

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.23

DOI: 10.17212/2782-2001-2025-2-53-80

О свойствах и проблемах применения критериев согласия, опирающихся на использование оценок энтропии и дивергенции Кульбака – Лейблера*

Б.Ю. ЛЕМЕШКО^{1,a}, С.Б. ЛЕМЕШКО^{1,2,b}

¹ 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет

² 630090, РФ, г. Новосибирск, ул. Пирогова, 1, Новосибирский национальный исследовательский государственный университет

^a lemeshko@ami.nstu.ru ^b skyer@mail.ru

Для проверки адекватности построенных моделей законов распределения случайных величин, как правило, применяются различные непараметрические критерии согласия, в частности критерии Колмогорова, Крамера – Мизеса – Смирнова, Андерсона – Дарлинга, Купера, Ватсона.

При справедливости простой проверяемой гипотезы непараметрические критерии согласия являются «свободными от распределения»: асимптотические распределения статистик не зависят от вида закона, относительно которого проверяется гипотеза. При проверке сложных гипотез, когда по выборке оцениваются параметры предполагаемого закона, свойство «свободы от распределения» теряется и распределения статистик становятся зависящими от целого ряда факторов. В таких ситуациях применение непараметрических критериев согласия возможно только при поддержке соответствующего программного обеспечения, позволяющего с использованием имитационного моделирования оценивать достигаемый уровень значимости P .

Распределения статистик критериев Жанга, представляющих собой развитие критериев Колмогорова, Крамера – Мизеса – Смирнова, Андерсона – Дарлинга, зависят от объемов выборок, поэтому их широкое применение при проверке простых и сложных гипотез возможно только с опорой на метод Монте – Карло.

Распределения статистик критериев согласия (при проверке простых и сложных гипотез) могут существенно изменяться вследствие естественного присутствия ошибок округления. Сигналом о возможности такой ситуации является наличие в анализируемых выборках значительного числа повторяющихся значений. В подобных ситуациях принятие решения о результатах проверки также невозможно без использования имитационного моделирования.

В последние годы предложено несколько критериев, ориентированных, например, на проверку принадлежности выборок нормальному или равномерному закону, статистики которых опираются на различные оценки энтропии. Как показывает опыт, относительно некоторых

* Статья получена 12 февраля 2025 г.

конкурирующих гипотез такие критерии демонстрируют более высокие оценки мощности по сравнению с классическими непараметрическими критериями согласия.

При построении статистики критерия Нугаби для различения двух гипотез использована дивергенция Кульбака – Лейблера, а в качестве оценки энтропии взята оценка, предложенная Васичеком. В настоящей работе показано, как распределения статистики критерия Нугаби зависят от объема выборок n и размера «окна» m , как меняются распределения статистик критерия при проверке различных сложных гипотез. Исследована мощность критерия при проверке нормальности относительно различных конкурирующих гипотез. Показано, как при данных n мощность зависит от размера «окна» m . Показано существование оптимального m , при котором мощность максимальна относительно рассматриваемой конкурирующей гипотезы. Показано, что при данном n оптимальные значения m , как правило, не совпадают для различных конкурирующих гипотез. Очевидно, что применение подобных критериев на практике также подразумевает использование соответствующего программного обеспечения и имитационного моделирования.

Ключевые слова: программная система, оценки параметров, непараметрические критерии согласия, распределение статистики, достигнутый уровень значимости, проверка нормальности, ошибка 1-го рода, ошибка 2-го рода, мощность критерия, имитационное моделирование

ВВЕДЕНИЕ

Проверка гипотез об адекватности построенных моделей законов распределения случайных величин, наблюдаемых в различных приложениях, как правило, осуществляется с использованием либо некоторого непараметрического критерия согласия, либо критериями согласия типа χ -квадрат. В случае определенных законов, наиболее востребованных на практике (например, нормального, экспоненциального, равномерного), могут применяться специальные параметрические критерии, ориентированные на проверку гипотезы только относительно этого закона. Зачастую выводы опираются на результаты применения нескольких критериев.

В классической ситуации, когда отсутствует какое-либо цензурирование или группирование данных и мы имеем дело с точечной выборкой $x_1, x_2, \dots, x_n$ объемом n , применяемый критерий согласия должен ответить на вопрос, верна ли гипотеза H_0 о принадлежности этой выборки закону с функцией распределения $F(x, \theta)$, где θ – скалярный или векторный параметр, или она должны быть отклонена.

При проверке согласия необходимо учитывать, какая гипотеза проверяется: простая или сложная. 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 . Предполагается, что оценка $\hat{\theta}$ скалярного или векторного параметра закона опирается на ту же выборку, по которой проверяют гипотезу.

Перечень существующих непараметрических критериев согласия, применяемых к произвольным законам распределения $F(x, \theta)$, достаточно ограничен. Исторически первым можно считать появление критерия Колмогорова [1]. Практически в то же время был построен критерий Крамера –

Мизеса – Смирнова [2–4]. Затем был предложен критерий Андерсона – Дарлинга [5, 6], чуть позже – критерии Купера [7] и Ватсона [8, 9]. При проверке простых гипотез, для которых они и были предложены, перечисленные критерии являются «свободными от распределения», т. е. известные асимптотические распределения $G(S|H_0)$ статистик S данных критериев не зависят от вида закона $F(x, \theta)$, соответствующего проверяемой гипотезе H_0 . При проверке сложных гипотез свойство «свободы от распределения», благодаря которому эти критерии удостоились названия «непараметрические», теряется [10].

В ситуации проверки сложной гипотезы на распределения статистик $G(S|H_0)$ критериев воздействует следующий ряд факторов [11]:

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

б) – тип оцениваемого параметра;

в) – число оцениваемых параметров;

г) – используемый метод оценивания параметров [12];

д) – в некоторых ситуациях распределения статистик зависят от конкретных значений параметров закона $F(x, \theta)$ (например, от значений параметров формы обобщенного нормального закона [13], гамма-распределения, бета-распределений и т. п.).

После публикации [10] усилия многих авторов были направлены на решение проблем, связанных с применением непараметрических критериев согласия в условиях проверки сложных гипотез. Но в рамках чисто аналитического подхода удавалось получать только частные результаты [14–16], не оказавшие заметного влияния на разрешение возникшей ситуации. Возможный же выход был подсказан еще в работах [17, 18], где для статистики критерия Колмогорова с использованием метода Монте-Карло были построены таблицы критических значений для проверки сложных гипотез о принадлежности выборок нормальному и экспоненциальному законам. Однако эффективное применение этого подхода в целях развития и расширения возможностей методов прикладной математической статистики задержалось до эпохи массового использования персональных компьютеров. В последние же 30 лет методы статистического моделирования в задачах исследования статистических и вероятностных закономерностей стали с успехом применяться особенно активно. Это касается и решения проблем применения непараметрических критериев согласия в ситуации проверки сложных гипотез.

В частности, на основании построенных нами таблиц критических значений и моделей распределений статистик критериев согласия Колмогорова, Крамера – Мизеса – Смирнова и Андерсона – Дарлинга, имеющих место в ситуациях проверки различных сложных гипотез [19, 20], были подготовлены методические рекомендации [21], на базе которых далее разработаны рекомендации Госстандарта РФ [22]. Позже в результате дальнейших исследований и уточнения моделей для критериев Колмогорова, Крамера – Мизеса – Смирнова и Андерсона – Дарлинга [23–27], построения аналогичных моделей для критериев Купера и Ватсона [28, 29] было подготовлено руководство [30].

Необходимо отметить, что при наличии фактора (д), когда мы знаем, что распределение статистики зависит от значения параметра (или параметров) формы предполагаемого закона, условия, полностью определяющие характер

сложной проверяемой гипотезы, становятся известными только после нахождения оценки этого параметра. Отсюда следует, что корректность выводов по непараметрическим критериям согласия при проверке соответствующей сложной гипотезы можно обеспечить только за счет интерактивного моделирования (методом Монте-Карло) распределения статистики применяемого критерия *в процессе самой проверки*. Это означает, что нельзя обойтись без соответствующей программной поддержки, реализующей возможность такого моделирования. В нашем случае мы опираемся на возможности развиваемой программной системы [31].

Прошло около 70 лет, как стало известно о существовании проблем, затрудняющих использование непараметрических критериев согласия для проверки сложных гипотез [10], но по-прежнему большинство совершаемых ошибок связано с примерами неправомерного применения классических результатов к проверке сложных гипотез [32, 33]. Проблема была, можно сказать, забыта, и о ней не упоминается в учебниках.

Через 40 лет после появления критерия Ватсона [8, 9] Жангом в работе [34] были предложены 3 критерия, представляющие собой дальнейшее развитие критериев Колмогорова, Крамёра – Мизеса – Смирнова и Андерсона – Дарлингга, а в [35] рассмотрено применение предложенных критериев для проверки нормальности. Жанг показал, что предложенные критерии, как правило, имеют некоторое преимущество в мощности по сравнению со своими предшественниками. В [28, 30] мы смогли убедиться в том, что это действительно имеет место. Распределения статистик, предложенных Жангом, при проверке простых и сложных гипотез зависят от объемов n анализируемых выборок. С позиций применения такая зависимость является недостатком. Однако, как справедливо полагал Жанг, этот недостаток не является существенным, так как для нахождения распределений статистик и вычисления достигнутого уровня значимости P_v можно воспользоваться интерактивным моделированием. Естественно, это требует соответствующего программного обеспечения (ПО), подобного упомянутой программной системе [31], в которой реализована возможность применения критериев Жанга. Несмотря на определенные преимущества в мощности и достаточно существенное число ссылок в научных публикациях на работы [34, 35], к сожалению, пока нельзя говорить о сколь-нибудь частом использовании критериев Жанга в приложениях. Если проверка по этим критериям будет реализована в некоторой популярной программной системе, то частота использования, конечно, возрастет.

Следует остановиться еще на одном факторе, воздействие которого приводит к изменению распределений статистик критериев согласия при проверке как простых, так и сложных гипотез. Последние 25–30 лет характеризуются взрывным накоплением наблюдаемых данных, фиксируемых в различных сферах человеческой деятельности. Вследствие этого особую актуальность приобретают методы, позволяющие извлекать из таких больших данных (Big Data) скрытые в них закономерности. Вполне естественно при анализе больших данных использовать различные критерии проверки статистических гипотез, так как асимптотические распределения статистик многих критериев при справедливости проверяемых гипотез имеют место именно при $n \rightarrow \infty$. Однако попытки непосредственного применения тех же критериев согласия к очень большим выборкам, извлекаемым из Big Data, всегда приводят к отклонению

даже справедливой проверяемой гипотезы. Причина этого явления лежит на поверхности. Да, мы имеем выборки очень большого объема, но данные в них представлены с определенными ошибками округления Δ . Вследствие этого в выборках оказывается много повторяющихся наблюдений. Поэтому реальные распределения $G(S_n|H_0)$ статистик S_n применяемых критериев при справедливости проверяемой гипотезы H_0 с ростом объемов выборок n всё больше отклоняются от асимптотических распределений $G(S|H_0)$ статистик этих критериев. В случае очень больших выборок выход из этой ситуации достаточно прост [36]: надо извлекать из Big Data выборки такого объема n , при котором распределение $G(S_n|H_0, \Delta)$ при ошибках округления Δ еще существенно не отклоняется от $G(S|H_0)$. Тогда выводы, формируемые на основании результатов проверки гипотезы по применяемым критериям, будут оставаться корректными.

