УДК 332.63

О влиянии ошибок округления на обеспечение корректности выводов в задачах проверки статистических гипотез

Б. Ю. Лемешко
ФГБОУ ВО «НГТУ»
г. Новосибирск, Россия
E-mail: Lemeshko@ami.nstu.ru
https://ami.nstu.ru/~headrd/

I. О состоянии использования статистических методов

Рассказывая на СУДОМЕТРИКЕ–2018 о проблемах, связанных с применением критериев проверки статистических гипотез в задачах метрологического обеспечения, анализа результатов измерений, контроля и испытаний, я отмечал следующее: существующий аппарат прикладной математической статистики крайне ограничено и/или неэффективно используется в задачах метрологии.

Из обширного множества критериев, предназначенных для проверки различных статистических гипотез, в задачах метрологического обеспечения, анализа результатов измерений, контроля и испытаний используется лишь очень ограниченный круг.

Ряд очень перспективных критериев, которые «буквально напрашиваются» для применения в приложениях (как, например, критерии однородности законов или однородности дисперсий в задачах лабораторных сличений), практически там не используются.

Зачастую критерии используются в условиях нарушения стандартных предположений, обуславливающих возможность их применения, что приводит к некорректности статистических выводов. К некорректности выводов может приводить и наличие ошибок округления, какие бы «сверхточные» измерения не проводились. За прошедшие 4 года существенных изменений с состоянием в лучшую сторону не произошло.

Объективными и субъективными причинами такого состояния являются следующие:

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

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

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

4. Отсутствует доступная информация о том, что происходит с критериями (с распределениями их статистик) в условиях нарушения стандартных предположений, в том числе под влиянием ошибок округления.

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

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

7. Негативное влияние «Руководства по выражению неопределенности измерения», тормозящее использование новых результатов в области прикладной математической статистики.

8. Человеческий фактор, не всегда способствующий лучшим решениям, а в задачах метрологии – использованию статистических методов. Лица, принимающие решения об использовании, не всегда обладают нужными компетенциями в оценке возможностей статистических методов.

II. Общие сведения о проверке статистических гипотез

С каждым из используемых для проверки гипотезы H_0 критериев связана статистика S, которая представляет собой некоторую меру для измерения вероятности соответствия (несоответствия) анализируемых выборок проверяемой гипотезе H_0. При справедливости проверяемой гипотезы H_0 статистика S подчиняется некоторому распределению $G(S|H_0)$.

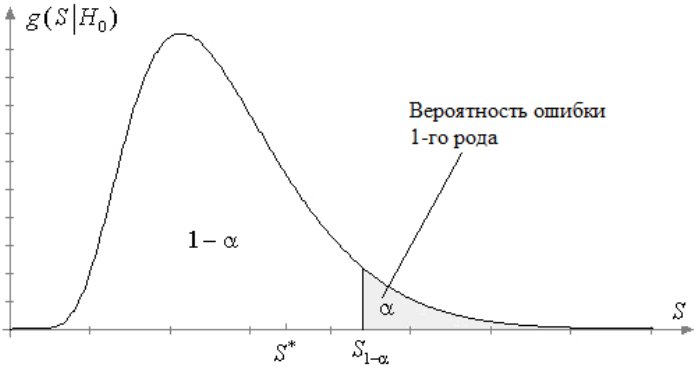


Рис. 1 – Плотность распределения статистики при справедливости гипотезы H_0 и критическое значение для правостороннего критерия

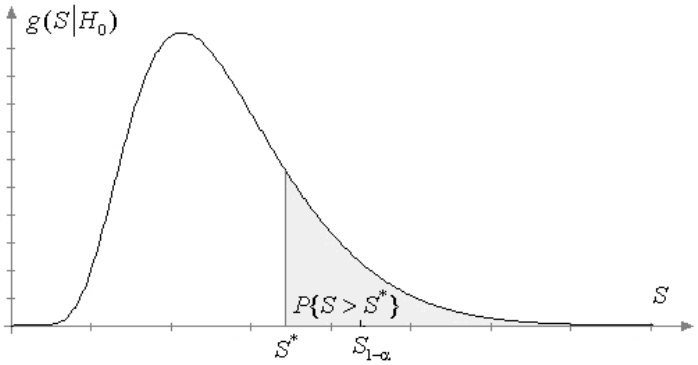


Рис. 2 – Плотность распределения статистики при справедливости гипотезы H_0 и достигнутый уровень значимости

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

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

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

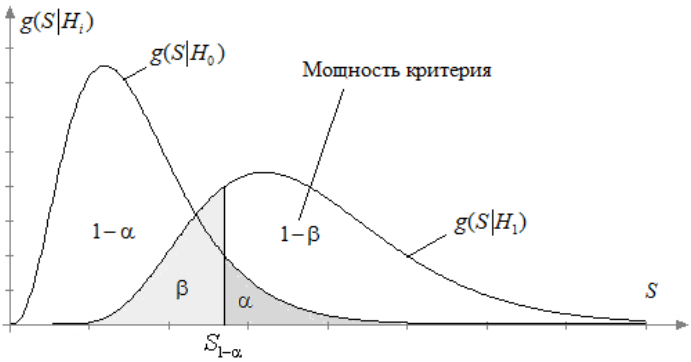


Рис. 3 – Плотности распределения статистик при справедливости соответственно гипотез H_0 и H_1 в случае правостороннего критерия

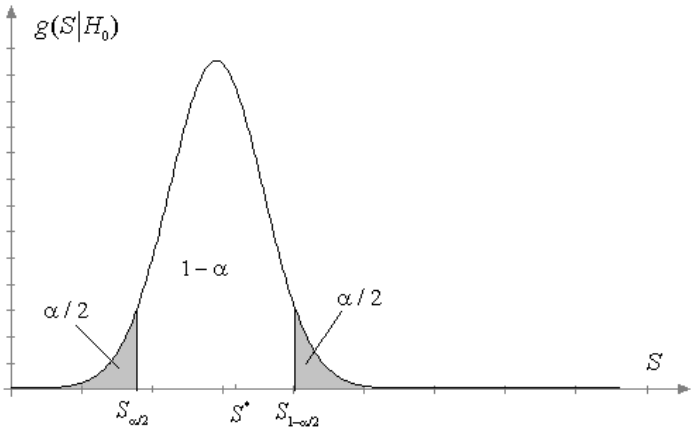


Рис. 4 – Плотность распределения статистики при справедливости гипотезы H_0 и критические значения для двустороннего критерия

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

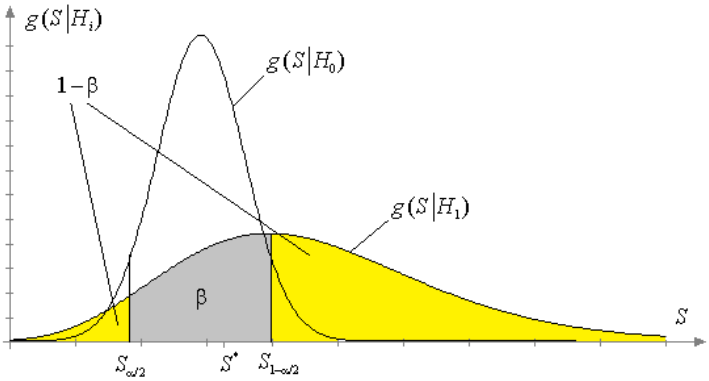


Рис. 5 – Плотности распределения статистик при справедливости соответственно гипотез H_0 и H_1 в случае двустороннего критерия

Предпочтительней использовать более мощные критерии.

III. Непараметрические критерии согласия

На «СУДОМЕТРИКЕ–2018» говорилось о проблемах, связанных с некорректным применением непараметрических критериев согласия, вследствие которых выводы о результатах проверки гипотез могут оказываться неверными. Речь шла о применении критериев согласия Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Купера, Ватсона, критериев Жанга.

При проверке согласия различают простые и сложные гипотезы.

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

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

Большинство ошибок применения совершается при проверке сложных, когда оценивание парметров и проверка гипотезы опираются на одну и ту же выборку.

