$. Общая область, которой принадлежат выборки, разбивается на r интервалов (групп). Пусть η_{ij} – количество элементов i-й выборки, попавших в j-й интервал, тогда $n_i = \sum_{j=1}^{r} \eta_{ij}$.

Статистика критерия однородности χ^2 имеет вид

$$\chi^2 = n \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{(\eta_{ij} - \nu_j n_i / n)^2}{\nu_j n_i} = n \left(\sum_{i=1}^{k} \sum_{j=1}^{r} \frac{\eta_{ij}^{\;2}}{\nu_j n_i} - 1 \right), \qquad (2.19)$$

где $\nu_j = \sum_{l=1}^{k} \eta_{lj}$ – общее число элементов всех выборок, попавших j-й интервал.

Асимптотическим распределением статистики (2.19) является χ^2-распределение с числом степеней свободы $(k-1)(r-1)$.

2.8. Примеры применения

Рассмотрим применение рассмотренных в разделе критериев проверки однородности законов на примере анализа 3-х нижеприведенных выборок, каждой объёмом в 40 наблюдений.

0.321	0.359	-0.341	1.016	0.207	1.115	1.163	0.900	-0.629	-0.524
-0.528	-0.177	1.213	-0.158	-2.002	0.632	-1.211	0.834	-0.591	-1.975
-2.680	-1.042	-0.872	0.118	-1.282	0.766	0.582	0.323	0.291	1.387
-0.481	-1.366	0.351	0.292	0.550	0.207	0.389	1.259	-0.461	-0.283

0.890	-0.700	0.825	1.212	1.046	0.260	0.473	0.481	0.417	1.825
1.841	2.154	-0.101	1.093	-1.099	0.334	1.089	0.876	2.304	1.126
-1.134	2.405	0.755	-1.014	2.459	1.135	0.626	1.283	0.645	1.100
2.212	0.135	0.173	-0.243	-1.203	-0.017	0.259	0.702	1.531	0.289

0.390	0.346	1.108	0.352	0.837	1.748	-1.264	-0.952	0.455	-0.072
-0.054	-0.157	0.517	1.928	-1.158	-1.063	-0.540	-0.076	0.310	-0.237
-1.109	0.732	2.395	0.310	0.936	0.407	-0.327	1.264	-0.025	-0.007
0.164	0.396	-1.130	1.197	-0.221	-1.586	-0.933	-0.676	-0.443	-0.101

Эмпирические распределения, соответствующие данным выборкам, представлены на рис. 2.32

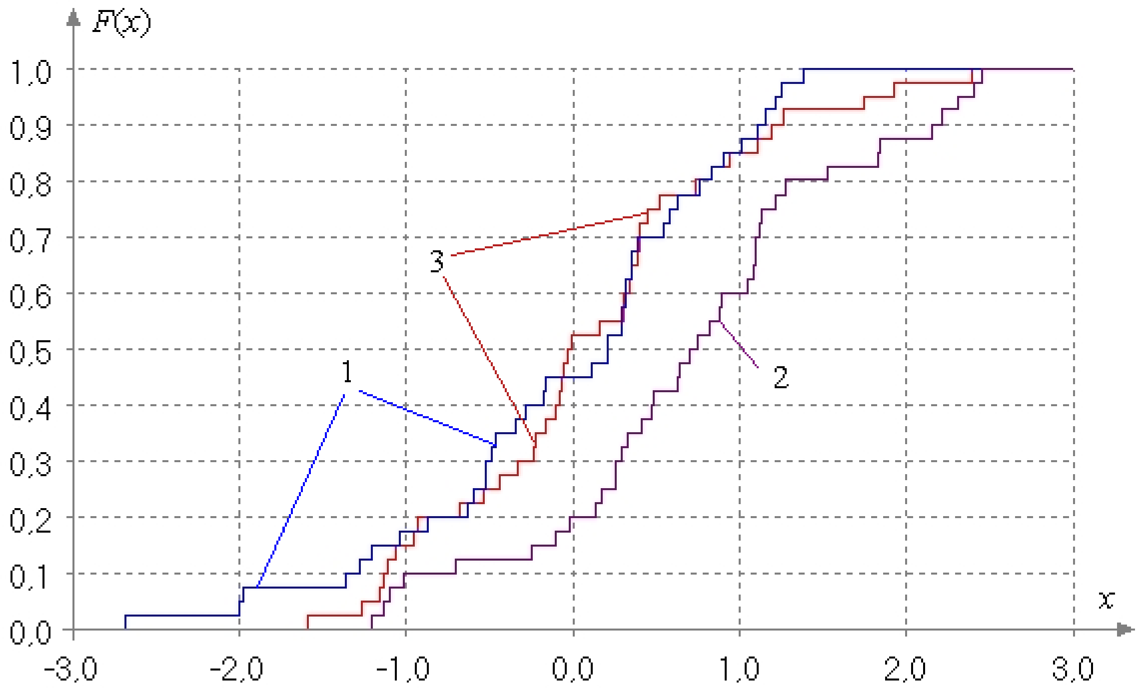


Рис. 2.32. Эмпирические распределения, соответствующие выборкам

Проверим гипотезу об однородности 1-й и 2-й выборок.

В таблице 2.35 приведены результаты проверки: значения статистик критериев и достигнутые уровни значимости p_{value}.

Оценки p_{value} вычислялись по значению статистики в соответствии с распределением $a2(t)$ для критерия Андерсона–Дарлинга, в соответствии с распределением $a1(t)$ для критерия Лемана–Розенблатта, в соответствии с распределением $K(t)$ для критерия Смирнова, в соответствии с бета-распределением 3-го рода из таблицы 2.8 при $k=2$ для k-выборочного критерия Андерсона–Дарлинга. Распределения статистик (2.12), (2.13) и (2.14) критериев Жанга и оценки p_{value} находились в результате моделирования.

По приведенным оценкам p_{value} очевидно, что гипотеза об однородности по всем критериям должна быть отклонена.

Таблица 2.35

Результаты проверки однородности 1-й и 2-й выборок

Критерий	Статистика	p_{value}
Андерсона–Дарлинга	5.19801	0.002314
k-выборочный Андерсона–Дарлинга	5.66112	0.003260
Лемана–Розенблатта	0.9650	0.002973
Смирнова	1.56525	0.014893
Смирнова модифицированный	1.61386	0.010933
Жанга Z_A	2.99412	0.0007
Жанга Z_C	2.87333	0.0008
Жанга Z_K	5.58723	0.0150
χ^2, $r=10$	16.3254	0.06039
χ^2, $r=8$	13.3293	0.06448

В таблице 2.36 приведены результаты проверки гипотезы об однородности 1-й и 3-й выборок.

Здесь оценки p_{value} по всем критериям очень высокие, поэтому проверяемая гипотеза об однородности не должна отклоняться.

Таблица 2.36

Результаты проверки однородности 1-й и 3-й выборок

Критерий	Статистика	p_{value}
Андерсона–Дарлинга	0.49354	0.753415
k -выборочный Андерсона–Дарлинга	-0.68252	0.767770
Лемана–Розенблатта	0.0500	0.876281
Смирнова	0.447214	0.989261
Смирнова модифицированный	0.495824	0.966553
Жанга Z_A	3.1998	0.332
Жанга Z_C	3.07077	0.384
Жанга Z_K	1.7732	0.531
χ^2, $r = 10$	11.6508	0.23372
χ^2, $r = 8$	4.0000	0.7798

В таблице 2.37 показаны результаты проверки гипотезы об однородности трёх рассматриваемых выборок по k-выборочному критерию Андерсона–Дарлинга, по критериям Жанга и по критериям со статистиками $S_{\max}^{Sm}$, $S_{\max}^{LR}$, $S_{\max}^{AD}$.

Т а б л и ц а 2.37

Результаты проверки однородности 3-х выборок

Критерий	Статистика	p_{value}
k-выборочный Андерсона–Дарлинга	4.73219	0.0028
Жанга Z_{A}	3.02845	0.0016
Жанга Z_C	2.92222	0.0017
Жанга Z_{K}	7.00231	0.0218
max Андерсона–Дарлинга	5.19801	0.0064
max Лемана–Розенблатта	0.9650	0.0094
max Смирнова модифицированный	1.72566	0.0144
χ^2, $r = 10$	25.556	0.1104
χ^2, $r = 8$	19.200	0.1574

В этом случае оценка p_{value}:

– для критерия Андерсона–Дарлинга вычислялась в соответствии с бета-распределением 3-го рода из таблицы 2.8 при $k=3$;
– для критериев Жанга – на основании статистического моделирования, проведенного в интерактивном режиме, при числе имитационных экспериментов $N=10^6$;
– для критерия со статистикой $S_{\max}^{Sm}$ оценка p_{value} при $k=3$ вычислялась в соответствии с бета-распределением 3-го рода из таблицы 2.21;
– для критерия со статистикой $S_{\max}^{LR}$ – в соответствии с распределением Sb-Джонсона из таблицы 2.25;
– для критерия со статистикой $S_{\max}^{AD}$ – в соответствии с бета-распределением 3-го рода из таблицы 2.29.

Заметим, что критерии Андерсона–Дарлинга $S_{\max}^{AD}$ и Лемана–Розенблатта $S_{\max}^{LR}$ зафиксировали максимальное отклонении между 1-й и 2-й выборками, а критерий Смирнова $S_{\max}^{Sm}$ – между 2-й и 3-й. Общий результат показывает, что проверяемая гипотеза об однородности 3-х выборок должна быть отклонена.

Можно обратить внимание на существенную зависимость результатов проверки по критерию однородности χ^2 от выбираемого числа интервалов r.

В данном случае результаты проверки были достаточно предсказуемы, так как 1-я и 3-я выборки были смоделированы в соответствии со стандартным нормальным законом, а полученные значения псевдослучайных величин округлены до 3-х значащих цифр после десятичной точки. В то время как 2-я выборка получена в соответствии с нормальным законом с параметром сдвига 0.5 и стандартным отклонением 1.1.

2.8. Сравнительный анализ мощности критериев

В случае k выборок относительно изменения параметра сдвига порядок предпочтения критериев имеет вид [136]:

$$S_{\max}^{AD} \succ \text{Андерсона–Дарлинга} \succ S_{\max}^{LR} \succ S_{\max}^{Sm} \succ \text{Жанга } Z_C \succ$$
$$\text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_{\text{K}} \succ \chi^2 .$$

Относительно изменения параметра масштаба –

$$\text{Жанга } Z_C \succ \text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_{\text{K}} \succ \text{Андерсона–Дарлинга} \succ$$
$$\chi^2 \succ S_{\max}^{AD} \succ S_{\max}^{Sm} \succ S_{\max}^{LR} .$$

При этом критерии Жанга со статистиками Z_{A} и Z_C практически эквивалентны по мощности, а критерий Андерсона–Дарлинга заметно уступает критериям Жанга.

Относительно ситуации, когда три выборки принадлежат нормальному закону, а четвёртая – логистическому, критерии располагаются по мощности в следующем порядке:

$$\text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\text{K}} \succ \chi^2 \succ \text{Андерсона–Дарлинга} \succ S_{\max}^{Sm} \succ S_{\max}^{AD} \succ S_{\max}^{LR} .$$

КРИТЕРИИ ПРОВЕРКИ ОДНОРОДНОСТИ СРЕДНИХ

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

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

при конкурирующей гипотезе

$$H_1\colon\ \mu_{i_1} \neq \mu_{i_2}\,,$$

где неравенство выполняется хотя бы для одной 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, критерий Ван дер Вардена, Критерий Фишера–Йэйтса–Терри–Гёфдинга, k - выборочный критерий Ван дер Вардена.

Основным предположением, обусловливающим возможность применения параметрических критериев, является принадлежность анализируемых выборок нормальному закону. Непараметрические критерии свободны от этого требования.

Параметрические критерии однородности средних устойчивы к нарушению предположения о нормальности. Лишь асимметричность законов распределения или наличие “тяжелых хвостов” приводит к существенному изменению предельных распределений параметрических критериев.

С устойчивостью критерия связана причина того, что параметрические критерии, как правило, имеют незначительное преимущество в мощности перед непараметрическими критериями однородности средних.

3.1. Параметрические критерии однородности средних

3.1.1. 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

$$z = \left(\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2\right) \Big/ \sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}\,, \qquad (3.1)$$

где $\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$, n_i – объем i -й выборки, $i = 1, 2$.

В случае принадлежности наблюдений (ошибок измерений) нормальным законам статистика z **подчиняется стандартному нормальному закону**. Если в статистике (3.1) опустить μ_1 и μ_2 , то будет проверяться гипотеза о равенстве математических ожиданий.

Критерий двусторонний. Проверяемая гипотеза отклоняется при больших по модулю значениях статистики (3.1).

3.1.2. Критерий Стьюдента

Это критерий проверки гипотезы о равенстве математических ожиданий при неизвестных, но равных дисперсиях. Критерий предусматривает вычисление статистики t:

$$t = \left(\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2\right) \Bigg/ \sqrt{\left[\frac{n_1 + n_2}{n_1 n_2}\right]\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]}, \qquad (3.2)$$

где

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

Аналогично, если в статистике (3.2) опустить μ_1 и μ_2, то проверяется гипотеза о равенстве математических ожиданий.

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

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

При неравных объемах выборок $n_1 \neq n_2$ статистика критерия имеет вид:

$$t = \left(\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2\right) \Big/ \sqrt{s_1^2/n_1 + s_2^2/n_2}\ . \tag{3.3}$$

Критерий также **двусторонний.** В случае нормального закона и справедливости гипотезы H_0 статистика (3.3) подчиняется распределению t_ν-Стьюдента с числом степеней свободы

$$\nu = \left(s_1^2/n_1 + s_2^2/n_2\right)^2 \Bigg/ \left[\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}\right]. \tag{3.4}$$

Задача, связанная с поиском распределения статистики (3.3), долгое время носила название проблемы Беренса–Фишера.

Иногда критерий называют критерием Крамера–Уэлча.

3.1.4. F-критерий однородности средних

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

Сумма квадратов отклонений по всем выборкам определяется соотношением

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

где

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

– среднее по всем выборкам.

Общая сумма Q_{kn} раскладывается на два компонента

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

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

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

Компонент Q_1, например, в задачах контроля качества 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

$$F = \frac{Q_1 / (k-1)}{Q_2 / [k(n-1)]} . \qquad (3.5)$$

Если все выборки извлекаются из нормальной генеральной совокупности, то при справедливости гипотезы H_0 статистка (3.5) подчиняется F_{ν_1,ν_2}-распределению Фишера со степенями свободы $\nu_1 = k-1$ и $\nu_2 = k(n-1)$.

Критерий правосторонний. Проверяемая гипотеза отклоняется при больших значениях статистики (3.5).

3.1.5. k-выборочный вариант критерия Стьюдента

В задачах метрологии предлагается развитие критерия, представленного в п. 3.1.3, на случай k выборок.

В случае анализа k выборок x_{ij}, $i=\overline{1,k}$, $j=\overline{1,n_i}$, для проверки гипотезы об однородности средних статистика критерия имеет вид

$$\chi^2 = \sum_{i=1}^{k} n_i \frac{(\bar{x}_i - y)^2}{s_i^2}, \qquad (3.6)$$

где $\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$, $s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$, $y = \left(\sum_{i=1}^{k} n_i \bar{x}_i / s_i^2\right) \Big/ \sum_{i=1}^{k} \frac{n_i}{s_i^2}$ – взвешенное среднее по всем выборкам.

Критерий правосторонний. Асимптотическим распределением статистики (3.6) является χ^2_{k-1}-распределение.

Распределение статистики (3.6) не отличается от χ^2_{k-1}-распределения лишь при $n_i \geq 215$.

Реально отличием можно пренебречь при $n_i \geq 100$.

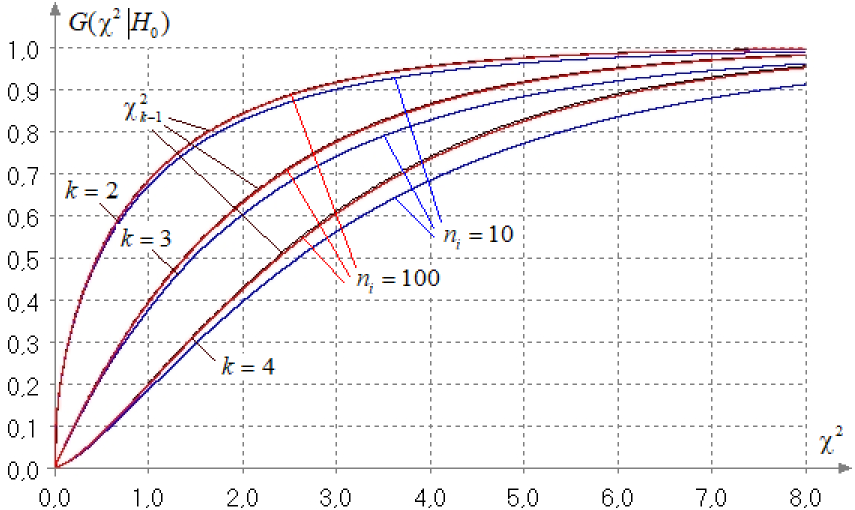


Рис. 3.2. Асимптотические и эмпирические распределения статистики (3.6) в случае справедливости гипотезы H_0 при $n_i = 10$ и $n_i = 100$ для $k = 2$ и $k = 3$

3.1.6. Об устойчивости параметрических критериев

Принадлежность выборок нормальному закону является основным предположением, обуславливающим возможность применения перечисленных выше параметрических критериев. Именно в предположении о нормальности имеют место указанные распределения статистик соответствующих критериев при справедливости проверяемой гипотезы H_0.

