

Научная статья

УДК 519.24

doi: 10.17223/19988605/64/8

О проблемах и ошибках применения критериев согласия

Борис Юрьевич Лемешко

Новосибирский государственный технический университет, Новосибирск, Россия, lemeshko@ami.nstu.ru

Аннотация. Рассматриваются изменения распределений статистик различных критериев согласия. Показывается, начиная с каких малых объемов выборок можно пренебречь отклонением распределений статистик от асимптотических в случае проверки простых гипотез. Приводятся факторы, под влиянием которых изменяются распределения статистик критериев при проверке сложных гипотез. Показывается, что при отсутствии ошибок округления применение критериев в условиях больших выборок не вызывает проблем. В этих же условиях наличие ошибок округления приводит к отклонению распределений статистик от асимптотических. При соизмеримости ошибок округления и ошибок измерений распределения статистик критериев могут не сходиться к асимптотическим и становятся зависящими от объемов выборок и от других факторов.

Ошибки применения критериев согласия в приложениях связаны не только с незнанием свойств критериев в конкретных (нестандартных) условиях, но и с присутствием ошибочных рекомендаций в литературе, в том числе учебного характера. В популярных программных системах статистического анализа, позволяющих применять рассматриваемые критерии, реальные свойства критериев, имеющие место в конкретной ситуации, как правило, не учитываются.

Показано, что корректность статистических выводов при использовании рассматриваемых критериев в нестандартных ситуациях можно гарантировать только за счет применения методов статистического моделирования в рамках соответствующего программного обеспечения.

Ключевые слова: непараметрические критерии согласия; статистическое моделирование; ошибка 1-го рода; ошибка 2-го рода; ошибки округления.

Для цитирования: 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

Original article

doi: 10.17223/19988605/64/8

About the problems and errors in applying the goodness of fit tests

Boris Yu. Lemesko

Novosibirsk State Technical University, Novosibirsk, Russian Federation, lemeshko@ami.nstu.ru

Abstract. Changes in the distributions of statistics for various goodness-of-fit tests are considered. It is shown, starting from what small sample sizes, it is possible to neglect the deviation of statistics distributions from asymptotic ones in the case of testing simple hypotheses. Factors are given, under the influence of which the distributions of test statistics change when testing composite hypotheses. It is shown that in the absence of rounding errors, the application of the tests in conditions of large samples does not cause problems. Under the same conditions, the presence of rounding errors leads to a deviation of the distributions of statistics from asymptotic ones. When rounding errors and measurement errors are commensurable, the distributions of test statistics may not converge to asymptotic ones and become dependent on sample sizes and other factors.

Errors in the application of goodness of fit criteria in applications are associated not only with ignorance of the properties of the tests in specific (non-standard) conditions, but also with the presence of erroneous recommendations in the literature, including educational ones. In popular statistical analysis software systems that allow the application of the tests under consideration, the real properties of the tests that take place in a particular situation, as a rule, are not taken into account.

It is shown that the correctness of statistical conclusions when using the tests under consideration in non-standard situations can only be guaranteed through the use of statistical simulating methods within the framework of the appropriate software.

Keywords: nonparametric goodness of fit tests; statistical simulating; type 1 error; type 2 error; rounding errors.

For citation: Lemesko, B.Yu. (2023) About the problems and errors in applying the goodness of fit tests. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 64. pp. 74–90. doi: 10.17223/19988605/64/8

Введение

Курс математической статистики, читаемый в университетах для инженерных специальностей, содержит лишь краткое изложение основ дисциплины и не дает представления о проблемах применения соответствующих методов на практике. Статистические методы и критерии обладают известными (асимптотическими) свойствами при выполнении определенных (стандартных) предположений. В приложениях эти предположения могут не выполняться, а в нестандартных условиях свойства применяемых методов и критериев могут существенно изменяться. В литературе учебного характера внимание на нюансах такого характера практически не акцентируется. В учебниках можно встретить неверные рекомендации по использованию и примеры некорректного применения различных критериев проверки гипотез. Не всегда корректно используются критерии проверки гипотез и в популярных программных системах.

Указанные причины часто приводят к неэффективному применению методов статистического анализа в приложениях и, более того, к неверным статистическим выводам. Такие случаи формируют отрицательное мнение относительно возможностей статистических методов и сокращают потенциальную область их применения, где в результате обедняется используемый в исследованиях математический аппарат. В качестве примера можно назвать сферу метрологии, где в силу объективных и субъективных причин преобладает негативное отношение к полезности методов статистического анализа вообще и к возможностям критериев согласия в частности.

В данном случае речь пойдет об изменении свойств критериев согласия (об изменении распределений статистик) под воздействием различных факторов. В таких ситуациях использование классических результатов, касающихся распределений статистик критериев, как правило, приводит к увеличению вероятностей ошибок или первого, или второго рода. А гарантировать корректность выводов можно только за счет использования методов статистического моделирования для исследования распределений статистик критериев в нестандартных условиях, что можно делать, опираясь на соответствующее программное обеспечение.

1. Непараметрические критерии согласия при проверке простых гипотез

Критерий **Колмогорова** опирается на статистику

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

где $F_n(x)$ – эмпирическая функция распределения; $F(x, \theta)$ – теоретическая функция распределения; n – объем выборки. Предельное распределение этой статистики для случая проверки простой гипотезы было получено Колмогоровым в [1]. При $n \rightarrow \infty$ функция распределения статистики $\sqrt{n} \cdot D_n$ сходится равномерно к функции распределения Колмогорова

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

Распределение статистики $\sqrt{n} \cdot D_n$ достаточно медленно сходится к асимптотическому. Поэтому при проверке гипотез с применением критерия Колмогорова рекомендуется использовать статистику с поправкой Большева [2, 3] в форме [4]

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

где $D_n = \max(D_n^+, D_n^-)$, $D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}$, $D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}$, n – объем выборки; $x_1, x_2, \dots, x_n$ здесь и далее – упорядоченные по возрастанию выборочные значения; $F(x, \theta)$ – функция закона распределения, согласие с которым проверяют. Распределение величины S_K при простой гипотезе в пределе подчиняется закону Колмогорова с функцией распределения $K(S)$.

Критерий Колмогорова является наиболее часто используемым из непараметрических критериев согласия. К сожалению, в подавляющем числе публикаций, в том числе учебного характера, а также в программных системах статистика критерия используется без поправки. В таких случаях отклонение реального распределения статистики $\sqrt{n} \cdot D_n$ от $K(S)$ заметно при величинах n порядка 100. Поэтому вычисление достигнутого уровня значимости p_{value} по распределению $K(S)$ при меньших значениях n приводит к увеличению вероятности ошибки второго рода β .

В ряде источников учебного характера, например [5, 6], в примерах, иллюстрирующих применение критерия Колмогорова, статистика D_n , в отличие от предусмотренного в [1], вычисляется по группированным данным. В обоснование ссылаются на программную систему STATISTICA, где предусмотрена такая возможность. Это неверная рекомендация, которая всегда приводит к меньшим значениям статистики и резко увеличивает вероятность ошибки второго рода, так как в этом случае асимптотическим распределением статистики $\sqrt{n} \cdot D_n$ уже не будет являться распределение $K(S)$.