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

Все непараметрические критерии первоначально строились для проверки простых гипотез. В этой ситуации все непараметрические критерии согласия «свободны от распределения» $F(x,\theta)$, относительно которого проверяется гипотеза.

IV. Непараметрические критерии согласия

Критерий Колмогорова

Критерий Колмогорова опирается на статистику:

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

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

В критерии Колмогорова рекомендуется использовать статистику с поправкой Большева в форме

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

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

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

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

Критерий Крамера-Мизеса-Смирнова

Статистика критерия Крамера–Мизеса–Смирнова имеет вид:

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

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

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

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

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

Критерий Андерсона-Дарлинга

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

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

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

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

Критерий Купера

В критерии Купера статистика V_n которого определяется соотношением

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

и используется в виде

$$V_n = D_n^+ + D_n^- \qquad (4)$$

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

Статистика $\sqrt{n}V_n$ подчиняется распределению:

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

В критерии используется либо модификация статистики:

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

(0.20)

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

(0.21)

где идея использования поправки вытекает из выражения для статистики критерия согласия Смирнова.

Критерий Ватсона

Статистика критерия Ватсона имеет вид

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

и используется в следующей удобной для расчетов форме

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

(5)

Предельное распределение $G\left(U_n^2 \middle| H_0\right)$ статистики U_n^2 приведено в виде:

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

Критерии Жанга

В диссертации Жанга и в последующих его работах предложены непараметрические критерии согла¬сия, статистики которых имеют вид:

$$Z_K = \max_{1\le i\le n}\left(\left(i-\frac{1}{2}\right)\log\left\{\frac{i-\frac{1}{2}}{nF(x_i,\theta)}\right\} + \left(n-i+\frac{1}{2}\right)\log\left[\frac{n-i+\frac{1}{2}}{n\{1-F(x_i,\theta)\}}\right]\right) \quad (6)$$

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log\{F(x_i,\theta)\}}{n-i+\frac{1}{2}} + \frac{\log\{1-F(x_i,\theta)\}}{i-\frac{1}{2}}\right] \quad (7)$$

$$Z_C = \sum_{i=1}^{n}\left[\log\left\{\frac{\left[F(x_i,\theta)\right]^{-1} - 1}{(n-\frac{1}{2})/(i-\frac{3}{4}) - 1}\right\}\right]^2 \quad (8)$$

Использование критериев со статистиками (6)–(8) осложняет сильная зависимость распределений статистик от объема выборки *n*.

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

V. Непараметрические критерии согласия при проверке сложных гипотез

Факторы, влияющие на распределения статистик критериев при проверке сложных гипотез

При проверке сложных гипотез, когда по той же самой выборке оценивают параметры наблюдаемого закона распределения вероятностей, все рассматриваемые непараметрические критерии согласия теряют свойство «свободы от распределения».

Более того, распределения статистик непараметрических критериев согласия при проверке сложных гипотез зависят от целого ряда факторов, определяющих «сложность» гипотезы.

На законы распределения $G(S|H_0)$ статистик критериев влияют следующие факторы:

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

– тип оцениваемого параметра и число оцениваемых параметров;

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

– используемый метод оценивания параметров.

Игнорирование того, что проверяют сложную гипотезу, и того, что сложные гипотезы могут быть различными, приводит к некорректному применению непараметриче¬ских 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.

Подробно о решении проблем применения критериев согласия при проверке сложных гипотез (как меняются распределения статистик, о построенных моделях этих распределений, о технологиях построения распределений статистик в интерактивном режиме) можно смотреть в руководстве:

Лемешко Б. Ю. Непараметрические критерии согласия. Руководство по применению: Монография. – М.: ИНФРА-М, 2014. – 163 с. DOI: 10.12737/11873(доп. тиражи 2018, 2019)

Предшественниками этого руководства были разработанные нами:

– Р 50.1.037–2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Не-

параметрические критерии. – М.: Изд-во стандартов. 2002. – 64 с.

– Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть II. Непараметрические критерии. – Новосибирск: Изд-во НГТУ, 1999. – 85 с.

VI. О СПЕЦИАЛЬНЫХ КРИТЕРИЯХ ПРОВЕРКИ РАЗЛИЧНЫХ СЛОЖНЫХ ГИПОТЕЗ

Для проверки гипотез о принадлежности выборок наиболее востребованным в приложениях параметрическим моделям законов распределения, таким как нормальный закон, равномерный закон, экспоненциальный закон, предложено множество специальных критериев, ориентированных на проверку сложных гипотез только относительно этого конкретного закона.

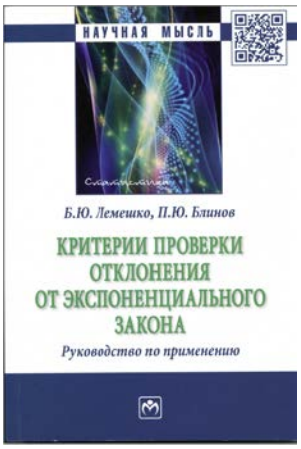
Имеется множество критериев, ориентированных на проверку гипотез об однороднсти законов, об однородности средних (о равенстве математических ожиданий), об однородности дисперсий. Об этих критериях говорилось на «СУДОМЕТРИКЕ-2018».

Достаточно представительное множество критериев предложено для проверки гипотез об отсутствии тренда в наблюдаемых рядах измерений.

Все эти критерии имеют свои достоинства и недостатки, обладают различной мощностью относительно разных конкурирующих гипотез. Эти индивидуальные особенности учтены в соответствующих руководствах, разработанных по результатам исследования критериев, и эти знания следует учитывать при применении критериев.

Лемешко Б. Ю., Блинов П. Ю. Критерии проверки отклонения распределения от равномерного закона. Руководство по применению: Монография. – М.: ИНФРА-М, 2015. – 183 с. – (Научная мысль). DOI: 10.12737/11304 (доп. тиражи 2016, 2020, 2022)

Показаны достоинства и недостатки 17 специальных критериев, ориентированных только на проверку равномерности (в ISW их около 25), а также применение для проверки равномерности непараметрических критериев согласия и критериев типа хи-квадрат. Приводятся результаты сравнительного анализа критериев.



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

В руководстве рассматриваются вопросы применения статистических критериев, ориентированных на проверку гипотезы о принадлежности анализируемой выборки экспоненциальному (показательному) закону распределения. В совокупности рассмотрено около 50 критериев. Указываются недостатки и преимущества различных критериев.

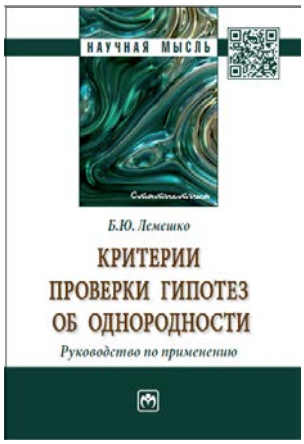
Приводятся оценки мощности критериев и результаты сравнительного анализа критериев, а также таблицы, содержащие процентные точки и модели распределений статистик, необходимые для применения критериев.

Рассматривается влияние ошибок округления на распределения статистик и на результаты статистических выводов.

Лемешко Б.Ю., Веретельникова И.В. Критерии про¬верки гипотез о случайности и отсутствии тренда. Руководство по применению: Монография. – Москва : ИНФРА-М. 2021. – 221 с. – (Научная мысль). DOI 10.12737/1587437

Рассматриваются вопросы применения статистических критериев, ориентированных на проверку гипотез об от¬сутствии тренда в анализируемых выборках. Неоткло¬нение такой гипотезы даёт основание рассматривать анализируемые данные как выборки независимых одинаково распределенных случайных величин.

Рассматривается множество специальных критериев, ориентированных на проверку такого рода гипотез, а также множество критериев однородности законов, однородности средних и однородности дисперсий, которые также могут применяться в указанных целях. Подчеркиваются недостатки и преимущества различных критериев, рассматривается применение критериев в условиях нарушения стандартных предположений (в том числе в условиях влияния на выводы ошибок округления). Приводятся оценки мощности критериев, что позволяет ориентироваться при выборе наиболее предпочтительных критериев.



