

*Компьютерные технологии
анализа данных и исследования
статистических
закономерностей*

Лемешко Борис Юрьевич

E-mail: Lemeshko@fpm.ami.nstu.ru

Тел. 346-37-54

1. Практика применения методов статистического анализа в приложениях богата постановками задач, формулировки которых не укладываются в рамки классических предположений. Широкий спектр методов статистического анализа базируется на предположении о принадлежности ошибок измерений нормальному закону. **В реальных условиях предположение “нормальности”, а часто и другие предположения, не выполняются. Использование классических методов математической статистики в таких ситуациях может оказаться некорректным.**

2. Нельзя забывать, что многие классические результаты имеют асимптотический характер, в то время как на практике обычно имеют дело с выборками конечных объемов. **Далеко не всегда применение асимптотических результатов правомерно при ограниченных объемах выборок.**

3. Форма представления (регистрации) данных (измерений) зачастую не соответствуют рассматриваемым в учебниках по математической статистике точечным выборкам. На самом деле реальные наблюдения (выборки) могут быть группированными, цензурированными, частично группированными, интервальными. **Математический аппарат должен давать возможность анализа данных при любой форме представления данных, должен учитывать эту форму, не пренебрегать ею.**

Выявление фундаментальных статистических закономерностей в таких нестандартных условиях, как правило, является сложной задачей для исследователя.

Аналитические методы исследования свойств статистических оценок и распределений статистик критериев чрезвычайно трудоемки и не позволяют, вследствие сложности, обеспечить решение всего множества задач. Реальный выход видится в широком использовании численного подхода, связанного с компьютерным моделированием статистических закономерностей в условиях, имитирующих реальную ситуацию проведения измерений, с последующим построением математических моделей, приближенно описывающих полученные закономерности. Такой подход позволяет добиться хороших результатов там, где их не удастся достичь одними аналитическими методами.

Поэтому в последнее время все большее распространение получают методы компьютерного моделирования и анализа статистических закономерностей.

В настоящее время известно множество систем статистического анализа, широко используемых в различных приложениях, например SPSS, STATISTICA, SAS, JMP, Minitab, SYSTAT, Statgraphics, BMDP, S-Plus, Stadia, SigmaStat, StatXact, Statistix, Statit, STATlab, WinSTAT, Stat Navigator, STATA и т.д.

В базе CTI (The Computers in Teaching Initiative - организация, объединяющая британские университеты) содержится более 100 пакетов, реализующих системы статистического анализа, в том числе Minitab, CLIM, SPSS, Genstat, SAS System, BMDP, Statgraphics, Instat, S-Plus и т.д. В России из упомянутых наиболее популярно применение систем STATISTICA, SPSS, SAS.

Одни программные системы статистического анализа представляют собой универсальные системы, ориентированные на максимально широкий спектр методов анализа, другие предназначены для решения относительно узкого класса задач.

Следует подчеркнуть, что, как правило, все системы дают инструментарий для решения задач статистического анализа в различных приложениях. Но сферой применения этих систем не является исследование самих методов математической статистики, методов статистического анализа, они не могут служить инструментом для исследования закономерностей в математической статистике, для развития её математического аппарата.

Если проанализировать публикации, например, в журналах «Journal of Statistical Software», «Journal of Computational and Graphical Statistics», «Communication in Statistics», «Computational Statistics & Data Analysis» и других, то становится очевидной тенденция увеличения работ, в которых для исследования свойств оценок и статистик, для подтверждения аналитических выводов используются методы численного анализа и, в частности, методы статистического моделирования. То есть, **все чаще компьютерные технологии используют для совершенствования аппарата прикладной математической статистики.**

В российских же публикациях, к сожалению как правило, ограничиваются применением статистического моделирования для проверки аналитических выводов или предположений, а, иногда, просто ограничиваются пожеланиями осуществить такую проверку.

С применением компьютерного подхода, его развитием нами был получен ряд полезных для практики результатов, в частности:

1) Результаты наших исследований распределений статистик непараметрических критериев согласия при простых и сложных гипотезах о согласии с рядом законов, наиболее часто используемых в приложениях, построенные модели распределений статистик, для различных сложных гипотез, таблицы процентных точек вошли в разработанные рекомендации по стандартизации Госстандарта России Р 50.1.037–2002 "Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии. - М.: Изд-во стандартов. 2002. - 64 с.". Рекомендации призваны ликвидировать случаи некорректного применения критериев согласия при статистической обработке результатов наблюдений в различных приложениях.

В настоящий момент эти результаты уточнены и расширены.

2) Построены таблицы асимптотически оптимального группирования для достаточно широкого круга распределений, наиболее часто используемых в приложениях. Применение таблиц асимптотически оптимального группирования обеспечивает максимальную мощность критериев типа χ^2 при близких альтернативах. Исследована зависимость мощности от числа интервалов, впервые было показано, что существует оптимальное число интервалов, зависящее от объема выборки, конкретных альтернатив и способа группирования. Часть результатов вошла в рекомендации Госстандарта Р 50.1.033–2001 «Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа хи-квадрат. - М.: Изд-во стандартов. 2002. - 87 с.»

3) Были исследованы потери в информации Фишера, связанные с цензурированием выборок. Оказалось, что даже при значительной степени цензурирования в некоторых случаях сохраняется достаточно много информации, позволяющей получать хорошие оценки параметров закона. Методами компьютерного моделирования были исследованы законы распределения оценок максимального правдоподобия (ОМП) параметров ряда распределений по цензурированным наблюдениям при различной степени цензурирования и различных объемах полных выборок. Оказалось, что при ограниченных объемах выборок распределения ОМП (асимптотически эффективных) оказываются далекими от асимптотически нормального. Более того, распределения оказались асимметричными, а ОМП – смещенными. Была показана возможность построения поправок для ОМП, ликвидирующих смещение.

Эти результаты были замечены специалистами в области теории надежности.

4) Была показана возможность применения непараметрических критериев согласия типа Колмогорова, типа омега-квадрат Мизеса и Андерсона-Дарлинга для проверки адекватности непараметрических моделей законов распределения. Показано, что при использовании непараметрических оценок на распределения статистик критериев согласия влияет ряд факторов, определяющих сложную проверяемую гипотезу: закон распределения наблюдаемой случайной величины, соответствующий этой гипотезе; вид используемой ядерной функции; объем выборки; метод оценивания параметров размытости (Постовалов С.Н., Лемешко Б.Ю., Французов А.В. К применению непараметрических критериев согласия для проверки адекватности непараметрических моделей // Автометрия. 2002. - № 2. - С.3-14).

5) Методами статистического моделирования исследованы распределения классических статистик, используемых при проверке гипотез о математических ожиданиях и дисперсиях. Показано, что при проверке гипотез о математических ожиданиях применение классических результатов оказывается корректным при существенных отклонениях наблюдаемого закона от нормального.

Этот вывод справедлив и для параметрических критериев типа Стьюдента, применяемых для проверки гипотез об однородности средних 2-х выборок.

б) Для статистик, используемых в критериях проверки гипотез о дисперсиях, получены таблицы процентных точек, применение которых правомерно при наблюдаемых законах, описываемых экспоненциальным семейством распределений. Исследованы распределения классических статистик, используемых при проверке гипотез о дисперсиях в серии выборок. Для статистик, используемых в критериях Бартлетта и Кокрена, получены таблицы процентных точек, применение которых правомерно при наблюдаемых законах, описываемых экспоненциальным семейством распределений.

Средства исследования ряда таких критериев встроены в программное обеспечение, используемое в научных исследованиях и учебном процессе.

7) Получены таблицы процентных точек для статистик критериев типа Граббса при проверке на выброс одновременно трех максимальных (трех минимальных) значений и одновременно минимального и максимального значений в выборке. Методами статистического моделирования исследованы распределения статистик критериев Граббса, используемых в задачах отбраковки аномальных измерений, при отклонениях наблюдаемого закона от нормального (Лемешко С.Б., Лемешко Б.Ю. Расширение области применения критериев типа Граббса, используемых при отбраковке аномальных измерений // Измерительная техника. 2005. № 6. – С. 13-19).

Разработанное программное обеспечение позволяет исследовать распределения данных статистик при различных законах распределения и различном количестве аномальных измерений.

8) Исследованы распределения статистик и мощность ряда критериев проверки отклонения от нормального закона, в том числе, не вошедших в ГОСТ Р ИСО 5479-2002. Показаны недостатки и преимущества различных критериев. Впервые выявлена смещенность при малых объемах выборок ряда критериев, в том числе критериев Шапиро-Уилка и Эппса-Палли, относительно симметричных альтернатив со значением эксцесса, меньшем трех. Проведено сравнение мощности рассмотренных критериев с критериями согласия (Лемешко Б.Ю., Лемешко С.Б. Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. № 2. – С. 3-24).

В настоящее время проводится исследование ещё ряда “критериев нормальности”.

9) Исследована мощность критериев однородности двух выборок Смирнова и Лемана-Розенблатта.

Недостатком критерия Смирнова является то, что дискретное распределение статистики сходится к предельному распределению Колмогорова **слева** и при малых объемах анализируемых выборок значительно отличается от предельного. Предложена поправка к статистике Смирнова, улучшающая сходимость распределения статистики к предельному (Лемешко С.Б., Лемешко Б.Ю. О сходимости распределений статистик и мощности критериев однородности Смирнова и Лемана-Розенблатта // Измерительная техника. 2005. № 12. – С.9-14).

10) Разработана методика моделирования распределений статистик многомерных случайных величин. Реализована универсальная процедура, позволяющая на базе экспоненциального семейства распределений моделировать псевдослучайные величины с заданными математическим ожиданием и ковариационной матрицей, распределенные как по многомерному нормальному закону, так и по законам отличным от нормального. Показано, что распределения статистик, используемых при проверке гипотез о векторе математических ожиданий, устойчивы к отклонениям многомерного закона от нормального в достаточно широких пределах: значимого изменения распределений статистик не происходит.

Показано, что распределения статистик критериев, используемых при проверке гипотез о ковариационной матрице, существенно зависят от вида наблюдаемого многомерного закона. В случае принадлежности наблюдений m -мерным законам, хорошо описываемым моделями, получаемыми в соответствии с разработанной процедурой моделирования, для распределений статистик критериев (при известном и неизвестном векторе математических ожиданий) найдены аналитические модели законов, описывающие распределения этих статистик при определенных значениях размерности многомерной величины и параметре формы экспоненциального семейства распределений.

Показано, что распределения статистик критериев, используемых при проверке гипотез вида о незначимости парных, частных и множественных коэффициентов корреляции, устойчивы к отклонениям наблюдаемого многомерного закона от нормального. Эмпирические распределения данных статистик по-прежнему хорошо описываются предельными законами, полученными в предположении о нормальности наблюдаемых величин. В то же время, в случае многомерных законов с «тяжелыми хвостами» наблюдается значимое отличие распределений статистик соответствующих критериев от предельных классических. Используемые в критериях проверки гипотез о равенстве заданному значению парного или частного коэффициента корреляции статистики соответствующих критериев существенно зависят от наблюдаемого многомерного закона.

11) Разработаны средства для моделирования и исследования законов распределения произвольных функций случайных величин и функций систем случайных величин (пока только независимых), и для построения для этих законов приближенных моделей (Огурцов Д.В., Лемешко Б.Ю. Статистическое моделирование как эффективный инструмент для исследования законов распределения функций случайных величин // Метрология. 2007. – № 5. – С. 3-13).

2. Проблемы применения непараметрических критериев согласия

2.1. Введение

К сожалению, практика применения непараметрических критериев согласия богата большим числом примеров их некорректного использования, особенно, в литературных источниках учебного характера. Наиболее типичные ошибки связаны с применением классических результатов, имеющих силу при проверке простых гипотез, для ситуаций, соответствующих проверке сложных гипотез [1].

Напомним, 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 . Проблемы возникают, если при проверке сложной гипотезы оценку параметра распределения θ вычисляют по той же самой выборке, по которой проверяют согласие. Если оценку θ вычисляют по некоторой другой выборке, то применение критерия не отличается от ситуации с проверкой простой гипотезы.

Далее, как правило, мы будем предполагать, что при проверке сложных гипотез оценка параметра θ вычисляется по этой же выборке.

Очевидно, что на практике при обработке результатов измерений с проблемой проверки сложных гипотез чаще всего сталкиваются именно в такой ситуации, так как сначала оценивают по выборке параметры модели, чтобы лучше подогнать ее к наблюдаемым данным, а потом проверяют адекватность полученной модели.

В процессе проверки согласия по выборке вычисляют значение S^* статистики используемого критерия. Затем для того, чтобы сделать вывод о принятии или отклонении гипотезы H_0 , необходимо знать условное распределение $G(S|H_0)$ статистики S критерия при справедливости H_0 . И если вероятность

$$P\{S > S^*\} = \int_{S^*}^{+\infty} g(s|H_0) ds \quad (1)$$

достаточно большая, по крайней мере $P\{S > S^*\} > \alpha$, где $g(s|H_0)$ – условная плотность, а α – задаваемый уровень значимости (вероятность ошибки 1-го рода – отклонить справедливую гипотезу H_0), то принято считать, что нет оснований для отклонения гипотезы H_0 .

Если в процессе анализа выборки рассматривают некоторую альтернативу H_1 : $F(x) = F_1(x, \theta)$, то с ней связывают условное распределение $G(S|H_1)$ и вероятность ошибки 2-го рода β (принять гипотезу H_0 , в то время как верна гипотеза H_1). Задание значения α для применяемого критерия согласия однозначно определяет и значение β :

$$\alpha = \int_{S_\alpha}^{+\infty} g(s|H_0) ds, \quad (2)$$

$$\beta = \int_0^{S_\alpha} g(s|H_1) ds. \quad (3)$$

При этом, чем больше мощность критерия $1 - \beta$, тем лучше он различает соответствующие гипотезы.

2.2. Распределения статистик непараметрических критериев при простых гипотезах

2.2.1. Критерий Колмогорова

В случае простых гипотез предельные распределения статистик рассматриваемых критериев согласия Колмогорова, Смирнова, ω^2 и Ω^2 Мизеса известны и не зависят от вида наблюдаемого закона распределения и, в частности, от его параметров. Считают, что эти критерии являются “свободными от распределения”. Это достоинство предопределило широкое использование данных критериев в различных приложениях.

Предельное распределение статистики

$$D_n = \sup_{|x| < \infty} |F_n(x) - F(x, \theta)|, \quad (4)$$

где $F_n(x)$ – эмпирическая функция распределения, $F(x, \theta)$ – теоретическая функция распределения, n – объем выборки, было получено Колмогоровым в [2]. При $n \rightarrow \infty$ функция распределения статистики $\sqrt{n}D_n$ сходится равномерно к функции распределения Колмогорова

$$K(s) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 s^2}. \quad (5)$$

Наиболее часто в критерии Колмогорова (Колмогорова-Смирнова) используют статистику с поправкой Большева вида [3]

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

где

$$D_n = \max(D_n^+, D_n^-), \quad (7) \quad D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}, \quad (8) \quad D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}, \quad (9)$$

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

Если для вычисленного по выборке значения статистики S_K^* выполняется неравенство

$$P\{S > S_K^*\} = 1 - K(S_K^*) > \alpha,$$

то нет оснований для отклонения гипотезы H_0 .

2.2.2. Критерий Смирнова

В критерии Смирнова используют статистику

$$D_n^+ = \sup_{|x| < \infty} (F_n(x) - F(x, \theta)) \quad (10)$$

или статистику

$$D_n^- = - \inf_{|x| < \infty} (F_n(x) - F(x, \theta)), \quad (11)$$

значения которых вычисляют по эквивалентным соотношениям (8), (9).

Реально в критерии обычно используют статистику [3]

$$S_m = \frac{(6nD_n^+ + 1)^2}{9n}, \quad (12)$$

которая при простой гипотезе в пределе подчиняется распределению χ^2 с числом степеней свободы, равным 2.

Гипотезу H_0 не отвергают, если для вычисленного по выборке значения статистики S_m^*

$$P\{S_m > S_m^*\} = \int_{S_m^*}^{\infty} \frac{1}{2} e^{-x/2} dx = e^{-S_m^*/2} > \alpha.$$

2.2.3. Критерии ω^2

В критериях типа ω^2 расстояние между гипотетическим и истинным распределениями рассматривают в квадратичной метрике.

Проверяемая гипотеза H_0 имеет вид [3]

$$H_0: \int_{-\infty}^{\infty} \{E[F_n(x)] - F(x)\}^2 \psi(F(x)) dF(x) = 0 \quad (13)$$

при альтернативной гипотезе

$$H_1: \int_{-\infty}^{\infty} \{E[F_n(x)] - F(x)\}^2 \psi(F(x)) dF(x) > 0, \quad (14)$$

где $E[\cdot]$ - оператор математического ожидания, $\psi(t)$ - заданная на отрезке $0 \leq t \leq 1$ неотрицательная функция, относительно которой предполагают, что $\psi(t)$, $t\psi(t)$, $t^2\psi(t)$ интегрируемы на отрезке $0 \leq t \leq 1$ [4]. Статистику критерия [3] выражают соотношением

$$\begin{aligned} \omega_n^2[\psi(F)] &= \int_{-\infty}^{\infty} \{E[F_n(x)] - F(x)\}^2 \psi(F(x)) dF(x) = \\ &= \frac{2}{n} \sum_{i=1}^n \left\{ g[F(x_i)] - \frac{2i-1}{2n} f[F(x_i)] \right\} + \int_0^1 (1-t)^2 \psi(t) dt, \end{aligned} \quad (15)$$

где

$$f(t) = \int_0^t \psi(s) ds, \quad g(t) = \int_0^t s\psi(s) ds.$$

При выборе $\psi(t) \equiv 1$ для критерия ω^2 Мизеса получают статистику вида (статистику **Крамера-Мизеса-Смирнова**)

$$S_{\omega} = n\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left\{ F(x_i, \theta) - \frac{2i-1}{2n} \right\}^2, \quad (16)$$

которая при простой гипотезе в пределе подчиняется закону с функцией распределения $a1(s)$, имеющей вид [3]

$$a1(s) = \frac{1}{\sqrt{2s}} \sum_{j=0}^{\infty} \frac{\Gamma(j+1/2)\sqrt{4j+1}}{\Gamma(1/2)\Gamma(j+1)} \exp\left\{-\frac{(4j+1)^2}{16s}\right\} \times \\ \times \left\{ I_{-\frac{1}{4}}\left[\frac{(4j+1)^2}{16s}\right] - I_{\frac{1}{4}}\left[\frac{(4j+1)^2}{16s}\right] \right\}, \quad (17)$$

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

$$I_{\nu}(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}, \quad |z| < \infty, \quad |\arg z| < \pi. \quad (18)$$

При выборе $\psi(t) \equiv 1/t(1-t)$ для критерия Ω^2 Мизеса статистика приобретает вид (статистика **Андерсона-Дарлинга**)

При выборе $\psi(t) \equiv 1/t(1-t)$ для критерия Ω^2 Мизеса статистика приобретает вид (статистика **Андерсона-Дарлинга**)

$$S_{\Omega} = n\Omega_n^2 = -n - 2 \sum_{i=1}^n \left\{ \frac{2i-1}{2n} \ln F(x_i, \theta) + \left(1 - \frac{2i-1}{2n}\right) \ln(1 - F(x_i, \theta)) \right\}. \quad (19)$$

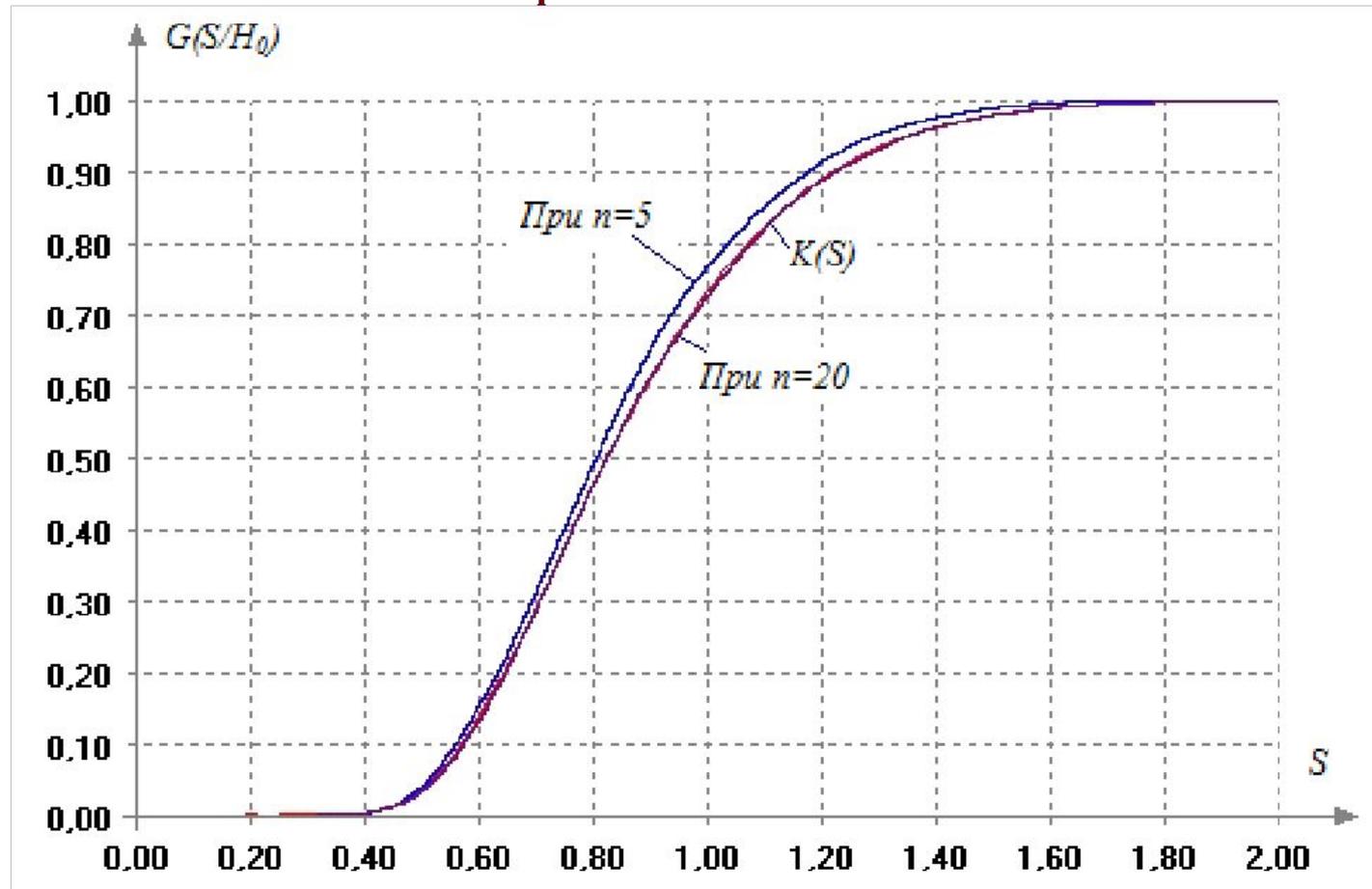
В пределе эта статистика подчиняется закону с функцией распределения $a2(s)$, имеющей вид [3]

$$a2(s) = \frac{\sqrt{2\pi}}{s} \sum_{j=0}^{\infty} (-1)^j \frac{\Gamma(j+1/2)(4j+1)}{\Gamma(1/2)\Gamma(j+1)} \exp\left\{-\frac{(4j+1)^2 \pi^2}{8s}\right\} \times \\ \times \int_0^{\infty} \exp\left\{\frac{s}{8(y^2+1)} - \frac{(4j+1)^2 \pi^2 y^2}{8s}\right\} dy. \quad (20)$$

Гипотезы о согласии не отвергают, если выполнены неравенства

$$P\{S_{\omega} > S_{\omega}^*\} = 1 - a1(S_{\omega}^*) > \alpha \quad \text{и} \quad P\{S_{\Omega} > S_{\Omega}^*\} = 1 - a2(S_{\Omega}^*) > \alpha.$$

Сходимость распределения статистики Колмогорова к предельному при проверке простых гипотез



2.3 Непараметрические критерии согласия при сложных гипотезах

2.3.1 Потеря критериями свойства “свободы от распределения”

При проверке сложных гипотез, когда по той же самой выборке оценивают параметры наблюдаемого закона распределения вероятностей, непараметрические критерии согласия Колмогорова, Смирнова, ω^2 и Ω^2 Мизеса теряют свойство “свободы от распределения”. В этом случае предельные распределения статистик этих критериев будут зависеть от закона, которому подчинена наблюдаемая выборка. Более того, распределения статистик непараметрических критериев согласия зависят и от используемого метода оценивания параметров. Следует также учитывать, что распределения статистик существенно зависят от объема выборки.

Игнорирование того, что проверяют сложную гипотезу, игнорирование различия в сложных гипотезах приводят 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 [5]-[7].

Точкой отсчета, с которой были начаты исследования предельных распределений статистик непараметрических критериев согласия при сложных гипотезах, послужила работа [8].

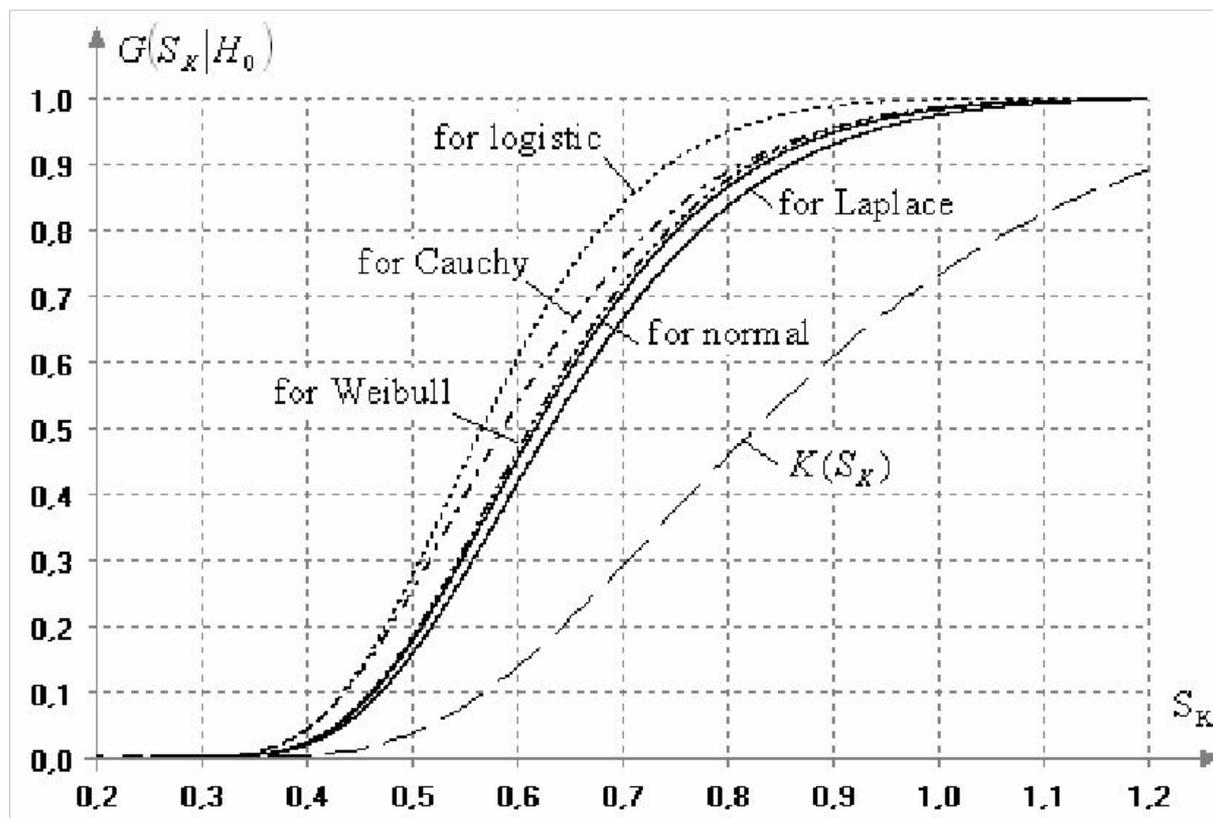
2.3.3. Факторы, влияющие на распределения статистик критериев при проверке сложных гипотез

Распределения статистик непараметрических критериев согласия при проверке сложных гипотез зависят от характера этой сложной гипотезы. На закон распределения статистики $G(S|H_0)$ влияют следующие факторы, определяющие “сложность” гипотезы:

- 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, бета-распределений);
- используемый метод оценивания параметров.

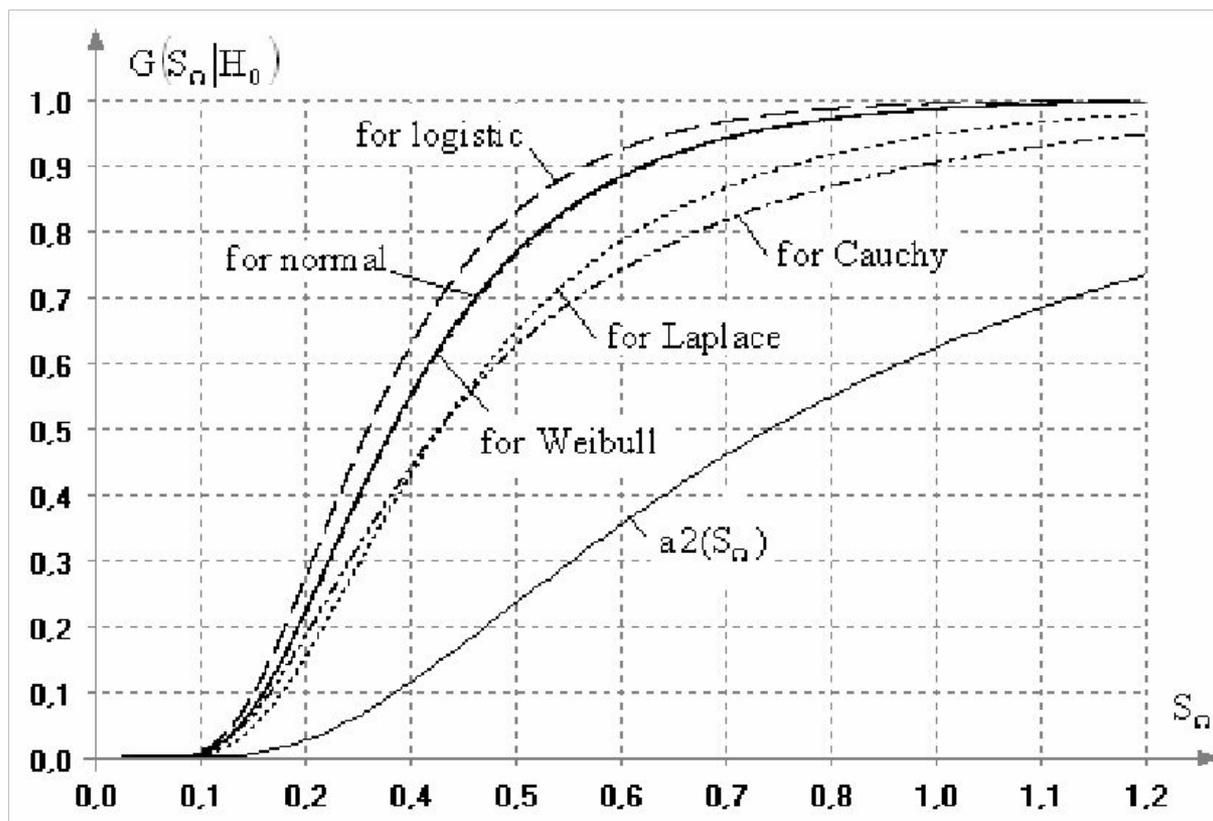
При малых объемах выборки n распределение $G(S_n|H_0)$ зависит от n . Однако существенная зависимость распределения статистики от n наблюдается только при небольших объемах выборки. Уже при $n \geq 15-20$ распределение $G(S_n|H_0)$ достаточно близко к предельному $G(S|H_0)$ и зависимостью от n можно пренебречь.