В критерии типа ω^2 **Крамера–Мизеса–Смирнова** статистика имеет вид:

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

и при справедливости простой проверяемой гипотезы в пределе подчиняется закону с функцией распределения $a1(s)$, приведенной в [4].

В критерии типа Ω^2 **Андерсона–Дарлинга** [7, 8] используется статистика

$$S_\Omega = -n - 2 \sum_{i=1}^n \left\{ \frac{2i-1}{2n} \ln F(x_i, \theta) + \left(1 - \frac{2i-1}{2n} \right) \ln(1 - F(x_i, \theta)) \right\}. \quad (3)$$

При проверке простой гипотезы в пределе эта статистика подчиняется закону с функцией распределения $a2(s)$ [4].

В критерии **Купера** [9] в качестве меры расхождения используется величина $V_n = D_n^+ + D_n^-$, где D_n^+ , D_n^- находятся так же, как и в критерии Колмогорова, а в качестве статистики используется

$$s = \sqrt{n} \cdot V_n. \quad (4)$$

Купером [9] в качестве предельного распределения $G(s|H_0)$ для случая проверки простой гипотезы дана следующая функция распределения [10]

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

Для статистики $\sqrt{n} \cdot V_n$ достаточно долго сохраняется зависимость от объема выборки n . В [11] построена модификация статистики

$$V = V_n \left(\sqrt{n} + 0,155 + \frac{0,24}{\sqrt{n}} \right),$$

для которой зависимостью распределения от объема выборки можно пренебречь при $n \geq 20$.

В [12] предложено применять в критерии Купера статистику вида:

$$V_n^{mod} = \sqrt{n} \cdot V_n + \frac{1}{3\sqrt{n}},$$

зависимостью распределения которой от объема выборки можно пренебречь при $n \geq 30$.

Статистика критерия **Ватсона** [13, 14] используется в следующей форме:

$$U_n^2 = \sum_{i=1}^n \left(F(x_i, \theta) - \frac{i-1/2}{n} \right)^2 - n \left(\frac{1}{n} \sum_{i=1}^n F(x_i, \theta) - \frac{1}{2} \right)^2 + \frac{1}{12n}. \quad (5)$$

Асимптотическим распределением статистики U_n^2 при проверке простой гипотезы является следующее распределение:

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

Зависимость распределения этой статистики от объема выборки выражена слабо. В [15] приведена модифицированная статистики Ватсона в форме

$$U_n^{2*} = (U_n^2 - 0.1/n + 0.1/n^2)(1 + 0.8/n),$$

отличием распределения которой от асимптотического можно пренебречь при $n \geq 20$.

На основании исследования распределений статистик критериев методами статистического моделирования в табл. 1 указаны объемы выборок, начиная с которых реальные распределения статистик (1)–(5) при справедливости простой проверяемой гипотезы практически (визуально) не отличаются от асимптотических, а также приведены объемы выборок, при которых отклонение этих распределений от асимптотических не отражается на статистических выводах. Последнее объясняется тем, что все эти критерии правосторонние, и при имеющемся отличии реального распределения статистики от асимптотического на правом «хвосте» эти распределения практически совпадают.

Таблица 1

Сходимость распределений статистик к асимптотическим

Тест	Статистика	Нет отклонения от асимптотического распределения	Можно пренебречь отклонением
Колмогорова (K)	(1)	при $n \geq 20$	при $n \geq 15$
Крамера–Мизеса–Смирнова (CMS)	(2)	при $n \geq 18$	при $n \geq 12$
Андерсона–Дарлинга (AD)	(3)	при $n \geq 10$	при $n \geq 6$
Купера (Ku)	(4)	при $n \geq 55$	при $n \geq 10$
Ватсона (W)	(5)	при $n \geq 75$	при $n \geq 10$

В диссертации **Жанга** и его последующих работах [16, 17] были предложены непараметрические критерии согласия, статистики которых имеют вид:

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

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

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

Справедливость утверждений автора критериев о их более высокой мощности по сравнению с критериями Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга была подтверждена исследованиями, проведенными в [12, 18]. Однако широкому применению критериев Жанга в приложениях препятствует зависимость распределений статистик от объема выборки n , которая, естественно, сохраняется и в ситуации проверки сложных гипотез.

2. Проблемы применения критериев при проверке сложных гипотез

При проверке сложных гипотез, когда по этой же выборке оцениваются параметры наблюдаемого закона распределения, все непараметрические критерии согласия теряют свойство «свободы от распределения» [19], и асимптотические распределения статистик $G(S|H_0)$ становятся зависящими от факторов, определяющих «сложность» гипотезы. В частности, на распределение статистики $G(S|H_0)$ критерия влияют следующие факторы:

- (a) – вид закона распределения $F(x, \theta)$, соответствующего проверяемой гипотезе H_0 ;
- (b) – тип оцениваемого параметра;
- (c) – число оцениваемых параметров;
- (d) – используемый метод оценивания параметров [20];
- (e) – в некоторых ситуациях распределения статистик зависят от конкретных значений параметров закона $F(x, \theta)$ (например, от значений параметров формы обобщенного нормального закона [21], гамма-распределения, бета-распределений и т.п.).

Хотя проблемы, связанные с применением непараметрических критериев согласия для проверки сложных гипотез, известны давно [19], с ними крайне редко знакомят в курсах лекций или освещают в учебной литературе. В то же время отличия в асимптотических распределениях тех же самых статистик при проверке простых и (различных) сложных гипотез настолько существенны, что пренебрегать этим абсолютно недопустимо. Большинство самых грубых и наиболее частых ошибок применения непараметрических критериев согласия, приводящих к существенному увеличению вероятностей ошибок 2-го рода, связано именно с незнанием этих обстоятельств.

Разрешение проблем, связанных с проверкой сложных гипотез аналитическими методами, связано с серьезными трудностями, а удачные попытки ряда авторов приводили и приводят к положительным результатам лишь частного характера. В то же время успешное использование в этих целях методов статистического моделирования, впервые продемонстрированное в [22, 23], позволяет получать полезные результаты с меньшими затратами. На базе результатов статистического моделирования распределений статистик критериев, имеющих место при проверке различных сложных гипотез, и построенных для них моделей были разработаны рекомендации Госстандарта РФ [24]. Дальнейшее развитие средств вычислительной техники и программного обеспечения исследований позволили уточнить модели и таблицы критических значений распределений статистик непараметрических критериев согласия [25–27], представленные ранее в [24], и подготовить руководство [28].