Лемешко Б. Ю. Критерии проверки гипотез об однородности. Руководство по применению: Моно-графия. – 2-е изд., перераб. и доп. – М.: ИНФРА-М, 2021. – 248 с. (Научная мысль) DOI: 10.12737/986695 (первое издание – 2017, 2018)

Рассмотрено применение двухвыборочных критериев однородности законов распределения Смирнова, Леманга–Розенблатта, Андрсона–Дарлинга–Петита, k-выборочного критерия Андерсона–Дарлинга. Для последнего приводятся модели предельных распределений статистики для различных *k*. Рассмотрено применение 3-х критериев однородности Жанга и предложены новые *k*-выборочные критерии на основе двухвыборочных с построенными моделями предельных распределений статистик.

Приведены результаты сравнительного анализа 9 параметрических и непараметрических критериев однородности средних, результаты сравнительного анализа более 20 параметрических и непараметрических критериев однородности дисперсий. В ISW реализуется корректное применение параметрических критериев в условиях нарушения стандартного предположения о нормальности.

Лемешко Б. Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению : Монография. – 2-е изд., перераб. и доп. – Москва : ИНФРА-М, 2023. – 353 с. – (Научная мысль). DOI 10.12737/1896110

(первое издание –2015, 2018,2019, 2021)

Рассматриваются вопросы применения статистических критериев, ориентированных на проверку гипотезы о принадлежности анализируемых данных нормальному закону распределения вероятностей. Рассматриваются и сравниваются специальные критерии, непараметричес¬кие критерии согласия и критерии типа χ^2. Указываются недостатки и преимущества различных критериев. Приводятся таблицы, процентных точек и модели распределений статистик, необходимые для корректного применения критериев.

По сравнению с первым изданием расширено мно¬жество рассмотренных критериев. Критерии проранжиро¬ваны (в совокупности 50 критериев) по мощности относительно близких конкурирующих гипотез. Ранжирование облегчает выбор предпочтительных критериев.

Показано, что вследствие наличия ошибок округления свойства критериев в приложениях могут существенно изменяться, и это необходимо учитывать при формировании статистических выводов.

VII. Ранжирование множества критериев нормальности по мощности

В качестве конкурирующих гипотез при исследовании мощности критериев рассматривалась принадлежность следующим законам:

H_1 – обобщённому нормальному закону с плотностью $f(x)=\dfrac{\theta_2}{2\theta_1\Gamma(1/\theta_2)}\exp\left\{-\left(\dfrac{|x-\theta_0|}{\theta_1}\right)^{\theta_2}\right\}$ при $\theta_2=4$;

H_2 – Лапласа с плотностью $f(x)=\dfrac{1}{2\theta_1}\exp\{-|x-\theta_0|/\theta_1\}$;

H_3 – логистическому с плотностью

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

(очень близкому к нормальному закону);

H_4 – минимального значения с плотностью $f(x)=\dfrac{1}{\theta_1}\exp\left\{\dfrac{x-\theta_0}{\theta_1}-\exp\left(\dfrac{x-\theta_0}{\theta_1}\right)\right\}$, при $\theta_0=0.38$, $\theta_1=0.8$;

H_5 –максимального значения с плотностью $f(x)=\dfrac{1}{\theta_1}\exp\left\{-\dfrac{x-\theta_0}{\theta_1}-\exp\left(-\dfrac{x-\theta_0}{\theta_1}\right)\right\}$ и $\theta_0=-0.38$, $\theta_1=0.8$.

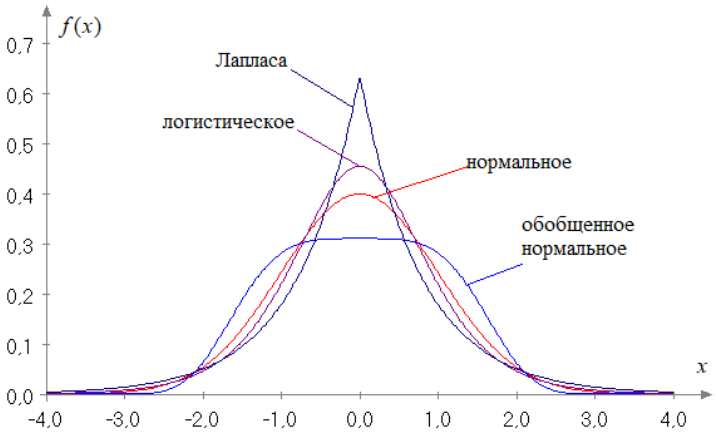


Рис. 6 – Плотности законов распределения, соответствующие рассматриваемым гипотезам H_0, H_1, H_2, H_3

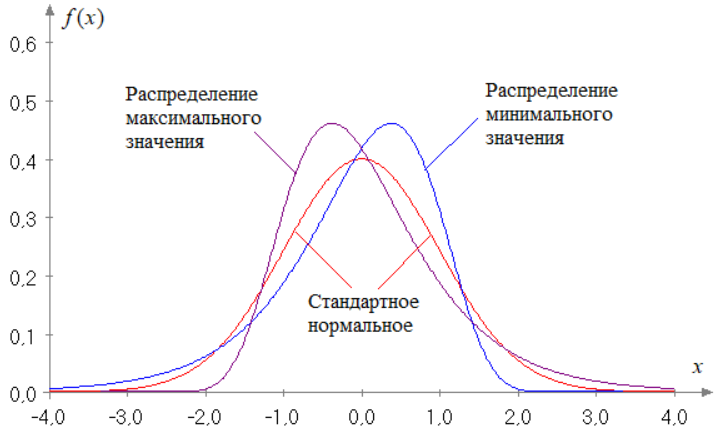


Рис. 7 – Плотности распределений, соответствующие гипотезам H_0, H_4, H_5

По мощности, проявленной относительно каждой конкурирующей гипотезы, множество критериев, применяемых для проверки нормальности, **упорядочивается различным образом**.

В такой ситуации для ранжирования критериев можно поступить следующим образом.

В убывающем ряду оценок мощности относительно рассматриваемой гипотезы находится ранг этого критерия. В качестве основы для рейтинга можно взять сумму рангов соответствующего критерия: сумму мест “занятых” критерием по величине мощности относительно H_i, $i=\overline{1,4}$. (относительно H_4 и H_5 критерии, как правило, демонстрируют одинаковую мощность). При упорядочивании критериев по возрастанию суммы, порядковые номера (ранги) этих сумм указывают рейтинг соответствующего критерия. При одинаковой сумме мест будут одинаковыми и рейтинги.

На результаты анализа и ранжирование (оценки рейтинга), предложенное в руководстве по критериям нормальности, можно ориентироваться в приложениях, отдавая предпочтение применению того или иного критерия для проверки гипотезы о принадлежности анализируемой выборки нормальному закону.

Результаты ранжирования приведены в таблице 1. Вид статистик рассмотренных критериев, положительные и отрицательные, связанные с их свойствами, отражены в руководстве.