В то же время давно известно, что параметрические критерии, предназначенные для проверки гипотез об однородности средних анализируемых выборок, достаточно устойчивы к нарушению стандартного предположения о нормальности.

Распределения статистик (3.1)–(3.3), (3.5) и (3.6) при справедливой проверяемой гипотезе H_0 исследовались для различных законов, в частности, в случае принадлежности наблюдений обобщённому нормальному закону с плотностью

$$f(x) = \frac{\theta_2}{2\theta_1 \Gamma(1/\theta_2)} \exp\left\{ -\left(\frac{|x-\theta_0|}{\theta_1} \right)^{\theta_2} \right\} \qquad (3.7)$$

с различными значениями параметра формы θ_2. При $\theta_2 = 2$ выражение (3.7) дает плотность нормального закона распределения. При больших значениях θ_2 распределение (3.7) стремится к равномерному, при малых θ_2 получаем симметричные законы с «тяжелыми хвостами».

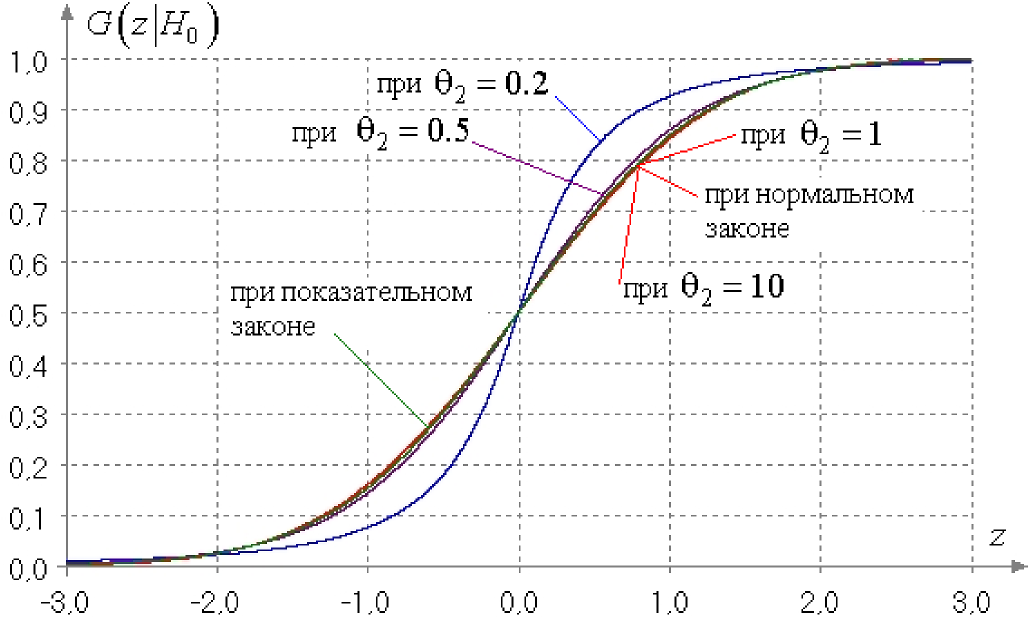


Рис. 3.3. Эмпирические распределения статистики (3.1) при различных законах распределения

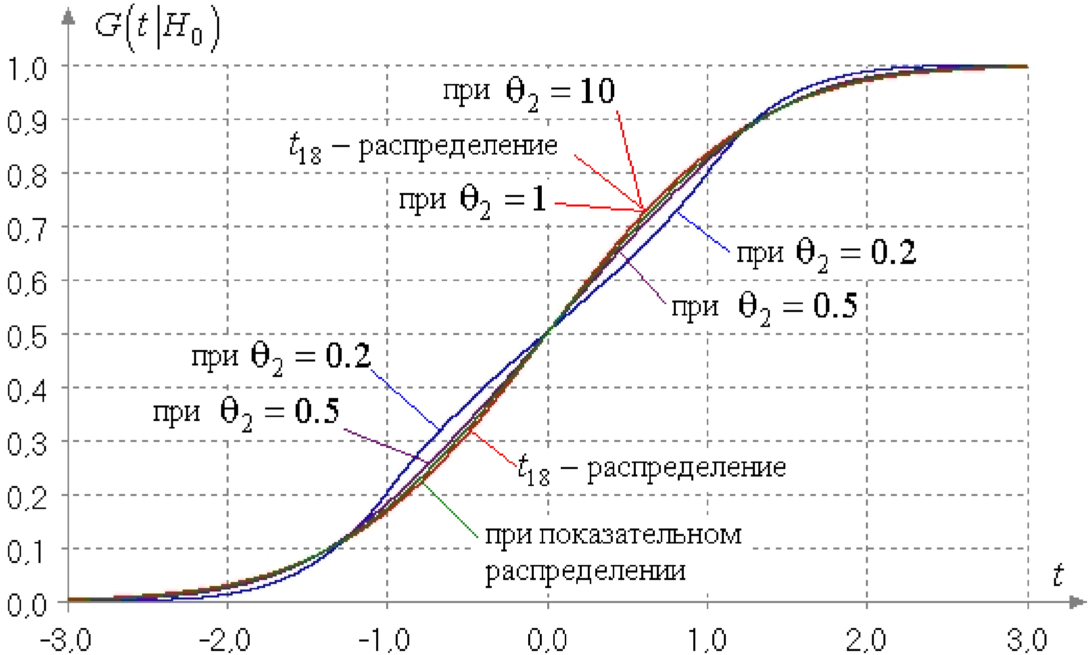


Рис. 3.4. Эмпирические распределения статистики (3.2) при различных законах распределения наблюдаемых величин и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 10$

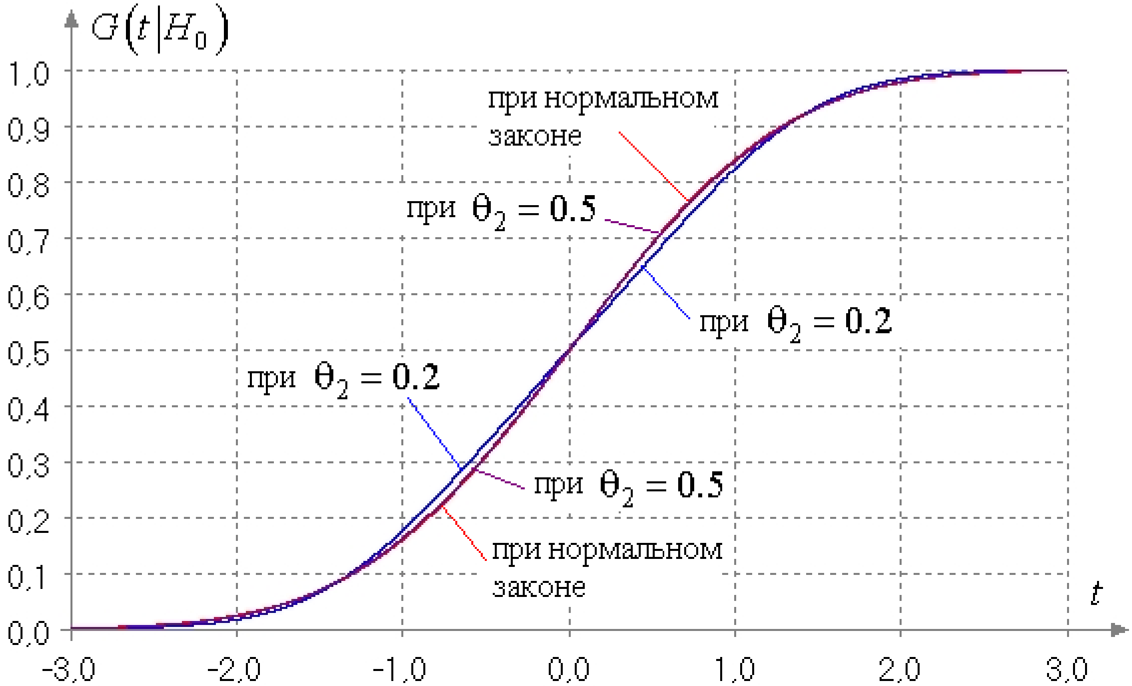


Рис. 3.5. Эмпирические распределения статистики (3.2) при различных законах распределения наблюдаемых величии и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 100$

3.2. Непараметрические критерии однородности средних

3.2.1. Критерии Уилкоксона и Манна–Уитни.

Ранговый критерий Манна и Уитни является непараметрическим аналогом t-критерия для сравнения двух средних значений непрерывных распределений. Критерий Манна и Уитни представляет собой развитие критерия Уилкоксона.

Для вычисления статистики Уилкоксона две независимые выборки объединяют в одну объёмом в n_1+n_2 значений и упорядочивают. По объединенной выборке определяют сумму рангов R_1, соответствующую элементам первой выборки, и сумму рангов второй R_2. Статистика критерия Уилкоксона имеет вид

$$U=\min\{U_1,U_2\}, \tag{3.8}$$

где

$$U_1=n_1n_2+n_1(n_1+1)/2-R_1, \qquad U_2=n_1n_2+n_2(n_2+1)/2-R_2.$$

Дискретные распределения U-статистики сильно зависят от размера выборок, что затрудняет применение критерия.

В критерии Манна–Уитни (Манна–Уитни–Уилкоксона) вместо U-статистики используется нормализованная статистика

$$\tilde{z}=\frac{U-n_1n_2/2}{\sqrt{n_1n_2(n_1+n_2+1)/12}}. \tag{3.9}$$

Дискретное распределение статистики (3.9) в случае справедливости проверяемой гипотезы H_0 хорошо приближается стандартным нормальным законом при $n_1+n_2>60$ (см. рис. 3.7), если объем каждой из выборок не слишком мал: $n_1\geq 8$, $n_2\geq 8$.

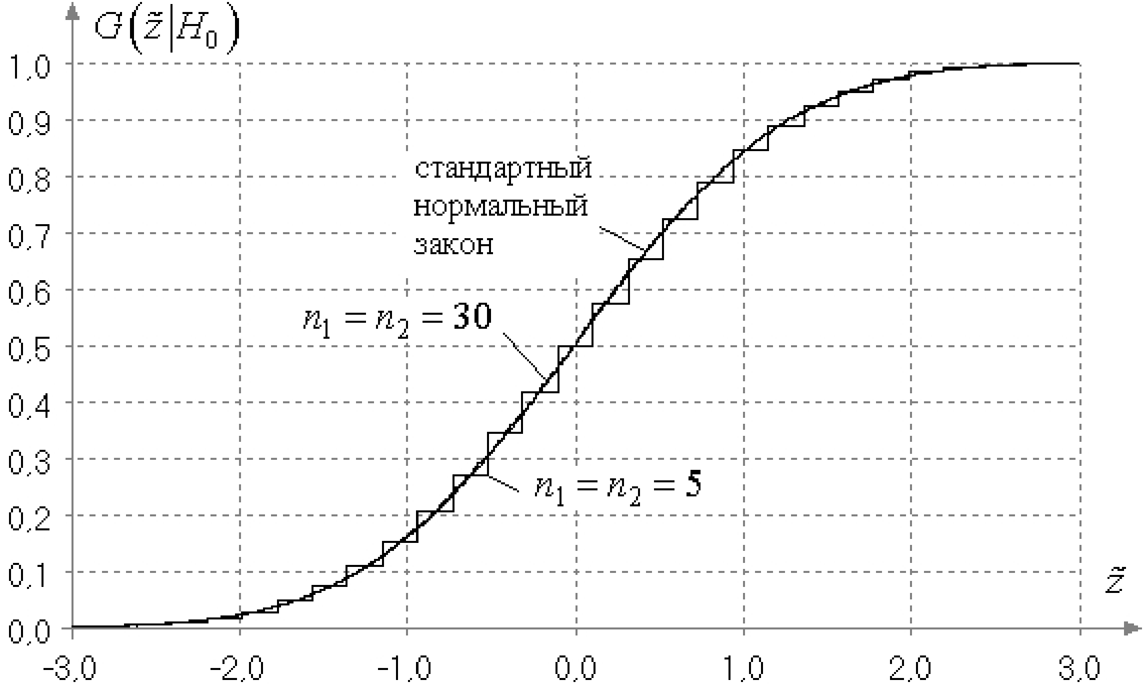


Рис. 3.7. Сходимость распределения статистики Манна–Уитни к стандартному нормальному закону

3.2.2. Критерий Краскела–Уаллиса

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

Статистика для проверки гипотезы H_0 имеет вид

$$H = \left[\frac{12}{n(n+1)}\right]\left[\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n+1) . \qquad (3.10)$$

Статистика H представляет собой дисперсию ранговых сумм.

При больших n_i и k в случае справедливости проверяемой гипотезы H_0 распределение статистики хорошо аппроксимируется χ^2_{k-1}-распределением. В описаниях критерия говорится, что χ^2_{k-1}-распределением практически можно пользоваться при $n_i \geq 5,\ k \geq 4$.

Исследования показали, что на практике в случае $k = 2$ можно пренебречь дискретностью при $n_i \geq 30$ (см. рис. 3.8).

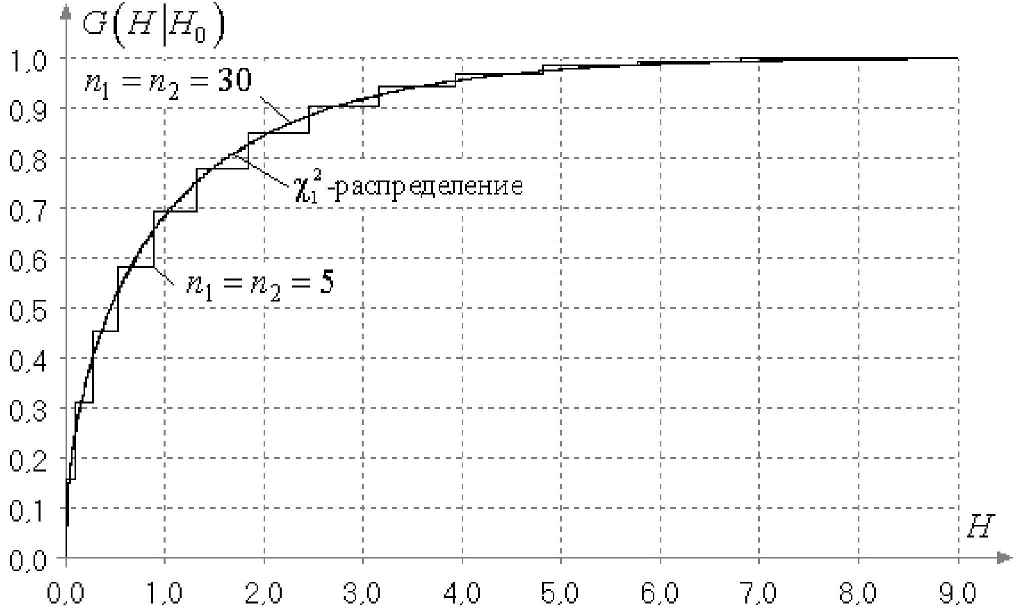


Рис. 3.8. Сходимость распределения статистики (3.10) Краскела–Уаллиса к χ_{k-1}^2-распределению при $k = 2$

3.2.3. Критерий Ван дер Вардена

Критерий предназначен для анализа 2-х выборок. Статистика непараметрического критерия Ван дер Вардена вычисляется в соответствии с выражением:

$$V = \sum_{i=1}^{n_2} u_{R_i/(n_1+n_2+1)}\,, \tag{3.11}$$

где u_γ – γ-квантиль стандартного нормального закона, R_i, $i=\overline{1,n_2}$ – ранг i-го наблюдения, например, как в (3.11), второй выборки в общем вариационном ряду объединенной выборки из n_1+n_2 наблюдений.

Считается, что при $n_1, n_2 \geq 20$ распределение статистики (3.11) удовлетворительно описывается нормальным законом с нулевым математическим ожиданием и дисперсией

$$D[V] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} u^2_{i/(n_1+n_2+1)}\,.$$

Нормализованная статистика

$$V^* = \frac{V}{\sqrt{D[V]}} \tag{3.12}$$

должна подчиняться **стандартному нормальному закону**.

Критерий двусторонний, проверяемая гипотеза отклоняется при больших по модулю значениях статистики (3.12).

Исследование распределений статистики (3.12) методами статистического моделирования показало, что при $n_1 + n_2 \geq 40$ отличием распределения $G(V^* | H_0)$ статистики (3.12) от стандартного нормального закона можно пренебречь.