Возможность построения моделей для асимптотических распределений $G(S|H_0)$ статистик критериев ограничено множеством тех параметрических законов $F(x, \theta)$, в случае которых проявляется влияние только комбинаций факторов (a)–(d). При наличии влияния фактора (e) возможность исследования неизвестного распределения статистики $G(S|H_0)$ или оценки по этому распределению достигнутого уровня значимости p_v появляется только после нахождения оценок тех параметров, от значений которых зависит $G(S|H_0)$. В такой ситуации применение критерия и реализация процедуры интерактивного исследования $G(S|H_0)$ возможны только в рамках соответствующего программного обеспечения. Таким примером является программная система [29]. При зависимости распределений статистик $G(S|H_0)$ от объемов выборок, а это мы имеем в случае применения критериев Жанга со статистиками (6)–(8), к интерактивному моделированию распределений статистик критериев приходится прибегать при любой комбинации факторов (a)–(e), определяющих сложность гипотезы.

3. Применение критериев при малых объемах выборок

Насколько быстро распределения статистик непараметрических критериев согласия в случае проверки сложных гипотез сходятся к асимптотическим (при их существовании), можно судить по

результатам, приведенным в табл. 1. При малых n распределения статистик также несколько отличаются от асимптотических. Значения p_v , вычисляемые по асимптотическим распределениям $G(S|H_0)$, как правило, несколько превышают реальные значения. Максимальная величина этой погрешности при задаваемом уровне значимости $\alpha = 0,1$ и $n \geq 30$ не превышает 0,005. При необходимости всегда имеется возможность, как и в [29], уточнить оценку p_v , воспользовавшись методами статистического моделирования.

Проблему применения критериев согласия при малых выборках видят в другом. Обычно критерии используют, чтобы подобрать некоторую удобную параметрическую модель закона $F(x, \theta)$. При малых n , как правило, не удастся однозначно идентифицировать вид модели $F(x, \theta)$, так как не отклоняются гипотезы о принадлежности выборки целому ряду законов. Причина в низкой мощности критериев согласия при малых n . Более того, при малых n (и малых α) можно столкнуться с ситуацией смещенности критериев относительно некоторых пар конкурирующих гипотез, когда мощность критерия $1 - \beta$, где β – вероятность ошибки 2-го рода (не отклонить проверяемую гипотезу при справедливости конкурирующей), оказывается меньше α [28, 30].

4. Критерии согласия при анализе больших выборок

В настоящее время значительный интерес вызывают вопросы анализа больших данных (Big Data) и извлечения из этих данных скрытых закономерностей. В процессе анализа больших выборок, в том числе для идентификации моделей законов распределения, применяют и критерии согласия. От специалистов, особенно занимающихся анализом данных из области экономики и опирающихся на свой опыт применения, можно услышать, что проверяемая гипотеза H_0 о принадлежности выборки любому закону $F(x, \theta)$ с ростом n всегда отклоняется: критерии не работают (!). Так ли это? Рассмотрим несколько примеров.

Пример 1. Будем моделировать выборки в соответствии со стандартным нормальным законом, проверяя простую гипотезу H_0 о принадлежности этих выборок моделируемому закону. Результаты проверки (значения статистик S и достигнутые уровни значимости p_v) по пяти критериям с известными асимптотическими распределениями статистик $G(S|H_0)$ приведены в табл. 2.

Как и следовало ожидать, с ростом объемов выборок n применяемые критерии не дают никаких оснований для отклонения справедливой гипотезы H_0 .

Таблица 2

Результаты проверки простой гипотезы о принадлежности выборок, смоделированных по стандартному нормальному закону, этому же закону

Test	$n = 10^2$		$n = 10^3$		$n = 10^4$		$n = 10^5$		$n = 10^6$		$n = 10^7$	
	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v
K	0,666	0,766	0,634	0,816	0,636	0,814	0,812	0,525	0,660	0,776	0,647	0,796
Ku	1,142	0,622	1,138	0,629	1,104	0,680	1,306	0,385	0,962	0,866	0,814	0,973
CMS	0,063	0,798	0,072	0,736	0,059	0,818	0,145	0,404	0,080	0,689	0,097	0,601
AD	0,384	0,864	0,568	0,679	0,565	0,682	1,209	0,264	0,611	0,638	0,585	0,662
W	0,053	0,668	0,067	0,523	0,059	0,610	0,068	0,512	0,030	0,926	0,034	0,887

Пример 2. Логистический закон распределения с плотностью

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

достаточно близок к нормальному, и с использованием критериев согласия при малых объемах выборок его достаточно сложно отличить от нормального.

Будем моделировать выборки в соответствии с логистическим законом (для определенности с параметрами $\theta_0 = 0$, $\theta_1 = 1$), проверяя простую или сложную гипотезу H_0 о принадлежности этих

выборки нормальному закону. В случае простой гипотезы H_0 рассматривается стандартный нормальный закон. Результаты проверки приведены в табл. 3.

Таблица 3

**Результаты проверки простой и сложной гипотез о нормальности
в случае принадлежности выборок логистическому закону**

	При проверке простой гипотезы						При проверке сложной гипотезы					
	$n = 10^2$		$n = 10^3$		$n = 10^4$		$n = 10^2$		$n = 10^3$		$n = 10^4$	
	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v
K	0,701	0,710	1,293	0,071	2,293	0	0,701	0,295	1,155	0,004	2,605	0
Ku	1,164	0,589	2,340	0,001	4,499	0	1,293	0,410	1,845	0,027	5,205	0
CMS	0,090	0,638	0,436	0,058	1,915	0	0,096	0,124	0,253	0,001	2,706	0
AD	0,634	0,616	2,744	0,037	13,08	0	0,675	0,078	1,573	0,000	17,29	0
W	0,082	0,393	0,433	0,000	1,914	0	0,096	0,305	0,235	0,020	2,698	0

В данном примере гипотеза H_0 неверна. Мы видим, что в случае простой гипотезы H_0 она отклоняется по всем критериям при $\alpha > 0,071$ и $n \geq 10^3$, а в случае сложной также по всем критериям и том же n отклоняется при $\alpha > 0,027$.

Пример 3. Пусть наблюдаемый закон представляет собой смесь двух нормальных законов: $F(x, \theta) = 0,9 \times N(0, 1) + 0,1 \times N(0, 1, 1)$, где 90% составляют наблюдения, принадлежащие стандартному нормальному закону. Проверяется простая гипотеза H_0 о принадлежности моделируемых выборок различного объема n стандартному нормальному закону. Результаты проверки представлены в табл. 4.

Таблица 4

**Результаты проверки простой гипотезы о принадлежности выборок
со смесями стандартному нормальному закону**

Test	$n = 10^2$		$n = 10^3$		$n = 10^4$		$n = 10^5$		$n = 10^6$	
	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v
K	0,706	0,701	0,550	0,923	1,015	0,254	1,698	0,006	4,188	0
Ku	1,146	0,616	1,073	0,724	1,387	0,285	1,976	0,012	4,207	0
CMS	0,0707	0,747	0,060	0,815	0,178	0,313	0,615	0,021	7,999	0
AD	0,418	0,830	0,543	0,703	1,150	0,287	2,761	0,036	40,61	0
W	0,055	0,645	0,060	0,599	0,084	0,381	0,238	0,018	1,264	0

В данном случае верна конкурирующая гипотеза H_1 , соответствующая рассматриваемой смеси и очень близкая к H_0 . Мы можем видеть, что при $n = 10^4$ гипотеза H_0 еще не отклоняется ни по одному из критериев, а при $n \geq 10^5$ и задании вероятности ошибки первого рода $\alpha = 0,04$ – отклоняется по всем.