На практике мы сталкиваемся с выборками небольшого объема, связанными в том числе с высокоточными измерениями, в которых есть значительная доля повторяющихся наблюдений. В таких ситуациях ошибки округления Δ соизмеримы с дисперсией σ^2 ошибок измерения, а распределения статистик критериев $G(S|H_0)$ (при проверке простых и сложных гипотез) становятся зависящими от Δ и n , даже если ранее они от n не зависели. И пренебрегать этим фактом нельзя [11, 37–40]. Получается, что к моменту проверки гипотезы (простой или сложной) распределение статистики $G(S_n|H_0, \Delta)$ применяемого критерия, на основании которого можно вычислить достигнутый уровень значимости P_v , всегда оказывается неизвестным. Следовательно, единственным корректным выходом остается использование интерактивного моделирования для оценки P_v или, при необходимости, – распределения статистики $G(S_n|H_0, \Delta)$. Именно таким образом подобные ситуации разрешаются в программной системе [31].

В течение достаточно продолжительного периода времени был предложен ряд специальных критериев, ориентированных на проверку нормальности [41–46], равномерности [47, 48] или на проверку принадлежности некоторым другим параметрическим законам распределения, статистики которых опираются на различные оценки энтропии. Сравнительный анализ мощности множества критериев нормальности, проведенный в [49], и критериев равномерности, представленный в [50], показал, что относительно ряда рассмотренных (соответственно в [49] и в [50]) конкурирующих гипотез энтропийные критерии, как правило, оказываются в группе критериев, демонстрирующих наиболее высокие оценки мощности.

В работе [51] рассмотрен общий подход к построению критериев согласия, использующих оценки информации Кульбака – Лейблера. В [52] при построении статистики критерия согласия рассматривалась пара конкурирующих гипотез вида:

$$H_0: F(x) = F_0(x, \theta) \text{ при некотором } \theta \in \Theta;$$

$$H_1: F(x) \neq F_0(x, \theta) \text{ при любом } \theta \in \Theta.$$

Для различения этих гипотез использовалась дивергенция Кульбака – Лейблера, которая при справедливости H_0 должна быть равна нулю, а в других случаях принимать положительные значения. В качестве оценки энтропии была взята оценка, предложенная Васичеком в [41].

Основная цель настоящей работы заключается в том, чтобы оценить реальные свойства критерия Нугаби (Noughabi), предложенного в [52], отметить положительные и отрицательные моменты, связанные с его применением, и спрогнозировать перспективы возможного использования критерия в приложениях.

1. ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ КРИТЕРИЯ НУГАБИ

Построенная в [52] статистика критерия Нугаби имеет вид

$$DA_{mn} = -\frac{1}{n} \sum_{i=1}^n \ln \left\{ \frac{n}{2m} \left(F(X_{(i+m)}, \theta) - F(X_{(i-m)}, \theta) \right) \right\}, \quad (1)$$

где «окно» m может принимать положительные целые значения, $m \leq n/2$, $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ – вариационный ряд, построенный по выборке $x_1, x_2, \dots, x_n$, и $X_{(i)} = X_{(1)}$, если $i < 1$, $X_{(i)} = X_{(n)}$, если $i > n$. При проверке простой гипотезы θ – известное значение вектора параметров закона, а при проверке сложной – его оценка $\hat{\theta}$, построенная по этой же выборке.

З а м е ч а н и е 1. С вычислением статистики (1) может возникнуть проблема, если при некотором i разность $F(X_{(i+m)}, \theta) - F(X_{(i-m)}, \theta) = 0$. Это возможно в следующих ситуациях: во-первых, когда из-за присутствия в выборке повторяющихся значений окажется $X_{(i+m)} = X_{(i-m)}$; во-вторых, когда значения функции распределения, соответствующие началу вариационного ряда, оказываются нулевыми, т. е. $F(X_{(i-m)}, \theta) = 0$ и $F(X_{(i+m)}, \theta) = 0$; в-третьих, когда в конце вариационного ряда окажется, что $F(X_{(i-m)}, \theta) = 1$ и $F(X_{(i+m)}, \theta) = 1$.

В выборке, соответствующей некоторой непрерывной случайной величине, не может быть повторяющихся значений, поэтому с позиции критерия при обнаружении первой ситуации проверяемая гипотеза H_0 о принадлежности наблюдаемой случайной величины закону с функцией распределения $F(x, \theta)$ должна быть отклонена с $P_v = 0$. Хотя, конечно, присутствие повторяющихся значений может быть следствием округления наблюдаемых значений (результатов измерений). И тогда, возможно, отклоняя гипотезу H_0 , совершаем ошибку 1-го рода.

Обнаружение второй и/или третьей ситуации свидетельствует о том, что область определения случайной величины, описываемой законом $F(x, \theta)$, не соответствует размаху выборки. С позиции критерия в таких ситуациях гипотеза H_0 также должна отклоняться с $P_v = 0$.

При наличии ошибок округления и наличии в выборке повторяющихся наблюдений можно предусмотреть возможность переполнения при вычислении статистики, выбирая подходящее значение «окна» m . Если максимальная длина серии повторяющихся (одинаковых) значений в вариационном ряду равна l , то необходимо выбирать m из условия $n/2 \geq m > l/2$ ($m > l/2$ при четном l , $m > (l+1)/2$ при нечетном l). При $l = n$ проверяемая гипотеза должна отклоняться со значением $P_v = 0$.

В [52] автором представлены краткие таблицы критических значений для статистики (1) при $n = 5 \div 50$ и $m = 1 \div 10$ для проверки сложных гипотез о принадлежности выборок закону нормальному, закону Лапласа и закону Вейбулла (с оцениванием двух параметров этих законов методом максимального правдоподобия). В случае проверки принадлежности экспоненциальному закону представлена таблица только для ситуации с вычислением оценки максимального правдоподобия (ОМП) параметра масштаба. Приведены также оценки мощности критерия относительно достаточно широкого круга конкурирующих законов при $n = 10, 20$ и $\alpha = 0.05$. Эти оценки сравнивались с мощностью, показанной критериями Колмогорова, Крамера – Мизеса – Смирнова, Андерсона – Дарлинга, Купера и Ватсона.

Когда исследования, результаты которых представляют собой содержание настоящей статьи, были практически завершены, стала доступна работа [53], где авторы привели эквивалентную (1) статистику, свободную от недостатка, указанного в замечании 1.

Статистика критерия Алави – Нугаби (Alavi – Noughabi), предложенного в [53], имеет вид

$$DS = \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^{\infty} \frac{1}{(2j-1)} \left\{ \frac{2m/n - (F(X_{(i+m)}, \theta) - F(X_{(i-m)}, \theta))}{2m/n + (F(X_{(i+m)}, \theta) - F(X_{(i-m)}, \theta))} \right\}^{(2j-1)} \quad (2)$$

при тех же обозначениях, что и в статистике (1).

Расхождения в критических значениях статистики (1) и статистики (2) и оценок мощности, представленных соответственно в [52] и [53], вполне объясняются различным количеством экспериментов N в методе Монте-Карло: в [52] $N = 20\,000$, в [53] $N = 50\,000$. А для того чтобы погрешность моделирования некоторой функции распределения (или оценки достигнутого уровня значимости P_v) укладывалась в интервал ± 0.001 , необходимо около $N = 1\,660\,000$ экспериментов. В качестве подтверждения практической эквивалентности статистик (1) и (2) в табл. 1 приведены критические значения статистик при проверке нормальности для $m = 2$, взятые из [52] и [53] и полученные в данном случае для статистики (1) при числе экспериментов $N = 1\,660\,000$.

При проверке простых гипотез распределения $G(DA|H_0)$ статистики критерия Нугаби, как и других непараметрических критериев согласия, от вида $F(x, \theta)$ не зависят. Распределения статистики (1) зависят от объемов выборок n при проверке простых и сложных гипотез. Конечно, это является недостатком, характерным для многих критериев, но это легко нивелируется наличием соответствующего ПО и использованием интерактивного моделирования.

Таблица 1

Table 1

Критические значения статистик (1) и (2) при проверке нормальности
($m = 2, \alpha = 0.05$)

Critical values of statistics (1) and (2) when testing normality ($m = 2, \alpha = 0.05$)

n	DA [53]	DS [54]	DA ($N = 1\,660\,000$)
10	0.545	0.5449	0.528
15	0.442	0.4406	0.433
20	0.380	0.3792	0.376
25	0.340	0.3396	0.338
30	0.314	0.3124	0.312
40	0.276	0.2788	0.277
50	0.253	0.2556	0.255

Для случая проверки простых гипотез рис. 1 демонстрирует зависимость распределения статистики (1) от объема выборки n при $m = 1$, а рис. 2 – зависимость от размера «окна» m при $n = 50$.