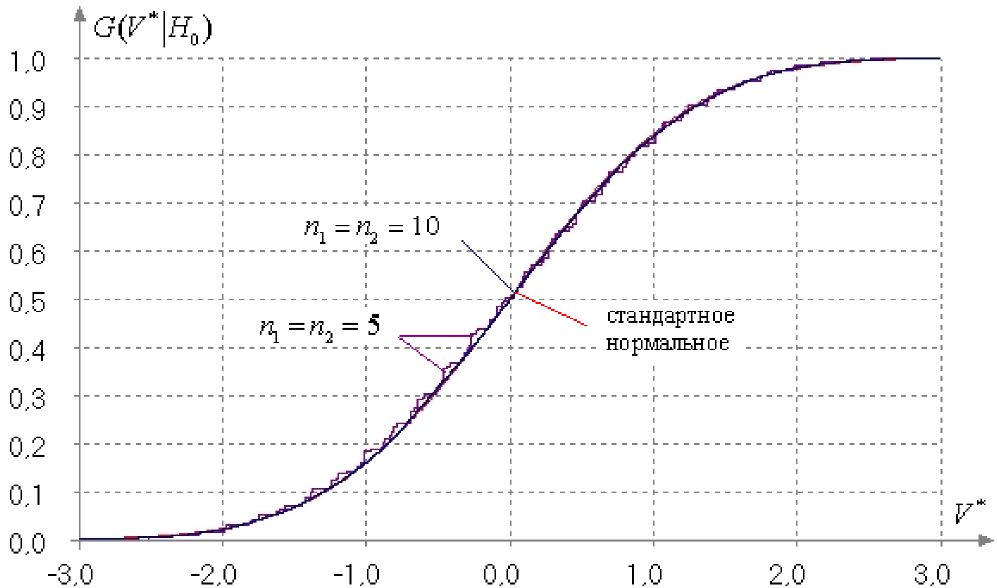


Рис. 3.9. Сходимость распределения нормализованной статистики (3.12) критерия Вандер Вардена к стандартному нормальному закону

Из непараметрических критериев однородности средних критерий Ван дер Вардена, по-видимому, является наиболее предпочтительным.

3.2.4. Критерий Фишера–Йэйтса–Терри–Гёфдинга

Этот критерий очень близок к критерию Ван дер Вардена. В качестве меток в критерии выбраны математические ожидания соответствующих порядковых статистик в выборке объёмом $n = n_1 + n_2$ из стандартного нормального закона. Статистика критерия имеет вид:

$$S = \sum_{i=1}^{n_2} u_{(R_i - 3/8)/(n+1/4)}, \tag{3.13}$$

где u_γ – γ-квантиль стандартного нормального закона, R_i, $i = \overline{1, n_2}$ – ранг i-го наблюдения, например, второй выборки в общем вариационном ряду объединенной выборки из $n_1 + n_2$ наблюдений.

Так же как и в случае критерия Ван дер Вардена, статистика (3.13) достаточно хорошо описывается нормальным законом с нулевым математическим ожиданием и дисперсией

$$D[S] = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \sum_{i=1}^{n_1+n_2} u^2_{(i-3/8)/(n_1+n_2+1/4)},$$

а нормализованная статистика

$$S^* = \frac{S}{\sqrt{D[S]}} \tag{3.14}$$

– стандартным нормальным законом.

3.2.5. Многовыборочный критерий Ван дер Вардена

Статистика критерия Ван дер Вардена для проверки гипотезы о равенстве средних по k выборкам имеет вид

$$T = (n-1)\sum_{i=1}^{k}\frac{1}{n_i}\left(\sum_{j=1}^{n_i} u_{R_{ij}/(n+1)}\right)^2 \Bigg/ \sum_{i=1}^{n} u_{i/(n+1)}^2 , \tag{3.15}$$

где $n=\sum_{i=1}^{k} n_i$, u_γ – γ-квантиль стандартного нормального закона, R_{ij}, $j=\overline{1,n_i}$, – ранг j-го элемента i-й выборки в вариационном ряду объединённой выборки объёма n.

При справедливости проверяемой гипотезы H_0 статистика (3.15) хорошо описывается χ^2_{k-1}-распределением. Отклонением распределения статистики от χ^2_{k-1}-распределения можно пренебречь при $n_i > 30$.

k-выборочный критерий Ван дер Вардена демонстрирует более высокую мощность по сравнению с критерием Краскела–Уаллиса.

3.3. Сравнительный анализ мощности критериев

Большая часть рассмотренных в руководстве критериев проверки гипотез об однородности средних предназначены для анализа двух выборок.

Относительно конкурирующей гипотезы H_1^1: $\mu_2 = \mu_1 + 0.1\sigma$ двухвыборочные критерии упорядочиваются следующим образом.

α	$n = 10$	$n = 20$	$n = 30$	$n = 50$	$n = 100$
z -критерий при известных дисперсиях					
0.1	0.108	0.117	0.126	0.142	0.184
t -критерий Стьюдента при неизвестных и равных дисперсиях					
0.1	0.108	0.117	0.126	0.141	0.183
t -критерий Стьюдента при неизвестных и неравных дисперсиях					
0.1	0.108	0.117	0.126	0.141	0.183
F-критерий, Стьюдента k-выборочный					
0.1	0.108	0.116	0.125	0.141	0.183
Критерий Ван дер Вардена					
0.1	0.108	0.116	0.125	0.141	0.183
Многовыборочный критерий Ван дер Вардена					
0.1	0.108	0.116	0.125	0.141	0.182
H-критерий Краскела–Уаллиса					
0.1	0.113	0.118	0.126	0.140	0.179
$\tilde{z}$ -критерий Манна–Уитни					
0.1	0.102	0.113	0.124	0.138	0.179

В случае k-выборочных критериев при $k = 3, 5$ относительно конкурирующих гипотез, при которых $k-1$ выборке соответствовало математическое ожидание μ, а k-й выборке – $\mu + 0.2\sigma$ и $\mu + 0.5\sigma$ (при одинаковых объемах сравниваемых выборок $n_i = n$). Можно заметить, что с ростом числа сравниваемых выборок k-выборочный критерий Стьюдента начинает уступать не только F-критерию, но и критерию Ван дер Вардена.

Мощность критериев относительно H_1^2 : $\mu_k = \mu_1 + 0.2\sigma$

Критерий	α	k = 3		k = 5	
		n = 30	n = 50	n = 30	n = 50
F-критерий	0.1	0.191	0.254	0.175	0.232
	0.05	0.112	0.160	0.101	0.142
	0.01	0.031	0.052	0.026	0.043
Стьюдента k-выборочный	0.1	0.190	0.253	0.174	0.230
	0.05	0.112	0.159	0.099	0.140
	0.01	0.031	0.052	0.026	0.042
Многовыборочный критерий Ван дер Вардена	0.1	0.190	0.252	0.174	0.231
	0.05	0.111	0.159	0.100	0.141
	0.01	0.031	0.051	0.026	0.043
H-критерий Краскела–Уаллиса	0.1	0.187	0.247	0.172	0.225
	0.05	0.109	0.155	0.098	0.137
	0.01	0.030	0.049	0.025	0.041

Аналогично, возрастает преимущество в мощности критерия Ван дер Вардена по сравнению с критерием Краскела–Уаллиса.

3.4. Выводы по разделу

Исследования подтвердили устойчивость параметрических критериев проверки однородности математических ожиданий.

Это означает, что если закон распределения, которому соответствуют анализируемые выборки, отличается от нормального, но нет оснований полагать, что наблюдаемые величины принадлежат законам с «тяжелыми хвостами», то применение параметрических критериев со статистиками остается корректным.

По крайней мере, такая ситуация не приводит к существенным погрешностям при оценке достигнутого уровня значимости.

4. КРИТЕРИИ ПРОВЕРКИ ОДНОРОДНОСТИ ДИСПЕРСИЙ

Применение классических критериев проверки однородности дисперсий всегда сопряжено с вопросом: насколько полученные выводы корректны в данной конкретной ситуации?

Одним из основных предположений при построении этих критериев является принадлежность наблюдаемых случайных величин (погрешностей измерений) нормальному закону распределения.

При этом давно известно, что параметрические критерии однородности дисперсий чрезвычайно чувствительны к малейшим отклонениям наблюдаемых случайных величин от нормального закона. При нарушении данного предположения условные распределения статистик критериев при справедливости проверяемой гипотезы, как правило, сильно изменяются.

В критериях проверки однородности дисперсий проверяемая гипотеза о постоянстве дисперсий k выборок имеет вид

$$H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 ,$$

а конкурирующая с ней гипотеза

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

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

Для проверки однородности дисперсий могут использоваться классические параметрические критерии Бартлетта, Кокрена, Фишера, Хартли, Левене, Неймана–Пирсона (критерия отношения правдоподобия), О`Брайена, Линка, Ньюмана, Блиса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Z–критерий Оверолла–Вудворда и его модификация, Миллера, Лайарда.

Непараметрические (ранговые) критерии: Ансари–Бредли, Муда, Сижела–Тьюки, Кейпена, Клотца, k-выборочный критерий Флайне–Киллина.

4.1. Параметрические критерии однородности дисперсий

4.1.1. Критерий Бартлетта

Статистика критерия Бартлетта [97] вычисляется в соответствии с соотношением

$$\chi^2 = M\left[1+\frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{\nu_i}-\frac{1}{N}\right)\right]^{-1}, \qquad (4.4)$$

где

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

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

Если гипотеза H_0 верна, все $\nu_i > 3$ и выборки извлекаются из нормальной генеральной совокупности, то статистика (4.4) приближенно подчиняется χ^2_{k-1}-распределению. Если вычисленное значение статистики $\chi^{2*} > \chi^2_{1-\alpha,k-1}$, то проверяемая гипотеза отклоняется на заданном уровне значимости α.

При нормально распределенных результатах измерений распределение статистики (4.4) практически не зависит от изменения объемов выборок, но зависит при нарушении предположения о нормальности.

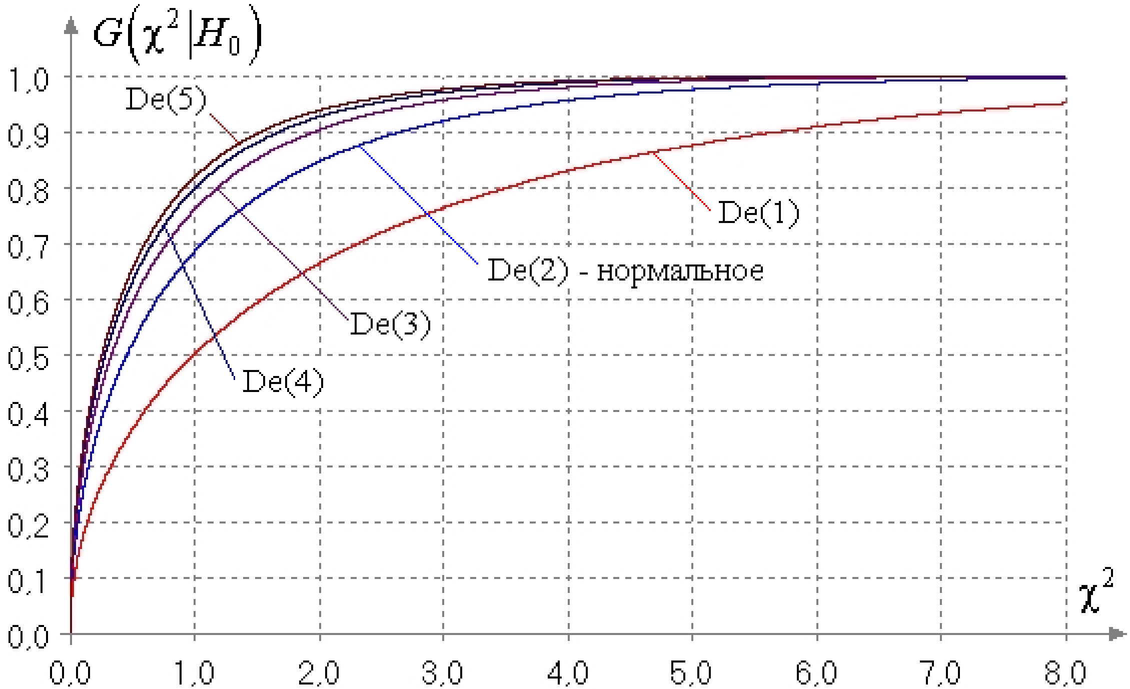


Рис. 4.2. Функции распределения статистики критерия Бартлетта в случае принадлежности выборок законам семейства (4.3) с различными значениями параметра формы ($n_i = 20,\quad i = \overline{1,k},\quad k = 2$)

4.1.2. Критерий Кокрена

Статистика критерия Кокрена выражается формулой

$$Q = \frac{S_{\max}^2}{S_1^2 + S_2^2 + \cdots + S_k^2}, \tag{4.5}$$

где $S_{\max}^2 = \max\left(S_1^2,\ S_2^2, \ldots,\ S_k^2\right)$; k – число независимых оценок дисперсий (число выборок); S_i^2, $i = \overline{1,k}$, – оценки выборочных дисперсий.

Проверяемая гипотеза отклоняется при больших значениях статистики.

В [Большев-СМирнов] критерий Кокрена называется более простым и менее мощным по сравнению с критерием Бартлетта. Второе не соответствует действительности: при выполнении стандартного предположения о нормальности при $k = 2$ эти критерии по мощности эквивалентны, а при $k > 2$ преимущество в мощности за критерием Кокрена.

Распределения статистики Кокрена сильно зависят от объемов наблюдаемых выборок (рис. 4.4) и их числа (рис. 4.5).

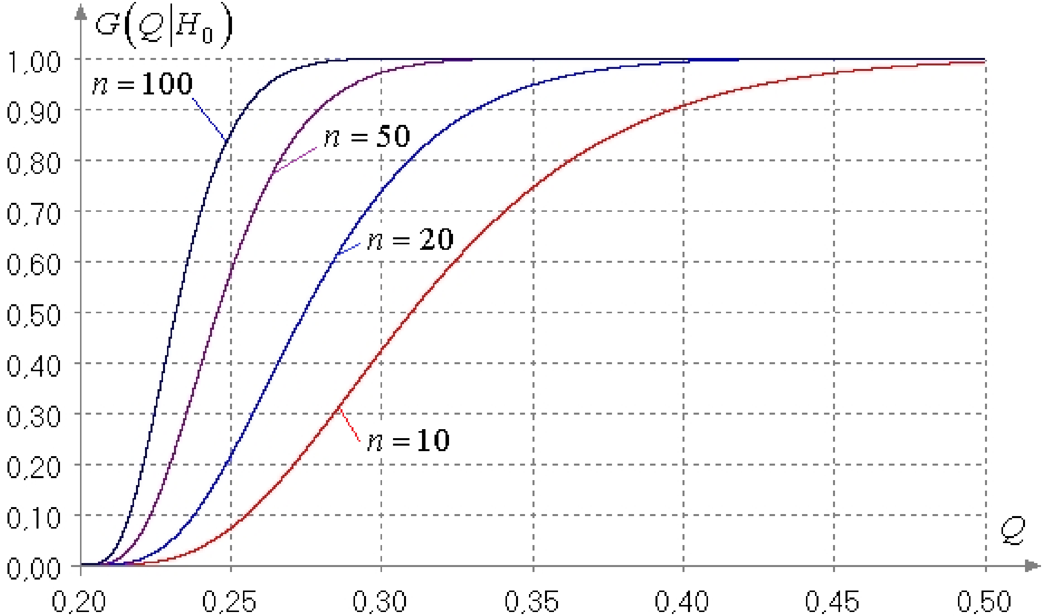


Рис. 4.4. Функции распределения статистики критерия Кокрена при равных объемах выборок $n_i = n$ в случае $k = 5$

4.1.3. Критерий Хартли

Статистика критерия Хартли, применяемого для проверки гипотезы об однородности дисперсий, имеет вид

$$F = \frac{S_{\max}^2}{S_{\min}^2}, \qquad (4.6)$$

где $S_{\max}^2 = \max\left(S_1^2,\ S_2^2, ...,\ S_k^2\right)$; $S_{\min}^2 = \min\left(S_1^2,\ S_2^2, ...,\ S_k^2\right)$; k – число независимых оценок дисперсий (число выборок); S_i^2, $i = \overline{1,k}$ – оценки выборочных дисперсий.

Проверяемая гипотеза отклоняется при больших значениях статистики.

Распределения статистики существенно зависят от k и объемов n_i, $i = \overline{1,k}$.

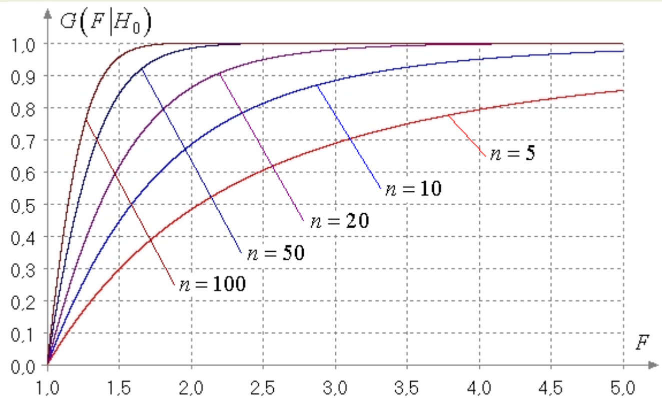


Рис. 4.6. Функции распределения статистики критерия Хартли при равных объемах выборок $n_i = n$ в случае $k = 2$

4.1.4. Критерий Левене

Статистика критерия Левене имеет вид

$$W = \frac{N-k}{k-1} \frac{\sum_{i=1}^{k} n_i \left(\overline{Z}_{i\bullet} - \overline{Z}_{\bullet\bullet} \right)^2}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Z_{ij} - \overline{Z}_{i\bullet} \right)^2}, \qquad (4.7)$$

где k – количество выборок; n_i − объем i-й выборки; $N = \sum_{i=1}^{k} n_i$; X_{ij} – j-е наблюдение в i-й выборке; $Z_{ij} = \left| X_{ij} - \overline{X}_{i\bullet} \right|$, в котором $\overline{X}_{i\bullet}$ − среднее в i-й выборке; $\overline{Z}_{i\bullet}$ − среднее Z_{ij} по i-й выборке; $\overline{Z}_{\bullet\bullet}$ − среднее Z_{ij} по всем выборкам. Проверяемая гипотеза отклоняется при больших значениях статистики.