Оценки мощности $1 - \beta$ критериев относительно H_1 приведены в табл. 5. Эти оценки подтверждают закономерность получения приводимых в табл. 4 результатов проверки.

Таблица 5

Оценки мощности критериев относительно H_1 при $\alpha = 0,05$

Test	$n = 10^2$	$n = 10^3$	$n = 10^4$	$n = 10^5$	$n = 2 \times 10^5$	$n = 10^6$
K	0,0549	0,0558	0,1376	0,7807	0,9743	1
Ku	0,0511	0,0514	0,0838	0,5319	0,8773	1
CMS	0,0532	0,0616	0,1527	0,8503	0,9886	1
AD	0,0528	0,0625	0,1587	0,8670	0,9915	1
W	0,0523	0,0540	0,0872	0,4864	0,8813	1

Подчеркнем, что в данном примере закон распределения в виде смеси, соответствующий конкурирующей гипотезе H_1 , более близок к стандартному нормальному закону, рассматриваемому в качестве H_0 при простой гипотезе, чем логистический закон в той же роли в предшествующем примере. Именно поэтому в данном случае H_0 отклоняется при существенно больших объемах выборок.

Обратим внимание, что проверка сложных гипотез о принадлежности тех же выборок со смесями нормальному закону (с вычислением по выборкам оценок максимального правдоподобия (ОМП) параметров) свидетельствует о том, что получаемые смеси очень хорошо описываются нормальными законами (табл. 6).

Таблица 6

Результаты проверки сложных гипотез о принадлежности выборок со смесями нормальным законам

Test	$n = 10^2$		$n = 10^3$		$n = 10^4$		$n = 10^5$		$n = 10^6$		$n = 10^7$	
	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v	S	p_v
K	0,572	0,633	0,591	0,578	0,503	0,820	0,997	0,020	0,693	0,311	0,407	0,974
Ku	1,112	0,670	0,996	0,826	0,839	0,962	1,444	0,231	1,067	0,732	0,775	0,959
CMS	0,052	0,482	0,063	0,352	0,027	0,887	0,119	0,062	0,048	0,547	0,023	0,942
AD	0,343	0,493	0,432	0,308	0,207	0,870	0,675	0,078	0,283	0,656	0,146	0,980
W	0,051	0,695	0,052	0,683	0,024	0,970	0,105	0,256	0,043	0,790	0,023	0,931

В частности, для смоделированной выборки объемом $n = 10^7$ ОМП параметров нормального закона принимают значения $\hat{\mu} = 0,010273$, $\hat{\sigma} = 1,00063$, а достигнутые уровни значимости p_v , приведенные в последней колонке, свидетельствуют об очень хорошем согласии. В то же время результаты проверки простой гипотезы о согласии этой же выборки со стандартным нормальным законом дают по всем критериям нулевые значения p_v .

Заметим, что множество параметрических моделей законов распределения, где-либо и когда-либо использованных для описания реально наблюдаемых случайных величин, представляет собой крайне ограниченный перечень, а множество законов, описывающих эти величины, бесконечно. Вид конкретного закона для ряда (ошибок) измерений наблюдаемой величины оказывается известен крайне редко. В большинстве случаев мы имеем дело с приближенными моделями, в той или иной мере отличающимися от истинного закона. В такой ситуации с ростом n проверяемая гипотеза о согласии даже с очень хорошей моделью, естественно, будет отклоняться. В то же время справедливая гипотеза H_0 отклоняться не должна! Однако именно с этим приходится сталкиваться, если опираться на классические результаты, характеризующие свойства критериев согласия при проверке простых и сложных гипотез. В чем же причина частой некорректности выводов при использовании критериев согласия для анализа больших выборок в приложениях?

Приведенные выше примеры, в которых рассматриваются выборки моделируемых непрерывных случайных величин, показывают отсутствие проблем принципиального характера. В отличие от того, что мы наблюдаем на практике в больших данных, в выборках, смоделированных в соответствии с приводимыми примерами, отсутствовали повторяющиеся значения. В Big Data мы имеем дело с выборками очень большого объема, но данные в таких выборках представлены с некоторой ошибкой округления Δ . Поэтому в них всегда присутствует значительная доля повторяющихся значений. Наличие ошибок округления отражается на распределениях статистик непараметрических критериев согласия.

На рис. 1 (без потери общности, при $\Delta = 10^{-3}\sigma$) показано, как с ростом n меняются распределения $G(S_n|H_0)$ статистики (3) критерия Андерсона–Дарлинга при проверке простой гипотезы о принадлежности выборок нормальному закону. При таком Δ в выборках из нормального закона может присутствовать порядка 9 830 уникальных значений. Как видим, при $n = 10^3$ распределение $G(S_n|H_0)$ еще не отличается от асимптотического (предельного) $a1(S)$, а при $n = 10^4$ отклонение начинает проявляться. Если при проверке опираться на распределение $a1(S)$, то очевидно, что при больших n справедливая гипотеза H_0 будет отвергаться.

Таким образом, при проверке простых гипотез наличие ошибок округления приводит к зависимости распределений статистик критериев от объемов выборок n и от величины Δ . При проверке сложных гипотез зависимость от n и Δ накладывается на изменения распределений статистик критериев, связанные с влиянием факторов (а)–(е), определяющих различие в сложных гипотезах.

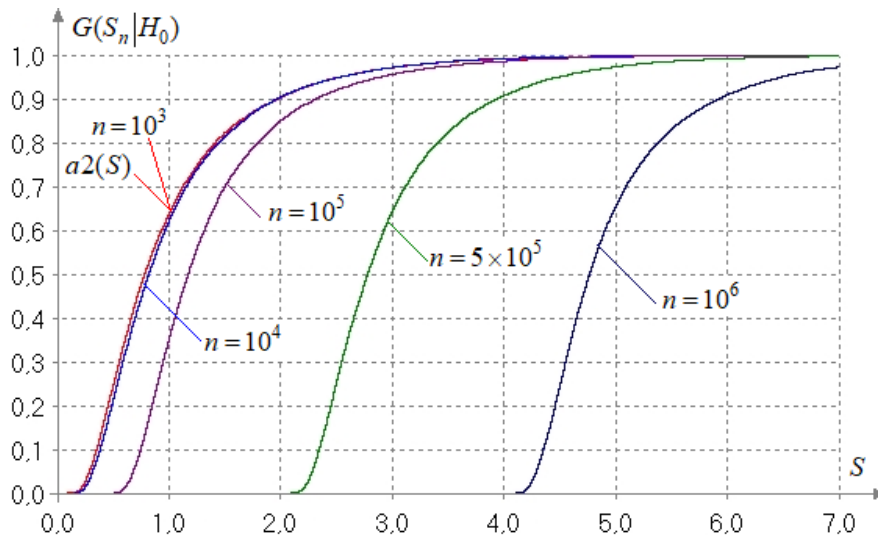


Рис. 1. Зависимость распределения статистики (3) критерия Андерсона–Дарлинга от объема выборки n при справедливости простой гипотезы H_0 о принадлежности выборки нормальному закону при $\Delta = 10^{-3}\sigma$
 Fig. 1. Dependence of the distribution of statistics (3) of the Anderson–Darling test on the sample size n under the validity of the simple hypothesis H_0 that the sample belongs to the normal law for $\Delta = 10^{-3}\sigma$