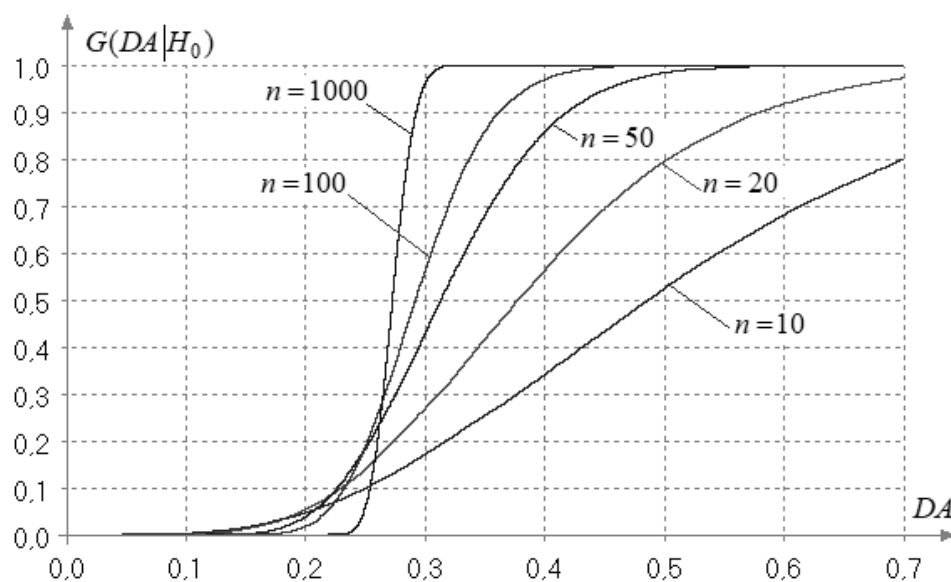


Рис. 1. Зависимость от n распределения статистики (1) при справедливости простой гипотезы и $m = 1$

Fig. 1. Dependence on the n distribution of statistics (1) when a simple hypothesis is true and $m = 1$

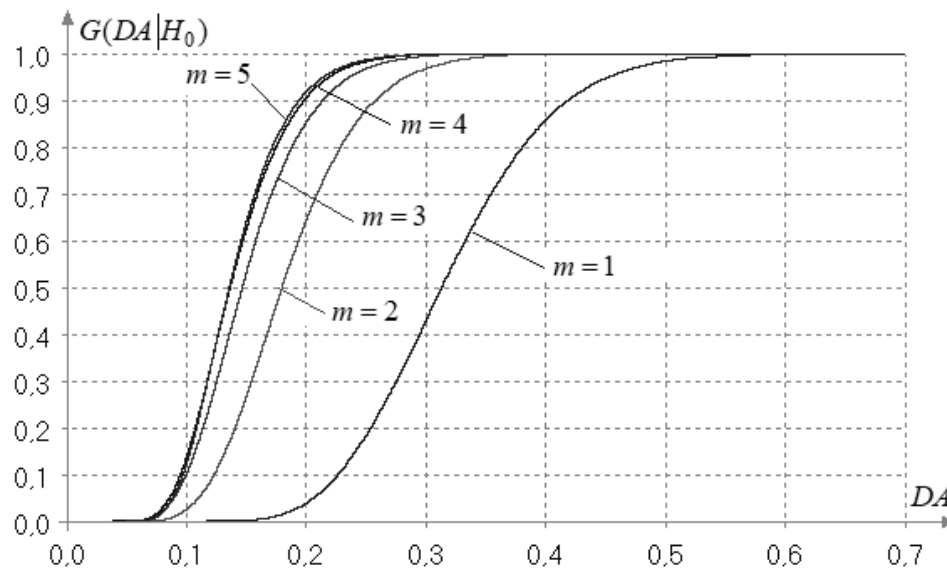


Рис. 2. Зависимость от m распределения статистики (1) при справедливости простой гипотезы и $n = 50$

Fig. 2. Dependence on the m distribution of statistics (1) when a simple hypothesis is true and $n = 50$

При проверке сложных гипотез с оцениванием параметров закона по анализируемой выборке сохраняется влияние на распределения $G(DA|H_0)$ статистики всех вышеуказанных факторов (а–д). Однако это влияние оказывается не таким «ярким», как в случае остальных классических критериев. Например, распределения $G(DA|H_0)$ статистики (1) при проверке сложной гипотезы о принадлежности законам (нормальному, Лапласа, Коши, логистическому) с вычислением ОМП двух параметров этого закона оказываются практически совпадающими, т. е. зависимость от вида закона может быть выражена очень слабо. Слабо выражена зависимость от типа оцениваемого параметра: распределения $G(DA|H_0)$ могут (практически) одинаково изменяться в результате оценивании только параметра сдвига, или только параметра масштаба. Но всё-таки зависимости от факторов (а–д) хорошо прослеживаются.

Пониманию нюансов того, как на распределениях $G(DA|H_0)$ статистики (1) отражается влияние факторов (а–д), должна способствовать картина, представленная на рис. 3 и 4, показывающая, как меняются распределения статистики в случае проверки (различных) сложных гипотез о принадлежности выборок обобщенному нормальному закону с плотностью

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

и значениями параметра формы $\theta_2 = 1$ (распределение Лапласа) и $\theta_2 = 7$.

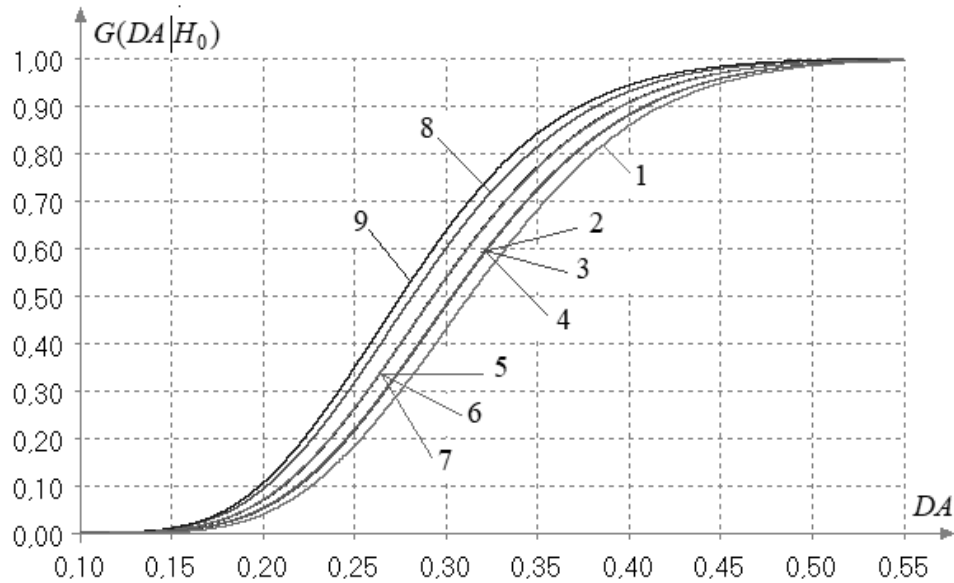


Рис. 3. Распределения статистики (1) при проверке сложных гипотез относительно обобщенного нормального закона с параметром формы $\theta_2 = 1$ и вычислением ОМП параметров (при $m = 1$, $n = 50$):

1 – при проверке простых гипотез; 2, 3, 4 – при оценивании только одного параметра из трех; 5, 6, 7 – при оценивании двух параметров из трех; 8 – при оценивании всех трех параметров закона; 9 – при оценивании всех трех параметров закона в случае $\theta_2 = 7$

Fig. 3. Distributions of statistics (1) when testing composite hypotheses with respect to the generalized normal law with a shape parameter $\theta_2 = 1$ and calculating the MLE of the parameters (for $m = 1$, $n = 50$):

1 is when testing simple hypotheses; 2, 3, 4 is when evaluating only one parameter out of three; 5, 6, 7 is when evaluating two parameters out of three; 8 is when evaluating all three parameters of the law; 9 is when evaluating all three parameters of the law in the case $\theta_2 = 7$

Очевидно, что корректное использование критерия со статистикой (1) возможно только в рамках некоторого ПО, позволяющего в ходе проверки гипотезы H_0 находить оценки достигнутого уровня P_v в условиях, соответствующих сложной проверяемой гипотезе, используя для этого методы статистического моделирования. При реализации процесса моделирования должно быть учтено, что распределение $G(DA|H_0)$ статистики (1) существенно зависит от объема выборки n и от размера «окна» m , в большей или меньшей степени – от вида закона $F(x, \theta)$, соответствующего проверяемой гипотезе H_0 , от типа и числа параметров, оцениваемых по выборке, от значений параметров формы закона, от выбранного метода оценивания параметров. Такая возможность реализована в программной системе [31].

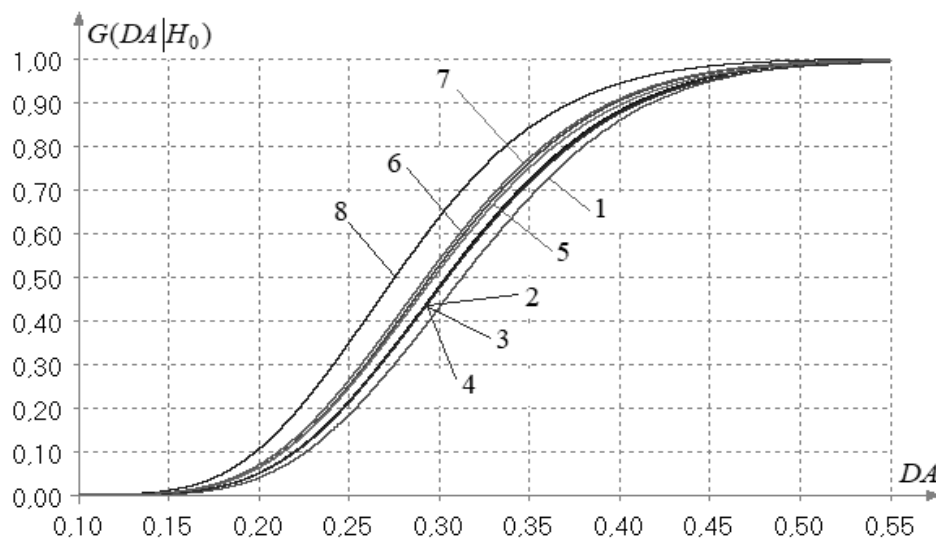


Рис. 4. Распределения статистики (1) при проверке сложных гипотез относительно обобщенного нормального закона с параметром формы $\theta_2 = 7$ и вычислением ОМП параметров (при $m = 1, n = 50$):

1 – при проверке простых гипотез; 2, 3, 4 – при оценивании только одного параметра из трех; 5 – при оценивании параметров формы и сдвига; 6 – при оценивании параметров формы и масштаба; 7 – при оценивании параметров сдвига и масштаба; 8 – при оценивании всех трех параметров закона

Fig. 4. Distributions of statistics (1) when testing composite hypotheses with respect to the generalized normal law with a shape parameter $\theta_2 = 7$ and calculating the MLE of the parameters (for $m = 1, n = 50$):

1 is when testing simple hypotheses; 2, 3, 4 is when evaluating only one parameter out of three; 5 is when evaluating the shape and shift parameters; 6 is when evaluating the shape and scale parameters; 7 is when evaluating the shift and scale parameters; 8 is when evaluating all three parameters of the law

2. ИССЛЕДОВАНИЕ МОЩНОСТИ КРИТЕРИЯ НУГАБИ

С проверкой гипотез связано два вида ошибок. Ошибка 1-го рода заключается в отклонении справедливой гипотезы H_0 . Вероятность такой ошибки, как правило, задают и обозначают α . Когда справедлива некоторая конкурирующая гипотеза H_1 , а мы в результате проверки не отклоняем H_0 , то совершаем ошибку 2-го рода. Вероятность такой ошибки обозначают β . Разность $1 - \beta$ называют мощностью применяемого критерия. По существу, мощность представляет собой вероятность отклонения H_0 при справедливости H_1 . Очевидно, что предпочтительней использовать критерии, обладающие большей мощностью относительно пары рассматриваемых гипотез H_0 и H_1 . Особенно важна способность критериев согласия различать близкие конкурирующие законы. Отметим, что для состоятельных и несмещенных критериев $1 - \beta \geq \alpha$.