ТАБЛИЦА 1 - РАНЖИРОВАНИЕ КРИТЕРИЕВ НОРМАЛЬНОСТИ

№	Критерий	Мощность относительно H_i				Место относительно H_i				Σ мест	R
		H_1	H_2	H_3	H_4	H_1	H_2	H_3	H_4		
1	Десгань–Мишо X_{APD}	0.304	0.705	0.318	0.744	14	7.5	13	16.5	51	1
2	Заманзаде–Аргами TZ_1	0.293	0.695	0.330	0.717	16.5	11	12	20	59.5	2
3	Десгань–Мишо X_{EPD}	0.451	0.705	0.299	0.356	1	7.5	15	43	66.5	3
4	Бонтемпса–Меддахи BM_{3_6}	0.216	0.689	0.342	0.744	31	15	8	16.5	70.5	4
5	Д`Агостино E_p	0.164	0.671	0.347	0.756	36	18	6	12.5	72.5	5
6	Шапиро-Франциа	0.139	0.691	0.333	0.760	39.5	13.5	10.5	9.5	73	7.5
7	Вайсберга-Бингема	0.139	0.691	0.333	0.760	39.5	13.5	10.5	9.5	73	7.5
8	Хегази–Грина T_2	0.061	0.723	0.359	0.735	47	5	3	18	73	7.5
9	Заманзаде–Аргами TZ_2	0.086	0.752	0.366	0.696	43	3	2	25	73	7.5
10	Ройстона	0.280	0.616	0.273	0.776	21	29	18.5	5.5	74	10
11	Филлибена	0.119	0.699	0.340	0.754	42	10	9	14	75	11
12	Чена–Шапиро	0.327	0.576	0.251	0.776	12	32	26	5.5	75.5	12
13	Гелы–Гаствирта	0.004	0.753	0.378	0.699	49	2	1	24	76	13
14	Гаствирта	0.384	0.731	0.273	0.354	11	4	18.5	44	77.5	14
15	Жанга Z_C	0.300	0.548	0.270	0.772	15	34	23	7	79	15
16	Гири	0.394	0.722	0.272	0.271	8.5	6.5	21	46	82	16
17	Харке–Бера	0.071	0.654	0.349	0.756	45	21	4.5	12.5	83	17.5
18	Бонетта–Сейер	0.394	0.722	0.272	0.271	8.5	6.5	21	47	83	17.5
19	Бонтемпса–Меддахи BM_{3_4}	0.066	0.653	0.349	0.757	46	22	4.5	11	83.5	19
20	Жанга Z_A	0.239	0.578	0.272	0.787	29	31	21	3	84	20
21	Шпигельхальтера	0.211	0.790	0.343	0.320	32	1	7	45	85	21
22	Оя $\tilde{T}_{12}$	0.406	0.694	0.228	0.455	6	12	32	38.5	88.5	22
23	Хегази–Грина T_1	0.218	0.674	0.267	0.724	30	17	24	19	90	23

Продолжение Таблицы 1.

24	Эппса–Пулли	0.275	0.623	0.249	0.751	23.5	26.5	27.5	15	92.5	24
25	Мартинеса–Иглевича	0.252	0.683	0.301	0.503	27	16	14	36	93	25
26	Шапиро–Уилка	0.389	0.502	0.203	0.765	10	38	39	8	95	26
27	Али–Чорго–Ревеса	0.281	0.643	0.240	0.702	20	23	30	22.5	95.5	27.5
28	Андерсона–Дарлинга	0.287	0.630	0.230	0.702	18	24	31	22.5	95.5	27.5
29	Десгань–Мишо R_n	0.275	0.670	0.287	0.398	23.5	19	17	42	101.5	29
30	Фросини	0.285	0.623	0.212	0.665	19	26.5	35	27	107.5	30
31	Ван Эса	0.150	0.667	0.298	0.551	38	20	16	34	108	31
32	Ватсона	0.293	0.626	0.204	0.583	16.5	25	38	32	111.5	32
33	Д`Агостино Z_2	0.428	0.489	0.241	0.402	5	40	29	41	115	33
34	Крамера–Мизеса–Смирнова	0.269	0.621	0.209	0.652	25	28	37	28	118	34
35	Жанга Z_K	0.186	0.569	0.249	0.679	34	33	27.5	26	120.5	35
36	Дэвида–Хартли–Пирсона	0.400	0.499	0.255	0.172	7	39	25	50	121	36
37	Васичека	0.434	0.397	0.137	0.611	3.5	43.5	46.5	29.5	123	37.5
38	Эбрахими	0.434	0.396	0.136	0.611	3.5	43.5	46.5	29.5	123	37.5
39	Локка–Сперриера T_{2n}	0.000	0.504	0.222	0.786	50	37	33	4	124	39
40	Оя $\tilde{T}_{34}$	0.262	0.518	0.172	0.715	26	36	43	21	126	40.5
41	Корреа	0.441	0.382	0.132	0.586	2	45	48	31	126	40.5
42	Купера	0.279	0.589	0.192	0.526	22	30	40	35	127	42
43	Лина–Мудхолкара	0.056	0.332	0.216	0.827	48	47	34	1	130	43
44	Хи-квадрат Пирсона	0.311	0.423	0.155	0.455	13	42	45	38.5	138.5	44
45	Локка–Сперриера T_{1n}	0.074	0.255	0.168	0.808	44	49	44	2	139	45
46	Колмогорова	0.208	0.540	0.181	0.564	33	35	42	33	143	46
47	Никулина–Рао–Робсона	0.240	0.473	0.188	0.431	28	41	41	40	150	47
48	Жанга	0.179	0.355	0.212	0.476	35	46	36	37	154	48
49	Чена	0.139	0.292	0.130	0.262	41	48	49	48	186	49.5
50	Брис–Хьюберт–Стройфа	0.162	0.147	0.100	0.247	37	50	50	49	186	49.5

VIII. Ранжирование множества критериев нормальности по мощности

В 2002 г. был введен в действие ГОСТ Р ИСО 5479–2002 , посвященный проверке отклонения распределения вероятностей от нормального закона. В марте 2020 года ГОСТ был переиздан.

По существу, стандарт предусматривает применение только 2-х критериев нормальности: Шапиро–Уилка и Эппса–Пулли (не свободных от недостатков). Не предусматривается ни использование непараметрических критериев согласия, ни критериев типа χ^2 . То, что связано с проверкой симметричности и значением эксцесса, не является проверкой нормальности. Модификация критерия Шапиро–Уилка для анализа совокупности малых выборок грешит тем, что распределение её статистики плохо описывается заявленным стандартным нормальным законом.

Если обратиться к результатам ранжирования множества критериев нормальности в нашем руководстве, то можно увидеть, что по предпочтительности критерий Эппса–Пулли находится только на 24-м месте, а Шапиро–Уилка – на 26-м в ряду из 50 критериев. При малых объёмах выборок из-за смещённости эти критерии не способны отличать от нормального закона некоторые близкие конкурирующие законы (в частности, гипотезу H_1).

«Оправдывает» существование этого стандарта лишь тот факт, что его текст представляет собой аутентичный перевод ISO 5479: 1997.

IX. О применении критериев согласия для анализа больших выборок

В приложениях всё чаще приходится сталкиваться с необходимостью анализа гигантских объёмов накапливаемых данных (Big Data). Естественно, для анализа применяют различные статистические методы, в том числе критерии проверки статистических гипотез и, в частности, критерии согласия. И здесь обнаруживается, что в случае “больших выборок” критерии “всегда отклоняют” проверяемую гипотезу.

Одновременно можно слышать заявления, что никаких проблем с применением статистических критериев для анализа больших выборок нет.

Проблемы, конечно же, есть, но они решаемы. Объёмы выборок, с которыми имеют дело в Big Data, принадлежащие некоторому непрерывному закону распределения, бывают практически неограничены. Но сами данные в них бывают представлены с ограниченной точностью (округлены с некоторым Δ). По сути, “нарушается предположение” о том, что наблюдается непрерывная случайная величина.

Статистики различных критериев согласия соответствующим образом измеряют отклонение эмпирического распределения $F_n(x)$ от теоретического $F(x)$. При наличии ошибок округления Δ с ростом n в выборках появляются повторяющиеся наблюдения. И, начиная с некоторого n, ступеньки $F_n(x)$ и расстояния $\max\left|F_n(x)-F(x)\right|$ перестают уменьшаться. А значения статистик критериев растут. Например, в критерии Колмогорова статистика имеет вид $S_K=\sqrt{n}\max\left|F_n(x)-F(x)\right|$.

Чтобы использовать классические результаты, рекомендуется применять критерий не ко всему массиву Big Data, а извлекать для анализа из этого массива выборки объёмов $n<n_{\max}$, при которых реальное распределение статистики критерия $G(S_n(\Delta)|H_0)$ ещё не отличается от асимптотического $G(S|H_0)$.