Чтобы можно было использовать классические результаты (при простых или сложных гипотезах), применяя соответствующие критерии, в [31] рекомендуется извлекать для анализа из Big Data выборки объемом не более $n_{\max}$, при котором распределение статистики $G(S_{n_{\max}}|H_0, \Delta)$ еще практически не отличается от асимптотического $G(S|H_0)$. В противном случае для обеспечения корректности выводов по результатам применения критерия потребуется нахождение реального распределения статистики $G(S_n|H_0, \Delta)$. Величина $n_{\max}$ зависит от Δ . Для ситуации, приведенной на рис. 1, в качестве $n_{\max}$ можно взять $n = 10^4$. Оценка $n_{\max}$ трудности не представляет, так как очевидно, что она не должна превышать количества уникальных значений случайной величины, которое может присутствовать в выборках из данного закона при данной величине Δ , уменьшенного в 3–4 раза.

При выполнении условия $\Delta \ll \sigma$ возможные изменения распределений статистик критериев следует предусматривать только при значительных объемах выборок, когда в выборках существенно возрастает доля повторяющихся значений.

5. Распределения статистик критериев при относительно больших ошибках округления

В приложениях различного характера достаточно часто приходится сталкиваться с выборками, полученными в том числе в результате высокоточных измерений, элементы которых отличаются друг от друга буквально единицами последнего десятичного разряда. В таких измерениях «на пределе точности» ошибка округления Δ оказывается достаточно близкой к σ ошибки измерения наблюдаемой случайной величины, а в выборках такого рода присутствует значительное количество повторяющихся значений. Как показывают исследования [32–34], распределения статистик критериев, применяемых к подобным выборкам, даже не пытаются сходиться к (существующим) асимптотическим $G(S|H_0)$, а с ростом n только удаляются от них.

В качестве примера на рис. 2 при $\Delta = 0,1\sigma$ показано изменение с ростом n распределения $G(S_n|H_0)$ статистики того же критерия Андерсона–Дарлинга при проверке сложной гипотезы о принадлежности выборок нормальному закону (при использовании оценок максимального правдоподобия (ОМП) для двух его параметров). В этом случае в выборках может присутствовать порядка 86 уникальных значений, а отклонением $G(S_n|H_0)$ от асимптотического распределения можно пре-

небрежь лишь при $n \approx 20 \div 30$. При $\Delta > 0,1\sigma$ изменения распределения $G(S_n|H_0)$ при тех же n будут еще более значительными.

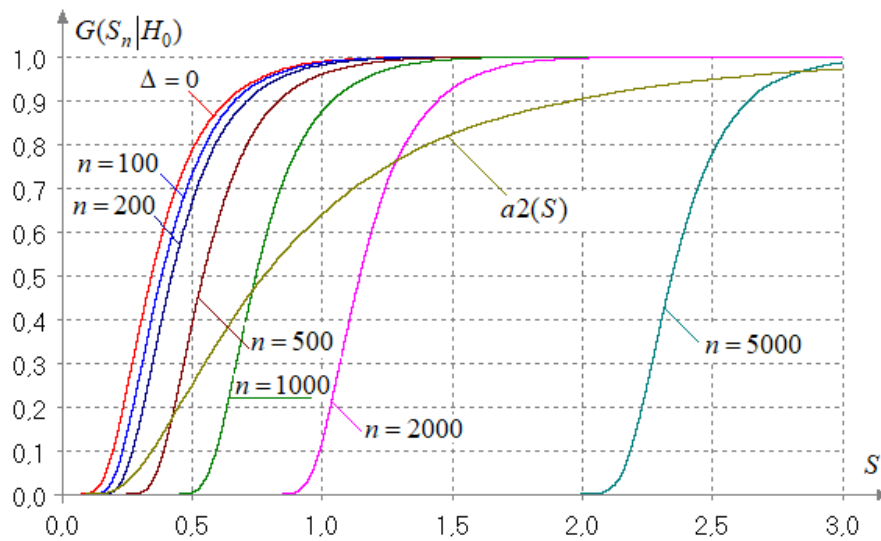


Рис. 2. Зависимость распределения статистики (3) критерия Андерсона–Дарлинга от объема выборки n при справедливости сложной гипотезы H_0 о принадлежности выборки нормальному закону при $\Delta = 0,1\sigma$
 Fig. 2. Dependence of the distribution of statistics (3) of the Anderson–Darling test on the sample size n under the validity of the composite hypothesis H_0 that the sample belongs to the normal law for $\Delta = 0,1\sigma$

Аналогичные картины наблюдаются для распределений статистик всех непараметрических критериев согласия.

Таким образом, наличие влияния ошибок округления для критериев со статистиками (1)–(5) в ситуации проверки простых гипотез приводит к появлению зависимости распределений статистик критериев от n , а при заданной величине Δ – от значения параметра масштаба σ предполагаемого закона [34]. В ситуации проверки сложных гипотез эти зависимости накладываются на зависимости, определяемые факторами (а)–(е). Ясно, что подобным же изменениям подвергаются распределения статистик (6)–(8) критериев Жанга, которые и так зависят от n .

Корректность выводов по всем непараметрическим критериям согласия при Δ , близким к σ , может быть обеспечена лишь за счет интерактивного моделирования распределений $G(S_n|H_0, \Delta)$ или оценивания достигнутого уровня значимости p_v в этих условиях. Такая возможность предусмотрена в программной системе [29].

6. О применении критерия согласия χ^2

Статистика критерия согласия χ^2 Пирсона, предложенного в [35], имеет вид:

$$\chi_n^2 = n \sum_{i=1}^k \frac{(n_i / n - P_i(\theta))^2}{P_i(\theta)}, \quad (9)$$

где $P_i(\theta)$ – вероятность попадания в i -й интервал в соответствии с теоретическим законом, относительно которого проверяется гипотеза. При справедливости простой проверяемой гипотезы асимптотическим распределением статистики является χ_{k-1}^2 -распределение. При справедливости сложной гипотезы и оценивании параметров в результате минимизации статистики (9) последняя подчиняется χ_{k-r-1}^2 -распределению [36–39], где r – число оцененных по выборке параметров.

Заметим, что отказываться от χ_{k-r-1}^2 -распределения как асимптотического распределения статистики нет оснований и в случае применения ряда оценок по группированным данным, в частности

ОМП и других, обладающих такими же асимптотическими свойствами? как оценки, минимизирующие статистику (9). По крайней мере результаты статистического моделирования это подтверждают.