В данном случае мы подробно исследовали мощность критерия со статистикой (1) при проверке гипотезы H_0 о принадлежности анализируемой

выборки нормальному закону относительно трех следующих конкурирующих гипотез.

В качестве конкурирующей гипотезы H_1 рассматривается обобщенный нормальный закон с плотностью (3) при значении параметра формы $\theta_2 = 4$.

В качестве конкурирующей гипотезы H_2 рассматривается распределение Лапласа. Это также частный случай закона (3) при значении параметра формы $\theta_2 = 1$.

В качестве гипотезы H_3 рассмотрено логистическое распределение с плотностью

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

Функция распределения этого закона при значениях параметров $\theta_0 = 0$, $\theta_1 = \sqrt{3} / \pi \approx 0.551328895$ оказывается наиболее близкой к функции распределения стандартного нормального закона. Этот закон трудно отличить от нормального с использованием различных критериев согласия.

Функции плотности законов, соответствующих гипотезам $H_0 \div H_3$, приведены на рис. 5.

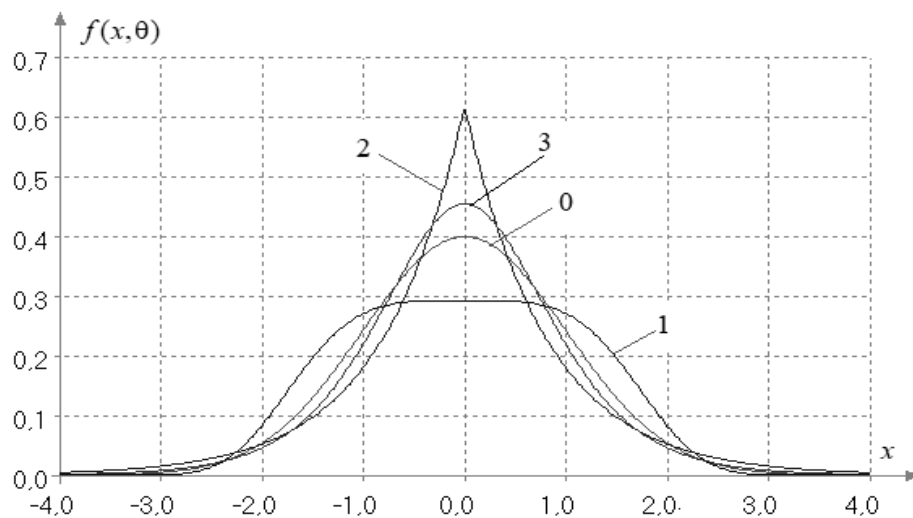


Рис. 5. Плотности распределения, соответствующие конкурирующим гипотезам: 0 – стандартный нормальный закон; 1 – обобщенный нормальный закон с параметром формы $\theta_2 = 4$, параметром масштаба $\theta_1 = 1.8878$; 2 – распределение Лапласа с параметром масштаба $\theta_1 = 1.80824$; 3 – логистическое распределение с параметром масштаба $\theta_1 = 0.551328895$

Fig. 5. Distribution densities corresponding to competing hypotheses:

0 is a standard normal law; 1 is a generalized normal law with shape parameters $\theta_2 = 4$, scale parameter $\theta_1 = 1.8878$; 2 is Laplace distribution with scale parameter $\theta_1 = 1.80824$; 3 is logistic distribution with scale parameter $\theta_1 = 0.551328895$

Конкурирующая гипотеза H_1 в настоящей статье рассматривается в связи с тем, что относительно именно этой гипотезы в свое время неожиданно проявилась смещенность некоторых специальных критериев нормальности [54, 49].

Полученные оценки мощности относительно конкурирующих гипотез H_1 , H_2 , H_3 представлены в табл. 2–4 соответственно. Как можно заметить, мощность критериев со статистикой (1) при одном и том же объеме выборки n зависит от размера «окна» m .

Зависимость мощности относительно гипотезы H_1 имеет следующий характер (см. табл. 2). С ростом m (при фиксированном n) мощность возрастает до максимальных значений при некотором m . Строки с размером «окна» m , которому соответствуют максимальные значения мощности при данном объеме n , в таблицах выделены **жирным** шрифтом. При дальнейшем росте m мощность начинает убывать, причем падает до значений $1 - \beta$, не превышающих величин α , что свидетельствует о смещенности критерия. В такой ситуации критерий уже в принципе не способен отличать гипотезу H_1 от гипотезы H_0 . Строки, где $1 - \beta < \alpha$, в таблицах выделены *курсивом* на темном фоне.

Относительно гипотезы H_2 (см. табл. 3) картина другая. С ростом m мощность сначала падает, достигая минимальных значений при некотором m , затем возрастает, принимая максимальные значения при максимально возможном m . Исключение из этого правила имеет место только при $n = 10$: при $m = 1$ критерий уже оказывается смещенным.

Относительно гипотезы H_3 (см. табл. 4) ситуация несколько иная. Также с ростом m мощность сначала падает, достигая минимальных значений при некотором m . Причем в точке m , соответствующей минимуму, и в близких к ней оказывается, что $1 - \beta < \alpha$. Затем мощность начинает возрастать, принимая максимальные значения при максимально возможном m . В случае $m = 1$ при $n = 10 \div 50$ имеем дело со смещенным критерием, а при $n \geq 100$ – критерий несмещенный (мощность $1 - \beta > \alpha$).

Если сравнить максимальные значения мощностей при соответствующих n , представленные в табл. 2–4, с мощностью классических непараметрических критериев согласия (относительно тех же конкурирующих гипотез), оценки которой представлены в разделе 3 руководства [49], то можно однозначно прийти к выводу о превосходстве в мощности критерия со статистикой (1).

Однако анализ мощности, которую критерий демонстрирует относительно конкурирующих гипотез H_1 и H_3 , позволяет увидеть следующую (неприятную) закономерность. Если при $n = 10$ максимальная мощность относительно H_1 (табл. 2) имеет место при $m = 2$, то при этом же m мощность относительно H_3 (табл. 4) оказывается меньше α . То есть критерий не может отличить H_3 от H_0 . Аналогично при $n = 20$ максимум мощности относительно H_1 оказывается при $m = 3$, а относительно H_3 в этом случае снова обнаруживается смещенность критерия (мощность меньше α). Практически то же самое наблюдается при остальных объемах выборок, представленных в табл. 2 и 4.

Если отталкиваться от максимальных значений мощности относительно H_3 и соответствующих им размеров «окна» m , приведенных в табл. 4, то при

тех же m и n в табл. 2 (к большому сожалению) обнаруживаем, что мощности относительно H_1 оказываются меньше α .

Вместе это означает следующее. Если мы оптимальным образом настроим критерий со статистикой (1) на максимальную мощность относительно конкурирующей гипотезы H_1 , то он не способен будет отличать гипотезу H_3 от гипотезы H_0 ; если оптимальным образом настроим критерий на максимальную мощность относительно конкурирующей гипотезы H_3 , – будет не способен отличать гипотезу H_1 от H_0 .

Таблица 2

Table 2

Мощность критерия относительно H_1 Power of the test relative to H_1

n	m	α				
		0.15	0.1	0.05	0.025	0.01
10	1	0.215	0.149	0.079	0.041	0.017
	2	0.227	0.159	0.084	0.044	0.018
	3	0.198	0.134	0.068	0.033	0.013
	4	<i>0.140</i>	<i>0.092</i>	<i>0.044</i>	<i>0.021</i>	<i>0.008</i>
	5	<i>0.118</i>	<i>0.077</i>	<i>0.037</i>	<i>0.018</i>	<i>0.007</i>
20	1	0.256	0.183	0.101	0.056	0.025
	2	0.319	0.236	0.137	0.078	0.036
	3	0.342	0.255	0.150	0.086	0.040
	5	0.313	0.229	0.132	0.073	0.032
	8	<i>0.129</i>	<i>0.083</i>	<i>0.039</i>	<i>0.018</i>	<i>0.007</i>
	10	<i>0.093</i>	<i>0.058</i>	<i>0.028</i>	<i>0.013</i>	<i>0.005</i>
30	1	0.283	0.205	0.116	0.065	0.029
	3	0.420	0.326	0.204	0.124	0.062
	5	0.449	0.352	0.225	0.138	0.070
	6	0.439	0.343	0.218	0.133	0.067
	9	0.336	0.250	0.146	0.079	0.034
	14	<i>0.082</i>	<i>0.051</i>	<i>0.024</i>	<i>0.011</i>	<i>0.004</i>
	15	<i>0.078</i>	<i>0.048</i>	<i>0.022</i>	<i>0.010</i>	<i>0.004</i>
40	1	0.302	0.221	0.128	0.073	0.034
	5	0.533	0.433	0.294	0.193	0.105
	6	0.541	0.442	0.301	0.198	0.108
	7	0.537	0.438	0.299	0.196	0.107
	10	0.477	0.382	0.252	0.159	0.083
	19	<i>0.069</i>	<i>0.042</i>	<i>0.019</i>	<i>0.009</i>	<i>0.003</i>
	20	<i>0.067</i>	<i>0.041</i>	<i>0.018</i>	<i>0.008</i>	<i>0.003</i>

Окончание табл. 2

End of the Tab. 2

n	m	α				
		0.15	0.1	0.05	0.025	0.01
50	1	0.318	0.235	0.139	0.081	0.038
	3	0.511	0.413	0.277	0.180	0.097
	6	0.611	0.513	0.366	0.252	0.145
	8	0.619	0.522	0.375	0.260	0.150
	10	0.598	0.502	0.359	0.246	0.142
	24	0.059	0.035	0.015	0.007	0.002
	25	0.058	0.034	0.015	0.006	0.002
100	1	0.383	0.294	0.183	0.111	0.056
	10	0.852	0.786	0.662	0.535	0.382
	11	0.857	0.793	0.671	0.546	0.392
	12	0.860	0.797	0.677	0.551	0.398
	13	0.860	0.798	0.679	0.555	0.402
	14	0.859	0.797	0.678	0.555	0.404
	50	0.027	0.015	0.006	0.002	0.001
200	18	0.983	0.970	0.934	0.881	0.786
	22	0.983	0.971	0.938	0.887	0.798
	26	0.982	0.968	0.935	0.884	0.795
	95	0.007	0.003	0.001	0.0004	0.0001
300	25	0.998	0.996	0.990	0.977	0.947
	30	0.998	0.996	0.990	0.978	0.9505
	35	0.998	0.996	0.989	0.977	0.949
	150	0.0016	0.0007	0.0002	0.00004	0.000