X. О ВЛИЯНИИ ОШИБОК ОКРУГЛЕНИЯ НА РАСПРЕДЕЛЕНИЯ СТАТИСТИК КРИТЕРИЕВ ПРИ АНАЛИЗЕ “ВЫСОКОТОЧНЫХ” ИЗМЕРЕНИЙ, ПРОВОДИМЫХ НА ПРЕДЕЛЕ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ ИЗМЕРИТЕЛЬНОЙ СИСТЕМЫ

Имеются в виду ситуации, когда ошибки округлении Δ соизмеримы с ошибками измерения ($\Delta\approx\sigma$, где σ среднее квадратичное отклонение ошибки измерения). При этом не важно, насколько уникальна измерительная система, и какие сверхточные измерения она обеспечивает. В выборках, соответствующих таким измерениям, как правило, присутствуют повторяющиеся значения.

Следует отметить, что есть целый ряд критериев (нормальности, равномерности, экспоненциальности), например, совокупность критериев, в статистиках которых использованы оценки энтропии, для которых присутствие в выборке даже пары одинаковых значений является основанием для отклонения проверяемой гипотезы (в выборке непрерывной случайной величины не могут присутствовать одинаковые значения). Далее речь будет идти не о таких критериях.

Ниже покажем, как под влиянием ошибок округления меняются распределения статистик критериев согласия, распределения статистик специальных критериев нормальности, экспоненциальности и других критериев.

Все приведенные ниже примеры говорят об одном: при соизмеримости Δ и σ нельзя пренебрегать фактом изменения распределений статистик критериев, так как в противном случае, как правило, возрастает вероятность ошибки 1-го рода α (отклонения справедливой гипотезы H_0), или реже – возрастает вероятность ошибки 2-го рода β (в случае двусторонних критериев).

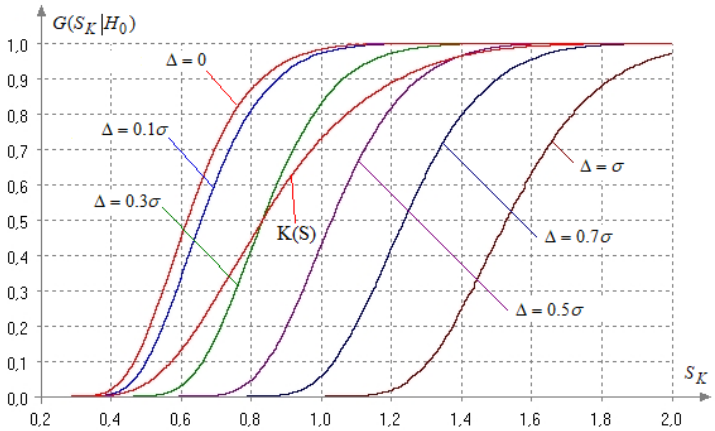


Рис. 8 – Зависимость распределения статистики критерия Колмогорова от Δ при справедливости сложной гипотезы H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $n=50$

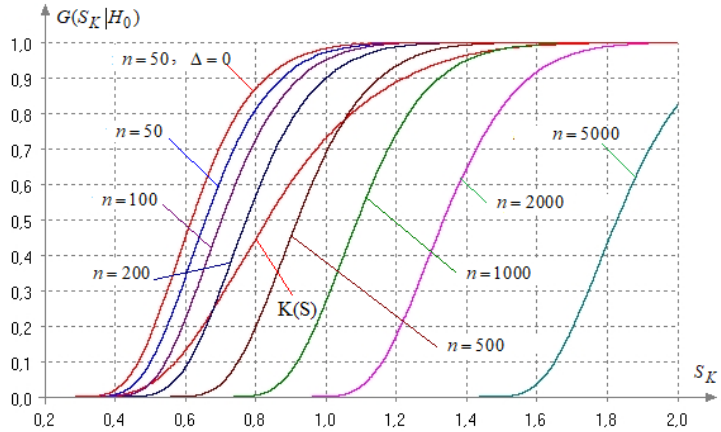


Рис. 9 – Зависимость распределения статистики критерия Колмогорова от n при справедливости сложной гипотезы H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $\Delta=0.1\sigma$

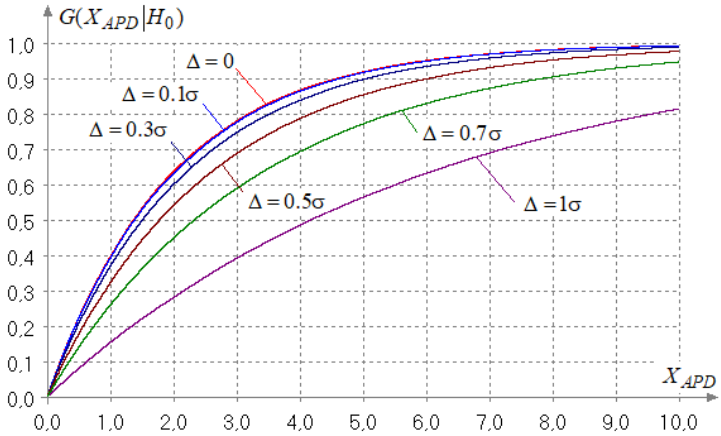


Рис. 9 – Зависимость от Δ распределения статистики X_{APD} критерия нормальности Десгань–Мишо при справедливости H_0 и $n=50$

Этот критерий занимает первую позицию в рейтинге 50-и критериев нормальности.

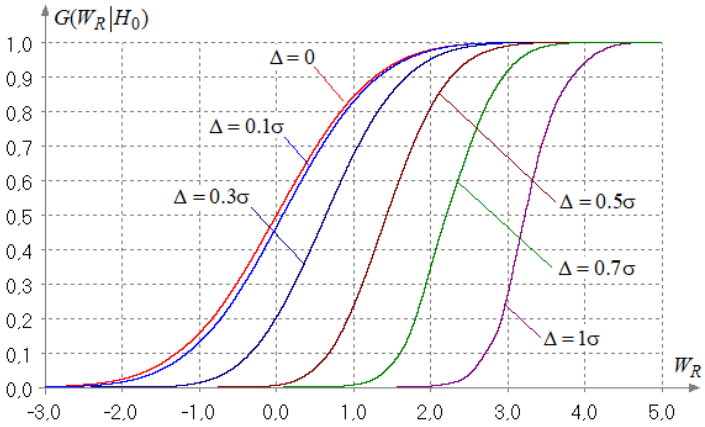


Рис. 10 – Зависимость от Δ распределения статистики W_R критерия нормальности Ройстона при справедливости H_0 и $n=50$

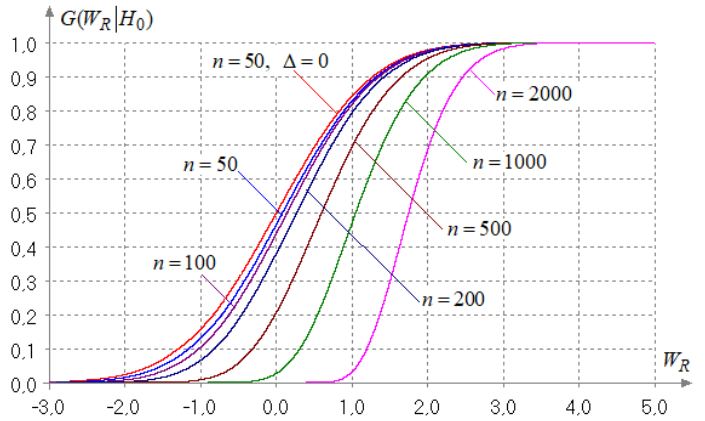


Рис. 11 – Зависимость распределения статистики W_R критерия нормальности Ройстона от n при справедливости H_0 и $\Delta=0.1\sigma$

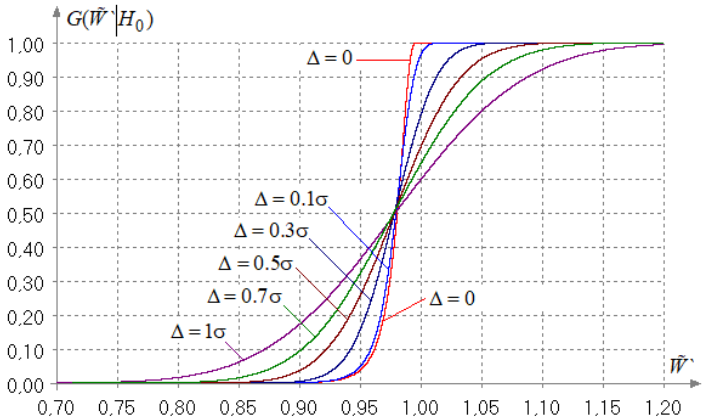


Рис. 11 – Зависимость от Δ распределения статистики $\tilde{W}`$ критерия нормальности Вайсберга–Биргема при справедливости H_0 и $n=50$

В случае правосторонних и левосторонних критериев пренебрежение фактом изменения распределений статистик вследствие влияния ошибок округления, как правило, приводит к занижению оценки p_{value} и росту вероятности ошибки 1-го рода.