Дискретные распределения $G(X_n^2|H_0)$ статистики сходятся к непрерывным χ^2 -распределениям при любом способе группирования выборки. При этом целесообразно следовать рекомендации [40], разбивая выборку на интервалы таким образом, чтобы ожидаемое число наблюдений $nP_i(\theta)$, попадающих в каждый интервал, было по возможности не меньше $5 \div 10$. Иногда предлагают объединять соседние интервалы, если в каких-то из них число попавших наблюдений оказывается нулевым. Это неверная рекомендация, так как такие действия приводят к увеличению вероятностей ошибок второго рода β .

Распределение статистики $G(X_n^2|H_1)$ при справедливости конкурирующей гипотезы H_1 зависит от способа разбиения на интервалы области определения случайной величины, соответствующей H_0 . То есть мощность критерия относительно H_1 зависит от способа группирования. При этом относительно конкурирующих гипотез H_i , близких к H_0 , критерий обладает максимальной мощностью в случае использования асимптотически оптимального группирования [41, 42], при котором минимизируются потери в информации Фишера, связанные с группированием. Мощность критерия Пирсона зависит также от выбора числа интервалов [43]. Очевидно, что относительно заданной конкурирующей гипотезы H_1 за счет выбора способа группирования и числа интервалов (при заданном n) всегда можно построить критерий χ^2 Пирсона максимальной мощности.

Заметим, что при проверке простых гипотез за счет выбора способа группирования и числа интервалов критерий χ^2 Пирсона может иметь значительное преимущество в мощности перед упомянутыми выше непараметрическими критериями согласия, а при проверке сложных гипотез такое преимущество уже отсутствует [44].

Если при проверке сложной гипотезы в критерии χ^2 Пирсона использовать различные оценки по негруппированным данным, то распределение статистики (9) будет существенно отличаться от χ_{k-r-1}^2 -распределения. В таких ситуациях распределения $G(X_n^2|H_0)$ зависят не только от метода оценивания, но и от способа группирования [45], а реально лежат между χ_{k-r-1}^2 - и χ_{k-1}^2 -распределениями. Надо полагать, что на распределениях статистики (9) в ситуации оценивания параметров по негруппированным данным сказывается влияние тех же факторов (а)–(е), что имеют место в случае непараметрических критериев согласия.

В [46] было показано, что если оценки параметров закона $F(x, \theta)$ находятся по негруппированным данным с использованием метода максимального правдоподобия, то статистика (9) распределена как сумма независимых слагаемых $\chi_{k-m-1}^2 + \sum_{j=1}^m \lambda_j \xi_j^2$, где $\xi_1, \dots, \xi_m$ – стандартные нормальные случайные величины, независимые одна от другой и от случайной величины χ_{k-m-1}^2 . В этой сумме числа $\lambda_1, \dots, \lambda_m$ принимают значения между 0 и 1 и представляют собой корни уравнения $|(1-\lambda)\mathbf{J}(\theta) - \mathbf{J}_\Gamma(\theta)| = 0$, где $\mathbf{J}(\theta)$ и $\mathbf{J}_\Gamma(\theta)$ – информационные матрицы Фишера по негруппированным и группированным данным соответственно. Однако этот результат не позволяет судить о конкретном виде распределения статистики (9) и не может быть использован в процессе формирования корректного вывода о проверяемой гипотезе.

Тем не менее, по-видимому, именно вследствие работы [46] появился ряд модификаций критерия χ^2 Пирсона, отличающихся от него лишь при проверке сложных гипотез и предусматривающих использование оценок по негруппированным данным. В первую очередь следует отметить критерий Никулина–Рао–Робсона [47–49], в котором предусмотрено использование ОМП. Позже появились и

другие критерии, большинство которых перечислено в [50]. Замечательно то, что вместе с модификациями были представлены и распределения их статистик, что облегчает применение соответствующих критериев. Кроме того, критерий Никулина–Рао–Робсона, как правило, демонстрирует преимущество в мощности по сравнению с критерием Пирсона [51]. К сожалению, совокупность этих модификаций не нашла широкого применения на практике. В первую очередь этот факт объясняется «сложностью подготовки» критериев, так как конкретные выражения статистик меняются в зависимости от вида параметрического закона, относительно которого проверяется гипотеза H_0 , и от совокупности оцениваемых параметров. Но при наличии программного обеспечения, предусматривающего проверку относительно различных моделей законов с использованием этих критериев (как в [29]), проблемы применения снимаются.

Подчеркнем, что фактом отличия распределения статистики (9) от χ^2_{k-r-1} -распределения очень часто пренебрегают, в том числе в популярных программных системах. Примером может служить широко применяемая система STATISTICA, где для вычисления p_v в подобной ситуации используют χ^2_{k-r-1} -распределения. Такие действия несколько увеличивают вероятность ошибок первого рода α .

В то же время корректность выводов по критерию χ^2 Пирсона при проверке сложных гипотез с оцениванием параметров по негруппированным данным всегда может быть обеспечена за счет моделирования в этих условиях распределения статистики $G(X_n^2|H_0)$ или оценки достигнутого уровня значимости p_v , что достаточно просто реализовать [29].

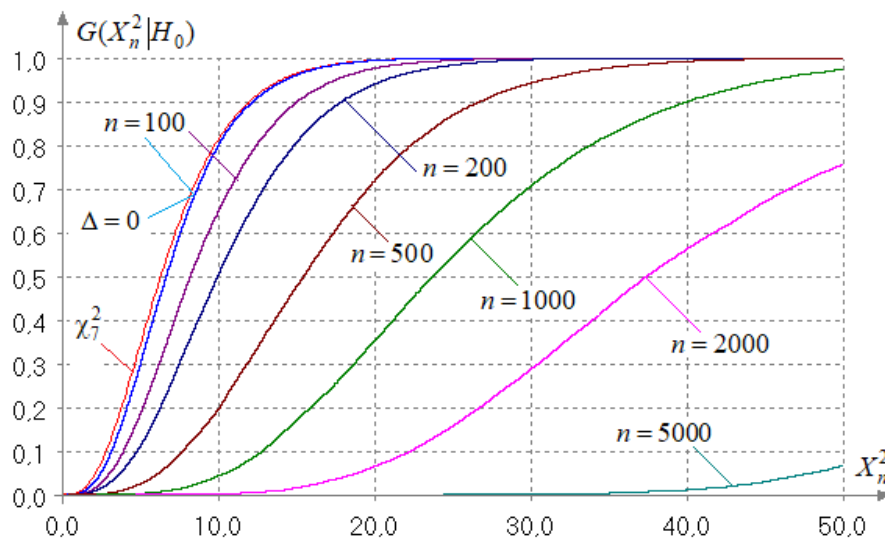


Рис. 3. Зависимость распределения статистики (9) критерия χ^2 Пирсона от объема выборки n при справедливости сложной гипотезы H_0 о принадлежности выборки нормальному закону при $\Delta = 0,1\sigma$

Fig. 3. Dependence of the distribution of statistics (9) of the χ^2 Pearson test on the sample size n under the validity of the composite hypothesis H_0 that the sample belongs to the normal law for $\Delta = 0,1\sigma$

Наличие ошибок округления Δ сказывается на распределениях $G(X_n^2|H_0)$ статистики критерия χ^2 Пирсона точно так же, как и на распределениях непараметрических критериев согласия. На рис. 3 показано, как изменяется распределение статистики $G(X_n^2|H_0)$ с ростом n при проверке сложной гипотезы о принадлежности выборок нормальному закону с вычислением ОМП двух его параметров по негруппированной выборке (число интервалов $k=10$, равновероятные интервалы, $\Delta = 0,1\sigma$).