В описаниях критерия говорится, что в случае принадлежности выборок нормальному закону и справедливости H_0 эта статистика подчинятся F_{ν_1,ν_2}-распределению Фишера с числом степеней свободы $\nu_1 = k-1$ и $\nu_2 = N-k$. На самом деле при объемах выборок в 10…20 элементов *распределения статистики* $G(W \mid H_0)$ *существенно отличаются от* F_{ν_1,ν_2} *-распределения*, и это надо иметь в виду при использовании критерия.

Вместо оценок среднего можно использовать выборочные медиану и усеченное среднее ($Z_{ij} = \left| X_{ij} - \tilde{X}_{i\bullet} \right|$, где $\tilde{X}_{i\bullet}$ − медиана в i-й выборке; $Z_{ij} = \left| X_{ij} - \overline{X}'_{i\bullet} \right|$, где $\overline{X}'_{i\bullet}$ − усеченное среднее в i-й выборке). Считается, что в этих случаях критерий становится еще устойчивей к нарушению предположений о нормальности.

4.1.5. Критерий Фишера

Критерий Фишера используется для проверки гипотезы о равенстве дисперсий *двух* выборок случайных величин. Статистика критерия имеет простой вид

$$F = \frac{s_1^2}{s_2^2}, \tag{4.8}$$

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

В отличие от критериев, рассмотренных выше, распределение статистики критерия Фишера при выполнении стандартного предположения известно точно. В случае принадлежности выборок нормальному закону и справедливости $H_0: \sigma_1^2 = \sigma_2^2$ эта статистика подчиняется F_{ν_1,ν_2} -распределению Фишера с числом степеней свободы $\nu_1 = n_1 - 1$ и $\nu_2 = n_2 - 1$, где n_1 и n_2 – объемы сравниваемых выборок. Проверяемая гипотеза отклоняется при малых $F^* < F_{\alpha/2,\nu_1,\nu_2}$ или больших $F^* > F_{1-\alpha/2,\nu_1,\nu_2}$ значениях статистики.

Как и остальные, критерий Фишера очень чувствителен к отклонениям от нормальности.

4.1.6. Критерий Неймана-Пирсона

Статистика критерия Неймана–Пирсона (критерия отношения правдоподобия) определяется отношением арифметического среднего всех оценок дисперсий s_i^2 к их геометрическому среднему:

$$h = \frac{1}{k}\sum_{i=1}^{k} s_i^2 \Bigg/ \left(\prod_{i=1}^{k} s_i^2\right)^{\frac{1}{k}}, \qquad (4.9)$$

где k – количество выборок, $s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(x_{ij} - \overline{x}_i\right)^2$ – оценки выборочных дисперсий, $\overline{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$ – выборочное среднее значение, x_{ij} – j-й элемент i-й выборки.

Обычно предполагается, что $n_1 = n_2 = \ldots = n_k = n$.

Критерий правосторонний. Проверяемая гипотеза H_0 отклоняется при больших значениях статистики (4.9), когда $h > h_{1-\alpha}$.

Распределения статистики (4.9) зависят от объёма n и от числа анализируемых выборок k, чувствительны к нарушению предположения о принадлежности выборок нормальным законам.

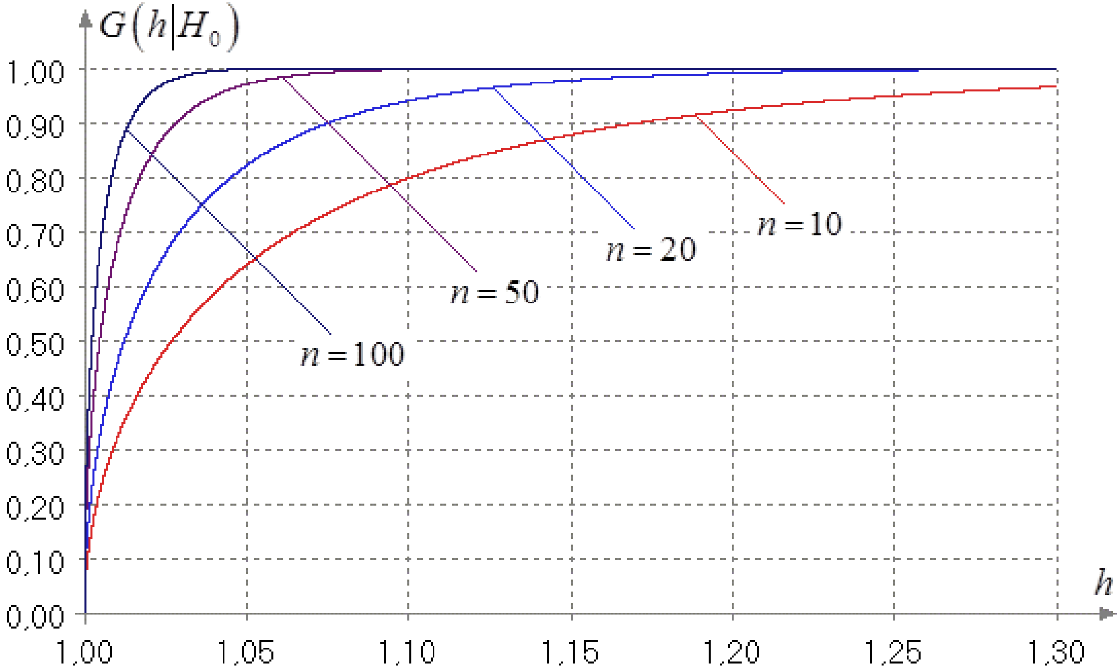


Рис. 4.14. Зависимость распределения статистики критерия Неймана-Пирсона **от объема выборки** (при $n_i = n$, $k = 2$)

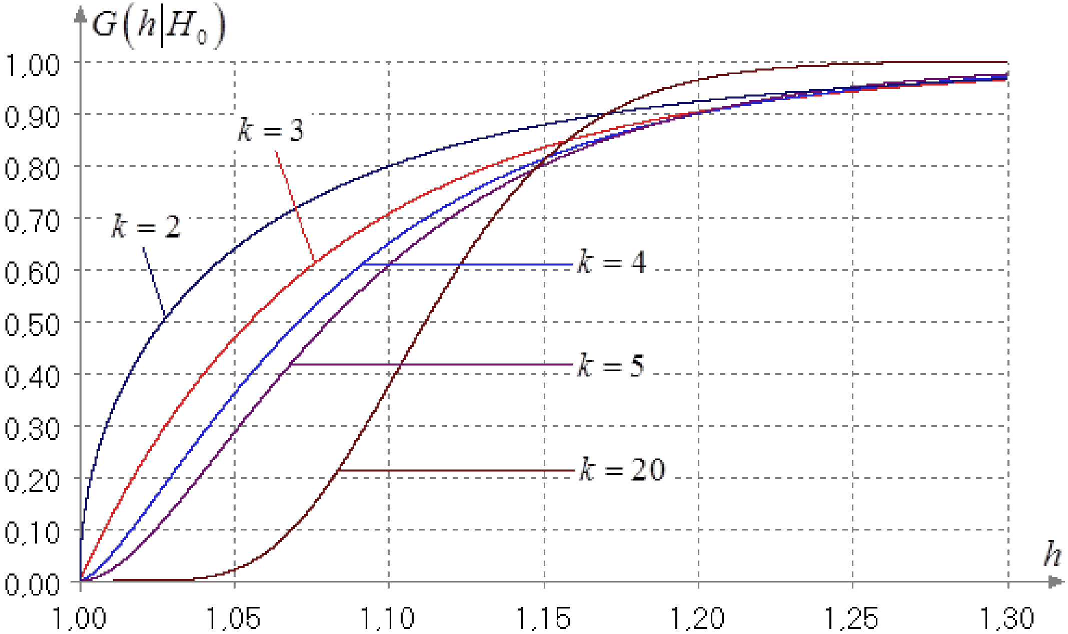


Рис. 4.15. Зависимость распределения статистики критерия
Неймана-Пирсона **от числа выборок** k (при $n = 10$)

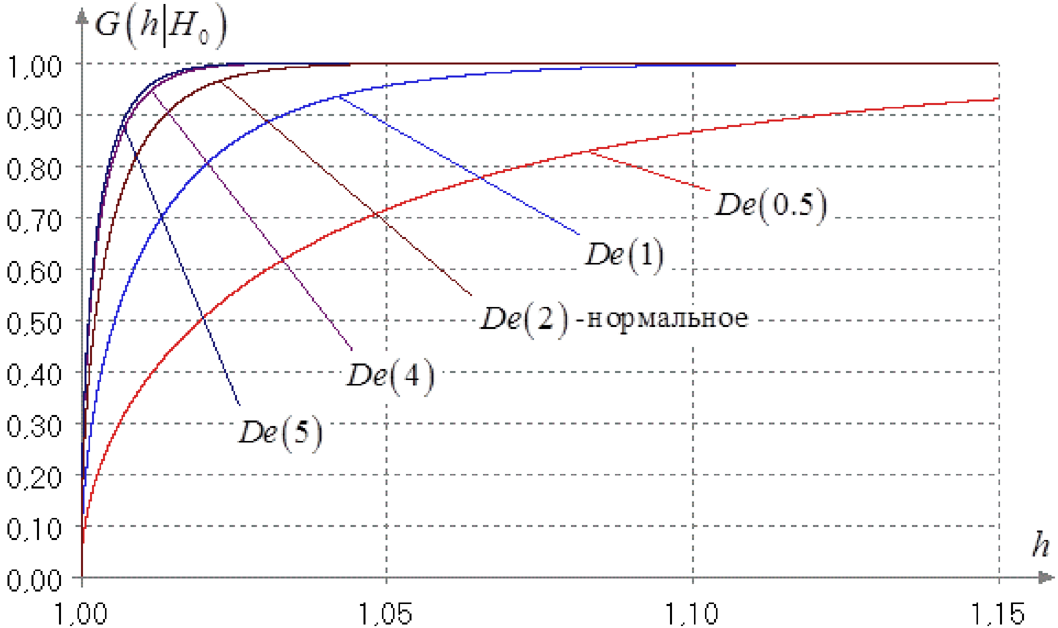


Рис. 4.16. Зависимость распределения статистики (4.9) критерия Неймана-Пирсона **от вида закона** (при $n = 100$, $k = 2$)

4.1.7. Критерий О`Брайена.

При формировании статистики критерия каждый j-й элемент i-й выборки x_{ij} преобразуется в соответствии с формулой

$$V_{ij} = \frac{(n_i - 1.5) n_i \left(x_{ij} - \bar{x}_i \right)^2 - 0.5 s_i^2 (n_i - 1)}{(n_i - 1)(n_i - 2)}, \tag{4.10}$$

где n_i – объём, $\bar{x}_i$ – среднее значение, s_i^2 – оценка дисперсии i-й выборки.

Статистика критерия имеет вид

$$V = \frac{\dfrac{1}{k-1} \sum_{i=1}^{k} n_i \left(\overline{V_i} - \overline{\overline{V_i}} \right)^2}{\dfrac{1}{N-k} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(V_{ij} - \overline{V_i} \right)^2}, \tag{4.11}$$

где $\overline{V_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} V_{ij}$, $\overline{\overline{V_i}} = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} V_{ij}$, $N = \sum_{i=1}^{k} n_i$.

Критерий правосторонний, и проверяемая гипотеза H_0 отклоняется при больших значениях статистики (4.11).

Предельным распределением статистики критерия О'Брайена при справедливости H_0 является $F_{k-1,N-k}$-распределение Фишера с $k-1$ и $N-k$ степенями свободы. Однако проведенные исследования показывают, что распределение статистики (4.11) критерия О'Брайена достаточно медленно сходится к соответствующему распределению Фишера.

4.1.8. Критерий Линка

Критерий Линка (критерий отношения размахов) является аналогом критерия Фишера и используется только при анализе 2-х выборок ($m=2$). Статистика критерия имеет вид:

$$F^* = \frac{\omega_{n_1}}{\omega_{n_2}}, \tag{4.12}$$

где $\omega_{n_1} = x_{1,\max} - x_{1,\min}$, $\omega_{n_1} = x_{2,\max} - x_{2,\min}$ – размахи, а $x_{1,\max}$, $x_{2,\max}$, $x_{1,\min}$, $x_{2,\min}$ – максимальные и минимальные элементы сравниваемых выборок.

Критерий двусторонний. Проверяемая гипотеза отклоняется с уровнем значимости α, если $F^* > F^*_{1-\alpha/2}$ или $F^* < F^*_{\alpha/2}$, где $F^*_{1-\alpha/2}$ и $F^*_{\alpha/2}$ – верхнее и нижнее критические значения статистики.

Распределение статистики критерия существенно зависит от объемов сравниваемых выборок.

Критерий Линка крайне чувствителен к любым отклонениям от нормальности.

4.1.9. Критерий Ньюмана

Статистика критерия Ньюмана (стьюдентизированного размаха) имеет вид:

$$q = \frac{\omega_{n_1}}{s_{n_2}}, \tag{4.13}$$

где $\omega_{n_1} = x_{1,\max} - x_{1,\min}$, $s_{n_2} = \sqrt{\frac{1}{n_2 - 1}\sum_{i=1}^{n_2}\left(x_{2i} - \bar{x}_2\right)^2}$.

Как и предыдущий, этот критерий также является двусторонним. Проверяемая гипотеза H_0 о равенстве дисперсий отклоняется, если $q < q_{\alpha/2}$ или $q > q_{1-\alpha/2}$, где $q_{\alpha/2}$ и $q_{1-\alpha/2}$ – нижнее и верхнее критические значения статистики при заданном уровне значимости α.

Распределения статистики (4.13) критерия Ньюмана при справедливости H_0 зависят от объёмов анализируемых выборок.

Критерий также очень чувствителен к любым отклонениям от нормальности.

4.1.10. Критерий Блиса–Кокрена–Тьюки

Статистика критерия, предложенного в качестве аналога критерия Кокрена, имеет вид

$$c = \frac{\max\limits_{1 \le i \le k} \omega_i}{\sum\limits_{i=1}^{k} \omega_i}, \qquad (4.14)$$

где k – количество сравниваемых выборок, $\omega_i = \max\limits_{1 \le j \le n_i} x_{ij} - \min\limits_{1 \le j \le n_i} x_{ij}$ – размах i-й выборки.

Критерий правосторонний. Если статистика $c > c_{1-\alpha}$, где $c_{1-\alpha}$ – верхнее критическое значение при заданном уровне значимости α, то проверяемая гипотеза H_0 о равенстве дисперсий отклоняется.

Распределение статистики критерия сильно зависит от объема выборок и числа сравниваемых выборок.

4.1.11. Критерий Кадуэлла–Лесли–Брауна

Этот критерий был предложен в качестве аналога критерия Хартли с заменой в статистике отношений оценок дисперсий на отношения размахов

$$K = \frac{\max\limits_{1 \le i \le k} \omega_i}{\min\limits_{1 \le i \le k} \omega_i}, \tag{4.15}$$

где k – количество выборок, ω_i – размах i-й выборки.

Критерий правосторонний. При $K > K_{1-\alpha}$, где $K_{1-\alpha}$ – верхнее критическое значение статистики при заданном уровне значимости α, проверяемая гипотеза H_0 отклоняется.

Распределения статистики критерия Кадуэлла–Лесли–Брауна так же, как и статистики критерия Блисса–Кокрена–Тьюки, существенно зависят и от объема выборок, и от закона распределения, которому подчиняются выборки.

4.1.12. Z–критерий Оверолла–Вудворда

Статистика критерия имеет вид:

$$Z = \frac{1}{k-1}\sum_{i=1}^{k} Z_i^2 , \qquad (4.16)$$

где $Z_i = \sqrt{\frac{c_i\left(n_i - 1\right)s_i^2}{MSE}} - \sqrt{c_i\left(n_i - 1\right) - \frac{c_i}{2}}$, $MSE = \frac{1}{N-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2$, k – количество выборок, $c_i = 2 + 1/n_i$, n_i – размер i-й выборки, s_i^2 – несмещенная оценка дисперсии i-й выборки, $N = \sum_{i=1}^{k} n_i$, x_{ij} – j-й элемент i-й выборки, $\bar{x}_i$ –среднее значение i-й выборки.

При справедливости проверяемой гипотезы H_0 об однородности дисперсий и принадлежности анализируемых выборок нормальному закону предельное распределение статистики (4.16) не зависит от размера выборки и подчиняется $F_{k-1,\infty}$-распределению Фишера.

Однако при малых объемах n_i распределение статистики Z–критерия Оверолла–Вудворда заметно отличается от предельного $F_{k-1,\infty}$-распределения Фишера.