Зависимость распределений статистик при проверке сложных гипотез от вида закона



Распределения статистики Колмогорова (1) при проверке сложных гипотез с вычислением ОМП 2-х параметров закона

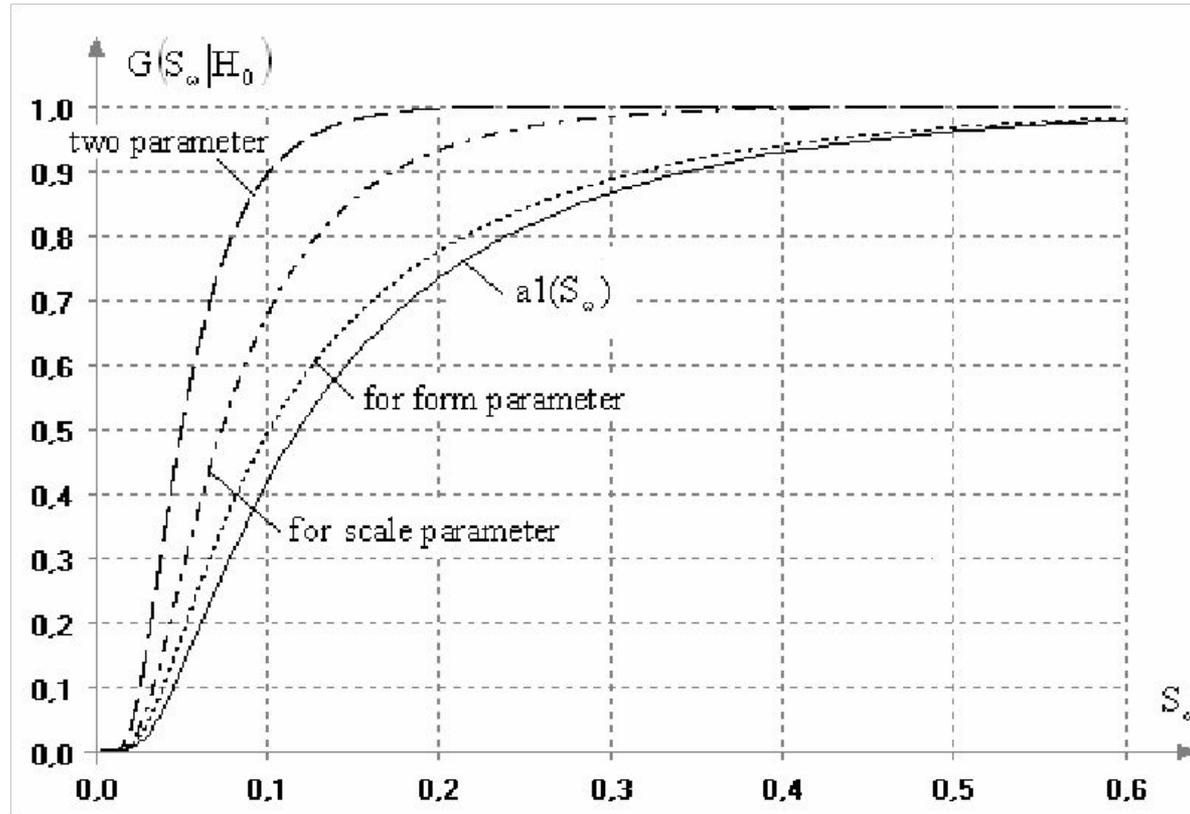
Зависимость распределений статистик при проверке сложных гипотез от вида закона



Распределения статистики Андерсона-Дарлингга при проверке сложных гипотез при вычислении ОМП 2-х параметров закона

Зависимость распределений статистик от количества и вида оцененных параметров

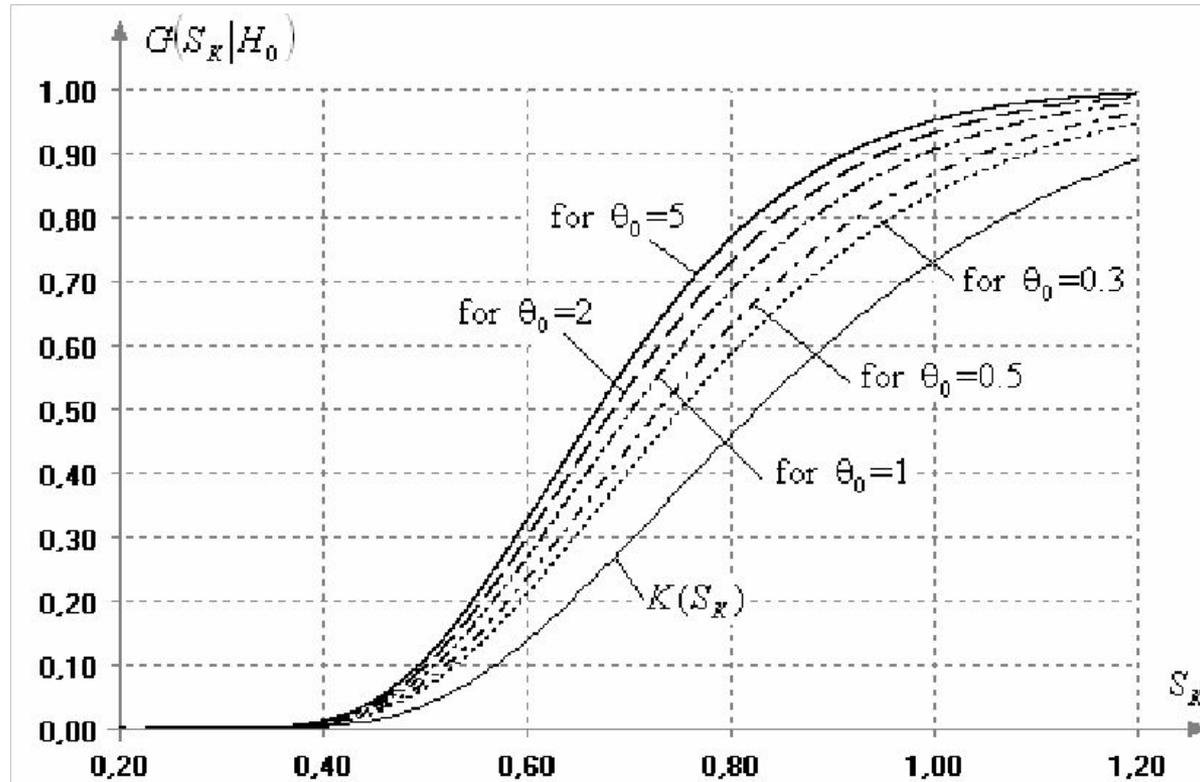
$$f(x, \theta) = \frac{\theta_0 x^{\theta_0 - 1}}{\theta_1^{\theta_0}} \exp\left\{-\left(\frac{x}{\theta_1}\right)^{\theta_0}\right\}$$



Распределения статистики Крамера-Мизеса-Смирнова при проверке сложных гипотез относительно распределения Вейбулла при вычислении ОМП различных параметров закона

Зависимость распределений статистик при проверке сложных гипотез от значения параметра (формы)

$$f(x, \theta) = \frac{x^{\theta_0 - 1}}{\theta_1^{\theta_0} \Gamma(\theta_0)} \exp\left(-\frac{x}{\theta_1}\right)$$



Распределения статистики Колмогорова при проверке сложных гипотез при вычислении ОМП только параметра масштаба гамма-распределения в зависимости от значения параметра формы

Распределения $G(S|H_0)$ статистики Колмогорова хорошо аппроксимируются семейством гамма-распределений с функцией плотности

$$\gamma(\theta_0, \theta_1, \theta_2) = \frac{1}{\theta_1^{\theta_0} \Gamma(\theta_0)} (x - \theta_2)^{\theta_0 - 1} e^{-(x - \theta_2)/\theta_1}.$$

А распределения статистик Крамера-Мизеса-Смирнова и Андерсона-Дарлинга неплохо приближаются семейством распределений Sb -Джонсона

$$Sb(\boldsymbol{\theta}) = \frac{\theta_1 \theta_2}{(x - \theta_3)(\theta_2 + \theta_3 - x)} \exp \left\{ -\frac{1}{2} \left[\theta_0 - \theta_1 \ln \frac{x - \theta_3}{\theta_2 + \theta_3 - x} \right]^2 \right\}.$$

Распределения всех статистик при проверке сложных гипотез относительно гамма-распределения хорошо приближаются семейством бета-распределений III-го рода

$$B_3(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4) = \frac{\theta_2^{\theta_0}}{\theta_3 B(\theta_0, \theta_1)} \frac{\left(\frac{x - \theta_4}{\theta_3} \right)^{\theta_0 - 1} \left(1 - \frac{x - \theta_4}{\theta_3} \right)^{\theta_1 - 1}}{\left[1 + (\theta_2 - 1) \frac{x - \theta_4}{\theta_3} \right]^{\theta_0 + \theta_1}}.$$

Table. Models of limiting statistic distributions of nonparametric goodness-of-fit when MLE are used.

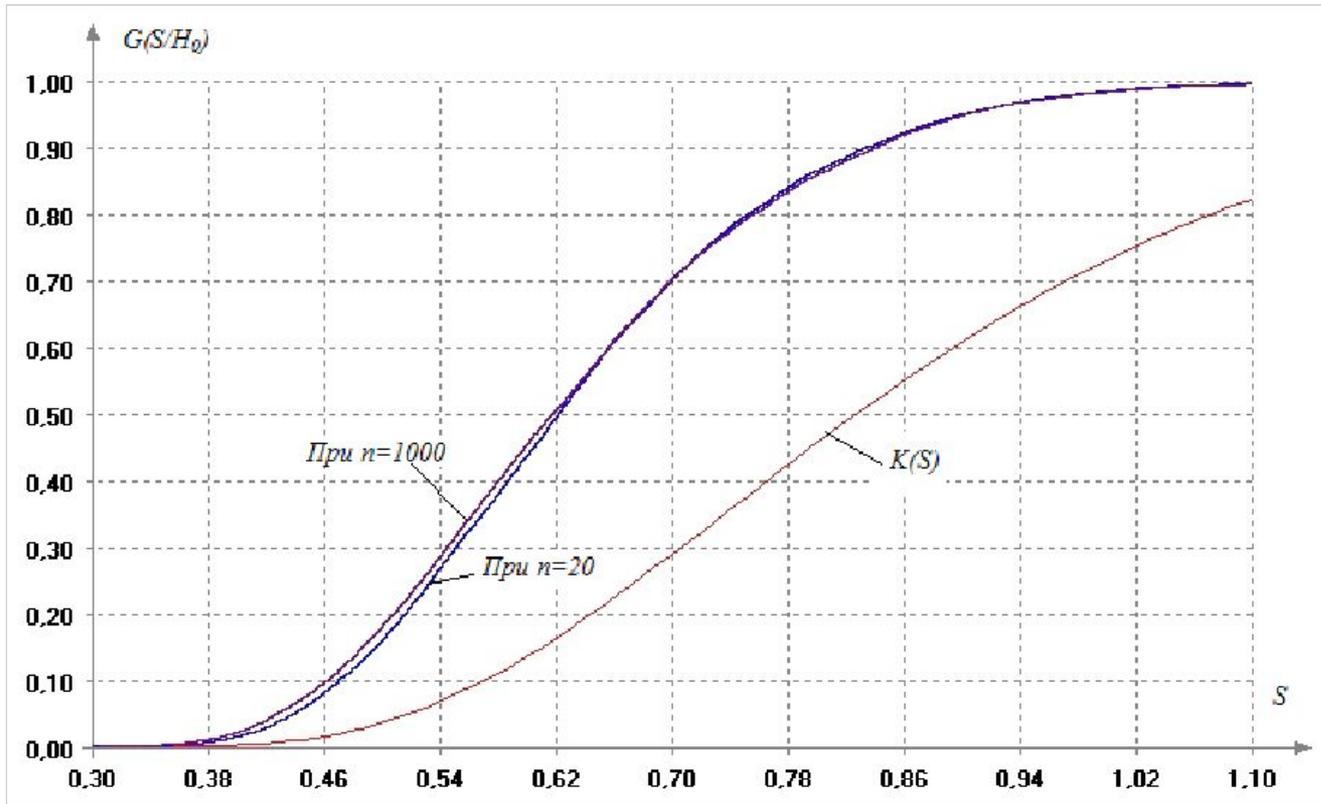
Test	Random variable distribution	Estimation of scale parameter	Estimation of shift parameter	Estimation of two parameters
Kolmogorov's	Exponential & Rayleigh	$\gamma(5.1092; 0.0861; 0.2950)$	–	–
	Seminormal	$\gamma(4.5462; 0.1001; 0.3100)$	–	–
	Maxwell	$\gamma(5.4566; 0.0794; 0.2870)$	–	–
	Laplace	$\gamma(3.3950; 0.1426; 0.3405)$	$\gamma(6.2887; 0.0718; 0.2650)$	$\gamma(6.2949; 0.0624; 0.2613)$
	Normal & Log-normal	$\gamma(3.5609; 0.1401; 0.3375)$	$\gamma(7.5304; 0.0580; 0.2400)$	$\gamma(6.4721; 0.0580; 0.2620)$
	Cauchy	$\gamma(3.0987; 0.1463; 0.3350)$	$\gamma(5.9860; 0.0780; 0.2528)$	$\gamma(5.3642; 0.0654; 0.2600)$
	Logistic	$\gamma(3.4954; 0.1411; 0.3325)$	$\gamma(7.6325; 0.0531; 0.2368)$	$\gamma(7.5402; 0.0451; 0.2422)$
	Extreme-value & Weibull	$\gamma(3.6805; 0.1355; 0.3350)^{1)}$	$\gamma(5.2194; 0.0848; 0.2920)^{2)}$	$\gamma(6.6012; 0.0563; .2598)$
Cramer-Mises-Smirnov's	Exponential & Rayleigh	$Sb(3.3738; 1.2145; 1.0792; 0.011)$	–	–
	Seminormal	$Sb(3.527; 1.1515; 1.5527; 0.012)$	–	–
	Maxwell	$Sb(3.353; 1.220; 0.9786; 0.0118)$	–	–
	Laplace	$Sb(3.2262; 0.9416; 2.703; 0.015)$	$Sb(2.9669; 1.2534; 0.6936; 0.01)$	$Sb(3.768; 1.2865; 0.8336; 0.0113)$
	Normal & Log-normal	$Sb(3.153; 0.9448; 2.5477; 0.016)$	$Sb(3.243; 1.315; 0.6826; 0.0095)$	$Sb(4.3950; 1.4428; 0.915; 0.009)$
	Cauchy	$Sb(3.1895; 0.9134; 2.690; 0.013)$	$Sb((2.359; 1.0732; 0.595; 0.0129)$	$Sb(3.4364; 1.0678; 1.000; 0.011)$
	Logistic	$Sb(3.264; 0.9581; 2.7046; 0.014)$	$Sb(4.0026; 1.2853; 1.00; 0.0122)$	$Sb(3.2137; 1.3612; 0.36; 0.0105)$
	Extreme-value & Weibull	$Sb(3.343; 0.9817; 2.753; 0.015)^{1)}$	$Sb(3.498; 1.2236; 1.1632; 0.01)^{2)}$	$Sb(3.3854; 1.4453; 0.4986; 0.007)$
Anderson-Darling's	Exponential & Rayleigh	$Sb(3.8386; 1.3429; 7.500; 0.090)$	–	–
	Seminormal	$Sb(4.2019; 1.2918; 11.500; 0.100)$	–	–
	Maxwell	$Sb(3.9591; 1.3296; 7.800; 0.1010)$	–	–
	Laplace	$Sb(4.3260; 1.0982; 27.00; 0.110)$	$Sb(3.1506; 1.3352; 4.9573; 0.096)$	$Sb(3,8071; 1,3531; 5,1809; 0,10)$
	Normal & Log-normal	$Sb(4.3271; 1.0895; 28.000; 0.120)$	$Sb(3.3085; 1.4043; 4.2537; 0.080)$	$Sb(3.5601; 1.4846; 3.0987; 0.08)$
	Cauchy	$Sb(3.7830; 1.0678; 18.0; 0.11)$	$Sb(3.4814; 1.2375; 7.810; 0.1)$	$Sb(3.290; 1.129; 5.837; 0.099)$
	Logistic	$Sb(3.516; 1.054; 14.748; 0.117)$	$Sb(5.1316; 1.5681; 10.0; 0.065)$	$Sb(3.409; 1.434; 2.448; 0.095)$
	Extreme-value & Weibull	$Sb(3.512; 1.064; 14.496; 0.125)^{1)}$	$Sb(4.799; 1.402; 13.0; 0.085)^{2)}$	$Sb(3.4830; 1.5138; 3.00; 0.07)$

Note. ¹⁾ - we estimated the Weibull distribution form parameter, ²⁾ - the Weibull distribution scale parameter.

Table. Models of limiting statistic distributions of the nonparametric goodness-of-fit when MLE are used in the case of gamma-distribution.

Test	Value of the form parameter	Estimation of scale parameter	Estimation of form parameter	Estimation of two parameters
Kolmogorov's	0.3	$B_3(6.3045; 5.9555; 3.0350; 1.3170; 0.281)$	$B_3(6.4536; 5.7519; 3.3099; 1.6503; 0.280)$	$B_3(6.9705; 5.6777; 3.6297; 1.5070; 0.270)$
	0.5	$B_3(6.9356; 5.0081; 4.3582; 1.8470; 0.280)$	$B_3(6.3860; 5.9685; 3.1228; 1.6154; 0.280)$	$B_3(6.4083; 5.9339; 3.2063; 1.4483; 0.2774)$
	1.0	$B_3(6.7187; 5.3740; 3.7755; 1.6875; 0.282)$	$B_3(6.1176; 6.4704; 2.6933; 1.5501; 0.280)$	$B_3(5.6031; 6.1293; 2.7065; 1.3607; 0.2903)$
	2.0	$B_3(5.8359; 22.6032; 2.1921; 4.00; 0.282)$	$B_3(6.1387; 6.5644; 2.6021; 1.4840; 0.280)$	$B_3(5.8324; 6.1446; 2.7546; 1.3280; 0.2862)$
	3.0	$B_3(5.9055; 24.4312; 2.0996; 4.00; 0.282)$	$B_3(6.1221; 6.6131; 2.5536; 1.4590; 0.280)$	$B_3(6.0393; 6.1276; 2.8312; 1.3203; 0.2827)$
	4.0	$B_3(5.9419; 27.1264; 1.9151; 4.00; 0.282)$	$B_3(6.0827; 6.7095; 2.4956; 1.4494; 0.280)$	$B_3(6.1584; 6.1187; 2.8748; 1.3170; 0.2807)$
Cramer-Mises-Smirnov's	0.3	$B_3(3.2722; 1.9595; 16.1768; 0.750; 0.013)$	$B_3(3.0247; 3.2256; 11.113; 0.7755; 0.0125)$	$B_3(2.3607; 4.0840; 7.0606; 0.6189; 0.0145)$
	0.5	$B_3(3.2296; 2.1984; 14.3153; 0.700; 0.013)$	$B_3(3.0143; 3.3504; 10.095; 0.7214; 0.0125)$	$B_3(2.7216; 3.9844; 7.4993; 0.5372; 0.013)$
	1.0	$B_3(3.1201; 2.5460; 11.1200; 0.600; 0.013)$	$B_3(2.9928; 3.4716; 8.8275; 0.6346; 0.0125)$	$B_3(3.0000; 3.8959; 7.3247; 0.4508; 0.012)$
	2.0	$B_3(2.9463; 3.1124; 9.1160; 0.600; 0.013)$	$B_3(2.9909; 3.5333; 8.2010; 0.5786; 0.0125)$	$B_3(3.0533; 3.9402; 7.1173; 0.4246; 0.0118)$
	3.0	$B_3(2.8840; 3.3796; 8.4342; 0.600; 0.013)$	$B_3(2.9737; 3.5528; 7.8843; 0.5549; 0.0125)$	$B_3(3.0703; 3.9618; 7.034; 0.4163; 0.0117)$
	4.0	$B_3(2.8522; 3.5285; 8.1044; 0.600; 0.013)$	$B_3(2.9677; 3.5426; 7.7632; 0.5418; 0.0125)$	$B_3(3.0967; 3.9539; 7.064; 0.4122; 0.0116)$
Anderson-Darling's	0.3	$B_3(3.3848; 2.8829; 14.684; 6.0416; 0.1088)$	$B_3(3.1073; 3.7039; 8.6717; 4.3439; 0.1120)$	$B_3(4.5322; 4.060; 10.0718; 2.9212; 0.078)$
	0.5	$B_3(5.0045; 2.9358; 18.8524; 5.2436; 0.077)$	$B_3(3.1104; 3.7292; 8.0678; 4.0132; 0.1120)$	$B_3(5.0079; 4.056; 10.0292; 2.5872; 0.073)$
	1.0	$B_3(5.0314; 3.1848; 15.4626; 4.3804; 0.077)$	$B_3(3.1149; 3.7919; 7.4813; 3.6770; 0.1120)$	$B_3(5.0034; 4.1093; 9.1610; 2.3427; 0.073)$
	2.0	$B_3(4.9479; 3.3747; 13.0426; 3.8304; 0.077)$	$B_3(3.0434; 4.1620; 7.1516; 3.8500; 0.1120)$	$B_3(4.9237; 4.2091; 8.6643; 2.2754; 0.073)$
	3.0	$B_3(5.0367; 3.4129; 12.9013; 3.6867; 0.077)$	$B_3(3.0565; 3.9092; 6.7844; 3.3972; 0.1120)$	$B_3(4.9475; 4.2070; 8.6686; 2.2512; 0.073)$
	4.0	$B_3(4.9432; 3.5038; 12.2240; 3.6302; 0.077)$	$B_3(3.0531; 3.9437; 6.7619; 3.3993; 0.1120)$	$B_3(4.9274; 4.2279; 8.5573; 2.2390; 0.073)$
	5.0	$B_3(4.8810; 3.5762; 11.7894; 3.6051; 0.077)$	$B_3(3.0502; 3.9640; 6.7510; 3.4024; 0.1120)$	$B_3(4.9207; 4.2432; 8.4881; 2.2314; 0.073)$

Сходимость распределения статистики Колмогорова к предельному при проверке сложных гипотез



При проверке сложной гипотезы о согласии с нормальным законом при вычислении ОМП 2-х параметров

Выводы

- 1. Надо быть очень внимательным к факту, какая гипотеза проверяется, простая или сложная.**
- 2. Если сложная и гипотеза проверяется по той же выборке, по которой проверяется согласие, то нельзя использовать классические результаты (предельные распределения и процентные точки).**
- 3. Следует помнить, что предельные распределения статистик непараметрических критериев согласия зависят от ряда факторов (типа оцениваемого параметра, количества оцениваемых параметров, возможно, от конкретных значений параметров формы, от метода оценивания параметров).**
- 4. Если модель предельного распределения статистики критерия для конкретной ситуации неизвестна, всегда можно воспользоваться технологией компьютерного моделирования и построить модель распределения статистики для этой ситуации.**
- 5. Остаются вопросы, связанные с мощностью критериев, предназначенных для проверки одних и тех же гипотез (?).**

Проблемы применения критериев типа хи-квадрат

Критерий χ^2 Пирсона. Применение критериев типа χ^2 предусматривает разбиение области определения случайной величины на k интервалов с подсчетом числа наблюдений n_i , попавших в них, и вероятностей попадания в интервалы $P_i(\theta)$, соответствующих теоретическому закону. Статистика критерия согласия χ^2 Пирсона имеет вид

$$X_n^2 = n \sum_{i=1}^k \frac{(n_i / n - P_i(\theta))^2}{P_i(\theta)}. \quad (4)$$

В случае проверки простой справедливой гипотезы в пределе эта статистика подчиняется χ_{k-1}^2 -распределению с $k-1$ степенями свободы.

Если верна конкурирующая гипотеза H_1 и выборка соответствует закону с распределением $F_1(x, \theta_1)$ с параметром θ_1 , то эта же статистика в пределе подчиняется нецентральному χ_{k-1}^2 -распределению с параметром нецентральности

$$v = n \sum_{i=1}^k \frac{(P_i^1(\theta_1) - P_i(\theta))^2}{P_i(\theta)},$$

где $P_i^1(\theta_1)$ – вероятность попадания в интервал при справедливой гипотезе H_1 .

В случае проверки сложной гипотезы, при справедливости H_0 и при условии, что оценки параметров находятся в результате минимизации статистики (4) по этой же самой выборке, статистика X_n^2 асимптотически распределена как χ_{k-r-1}^2 , где r – число оцененных по выборке параметров.

Ошибки и неверные действия при использовании критериев типа хи-квадрат

Анализ примеров “неудачного” применения критериев типа χ^2 позволяет выделить две группы причин, которые могут приводить к неверным статистическим выводам.

Во-первых, это часто совершаемые принципиальные ошибки, при которых использование в качестве предельного χ_{k-r-1}^2 -распределения оказывается неправомерным.

Во-вторых, действия, использующие возможности критерия не наилучшим образом.

В первом случае возрастает вероятность ошибки первого рода α (отклонить верную проверяемую гипотезу), во втором – вероятность ошибки второго рода β (принять проверяемую гипотезу при справедливости альтернативы).

Первоначально предполагалось, что в случае проверки сложных гипотез и оценивании по выборке параметров наблюдаемого закона использование в качестве предельных χ_{k-r-1}^2 -распределений справедливо лишь при определении оценок минимизацией статистики X_n^2 .

Позднее было доказано, что статистика X_n^2 подчиняется χ_{k-r-1}^2 -распределению и в том случае, если используются оценки максимального правдоподобия (ОМП) по группированным наблюдениям.

Наши исследования методами статистического моделирования распределений данной статистики при проверке сложных гипотез и использовании ОМП по группированным наблюдениям (при конечных объемах выборок) также подтвердили хорошее согласие получаемых эмпирических распределений статистики с χ_{k-r-1}^2 -распределениями. Кроме того, наши исследования показали, что есть все основания использовать χ_{k-r-1}^2 -распределения в качестве предельных распределений статистики X_n^2 и в том случае, если параметры сдвига и масштаба наблюдаемых законов случайных величин будут находиться в виде линейных комбинаций выборочных квантилей (L -оценок и оптимальных L -оценок).

Следует полагать, что применение χ_{k-r-1}^2 -распределений в качестве предельных распределений оказывается оправданным и при использовании ряда других оценок, предусматривающих группирование наблюдений, в частности, при нахождении оценок в результате минимизации модифицированной статистики X_n^2

$$\text{mod } X_n^2 = n \sum_{i=1}^k \frac{(n_i - nP_i(\theta))^2}{n_i},$$

где n_i заменяется на 1, если $n_i = 0$, в результате минимизации расстояния Хеллингера

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

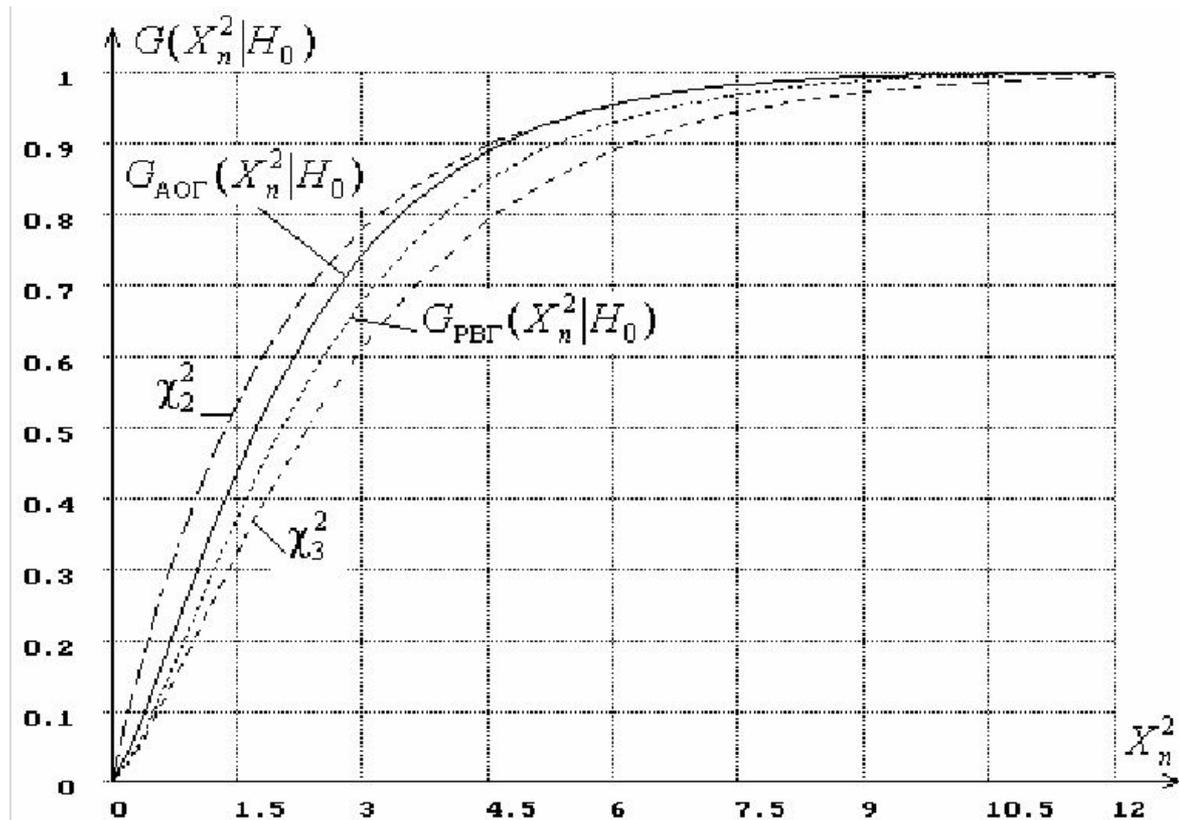
в результате минимизации дивергенции Кульбака-Лейблера (информации Кульбака-Лейблера)

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

Если же оценки параметров искать по точечным выборкам (по исходным негруппированным наблюдениям), то *предельные распределения статистики X_n^2 не являются χ_{k-r-1}^2 -распределениями.*

При вычислении оценок максимального правдоподобия (ОМП) по негруппированным данным эта же статистика распределена как сумма независимых слагаемых $\chi_{k-r-1}^2 + \sum_{j=1}^r \lambda_j \xi_j^2$, где ξ_1, \dots, ξ_r – стандартные нормальные случайные величины, независимые одна от другой и от χ_{k-r-1}^2 , а $\lambda_1, \dots, \lambda_r$ – некоторые числа между 0 и 1.

Более того, распределения статистики X_n^2 становятся зависящими от того, как разбивается область определения случайной величины на интервалы.



Распределения $G(X_n^2 | H_0)$ при асимптотически оптимальном группировании (АОГ) и при разбиении на интервалы равной вероятности (РВГ) в случае проверки согласия с нормальным распределением с оцениванием двух его параметров и числе интервалов $k = 5$.

При оценивании параметров нормального закона по группированной выборке статистика X_n^2 подчинялась бы в данном случае χ_2^2 -распределению. Как подчеркивает рисунок, распределения статистики $G_{\text{АОГ}}(X_n^2 | H_0)$ и $G_{\text{РВГ}}(X_n^2 | H_0)$ очень существенно отличаются от χ_2^2 -распределения. Игнорирование этого факта на практике часто **приводит к неоправданному отклонению** проверки гипотезы, к увеличению числа ошибок первого рода.