Таблица 3

Table 3

Мощность критерия относительно H_2 Power of the test relative to H_2

n	m	α				
		0.15	0.1	0.05	0.025	0.01
10	1	0.141	0.094	0.048	0.024	0.010
	2	0.152	0.105	0.057	0.031	0.014
	3	0.202	0.149	0.091	0.056	0.029
	4	0.288	0.219	0.137	0.086	0.046
	5	0.325	0.249	0.157	0.100	0.054

Окончание табл. 3

End of the Tab. 3

n	m	α				
		0.15	0.1	0.05	0.025	0.01
20	1	0.171	0.117	0.062	0.033	0.014
	2	0.157	0.108	0.057	0.031	0.014
	3	0.153	0.106	0.058	0.033	0.016
	4	0.165	0.118	0.069	0.041	0.021
	5	0.196	0.145	0.089	0.056	0.032
	8	0.428	0.339	0.223	0.146	0.084
	9	0.466	0.376	0.254	0.167	0.096
	10	0.480	0.391	0.270	0.181	0.105
	30	0.208	0.148	0.083	0.046	0.021
	3	0.183	0.130	0.073	0.041	0.020
30	4	0.177	0.125	0.071	0.041	0.020
	5	0.179	0.128	0.074	0.044	0.023
	6	0.192	0.140	0.084	0.051	0.028
	9	0.316	0.249	0.166	0.110	0.065
	14	0.601	0.512	0.378	0.269	0.163
	15	0.606	0.518	0.387	0.278	0.173
	40	0.244	0.179	0.104	0.061	0.029
	5	0.209	0.152	0.089	0.053	0.027
	6	0.207	0.151	0.090	0.055	0.029
	7	0.212	0.156	0.095	0.058	0.031
40	10	0.281	0.218	0.141	0.093	0.054
	19	0.705	0.626	0.496	0.379	0.252
	20	0.707	0.629	0.501	0.385	0.259
	50	0.280	0.209	0.126	0.076	0.039
	3	0.272	0.205	0.126	0.077	0.040
	6	0.244	0.182	0.111	0.068	0.036
	7	0.241	0.179	0.110	0.068	0.036
	8	0.243	0.181	0.112	0.070	0.038
	9	0.250	0.188	0.117	0.074	0.041
	10	0.264	0.201	0.127	0.082	0.046
50	24	0.785	0.717	0.599	0.484	0.348
	25	0.787	0.719	0.601	0.488	0.354
	100	0.434	0.350	0.239	0.161	0.094
	50	0.961	0.940	0.894	0.837	0.748
	200	0.655	0.573	0.447	0.341	0.233
	100	0.999	0.998	0.995	0.991	0.979
	300	0.793	0.727	0.612	0.505	0.380
	150	1	1	1	1	0.999

Таблица 4

Table 4

Мощность критерия относительно H_3 Power of the test relative to H_3

n	m	α				
		0.15	0.1	0.05	0.025	0.01
10	1	0.137	0.091	0.045	0.022	0.009
	2	0.142	0.095	0.048	0.024	0.010
	3	0.164	0.114	0.061	0.033	0.014
	4	0.198	0.140	0.078	0.043	0.019
	5	0.208	0.149	0.083	0.047	0.022
20	1	0.134	0.088	0.044	0.022	0.009
	2	0.125	0.082	0.039	0.019	0.008
	3	0.126	0.083	0.041	0.020	0.020
	4	0.135	0.090	0.046	0.023	0.010
	5	0.151	0.103	0.055	0.030	0.014
	8	0.240	0.176	0.103	0.061	0.030
	9	0.250	0.185	0.110	0.065	0.033
	10	0.254	0.188	0.113	0.068	0.034
	30	0.136	0.090	0.045	0.022	0.009
	3	0.114	0.074	0.035	0.017	0.007
30	4	0.114	0.074	0.036	0.017	0.007
	5	0.118	0.077	0.038	0.019	0.008
	6	0.127	0.084	0.043	0.022	0.009
	9	0.189	0.135	0.077	0.045	0.022
	14	0.286	0.216	0.133	0.082	0.043
	15	0.288	0.218	0.135	0.083	0.044
	40	0.140	0.093	0.046	0.023	0.009
	5	0.105	0.067	0.032	0.016	0.006
	6	0.107	0.069	0.033	0.016	0.007
	7	0.113	0.073	0.036	0.018	0.007
40	10	0.154	0.105	0.056	0.031	0.015
	19	0.318	0.245	0.156	0.098	0.053
	20	0.319	0.246	0.157	0.099	0.054
	50	0.144	0.096	0.048	0.024	0.010
	3	0.111	0.072	0.034	0.017	0.006
	6	0.097	0.061	0.029	0.014	0.005
	7	0.099	0.062	0.029	0.014	0.006
	8	0.102	0.065	0.031	0.015	0.006
	9	0.109	0.070	0.034	0.017	0.007
	10	0.119	0.078	0.039	0.020	0.009
50	24	0.348	0.271	0.176	0.113	0.063
	25	0.349	0.272	0.177	0.114	0.063
	100	0.167	0.114	0.060	0.031	0.013
	10	0.078	0.048	0.021	0.010	0.004
	50	0.480	0.395	0.280	0.196	0.120

Окончание табл. 4

End of the Tab. 4

n	m	α				
		0.15	0.1	0.05	0.025	0.01
200	1	0.205	0.144	0.079	0.043	0.019
	20	0.068	0.041	0.018	0.008	0.003
	100	0.678	0.597	0.471	0.366	0.255
300	1	0.237	0.171	0.097	0.055	0.026
	30	0.070	0.042	0.019	0.009	0.003
	150	0.804	0.739	0.628	0.521	0.395

В данном случае мы рассмотрели применение критерия со статистикой (1) для проверки нормальности. Понятно, что с подобными же ситуациями мы будем сталкиваться при использовании критерия для проверки принадлежности другим параметрическим законам распределения, когда при одном и том же значении «окна» m одни близкие конкурирующие законы можно отличать от закона, соответствующего H_0 , а другие отличать не удастся.

В частности, с такими ситуациями приходится иметь дело, используя критерий со статистикой (1) для проверки равномерности. В [50] при исследовании свойств различных критериев, ориентированных на проверку принадлежности выборок равномерному закону, в качестве достаточно близких к H_0 конкурирующих гипотез H_1 , H_2 , H_3 рассматривались 3 бета-распределения 1-го рода:

$$H_1 : F(x) = B_I(1.5, 1.5, 1, 0), \quad x \in [0, 1];$$

$$H_2 : F(x) = B_I(0.8, 1, 1, 0), \quad x \in [0, 1];$$

$$H_3 : F(x) = B_I(1.1, 0.9, 1, 0), \quad x \in [0, 1].$$

Здесь первые 2 параметра – это параметры формы, 3-й – параметр масштаба, 4-й – параметр сдвига. В данном случае рассматриваем простую гипотезу о принадлежности выборки равномерному закону на интервале $[0, 1]$. Проведенные исследования мощности критерия Нугаби на выборках объемом $n = 20$ и $m = 1 \div 10$ показали, что относительно H_1 максимум мощности достигается при $m = 10$, относительно H_2 – при $m = 2$, относительно H_3 – при $m = 6$, а относительно H_2 при $m = 8 \div 10$ критерий оказался смещенным.

В качестве одного из вариантов разрешения очерченной проблемы можно рассматривать предложение о применении критерия со статистикой (1) не при одном, а при нескольких отличающихся значениях m с надеждой на то, что при разных m будут распознаваться различные конкурирующие гипотезы.

Пример. Анализируемая выборка объемом $n = 50$ смоделирована в соответствии с законом (3), соответствующим гипотезе H_1 , при $\sigma^2 = 1$. Наблюдения фиксировались с ошибкой округления $\Delta = 0.01$. Полученная выборка имеет следующий вид:

0.50	1.40	-0.89	-1.09	-0.88	0.64	-0.70	-0.58	0.25	0.61
0.63	-1.10	-0.02	1.20	-1.47	1.02	-0.00	0.18	-0.74	0.21
-1.46	-0.47	-0.72	1.25	1.44	0.65	1.82	1.16	-2.03	-1.45
-0.85	0.04	0.89	0.23	0.82	0.64	-0.71	1.52	0.26	1.70
-0.80	-1.02	-0.57	-1.39	0.04	2.16	1.02	0.40	0.31	0.40

Гипотеза о принадлежности выборки нормальному закону проверялась при $m = 1$, $m = 8$ и $m = 25$. Оценки параметров находились методом максимального правдоподобия. При $n = 50$ относительно гипотезы H_1 максимальную мощность критерий имеет для $m = 8$ (см. табл. 2). Результаты проверки (значения статистики и P_v) представлены в табл. 5. Здесь же приведены мощности критерия при $\alpha = 0.1$.

Таблица 5

Table 5

Результаты проверки нормальности

Normality test results

m	Значение статистики DA	P_v	Мощность
1	0.377736	0.1438	0.235
8	0.157715	0.1897	0.522
25	0.312375	0.8569	0.034

Полную картину, объясняющую результаты проверки, отражает рис. 6, на котором показаны пары распределений статистики (1) при справедливости гипотезы о нормальности $G(DA|H_0, m)$ и при справедливости конкурирующей гипотезы $G(DA|H_1, m)$ при соответствующих значениях «окна» m . Видим, что при $m = 25$ критерий является смещенным: распределение $G(DA|H_1, m = 25)$ оказывается левее $G(DA|H_0, m = 25)$.

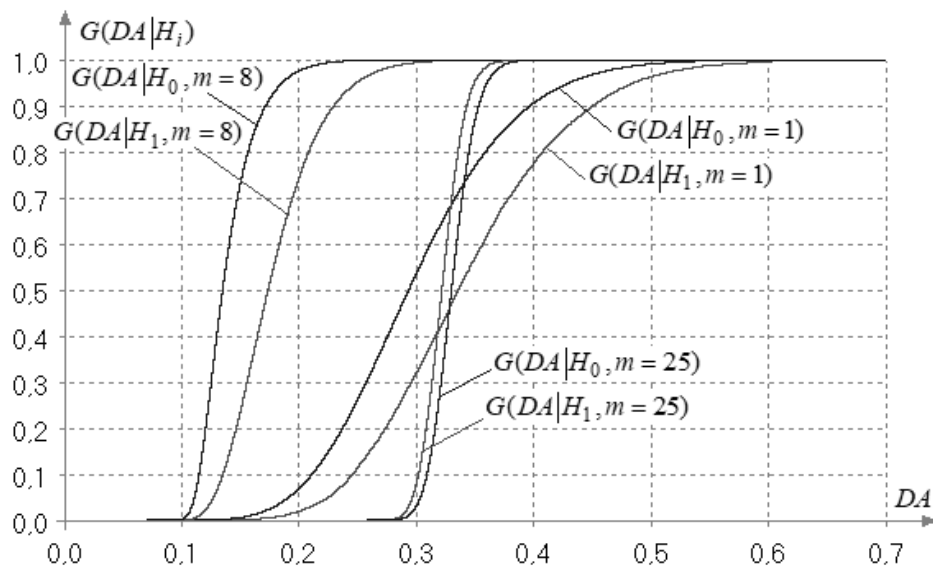


Рис. 6. Распределения статистики (1) для проверки нормальности при справедливости гипотезы H_0 и справедливости гипотезы H_1 при разных m и $n = 50$

Fig. 6. Distributions of statistics (1) when testing normality under the hypothesis H_0 and under the hypothesis H_1 for different m and $n = 50$

Как можно видеть, при задании уровня значимости $\alpha = 0.1$ гипотеза о нормальности не отклоняется ни в одном случае.