Заключение

Таким образом, при различии всех обстоятельств, причина некорректности вывода по результатам применения любого критерия согласия оказывается всегда одной и той же. И заключается она в использовании в качестве $G(S|H_0)$ не того распределения.

Очевидно, что область использования классических результатов, касающихся проверки простых гипотез, а также частных результатов относительно проверки сложных гипотез, учитывая возможное влияние ошибок округления, представляется достаточно узкой. Следовательно, в большинстве приложений мы сталкиваемся с применением критериев в нестандартных условиях, когда распределения статистик $G(S_n|H_0, \Delta)$ в той или иной мере отличаются от асимптотических (или от известных). В современных условиях это не представляет серьезной проблемы, так как, используя методы статистического моделирования, мы в интерактивном режиме проверки можем исследовать $G(S_n|H_0, \Delta)$ и оценить достигнутый уровень значимости p_v [29].

Список источников

1. Kolmogoroff A.N. Sulla determinazione empirica di una legge di distribuzione // G. Ist. Ital. Attuar. 1933. V. 4, № 1. P. 83–91.
2. Большев Л.Н. Асимптотические пирсоновские преобразования // Теория вероятностей и ее применение. 1963. Т. 8, № 2. С. 129–155.
3. Большев Л.Н. Теория вероятностей и математическая статистика : избр. тр. / под ред. Ю.В. Прохорова. М. : Наука, 1987. 286 с.
4. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М. : Наука, 1983. 416 с.
5. Кремер Н.Ш. Теория вероятностей и математическая статистика : учебник. 3-е изд., перераб. и доп. М. : ЮНИТИ-ДАНА, 2010. 551 с.
6. Шпаков П.С., Юнаков П.Л. Математическая обработка результатов измерений : учеб. пособие. Красноярск : Сиб. федер. ун-т, 2014. 410 с.
7. Anderson T.W., Darling D.A. Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes // The Annals of Mathematical Statistics. 1952. V. 23, № 2. P. 193–212. URL: <https://www.jstor.org/stable/2236446>
8. Anderson T.W., Darling D.A. A test of goodness of fit // Journal of the American Statistical Association. 1954. V. 29. P. 765–769.
9. Kuiper N.H. Tests concerning random points on a circle // Proc. Koninkl. Nederl. Akad. Van Wetenschappen. Series A. 1960. V. 63. P. 38–47.
10. Stephens M.A. The goodness-of-fit statistic V_N : distribution and significance points // Biometrika. 1965. V. 52, № 3/4. P. 309–321.
11. Stephens M.A. EDF statistics for goodness of fit and some comparisons // Journal of the American Statistical Association. 1974. V. 69, № 347. P. 730–737.
12. Лемешко Б.Ю., Горбунова А.А. О применении и мощности непараметрических критериев согласия Купера, Ватсона и Жанга // Измерительная техника. 2013. № 5. С. 3–9.
13. Watson G.S. Goodness-of-fit tests on a circle. I // Biometrika. 1961. V. 48, № 1/2. P. 109–114.
14. Watson G.S. Goodness-of-fit tests on a circle. II // Biometrika. 1962. V. 49, № 1/2. P. 57–63.
15. Biometrika tables for Statistics / E.S. Pearson, H.O. Hartley (eds.). Cambridge : Cambridge University Press, 1972. V. 2. 385 p.
16. Zhang J. Powerful goodness-of-fit tests based on the likelihood ratio // Journal of the Royal Statistical Society. Series B. 2002. V. 64, № 2. P. 281–294.
17. Zhang J., Wu Yu. Likelihood-ratio tests for normality // Computational Statistics & Data Analysis. 2005. V. 49, № 3. P. 709–721.
18. Лемешко Б.Ю., Горбунова А.А., Лемешко С.Б., Рогожников А.П. О решении проблем применения некоторых непараметрических критериев согласия // Автометрия. 2014. № 1. С. 26–43.
19. Кас М., 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. V. 26. P. 189–211. URL: <https://www.jstor.org/stable/2236876>
20. 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.
21. Лемешко Б.Ю., Маклаков А.А. Непараметрические критерии при проверке сложных гипотез о согласии с распределениями экспоненциального семейства // Автометрия. 2004. № 3. С. 3–20.
22. Lilliefors H.W. On the Kolmogorov–Smirnov test for normality with mean and variance unknown // Journal of the American Statistical Association. 1967. V. 62. P. 399–402.
23. Lilliefors H.W. On the Kolmogorov–Smirnov test for the exponential distribution with mean unknown // Journal of the American Statistical Association. 1969. V. 64. P. 387–389.