Критерий Никулина. В данном случае рассматривалась статистика, предложенная Никулиным [31,40-43]. Критерий предусматривает оценивание неизвестных параметров распределения $F(x, \theta)$ методом максимального правдоподобия по негруппированным данным. При этом вектор вероятностей попадания в интервалы $P = (P_1, \dots, P_k)^T$ предполагается заданным, и граничные точки интервалов определяют по соотношениям $x_i(\theta) = F^{-1}(P_1 + \dots + P_i)$, $i = \overline{1, (k-1)}$. Предложенная статистика имеет вид [25]

$$Y_n^2 = X_n^2 + n^{-1} a^T(\theta) \Lambda(\theta) a(\theta), \quad (5)$$

где X_n^2 вычисляется по формуле (4); матрица $\Lambda(\theta) = \left\| J(\theta_l, \theta_j) - \sum_{i=1}^k \frac{w_{\theta_i} w_{\theta_j}}{P_i} \right\|^{-1}$, элементы и

размерность которой определяются оцениваемыми компонентами вектора параметров

θ ; $J(\theta_l, \theta_j) = \int \left(\frac{\partial f(x, \theta)}{\partial \theta_l} \frac{\partial f(x, \theta)}{\partial \theta_j} \right) f(x, \theta) dx$ – элементы информационной матрицы по

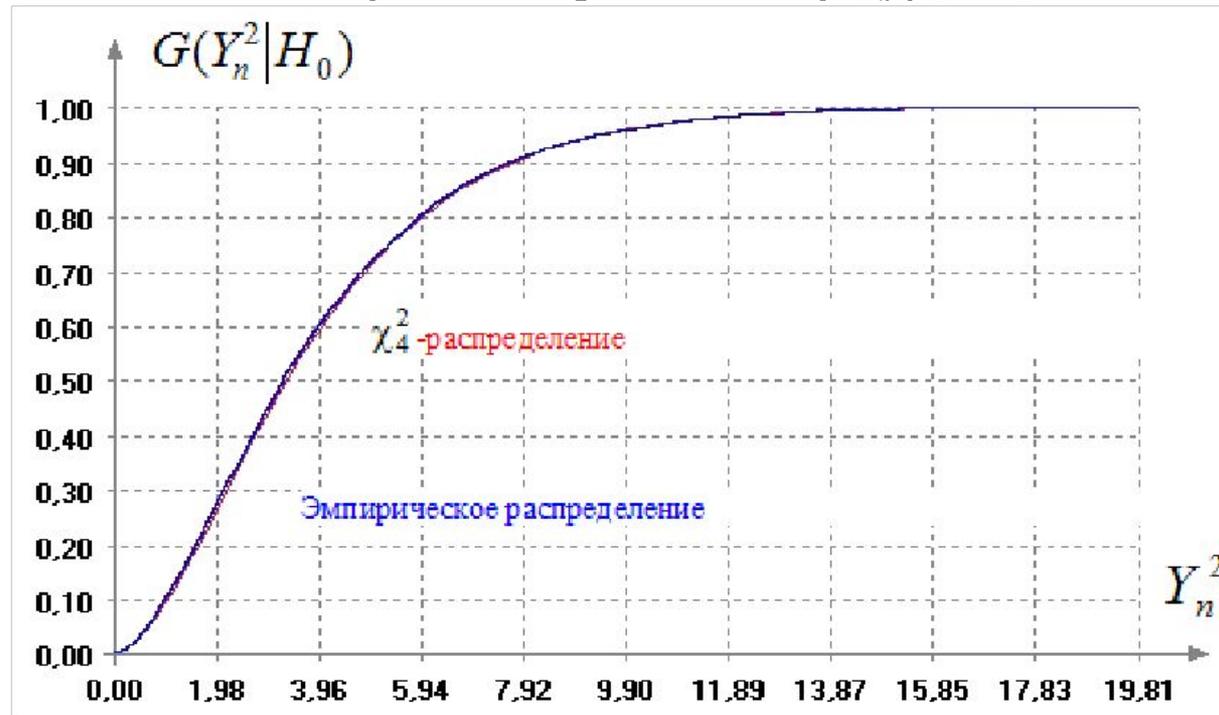
негруппированным данным; компоненты вектора $a(\theta)$ имеют вид

$$a_{\theta_l} = w_{\theta_l 1} n_1 / P_1 + \dots + w_{\theta_l k} n_k / P_k;$$

$$w_{\theta_l i} = -f[x_i(\theta), \theta] \frac{\partial x_i(\theta)}{\partial \theta_l} + f[x_{i-1}(\theta), \theta] \frac{\partial x_{i-1}(\theta)}{\partial \theta_l}.$$

Замечательным фактом, отличающим эти критерии, является то, что статистика критерия при справедливости проверяемой гипотезы в пределе подчиняется χ_{k-1}^2 -распределению независимо от числа параметров закона, оцененных методом максимального правдоподобия. Во-вторых, мощность критерия, как правило, выше мощности критерия χ^2 Пирсона.

Этот рисунок иллюстрирует близость смоделированного распределения статистики Никулина теоретическому χ_{k-1}^2 -распределению



При 5 интервалах и вычислении 2-х параметров логистического распределения

Факторы, влияющие на мощность критериев типа хи-квадрат

1. Способ группирования, способ разбиения области определения случайной величины на интервалы.
2. Выбор числа интервалов.

Зависимость мощности от способа группирования

Практики чаще всего разбивают область, которой принадлежит выборка, на интервалы равной длины, теоретики предпочитают разбиение на интервалы равной вероятности.

С позиции мощности критериев, то есть способности различать, в том числе, близкие альтернативы **эти подходы не являются оптимальными.**

Как упоминалось, если верна конкурирующая гипотеза H_1 и выборка соответствует закону с распределением $F_1(x, \theta_1)$ с параметром θ_1 , то статистика критерия Пирсона при проверке простой гипотезы в пределе подчиняется нецентральному χ_{k-1}^2 -распределению с параметром нецентральности

$$v = n \sum_{i=1}^k \frac{(P_i^1(\theta_1) - P_i(\theta))^2}{P_i(\theta)},$$

где $P_i^1(\theta_1)$ – вероятность попадания в интервал при справедливой гипотезе H_1 .

В случае проверки сложной гипотезы, при справедливости H_0 и при условии, что оценки параметров находятся в результате минимизации статистики (4) по этой же самой выборке, статистика X_n^2 асимптотически распределена как χ_{k-r-1}^2 , где r – число оцененных по выборке параметров.

Можно показать, разлагая $P_i(\theta_1)$ в соотношении (3) в ряд Тейлора при малых $\delta\theta = \theta_1 - \theta$ и пренебрегая членами высшего порядка, что

$$\begin{aligned} v &\approx N \sum_{i=1}^k \frac{[P_i(\theta) + \nabla^T P_i(\theta) \delta\theta - P_i(\theta)]^2}{P_i(\theta)} = N \sum_{i=1}^k \frac{\delta\theta^T \nabla P_i(\theta) \nabla^T P_i(\theta) \delta\theta}{P_i(\theta)} = \\ &= N \delta\theta^T \left(\sum_{i=1}^k \frac{\nabla P_i(\theta) \nabla^T P_i(\theta)}{P_i(\theta)} \right) \delta\theta = N \delta\theta^T \mathbf{J}_\Gamma(\theta) \delta\theta, \end{aligned}$$

где

$$\mathbf{J}_\Gamma(\theta) = \sum_{i=1}^k \frac{\nabla P_i(\theta) \nabla^T P_i(\theta)}{P_i(\theta)}$$

– информационная матрица Фишера по группированным данным. Мощность критерия χ^2 Пирсона представляет собой неубывающую функцию от v . Матрица потерь информации, вызванных группированием, $\Delta\mathbf{J} = \mathbf{J}(\theta) - \mathbf{J}_\Gamma(\theta)$, где $\mathbf{J}(\theta)$ – информационная матрица Фишера по негруппированным наблюдениям, является неотрицательно определенной, и, следовательно, $\delta\theta^T \Delta\mathbf{J} \delta\theta \geq 0$. Так как $\delta\theta^T \mathbf{J}_\Gamma(\theta) \delta\theta = \delta\theta^T \mathbf{J}(\theta) \delta\theta - \delta\theta^T \Delta\mathbf{J} \delta\theta$, то очевидно, что с ростом потерь информации падает и мощность критерия при близких конкурирующих гипотезах.

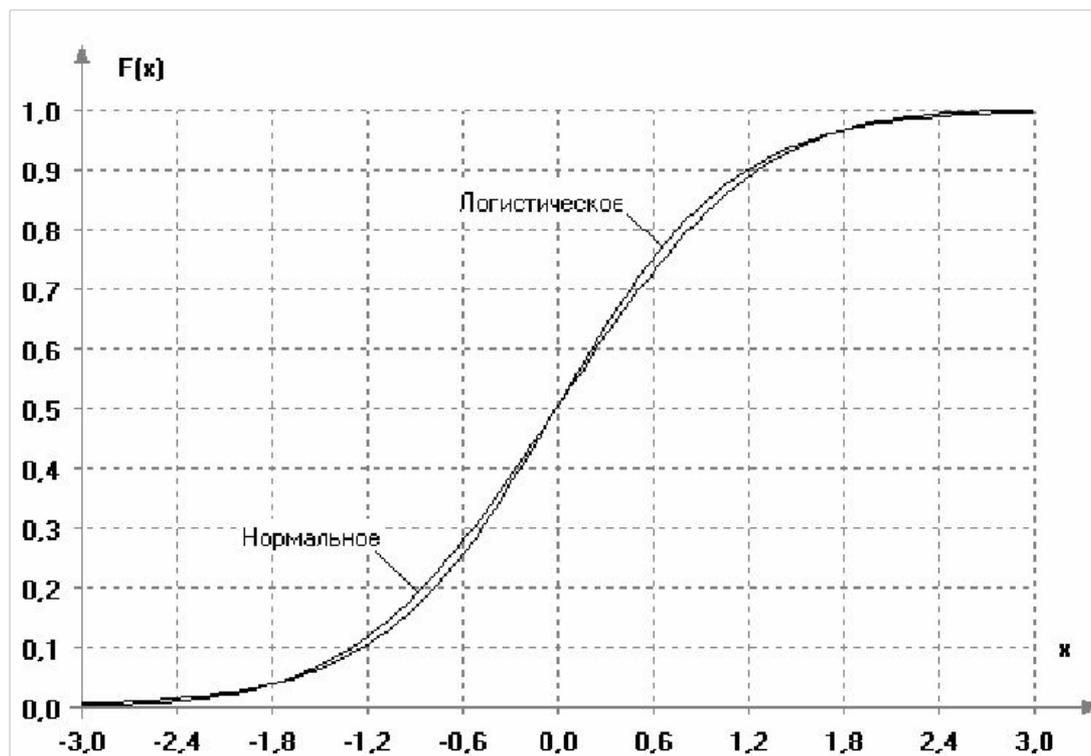
Потери от группирования можно уменьшить, решая задачу асимптотически оптимального группирования и подбирая граничные точки так, чтобы $\mathbf{J}_\Gamma(\theta)$ стремилась к информационной матрице по негруппированным данным $\mathbf{J}(\theta)$. В случае скалярного параметра эта задача сводится к максимизации количества информации Фишера о параметре по группированной выборке

$$\max_{x_0 < x_1 < \dots < x_{k-1} < x_k} \sum_{i=1}^k \left(\frac{\partial \ln P_i(\theta)}{\partial \theta} \right)^2 P_i(\theta).$$

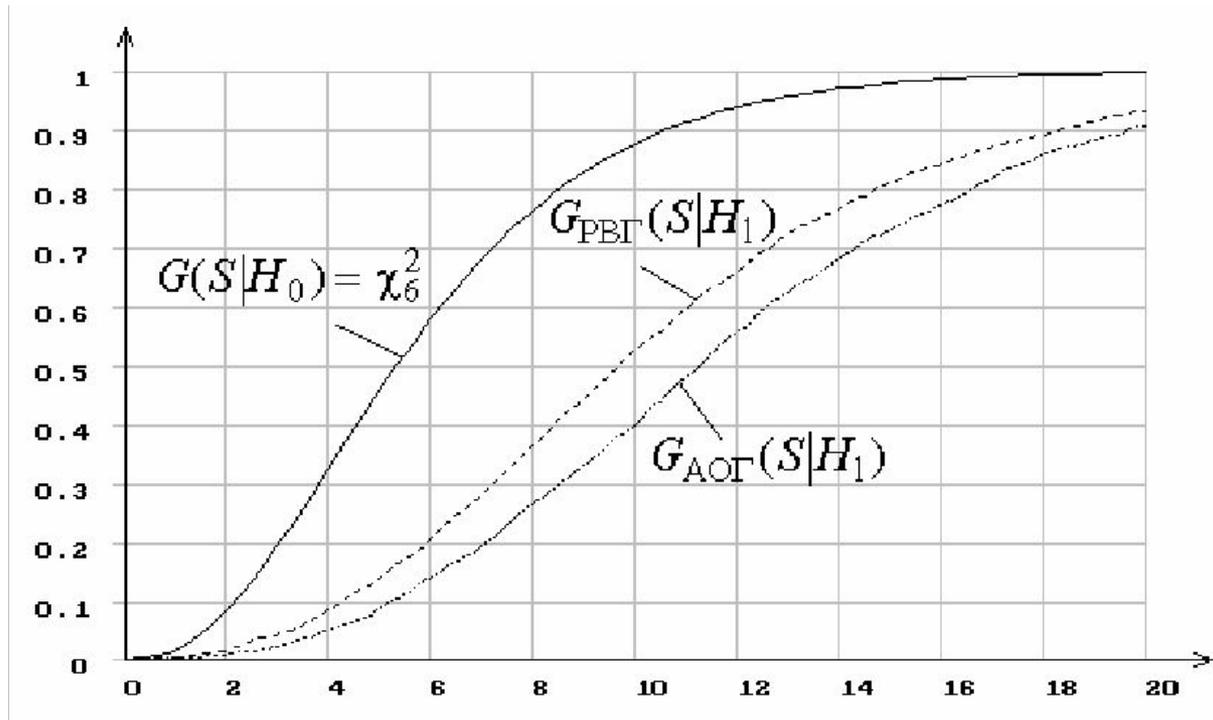
А в случае вектора параметров в качестве критериев оптимальности могут быть выбраны различные функционалы от информационной матрицы Фишера. Наиболее естественно максимизировать определитель информационной матрицы, т.е. решать задачу

$$\max_{x_0 < x_1 < \dots < x_{k-1} < x_k} \det \mathbf{J}_\Gamma(\theta).$$

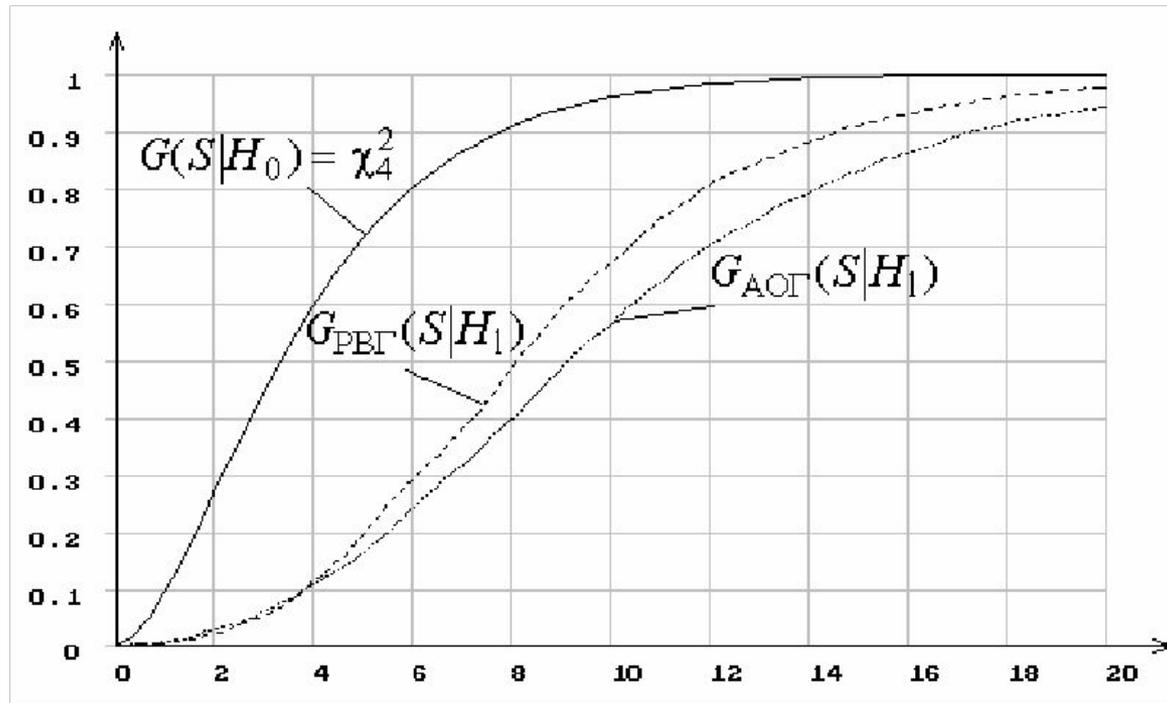
Рассмотрим два близких распределения, нормальное и логистическое, соответствующие конкурирующим гипотезам H_0 и H_1 . Эти два закона достаточно близки и трудноразличимы с помощью критериев согласия.



Нормальное и логистическое распределения, соответствующие H_0 и H_1



Распределения статистики χ^2 Пирсона при проверке простой гипотезы H_0 .
 Число интервалов группирования – 7.



Распределения статистики χ^2 Пирсона при проверке простой гипотезы H_0 .
Число интервалов группирования – 7.

Зависимость мощности от числа интервалов

Исторически было предложено много формул для определения числа интервалов.

Во многих источниках, можно найти упоминание эвристической формулы Старджесса для определения “оптимального” числа интервалов

$$k = \log_2 N + 1 = 3,3 \lg N + 1.$$

Для определения “оптимального” числа интервалов рекомендуют формулу Брукса и Каррузера

$$k = 5 \lg N.$$

Рекомендуют соотношение

$$k = \sqrt{N}.$$

Для равновероятных интервалов их число устанавливают порядка

$$k \approx 4\sqrt[5]{2}(N/t)^{0.4},$$

где t – квантиль стандартного нормального распределения для заданного уровня значимости. В ряде работ приводят модификации данной формулы. Предлагают значение

$$k = 4 \lg N,$$

и её дальнейшее развитие

$$k = 5 \lg N - 5.$$

Предлагают соотношение

$$k = \frac{4}{\aleph} \lg \frac{N}{10},$$

где $\aleph = 1/\sqrt{\mu_4/\sigma^4}$ – значение контрэксцесса, μ_4 – четвертый центральный момент, σ – среднее квадратичное отклонение.

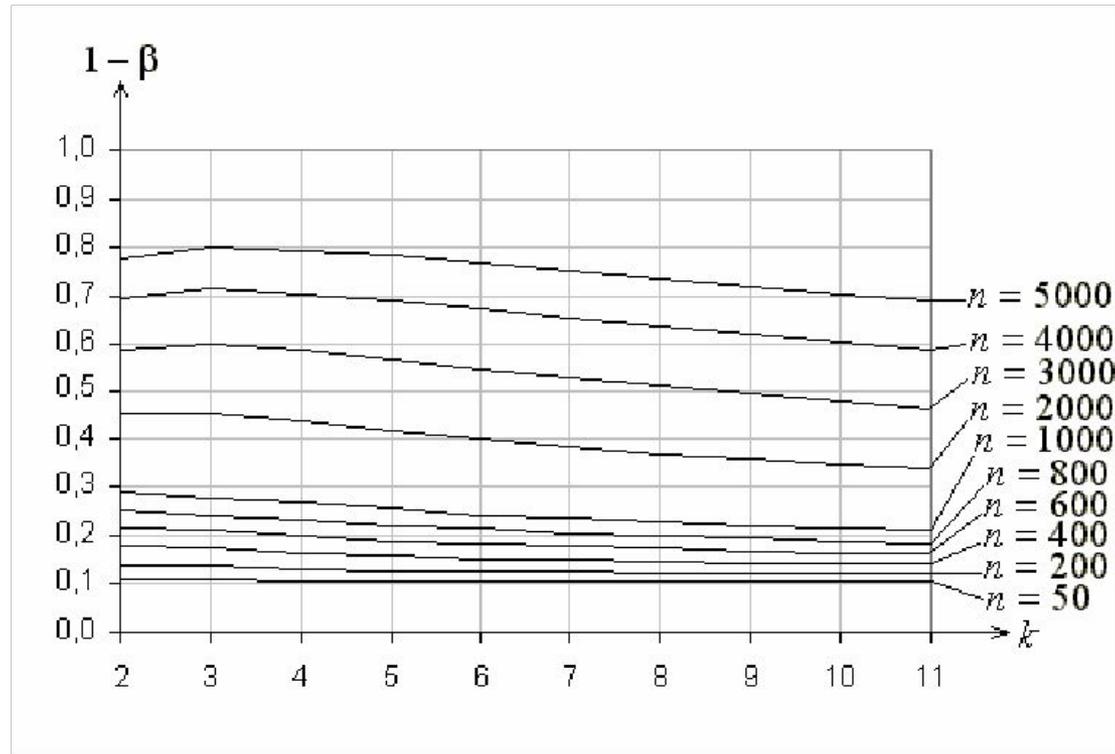
При больших объемах выборок N разброс значений k , задаваемых различными формулами, достаточно велик. Поэтому на практике при выборе числа интервалов больше руководствуются тем, чтобы в интервалы попадало число наблюдений не менее 5-10. Так, например, в рекомендациях ВНИИМетрологии [34] в зависимости от N предлагают следующие значения k :

N	k
40–100	7–9
100–500	8–12
500–1000	10–16
1000–10000	12–22

Все вышеперечисленные рекомендации опирались на предположение, что k следует выбирать таким образом, чтобы вид гистограммы был как можно ближе к плавной кривой плотности распределения генеральной совокупности. Ченцовым Н.Н. показано, что уклонение гистограммы от плотности распределения в лучшем случае имеет порядок $1/\sqrt[3]{N}$, достигаемый при числе интервалов k порядка $\sqrt[3]{N}$.

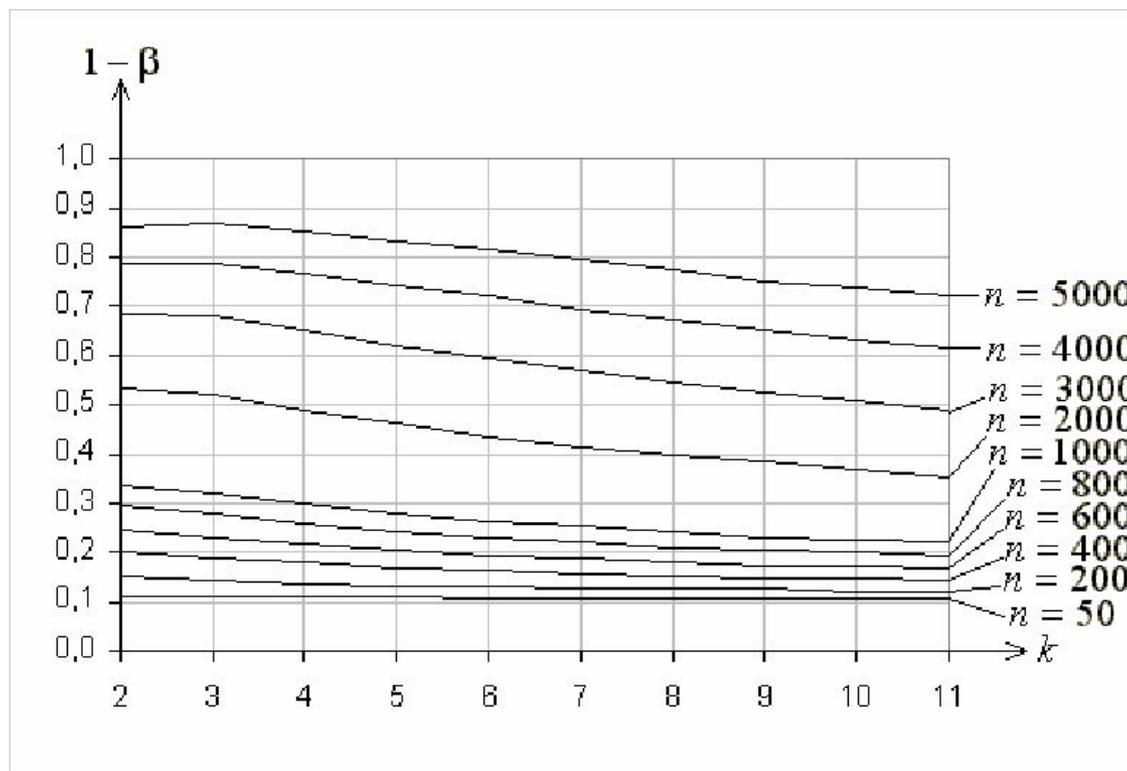
Очевидно, что “оптимальное” значение k зависит не только от объема выборки, но и от вида закона распределения и от способа группирования.

На рисунке в зависимости от числа k равновероятных интервалов при различных n представлены функции мощности критерия χ^2 Пирсона при проверке простой гипотезы о согласии с экспоненциальным законом ($H_0: f(x) = \theta \exp\{-\theta x\}$ при $\theta = 1$; $H_1: f(x) = \theta \exp\{-\theta x\}$ при $\theta = 1.05$).



Функции мощности критерия χ^2 Пирсона при проверке простой гипотезы о согласии с экспоненциальным законом при равновероятном группировании

Аналогичные функции при использовании асимптотически оптимального группирования. И в том и в другом случае с ростом k мощность падает, но в случае асимптотически оптимального группирования она выше, чем при равновероятном.

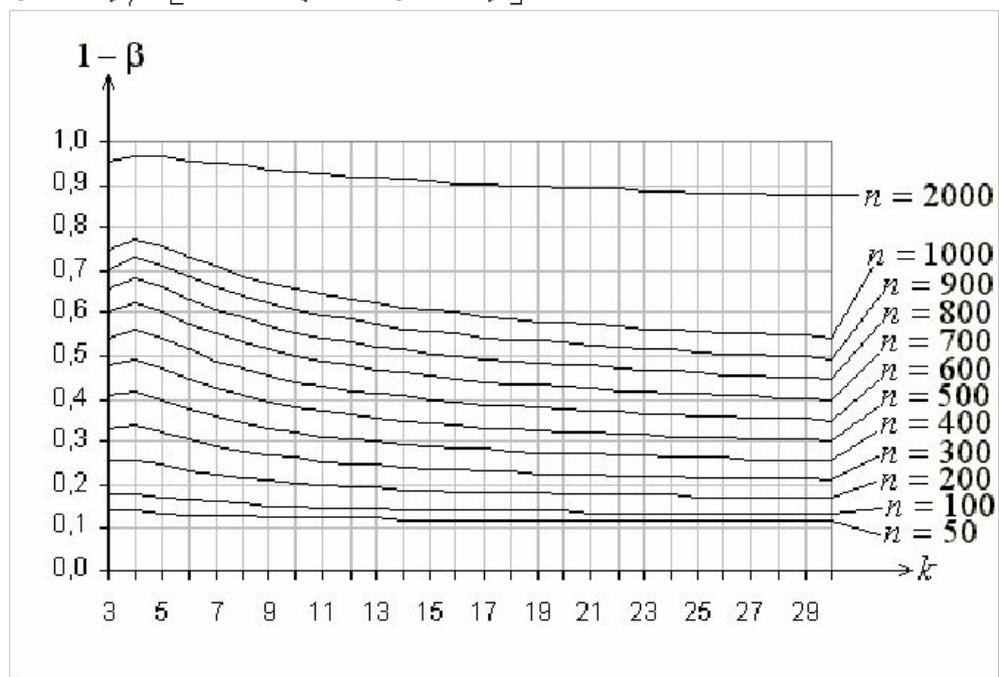


Функции мощности критерия χ^2 Пирсона при проверке простой гипотезы о согласии с экспоненциальным законом при асимптотически оптимальном группировании

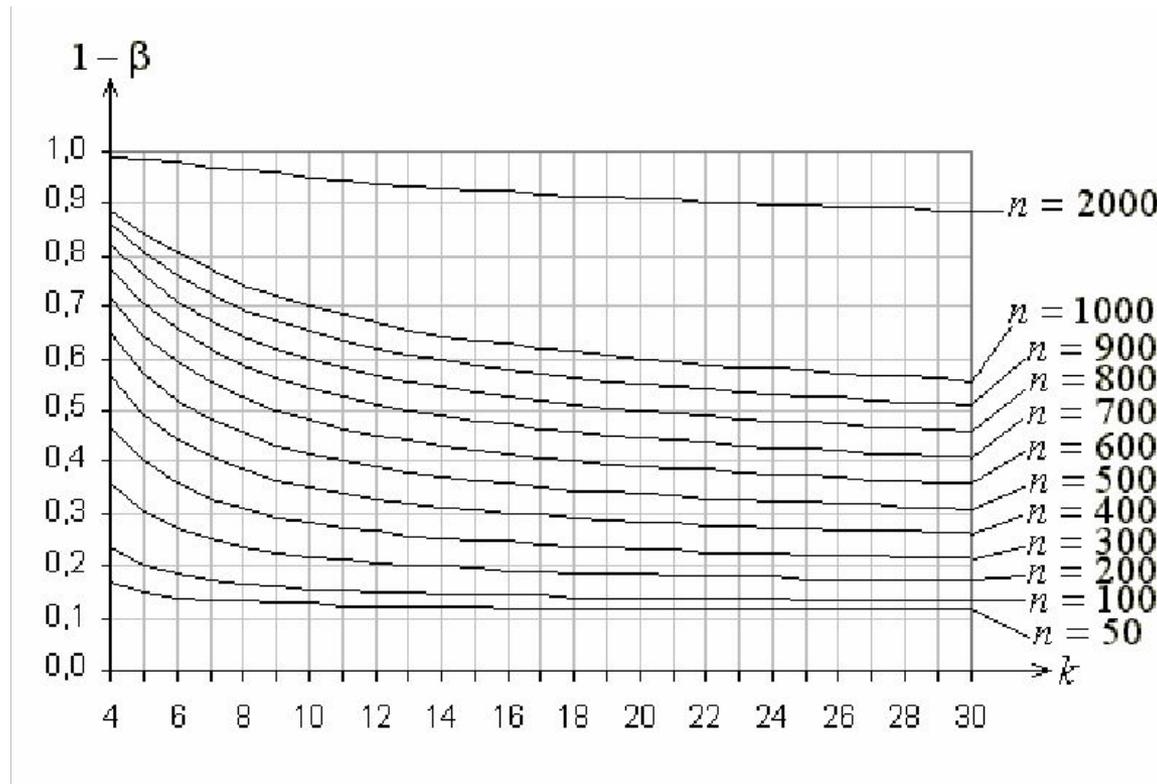
Здесь и на следующем рисунке представлены функции мощности критерия χ^2 Пирсона при проверке простых и сложных гипотез о согласии с нормальным законом H_0 :

$$f(x) = \frac{1}{\theta_1 \sqrt{2\pi}} \exp\left\{-\frac{(x - \theta_0)^2}{2\theta_1^2}\right\} \quad \text{при логистическом законе в качестве альтернативы } H_1:$$

$$f(x) = \frac{\pi}{\theta_1 \sqrt{3}} \exp\left\{-\frac{\pi(x - \theta_0)}{\theta_1 \sqrt{3}}\right\} / \left[1 + \exp\left\{-\frac{\pi(x - \theta_0)}{\theta_1 \sqrt{3}}\right\}\right]^2 \quad \text{при значениях параметров } \theta_0 = 0, \theta_1 = 1.$$