А если в соответствии с H_1 смоделировать, например, 1000 выборок объемом $n = 50$, то при $m = 1$ проверяемая гипотеза о нормальности будет отклонена в среднем в 23.5 % случаев, при $m = 8$ – в 52.2 % случаев, при $m = 25$ – лишь в 3.4 % случаев.

ЗАКЛЮЧЕНИЕ

Таким образом, при проверке гипотезы H_0 о принадлежности анализируемой выборки закону $F(x, \theta)$ критерий Нугаби за счет выбора оптимального размера «окна» m может быть настроен на максимальную мощность относительно некоторого заданного конкурирующего закона, близкого в некотором смысле к закону, соответствующему проверяемой гипотезе H_0 . Как правило, но не всегда, в такой ситуации критерий оказывается мощнее других, ставших классическими, непараметрических критериев согласия.

В табл. 5.1 руководства [49] множество из 50 различных критериев проверки нормальности упорядочено по мощности относительно тех же, что и здесь, конкурирующих гипотез H_1 , H_2 , H_3 (при объемах выборок $n = 50$ и заданном уровне значимости $\alpha = 0.1$).

Если сравнить полученные оценки мощности критерия Нугаби при оптимальных значениях m , представленные в табл. 2–4, с приведенными в [49] оценками мощности множества других критериев, то можно констатировать следующее. По мощности относительно гипотезы H_1 критерий со статистикой (1) оказывается на первом месте (с мощностью 0.522 против 0.451 у лучшего критерия в [49]), относительно H_2 окажется на 8-м месте (среди 51 критерия), относительно H_3 будет делить 21–24-е места вместе с тремя другими критериями.

С другой стороны, очевидно, что высокая мощность критерия при оптимальном значении m относительно одной конкурирующей гипотезы не гарантирует того, что при этом же m критерий будет способен отличать от H_0 другие близкие конкурирующие гипотезы.

Также очевидно, что широкое применение критериев, подобных критериям Нугаби [52, 55] и Алави – Нугаби [53], возможно только при наличии программной поддержки, позволяющей с опорой на методы статистического моделирования (как, например, в [31]) формировать корректные выводы о результатах проверки гипотез.

СПИСОК ЛИТЕРАТУРЫ

1. Kolmogoroff A.N. Sulla determinazione empirica di una legge di distribuzione // *Giornale del Istituto* 1. Kolmogoroff A.N. Sulla determinazione empirica di una legge di distribuzione // *Giornale del Istituto Italiano degli Attuari*. – 1933. – Vol. 4, N 1. – P. 83–91.
2. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.
3. Mises R. von. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. – Leipzig; Wien: Deuticke, 1931. – 574 S.

4. Smirnov N. Sur la distribution de ω^2 // Comptes Rendus de l'Academie des Sciences. – Paris, 1936. – Vol. 202. – P. 449–452.
5. Anderson T.W., Darling D.A. Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes // The Annals of Mathematical Statistics. – 1952. – Vol. 23 (2). – P. 193–212. – DOI: 10.1214/aoms/1177729437.
6. Anderson T.W., Darling D.A. A test of goodness of fit // Journal of the American Statistical Association. – 1954. – Vol. 49 (268). – P. 765–769. – DOI: 10.1080/01621459.1954.10501232.
7. Kuiper N.H. Tests concerning random points on a circle // Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Proceedings. Series A. – 1960. – Vol. 63. – P. 38–47.
8. Watson G.S. Goodness-of-fit tests on a circle // Biometrika. – 1961. – Vol. 48 (1–2). – P. 109–114.
9. Watson G.S. Goodness-of-fit tests on a circle. II // Biometrika. – 1962. – Vol. 49 (1–2). – P. 57–63.
10. Kac M., Kiefer J., Wolfowitz J. On tests of normality and other tests of goodness of fit based on distance methods // The Annals of Mathematical Statistics. – 1955. – Vol. 26 (2). – P. 189–211. – DOI: 10.1214/aoms/1177728538.
11. Лемешко Б.Ю. Непараметрические критерии согласия. Руководство по применению: монография. – 2-е изд., перераб. и доп. – М.: Инфра-М, 2024. – 201 с. – (Научная мысль). – DOI: 10.12737/2058731.
12. 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. – С. 62–71.
13. Лемешко Б.Ю., Маклаков А.А. Непараметрические критерии при проверке сложных гипотез о согласии с распределениями экспоненциального семейства // Автотметрия. – 2004. – № 3. – С. 3–20.
14. Мартынов Г.В. Критерии омега-квадрат. – М.: Наука, 1978. – 80 с.
15. Тюрин Ю.Н. О предельном распределении статистик Колмогорова – Смирнова для сложной гипотезы // Известия Академии наук СССР. Серия математическая. – 1984. – Т. 48, № 6. – С. 1314–1343.
16. Тюрин Ю.Н., Саввушкина Н.Е. Критерии согласия для распределения Вейбулла – Гнеденко // Известия Академии наук СССР. Техническая кибернетика. – 1984. – № 3. – С. 109–112.
17. Lilliefors H.W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown // Journal of the American Statistical Association. – 1967. – Vol. 62 (318). – P. 399–402.
18. Lilliefors H.W. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown // Journal of the American Statistical Association. – 1969. – Vol. 64. – P. 387–389.
19. 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.
20. Лемешко Б.Ю., Постовалов С.Н. О распределениях статистик непараметрических критериев согласия при оценивании по выборкам параметров наблюдаемых законов // Заводская лаборатория. Диагностика материалов. – 1998. – Т. 64, № 3. – С. 61–72.
21. Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Ч. 2. Непараметрические критерии. – Новосибирск: Изд-во НГТУ, 1999. – 85 с.
22. Р 50.1.037–2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Ч. 2. Непараметрические критерии. – М.: Изд-во стандартов, 2002. – 64 с.
23. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. 1 // Измерительная техника. – 2009. – № 6. – С. 3–11.
24. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. 2 // Измерительная техника. – 2009. – № 8. – С. 17–26.

25. *Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N.* Statistic distribution models for some non-parametric goodness-of-fit tests in testing composite hypotheses // *Communications in Statistics – Theory and Methods*. – 2010. – Vol. 39 (3). – P. 460–471.
26. *Lemeshko B.Yu., Lemeshko S.B.* Models of statistic distributions of nonparametric goodness-of-fit tests in composite hypotheses testing for double exponential law cases // *Communications in Statistics – Theory and Methods*. – 2011. – Vol. 40 (16). – P. 2879–2892.
27. *Lemeshko B.Yu., Lemeshko S.B.* Construction of statistic distribution models for nonparametric goodness-of-fit tests in testing composite hypotheses: the computer approach // *Quality Technology & Quantitative Management*. – 2011. – Vol. 8 (4). – P. 359–373.
28. *Лемешко Б.Ю., Горбунова А.А.* О применении и мощности непараметрических критериев согласия Купера, Ватсона и Жанга // *Измерительная техника*. – 2013. – № 5. – С. 3–9.
29. *Лемешко Б.Ю., Горбунова А.А.* Применение непараметрических критериев согласия Купера и Ватсона при проверке сложных гипотез // *Измерительная техника*. – 2013. – № 9. – С. 14–21.
30. *Лемешко Б.Ю.* Непараметрические критерии согласия: руководство по применению. – М.: Инфра-М, 2014. – 163 с. – DOI: 10.12737/11873.
31. Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин «Интервальная статистика 5.4»: свидетельство о гос. регистрации программы для ЭВМ № 2018666213 / Лемешко Б.Ю., Лемешко С.Б., Блинов П.Ю., Веретельникова И.В., Новикова А.Ю. – Заявка № 2018663206; заявл. 22.11.2018; зарег. 13.12.2018. – URL: <https://ami.nstu.ru/~headrd/ISW.htm> (дата обращения: 16.11.2024).
32. *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.
33. *Лемешко Б.Ю.* О проблемах и ошибках применения критериев согласия // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. – 2023. – № 64. – С. 74–90. – DOI: 10.17223/19988605/64/8.
34. *Zhang J.* Powerful goodness-of-fit tests based on the likelihood ratio // *Journal of the Royal Statistical Society: Series B*. – 2002. – Vol. 64 (2). – P. 281–294.
35. *Zhang J., Wu Yu.* Likelihood-ratio tests for // *Computational Statistics & Data Analysis*. – 2005. – Vol. 49 (3). – P. 709–721.
36. *Лемешко Б.Ю., Лемешко С.Б., Семёнова М.А.* К вопросу статистического анализа больших данных // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. – 2018. – № 44. – С. 40–49. – DOI: 10.17223/19988605/44/5.
37. *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.
38. *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.
39. *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. – P. 012063. – DOI: 10.1088/1742-6596/1715/1/012063.
40. *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.
41. *Vasicek O.* A test for normality based on sample entropy // *Journal of the Royal Statistical Society. Series B, Methodological*. – 1976. – Vol. 38 (1). – P. 54–59. – DOI: 10.1111/j.2517-6161.1976.tb01566.x.
42. *Arizono I., Ohta H.* A test for normality based on Kullback-Leibler information // *The American Statistician*. – 1989. – Vol. 43 (1). – P. 20–22.
43. *Van Es B.* Estimating functionals related to a density by a class of statistics based on spacings // *Scandinavian Journal of Statistics*. – 1992. – Vol. 19 (1). – P. 61–72.