4.1.13. Модифицированный Z–критерий

В целях построения критерия, более устойчивого к нарушению стандартного предположения о нормальности, предложена модификация Z–критерия, статистика которого отличается вычислением величин c_i:

$$c_i = 2.0\left[\frac{1}{K_i}\left(2.9+\frac{0.2}{n_i}\right)\right]^{\frac{1.6(n_i-1.8K_i+14.7)}{n_i}}, \tag{4.17}$$

где $K_i = \frac{1}{n_i-2}\sum_{j=1}^{n_i} G_{ij}^4$ – оценка коэффициента эксцесса i -й выборки, $G_{ij} = \left(x_{ij}-\bar{x}_i\right)\Big/\sqrt{\frac{n_i-1}{n_i}s_i^2}$.

Распределение статистики модифицированного Z–критерия с ростом объема выборок очень медленно сходится к $F_{k-1,\infty}$ -распределению Фишера. Даже при больших объемах выборок распределение статистики критерия не согласуется с $F_{k-1,\infty}$ -распределением, хотя в области больших значений статистики отличие её распределения от $F_{k-1,\infty}$ -распределения не играет существенного значения.

4.1.14. Критерий Миллера

Миллер предложил критерий однородности дисперсий, базирующийся на F-преобразовании Фишера для выборочных дисперсий. Лайард обобщил двухвыборочный критерий Миллера на случай k выборок.

Статистика k-выборочного критерия Миллера имеет вид

$$M = \frac{\sum_{i=1}^{k} n_i \left(\bar{U}_{i\bullet} - \bar{U}_{\bullet\bullet}\right)^2 \Big/ (k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(U_{ij} - \bar{U}_{i\bullet}\right)^2 \Big/ (n-k)}, \tag{4.18}$$

где

$$U_{ij} = n_i \ln S_i^2 - (n_i - 1)\ln S_{i(j)}^2;$$

$$S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2; \qquad \bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij};$$

$$S_{i(j)}^2 = \frac{1}{n_i - 2}\sum_{l \ne j}\left(x_{il} - \bar{x}_{i(j)}\right)^2; \qquad \bar{x}_{i(j)} = \frac{1}{n_i - 1}\sum_{l \ne j} x_{il};$$

$$\bar{U}_{i\bullet} = \frac{1}{n_i}\sum_{j=1}^{n_i} U_{ij}; \qquad \bar{U}_{\bullet\bullet} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} U_{ij}; \qquad n = \sum_{i=1}^{k} n_i.$$

При справедливости проверяемой гипотезы H_0 об однородности дисперсий и выполнении предположений о принадлежности выборок нормальным законам статистика (4.18) должна подчиняться F-распределению Фишера с числами степеней свободы ($k-1$) и ($n-k$).

4.1.15. Критерий Лайарда

Лайард представил критерий со статистикой, в которой используется функция эксцесса нескольких выборок для проверки однородности дисперсий.

Статистика критерия Лайарда определяется выражением

$$L = \sum_{i=1}^{k}(n_i - 1)\frac{\left(\ln S_i^2 - T\right)^2}{\delta^2}, \tag{4.19}$$

где

$$T = \left[\sum_{i=1}^{k}(n_i - 1)\ln S_i^2\right] \Big/ (n-k)\,; \qquad n = \sum_{i=1}^{k} n_i\,; \;\; \delta^2 = 2 + \gamma\left[1 - k/n\right];$$

$$\gamma = \frac{1}{n}\sum_{i=1}^{k} n_i^2 \left[\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^4 \Bigg/ \left(\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2\right)^2\right] - 3. \tag{4.20}$$

Здесь γ – взвешенное среднее коэффициентов эксцесса k выборок.

Проверяемая гипотеза отклоняется при больших значениях статистики (4.19), которая при справедливости проверяемой гипотезы об однородности дисперсий и выполнении предположений о принадлежности выборок нормальным законам асимптотически подчиняется χ^2_{k-1}-распределению.

Однако необходимо отметить, что сходимость распределения статистики $G(L_n | H_0)$ к χ^2_{k-1}-распределению достаточно медленная.

4.2. Непараметрические критерии однородности дисперсий

4.2.1. Критерий Ансари–Бредли

Непараметрические аналоги критериев проверки однородности дисперсий предназначены для проверки гипотез о принадлежности двух выборок с объемами n_1 и n_2 общей генеральной совокупности с одинаковыми характеристиками рассеяния. При этом, как правило, предполагается равенство средних.

Статистика критерия Ансари–Бредли может быть вычислена следующим образом:

$$S = \sum_{i=1}^{n_1} \left\{ \frac{n_1 + n_2 + 1}{2} - \left| R_i - \frac{n_1 + n_2 + 1}{2} \right| \right\}, \qquad (4.22)$$

где R_i – ранги элементов первой выборки в общем вариационном ряду. Проверяемая гипотеза не отклоняется при $S_{\alpha/2} < S < S_{1-\alpha/2}$.

В случае принадлежности выборок случайных величин одному и тому же закону распределение $G\left(S \middle| H_0\right)$ статистики (4.22) при справедливости проверяемой гипотезы H_0 не зависит от вида этого закона.

Математическое ожидание и дисперсия статистики (4.22) имеют вид

$$E[S] = \begin{cases} \dfrac{n_1(n_1 + n_2 + 2)}{4} & \text{при четном } (n_1 + n_2), \\ \dfrac{n_1(n_1 + n_2 + 1)^2}{4(n_1 + n_2)} & \text{при нечетном } (n_1 + n_2); \end{cases} \quad D[S] = \begin{cases} \dfrac{n_1 n_2 (n_1 + n_2 - 2)(n_1 + n_2 + 2)}{48(n_1 + n_2 - 1)} & \text{при четном } (n_1 + n_2), \\ \dfrac{n_1 n_2 (n_1 + n_2 + 1)\left[(n_1 + n_2)^2 + 3\right]}{48(n_1 + n_2)^2} & \text{при нечетном } (n_1 + n_2). \end{cases}$$

При объемах выборок $n_1,\ n_2 > 10$ дискретное распределение нормированной статистики

$$S^* = (S - E[S]) / \sqrt{D[S]} \tag{4.23}$$

достаточно хорошо приближается стандартным нормальным законом.

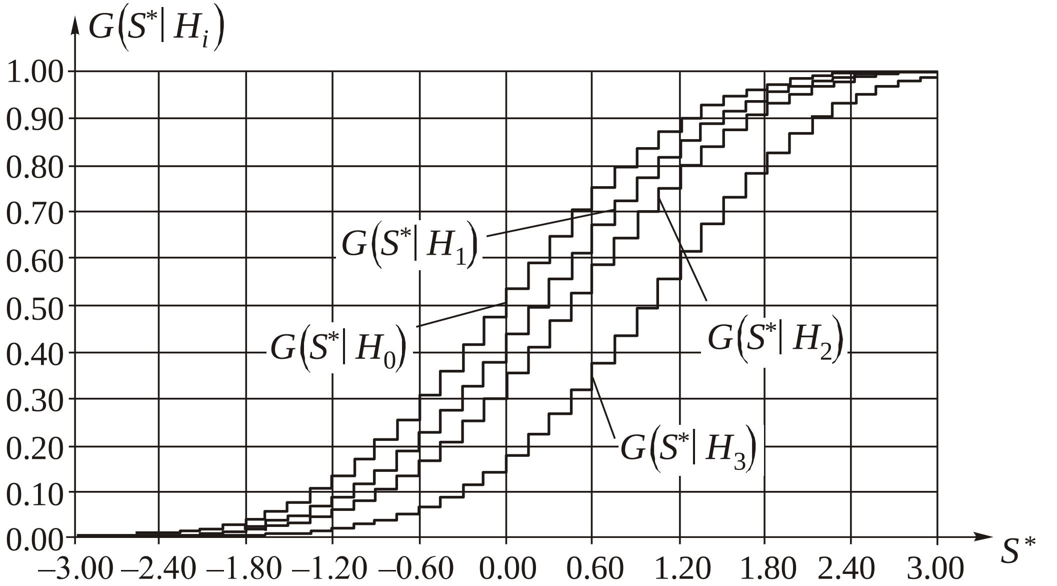


Рис. 4.33. Распределения $G(S^*|H_i)$ нормированной статистики критерия Ансари–Бредли [выборки объемом $n_1 = n_2 = 10$ принадлежат De(3)]

4.2.2. Критерий Муда

Статистика критерия имеет вид:

$$M=\sum_{i=1}^{n_1}\left(R_i-\frac{n_1+n_2+1}{2}\right)^2, \tag{4.24}$$

где R_i – ранги элементов первой выборки в общем вариационном ряду двух выборок. Проверяемая гипотеза не отклоняется при $M_{\alpha/2}<M<M_{1-\alpha/2}$.

При $n_1, n_2>10$ распределение нормированной статистики

$$M^*=\frac{\left(M-E[M]+\frac{1}{2}\right)}{\sqrt{D[M]}}, \tag{4.25}$$

где

$$E[M]=\frac{n_1(n_1+n_2+1)(n_1+n_2-1)}{12},$$

$$D[M]=\frac{n_1 n_2\left(n_1+n_2+1\right)(n_1+n_2+2)(n_1+n_2-2)}{180},$$

хорошо приближается стандартным нормальным законом, а при $n_1, n_2>20$, как показали исследования, дискретностью распределений статистик (4.24), (4.25) вообще можно пренебречь.

4.2.3. Критерий Сижела–Тьюки

Статистика критерия строится следующим образом. Вариационный ряд, построенный по объединенной выборке, $x_1 \le x_2 \le ... \le x_n$, где $n = n_1 + n_2$, преобразуется в последовательность вида

$$x_1, x_n, x_{n-1}, x_2, x_3, x_{n-2}, x_{n-3}, x_4, x_5, \ldots,$$

т. е. оставшийся ряд «переворачивается» каждый раз после приписывания рангов паре крайних значений. В качестве статистики критерия используется сумма рангов элементов первой выборки.

Статистика критерия имеет вид

$$R = \sum_{i=1}^{n_1} R_i . \tag{4.26}$$

Проверяемая гипотеза не отклоняется при $R_{\alpha/2} < R < R_{1-\alpha/2}$.

При $n_1, n_2 > 10$ распределение нормированной статистики

$$R^* = (R - E[R]) / \sqrt{D[R]}, \tag{4.27}$$

где

$$E[R] = \frac{n_1(n_1 + n_2 + 1)}{2}, \quad D[R] = \frac{n_1 n_2 \left(n_1 + n_2 + 1\right)}{12},$$

достаточно хорошо приближается стандартным нормальным законом. В этом случае проверяемая гипотеза не отклоняется при $N^*_{\alpha/2} < R^* < N^*_{1-\alpha/2}$.

При этом дискретностью распределения статистики можно практически пренебречь с $n_1, n_2 > 30$.

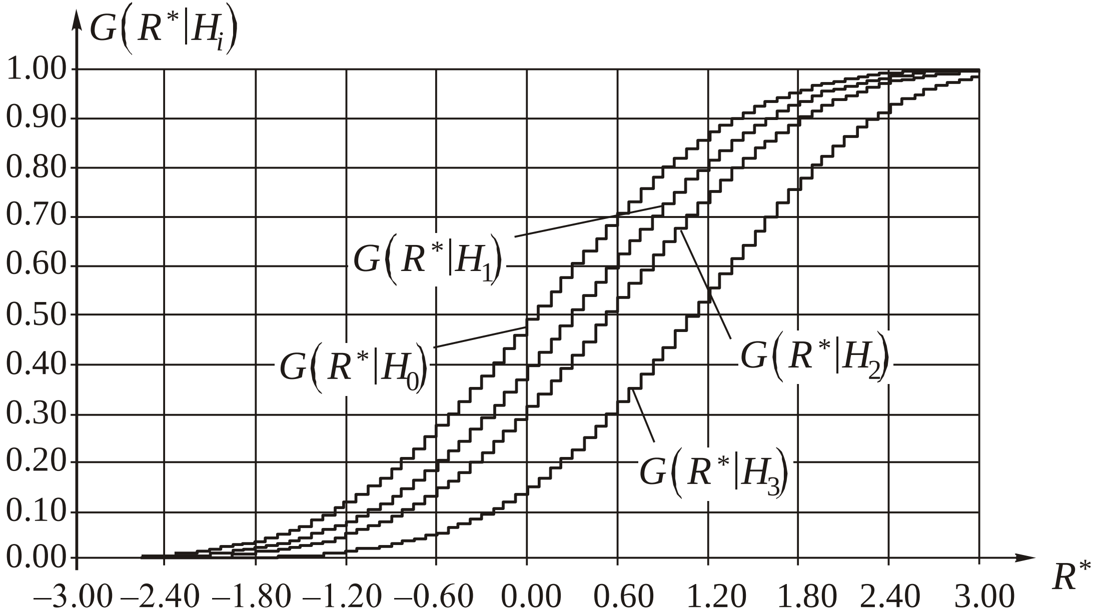


Рис. 4.35. Распределения $G(R^* \mid H_i)$ статистики нормированного критерия Сижела–Тьюки [выборки объемом $n_1 = n_2 = 10$ принадлежат De(3)]

По мощности критерний эквивалентен критерию Ансари–Бредли.

4.2.4. Критерий Клотца

Статистика критерия имеет вид:

$$L = \sum_{i=1}^{n_1} u^2_{R_i/(n_1+n_2+1)} , \qquad (4.28)$$

где n_1 и n_2 – объемы сравниваемых выборок; R_i – ранг i-го элемента первой выборки в общем упорядоченном по возрастанию ряду (n_1+n_2) значений объединенной выборки; u_γ – γ-квантиль стандартного нормального распределения. Значения $u^2_{i/(N+1)}$, называемые метками критерия, которые несложно вычислить.

Гипотеза о равенстве параметров масштаба не отклоняется с достоверностью α, если $L_{\alpha/2} < L < L_{1-\alpha/2}$.

При $n_1, n_2 > 10$ нормализованная статистика

$$L^* = \frac{L - E[L]}{\sqrt{D[L]}} , \qquad (4.29)$$

где

$$E[L] = \frac{n_1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} u^2_{\frac{i}{n_1+n_2+1}} , \quad D[L] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} u^4_{\frac{i}{n_1+n_2+1}} - \frac{n_2}{n_1(n_1+n_2-1)} \left[\frac{n_1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} u^2_{\frac{i}{n_1+n_2+1}} \right]^2 ,$$

хорошо аппроксимируется стандартным нормальным законом. Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $N^*_{\alpha/2} < L^* < N^*_{1-\alpha/2}$, где N^*_α – соответствующая квантиль стандартного нормального закона.

Дискретностью распределений статистик критерия можно пренебречь при $n_1, n_2 > 10$ (см. рис. 4.36). При таких же объёмах выборок можно считать несущественным отклонение распределения $G(L^* | H_0)$ статистики (4.25) от аппроксимирующего стандартного нормального закона.

Критерий Клотца демонстрирует мощность, превышающую мощность критерия Муда, который в свою очередь имеет преимущество перед критериями Ансари–Бредли и Сижела–Тьюки.

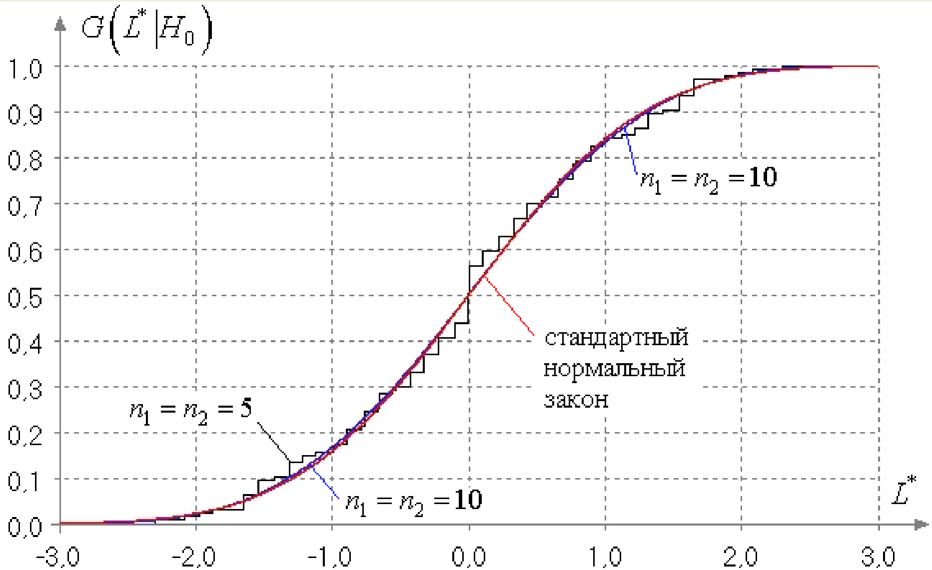


Рис. 4.36. Сходимость распределения $G(L^* | H_0)$ статистики (4.29) к стандартному нормальному закону

4.2.5. Критерий Кейпена

Статистика критерия Кейпена задается соотношением

$$K = \sum_{i=1}^{n_1} a_{n_1+n_2}\left(R_i\right), \qquad (4.30)$$

где n_1 и n_2 – объемы сравниваемых выборок; R_i – ранг i-го элемента первой (меньшей по объёму) выборки в общем упорядоченном по возрастанию ряду $n = n_1 + n_2$ значений объединенной выборки. Метки $a_n(i)$ представляют собой квадрат математического ожидания i-й порядковой статистики в выборке объема n из стандартного нормального закона и приближённо определяются соотношением $a_n(i) = u^2_{(i-3/8)/(n+1/4)}$, где u_γ – γ-квантиль стандартного нормального распределения.