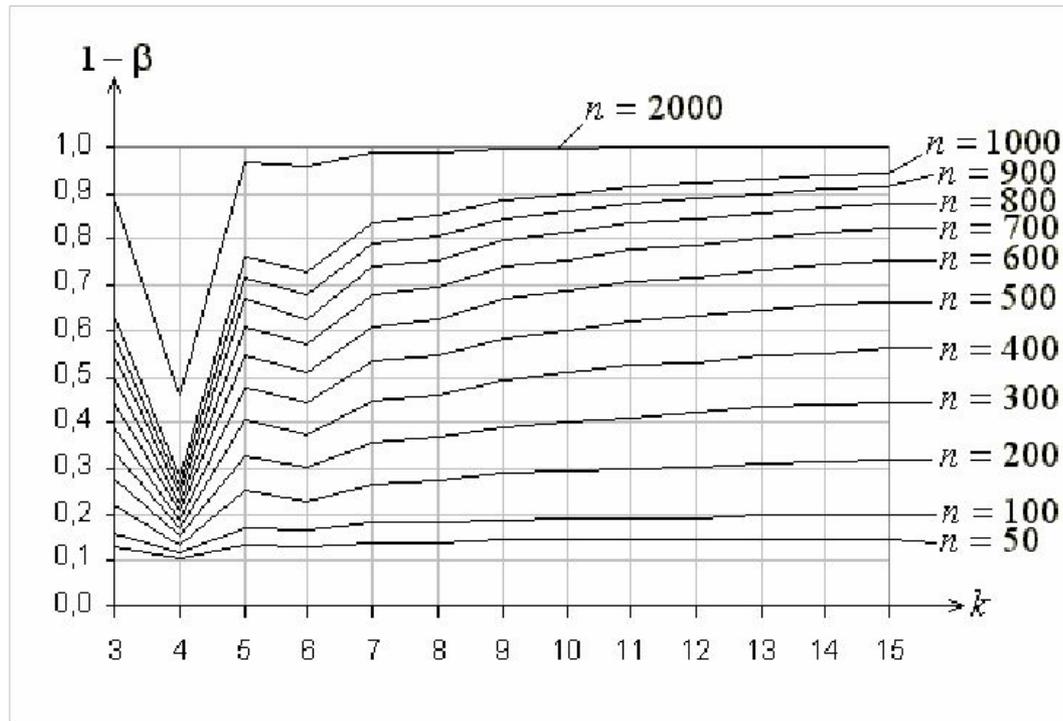


Функции мощности критерия χ^2 Пирсона при проверке простой гипотезы о согласии с нормальным законом при **равновероятном** группировании и альтернативе, соответствующей логистическому закону

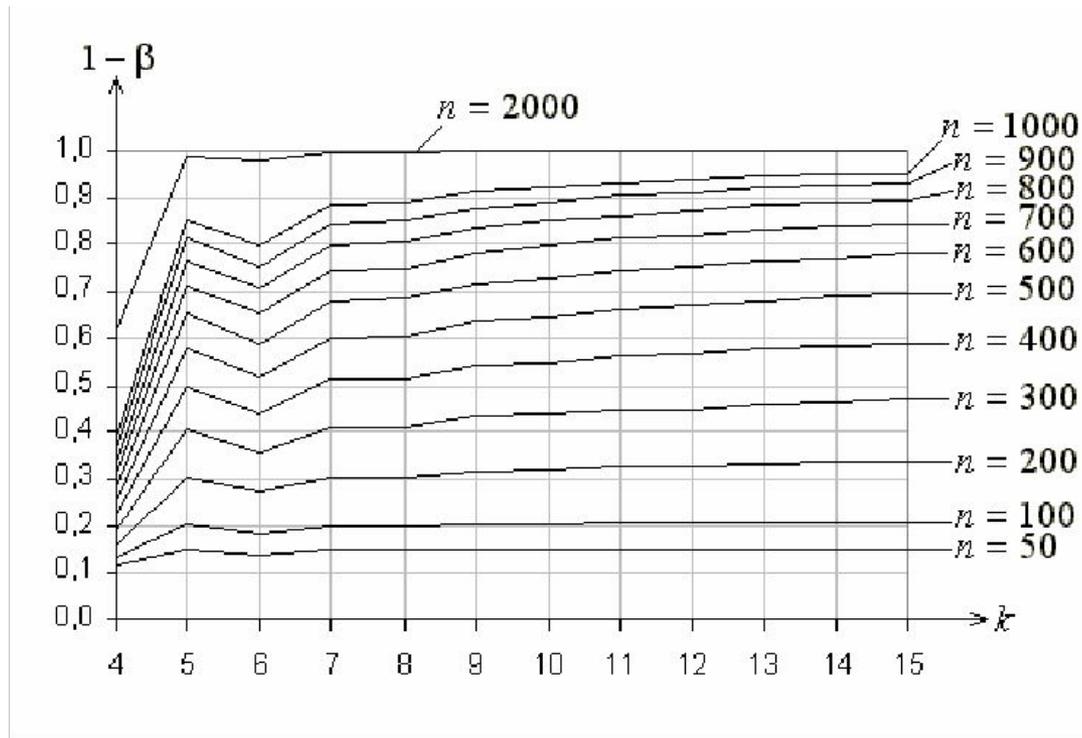


Функции мощности критерия χ^2 Пирсона при проверке сложной гипотезы о согласии с нормальным законом при **равновероятном** группировании и альтернативе, соответствующей логистическому закону

Аналогично при **асимптотически оптимальном** группировании

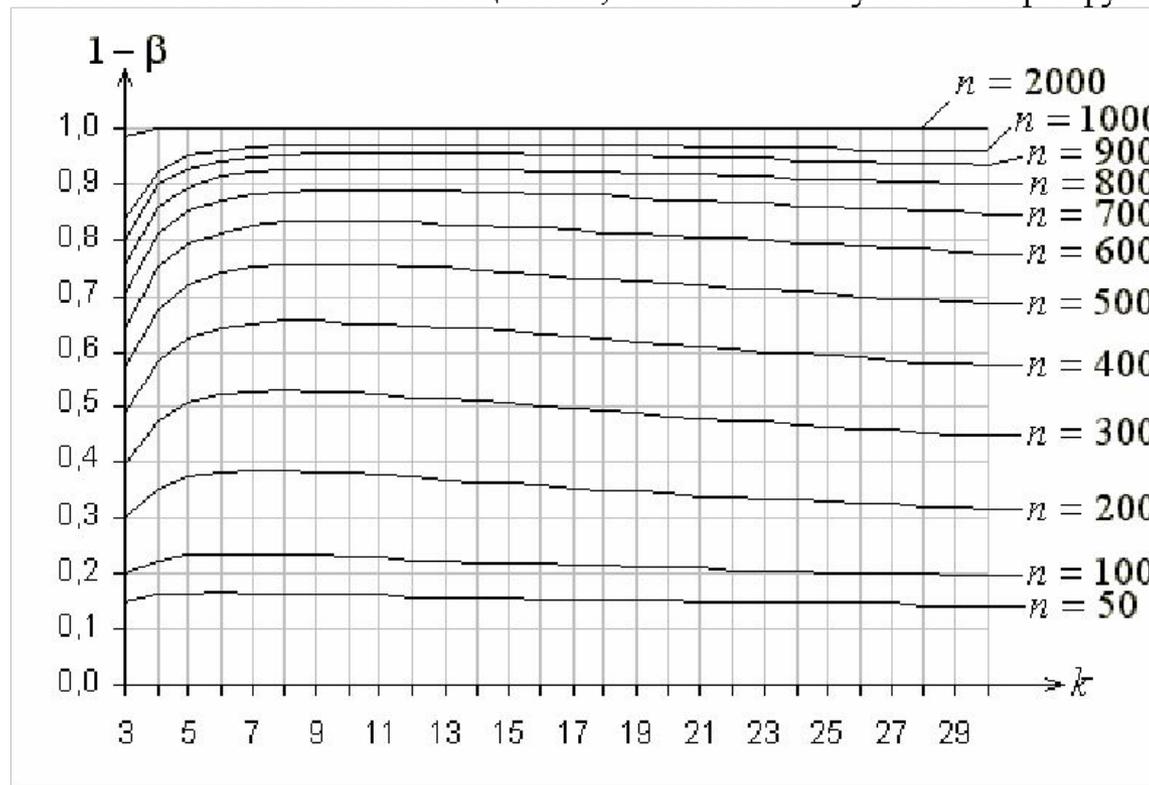


Функции мощности критерия χ^2 Пирсона при проверке **простой** гипотезы о согласии с нормальным законом при асимптотически оптимальном группировании и альтернативе, соответствующей логистическому закону



Функции мощности критерия χ^2 Пирсона при проверке **сложной** гипотезы о согласии с нормальным законом при **асимптотически оптимальном** группировании и альтернативе, соответствующей логистическому закону

Функция мощности критерия типа χ^2 Никулина, как следует из рисунка на области значений k , содержащей максимальное значение мощности, является выпуклой вверх функцией.



Функции мощности критерия типа χ^2 Никулина при проверке сложной гипотезы о согласии с нормальным законом при **равновероятном** группировании и альтернативе, соответствующей логистическому закону

Таким образом, выбирая число интервалов в критериях типа χ^2 , следует осознавать, что увеличение их числа не приводит к росту мощности критерия.

Если заданы конкретная альтернатива и объем выборки N , можно выбрать оптимальное число интервалов k так, чтобы максимизировать мощность.

Анализ мощности критериев согласия при близких альтернативах

Отдавая при проведении статистического анализа данных предпочтение некоторому критерию, экспериментатор хотел бы иметь уверенность в том, что для заданной вероятности ошибки первого рода α гарантируется минимальная вероятность ошибки 2-го рода β . Другими словами, хотелось бы отдать предпочтение критерию, наиболее мощному относительно интересующей нас пары альтернатив H_0 и H_1 .

Информация, содержащаяся в различных источниках, о преимуществах в определенных ситуациях того или иного критерия согласия неоднозначна и зачастую противоречива. Результаты исследования асимптотической мощности критериев, например [1-4], трудно использовать вследствие ограниченных объемов выборок, с которыми приходится иметь дело практику. Рекомендации различных авторов носят субъективный характер, отражают сложившиеся стереотипы, базируются на конкретных частных примерах и ограниченном опыте практического применения.

Исследования мощности затруднены отсутствием результатов, связанных с аналитическим представлением функций распределения $G(S|H_1)$ для конкретных критериев согласия при проверке сложных гипотез, в частности, для непараметрических критериев и для критериев типа χ^2 при оценивании параметров по точечным выборкам (по негруппированным наблюдениям).

Цель исследований, результаты которых приводятся ниже, заключалась в сравнительном анализе мощности наиболее часто используемых критериев согласия на некоторых парах достаточно близких конкурирующих гипотез H_0 и H_1 . **Интерес представляет способность критериев различать именно близкие гипотезы, так как распознавание отличия в далеких законах распределения, как правило, не составляет проблем.**

Рассматриваемые альтернативы Результаты сравнительного анализа мощности критериев согласия в работе иллюстрируются на двух парах альтернатив. Первую пару составили нормальный и логистический законы: проверяемой гипотезе H_0 соответствовал нормальный закон с плотностью

$$f(x) = \frac{1}{\theta_0 \sqrt{2\pi}} \exp\left\{-\frac{(x - \theta_1)^2}{2\theta_0^2}\right\},$$

а конкурирующей гипотезе H_1 – логистический с функцией плотности

$$f(x) = \frac{\pi}{\theta_0 \sqrt{3}} \exp\left\{-\frac{\pi(x - \theta_1)}{\theta_0 \sqrt{3}}\right\} / \left[1 + \exp\left\{-\frac{\pi(x - \theta_1)}{\theta_0 \sqrt{3}}\right\}\right]^2$$

и параметрами $\theta_0 = 1$, $\theta_1 = 0$. В случае простой гипотезы H_0 параметры нормального закона имеют те же значения. Эти два закона близки и трудно различимы с помощью критериев согласия.

Вторую пару составили: H_0 – распределение Вейбулла с плотностью

$$f(x) = \frac{\theta_0 (x - \theta_2)^{\theta_0 - 1}}{\theta_1^{\theta_0}} \exp\left\{-\left(\frac{x - \theta_2}{\theta_1}\right)^{\theta_0}\right\}$$

и параметрами $\theta_0 = 2$, $\theta_1 = 2$, $\theta_2 = 0$; H_1 – гамма-распределение с плотностью

$$f(x) = \frac{1}{\theta_1 \Gamma(\theta_0)} \left(\frac{x - \theta_2}{\theta_1}\right)^{\theta_0 - 1} e^{-(x - \theta_2)/\theta_1}$$

и параметрами $\theta_0 = 3.12154$, $\theta_1 = 0.557706$, $\theta_2 = 0$, при которых гамма-распределение наиболее близко к данному распределению Вейбулла.

В работе исследовалась мощность при проверке простых и сложных гипотез H_0 против простой альтернативы H_1 .

Мощность критериев в случае пары альтернатив “нормальный-логистический”

О близости распределений нормального и логистического, соответствующих конкурирующим гипотезам H_0 и H_1 , свидетельствует рис. 1.

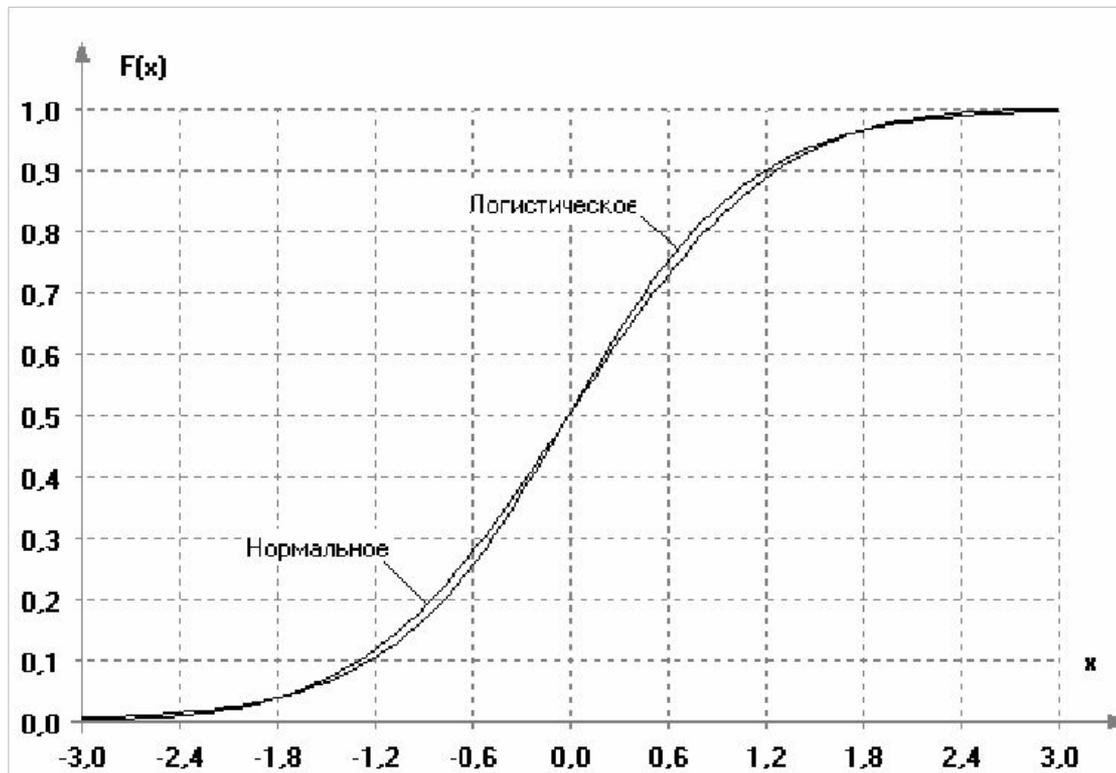


Рис. 1. Нормальное и логистическое распределения, соответствующие H_0 и H_1

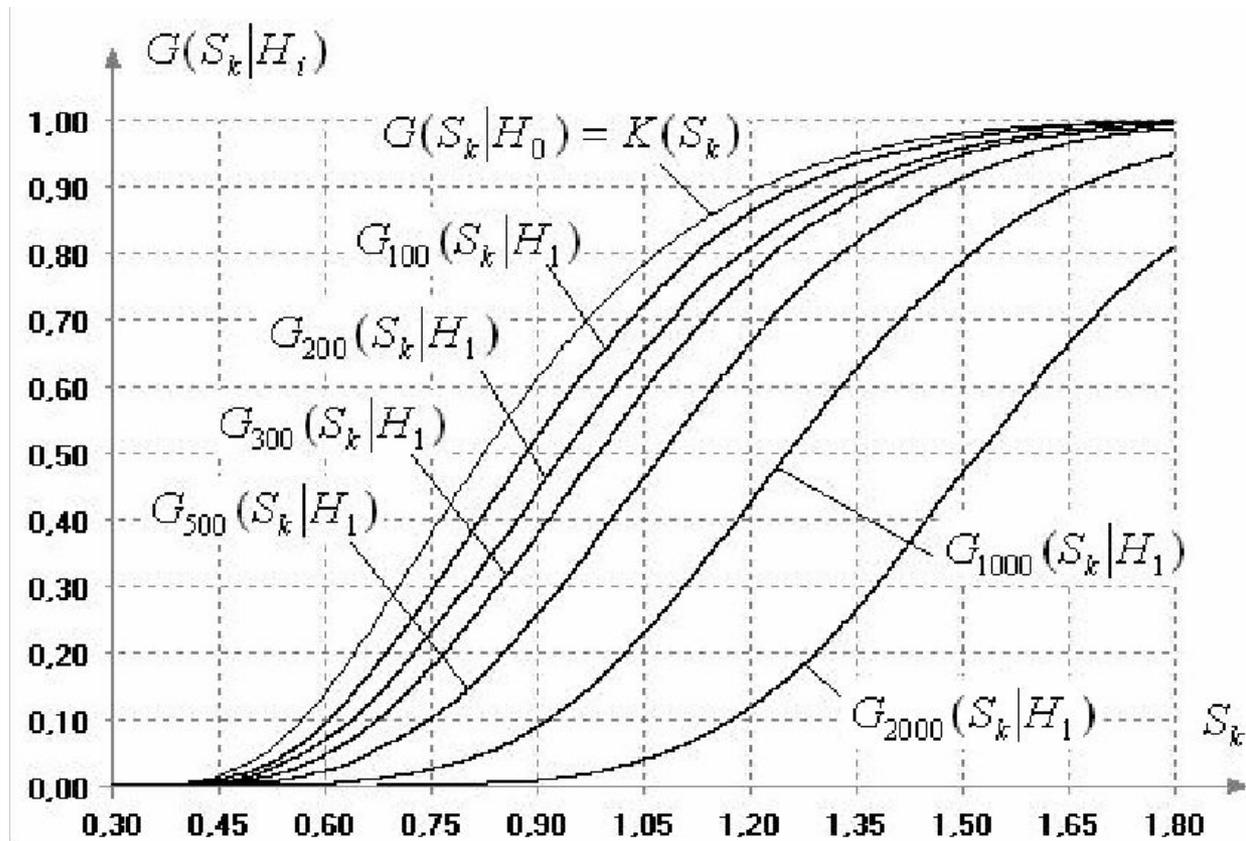


Рис. 2. Распределения статистики (1) типа Колмогорова $G(S_k|H_0)=K(S_k)$ и $G_n(S_k|H_1)$ при проверке простой гипотезы H_0 о согласии с нормальным законом при альтернативе H_1

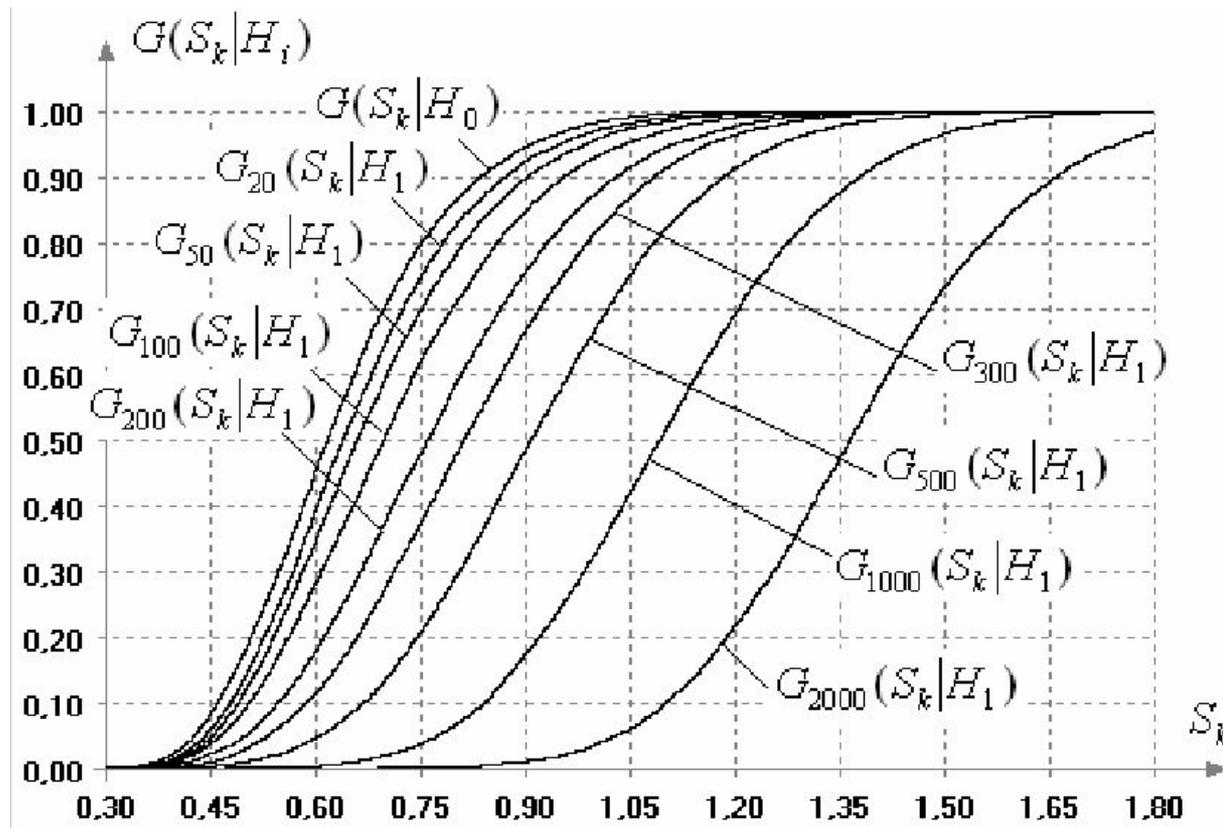


Рис.3. Распределения статистики (1) типа Колмогорова $G(S_k|H_0)$ и $G_n(S_k|H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования ОМП при альтернативе H_1

Мощность непараметрических критериев сильно зависит от используемого метода оценивания

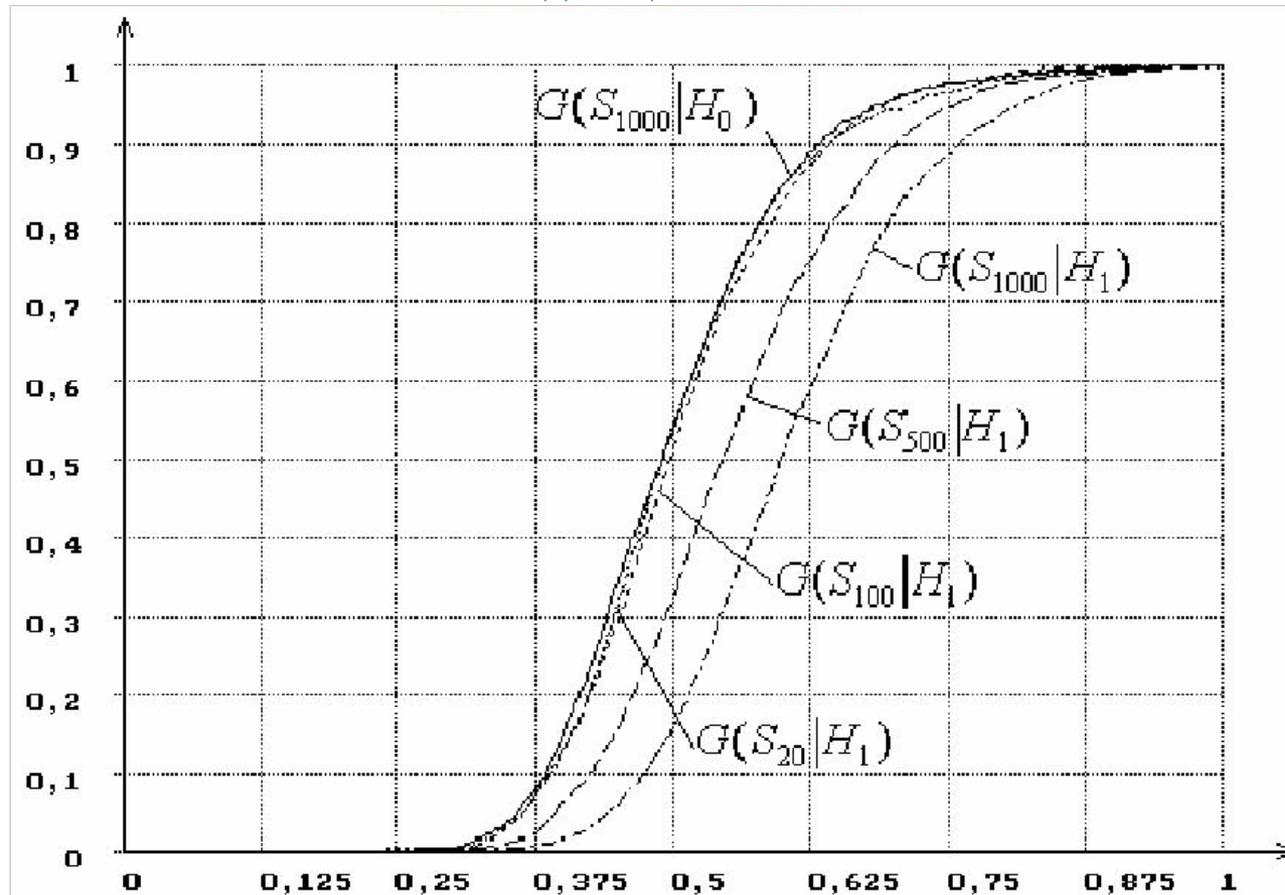


Рис. За. Распределения статистики (1) типа Колмогорова $G(S_k|H_0)$ и $G_n(S_k|H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования MD -оценок, минимизирующих статистику (1), при альтернативе H_1 (логистическое распределение)

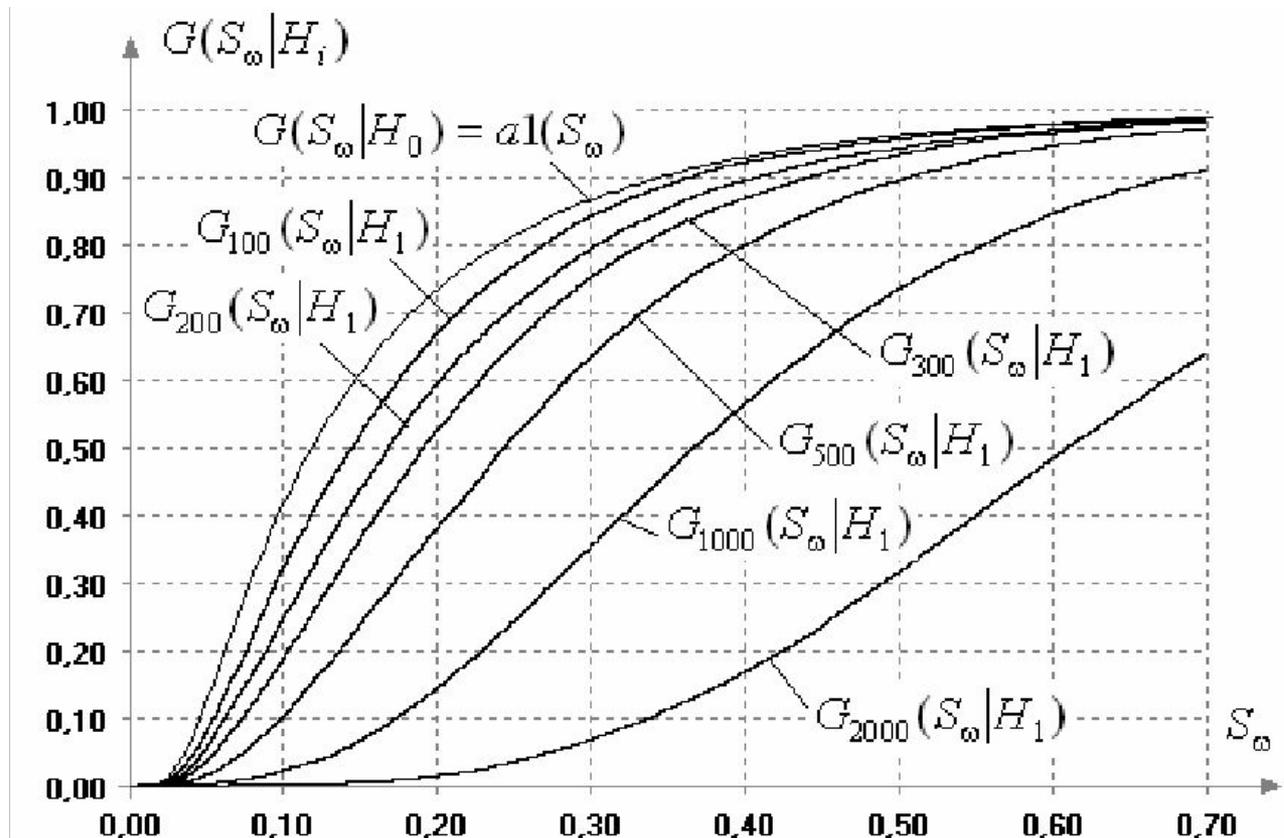


Рис. 4. Распределения статистики (2) типа ω^2 Крамера-Мизеса-Смирнова $G(S_\omega|H_0) = a1(S_\omega)$ и $G_n(S_\omega|H_1)$ при проверке простой гипотезы H_0 о согласии с нормальным законом при альтернативе H_1