44. Ebrahimi N., Pflughoeft K., Soofi E.S. Two measures of sample entropy // *Statistics & Probability Letters*. – 1994. – Vol. 20 (3). – P. 225–234.
45. Correa J.C. A new estimator of entropy // *Communication in Statistics – Theory and Methods*. – 1995. – Vol. 24 (10). – P. 2439–2449.
46. Zamanzade E., Arghami N.R. Testing normality based on new entropy estimators // *Journal of Statistical Computation and Simulation*. – 2012. – Vol. 82 (11). – P. 1701–1713. – DOI: 10.1080/00949655.2011.592984.
47. Dudewicz E.J., Van der Meulen E.C. Entropy-based tests of uniformity // *Journal of the American Statistical Association*. – 1981. – Vol. 76 (376). – P. 967–974.
48. Zamanzade E. Testing uniformity based on new entropy estimators // *Journal of Statistical Computation and Simulation*. – 2015. – Vol. 85 (16). – DOI: 10.1080/00949655.2014.958085.
49. Лемешко Б.Ю. Критерии проверки отклонения распределения от нормального закона: руководство по применению. – 2-е изд., перераб. и доп. – М.: Инфра-М, 2023. – 353 с. – (Научная мысль). – DOI: 10.12737/1896110.
50. Лемешко Б.Ю., Блинов П.Ю. Критерии проверки отклонения распределения от равномерного закона: руководство по применению. – М.: Инфра-М, 2015. – 183 с. – (Научная мысль). – DOI: 10.12737/11304.
51. Noughabi H.A., Arghami N.R. General treatment of goodness-of-fit tests based on Kullback–Leibler information // *Journal of Statistical Computation and Simulation*. – 2012. – Vol. 83 (8). – P. 1556–1569. – DOI: 10.1080/00949655.2012.667100.
52. Noughabi H.A. A new estimator of Kullback–Leibler information and its application in goodness of fit tests // *Journal of Statistical Computation and Simulation*. – 2019. – Vol. 89 (10). – P. 1914–1934. – DOI: 10.1080/00949655.2019.1602870.
53. Alavi S.Q., Noughabi H.A. A novel estimator of Kullback–Leibler information with its application to goodness of fit tests // *Statistics, Optimization and Information Computing*. – 2025. – Vol. 13 (5). – P. 1800–1818. – DOI: 10.19139/soic-2310-5070-2032.
54. Лемешко Б.Ю., Лемешко С.Б. Сравнительный анализ критериев проверки отклонения распределения от нормального закона // *Метрология*. – 2005. – № 2. – С. 3–24.
55. Noughabi H.A. A comparative study of goodness-of-fit tests for the Gumbel distribution // *Statistica Applicata – Italian Journal of Applied Statistics*. – 2024. – P. 1–17. – DOI: 10.26398/IJAS.267.

Лемешко Борис Юрьевич, доктор технических наук, профессор, профессор кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Основное направление научных исследований – компьютерные технологии исследования статистических и вероятностных закономерностей. Имеет более 400 печатных работ, в том числе 20 монографий и учебных пособий. E-mail: lemeshko@ami.nstu.ru

Лемешко Станислав Борисович, кандидат технических наук, старший научный сотрудник Центра статистических технологий Новосибирского государственного технического университета. Основное направление научных исследований – компьютерные технологии моделирования и исследования статистических закономерностей. Имеет 69 печатных работ, в том числе одну монографию. E-mail: skyer@mail.ru

Lemeshko Boris Yu., D.Sc. (Eng.), professor, professor at the Department of Theoretical and Applied Informatics, Novosibirsk State Technical University. The main field of is scientific research is computer technologies for the study of statistical and probabilistic laws. He has more than 400 publications, including 20 monographs and textbooks. E-mail: lemeshko@ami.nstu.ru

Lemeshko Stanislav B., PhD (Eng.), senior researcher at Center for Statistical Technology, Novosibirsk State Technical University. The main field of his scientific research is computer modeling technologies and research of statistical laws. He has 69 publications, including 1 monograph. E-mail: skyer@mail.ru

On the properties and problems of applying goodness-of-fit tests based on the use of entropy and Kullback – Leibler divergence estimates*

B.Yu. LEMESHKO^{1, a}, *S.B. LEMESHKO*^{1, 2, b}

¹ *Novosibirsk State Technical University, 20, K. Marx Prospekt, Novosibirsk, 630073, Russian Federation*

² *Novosibirsk National Research State University, 1, Pirogov Street, Novosibirsk, 630090, Russian Federation*

^a *lemeshko@ami.nstu.ru* ^b *skyer@mail.ru*

Abstract

To verify the adequacy of the constructed models of distribution laws of random variables, various nonparametric goodness-of-fit tests are usually used, in particular, Kolmogorov, Kramer – Mises – Smirnov, Anderson – Darling, Kuiper, and Watson.

When a simple hypothesis is tested, nonparametric goodness-of-fit tests are “distribution-free”: the asymptotic distributions of statistics do not depend on the type of law against which the hypothesis is tested.

When testing composite hypotheses, when the parameters of the assumed law are estimated from a sample, the property of “freedom from distribution” is lost, and the distributions of statistics become dependent on a number of factors. In such situations, the use of nonparametric goodness-of-fit tests is possible only with the support of appropriate software that allows the achieved significance level P_v to be assessed using simulation modeling.

The distributions of the statistics of the Zhang tests, which are a development of the Kolmogorov, Kramer – Mises – Smirnov, and Anderson – Darling tests, respectively, depend on the sample sizes, so their wide application in testing simple and complex hypotheses is possible only with the support of the Monte Carlo method.

Distributions of goodness-of-fit tests statistics (when testing simple and composite hypotheses) can vary significantly due to the natural presence of rounding errors. A signal about the possibility of such a situation is the presence of a significant number of repeating values in the analyzed samples. In such situations, making a decision on the results of the inspection is also impossible without the use of simulation modeling.

In recent years, several criteria have been proposed, aimed, for example, at checking whether samples belong to a normal or uniform law, the statistics of which are based on various entropy estimates. As experience shows, with respect to some competing hypotheses, such criteria demonstrate higher power estimates compared to classical nonparametric goodness-of-fit tests.

When constructing the statistics of the Noughabi test to distinguish between two hypotheses, the Kullback-Leibler divergence was used, and the estimate proposed by Vasicek was taken as an estimate of entropy. This paper shows how the distributions of the Noughabi test statistics depend on the sample size n and the window size m , and how the distributions of the test statistics change when testing various composite hypotheses. The power of the criterion in testing normality against various competing hypotheses was investigated. It is shown how, for given n , the power depends on the size of the “window” m . The existence of an optimal m is shown, at which the power is maximum relative to the competing hypothesis under consideration. It is shown that for a given n , the optimal values of m , as a rule, do not coincide for different competing hypotheses. Obviously, the application of such criteria in practice also implies the use of appropriate software and simulation modeling.

* Received 12 February 2025.

Keywords: software system, parameter estimates, nonparametric goodness-of-fit tests, statistics distribution, achieved significance level, normality test, error of the 1st kind, error of the 2nd kind, power of test, simulation modeling

REFERENCES

1. Kolmogoroff A.N. Sulla determinazione empirica di una legge di distribuzione. *Giornale del Istituto Italiano degli Attuari*, 1933, vol. 4, no. 1, pp. 83–91.
2. Cramèr H. *Mathematical methods of statistics*. Princeton University Press, 1946 (Russ. ed.: Kramer G. *Matematicheskie metody statistiki*. Moscow, Mir Publ., 1975. 648 p.).
3. Mises R. von. *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Leipzig, Wien, Deuticke, 1931. 574 S.
4. Smirnov N. Sur la distribution de ω^2 . *Comptes Rendus de l'Académie des Sciences*. Paris, 1936, vol. 202, pp. 449–452.
5. Anderson T.W., Darling D.A. Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 1952, vol. 23 (2), pp. 193–212. DOI: 10.1214/aoms/1177729437.
6. Anderson T.W., Darling D.A. A test of goodness of fit. *Journal of the American Statistical Association*, 1954, vol. 49 (268), pp. 765–769. DOI: 10.1080/01621459.1954.10501232.
7. Kuiper N.H. Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Proceedings. Series A*, 1960, vol. 63, pp. 38–47.
8. Watson G.S. Goodness-of-fit tests on a circle. *Biometrika*, 1961, vol. 48 (1–2), pp. 109–114.
9. Watson G.S. Goodness-of-fit tests on a circle. II. *Biometrika*, 1962, vol. 49 (1–2), pp. 57–63.
10. Kac M., Kiefer J., Wolfowitz J. On tests of normality and other tests of goodness of fit based on distance methods. *The Annals of Mathematical Statistics*, 1955, vol. 26 (2), pp. 189–211. DOI: 10.1214/aoms/1177728538.
11. Lemesko B.Yu. *Neparametricheskie kriterii soglasiya. Rukovodstvo po primeneniyu* [Nonparametric Goodness-of-Fit Tests. Guide on the application]. 2nd ed., rev. Moscow, Infra-M Publ., 2024. 201 p. DOI 10.12737/2058731.
12. Lemesko B.Yu., Postovalov S.N. O zavisimosti raspredelenii statistik neparametricheskikh kriteriev i ikh moshchnosti ot metoda otsenivaniya parametrov [On the dependence of the distributions of statistics of nonparametric tests and their power on the method of parameter estimation]. *Zavodskaya laboratoriya. Diagnostika materialov = Industrial laboratory. Diagnostics of materials*, 2001, vol. 67, no. 7, pp. 62–71.
13. Lemesko B.Yu., Maklakov A.A. Nonparametric test in testing composite hypotheses on goodness of fit exponential family distributions. *Optoelectronics, Instrumentation and Data Processing*, 2004, vol. 40 (3), pp. 3–18. Translated from *Avtometriya*, 2004, no. 3, pp. 3–20.
14. Martynov G.V. *Kriterii omega-kvadrat* [Omega-square criteria]. Moscow, Nauka Publ., 1978. 80 p.
15. Tyurin Yu.N. O predel'nom raspredelenii statistik Kolmogorova – Smirnova dlya slozhnoi gipotezy [On the limit distribution of Kolmogorov–Smirnov statistics for a composite hypothesis]. *Izvestiya Akademii nauk SSSR. Seriya matematicheskaya = Mathematics of the USSR-Izvestiya*, 1984, vol. 48, no. 6, pp. 1314–1343. (In Russian).
16. Tyurin Yu.N., Savvushkina N.E. Kriterii soglasiya dlya raspredeleniya Veibulla – Gnedenko [Goodness of fit for the Weibull–Gnedenko distribution]. *Izvestiya Akademii nauk SSSR. Tekhnicheskaya kibernetika = Soviet Journal of Computer and Systems Sciences*, 1984, no. 3, pp. 109–112. (In Russian).
17. Lilliefors H.W. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 1967, vol. 62, pp. 399–402.
18. Lilliefors H.W. On the Kolmogorov–Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 1969, vol. 64, pp. 387–389.
19. Lemesko B.Yu., Postovalov S.N. Prikladnye aspekty ispol'zovaniya kriteriev soglasiya v sluchae proverki slozhnykh gipotez [Applied aspects of using goodness-of-fit criteria in the case of

testing complex hypotheses]. *Nadezhnost' i kontrol' kachestva = Reliability and Quality Control*, 1997, no. 11, pp. 3–17.

20. Lemesenko B.Yu., Postovalov S.N. Statistical distributions of nonparametric goodness-of-fit tests as estimated by the sample parameters of experimentally observed laws. *Industrial laboratory (Diagnostics of Materials)*, 1998, vol. 64, no. 3, pp. 197–208. Translated from *Zavodskaya laboratoriya. Diagnostika materialov*, 1998, vol. 64, no. 3, pp. 61–72.