В случае двусторонних критериев поведение распределений $G(S_n|H_0)$ статистик под влиянием Δ менее предсказуемо: использование классических результатов, не учитывающих влияния Δ, может приводить как к занижению, так и к завышению оценок p_{value}.

В качестве примера продемонстрируем изменение распределений $G(d|H_0)$ статистики двустороннего критерия Гири в зависимости от Δ при объёме выборок $n=50$ (см. рис. 12).

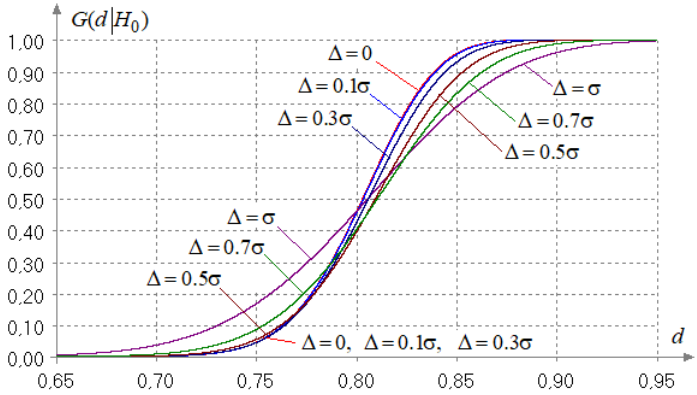


Рис. 12 – Зависимость от Δ распределения статистики d критерия Гири при справедливости гипотезы H_0 о нормальности при $n=50$

Значения p_{value}, получаемые по реальным распределениям статистик, учитывающим влияние Δ, в случае двусторонних критериев могут быть как выше, так и ниже их оценок, вычисляемых на основании классических результатов. Это можно увидеть далее в таблицах 3-5, представленных в следующем разделе.

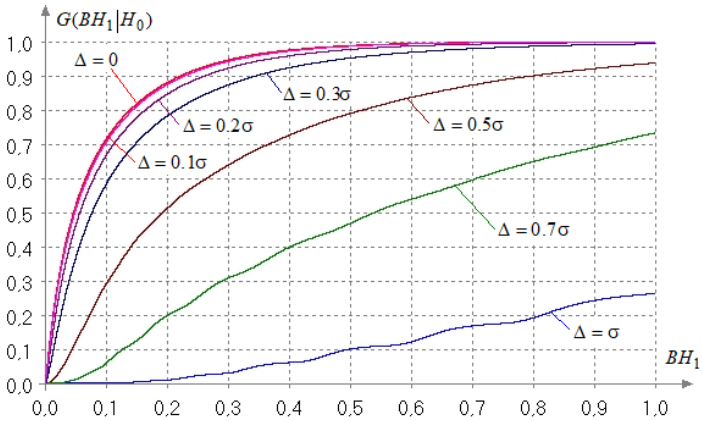


Рис. 13 – Зависимость распределения статистики (2.3) критерия BH_1 Баринхауса–Хензе при проверке экспоненциальности в зависимости от ошибки округления Δ при $n=50$

Все приведенные выше примеры говорят об одном: при соизмеримости Δ и σ нельзя пренебрегать фактом изменения распределений статистик критериев, так как в противном случае, как правило, возрастает вероятность ошибки 1-го рода α (отклонения справедливой гипотезы H_0), или реже – возрастает вероятность ошибки 2-го рода β.

Решать проблему применения критериев проверки различных гипотез в условиях влияния ошибок округления можно единственным способом, разрабатывая программное обеспечение, позволяющее методами статистического моделирования исследовать распределения

статистик критериев (или находить оценки p_{value}) в конкретных условиях приложения и при конкретном значении Δ.

Таким примером является система ISW, в рамках которой проведены настоящие исследования.

Покажем, как отражается наличие ошибок округления в хорошо известных рядах измерений на результатах проверки принадлежности нормальному закону.

XI. Применение критериев нормальности в условиях округления измерений

В работе Р. Фишера результаты измерений характеристик ирисов были использованы для решения задачи таксономии. В нашем случае рассмотрим, насколько хорошо эти измерения описываются нормальными законами распределения.

Результаты измерений в сантиметрах представлены в таблице 2.

В таблице для каждого из 3-х видов ириса (Iris setosa, Iris versicolor, Iris virginica) представлены измерения 4-х характеристик для 50 представителей каждого вида: S_l – Sepal length – длина чашелистика, S_w – Sepal width – ширина чашелистика, P_l – Petal length – длина лепестка, P_w – Petal width – ширина лепестка. Погрешность округления Δ=0.1 одна и та же для всех измерений.

Таблица 2 - Результаты измерений характеристик ирисов

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

ПРОДОЛЖЕНИЕ ТАБЛИЦЫ 2

42	4.5	2.3	1.3	0.3	6.1	3.0	4.6	1.4	6.9	3.2	5.1	2.3
43	4.4	3.2	1.3	0.2	5.8	2.6	4.0	1.2	5.8	3.3	5.1	1.9
44	5.0	3.5	1.6	0.6	5.0	2.3	3.3	1.0	6.8	3.3	5.9	2.3
45	5.1	3.8	1.9	0.4	5.6	2.7	4.2	1.3	6.7	3.3	5.7	2.5
46	4.8	3.0	1.4	0.3	5.7	3.0	4.2	1.2	6.7	3.4	5.2	2.3
47	5.1	3.8	1.6	0.2	5.7	2.9	4.2	1.3	6.3	3.4	5.0	1.9
48	4.6	3.2	1.4	0.2	6.2	2.9	4.3	1.3	6.5	3.6	5.2	2.0
49	5.3	3.7	1.5	0.2	5.1	2.5	3.0	1.1	6.2	3.8	5.4	2.3
50	5.0	3.3	1.4	0.2	5.7	2.8	4.1	1.3	5.9	3.8	5.1	1.8

Вследствие естественного округления результатов измерений в столбцах таблицы 2 наблюдаются повторяющиеся значения.

Посмотрим, как это отражается на результатах проверки гипотезы о принадлежности измерений характеристик ирисов нормальному закону.Результаты проверки принадлежности результатов измерений 4-х характеристик для каждого из 3-х видов ириса нормальным законам представлены в таблицах 3-5.

Всего анализировалось 12 выборок. Использовались:

– 8 критериев согласия (Колмогорова (K), Крамера–Мизеса–Смирнова (CMS), Андерсона–Дарлинга (AD), Купера (Ku), Ватсона (W), Жанга (Z_A , Z_C и Z_K));

– 6 специальных критериев нормальности (Бонтемпса–Меддахи со статистикой BM_{3-6} (BM), Десгань–Мишо со статистикой X_{APD} (DM), Филлибена (Fb), Гири (Gr), Ройстона (Rn), Вайсберга–Биргема (WB)).