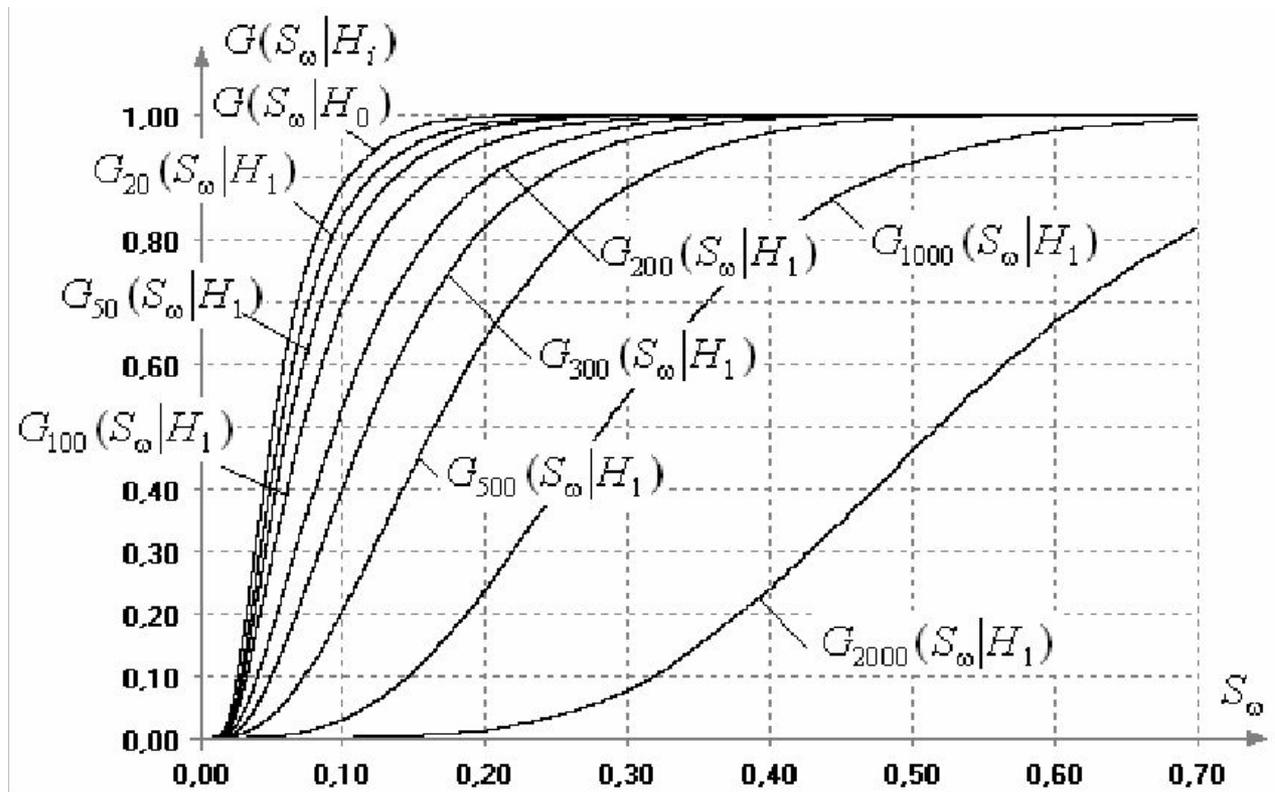


Рис. 5. Распределения статистики (2) типа ω^2 Крамера-Мизеса-Смирнова $G(S_\omega | H_0)$ и $G_n(S_\omega | H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования ОМП при альтернативе H_1

Таблица 2. Мощность критериев согласия при проверке простой гипотезы H_0 (нормальное распределение) против альтернативы H_1 (логистическое)

α	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия χ^2 Пирсона при $k=15$ и АОГ						
0.15	0.349	0.459	0.565	0.737	0.946	0.999
0.1	0.290	0.388	0.490	0.671	0.922	0.998
0.05	0.210	0.292	0.385	0.565	0.871	0.996
0.025	0.154	0.222	0.302	0.472	0.813	0.992
0.01	0.107	0.159	0.221	0.369	0.729	0.983
Мощность критерия Ω^2 Андерсона-Дарлинга						
0.15	0.194	0.258	0.328	0.472	0.776	0.982
0.1	0.125	0.169	0.222	0.343	0.654	0.957
0.05	0.057	0.079	0.107	0.181	0.439	0.869
0.025	0.026	0.036	0.049	0.088	0.261	0.724
0.01	0.010	0.013	0.017	0.031	0.114	0.491
Мощность критерия Колмогорова						
0.15	0.190	0.246	0.303	0.415	0.662	0.922
0.1	0.127	0.170	0.215	0.309	0.544	0.861
0.05	0.062	0.088	0.116	0.179	0.365	0.721
0.025	0.031	0.044	0.061	0.100	0.231	0.560
0.01	0.012	0.018	0.026	0.044	0.119	0.366
Мощность критерия ω^2 Крамера-Мизеса-Смирнова						
0.15	0.178	0.228	0.283	0.401	0.680	0.947
0.1	0.114	0.147	0.186	0.277	0.542	0.892
0.05	0.052	0.067	0.086	0.136	0.324	0.742
0.025	0.024	0.030	0.039	0.062	0.171	0.548
0.01	0.010	0.011	0.014	0.021	0.065	0.307

Таблица 3. Мощность критерия согласия χ^2 Пирсона при проверке простой гипотезы H_0 (нормальное распределение) против альтернативы H_1 (логистическое) в зависимости от способа группирования и числа интервалов

α	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия χ^2 Пирсона при $k=9$ и АОГ						
0.15	0.269	0.381	0.488	0.670	0.917	0.998
0.1	0.204	0.302	0.403	0.589	0.880	0.995
0.05	0.129	0.203	0.287	0.464	0.806	0.989
0.025	0.084	0.136	0.203	0.359	0.723	0.979
0.01	0.050	0.081	0.127	0.249	0.608	0.957
Мощность критерия χ^2 Пирсона при $k=9$ и РВГ						
0.15	0.210	0.282	0.349	0.483	0.747	0.960
0.1	0.152	0.208	0.270	0.392	0.673	0.938
0.05	0.083	0.123	0.170	0.273	0.547	0.890
0.025	0.046	0.072	0.105	0.186	0.435	0.828
0.01	0.020	0.036	0.056	0.109	0.310	0.734
Мощность критерия χ^2 Пирсона при $k=15$ и РВГ						
0.15	0.192	0.257	0.312	0.432	0.690	0.941
0.1	0.139	0.187	0.237	0.343	0.607	0.911
0.05	0.073	0.106	0.144	0.227	0.477	0.848
0.025	0.040	0.061	0.085	0.149	0.365	0.772
0.01	0.018	0.029	0.043	0.083	0.247	0.662

Таблица 4. Мощность критериев согласия при проверке сложной гипотезы H_0 (нормальное распределение) против альтернативы H_1 (логистическое)

α	$n=20$	$n=50$	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия Ω^2 Андерсона-Дарлинга								
0.15	0.222	0.297	0.400	0.575	0.708	0.873	0.989	1.000
0.1	0.164	0.230	0.324	0.496	0.636	0.828	0.981	1.000
0.05	0.098	0.149	0.224	0.377	0.519	0.741	0.963	1.000
0.025	0.060	0.096	0.152	0.282	0.414	0.649	0.935	0.999
0.01	0.031	0.054	0.091	0.186	0.297	0.525	0.885	0.998
Мощность критерия типа χ^2 Никулина при $k=15$ и АОГ								
0.15	0.245	0.320	0.395	0.536	0.646	0.806	0.967	1.000
0.1	0.195	0.249	0.332	0.466	0.579	0.755	0.952	0.999
0.05	0.137	0.165	0.248	0.368	0.480	0.669	0.921	0.998
0.025	0.077	0.112	0.184	0.291	0.395	0.587	0.883	0.996
0.01	0.036	0.071	0.125	0.213	0.304	0.488	0.825	0.992
Мощность критерия ω^2 Крамера-Мизеса-Смирнова								
0.15	0.210	0.273	0.366	0.529	0.659	0.836	0.980	1.000
0.1	0.153	0.208	0.291	0.447	0.582	0.781	0.968	1.000
0.05	0.090	0.130	0.194	0.329	0.458	0.678	0.939	0.999
0.025	0.053	0.082	0.128	0.237	0.353	0.573	0.897	0.998
0.01	0.027	0.044	0.074	0.150	0.243	0.445	0.825	0.994
Мощность критерия χ^2 Пирсона при $k=15$ и АОГ								
0.15	0.243	0.295	0.342	0.467	0.579	0.751	0.950	0.999
0.1	0.194	0.220	0.280	0.393	0.502	0.688	0.928	0.998
0.05	0.140	0.133	0.199	0.291	0.391	0.583	0.882	0.996
0.025	0.081	0.080	0.137	0.214	0.303	0.486	0.827	0.992
0.01	0.036	0.043	0.079	0.139	0.213	0.376	0.745	0.984
Мощность критерия Колмогорова								
0.15	0.200	0.246	0.313	0.440	0.554	0.732	0.941	0.999
0.1	0.142	0.181	0.236	0.351	0.459	0.646	0.905	0.997
0.05	0.080	0.105	0.143	0.230	0.322	0.502	0.823	0.990
0.025	0.045	0.061	0.086	0.149	0.219	0.376	0.721	0.975
0.01	0.021	0.029	0.043	0.081	0.127	0.244	0.575	0.938

Таблица 5. Мощность критериев согласия Никулина и χ^2 Пирсона при проверке сложной гипотезы H_0 (нормальное распределение) против альтернативы H_1 (логистическое) в зависимости от способа группирования и числа интервалов

α	$n=20$	$n=50$	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия типа χ^2 Никулина при $k=9$ и АОГ								
0.15	0.203	0.269	0.339	0.480	0.599	0.774	0.958	0.999
0.1	0.144	0.204	0.271	0.406	0.525	0.714	0.940	0.999
0.05	0.080	0.129	0.186	0.304	0.417	0.617	0.901	0.997
0.025	0.046	0.084	0.129	0.227	0.329	0.525	0.854	0.994
0.01	0.023	0.049	0.081	0.155	0.239	0.419	0.784	0.988
Мощность критерия типа χ^2 Никулина при $k=9$ и РВГ								
0.15	0.200	0.258	0.333	0.477	0.600	0.782	0.964	1.000
0.1	0.143	0.195	0.262	0.399	0.523	0.722	0.947	0.999
0.05	0.081	0.121	0.175	0.292	0.410	0.621	0.910	0.998
0.025	0.047	0.076	0.117	0.213	0.318	0.525	0.865	0.996
0.01	0.023	0.042	0.070	0.140	0.225	0.415	0.796	0.991
Мощность критерия типа χ^2 Никулина при $k=15$ и РВГ								
0.15	0.197	0.263	0.328	0.465	0.584	0.765	0.959	0.999
0.1	0.141	0.198	0.259	0.389	0.508	0.705	0.940	0.999
0.05	0.081	0.126	0.174	0.286	0.399	0.604	0.901	0.998
0.025	0.048	0.080	0.119	0.211	0.311	0.512	0.855	0.995
0.01	0.024	0.045	0.072	0.140	0.222	0.404	0.785	0.990
Мощность критерия типа χ^2 Никулина при $k=7$ и РВГ								
0.15	0.199	0.256	0.332	0.476	0.598	0.780	0.962	0.999
0.1	0.145	0.193	0.260	0.397	0.522	0.719	0.944	0.999
0.05	0.082	0.120	0.172	0.290	0.407	0.616	0.905	0.997
0.025	0.047	0.075	0.114	0.210	0.314	0.520	0.858	0.995
0.01	0.023	0.041	0.066	0.136	0.219	0.406	0.785	0.989
Мощность критерия типа χ^2 Никулина при $k=7$ и АОГ								
0.15	0.189	0.250	0.320	0.457	0.573	0.751	0.948	0.999
0.1	0.136	0.187	0.249	0.380	0.497	0.688	0.926	0.998
0.05	0.077	0.115	0.164	0.276	0.386	0.584	0.880	0.995
0.025	0.042	0.072	0.110	0.201	0.298	0.490	0.826	0.991
0.01	0.018	0.040	0.066	0.133	0.210	0.382	0.746	0.982

Мощность критериев согласия при малых объемах выборок n можно сравнить с мощностью критериев, построенных специально для проверки отклонения распределения от нормального закона.

Таблица 8. Мощности критериев проверки отклонения распределения от нормального закона (Шапиро-Уилка, Эпса-Палли и Д'Агостино со статистикой z_2) относительно альтернативы H_1 (логистический закон)

α	Шапиро-Уилка		Эпса-Палли		Д'Агостино z_2	
	$n=20$	$n=50$	$n=20$	$n=50$	$n=20$	$n=50$
0.1	0.181	0.202	0.178	0.249	0.189	0.327
0.05	0.117	0.141	0.111	0.165	0.111	0.223
0.01	0.044	0.067	0.037	0.062	0.032	0.089

Как видим, “специальные” критерии относительно рассматриваемой пары конкурирующих гипотез в среднем оказываются несколько мощнее.

Мощности критериев в случае пары альтернатив “Вейбулла – гамма-распределение”

Таблица 9. Мощность критериев согласия при проверке простой гипотезы H_0 (распределение Вейбулла с параметрами 2, 2, 0) относительно альтернативы H_1 (гамма-распределение с параметрами 3.12154, 0.557706, 0)

α	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия χ^2 Пирсона при $k=15$ и АОГ						
0.15	0.486	0.621	0.757	0.909	0.996	1.000
0.1	0.418	0.556	0.701	0.876	0.993	1.000
0.05	0.324	0.469	0.611	0.815	0.986	1.000
0.025	0.254	0.403	0.529	0.751	0.974	1.000
0.01	0.191	0.332	0.437	0.668	0.954	1.000
Мощность критерия Ω^2 Андерсона-Дарлингга						
0.15	0.302	0.446	0.577	0.781	0.976	1.000
0.1	0.223	0.348	0.473	0.689	0.951	1.000
0.05	0.131	0.224	0.326	0.533	0.882	0.998
0.025	0.076	0.141	0.220	0.396	0.785	0.993
0.01	0.037	0.075	0.126	0.257	0.636	0.975
Мощность критерия ω^2 Крамера-Мизеса-Смирнова						
0.15	0.295	0.425	0.539	0.716	0.931	0.998
0.1	0.224	0.343	0.453	0.637	0.894	0.995
0.05	0.138	0.233	0.329	0.508	0.816	0.987
0.025	0.084	0.155	0.233	0.393	0.725	0.970
0.01	0.043	0.088	0.142	0.270	0.597	0.934
Мощность критерия Колмогорова						
0.15	0.294	0.421	0.531	0.700	0.915	0.995
0.1	0.225	0.342	0.450	0.628	0.879	0.992
0.05	0.141	0.237	0.332	0.508	0.806	0.981
0.025	0.087	0.160	0.239	0.401	0.723	0.964
0.01	0.045	0.093	0.150	0.282	0.606	0.930

Таблица 10. Мощность критерия согласия χ^2 Пирсона при проверке простой гипотезы H_0 (распределение Вейбулла с параметрами 2, 2, 0) относительно альтернативы H_1 (гамма-распределение с параметрами 3.12154, 0.557706, 0) в зависимости от способа группирования и числа интервалов

α	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия χ^2 Пирсона при $k=9$ и АОГ						
0.15	0.427	0.608	0.748	0.910	0.996	1.000
0.1	0.353	0.534	0.684	0.874	0.993	1.000
0.05	0.261	0.429	0.581	0.807	0.985	1.000
0.025	0.202	0.343	0.488	0.734	0.973	1.000
0.01	0.152	0.255	0.384	0.637	0.950	1.000
Мощность критерия χ^2 Пирсона при $k=15$ и РВГ						
0.15	0.234	0.347	0.446	0.637	0.908	0.998
0.1	0.174	0.266	0.361	0.549	0.867	0.996
0.05	0.097	0.164	0.245	0.417	0.785	0.990
0.025	0.056	0.102	0.161	0.311	0.695	0.979
0.01	0.026	0.054	0.092	0.203	0.574	0.958
Мощность критерия χ^2 Пирсона при $k=9$ и РВГ						
0.15	0.240	0.344	0.440	0.616	0.883	0.995
0.1	0.177	0.262	0.354	0.528	0.835	0.990
0.05	0.100	0.164	0.238	0.399	0.743	0.979
0.025	0.057	0.100	0.157	0.294	0.646	0.960
0.01	0.026	0.053	0.090	0.191	0.520	0.924

Таблица 11. Мощность критериев согласия при проверке сложной гипотезы H_0 (распределение Вейбулла 2, 2, 0) относительно альтернативы H_1 (гамма-распределение с параметрами 3.12154, 0.557706, 0)

α	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия Ω^2 Андерсона-Дарлинга						
0.15	0.435	0.667	0.817	0.952	0.999	1.000
0.1	0.353	0.589	0.757	0.928	0.998	1.000
0.05	0.244	0.466	0.650	0.876	0.995	1.000
0.025	0.167	0.361	0.547	0.811	0.990	1.000
0.01	0.100	0.252	0.424	0.715	0.977	1.000
Мощность критерия ω^2 Крамера-Мизеса-Смирнова						
0.15	0.396	0.603	0.750	0.913	0.996	1.000
0.1	0.316	0.520	0.679	0.875	0.993	1.000
0.05	0.212	0.394	0.560	0.797	0.984	1.000
0.025	0.143	0.295	0.452	0.712	0.968	1.000
0.01	0.082	0.196	0.330	0.593	0.936	1.000
Мощность критерия χ^2 Никулина при $k=9$ и АОГ						
0.15	0.324	0.511	0.665	0.869	0.993	1.000
0.1	0.246	0.423	0.584	0.818	0.987	1.000
0.05	0.153	0.299	0.454	0.720	0.973	1.000
0.025	0.096	0.209	0.347	0.619	0.951	1.000
0.01	0.051	0.129	0.238	0.492	0.909	0.999
Мощность критерия χ^2 Пирсона при $k=9$ и АОГ						
0.15	0.347	0.525	0.678	0.868	0.992	1.000
0.1	0.273	0.439	0.596	0.818	0.986	1.000
0.05	0.172	0.311	0.463	0.719	0.970	1.000
0.025	0.104	0.218	0.352	0.617	0.946	1.000
0.01	0.053	0.133	0.237	0.483	0.898	0.999
Мощность критерия Колмогорова						
0.15	0.340	0.510	0.646	0.830	0.981	1.000
0.1	0.262	0.420	0.558	0.762	0.965	1.000
0.05	0.164	0.293	0.420	0.640	0.925	0.999
0.025	0.101	0.200	0.306	0.519	0.867	0.997
0.01	0.052	0.115	0.193	0.375	0.763	0.988

Таблица 12. Мощность критериев согласия Никулина и χ^2 Пирсона при проверке сложной гипотезы H_0 (распределение Вейбулла 2, 2, 0) относительно альтернативы H_1 (гамма-распределение с параметрами 3.12154, 0.557706, 0) в зависимости от способа группирования и числа интервалов

α	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
Мощность критерия χ^2 Никулина при $k=15$ и АОГ						
0.15	0.365	0.487	0.634	0.828	0.986	1.000
0.1	0.291	0.406	0.558	0.770	0.976	1.000
0.05	0.195	0.302	0.443	0.666	0.952	1.000
0.025	0.131	0.230	0.348	0.569	0.919	1.000
0.01	0.078	0.165	0.250	0.455	0.862	0.999
Мощность критерия χ^2 Пирсона при $k=15$ и АОГ						
0.15	0.385	0.491	0.637	0.839	0.988	1.000
0.1	0.305	0.405	0.559	0.780	0.980	1.000
0.05	0.196	0.293	0.442	0.676	0.959	1.000
0.025	0.124	0.218	0.344	0.574	0.928	1.000
0.01	0.065	0.151	0.234	0.448	0.868	0.999
Мощность критерия χ^2 Никулина при $k=9$ и РВГ						
0.15	0.295	0.455	0.599	0.806	0.981	1.000
0.1	0.220	0.367	0.509	0.740	0.968	1.000
0.05	0.133	0.250	0.378	0.624	0.938	1.000
0.025	0.080	0.167	0.276	0.512	0.894	0.999
0.01	0.040	0.097	0.176	0.380	0.822	0.997
Мощность критерия χ^2 Никулина при $k=15$ и РВГ						
0.15	0.273	0.421	0.558	0.774	0.975	1.000
0.1	0.202	0.335	0.468	0.702	0.960	1.000
0.05	0.120	0.224	0.341	0.582	0.923	1.000
0.025	0.071	0.147	0.244	0.469	0.874	0.999
0.01	0.036	0.085	0.153	0.343	0.796	0.997
Мощность критерия χ^2 Пирсона при $k=9$ и РВГ						
0.15	0.259	0.361	0.465	0.639	0.896	0.996
0.1	0.187	0.282	0.376	0.552	0.851	0.992
0.05	0.114	0.181	0.257	0.421	0.764	0.982
0.025	0.062	0.119	0.170	0.310	0.664	0.964
0.01	0.028	0.057	0.096	0.198	0.529	0.926

Выводы по мощности критериев согласия

Для случая проверки *простых гипотез* можно упорядочить критерии по мощности следующим образом:

χ^2 Пирсона (АОГ) \boxtimes Ω^2 Андерсона-Дарлинга \boxtimes ω^2 Мизеса \boxtimes Колмогорова

Такая шкала справедлива при использовании в критерии χ^2 Пирсона АОГ, при котором минимизируются потери в информации Фишера. При очень близких гипотезах может быть:

Колмогорова \boxtimes ω^2 Мизеса.

При проверке *сложных гипотез* градация по мощности оказывается существенно иной:

Ω^2 Андерсона-Дарлинга \boxtimes ω^2 Мизеса \boxtimes Y_n^2 (АОГ) \boxtimes χ^2 Пирсона (АОГ) \boxtimes Колмогорова.

При очень близких гипотезах может быть:

Ω^2 Андерсона-Дарлинга \boxtimes Y_n^2 (АОГ) \boxtimes ω^2 Мизеса \boxtimes χ^2 Пирсона (АОГ) \boxtimes Колмогорова.

Указанные выводы носят интегрированный характер. Такое упорядочение не является жёстким. Как видно из таблиц с приведенными значениями мощности, иногда критерий имеет преимущества по мощности при одних значениях α и объемах выборок n и уступает при других значениях α и n .

Надо иметь в виду, что мощность критериев типа χ^2 (Пирсона и Никулина) зависит не только от гипотез H_0 , H_1 и объема выборок n , но при заданных H_0 и H_1 — от способа группирования и числа интервалов.

Литература

1. Чиби́сов Д.М. Об асимптотической мощности и эффективности критерия ω_n^2 // ДАН СССР, 1961. 138, 2. – С.322-325.
2. Чиби́сов Д.М. Исследования мощности некоторых непараметрических критериев // Теория вероятностей и ее применения. 1962. – Т. 7. – № 3. – С. 355-356.
3. Чиби́сов Д.М. Об асимптотической мощности критериев согласия при близких альтернативах // Теория вероятностей и ее применения. 1964. – Т. 9. – № 3. – С. 561-566.
4. Чиби́сов Д.М. К исследованию асимптотической мощности критериев согласия // Теория вероятностей и ее применения. 1965. – Т. 10. – № 3. – С. 460-478.
5. Kolmogoroff A.N. Sulla determinazione empirica di una legge di distribuzione // G. Ist. Ital. attuar. 1933. – Vol. 4. – № 1. – P. 83-91.
6. Больше́в Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.
7. Больше́в Л.Н. Асимптотические пирсоновские преобразования // Теория вероятностей и ее применения. 1963. – Т. 8. – № 2. – С. 129-155.
8. Больше́в Л.Н. Теория вероятностей и математическая статистика / Избранные труды. Под ред. Ю.В. Прохорова. 6– М.: Наука, 1987. – 286 с.
9. Кас М., Kiefer J., Wolfowitz J. On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. – V.26. – P.189-211.
10. Р 50.1.037-2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии. - М.: Изд-во стандартов. 2002. - 64 с.
11. Мартынов Г.В. Критерии омега-квадрат. – М.: Наука, 1978. – 80 с.
12. Pearson E.S., Hartley H.O. Biometrika tables for Statistics. V.2. – Cambridge: University Press, 1972. – 634 p.
13. Stephens M.A. Use of Kolmogorov–Smirnov, Cramer – von Mises and related statistics – without extensive table // J. R. Stat. Soc. – 1970. – V. 32. – P. 115-122.

14. Stephens M.A. EDF statistics for goodness of fit and some comparisons // J. Am. Statist. Assoc. 1974. – V.69. – P. 730-737.
15. Chandra M., Singpurwalla N.D., Stephens M.A. Statistics for Test of Fit for the Extrem-Value and Weibull Distribution // J. Am. Statist. Assoc. 1981. – V.76. – P. 375.
16. Тюрин Ю.Н. О предельном распределении статистик Колмогорова–Смирнова для сложной гипотезы // Изв. АН СССР. Сер. Матем. 1984. – Т.48. – № 6. – С. 1314-1343.
17. Тюрин Ю.Н., Саввушкина Н.Е. Критерии согласия для распределения Вейбулла–Гнеденко. // Изв. АН СССР. Сер. Техн. Кибернетика. 1984. – № 3. – С. 109-112.
18. Тюрин Ю.Н. Исследования по непараметрической статистике (непараметрические методы и линейная модель): Автореф. дисс. д-ра физ.-мат. наук. – М., 1985. – 33 с. – (МГУ).
19. Саввушкина Н.Е. Критерий Колмогорова–Смирнова для логистического и гамма-распределения // Сб. тр. ВНИИ систем. исслед. – 1990, № 8.
20. Лемешко Б.Ю., Постовалов С.Н. Прикладные аспекты использования критериев согласия в случае проверки сложных гипотез // Надежность и контроль качества. – 1997. – № 11. – С. 3-17.
21. Лемешко Б.Ю., Постовалов С.Н. О распределениях статистик непараметрических критериев согласия при оценивании по выборкам параметров наблюдаемых законов // Заводская лаборатория. 1998. – Т. 64. – № 3. – С. 61-72.
22. Лемешко Б.Ю., Постовалов С.Н. О зависимости распределений статистик непараметрических критериев и их мощности от метода оценивания параметров // Заводская лаборатория. Диагностика материалов. 2001. – Т. 67. – № 7.
23. Лемешко Б.Ю., Постовалов С.Н. Применение непараметрических критериев согласия при проверке сложных гипотез // Автометрия. 2001. – № 2. – С. 88-102.
24. Лемешко Б.Ю., Постовалов С.Н. Непараметрические критерии при проверке сложных гипотез о согласии с распределениями Джонсона // Доклады СО АН ВШ. 2002. - № 1(5). - С.65-74.
25. Лемешко Б.Ю., Маклаков А.А. Непараметрические критерии при проверке сложных гипотез о согласии с распределениями экспоненциального семейства // Автометрия. 2004. №3. - С. 3-20.
26. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.

27. Birch M.W. A new proof of the Pearson–Fisher theorem // Ann. Math. Statist. – 1964. V. 35. – P. 817.
28. Лемешко Б.Ю., Чимитова Е.В. Оптимальные L -оценки параметров сдвига и масштаба распределений по выборочным квантилям // Заводская лаборатория. Диагностика материалов. 2004. – Т.70. – №1. – С. 54-66.
29. Chernoff H., Lehmann E.L. The use of maximum likelihood estimates in χ^2 test for goodness of fit // Ann. Math. Stat., 1954. – V. 25. – P. 579-586.
30. Чибисов Д.М. Некоторые критерии типа хи-квадрат для непрерывных распределений // Теория вероятностей и ее применения. 1971. – Т. XVI. – № 1. – С. 3-20.
31. Никулин М.С. Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // Теория вероятностей и ее применение. 1973. – Т. XVIII. – № 3. – С. 583-591.
32. Лемешко Б.Ю. Постовалов С.Н. Прикладные аспекты использования критериев согласия в случае проверки сложных гипотез // Надежность и контроль качества. – 1997. – № 11. – С. 3-17.
33. Лемешко Б.Ю., Постовалов С.Н. О зависимости предельных распределений статистик χ^2 Пирсона и отношения правдоподобия от способа группирования данных // Заводская лаборатория. 1998. – Т. 64. – № 5. – С.56-63.
34. Денисов В.И., Лемешко Б.Ю. Оптимальное группирование при обработке экспериментальных данных // Измерительные информационные системы. – Новосибирск, 1979. – С. 5-14.
35. Денисов В.И., Лемешко Б.Ю., Цой Е.Б. Оптимальное группирование, оценка параметров и планирование регрессионных экспериментов: В 2-х ч. / Новосиб. гос. техн. ун-т. – Новосибирск, 1993. – 346 с.
36. Лемешко Б.Ю. Асимптотически оптимальное группирование наблюдений – это обеспечение максимальной мощности критериев // Надежность и контроль качества. 1997. – № 8. – С. 3–14.
37. Лемешко Б.Ю. Асимптотически оптимальное группирование наблюдений в критериях согласия // Заводская лаборатория. 1998. – Т. 64. – №1. – С. 56-64.
38. 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.

39. Van der Vaart A.W. Asymptotic Statistics. – Cambridge University Press. 1998. –443 p.
40. Никулин М.С. О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. – 1973. – Т. XVIII. – № 3. – С.675-676.
41. Мирвалиев М., Никулин М.С. Критерии согласия типа хи-квадрат // Заводская лаборатория. 1992. – Т. 58. – № 3. – С.52-58.
42. Aguirre N., Nikulin M. Chi-squared goodness-of-fit test for the family of logistic distributions // Kybernetika. – 1994. – V. 30. – № 3. – P.214-222.
43. Greenwood P.E., Nikulin M.S. A Guide to Chi-Squared Testing. – John Wiley & Sons, Inc. 1996. – 280 p.
44. Лемешко Б.Ю., Чимитова Е.В. О выборе числа интервалов в критериях согласия типа χ^2 // Заводская лаборатория. Диагностика материалов. 2003. – Т. 69. – № 1. – С. 61-67.
45. Воинов В.Г. Об оптимальных свойствах критерия Рао-Робсон-Никулина // Заводская лаборатория. Диагностика материалов. 2006. – Т.72. – № 3. – С.65-70.
46. Лемешко Б.Ю., Постовалов С.Н., Чимитова Е.В. О распределениях статистики и мощности критерия типа χ^2 Никулина //Заводская лаборатория. Диагностика материалов. 2001. – Т.67. – №3. – С.52-58.
47. Лемешко Б.Ю., Чимитова Е.В. Об ошибках и неверных действиях, совершаемых при использовании критериев согласия типа χ^2 // Измерительная техника. 2002. - № 6. - С. 5-11.

Некоторые проблемы критериев проверки нормальности

При проверке отклонения распределения от нормального закона используются:

1. Критерий проверки на симметричность.
2. Критерий проверки на эксцесс.
3. Критерий Шапиро-Уилка.
4. Критерий Эппса-Палли.
5. Модифицированный критерий Шапиро-Уилка.
6. Совместный критерий проверки на симметричность и нулевой коэффициент эксцесса.
7. Модификация D'Agostino критерия проверки на симметричность.
8. Модификация D'Agostino критерия проверки на симметричность и эксцесс.
9. Совместный критерий проверки на симметричность и нулевой коэффициент эксцесса D'Agostino.

И ещё целый ряд критериев.