21. Lemesenko B.Yu., Postovalov S.N. *Prikladnaya statistika. Pravila proverki soglasiya opytного raspredeleniya s teoreticheskim. Metodicheskie rekomendatsii. Ch. 2. Neparametricheskie kriterii* [Applied statistics. Rules for testing the agreement of the experimental distribution with the theoretical one. Methodological recommendations. Part 2. Nonparametric tests]. Novosibirsk, NSTU Publ., 1999. 85 p.

22. R 50.1.037–2002. *Rekomendatsii po standartizatsii. Prikladnaya statistika. Pravila proverki soglasiya opytного raspredeleniya s teoreticheskim. Ch. 2. Neparametricheskie kriterii* [Recommendations for standardization. Applied statistics. Rules for checking the agreement of the experimental distribution with the theoretical one. Pt. 2. Nonparametric tests]. Moscow, Standards Publ., 2002. 64 p.

23. Lemesenko B.Yu., Lemesenko S.B. Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. Pt. 1. *Measurement Techniques*, 2009, vol. 52 (6), pp. 555–565. Translated from *Izmeritel'naya tekhnika*, 2009, no. 6, pp. 3–11.

24. Lemesenko B.Yu., Lemesenko S.B. Models for statistical distributions in nonparametric fitting tests on composite hypotheses based on maximum-likelihood estimators. Pt. 2. *Measurement Techniques*, 2009, vol. 52 (8), pp. 799–812. Translated from *Izmeritel'naya tekhnika*, 2009, no. 8, pp. 17–26.

25. Lemesenko B.Yu., Lemesenko S.B., Postovalov S.N. Statistic distribution models for some nonparametric goodness-of-fit tests in testing composite hypotheses. *Communications in Statistics – Theory and Methods*, 2010, vol. 39 (3), pp. 460–471.

26. Lemesenko B.Yu., Lemesenko S.B. Models of statistic distributions of nonparametric goodness-of-fit tests in composite hypotheses testing for double exponential law cases. *Communications in Statistics. Theory and Methods*, 2011, vol. 40 (16), pp. 879–2892.

27. Lemesenko B.Yu., Lemesenko S.B. Construction of statistic distribution models for nonparametric goodness-of-fit tests in testing composite hypotheses: the computer approach. *Quality Technology & Quantitative Management*, 2011, vol. 8 (4), pp. 359–373.

28. Lemesenko B.Yu., Gorbunova A.A. Application and power of the nonparametric Kuiper, Watson, and Zhang tests of goodness-of-fit. *Measurement Techniques*, 2013, vol. 56 (5), pp. 465–475. DOI: 10.1007/s11018-013-0229-7. Translated from *Izmeritel'naya tekhnika*, 2013, no. 5, pp. 3–9.

29. Lemesenko B.Yu., Gorbunova A.A. Application of nonparametric Kuiper and Watson tests of goodness-of-fit for composite hypotheses. *Measurement Techniques*, 2013, vol. 56 (9), pp. 965–973. Translated from *Izmeritel'naya tekhnika*, 2013, no. 9, pp. 14–21.

30. Lemesenko B.Yu. *Neparametricheskie kriterii soglasiya: rukovodstvo po primeneniyu* [Nonparametric criteria of consent: application guide]. Moscow, Infra-M Publ., 2014. 163 p. DOI: 10.12737/11873.

31. Lemesenko B.Yu., Lemesenko S.B., Blinov P.Yu., Veretel'nikova I.V., Novikova A.Yu. Statisticheskii analiz interval'nykh nablyudenii odnomernykh nepreryvnykh sluchainykh velichin "Interval'naya statistika 5.4" [Statistical analysis of interval observations of one-dimensional continuous random variables "Interval statistics 5.4"]. The certificate on official registration of the computer program. No. 2018666213, 2018. Available at: <https://ami.nstu.ru/~headrd/ISW.htm> (accessed 22.11.2024).

32. Lemesenko B.Yu., Lemesenko S.B. Problemy primeneniya neparametricheskikh kriteriev soglasiya v zadachakh obrabotki rezul'tatov izmerenii [Problems of application of nonparametric criteria of agreement in problems of processing measurement results]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2021, no. 2 (82), pp. 47–66. DOI: 10.17212/2782-2001-2021-2-47-66.

33. Lemesenko B.Yu. O problemakh i oshibkakh primeneniya kriteriev soglasiya [On the problems and errors in the application of consent criteria]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika = Tomsk State University Journal of Control and Computer Science*, 2023, vol. 64, pp. 74–90. DOI: 10.17223/19988605/64/8.

34. Zhang J. Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B*, 2002, vol. 64 (2), pp. 281–294.
35. Zhang J., Wu Yu. Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis*, 2005, vol. 49 (3), pp. 709–721.
36. Limeshko B.Yu., Limeshko S.B., Semenova M.A. K voprosu statisticheskogo analiza bol'shikh dannykh [To question of the statistical analysis of big data]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika = Tomsk State University Journal of Control and Computer Science*, 2018, vol. 44, pp. 40–49. DOI: 10.17223/19988605/44/5.
37. Limeshko B.Yu., Limeshko S.B. Effect of the Roundoff on the properties of criteria for testing statistical hypotheses. *Optoelectronics, Instrumentation and Data Processing*, 2020, vol. 56 (3), pp. 242–250. DOI: 10.3103/S8756699020030103. Translated from *Avtometriya*, 2020, vol. 56, no. 3, pp. 35–45. DOI: 10.15372/AUT20200305.
38. Limeshko B.Yu., Limeshko S.B. O vliyani oshibok okrugleniya na raspredeleniya statistik kriteriev soglasiya [About the influence of rounding errors on distributions of statistics of the goodness-of-fit tests]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika = Tomsk State University Journal of Control and Computer Science*, 2020, vol. 53, pp. 47–60. DOI: 10.17223/19988605/53/5.
39. Limeshko B.Y., Limeshko S.B. About the effect of rounding on the properties of tests for testing statistical hypotheses. *Journal of Physics: Conference Series*, 2021, vol. 1715, p. 012063. DOI: 10.1088/1742-6596/1715/1/012063.
40. Limeshko B.Yu., Limeshko S.B. Neparametricheskie kriterii soglasiya pri proverke normal'nosti v usloviyakh okrugleniya izmerenii [Nonparametric goodness-of-fit tests for normality testing under rounding-off measurements]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2022, no. 2 (86), pp. 21–38. DOI: 10.17212/2782-2001-2022-2-21-38.
41. Vasicek O. A test for normality based on sample entropy. *Journal of the Royal Statistical Society, Series B, Methodological*, 1976, vol. 38 (1), pp. 54–59. DOI: 10.1111/j.2517-6161.1976.tb01566.x.
42. Arizono I., Ohta H. A test for normality based on Kullback-Leibler information. *The American Statistician*, 1989, vol. 43 (1), pp. 20–22.
43. Van Es B. Estimating functionals related to a density by a class of statistics based on spacings. *Scandinavian Journal of Statistics*, 1992, vol. 19 (1), pp. 61–72.
44. Ebrahimi N., Pflughoeft K., Soofi E.S. Two measures of sample entropy. *Statistics & Probability Letters*, 1994, vol. 20 (3), pp. 225–234.
45. Correa J.C. A new estimator of entropy. *Communication in Statistics – Theory and Methods*, 1995, vol. 24 (10), pp. 2439–2449.
46. Zamanzade E., Arghami N.R. Testing normality based on new entropy estimators. *Journal of Statistical Computation and Simulation*, 2012, vol. 82 (11), pp. 1701–1713. DOI: 10.1080/00949655.2011.592984.
47. Dudewicz E.J., Van der Meulen E.C. Entropy-based tests of uniformity. *Journal of the American Statistical Association*, 1981, vol. 76 (376), pp. 967–974.
48. Zamanzade E. Testing uniformity based on new entropy estimators. *Journal of Statistical Computation and Simulation*, 2015, vol. 85 (16). – DOI: 10.1080/00949655.2014.958085.
49. Limeshko B.Yu. *Kriterii proverki otkloneniya raspredeleniya ot normal'nogo zakona: rukovodstvo po primeneniyu* [Criteria for checking the deviation of the distribution from the normal law. Application Guide]. 2nd ed., rev. Moscow, Infra-M Publ., 2023. 353 p. DOI: 10.12737/1896110.
50. Limeshko B.Yu., Blinov P.Yu. *Kriterii proverki otkloneniya raspredeleniya ot ravnomernogo zakona: rukovodstvo po primeneniyu* [Tests for checking the deviation from uniform distribution law. Guide on the application]. Moscow, Infra-M Publ., 2015. 183 p. DOI: 10.12737/11304.
51. Noughabi H.A., Arghami N.R. General treatment of goodness-of-fit tests based on Kullback–Leibler information. *Journal of Statistical Computation and Simulation*, 2012, vol. 83 (8), pp. 1556–1569. DOI: 10.1080/00949655.2012.667100.
52. Noughabi H.A. A new estimator of Kullback–Leibler information and its application in goodness of fit tests. *Journal of Statistical Computation and Simulation*, 2019, vol. 89 (10), pp. 1914–1934. DOI: 10.1080/00949655.2019.1602870.

53. Alavi S.Q., Noughabi H.A. A novel estimator of Kullback–Leibler information with its application to goodness of fit tests. *Statistics, Optimization and Information Computing*, 2025, vol. 13 (5), pp. 1800–1818. DOI: 10.19139/soic-2310-5070-2032.

54. Lemesko B.Yu., Lemesko S.B. Sravnitel'nyi analiz kriteriev proverki otkloneniya raspredeleniya ot normal'nogo zakona [Comparative analysis of criteria for testing the deviation of distribution from the normal law]. *Metrologiya = Metrology*, 2005, no. 2, pp. 3–24. (In Russian).

55. Noughabi H.A. A comparative study of goodness-of-fit tests for the Gumbel distribution. *Statistica Applicata – Italian Journal of Applied Statistics*, 2024, pp. 1–17. DOI: 10.26398/IJAS.267.

Для цитирования:

Лемешко Б.Ю., Лемешко С.Б. О свойствах и проблемах применения критериев согласия, опирающихся на использование оценок энтропии и дивергенции Кульбака–Лейблера // Системы анализа и обработки данных. – 2025. – № 2 (98). – С. 53–80. – DOI: 10.17212/2782-2001-2025-2-53-80.

For citation:

Lemesko B.Yu., Lemesko S.B. O svoistvakh i problemakh primeneniya kriteriev soglasiya, opirayushchikhsya na ispol'zovanie otsenok entropii i divergentsii Kul'baka – Leiblera [On the properties and problems of applying goodness-of-fit tests based on the use of entropy and Kullback–Leibler divergence estimates]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2025, no. 2 (98), pp. 53–80. DOI: 10.17212/2782-2001-2025-2-53-80.