ТАБЛИЦА 3 - ПРОВЕРКА ГИПОТЕЗ О ПРИНАДЛЕЖНОСТИ НОРМАЛЬНОМУ ЗАКОНУ ХАРАКТЕРИСТИК ДЛЯ IRIS SETOSA

Test	Sepal length			Sepal width			Petal length			Petal width		
	$\mu = 5.006$, $\sigma = 0.3489$			$\mu = 3.428$, $\sigma = 0.3753$			$\mu = 1.462$, $\sigma = 0.1719$			$\mu = 0.246$, $\sigma = 0.1043$		
	S	p_{value}		S	p_{value}		S	p_{value}		S	p_{value}	
		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$
K	0.828	0.106	0.479	0.758	0.192	0.614	1.102	0.006	0.521	2.501	0.000	0.000
CMS	0.071	0.269	0.659	0.074	0.248	0.558	0.187	0.009	0.429	0.982	0.000	0.0001
AD	0.414	0.339	0.731	0.484	0.231	0.496	0.999	0.013	0.522	4.747	0.000	0.0001
Ku	1.511	0.047	0.388	1.420	0.086	0.485	2.110	0.000	0.343	4.148	0.000	0.0004
W	0.070	0.229	0.627	0.072	0.217	0.540	0.188	0.004	0.402	0.940	0.000	0.0001
Z_A	3.312	0.636	0.828	3.356	0.148	0.220	3.365	0.107	0.520	3.706	0.000	0.001
Z_C	5.643	0.518	0.689	6.690	0.376	0.515	8.390	0.216	0.756	41.661	0.0002	0.002
Z_K	1.175	0.262	0.630	1.168	0.078	0.249	0.206	0.019	0.610	12.067	0.000	0.000
BM	0.374	0.924	0.928	2.799	0.282	0.288	3.525	0.196	0.221	16.154	0.010	0.020
DM	0.243	0.884	0.893	2.108	0.345	0.370	3.410	0.169	0.299	17.185	0.000	0.022
Fb	0.991	0.543	0.833	0.981	0.110	0.191	0.974	0.033	0.365	0.891	0.000	0.001
Gr	0.776	0.371	0.351	0.766	0.227	0.218	0.765	0.219	0.156	0.791	0.692	0.120
Rn	0.102	0.462	0.735	0.608	0.273	0.461	1.600	0.055	0.611	4.782	0.000	0.0005
WB	0.982	0.537	0.547	0.964	0.119	0.283	0.949	0.035	0.278	0.795	0.000	0.011

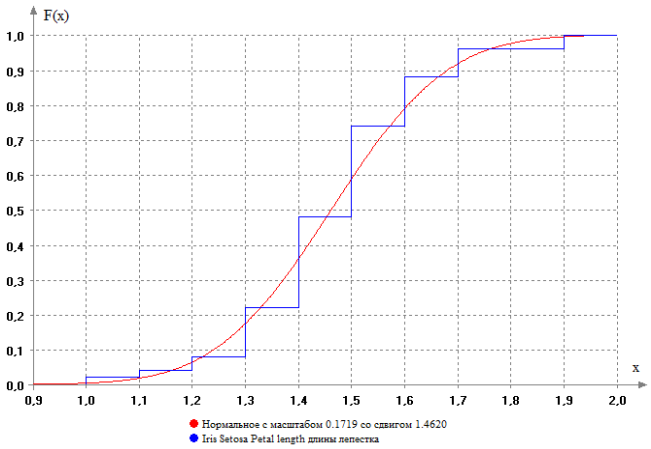


Рис. 14 – Эмпирическое распределение Petal length для Iris setosa и соответствующее нормальное распределение. Если бы мы не учли влияния округлений, то при зада-нии уровня значимости α=0.1 по 9 критериям из 14 гипоте-за о нормальности была бы отклонена!

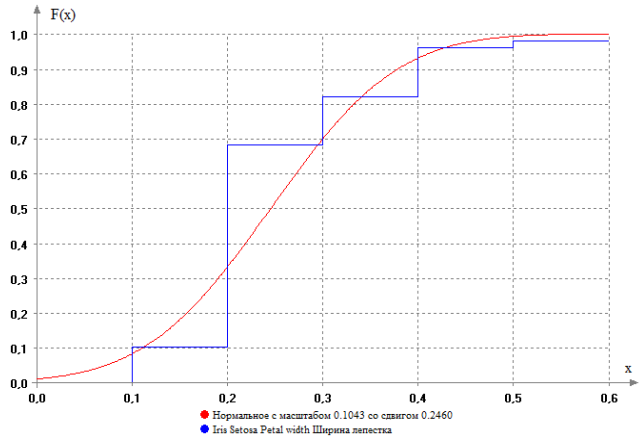


Рис. 15 – Эмпирическое распределение Petal width для Iris setosa и соответствующее нормальное распределение. При задании уровня значимости α=0.1 гипотеза о нормальности не отклоняется только по критерию Гири!

ТАБЛИЦА 4 – ПРОВЕРКА ГИПОТЕЗ О ПРИНАДЛЕЖНОСТИ НОРМАЛЬНОМУ ЗАКОНУ ХАРАКТЕРИСТИК ДЛЯ IRIS VERSICOLOR

Критерий Test	Sepal length			Sepal width			Petal length			Petal width		
	$\mu = 5.936$, $\sigma = 0.5110$			$\mu = 2.770$, $\sigma = 0.3106$			$\mu = 4.260$, $\sigma = 0.4652$			$\mu = 1.3260$, $\sigma = 0.1958$		
	S	p_{value}		S	p_{value}		S	p_{value}		S	p_{value}	
		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$
K	0.716	0.265	0.557	0.888	0.061	0.415	0.860	0.079	0.259	1.065	0.010	0.453
CMS	0.059	0.395	0.602	0.105	0.094	0.345	0.091	0.147	0.262	0.154	0.023	0.467
AD	0.374	0.421	0.613	0.573	0.139	0.450	0.562	0.149	0.254	0.975	0.015	0.319
Ku	1.282	0.199	0.519	1.403	0.097	0.678	1.252	0.232	0.623	1.886	0.001	0.432
W	0.057	0.362	0.576	0.098	0.093	0.373	0.076	0.191	0.359	0.153	0.014	0.451
Z_A	3.313	0.624	0.717	3.320	0.500	0.742	3.354	0.155	0.201	3.387	0.050	0.214
Z_C	6.214	0.441	0.515	5.900	0.483	0.691	8.593	0.203	0.256	12.31	0.055	0.226
Z_K	0.849	0.520	0.742	1.348	0.176	0.568	1.285	0.204	0.399	2.563	0.008	0.283
BM	1.168	0.639	0.644	2.464	0.333	0.343	4.062	0.151	0.154	0.913	0.721	0.744
DM	0.682	0.708	0.716	2.379	0.302	0.340	3.771	0.150	0.164	0.119	0.942	0.955
Fb	0.992	0.640	0.781	0.988	0.351	0.679	0.984	0.168	0.239	0.976	0.052	0.342
Gr	0.825	0.444	0.482	0.820	0.548	0.658	0.815	0.674	0.726	0.803	0.986	0.786
Rn	0.089	0.468	0.594	0.418	0.340	0.658	1.014	0.156	0.229	1.922	0.027	0.243
WB	0.984	0.620	0.597	0.976	0.346	0.471	0.968	0.166	0.306	0.953	0.047	0.283