Критерий Шапиро-Уилка

Статистика критерия имеет вид

$$W = S^2 / \sum_{i=1}^n (X_i - \bar{X})^2 ,$$

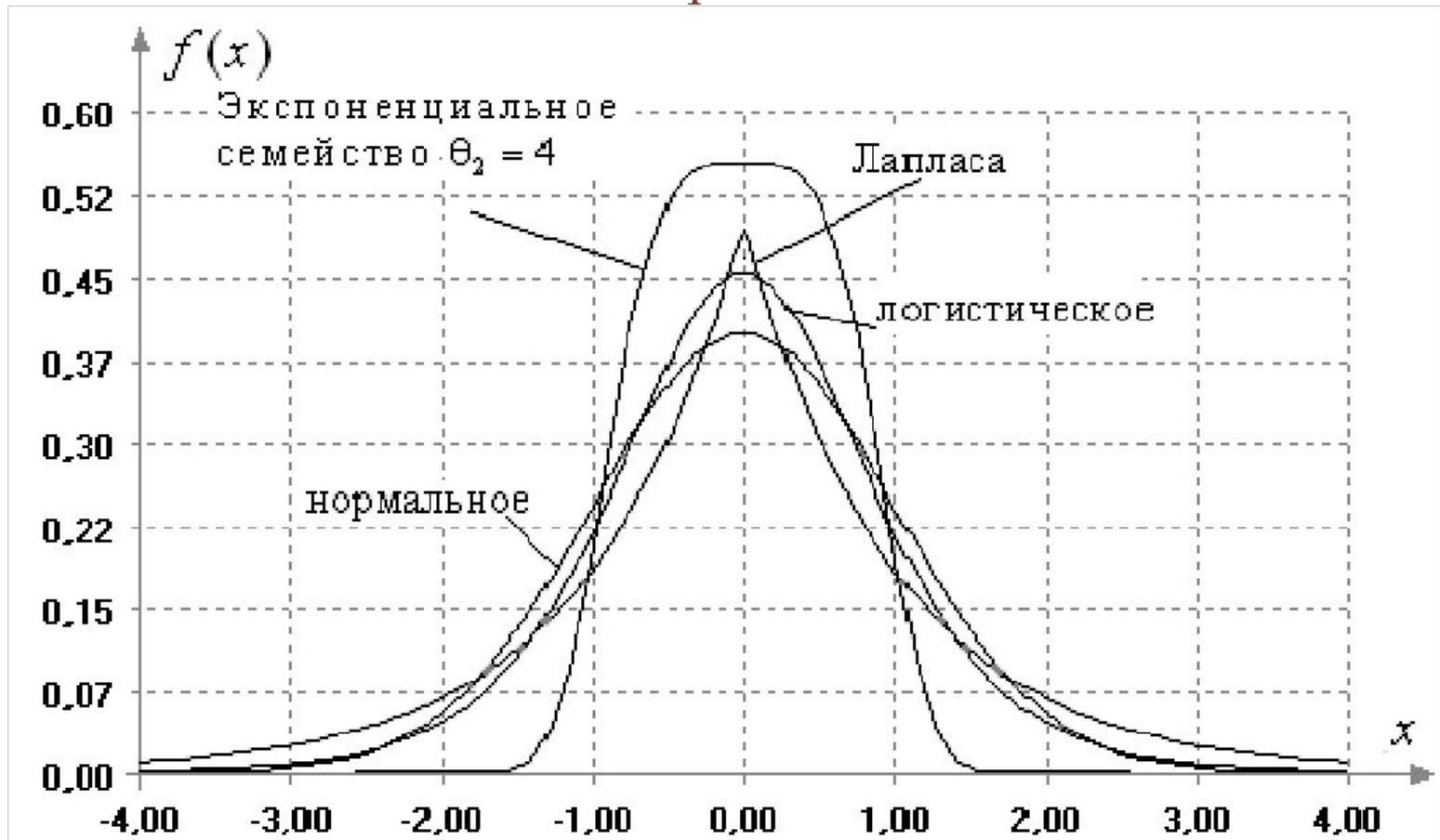
где

$$S = \sum_k a_k [X_{(n+1-k)} - X_{(k)}] ,$$

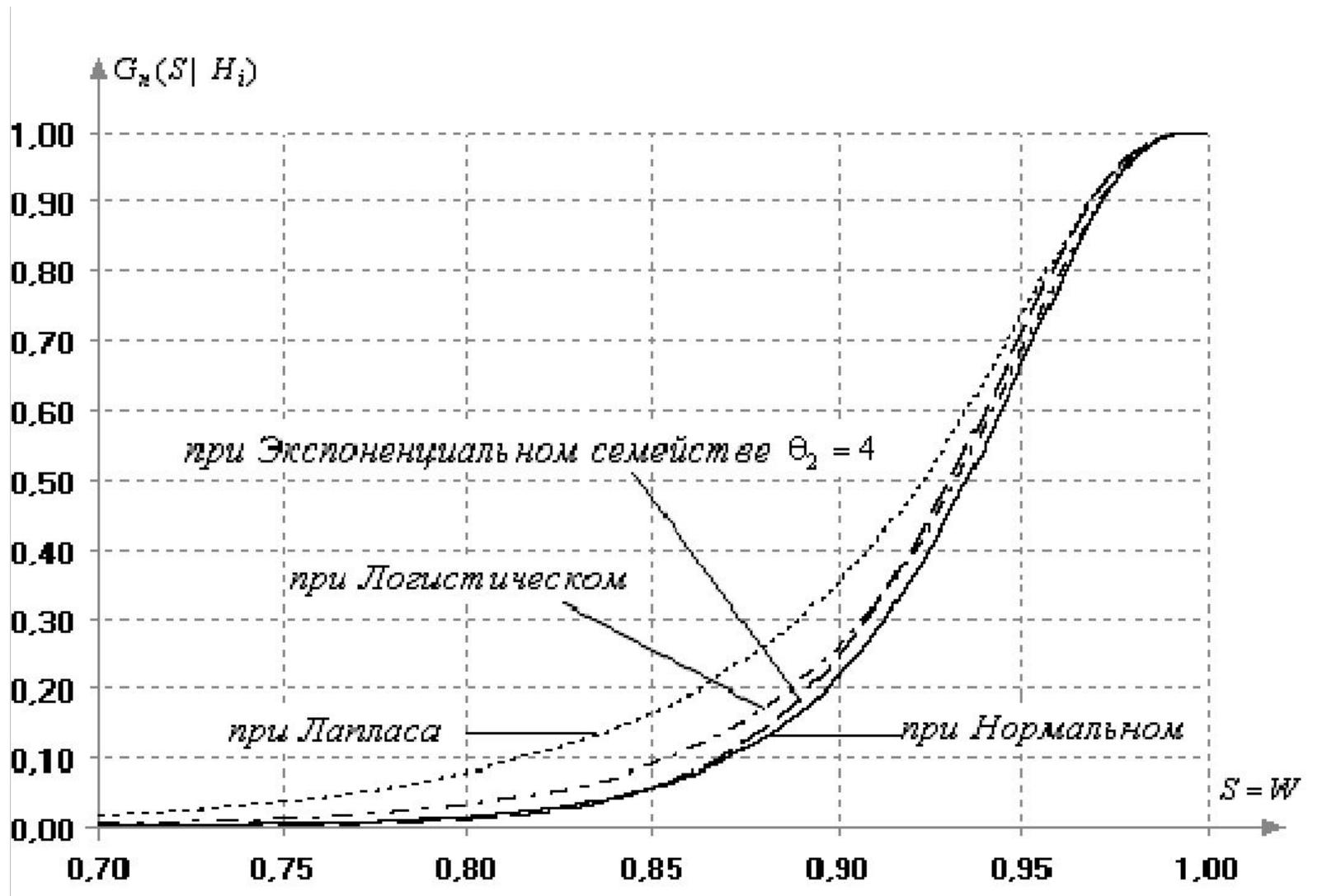
коэффициенты a_k приведены в стандарте и первоисточниках.

Гипотеза о нормальности отвергается при малых значениях статистики W .

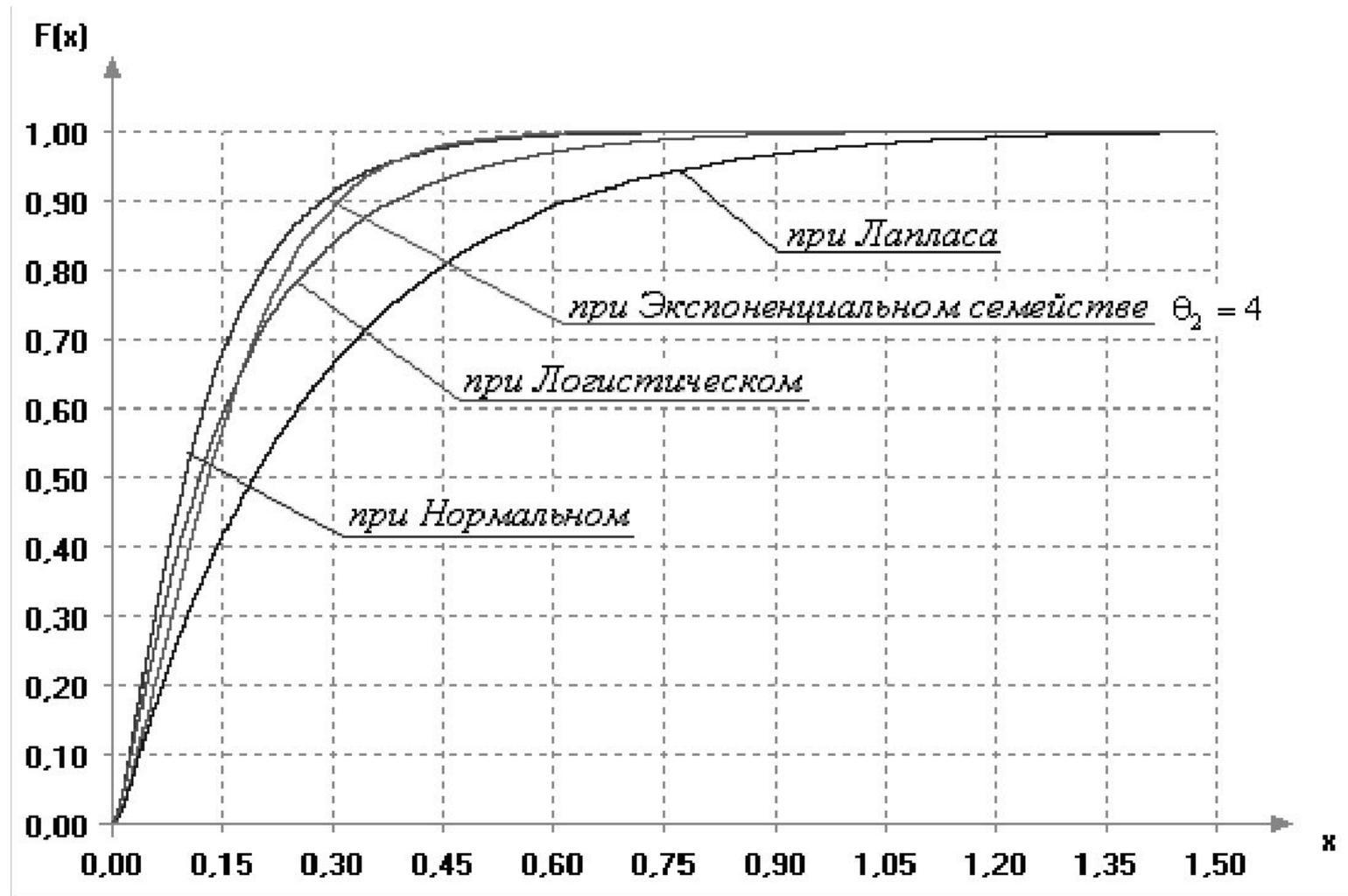
Рассмотрим способность критерия отличать некоторые альтернативы



Плотности распределений, соответствующих рассматриваемым гипотезам H_i



Распределения статистики критерия Шапиро-Уилка в зависимости от наблюдаемого закона при $n=20$



Распределения статистики критерия Эппса-Палли в зависимости от наблюдаемого закона при $n=20$

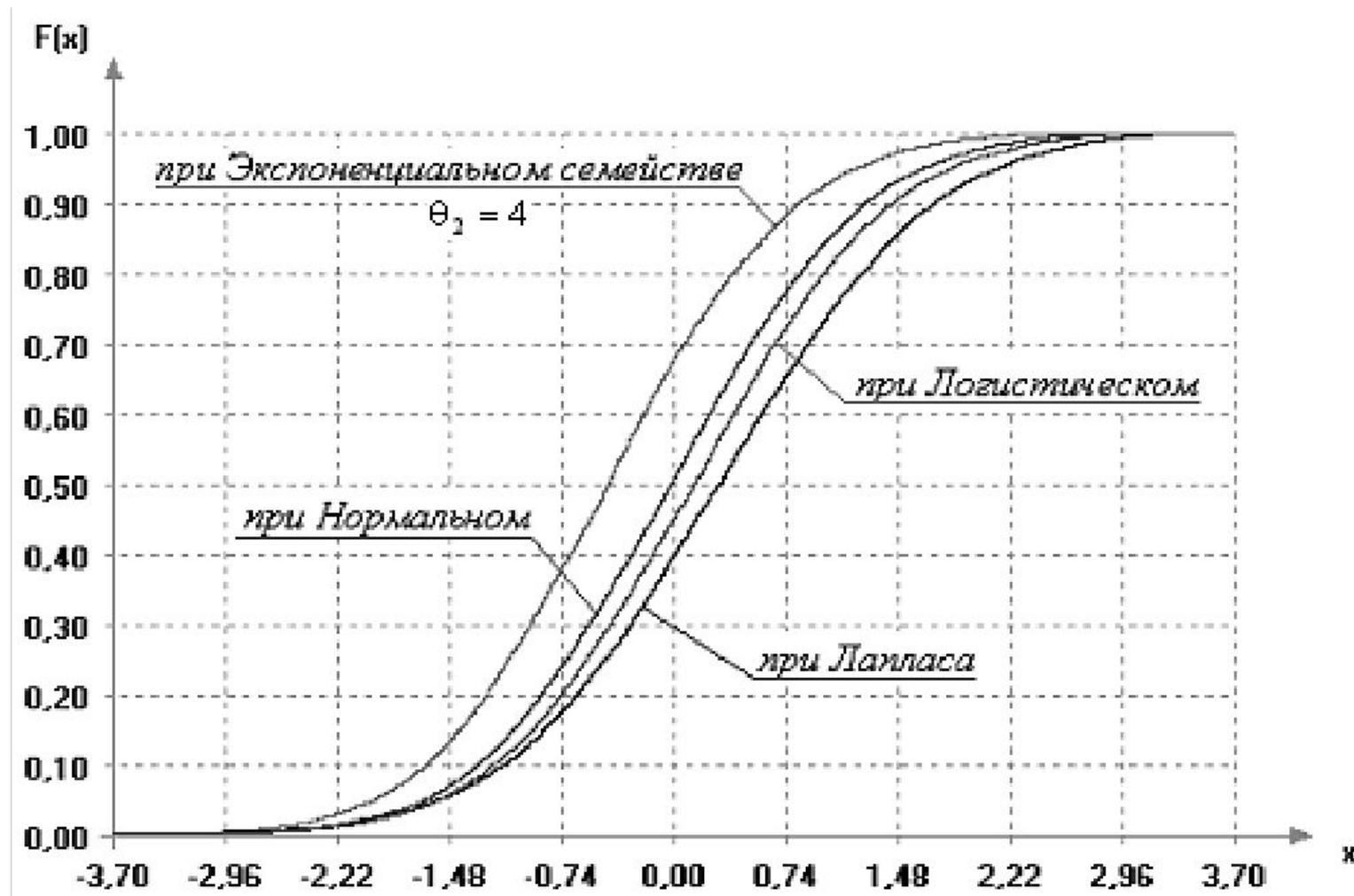
Эти рисунки говорят о смещенности критериев относительно некоторых альтернатив.

Модификация D'Agostino критерия проверки отклонения от нормального закона

При помощи следующих преобразований статистики критериев симметричности и эксцесса преобразуются в статистику, приближенно распределенную по стандартному нормальному закону:

$$\begin{aligned}\delta &= (n-3)(n+1)(n^2 + 15n - 4), \\ a &= \frac{(n-2)(n+5)(n+7)(n^2 + 27n - 70)}{6\delta}, \\ c &= \frac{(n-7)(n+5)(n+7)(n^2 + 2n - 5)}{6\delta}, \\ k &= \frac{(n+5)(n+7)(n^3 + 37n^2 + 11n - 313)}{12\delta}, \\ \alpha &= a + \beta_1 c, \\ \chi &= (\beta_2 - 1 - \beta_1)2k, \\ z_2 &= \left\{ \left(\frac{\chi}{2\alpha} \right)^{1/3} - 1 + \frac{1}{9\alpha} \right\} (9\alpha)^{1/2}, \quad (1)\end{aligned}$$

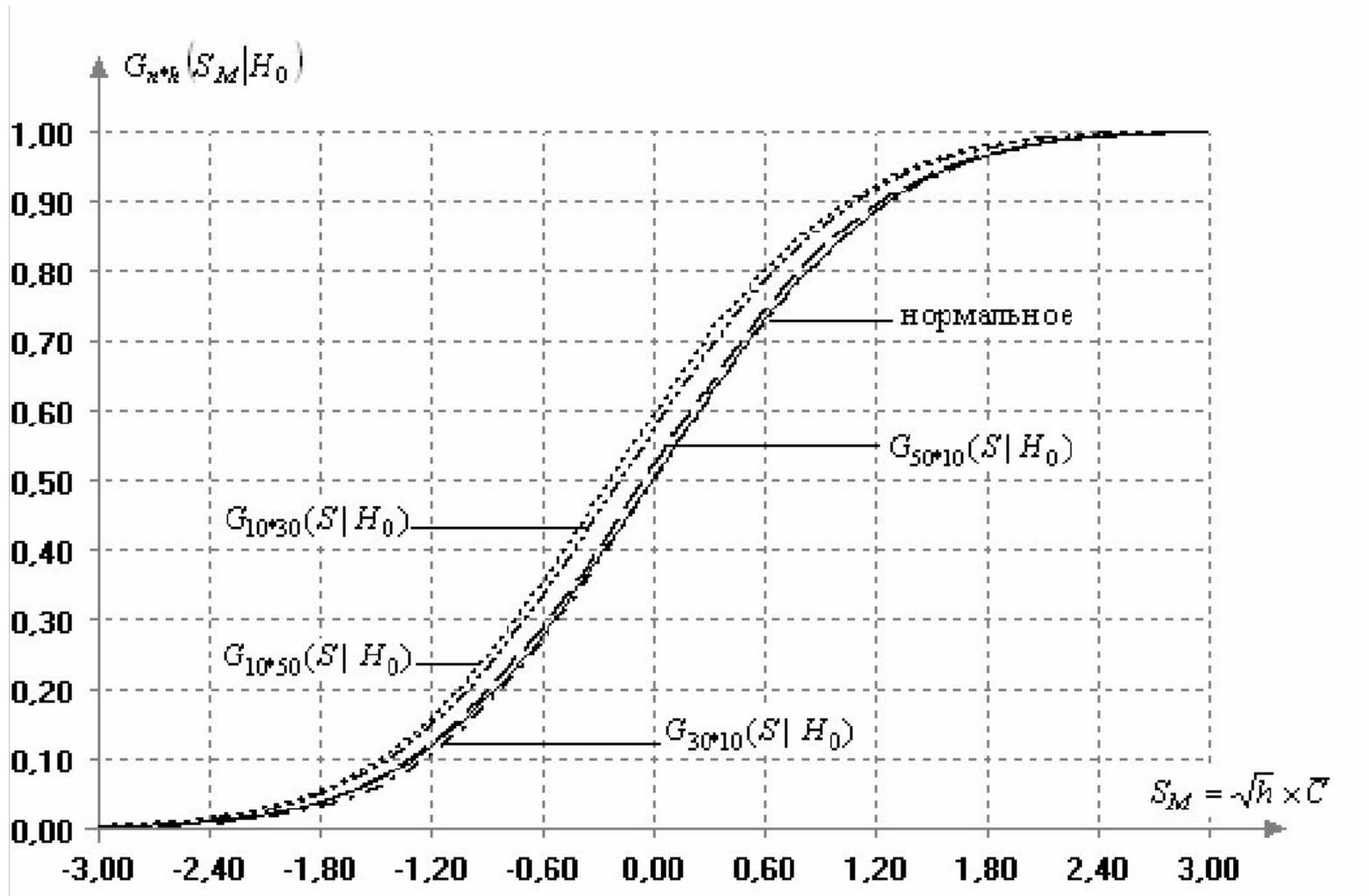
где β_1 - оценка коэффициента симметрии, β_2 - оценка коэффициента эксцесса



Условные распределения статистики D'Agostino в зависимости от вида наблюдаемого закона при $n=10$

Данный критерий отличает все альтернативы

Распределения статистики модифицированного критерия Шапиро-Уилка плохо сходятся к «предельному»



Условные распределения $G_{n \cdot h}(S_M | H_0)$ статистики модифицированного критерия Шапиро-Уилка при различных комбинациях n и h

Критерии проверки гипотез об однородности средних

*при нарушении классических
предположений о принадлежности
наблюдений (ошибок измерений)
нормальному закону*

Об устойчивости распределений статистик критериев проверки гипотез о средних

Проверяемая гипотеза о равенстве математических ожиданий задается в виде

$$H_0 : \mu_1 = \mu_2,$$

конкурирующая –

$$H_1 : \mu_1 \neq \mu_2.$$

Для проверки гипотезы H_0 могут использоваться *параметрические критерии*, предназначенные для проверки гипотезы в случае принадлежности наблюдений нормальному закону:

- критерий сравнения двух выборочных средних при известных дисперсиях;
- критерий сравнения двух выборочных средних при неизвестных, но равных дисперсиях (критерий Стьюдента);
- критерий сравнения двух выборочных средних при неизвестных и неравных дисперсиях (проблема Беренса-Фишера).

Непараметрические критерии:

U -критерий Уилкоксона, Манна–Уитни;

H -критерий Краскела–Уаллиса.

Статистики непараметрических критериев

№	Статистика	Условия	Распределение статистики
1	$z = \frac{\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$ $\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j$	σ_1^2, σ_2^2 - известны	Стандартное нормальное $N(0,1)$
2	$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left[\frac{n_1 + n_2}{n_1 n_2} \right] \left[\frac{Q_1 + Q_2}{n_1 + n_2 - 2} \right]}}$ $Q_i = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$	$\sigma_1^2 = \sigma_2^2$ - неизвестны	t_v -распределение $v = n_1 + n_2 - 2$
3	$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]}}$ $s_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$	$\sigma_1^2 \neq \sigma_2^2$ - неизвестны	t_v -распределение $v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1}}$

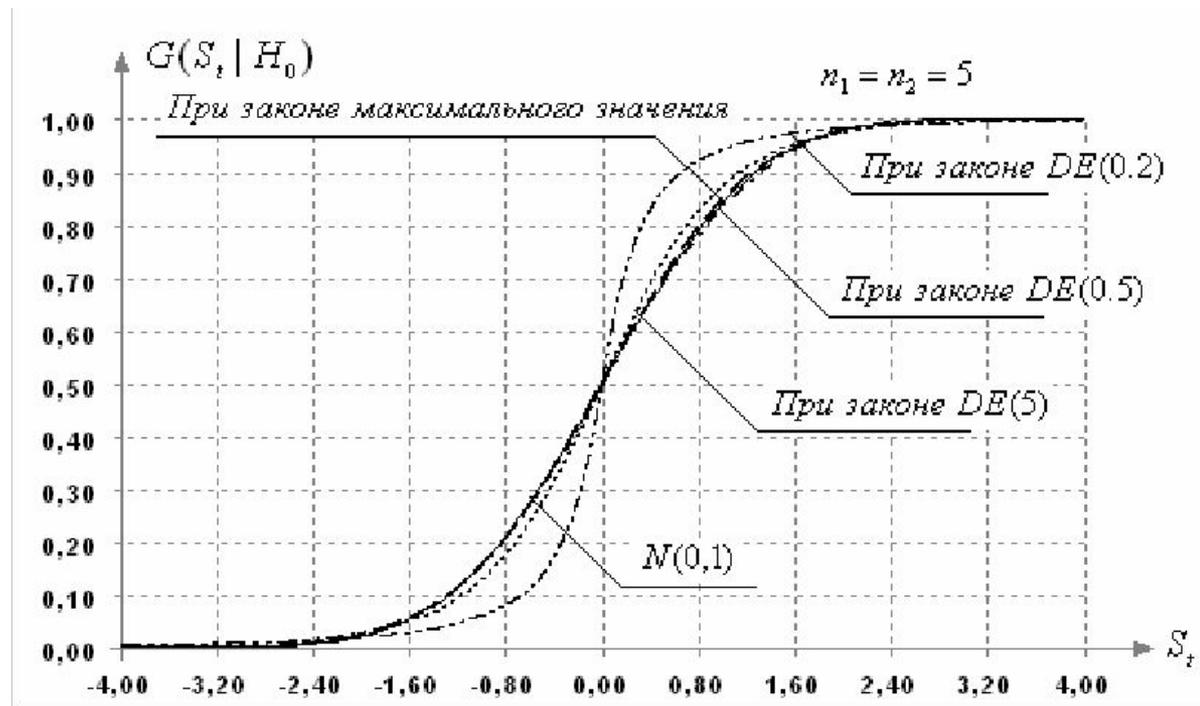


Рис. Эмпирические распределения статистики (1) при различных законах распределения наблюдаемых случайных величин при объемах выборок $n_1 = n_2 = 5$

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

Распределение статистики существенно отличается от стандартного нормального только в случае принадлежности наблюдений законам с очень “тяжелыми” хвостами.

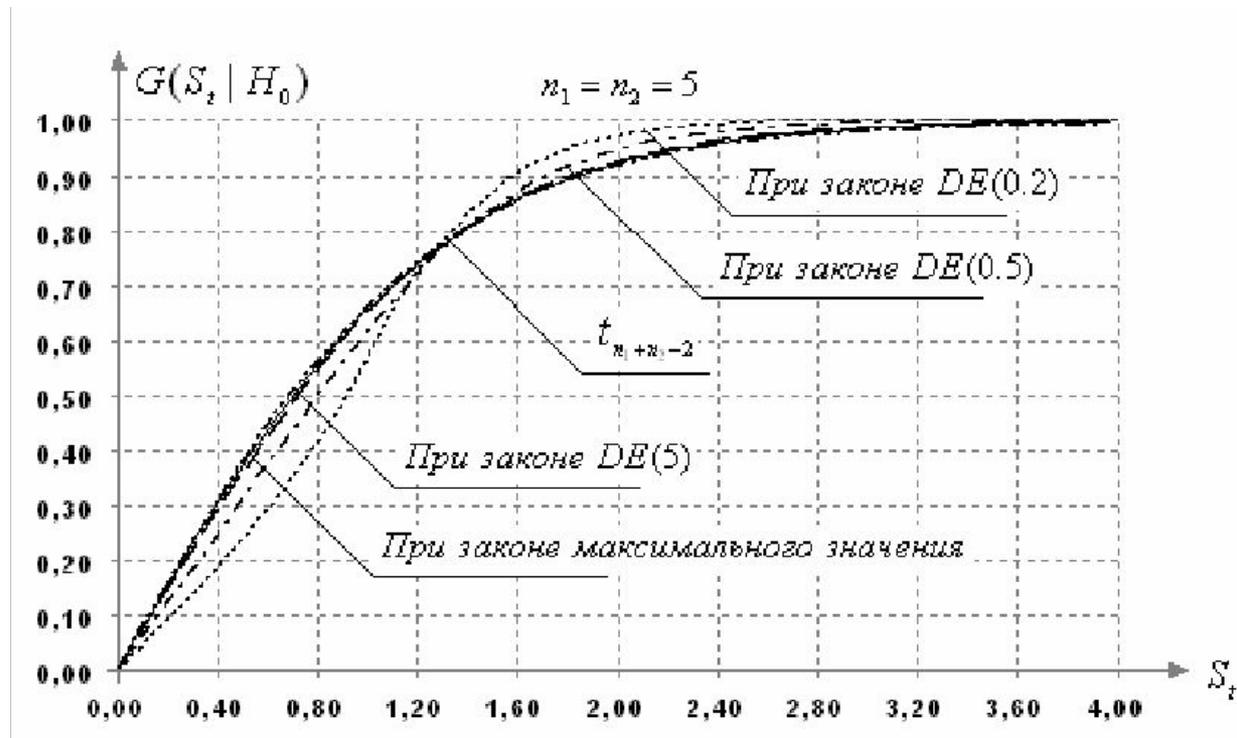


Рис. Эмпирические функции распределения статистики (2) при различных законах распределения наблюдаемых случайных величин при объемах выборок $n_1 = n_2 = 5$

Аналогично, распределение статистики существенно отличается от соответствующего t -распределения только в случае принадлежности наблюдений законам с очень “тяжелыми” хвостами.

Мощность критериев однородности

		$1 - \beta$			
		$H_1^1: \mu_1 = \mu_0 + 0,1 * \sigma$			
n	α	При неизвестных, но равных дисперсиях	При неизвестных и неравных дисперсиях	U-критерий Уилкоксона, Манна и Уитни	H-критерий Краскела-Уаллиса
10	0.1	0.144	0.133	0.113	0.101
	0.05	0.076	0.069	0.057	0.047
	0.01	0.017	0.014	0.008	0.007
20	0.1	0.165	0.149	0.120	0.111
	0.05	0.090	0.080	0.060	0.055
	0.01	0.021	0.018	0.012	0.010
50	0.1	0.216	0.186	0.138	0.129
	0.05	0.125	0.104	0.076	0.068
	0.01	0.033	0.026	0.018	0.015

Выводы

- 1. Параметрические критерии проверки однородности средних устойчивы к отклонениям законов распределения наблюдаемых случайных величин от нормального.**
- 2. Отличия распределений статистик от “классических” становятся заметными в случае асимметричности наблюдаемых законов, но эти отклонения не очень значительны.**
- 3. Существенные отклонения распределений статистик от “классических” наблюдаются в случае законов с тяжелыми хвостами.**
- 4. Мощность параметрических критериев выше непараметрических. Следовательно, их применение в свете устойчивости к отклонениям от классических предположений о нормальности является предпочтительным.**

Устойчивость критерия независимости Аббе

Критерий Аббе предназначен для проверки гипотез вида

$$H_0: E[\xi_1] = E[\xi_2] = \dots = E[\xi_n].$$

То есть для проверки того, что все наблюдаемые величины $\xi_1, \xi_2, \dots, \xi_n$ в выборке объема n имеют одинаковые математические ожидания. Конкурирующая гипотеза (альтернатива) заключается в справедливости $|E[\xi_{i+1}] - E[\xi_i]| > 0$ для всех или некоторых значений $i = 1, 2, \dots, n-1$. Критерий часто используют для проверки независимости последовательности измерений, для проверки отсутствия тренда.

Статистика критерия Аббе в современной форме представляет собой отношение

$$S_A = \frac{1}{2} \left(\frac{\sum_{i=1}^{n-1} (\xi_{i+1} - \xi_i)^2}{\sum_{i=1}^n (\xi_i - \bar{\xi})^2} \right),$$

где $\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$.

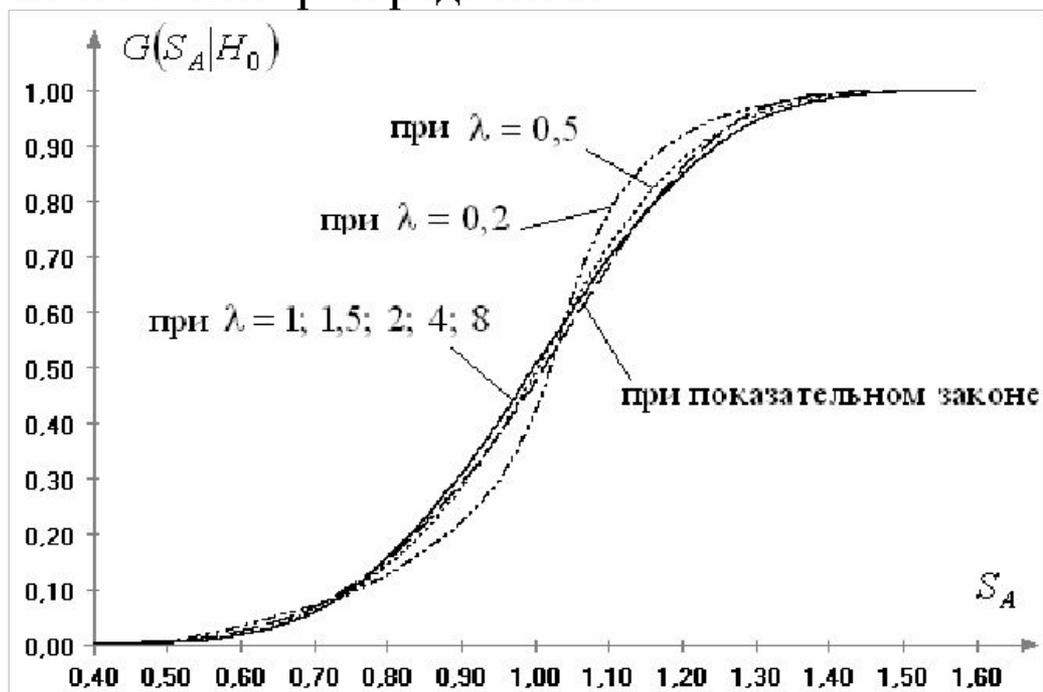
Условное распределение статистики $G(S_{A_n} | H_0)$ при справедливости H_0 зависит от объема выборки n , симметрично относительно 1 и определено на интервале $1 \pm \cos(\pi/n)$. С ростом объемов выборок распределение статистики хорошо приближается нормальным законом с параметром сдвига 1 и со стандартным отклонением при объемах $n > 20$, равным $[(n-2)/(n^2-1)]^{1/2}$.