При $n_1, n_2 > 10$ справедлива нормальная аппроксимация

$$K^* = \frac{K - E[K]}{\sqrt{D[K]}}, \qquad (4.31)$$

где

$$E[K] = n_1, \quad D[K] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} a^2_{n_1+n_2}(i) - \frac{n_1 n_2}{n_1+n_2-1}.$$

Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $N^*_{\alpha/2} < K^* < N^*_{1-\alpha/2}$.

Распределения статистики (4.31) критерия Кейпена, как и критерия Клотца (4.29), с ростом объёмов выборок быстро сходятся к стандартному нормальному закону.

По своим свойствам и мощности критерий эквивалентен критерию Клотца.

4.2.6. k-выборочный критерий Флайне–Киллина

Модифицированный непараметрический критерий Флайне–Киллина предназначен для проверки однородности дисперсий $k \ge 2$ выборок с объёмами n_i, $i = \overline{1,k}$. Статистика критерия формируется следующим образом.

По исходным выборкам вычисляются абсолютные значения $z_{ji} = \left| x_{ji} - \tilde{x}_i \right|$, где $\tilde{x}_i$ – выборочная медиана i-й выборки, $i = \overline{1,k}$. Далее строится вариационный ряд объединённой выборки z_{ji}, $j = \overline{1,n_i}$, $i = \overline{1,k}$. Для элементов i-й выборки на основании рангов R_{ji} её элементов z_{ji} в объединённой выборке строятся метки

$$a_{n,R_{ji}} = \Phi^{-1}\left(\frac{1 + R_{ji} / (n+1)}{2} \right), \; j = \overline{1,n_i},$$

где $n = \sum_{i=1}^{k} n_i$, находятся $\bar{A}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} a_{n,R_{ji}}$.

Статистика критерия имеет вид:

$$\chi_0^2 = \frac{\sum_{i=1}^{k} n_i \left(\bar{A}_i - \bar{a} \right)^2}{V^2}, \qquad (4.32)$$

где $\bar{a} = \frac{1}{n} \sum_{j=1}^{n} a_{n,j}$, $V^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(a_{n,j} - \bar{a} \right)^2$.

Асимптотическим распределением статистики (4.32) при справедливости H_0 и больших объёмах выборок является χ_{k-1}^2-распределение. При малых объёмах выборок распределение статистики дискретное.

Сходимость распределений $G(\chi_0^2 | H_0)$ статистики (4.32) с ростом объёмов выборок к χ_{k-1}^2-распределению в случае принадлежности сравниваемых выборок нормальному закону при $k=2$ и равных n_i иллюстрирует рис. 4.36.

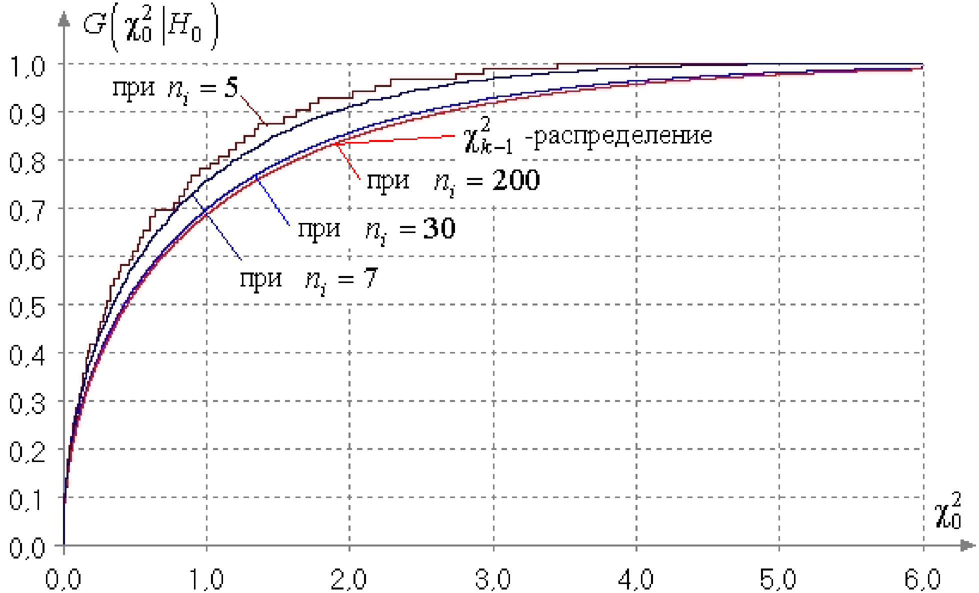


Рис. 4.36. Сходимость распределения $G(\chi_0^2 | H_0)$ статистики (4.32) к χ_{k-1}^2-распределению при $k=2$ и $n_1 = n_2$

При ограниченных объёмах выборок распределения $G(\chi_0^2 | H_0)$ статистики (4.32) зависят от закона, которому принадлежат выборки.

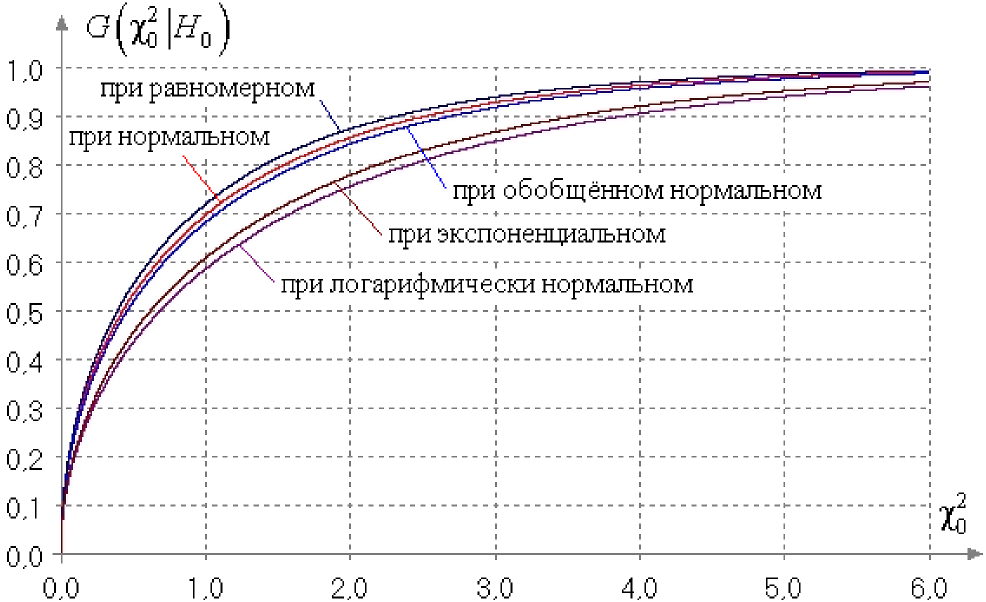


Рис. 4.37. Зависимость распределений $G(\chi_0^2 | H_0)$ статистики (4.32) от закона, которому принадлежат выборки, при $k = 2$ и $n_1 = n_2 = 30$

Исходный критерий Флайне–Киллина был построен в предположении о принадлежности анализируемых выборок симметричным законам. Именно в такой ситуации асимптотическим распределением статистики (4.32) оказывается χ^2_{k-1}-распределение.

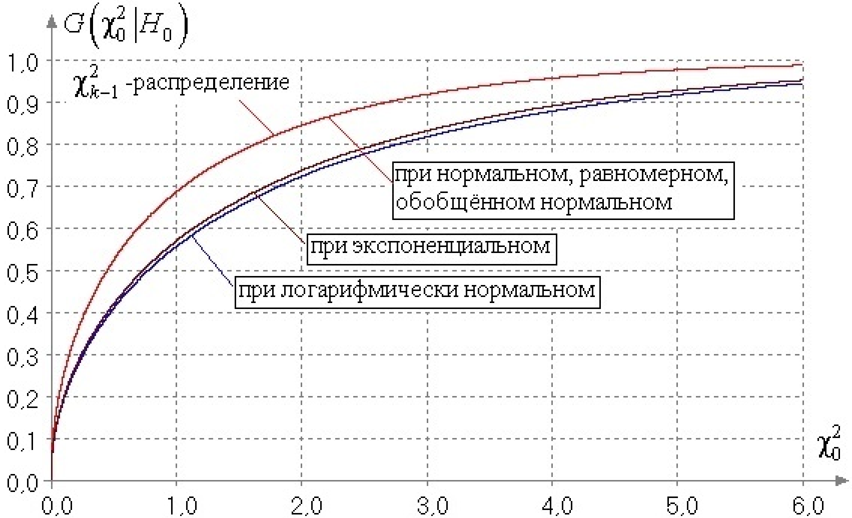


Рис. 4.38. Распределения $G(\chi_0^2 \mid H_0)$ статистики (4.32) при различных законах, которым принадлежат пары сравниваемых выборок, при $k = 2$ и $n_1 = n_2 = 200$

Предположение о симметричности закона, а также зависимость распределения статистики при ограниченных объёмах выборок от n_i и от вида симметричного закона, существенно ограничивают возможность формирования корректного статистического вывода по результатам проверки гипотезы.

4.3. Сравнительный анализ мощности критериев

В случае анализа двух выборок и выполнении стандартного предположения о нормальности наибольшей и одинаковой мощностью относительно тех же конкурирующих гипотез обладают критерии Фишера, Бартлетта, Кокрена, Хартли, Неймана–Пирсона и Z–критерий Оверолла–Вудворда.

Далее в порядке убывания мощности следует группа устойчивых критериев О`Брайена, модифицированный Z-критерий, Левене.

Наименее перспективна для применения группа параметрических критериев Ньюмана, Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Линка, которая уступает в мощности даже непараметрическим критериям, имея перед последними некоторое преимущество в мощности лишь при очень малых объёмах выборок.

Среди непараметрических критериев наибольшей мощностью обладает критерий Клотца, затем идёт критерий Флайне–Киллина, потом критерий Муда, который уже заметно уступает критериям Фишера, Бартлетта, Кокрена, Хартли, Неймана–Пирсона, Z–критерию Оверолла–Вудворда, О`Брайена, модифицированному Z-критерию и Левене.

В случае нарушения стандартного предположения и принадлежности двух выборок законам с более легкими хвостами, чем у нормального закона, вышеуказанный порядок предпочтительности сохраняется.

При симметричных законах с более тяжелыми хвостами по сравнению с нормальным законом (в случае двух выборок) критерии упорядочиваются следующим образом:

Флайне–Киллина $\succ$ *Клотца* $\succ$ *Муда* $\succ$ *Левене* $\succ$ *Сижела–Тьюки* ~ *Ансари–Бредли* $\succ$ *Лайарда* $\succ$ *О'Брайена* $\succ$ *Миллера* $\succ$ *Модифицированный Z–критерий* $\succ$ ***группа критериев*** *(Бартлетта, Кокрена, Хартли, Фишера, Неймана–Пирсона, Z–критерий Оверолла–Вудворда)* $\succ$ *Ньюмана* $\succ$ ***группа критериев*** *(Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Линка)*

Если анализируется большее число выборок, то при выполнении стандартного предположения или в случае принадлежности выборок законам с более лёгкими хвостами по сравнению с нормальным законом критерии по убыванию мощности располагаются в следующем порядке:

Кокрена $\succ$ *О'Брайена* $\succ$ *Z–критерий Оверолла–Вудворда* $\succ$ *Модифицированный Z–критерий* $\succ$ ***группа критериев*** *(Бартлетта, Неймана–Пирсона)* $\succ$ *Хартли* $\succ$ *Миллера* $\succ$ *Лайарда* $\succ$ *Левене* $\succ$ *Флайне–Киллина* $\succ$ *Блисса–Кокрена–Тьюки* $\succ$ *Кадуэлла–Лесли–Брауна.*

Надо отметить, что при симметричных законах с более лёгкими хвостами критерий Левене уступает в мощности критерию Флайне–Киллина.

При симметричных законах с более тяжелыми хвостами ситуация меняется:

Флайне–Киллина $\succ$ *Левене* $\succ$ *О'Брайена* $\succ$ *Лайарда* $\succ$ *Модифицированный Z–критерий* $\succ$ *Миллера* $\succ$ ***группа критериев*** *(Бартлетта, Неймана–Пирсона)* $\succ$ *Z–критерий Оверолла–Вудворда* $\succ$ *Хартли* $\succ$ *Кокрена* $\succ$ *Кадуэлла–Лесли–Брауна* $\succ$ *Блисса–Кокрена–Тьюки.*

Основной недостаток параметрических критериев связан с тем, что классические результаты обеспечивают корректность применения данных критериев лишь при выполнении предположения о принадлежности анализируемых выборок нормальным законам, так как только для этой ситуации известны распределения статистик рассмотренных параметрических критериев или таблицы процентных точек.

Этот недостаток можно преодолеть, используя компьютерные технологии для исследования распределений статистик и оценки p_{value}.

Непараметрические критерии также не свободны от недостатков.

Во-первых, при их использовании, как правило, предполагается равенство средних.

Во-вторых, корректность их применения обеспечивается в случае принадлежности анализируемых выборок закону распределения вероятностей одного и того же вида.

По существу, при справедливости проверяемой гипотезы о равенстве дисперсий, соответствующих двум выборкам, распределение статистики известно для случая однородности законов этих выборок.

Выполнение этих условий несколько сужает область корректного использования непараметрических критериев.

В-третьих, непараметрические критерии обладают заметно меньшей мощностью по сравнению с параметрическими.

Действие параметрических критериев при необходимости можно распространить на ситуации, когда выборки описываются законами, отличающимися от нормального, воспользовавшись методикой компьютерного моделирования для исследования распределений статистик и построения для этих распределений моделей или таблиц процентных точек. Из рассмотренных критериев в наибольшей степени на эту роль подходит критерий Кокрена.

4.4. О применении критериев однородности в условиях округления данных

Влияние ошибок округления на распределения статистик критериев однородности законов покажем на примере двухвыборочного критерия Андерсона–Дарлинга–Петита со статистикой

$$A^2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1+n_2-1} \frac{\left(M_i(n_1+n_2) - n_1 i\right)^2}{i(n_1+n_2-i)}, \tag{3}$$

где M_i – число элементов первой выборки, меньших или равных i-му элементу вариационного ряда объединенной выборки.

Влияние погрешности округления на распределения статистик критериев однородности законов при справедливости H_0 без потери общности рассмотрим в случае принадлежности анализируемых выборок стандартному нормальному закону.

При одинаковых ошибках округления анализируемых выборок $\Delta_1 = \Delta_2$ и объёмах $n_i = 50$ и при $\Delta_i \le 0.01\sigma$ распределения $G(A^2 | H_0)$ практически не отклоняются от распределения $a2(S)$, но отклоняются при неравных Δ_i.

На рис. 6.1 распределения $G(A^2 | H_0)$ статистики критерия при $n_i = 50$ показаны в зависимости от Δ_2 при $\Delta_1 = 0.01\sigma$. В этом случае отклонение распределения $G(A^2 | H_0)$ от $a2(S)$ при $\Delta_1 = 0.1\sigma$ еще практического значения не имеет. При тех же Δ_i с ростом n_i отклонения $G(A^2 | H_0)$ от $a2(S)$ увеличивается.

Таким же образом ошибки округления сказываются на распределениях статистик других двухвыборочных (Смирнова, Лемана–Розенблатта) и k-выборочных (Андерсона–Дарлинга, Жанга и других) критериев однородности законов.

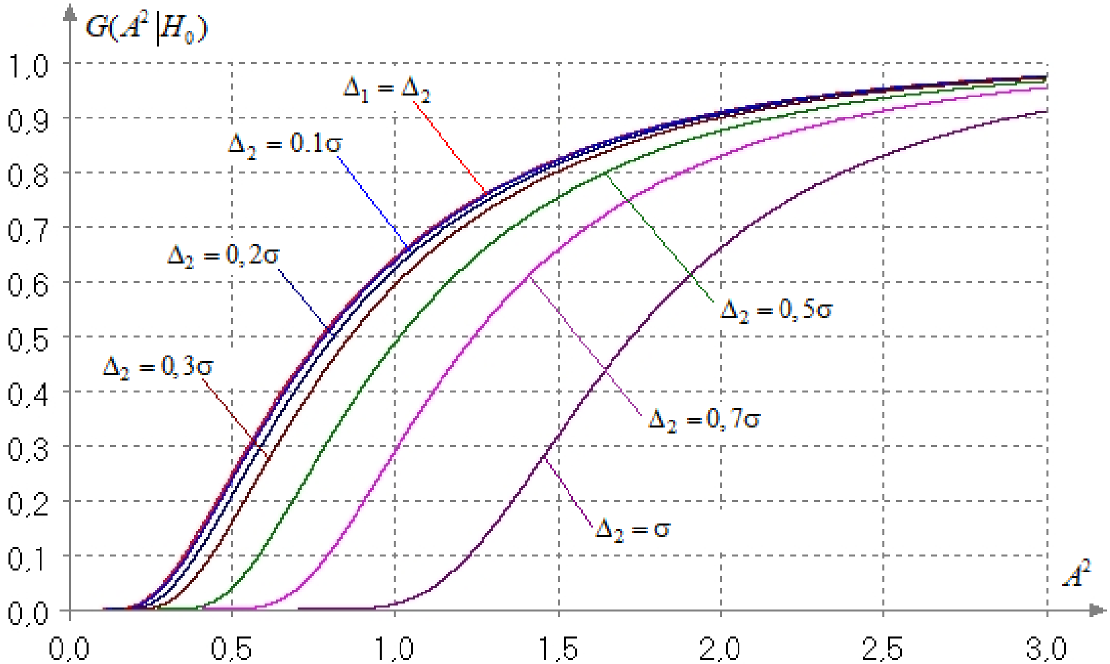


Рис. 6.1. Зависимость распределений статистики (3) критерия Андерсона–Дарлинга–Петита от Δ_2 (при $\Delta_1 = 0.01\sigma$)

Исследование распределений статистик 2-х и k-выборочных параметрических критериев **однородности средних показало, что ошибки округления результатов измерений не оказывают на них какого-либо заметного влияния**.