ТАБЛИЦА 5 – ПРОВЕРКА ГИПОТЕЗ О ПРИНАДЛЕЖНОСТИ НОРМАЛЬНОМУ ЗАКОНУ ХАРАКТЕРИСТИК ДЛЯ IRIS VIRGINICA

Критерий Test	Sepal length			Sepal width			Petal length			Petal width		
	$\mu = 6.5880$, $\sigma = 0.6295$			$\mu = 2.9740$, $\sigma = 0.3193$			$\mu = 5.552$, $\sigma = 0.5463$			$\mu = 2.0260$, $\sigma = 0.2719$		
	S	p_{value}		S	p_{value}		S	p_{value}		S	p_{value}	
		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$		$\Delta = 0$	$\Delta = 0.1$
K	0.841	0.095	0.202	0.925	0.042	0.314	0.843	0.092	0.238	0.895	0.057	0.504
CMS	0.089	0.155	0.210	0.106	0.093	0.313	0.087	0.165	0.250	0.121	0.059	0.308
AD	0.557	0.153	0.204	0.611	0.112	0.355	0.619	0.108	0.159	0.760	0.048	0.245
Ku	1.331	0.151	0.336	1.744	0.008	0.170	1.322	0.159	0.411	1.746	0.008	0.259
W	0.085	0.140	0.198	0.102	0.081	0.308	0.074	0.201	0.322	0.121	0.044	0.273
Z_A	3.332	0.334	0.377	3.342	0.238	0.389	3.356	0.146	0.175	3.356	0.146	0.296
Z_C	6.711	0.373	0.420	7.263	0.314	0.486	9.466	0.148	0.178	9.737	0.133	0.275
Z_K	1.306	0.194	0.296	1.496	0.124	0.439	1.554	0.107	0.201	2.076	0.028	0.230
BM	1.605	0.514	0.517	2.839	0.276	0.285	3.994	0.156	0.157	1.475	0.550	0.560
DM	0.824	0.660	0.667	2.762	0.249	0.283	3.038	0.216	0.228	1.233	0.536	0.579
Fb	0.985	0.220	0.267	0.982	0.116	0.244	0.983	0.137	0.176	0.983	0.147	0.396
Gr	0.798	0.871	0.854	0.759	0.152	0.146	0.805	0.938	0.970	0.839	0.200	0.317
Rn	0.649	0.259	0.316	0.912	0.181	0.389	1.238	0.108	0.144	1.360	0.088	0.268
WB	0.971	0.221	0.324	0.964	0.122	0.309	0.965	0.131	0.240	0.966	0.136	0.350

Для каждой выборки в таблицах приводятся ОМП параметров μ и σ нормального закона и вычисленные значения S статистик применяемых критериев. Значения достигнутого уровня значимости p_{value} в предположении об отсутствии ошибок округления (при $\Delta = 0$) для критериев Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Купера, Ватсона, Десгань–Мишо и Ройстона могут быть рассчитаны по известным асимптотическим распределениям статистик. Но распределения статистик остальных критериев зависят от n. Для чистоты эксперимента оценки p_{value} в отсутствие ошибок округления (при $\Delta = 0$) находились по распределениям статистик $G(S_n|H_0)$, моделируемым при $n = 50$.

Оценки p_{value} в условиях влияния ошибок округления (при $\Delta = 0.1$ и при соответствующей ОМП для σ) вычислялись по реальным распределениям $G(S_n|H_0)$ статистик критериев, моделируемым в интерактивном режиме. Такая возможность реализована в нашей программной системе ISW.

Как можно видеть, оценки p_{value}, вычисленные по реальным распределениям $G(S_n|H_0)$ статистик, имеющим место в условиях наличия ошибок округления ($\Delta = 0.1$), кардинально отличаются от значений p_{value}, вычисленных по распределениям статистик этих же критериев в условиях отсутствия ошибок округления ($\Delta = 0$). И если пренебречь влиянием ошибок округления на распределения ста-

тистик критериев, то во многих случаях гипотеза о нормальности будет несправедливо отклоняться.

В данном случае надо обратить внимание на то, что каждой проверке по каждому применяемому критерию при ($\Delta = 0.1$ и $n = 50$) соответствует своё распределение статистики $G(S_n|H_0)$, зависящее от σ нормального закона. То есть, для анализа 12 выборок по каждому из 14 критериев мы должны использовать 12 различных распределений $G(S_{50}|H_0)$ статистики применяемого критерия, по которому и вычисляется p_{value}.

Следует сделать ещё одно важное замечание. Среди специальных критериев нормальности есть правосторонние, левосторонние и двусторонние критерии. Вследствие влияния ошибок округления распределения статистик правосторонних критериев сдвигаются вправо, а левосторонних – влево. При этом реальный достигнутый уровень значимости p_{value}, учитывающий влияние Δ, всегда оказывается не меньше того, что мы имеем при его вычислении без учёта такого влияения. В случае двусторонних критериев вследствие влияния Δ область определения статистик критериев также меняется: при этом она может изменяться по масштабу и сдвигаться влево или вправо. Поэтому реальный p_{value}, учитывающий влияние Δ, может возрастать, а может и уменьшаться. Чтобы подчеркнуть этот факт, в строках таблиц 3-5 для двустороннего критерия Гири ситуации с уменьшением p_{value} из-за влияния Δ выделены цветом.

В отсутствие влияния округлений распределения $G(S_n|H_0)$ статистик критериев Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Купера, Ватсона, Десгань–Мишо и Ройстона быстро сходятся к асимптотическим $G(S|H_0)$ распределениям этих статистик: отклонением $G(S_n|H_0)$ от $G(S|H_0)$ можно пренебречь, как правило, при $n \geq 25 \div 30$. При наличии влияния ошибок округления (как в данном случае) распределения $G(S_n|H_0)$ могут не сходиться к асимптотическим $G(S|H_0)$, а с ростом n всё дальше отклоняться от них.

Проведенные ранее исследования показали, что и при проверке простых гипотез о принадлежности выборок нормальному закону (в условиях влияния Δ) распределения статистик $G(S_n|H_0)$ непараметрических критериев согласия (Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Купера, Ватсона) становятся зависящими от n, от Δ и от значения параметра масштаба σ, а с ростом n всё больше отклоняются от асимптотических $G(S|H_0)$.

В общем случае проверки сложной гипотезы **о принадлежности выборки некоторому закону $F(x,\theta)$ к факторам, влияющим на распределения статистик $G(S|H_0)$ при сложной гипотезе,** добавляется зависимость от n, Δ и от значений оценок параметров формы и масштаба закона $F(x,\theta)$.

XII. Заключение

Да, реальные свойства статистических критериев проверки гипотез порой отличаются от заявленных. Да, применение критериев в приложениях зачастую оказывается в условиях нарушения стандартных предположений. Да, естественное присутствие ошибок округления также может приводить к изменению свойств критериев.

Но всё это вместе не исключает возможность корректного применения критериев в нестандартных условиях различных приложений, опираясь на знание реальных свойств критериев и компьютерные технологии. А корректное применение методов статистического анализа способствует принятию обоснованных решений, в том числе в задачах метрологического обеспечения.

Литература

[1] Лемешко Б. Ю. Непараметрические критерии согласия. Руководство по применению: Монография. – М.: ИНФРА-М, 2014. – 163 с. DOI: 10.12737/11873.

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

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

[4] Лемешко Б.Ю., Веретельникова И.В. Критерии проверки гипотез о случайности и отсутствии тренда. Руководство по применению: Монография. – Москва : ИНФРА-М. 2021. – 221 с. – (Научная мысль). DOI 10.12737/1587437.

[5] Лемешко Б. Ю. Критерии проверки гипотез об однородности. Руководство по применению: Моно¬графия. – 2-е изд., перераб. и доп. – М.: ИНФРА-М, 2021. – 248 с. (Научная мысль) DOI: 10.12737/986695.

[6] Лемешко Б. Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по примене-нию : Монография. – 2-е изд., перераб. и доп. – Москва : ИНФРА-М, 2023. – 353 с. – (Научная мысль). DOI 10.12737/1896110.

[7] Лемешко Б. Ю., Лемешко С. Б. Непараметрические критерии согласия при проверке нормальности в условиях округления измерений // Системы анализа и обработки данных. – 2022. – № 2 (86). – С. 21-38. – DOI: 10.17212/2782-2001-2022-2-21-38.

[8] Лемешко Б. Ю., Лемешко С. Б. Проблемы применения непараметрических критериев согласия в задачах обработки результатов измерений // Системы анализа и обработки данных. – 2021. – № 2 (82). – С. 47-66. – DOI: 10.17212/2782-2001-2021-2-47-66.

[9] Lemeshko B. Y., Lemeshko S. B. About the effect of rounding on the properties of tests for testing statistical hypotheses // Journal of Physics: Conference Series. – 2021. – Vol. 1715. 012063. DOI: 10.1088/1742-6596/1715/1/012063.

[10] Лемешко Б. Ю., Лемешко С. Б. О влиянии ошибок округления на распределения статистик критериев согласия // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2020. – № 53. – С. 47-60. DOI: 10.17223/19988605/53/5.

[11] Лемешко Б. Ю., Лемешко С. Б. Влияние округления на свойства критериев проверки статистических гипотез // Автометрия. – 2020. – Т. 56, № 3. – С. 35-45. - DOI: 10.15372/AUT20200305.