Решение об отклонении проверяемой гипотезы о равенстве средних принимается, если значение статистики S_A оказывается меньше критического значения $S_{A_n}(\alpha)$, где α – заданный уровень значимости.

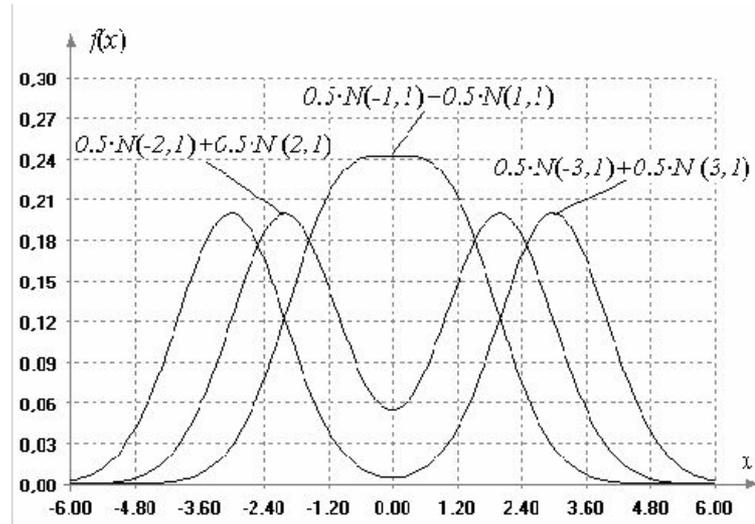
На рис. представлены распределения статистики критерия Аббе в случае принадлежности наблюдений различным законам, в частности, семейству распределений с плотностью

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

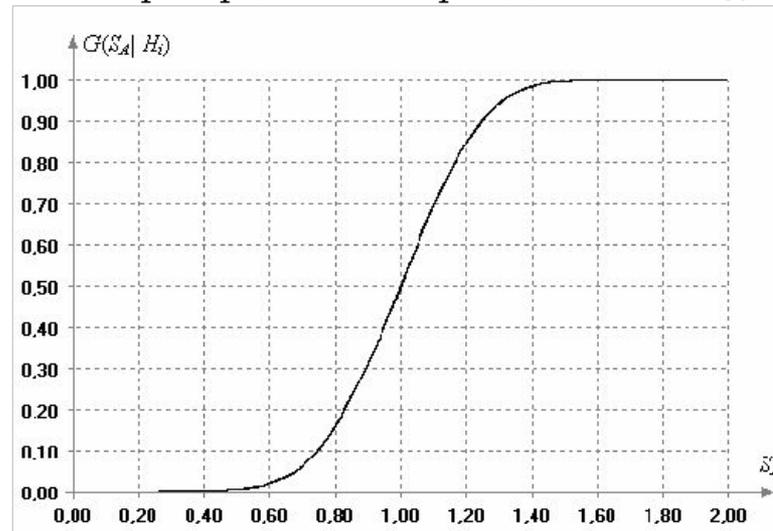
и параметрами формы $\lambda = 0,2; 0,5; 1; 1,5; 2; 4; 8$. В случае $\lambda = 2$ выражение дает плотность нормального закона распределения.



Распределения статистики критерия Аббе в зависимости от параметра формы распределения семейства при $n = 25$



Смеси нормальных распределений при сдвигах $\pm 1\theta_1$, $\pm 2\theta_1$ и $\pm 3\theta_1$



Распределения статистики критерия Аббе в случае нормального закона и в случае смесей нормальных распределений при $n = 25$ (совпадают)

Критерии проверки гипотез об однородности дисперсий

*при нарушении классических
предположений о принадлежности
наблюдений (ошибок измерений)
нормальному закону*

Критерий Бартлетта

Статистика критерия Бартлетта вычисляется в соответствии с соотношением [2]:

$$\chi^2 = M \left[1 + \frac{1}{3(m-1)} \left(\sum_{i=1}^m \frac{1}{v_i} - \frac{1}{N} \right) \right]^{-1} \quad (5)$$

где n_i – объемы выборок, $v_i = n_i$, если математическое ожидание известно, и $v_i = n_i - 1$, если

неизвестно, $N = \sum_{i=1}^m v_i$,

$$M = N \ln \left(\frac{1}{N} \sum_{i=1}^m v_i S_i^2 \right) - \sum_{i=1}^m v_i \ln S_i^2,$$

S_i^2 – оценки выборочных дисперсий. При неизвестном математическом ожидании оценки

$S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ji} - \bar{X}_i)^2$, где $\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ji}$ и $v_i = n_i - 1$. Если гипотеза H_0 верна, все $v_i > 3$ и вы-

борки извлекаются из нормальной генеральной совокупности, то статистика (5) приближенно

подчиняется χ_{m-1}^2 -распределению.

При нормально распределенных наблюдаемых (измеряемых) величинах распределение статистики (5) практически не зависит от изменения объема выборки. Например, на рис. 1 приведены практически совпадающие функции распределения статистики критерия Бартлетта (5) при различных объемах выборок ($n = 10, 50, 100$). Это означает, что в случае нормальности наблюдаемых величин выводы остаются корректными и при очень малых объемах анализируемых выборок.

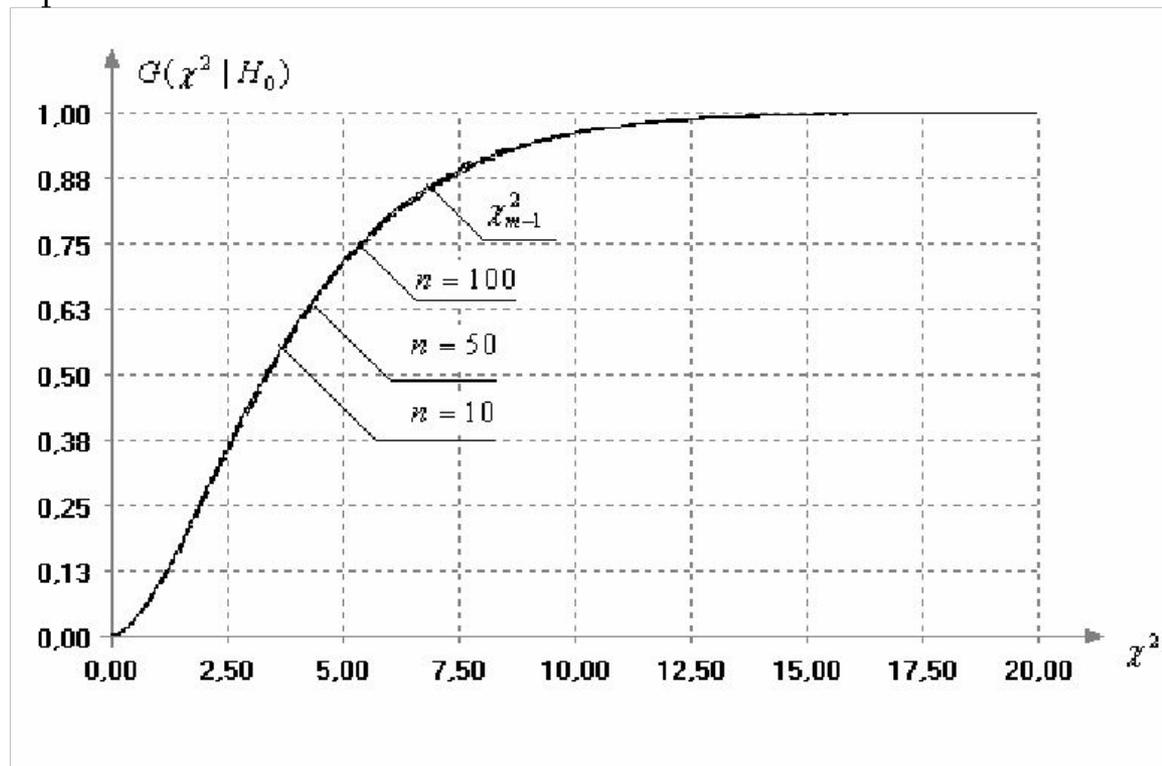


Рис. 1. Функции распределения статистики классического критерия Бартлетта при различных объемах выборок при $m = 5$

В то же время, распределения статистики (5) очень чувствительны к отклонениям наблюдаемого закона от нормального. Вид распределения статистики (5) исследовался при различных наблюдаемых законах, в частности, в случае принадлежности моделируемых выборок законам логистическому с плотностью

$$f(x) = \frac{\pi}{\theta_0 \sqrt{3}} \exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0 \sqrt{3}}\right\} / \left[1 + \exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0 \sqrt{3}}\right\}\right]^2,$$

Лапласа с плотностью

$$f(x) = \frac{\theta_0}{2} \exp\{-\theta_0|x-\theta_1|\},$$

экспоненциальному семейству распределений с различными параметрами формы с плотностью

$$De(\lambda) = f(x) = \frac{\lambda}{2\sqrt{2}\theta_1\Gamma(1/\lambda)} \exp\left(-\left(\frac{|x-\theta_0|}{\sqrt{2}\theta_1}\right)^\lambda\right), \quad (6)$$

где λ – параметр формы. Законы нормальный и Лапласа являются частными случаями данного семейства распределений при значениях параметра формы 2 и 1 соответственно. Семейство (6) может быть хорошей моделью для законов распределения ошибок различных измерительных систем.

Рис. 2 отражает зависимость распределений статистики (5) от вида наблюдаемого закона при различных объемах выборок. Видно, что при отклонении закона распределения наблюдаемого показателя от нормального закона распределение статистики критерия Бартлетта (5) существенно отличается от χ^2_{m-1} -распределения. При этом распределения статистики становятся более зависимыми от объема выборки, чем в случае нормального закона.

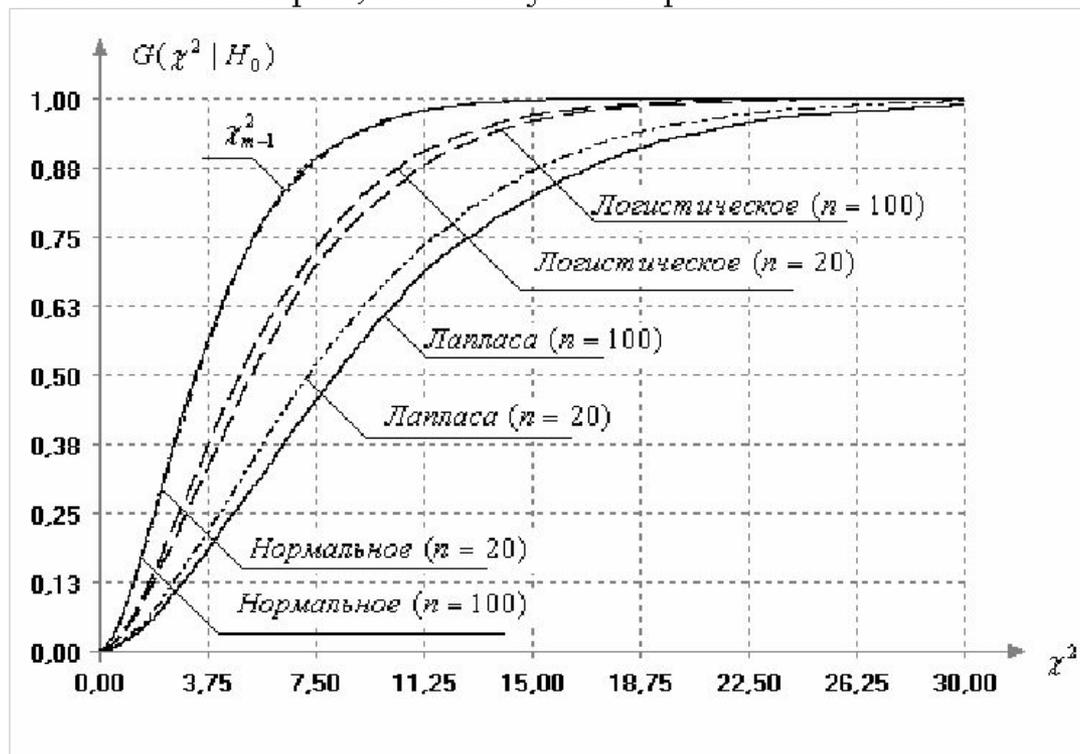


Рис. 2. Функции распределения статистики Бартлетта при отклонении закона распределения наблюдаемого показателя от нормального при различных объемах выборки и $m = 5$

На рис. 3 показано, как меняется распределение статистики Бартлетта, если наблюдаемая случайная величина подчиняется экспоненциальному семейству распределений с различными значениями параметра формы.

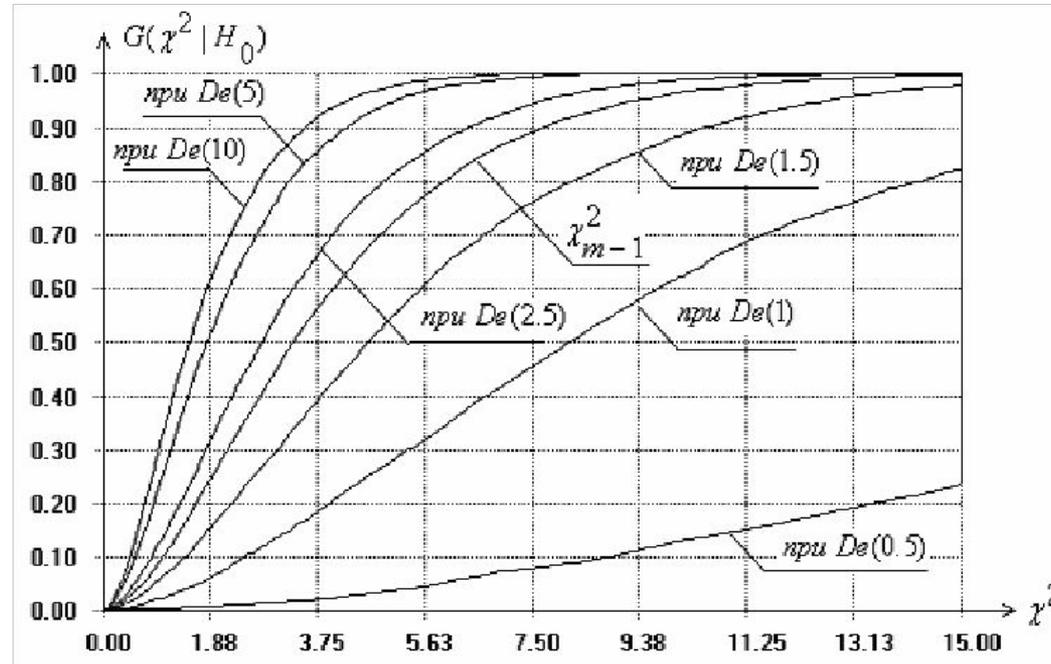


Рис. 3. Функции распределения статистики критерия Бартлетта в случае распределений экспоненциального семейства $De(\lambda)$ с различными значениями параметра формы при $n = 100$ и $m = 5$

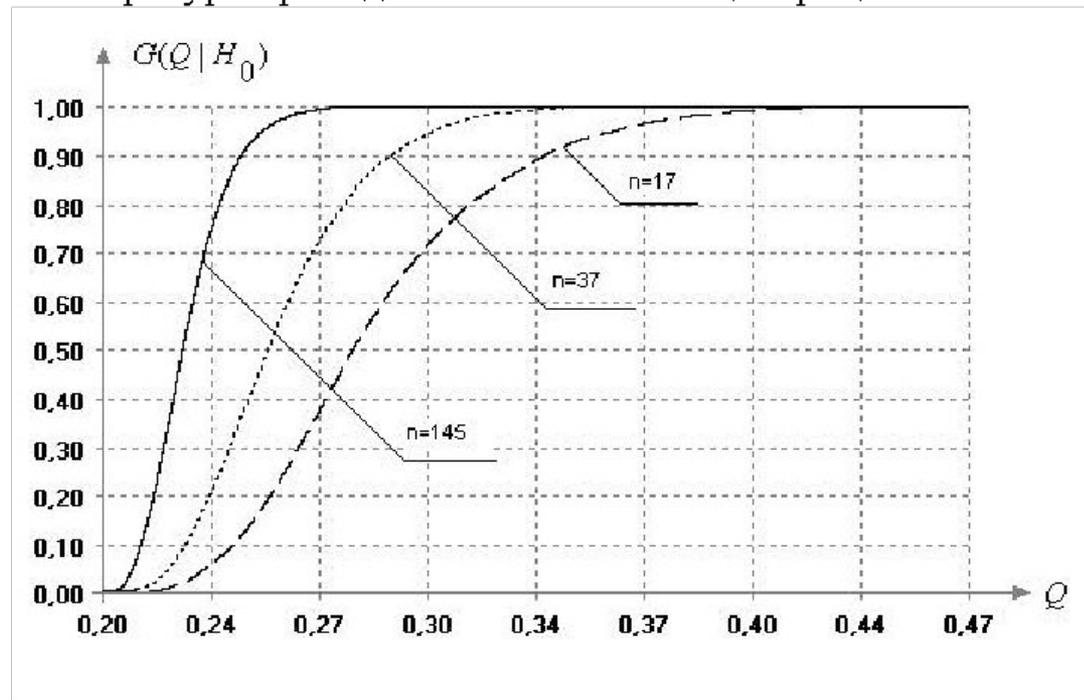
Критерий Кокрена

В том случае, когда все n_i одинаковы, $n_1 = n_2 = \dots = n_m = n$, возможно использование более простого критерия Кокрена. Статистика Q критерия Кокрена выражается формулой [2]

$$Q = \frac{S_{\max}^2}{S_1^2 + S_2^2 + \dots + S_m^2}, \quad (7)$$

где $S_{\max}^2 = \max(S_1^2, S_2^2, \dots, S_m^2)$, где m – число независимых оценок дисперсий (число выборок).

Распределения статистики Кокрена сильно зависят от объема наблюдаемых выборок. Поэтому в справочной литературе приводятся только таблицы процентных точек.



Распределения статистики критерия Кокрена при различных объемах выборок при $m = 5$

Как и критерий Бартлетта, критерий Кокрена используется в предположении, что наблюдаемая случайная величина принадлежит нормальному закону. Насколько сильно меняется распределение статистики Кокрена (7) в случае определенных отклонений закона распределения наблюдаемой случайной величины (контролируемого показателя) от нормального показывает следующий рисунок.

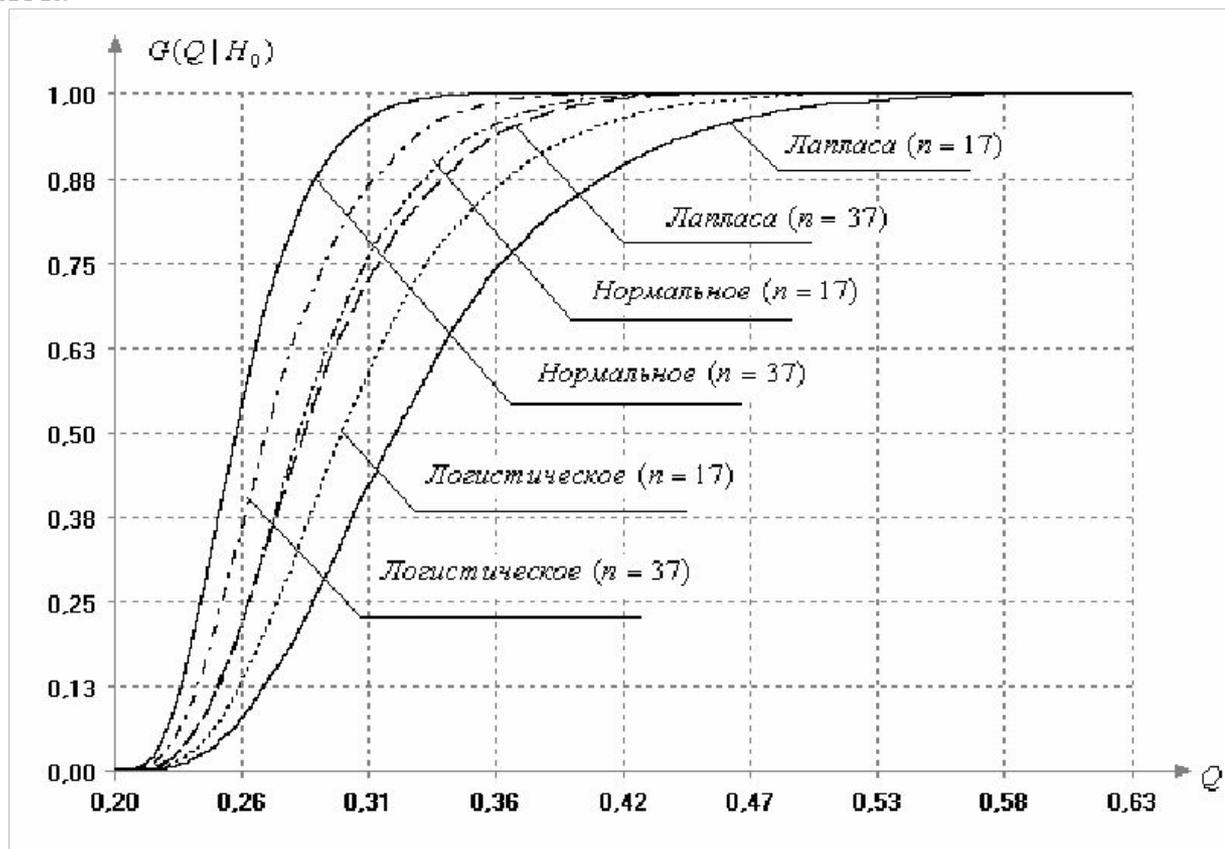
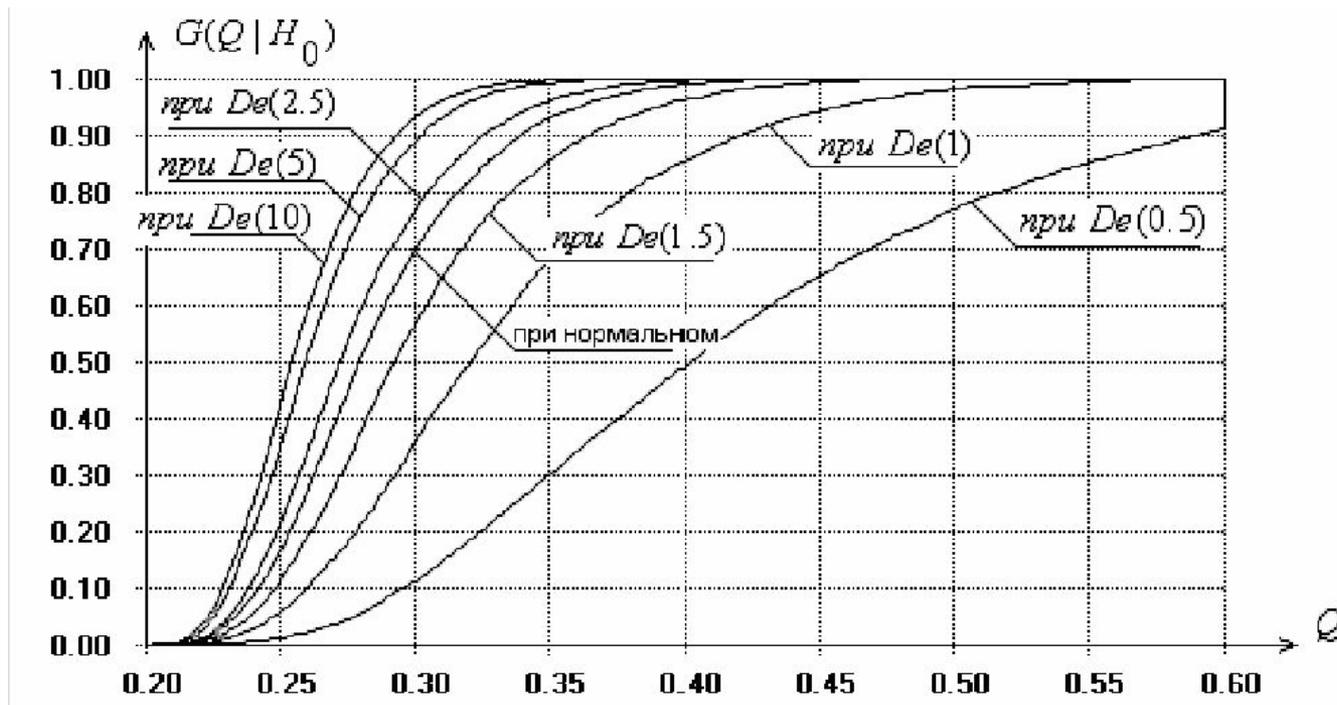


Рис. 5. Функции распределения статистики критерия Кокрена при отклонении закона распределения наблюдаемого показателя от нормального при различных объемах выборки при $m = 5$



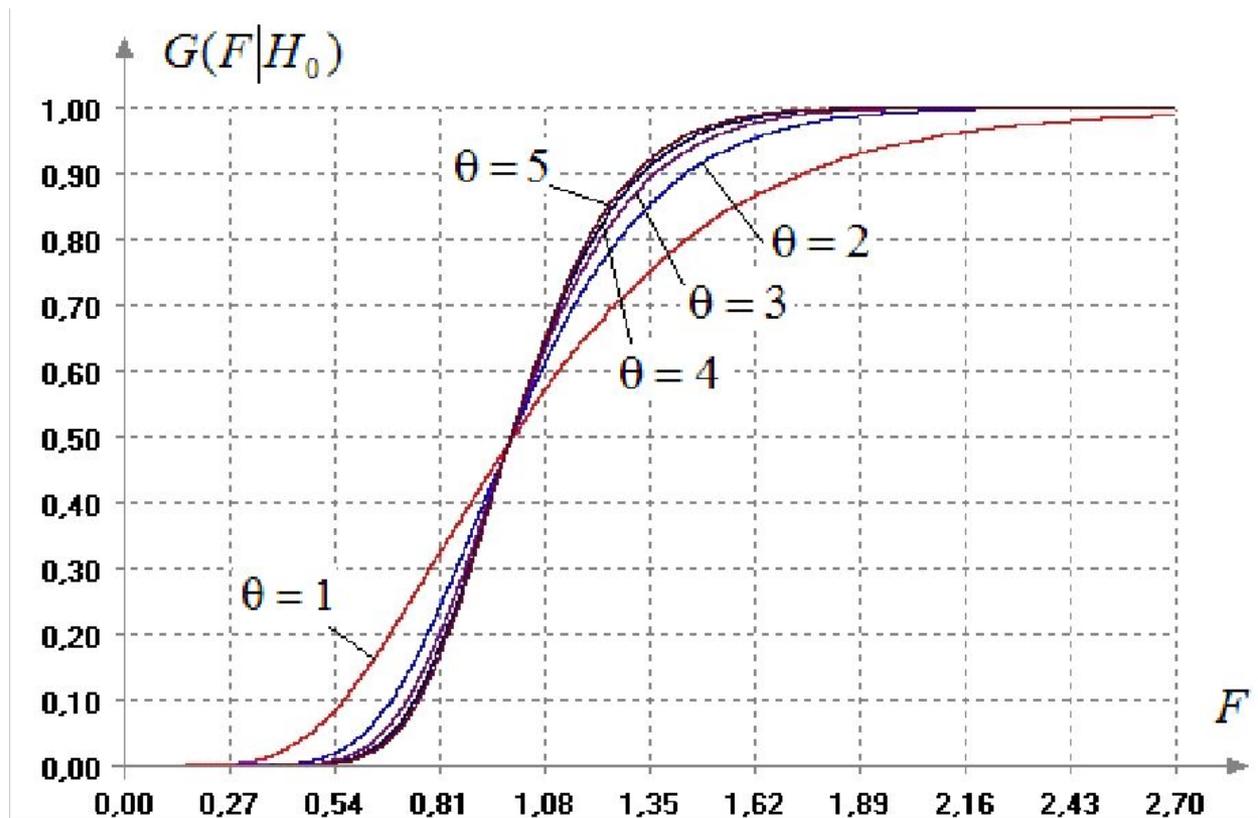
Функции распределения статистики критерия Кокрена в случае распределений экспоненциального семейства с различными значениями параметра формы при $n = 17$ и $m = 5$

Критерий Фишера

Для сравнения двух выборочных дисперсий из нормальных совокупностей. Для определения того, относятся ли две выборки к одной и той же генеральной совокупности, проверяется гипотеза вида $H_0 : \sigma_1^2 = \sigma_2^2$. Статистика для проверки гипотезы имеет вид

$$F = \frac{s_1^2}{s_2^2}.$$

В случае принадлежности выборок нормальному закону и справедливости H_0 эта статистика подчиняется F_{v_1, v_2} -распределению Фишера с числом степеней свободы $v_1 = n_1 - 1$ и $v_2 = n_2 - 1$.



Распределения статистики Фишера при справедливой гипотезе H_0 объеме 2-х выборок по $n = 50$ в случае принадлежности выборок различным законам из экспоненциального семейства: $\theta = 1$, $\theta = 2$ (нормальный), $\theta = 3$, $\theta = 4$, $\theta = 5$.