Напротив, на распределения статистик параметрических критериев, используемых для проверки аналогичных гипотез **относительно дисперсий**, **ошибки округления могут оказывать существенное влияние**.

Покажем это на критерии Бартлетта со статистикой (4.4). Асимптотическим распределением статистики критерия Бартлетта при выполнении стандартного предположения о нормальности и числе сравниваемых выборок k является χ^2_{k-1}-распределение. При наличии округления и равных Δ_i реальные распределения $G(\chi^2 | H_0)$ статистики критерия не отклоняются от χ^2_{k-1}-распределения.

При существенных и различающихся Δ_i реальные распределения $G(\chi^2 | H_0)$ отклоняются от χ^2_{k-1}-распределения. Рис. 6.2 иллюстрирует зависимость распределения $G(\chi^2 | H_0)$ статистики (4.4) критерия Бартлетта от погрешности округления Δ_2 наблюдений во второй выборке при $\Delta_1 = 0.01\sigma$ и при объёмах выборок $n_i = 1000$ в случае $k = 2$.

Как можно видеть, при существенно отличающихся Δ_1 и Δ_2 распределения $G(\chi^2 | H_0)$ значительно отклоняются от χ^2_1-распределения. Заметим, что при меньших объёмах выборок, например, при $n_i = 100$ отклонением реального распределения от асимптотического при $\Delta_2 = 0.5\sigma$ ещё можно пренебречь.

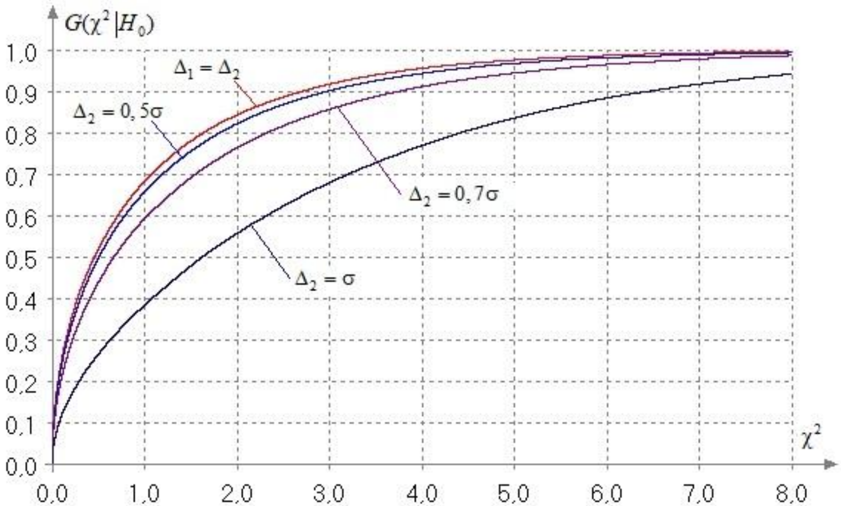


Рис. 6.2. Зависимость распределения статистики (5.1) критерия Бартлетта от Δ_2 (при $\Delta_1 = 0.01\sigma$ и $n_i = 1000$)

Если реальные распределения $G(\chi^2|H_0)$ статистики при наличии округления и равных Δ_i не отклоняются от асимптотических, то при справедливости конкурирующих гипотез H_i распределения $G(\chi^2|H_i)$ зависят от величин Δ_i. И с ростом Δ_i мощность критериев снижается даже при равных Δ_i.

Аналогичным образом при неравных Δ_i с ростом ошибок округления меняются распределения статистик других параметрических критериев однородности дисперсий.

References

1. *Лемешко Б. Ю.* Об ошибках, совершаемых при использовании непараметрических критериев согласия / Б. Ю. Лемешко // Измер. техника. 2004. – № 2. – С. 15–20.
2. *Lemeshko B. Yu.* Errors when using nonparametric fitting criteria / B. Yu. Lemeshko // Measurement Techniques. – 2004. – Vol. 47, №. 2. – P. 134–142.
3. *Kolmogoroff A. N.* Sulla determinazione empirica di una legge di distribuzione / A. N. Kolmogoroff // G. Ist. Ital. attuar. – 1933. – Vol. 4. – № 1. – P. 83–91.
4. *Большев Л. Н.* Асимптотические пирсоновские преобразования / Л. Н. Большев // Теория вероятностей и ее применение. – 1963. – Т. 8, № 2. – С. 129–155.
5. *Большев Л. Н.* Теория вероятностей и математическая статистика : избр. тр. / Л. Н. Большев ; под ред. Ю. В. Прохорова. – М. : Наука, 1987. – 286 с.
6. *Большев Л. Н.* Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – М. : Наука, 1983. – 416 с.
7. *Anderson T. W.* A test of goodness of fit / T. W. Anderson, D. A. Darling // J. Amer. Stist. Assoc. – 1954. – Vol. 29. – P. 765–769.
8. *Anderson T. W.* Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes / T. W. Anderson, D. A. Darling // AMS. – 1952. – Vol. 23. – P. 193–212.
9. *Kuiper N.H.* Tests concerning random points on a circle / N. H. Kuiper // Proc. Konikl. Nederl. Akad. Van Wettenschappen. – 1960. Series A – V. 63. – P.38-47.
10. *Stephens M. A.* EDF statistics for goodness of fit and some comparisons / M. A. Stephens // J. Am. Statist. Assoc. – 1974. – Vol. 69. – No. 347. – P. 730–737.
11. *Лемешко Б. Ю.* О применении и мощности непараметрических критериев согласия Купера, Ватсона и Жанга / Б. Ю. Лемешко, А. А. Горбунова // Измерительная техника. – 2013. – № 5. – С. 3-9.
12. *Watson G. S.* Goodness-of-fit tests on a circle. I. / G. S. Watson // Biometrika. – 1961. – V. 48. – No. 1-2. – P.109-114.
13. *Watson G. S.* Goodness-of-fit tests on a circle. II. / G. S. Watson // *Biometrika*. – 1962. – V. 49. – No. 1-2. – P.57- 63.
14. *Zhang J.* Powerful goodness-of-fit and multi-sample tests / J. Zhang // PhD Thesis. York University, Toronto. 2001. – 113 p. URL: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (дата обращения 28.01.2013).

15. *Zhang J.* Powerful goodness-of-fit tests based on the likelihood ratio / J. Zhang // Journal of the Royal Statistical Society: Series B. – 2002. – V. 64. – № 2. – P.281-294.
16. *Zhang J.* Likelihood-ratio tests for normality / J. Zhang, Yu. Wub // Computational Statistics & Data Analysis. –. V. 49. – No. 3. – P.709-721.
17. *Kac M.* On tests of normality and other tests of goodness of fit based on distance methods / M. Kac, J. Kiefer, J. Wolfowitz // Ann. Math. Stat. – 1955. – Vol. 26. – P. 189–211.
18. *Никулин М. С.* О критерии хи-квадрат для непрерывных распределений / М. С. Никулин // Теория вероятностей и ее применение. – 1973. – Т. 18. – № 3. – С.675-676.
19. *Никулин М. С.* Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба / М. С. Никулин // Теория вероятностей и ее применение. – 1973. – Т. 18. – № 3. – С.583-591.
20. *Rao K.C.* A chi-squared statistic for goodness-of-fit tests within the exponential family / K. C. Rao, D. S. Robson // Comm. Statist. – 1974. – V. 3. – № 12. – P.1139-1153.
21. *Ansari A. R.* Rank-tests for dispersions / A. R. Ansari, R. A. Bradley // AMS. – 1960. – Vol. 31, № 4. – P. 1174–1189.
22. *Bartlett M. S.* Properties of sufficiency of statistical tests / M. S. Bartlett // Proc. Roy. Soc. – 1937. – A 160. – P. 268–287.
23. *Behrens W. U.* Ein Beitrag Zur Fehlerberechnung bei wenigen Beobachtungen // Landw. Jb., 1929. B. 68, – S. 807-837.
24. *Bliss C.I., Cochran W.G., Tukey J.W.* A rejection criterion based upon the range // Biometrika. 1956. Vol. 43. No. 3/4. – P. 418-422.
25. *Brown M. B.* Robust Tests for Equality of Variances / M. B. Brown, A. B. Forsythe // J. Amer. Statist. Assoc. – 1974. – Vol. 69. – P. 364–367.
26. *Capon J.* Asymptotic efficiency of certain locally most powerful rank tests / J. Capon // AMS. – 1961. – Vol. 32, № 1. – P. 88–100.
27. *Chen S., Pokojovy M.* Modern and classical *k*-sample omnibus tests // Wiley Online Library, 2017. DOI: 10.1002/wics.1418
28. *Cochran W. G.* The distribution of the largest of a set of estimated variances as a fraction of their total / W. G. Cochran // Annals of Eugenics. – 1941. – Vol. 11. – P. 47–52.
29. *Conover W. J.* Several k-sample Kolmogorov-Smirnov tests // The Annals of Mathematical Statistics. – 1965. – Vol. 36, No. 3. – P.1019-1026.
30. *Conover W. J.* Practical Nonparametric Statistics / W. J. Conover. – 3d ed. – Wiley, 1999. – 584 p.

31. *Conover W.J., Johnson M.E., Johnson M.M.* A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data // Technometrics. – 1981. – Vol. 23, No. 4. – P. 351-361.
32. *Cox M.G.* The evaluation of key comparison data // Metrologia.2002. Vol. 39. No. 6. – P.589-585. DOI: 10.1088/0026-1394/39/6/10
33. *Cox M.G.* The evaluation of key comparison data: determining the largest consistent subset // Metrologia. 2007. Vol. 44. No. 3. – P.187-200. DOI: 10.1088/0026-1394/44/3/005
34. *Fisher R. A.* The fiducial argument in statistical inference // Annals of Eugenics, Lond., 1935. Vol. **6**, No. 4. – P. 391-398.
35. *Fisher R. A.* The asymptotic approach to Behrens's integral, with further tables for the *d* test of significance // Annals of Eugenics, Lond., 1941. Vol. **11**. – P. 141-172.
36. *Fisher R.A., Yates F.* Statistical tables for biological, agricultural and medical research. London & Edinburgh: Oliver and Boyd, 1948.
37. *Fligner M.A., Killeen T.J.* Distribution-Free Two-Sample Tests for Scale // Journal of American Statistical Association. 1976. Vol. 71. No. 353. – P.210-213.
38. *Gorbunova A. A.* Classical tests of variances homogeneity for non-normal distributions / A. A. Gorbunova, S. B. Lemeshko, B. Yu. Lemeshko // Proceedings Third International Conference on Accelerated Life Testing, Reliability-based Analysis and Design. 19-21 May 2010. – Clermont-Ferrand, France. – P. 117–124.
39. *Gorbunova A.A., Lemeshko B.Yu.* Application of Variance Homogeneity Tests Under Violation of Normality Assumption // Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Simulations and Statistical Inference" – AMSA'2011, Novosibirsk, Russia, 20-22 September, 2011. P. 28-36.
40. *Gorbunova A.A., Lemeshko B.Yu.* Application of Parametric Homogeneity of Variances Tests under Violation of Classical Assumption // Proceedings, 2nd Stochastic Modeling Techniques and Data Analysis International Conference. 5 - 8 June 2012, Chania, Crete, Greece. P.253-260. (http://www.smtda.net/images/1_SMTDA2012_Proceedings_D-J_119-338.pdf)
41. *Hartley H. O.* The maximum F-ratio as a short-cut test of heterogeneity of variance / H. O. Hartley // Biometrika. – 1950. – Vol. 37. – P. 308–312.
42. *Hines W.* Increased power with modified forms of the Levene (Med) test for heterogeneity of variance / W. Hines, R. Hines // Biometrics. – 2000. – Vol. 56. – P. 451–454.
43. *Hoeffding W.* Optimum non-parametric test // Proc.11 th Berkeley Symp., 1950. – P.83-82.

44. *Hollander M.* Non-parametric Statistical Methods / M. Hollander, D. A. Wolfe. – 2nd ed. – New York : Wiley, 1999.
45. *Kiefer J.* K-Sample Analogues of the Kolmogorov-Smirnov and Cramer-v. Mises Tests // Annals of Mathematical Statistics, 1959. Vol. 30. No. 2. – P. 420-447.
46. *Klotz J.* Nonparametric tests for scale / J. Klotz // AMS. – 1962. – Vol. 33. – P. 498–512.
47. *Kruskal W. H.* Use of ranks in one-criterion variance analysis / W. H. Kruskal, W. A. Wallis // J. Amer. Statist. Assoc. – 1952. – Vol. 47. – P. 583–621.
48. *Kruskal W. H.* Use of ranks in one-criterion variance analysis / W. H. Kruskal, W. A. Wallis // J. Amer. Statist. Assoc. – 1953. – Vol. 48. – P. 907–911.
49. *Laubsher N. F.* Exact critical Values for Mood`s distribution-free test statistic for dispersion and its normal approximation / N. F. Laubsher, F. E. Steffens, E. M. De Lange // Technometrics. – 1968. – Vol. 10, № 3. – P. 497–508.
50. *Layard M.W.J.* Robust large-sample tests for homogeneity of variances // Journal of the American Statistical Association. – 1973. – Vol. 68. – P. 195–198.
51. *Lee H.B., Katz G.S., Restori A.F.* A Monte Carlo Study of Seven Homogeneity of Variance Tests // Journal of Mathematics and Statistics. 2010. Vol. 6, No. 3. P. 359-366.
52. *Lehman S.* Exact and approximate distributions for the Wilcoxon statistic with ties // Journal of the American Statistical Association. 1961. Vol. 56. – P. 293-988.
53. *Lehmann E. L.* Consistency and unbiasedness of certain nonparametric tests / E. L. Lehmann // Ann. Math. Statist. – 1951. – Vol. 22, № 1. – P. 165–179.
54. *Lemeshko B. Y.* Application of Homogeneity Tests: Problems and Solution / B. Y. Lemeshko, I. V. Veretelnikova, S. B. Lemeshko, A. Y. Novikova // In: Rykov V., Singpurwalla N., Zubkov A. (eds) Analytical and Computational Methods in Probability Theory. ACMPT 2017. Lecture Notes in Computer Science. : monograph. - Cham : Springer, 2017. - 10684. - P. 461-475.
55. *Lemeshko B.* Bartlett and Cochran tests in measurements with probability laws different from normal / B. Lemeshko, E. Mirkin // Measurement Techniques. – 2004. – Vol. 47, № 10. – P. 960–968.
56. *Lemeshko B. Yu.* Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. P. 1 / B. Yu. Lemeshko, S. B. Lemeshko // Measurement Techniques. – 2009. – Vol. 52, № 6. – P. 555–565.

57. *Lemeshko B. Yu.* Models for statistical distributions in nonparametric fitting tests on composite hypotheses based on maximum-likelihood estimators. P. II / B. Yu. Lemeshko, S. B. Lemeshko // Measurement Techniques. – 2009. – Vol. 52, № 8. – P. 799–812.
58. *Lemeshko B. Yu.* Power and robustness of criteria used to verify the homogeneity of means / B. Yu. Lemeshko, S. B. Lemeshko // Measurement Techniques. – 2008. – Vol. 51, № 9. – P. 950–959.
59. *Lemeshko B. Yu.* Statistical distribution convergence and homogeneity test power for Smirnov and Lehmann–Rosenblatt tests / B. Yu. Lemeshko, S. B. Lemeshko // Measurement Techniques – 2005. – Vol. 48, № 12. – P. 1159–1166.
60. *Lemeshko B.Yu.* Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria / B. Yu. Lemeshko, S. B. Lemeshko, A. A. Gorbunova // Measurement Techniques. – 2010. – Vol. 53, № 3. – P. 237–246.
61. *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. DOI: 10.1007/s11018-010-9530-x
62. *Lemeshko B.Y., Sataeva T.S.* Application and Power of Parametric Criteria for Testing the Homogeneity of Variances. Part III // Measurement Techniques, 2017. Vol. 60. No. 1. – P. 7-14. DOI: 10.1007/s11018-017-1141-3
63. *Lemeshko B.Y., Sataeva T.S.* Application and Power of Parametric Criteria for Testing the Homogeneity of Variances. Part IV // Measurement Techniques, 2017. Vol. 60. No. 5. – P. 425-431. DOI: 10.1007/s11018-017-1213-4
64. *Lemeshko B.Yu., Sataeva T.S.* On the Properties and Application of Tests for Homogeneity of Variances in the Problems of Metrology and Control // Advances in Intelligent Systems and Computing, Vol. 543, 2017. – P. 784-798. DOI: 10.1007/978-3-319-48923-0_84
65. *Lemeshko B.Yu., Gorbunova A.A., Lemeshko S.B., Rogozhnikov A.P*. Application of Nonparametric Goodness-of-fit tests for Composite Hypotheses in Case of Unknown Distributions of Statistics // Proceedings of the International Workshop “Applied Methods of Statistical Analysis. Applications in Survival Analysis, Reliability and Quality Control” – AMSA’2013, Novosibirsk, Russia, 25-27 September, 2013. P. 8-24.
66. *Lemeshko B.Yu., Lemeshko S.B., Rogozhnikov A.P*. Real-Time Studying of Statistic Distributions of Non-Parametric Goodness-of-Fit Tests when Testing Complex Hypotheses // Proceedings of the International Workshop “Applied Methods of Statistical Analysis. Simulations and Statistical Inference” – AMSA’2011, Novosibirsk, Russia, 20-22 September, 2011. P. 19-27.
67. *Leslie R.T., Brown B.M.* Use of range in testing heterogeneity of variance // Biometrika. 1966. Vol. 53. No.1/2. – P. 221-227.