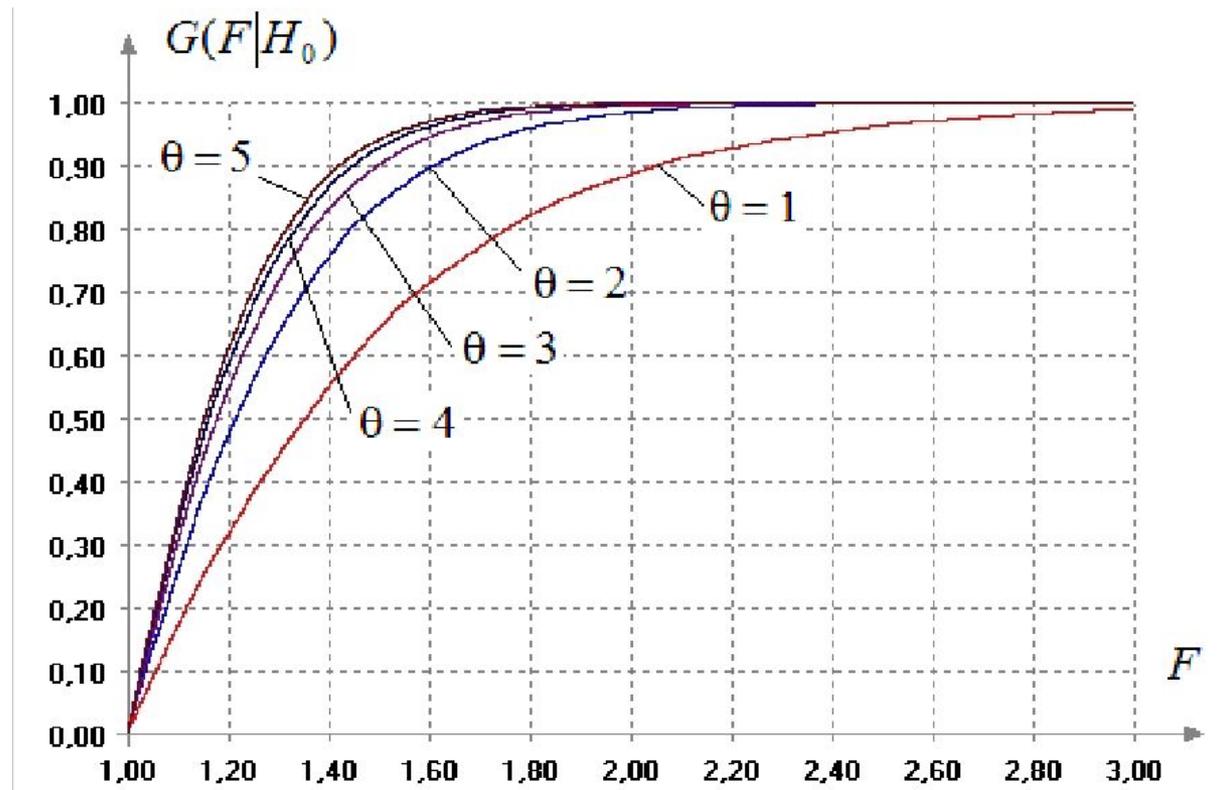
Критерий Хартли

Проверка равенства нескольких дисперсий для выборок равного объема по Хартли.

Статистика для проверки гипотезы имеет вид

$$F = \frac{S_{\max}^2}{S_{\min}^2}.$$

Степенями свободы для распределения статистики являются число выборок $\nu_1 = m$ и $\nu_2 = n - 1$. В литературе для статистики приводятся лишь таблицы процентных точек.



Распределения статистики Хартли при справедливой гипотезе H_0 объеме 2-х выборок по $n = 50$ в случае принадлежности выборок различным законам из экспоненциального семейства: $\theta = 1$, $\theta = 2$ (нормальный), $\theta = 3$, $\theta = 4$, $\theta = 5$.

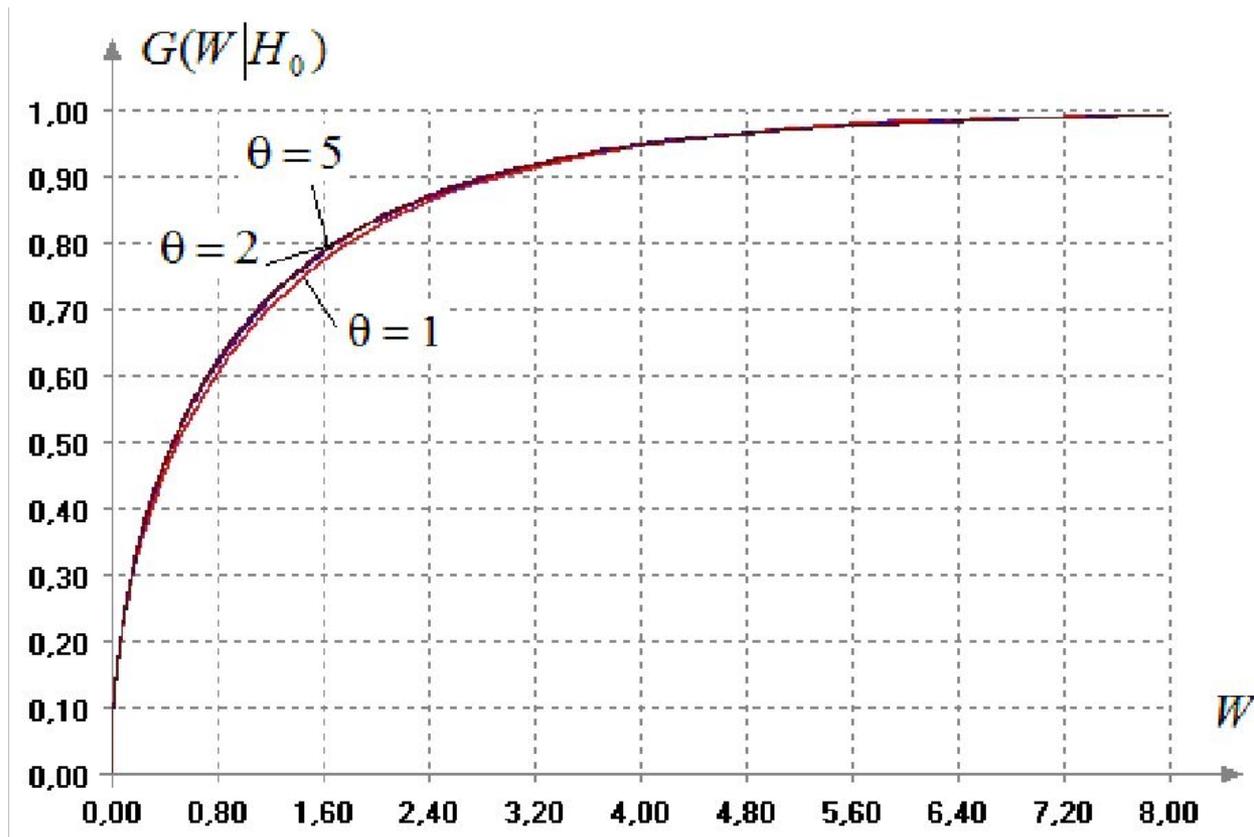
Критерий Левене

Считается, что критерий Левене менее чувствителен к отклонениям от нормальности, чем критерий Бартлетта.

Пусть n_i – объем i -й выборки, $N = \sum_{i=1}^m n_i$, X_{ij} – j -е наблюдение в i -й выборке. Статистика критерия Левене имеет вид:

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

где Z_{ij} определяется выражением $Z_{ij} = |X_{ij} - \bar{X}_{i\cdot}|$, в котором $\bar{X}_{i\cdot}$ – среднее в i -й выборке. $\bar{Z}_{i\cdot}$ – среднее Z_{ij} по i -й выборке, $\bar{Z}_{\cdot\cdot}$ – среднее Z_{ij} по всем выборкам. В случае принадлежности выборок нормальному закону и справедливости H_0 эта статистика подчиняется F_{v_1, v_2} - распределению Фишера с числом степеней свободы $v_1 = m - 1$ и $v_2 = N - m$.



Распределения статистики Левене при справедливой гипотезе H_0 объеме 2-х выборок по $n = 50$ в случае принадлежности выборок различным законам из экспоненциального семейства: $\theta = 1$, $\theta = 2$ (нормальный), $\theta = 3$, $\theta = 4$, $\theta = 5$.

Как видим, критерий Левене оказывается устойчивым к нарушению предположения нормальности.

В оригинальном критерии Левене предусмотрено использование только выборочных средних. Brown и Forsythe расширили критерий Левене на случай использования выборочных медиан и усеченного среднего ($Z_{ij} = |X_{ij} - \tilde{X}_{i\cdot}|$, где $\tilde{X}_{i\cdot}$ – медиана в i -й выборке; $Z'_{ij} = |X_{ij} - \bar{X}'_{i\cdot}|$, где $\bar{X}'_{i\cdot}$ – усеченное среднее в i -й выборке). **Говорится**, что в этих случаях критерий становится устойчивей к нарушению предположений о нормальности, однако, не приводит ли это к тому, что распределение статистики будет несколько отличаться от F_{v_1, v_2} -распределения.

Выводы

Критерии проверки однородности дисперсий, как правило, очень чувствительны к нарушению предположения о принадлежности наблюдений нормальному закону.

Исключение составляет критерий Левене. Этому есть объяснение. В критерии Левене проверка однородности дисперсий сводится к проверке «однородности средних».

Аналогичная ситуация характерна для F-критерия, применяемого для проверки гипотез об однородности средних.

Критерий Шеффе, используемый в дисперсионном анализе в аналогичной ситуации проверки однородности средних, также обладает известной устойчивостью.

Модификация критерия однородности Смирнова

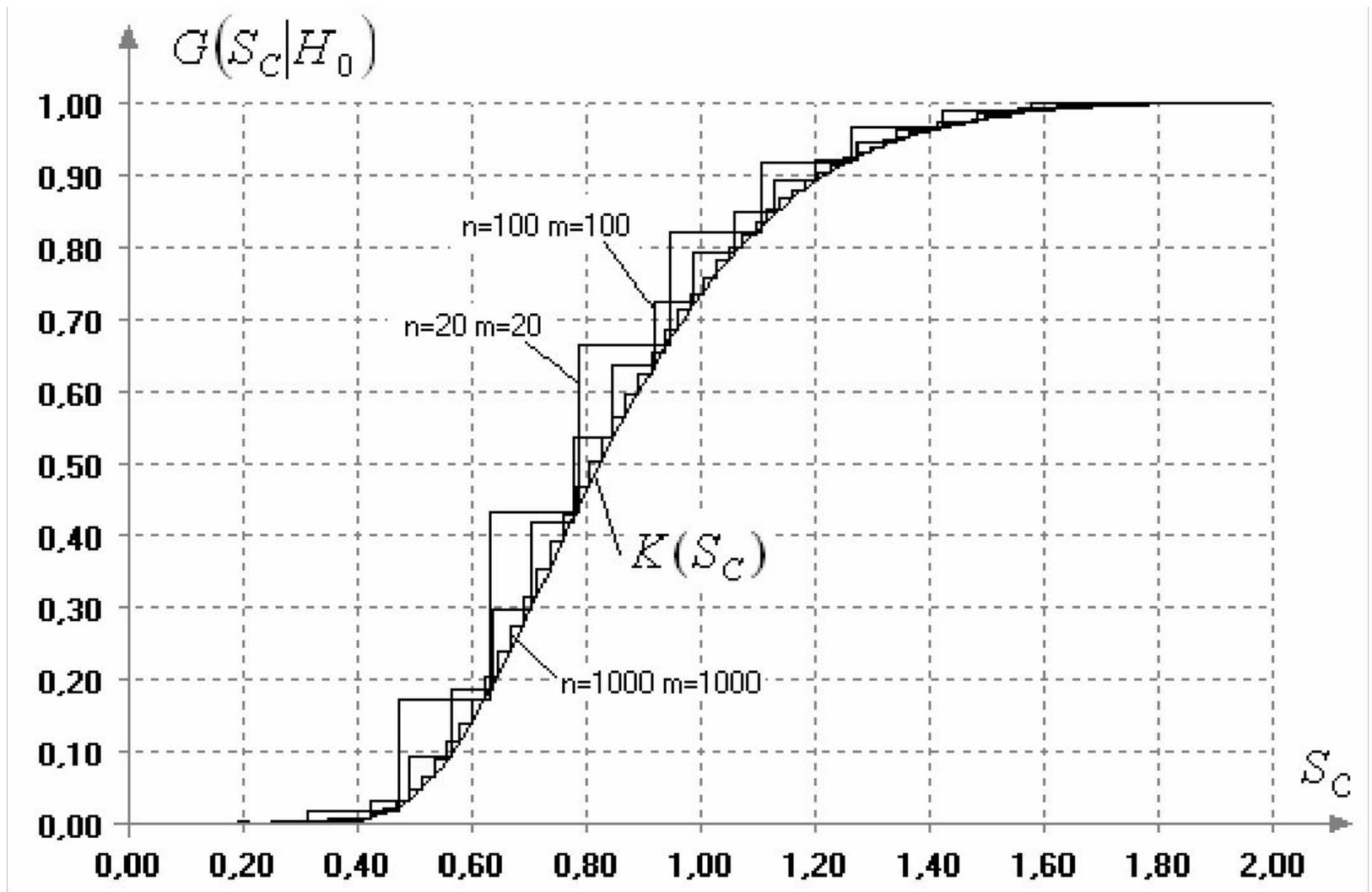
$F(x)$ и $G(x)$ - непрерывны. Статистика критерия Смирнова

$$S_C = \sqrt{\frac{mn}{m+n}} D_{m,n},$$
$$D_{m,n} = \sup_x |G_m(x) - F_n(x)|.$$

При практическом использовании критерия значение статистики $D_{m,n}$ рекомендуется вычислять в соответствии с соотношениями

$$D_{m,n}^+ = \max_{1 \leq r \leq m} \left[\frac{r}{m} - F_n(x_r) \right] = \max_{1 \leq s \leq n} \left[G_m(y_s) - \frac{s-1}{n} \right],$$
$$D_{m,n}^- = \max_{1 \leq r \leq m} \left[F_n(x_r) - \frac{r-1}{m} \right] = \max_{1 \leq s \leq n} \left[\frac{s}{n} - G_m(y_s) \right],$$
$$D_{m,n} = \max(D_{m,n}^+, D_{m,n}^-).$$

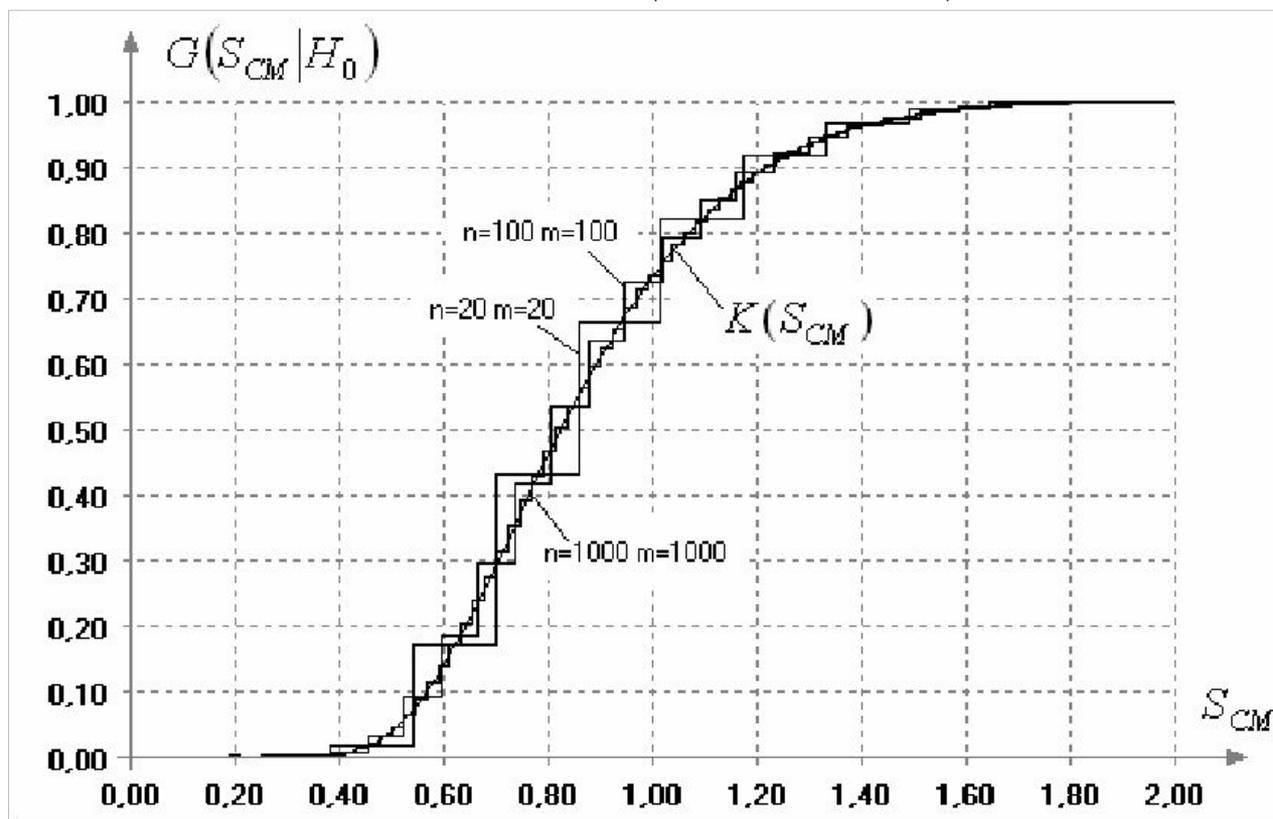
Статистика в пределе подчиняется распределению Колмогорова $K(s)$. Однако при ограниченных значениях m и n случайные величины $D_{m,n}^+$ и $D_{m,n}^-$ являются дискретными, и множество их возможных значений представляет собой решетку с шагом $1/k$, где k наименьшее общее кратное m и n .



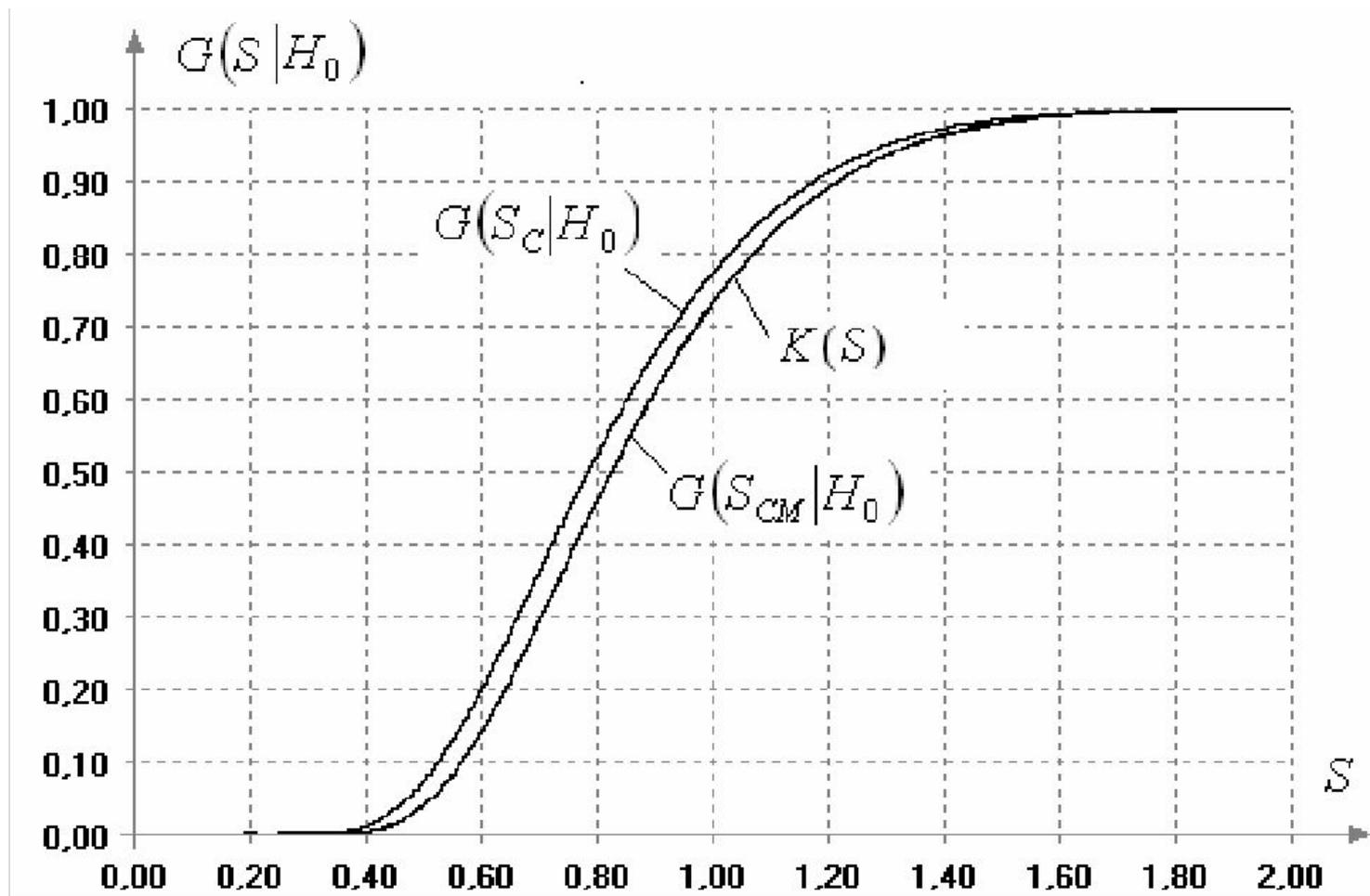
Распределения статистики критерия Смирнова при справедливости H_0 в зависимости от m и n

Распределение статистики Смирнова заметно сдвинуто влево от $K(s)$. В этой связи **предложена простая модификация статистики**,

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



Распределения модифицированной статистики при справедливости H_0 в зависимости от m и n



Распределения статистик Смирнова и модифицированной при справедливости H_0 , и взаимно простых $m = 61$ и $n = 53$

Исследование мощности критериев однородности

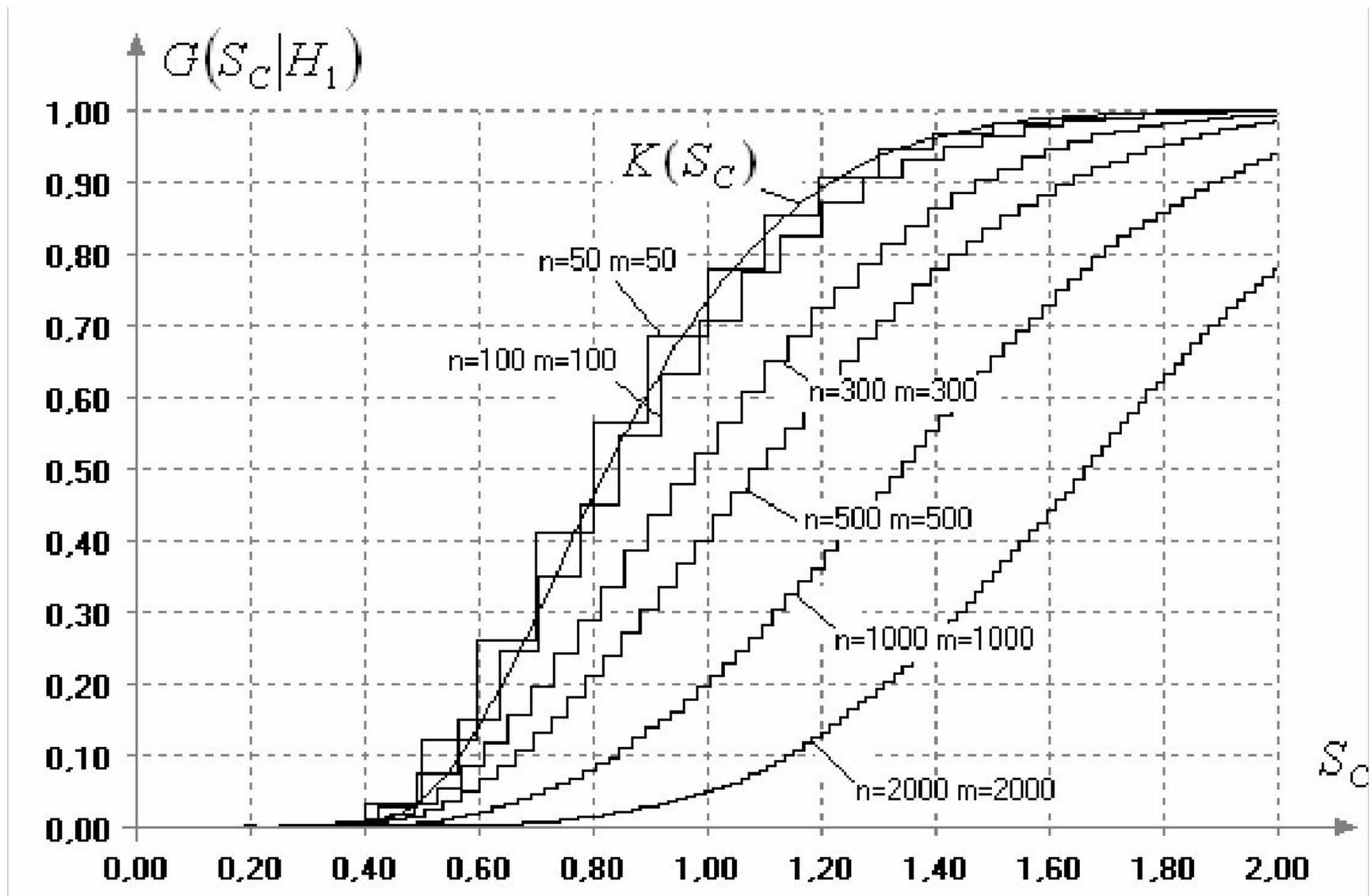
Для определенности гипотезе H_0 соответствовала принадлежность выборок одному и тому же стандартному нормальному закону распределения с плотностью

$$f(x) = \frac{1}{\theta_1 \sqrt{2\pi}} \exp\left\{-\frac{(x - \theta_0)^2}{2\theta_1^2}\right\}$$

с параметрами сдвига $\theta_0 = 0$ и масштаба $\theta_1 = 1$. При всех альтернативах первая выборка всегда соответствовала стандартному нормальному закону, а вторая – некоторому другому. В частности, в случае гипотезы H_1 вторая выборка соответствовала нормальному закону с параметром сдвига $\theta_0 = 0.1$ и параметром масштаба $\theta_1 = 1$. В случае гипотезы H_2 – нормальному закону с параметрами $\theta_0 = 0.5$ и $\theta_1 = 1$. В случае гипотезы H_3 – нормальному закону с параметрами $\theta_0 = 0$ и $\theta_1 = 1.1$. В случае гипотезы H_4 – нормальному закону с параметрами $\theta_0 = 0$ и $\theta_1 = 1.5$. В случае гипотезы H_5 – вторая выборка соответствовала логистическому закону с плотностью

$$f(x) = \frac{\pi}{\theta_1 \sqrt{3}} \exp\left\{-\frac{\pi(x - \theta_0)}{\theta_1 \sqrt{3}}\right\} / \left[1 + \exp\left\{-\frac{\pi(x - \theta_0)}{\theta_1 \sqrt{3}}\right\}\right]^2$$

и параметрами $\theta_0 = 0$ и $\theta_1 = 1$.



Критерий однородности Лемана-Розенблатта

Критерий однородности представляет собой критерий типа ω^2 . Статистика критерия имеет вид

$$T = \frac{mn}{m+n} \int_{-\infty}^{\infty} [G_m(x) - F_n(x)]^2 dH_{m+n}(x),$$

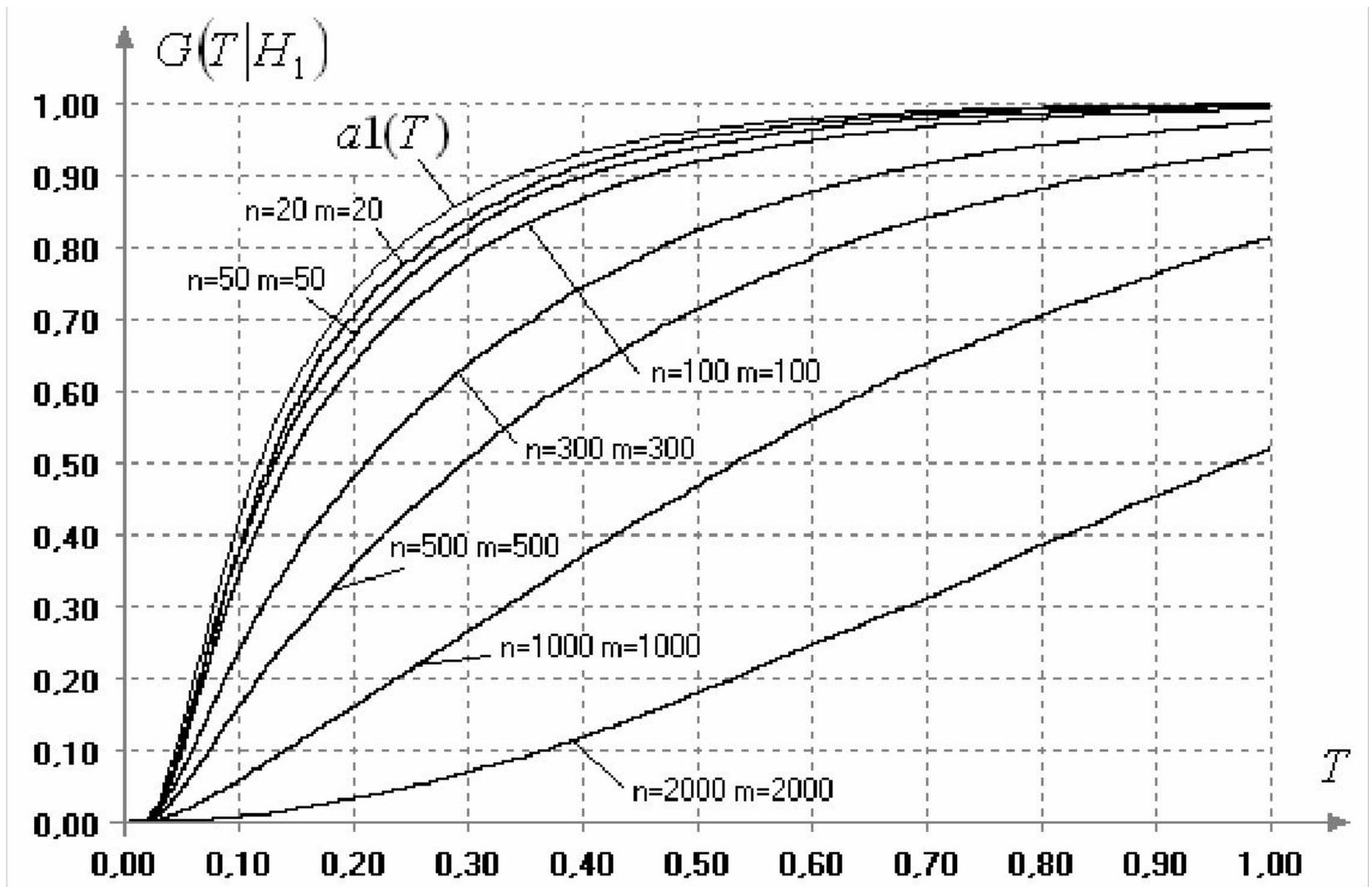
где $H_{m+n}(x) = \frac{m}{m+n} G_m(x) + \frac{n}{m+n} F_n(x)$ – эмпирическая функция распределения, построенная по вариационному ряду объединения двух выборок.

Статистика T используется в форме

$$T = \frac{1}{mn(m+n)} \left[n \sum_{i=1}^n (r_i - i)^2 + m \sum_{j=1}^m (s_j - j)^2 \right] - \frac{mn-1}{6(m+n)}, \quad (4)$$

где r_i – порядковый номер (ранг) y_i , s_j – порядковый номер (ранг) x_j в объединенном вариационном ряду.

Статистика (4) в пределе распределена как $al(t)$ и, в отличие от статистики критерия Смирнова, быстро сходится к предельному.



Распределения статистики Лемана-Розенблатта при справедливости H_1

07/30/2023 «ИТ анализа данных и моделирования ...» 128

Выводы

Сравнивая мощность критериев относительно различных альтернатив с учетом действительных уровней значимости критерия Смирнова, можно отметить, что, как правило, мощность критерия Лемана-Розенблатта заметно выше мощности критерия Смирнова.

Однако относительно очень близких альтернатив несколько выше оказывается мощность критерия Смирнова. Последнее интуитивно понятно, если вспомнить, что в критерии Смирнова мера отклонения линейная, а в критерии Лемана-Розенблатта – квадратичная.

Заключение

Компьютерные технологии анализа данных и исследования вероятностных и статистических закономерностей представляют собой мощный инструмент для развития и совершенствования аппарата прикладной математической статистики.