68. *Levene H.* Robust tests for equality of variances // Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. – 1960. – P. 278-292.
69. *Levene* Test for Equality of Variances [Электронный ресурс] // e-Handbook of Statistical Methods. – Режим доступа : http://www.itl.nist.gov/div898/handbook/ eda/section3/eda35a.htm. –Загл. с экрана.
70. *Lim T.-S., Loh W.-Y.* A comparison of tests of equality of variances // Computational Statistics & Data Analysis. 1996. Vol. 22, No. 5. P. 287-301.
71. *Link R.F.* The sampling distribution of the ratio of two ranges from independent samples // The annals of mathematical statistics. 1950. Vol. 21, No. 1. – P. 112-116.
72. *Mann H. B.* On a test of whether one of two random variables is stochastically larger than the other / H. B. Mann, D. R. Whitney // Ann. Math. Statist. – 1947. – Vol. 18. – P. 50–60.
73. *Miller R.G.* Jackknifing variances // The Annals of Mathematical Statistics – 1968. – Vol. 39. – P. 567–582.
74. *Milton R. C.* An extended table of critical values for the Mann –Whitney (Wilcoxon) two-sample statistic / R. C. Milton // J. Amer. Statist. Ass. – 1964. – Vol. 59. – P. 925–934.
75. *Mood A.* On the asymptotic efficiency of certain nonparametric tests / A. Mood // AMS. – 1954. – Vol. 25. No. 3. – P. 514–522.
76. *Neel J. H.* A Monte Carlo Study of Levene`s Test of Homogeneity of Variance: Empirical Frequencies of Type I Error in Normal Distributions Paper : presented at the Annual Meeting of the American Educational Research Association Convention / J. H. Neel, W. M. Stallings. – Chicago, Illinois, 1974. – April.
77. *Newman D.* The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation // Biometrika. 1939. Vol. 31. No.1/2. – P. 20-30.
78. *Pettitt A.N.* A two-sample Anderson-Darling rank statistic // Biometrika. 1976. Vol. 63. No.1. P. 161-168.
79. *Rosenblatt M.* Limit theorems associated with variants of the von Mises statistic / M. Rosenblatt // Ann. Math. Statist. – 1952. – Vol. 23. – P. 617–623.
80. *Sataeva T. S., Lemeshko B.Yu.* About properties and power of classical tests of homogeneity of variances // Proceedings 2016 11th International Forum on Strategic Technology (IFOST), June 1-3, 2016, Novosibirsk, Russia. Part 1. – P. 350-354. DOI: 10.1109/IFOST.2016.7884125
81. *Scheffe H.* On Solutions of the Behrens-Fisher Problem, Based on the t-Distribution // The Annals of Mathematical Statistics. 1943. Vol. 14, No. 1. – P. 35-44.

82. *Scheffe H.* Practical Solutions of the Behrens-Fisher Problem // Journal of the American Statistical Association. 1970. Vol. 65. No. 332. – P. 1501-1508.
83. *Scholz F.W., Stephens M.A.* K-Sample Anderson–Darling Tests // Journal of the American Statistical Association. 1987. Vol. 82. No. 399. – P. 918-924.
84. *Siegel S.* A nonparametric sum of rank procedure for relative spread in unpaired samples / S. Siegel, J. W. Tukey // JASA. – 1960. – Vol. 55, № 291. – P. 429–445.
85. *Sukhatme B. V.* On certain Two-sample nonparametric tests for variances / B. V. Sukhatme // AMS. – 1957. – Vol. 28, № 1. – P. 188–194.
86. *Terry M.E.* Some rank order test which are most powerful against specific parametric alternatives // The Annals of Mathematical Statistics. 1952. Vol. 23. – P.346-366.
87. *O'Brien R.G.* Robust techniques for testing heterogeneity of variance effects in factorial designs // Psychometrika. 1978. Vol. 43, No. 3. P. 327-342.
88. *Overall J.E., Woodward J.A.* A simple test for heterogeneity of variance in complex factorial design // Psychometrika. 1974. Vol. 39. No. 3. – P. 311-318.
89. *Overall J.E., Woodward J.A.* A robust and powerful test for heterogeneity of variance // University of Texas Medical Branch Psychometric Laboratory. 1976.
90. *Parra-Frutos I.* The behaviour of the modified Levene's test when data are not normally distributed // Computational Statistics. – 2009. – Vol. 24. – P. 671–693.
91. *Welch B. L.* The Significance of the Difference Between Two Means when the Population Variances are Unequal // Biometrika. 1938. Vol. 29, No. 3/4. – P. 350-362.
92. *Welch B. L.* The generalization of "Student's" problem when several different population variances are involved // Biometrika. 1947. Vol. 34, No. 1/2. – P. 28-35.
93. *Wilcoxon F.* Individual comparisons by ranking methods / F. Wilcoxon // Biometrics Bulletin. – 1945. – № 1. – P. 80–83.
94. *Zhang J.* Powerful goodness-of-fit and multi-sample tests / J. Zhang // PhD Thesis. York University, Toronto. 2001. – 113 p. URL: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (дата обращения 28.01.2013).
95. *Zhang J.* Powerful Two-Sample Tests Based on the Likelihood Ratio / J. Zhang // Technometrics. – 2006. – V. 48. – No. 1. – P.95-103. DOI 10.1198/004017005000000328

96. *Zhang J., Wu Y.* k-Sample tests based on the likelihood ratio // Computational Statistics & Data Analysis. – 2007. – V. 51. – No. 9. – P. 4682-4691.
97. *Большев Л. Н.* Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – М. : Наука, 1983. – 416 с.
98. *Боровков А. А.* К задаче о двух выборках / А. А. Боровков // Изв. АН СССР, Сер. матем. – 1962. – Т. 26. – С. 605–624.
99. *Ван дер Варден Б.Л*. Математическая статистика. – М.: Иностранная литература, 1960. – 435 с.
100. *Гаек Я., Шидак З.* Теория ранговых критериев. – Главная редакция физико-математической литературы изд-ва «Наука», 1971. – 376 с.
101. *Закс Л.* Статистическое оценивание / Л. Закс. – М. : Статистика, 1976. – 598 с.
102. *Кобзарь А. И.* Прикладная математическая статистика. Для инженеров и научных работников / А. И. Кобзарь. – М. : Физматлит, 2006. – 816 с.
103. *Королюк В. С.* Асимптотический анализ распределений максимальных уклонений в схеме Бернулли // Теория вероятностей и ее применения. – 1959. – Т.4. – С. 369-397.
104. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.
105. *Лемешко Б.Ю.* О задаче идентификации закона распределения случайной составляющей погрешности измерений // Метрология. 2004. № 7. С. 8-17.
106. *Лемешко Б.Ю.* Непараметрические критерии согласия. Руководство по применению. М.: ИНФРА-М, 2014. 163 с. DOI: 10.12737/11873
107. *Лемешко Б.Ю.* Критерии проверки отклонения распределения от нормального закона. Руководство по применению. М.: ИНФРА-М, 2015. 160 с. (Научная мысль). DOI: 10.12737/6086
108. *Лемешко Б.Ю., Блинов П.Ю.* Критерии проверки отклонения распределения от равномерного закона. Руководство по применению. М.: ИНФРА-М, 2015. 183 с. (Научная мысль). DOI: 10.12737/11304
109. *Лемешко Б. Ю.* Статистический анализ одномерных наблюдений случайных величин : программная система / Б. Ю. Лемешко. – Новосибирск : Изд-во НГТУ, 1995. – 125 с.
110. *Лемешко Б. Ю.* Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. I / Б. Ю. Лемешко, С. Б. Лемешко // Измерительная техника. – 2009. – № 6. – С. 3–11.

111. *Лемешко Б. Ю.* Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. II / Б. Ю. Лемешко, С. Б. Лемешко // Измерительная техника. – 2009. – № 8. – С. 17–26.
112. *Лемешко Б. Ю.* О сходимости распределений статистик и мощности критериев однородности Смирнова и Лемана–Розенблатта / Б. Ю. Лемешко, С. Б. Лемешко // Измерительная техника. – 2005. – № 12. – С. 9–14.
113. *Лемешко Б.Ю.* О применении критериев проверки однородности законов распределения / Б.Ю. Лемешко, С.Б. Лемешко, И.В. Веретельникова // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – № 41. – С. 24-31.
114. *Лемешко Б. Ю.* Об устойчивости и мощности критериев проверки однородности средних / Б. Ю. Лемешко, С. Б. Лемешко // Измерительная техника. – 2008. – № 9. – С. 23–28.
115. *Лемешко Б. Ю.* О применении и мощности критериев проверки однородности дисперсий. Ч. I. Параметрические критерии / Б. Ю. Лемешко, С. Б. Лемешко, А. А. Горбунова // Измерительная техника. – 2010. – № 3. – С. 10–16.
116. *Лемешко Б. Ю.* О применении и мощности критериев проверки однородности дисперсий. Ч. II. Непараметрические критерии / Б. Ю. Лемешко, С. Б. Лемешко, А. А. Горбунова // Измерительная техника. – 2010. – № 5. – С. 11–18.
117. *Лемешко Б.Ю., Сатаева Т.С.* Применение и мощность параметрических критериев проверки однородности дисперсий. Ч. 4 // Измерительная техника. – 2017. – № 5. – С. 12-17.
118. *Лемешко Б.Ю., Сатаева Т.С.* Применение и мощность параметрических критериев проверки однородности дисперсий. Ч. 4 // Измерительная техника. – 2017. – № 5. – С. 12-17.
119. *Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н., Чимитова Е.В.* Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: Монография. Новосибирск : Изд-во НГТУ, 2011. 888 с.
120. *Лемешко Б. Ю.* Исследование критериев проверки гипотез, используемых в задачах управления качеством / Б. Ю. Лемешко, С. Б. Лемешко, Е. П. Миркин // АПЭП-2004. Актуальные проблемы электронного приборостроения : материалы VII междунар. конф. – Новосибирск, 2004. – Т. 6. – С. 269–272.
121. *Лемешко Б. Ю.* Критерии Бартлетта и Кокрена в измерительных задачах при вероятностных законах, отличающихся от нормального / Б. Ю. Лемешко, Е. П. Миркин // Измерительная техника. – 2004. – № 10. – С. 10–16.

122. *Лемешко Б. Ю.* Корреляционный анализ наблюдений многомерных случайных величин при нарушении предположений о нормальности / Б. Ю. Лемешко, С. С. Помадин // Сибирский журнал индустриальной математики. 2002. – Т. 5, № 3. – С. 115–130.
123. *Лемешко Б. Ю.* Проверка гипотез о математических ожиданиях и дисперсиях в задачах метрологии и контроля качества при вероятностных законах, отличающихся от нормального / Б. Ю. Лемешко, С. С. Помадин // Метрология, 2004. – № 3. – С. 3–15.
124. *Лемешко Б. Ю.* Исследование распределений статистик, используемых для проверки гипотез о равенстве дисперсий при законах ошибок, отличных от нормального / Б. Ю. Лемешко, В. М. Пономаренко // Науч. вест. НГТУ. – 2006. – № 2(23). – С. 21–33.
125. *Лемешко Б. Ю.* Компьютерные технологии анализа данных и исследования статистических закономерностей : учеб. пособие / Б. Ю. Лемешко, С. Н. Постовалов. – Новосибирск : Изд-во НГТУ, 2004. – 120 с.
126. *Лемешко Б. Ю.* Компьютерные технологии анализа данных и исследования статистических закономерностей : метод. указания к выполнению лаб. работ / Б. Ю. Лемешко, С. Н. Постовалов, С. Б. Лемешко. – Новосибирск : Изд-во НГТУ, 2007. – 71 с.
127. *Макаров А.А., Симонова Г.И.* Исследование мощности двухвыборочного критерия Андерсена–Дарлинга в случае засорения одной из выборок. // Статистические методы оценивания и проверки гипотез. Межвуз. сб. науч. тр. № 20, Перм. ун-т. – Пермь, 2007 – с. 40-52.
128. *Миттаг Х.-Й.* Статистические методы обеспечения качества / Х.-Й. Миттаг, Х. Ринне. – М. : Машиностроение. 1995. – 600 с.
129. *Орлов А. И.* О проверке однородности двух независимых выборок / А. И. Орлов // Завод. лаб. – 2003. – Т. 69, №. 1. – С. 55–60.
130. *Постовалов С.Н.* Применение компьютерного моделирования для расширения прикладных возможностей классических методов проверки статистических гипотез. Диссертация на соискание ученой степени доктора технических наук. Новосибирск. 2013. 298 с.
131. Программная система статистического анализа одномерных случайных величин – ISW. URL: http://ami.nstu.ru/~headrd/ISW.htm (дата обращения 12.07.2016)
132. Р 50.1.033–2001. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Ч. I. Критерии типа хи-квадрат. – М. : Изд-во стандартов, 2002. – 87 с.

133. Р 50.1.037–2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Ч. II. Непараметрические критерии. – М. : Изд-во стандартов, 2002. – 64 с.
134. *Смирнов Н. В.* Вероятности больших значений непараметрических односторонних критериев согласия / Н. В. Смирнов // Тр. матем. ин-та АН СССР. – 1961. – Т. 64. – С. 185–210.
135. *Смирнов Н. В.* Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках / Н. В. Смирнов // Бюл. МГУ, Серия А. – 1939. – Т. 2, № 2. – С. 3–14.
136. *Лемешко Б. Ю.* Мощность *k*-выборочных критериев проверки однородности законов / Б. Ю. Лемешко, И. В. Веретельникова // Измерительная техника. 2018. № 7. – С. 3-7.
137. *Pearson E.S., D'Agostino R.B., Bowman K.O.* Tests for departure from normality: Comparison of powers // Biometrika. 1977. – Vol. 64. – P. 231-246. DOI: 10.1093/biomet/64.2.427-a
138. *Tricker A.R.* The effect of rounding on the significance level of certain normal test statistics // Journal of Applied Statistics. – 1990. – Vol. 17. – No. 1. – P. 31-38. DOI: 10.1080/757582644
139. *Tricker A.R.* The effect of rounding on the power level of certain normal test statistics // Journal of Applied Statistics. – 1990. – Vol. 17. – No. 2. – P. 219-228. DOI: 10.1080/757582833
140. *Deidda R., Puliga M.* Sensitivity of goodness-of-fit statistics to rainfall data rounding off // Physics and Chemistry of the Earth. – 2006. – Vol. 31. – P. 1240-1251. DOI: 10.1016/j.pce.2006.04.041
141. *Choulakian V., Stephens M.A.* Goodness-of-Fit Tests for the Generalized Pareto Distribution // Technometrics. – 2001. – Vol. 43. – P. 478-484. DOI: 10.1198/00401700152672573
142. Лемешко Б. Ю. Лемешко С.Б., Семёнова М.А. К вопросу статистического анализа больших данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2018. № 44. С. 40-49. DOI: 10.17223/19988605/44/5
143. *Лемешко Б.Ю., Лемешко С.Б.* Влияние округления на свойства критериев проверки статистических гипотез // Автометрия. – 2020. – Т. 56, – № 3. – С. 35-45. DOI: 10.15372/AUT20200305
144. *Лемешко Б.Ю., Лемешко С.Б.* О влиянии ошибок округления на распределения статистик критериев согласия // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2020. – № 53. – С. 47-60. DOI: 10.17223/19988605/53/5
145. *Lemeshko B.Y., Lemeshko S.B.* About the effect of rounding on the properties of tests for testing statistical hypotheses // Journal of Physics: Conference Series. – 2021. – Vol. 1715. 012063. DOI: 10.1088/1742-6596/1715/1/012063

146. *Лемешко Б.Ю., Лемешко С.Б.* Проблемы применения непараметрических критериев согласия в задачах обработки результатов измерений // Системы анализа и обработки данных. – 2021. – № 2 (82). – С. 47-66. – DOI: 10.17212/2782-2001-2021-2-47-66.
147. *Fisher R.A.* The Use of Multiple Measurements in Taxonomic Problems // Annals of Eugenics. 1936. – Vol. 7. – P. 179-188.