24. Р 50.1.037–2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. М. : Изд-во стандартов, 2002. Ч. II: Непараметрические критерии. 64 с.
25. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. I // Измерительная техника. 2009. № 6. С. 3–11.
26. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. II // Измерительная техника. 2009. № 8. С. 17–26.
27. Lemeschko B.Yu., Lemeschko 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. V. 39, № 3. P. 460–471.
28. Лемешко Б.Ю. Непараметрические критерии согласия : руководство по применению. М. : ИНФРА-М, 2014. 163 с.
29. Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин «Интервальная статистика 5.4» : св-во о гос. регистрации программы для ЭВМ № 2018666213 / Лемешко Б.Ю., Лемешко С.Б., Блинов П.Ю., Веретельникова И.В., Новикова А.Ю. Заявка № 20186663206; заявл. 22.11.2018; запат. 13.12.2018. URL: <https://ami.nstu.ru/~headrd/ISW.htm> (дата обращения: 03.04.2023).
30. Лемешко Б.Ю., Блинов П.Ю., Лемешко С.Б. Смещенность непараметрических критериев согласия относительно некоторых пар конкурирующих гипотез // Измерительная техника. 2016. № 5. С. 16–20.
31. 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.
32. Лемешко Б. Ю., Лемешко С.Б. О влиянии ошибок округления на распределение статистик критериев согласия // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2020. № 53. С. 47–60.
33. 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.
34. 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.
35. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably to have arisen from random sampling // The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1900. V. 50. P. 157–175.
36. Fisher R.A. On the interpretation of χ^2 from contingency tables, and the calculation of P // Journal of the Royal Statistical Society. 1922. V. 85, № 1. P. 87–94.
37. Fisher R.A. The conditions under which χ^2 measures the discrepancy between observation and hypothesis // Journal of the Royal Statistical Society. 1924. V. 87, № 3. P. 442–450. URL: <https://www.jstor.org/stable/2341149>
38. Neyman J., Pearson E.S. On the use and interpretation of certain test criteria for purposes of statistical inference: Part I // Biometrika. 1928. V. 20A, № 1/2. P. 175–240.
39. Neyman J., Pearson E.S. On the use and interpretation of certain test criteria for purposes of statistical inference: Part II // Biometrika. 1928. V. 20A, № 3/4. P. 263–294.
40. Mann H.B., Wald A. On the choice of the number of class intervals in the application of the chi square test // The Annals of Mathematical Statistics. 1942. V. 13, № 3. P. 306–317. URL: <https://www.jstor.org/stable/2235942>
41. Денисов В.И., Лемешко Б.Ю. Оптимальное группирование при обработке экспериментальных данных // Измерительные информационные системы. Новосибирск : Новосиб. электротехн. ин-т, 1979. С. 5–14.
42. Лемешко Б.Ю. Асимптотически оптимальное группирование наблюдений в критериях согласия // Заводская лаборатория. 1998. Т. 64, № 1. С. 56–64.
43. Лемешко Б.Ю., Чимитова Е.В. О выборе числа интервалов в критериях согласия типа χ^2 // Заводская лаборатория. Диагностика материалов. 2003. Т. 69, № 1. С. 61–67.
44. Chimitova E.V., Lemeschko B.Y. Chi-Squared Goodness-of-Fit Tests: The Optimal Choice of Grouping Intervals // Recent Advances in Systems, Control and Information Technology. SCIT 2016. Advances in Intelligent Systems and Computing / R. Szweczyk, M. Kaliczyńska (eds). Cham : Springer, 2017. V. 543. P. 760–774.
45. Лемешко Б.Ю., Постовалов С.Н. О зависимости предельных распределений статистик χ^2 Пирсона и отношения правдоподобия от способа группирования данных // Заводская лаборатория. 1998. Т. 64, № 5. С. 5–63.
46. Chernoff H., Lehmann E.L. The use of maximum likelihood estimates in χ^2 test for goodness of fit // The Annals of Mathematical Statistics. 1954. V. 25. P. 579–586.
47. Никулин М.С. Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // Теория вероятностей и ее применение. 1973. Т. 18, № 3. С. 583–591.
48. Никулин М.С. О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. 1973. Т. 18, № 3. С. 675–676.
49. Rao K.C. Robson B.S. A chi-squared statistic for goodness-of-fit tests within the exponential family // Communications in Statistics – Theory and Methods. 1974. V. 3. P. 1139–1153.
50. Greenwood P.E., Nikulin M.S. A guide to chi-squared testing. New York : John Wiley & Sons, 1996. 280 p.
51. Лемешко Б.Ю., Постовалов С.Н., Чимитова Е.В. О распределениях статистики и мощности критерия типа χ^2 Никулина // Заводская лаборатория. Диагностика материалов. 2001. Т. 67, № 3. С. 52–58.

References

1. Kolmogoroff, A.N. (1933) Sulla determinazione empirica di una legge di distribuzione. *Giornale del Istituto Italiano degli Attuari*. 4(1). pp. 83–91.
2. Bolshev, L.N. (1963) Asimptoticheskie pirsonovskie preobrazovaniya [Asymptotic Pearson transformations]. *Teoriya veroyatnostey i ee primeneniye*. 8(2). pp. 129–155.
3. Bolshev, L.N. (1987) *Teoriya veroyatnostey i matematicheskaya statistika* [Theory of Probability and Mathematical Statistics]. Moscow: Nauka.
4. Bolshev, L.N. & Smirnov, N.V. (1983) *Tablitsy matematicheskoy statistiki* [Tables for Mathematical Statistics]. Moscow: Nauka.
5. Kremer N.Sh. (2010) *Teoriya veroyatnostey i matematicheskaya statistika* [Probability Theory and Mathematical Statistics]. 3rd ed. Moscow: UNITY-DANA.
6. Shpakov, P.S. & Yunakov, P.L. (2014) *Matematicheskaya obrabotka rezul'tatov izmereniy* [Mathematical Processing of Measurement Results]. Krasnoyarsk: SFU.
7. Anderson T.W. & Darling D.A. (1952) Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*. 23(2). pp. 193–212.
8. Anderson, T.W. & Darling, D.A. (1954) A test of goodness of fit. *Journal of the American Statistical Association*. 29. pp. 765–769.
9. Kuiper, N.H. (1960) Tests concerning random points on a circle. *Proc. Koninkl. Nederl. Akad. Van Wetenschappen. Series A*. 63. pp. 38–47.
10. Stephens, M.A. (1965) The goodness-of-fit statistic V_N : distribution and significance points. *Biometrika*. 52(3/4). pp. 309–321. DOI: 10.2307/2333685
11. Stephens, M.A. (1974) EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*. 69(347). pp. 730–737.
12. Lemesenko, B.Yu. & Gorbunova, A.A. (2013) Application and Power of the Nonparametric Kuiper, Watson, and Zhang Tests of Goodness-of-Fit. *Izmeritel'naya tekhnika – Measurement Techniques*. 56(5). pp. 465–475. DOI: 10.1007/s11018-013-0229-7
13. Watson, G.S. (1961) Goodness-of-fit tests on a circle. I. *Biometrika*. 48(1/2). pp. 109–114. DOI: 10.2307/2333135
14. Watson, G.S. (1962) Goodness-of-fit tests on a circle. II. *Biometrika*. 49(1/2). pp. 57–63. DOI: 10.2307/2333467
15. Pearson, E.S. & Hartley, H.O. (eds) (1972) *Biometrika tables for Statistics*. Vol. 2. Cambridge: University Press.
16. Zhang, J. (2002) Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B*. 64(2). pp. 281–294. DOI: 10.1111/1467-9868.00337
17. Zhang, J. & Wu, Yu. (2005) Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis*. 49(3). pp. 709–721. DOI: 10.1016/j.csda.2004.05.034
18. Lemesenko, B.Yu., Gorbunova, A.A., Lemesenko, S.B. & Rogozhnikov, A.P. (2014) Solving problems of using some nonparametric goodness-of-fit tests. *Optoelectronics, Instrumentation and Data Processing*. 50(1). pp. 21–35. DOI: 10.3103/S875669901401004X
19. Kac, M., Kiefer, J. & Wolfowitz, J. (1955) On tests of normality and other tests of goodness of fit based on distance methods. *The Annals of Mathematical Statistics*. 26. pp. 189–211.
20. Lemesenko, B.Yu. & Postovalov, S.N. (2001) O zavisimosti raspredeleniy statistik neparametricheskikh kriteriyev i ikh moshchnosti ot metoda otsenivaniya parametrov [On the dependence of the distributions of statistics of non-parametric tests and their power on the parameter estimation method]. *Zavodskaya laboratoriya. Diagnostika materialov*. 67(7). pp. 62–71.
21. Lemesenko, B.Yu. & Maklakov, A.A. (2004) Nonparametric Test in Testing Composite Hypotheses on Goodness of Fit Exponential Family Distributions. *Optoelectronics, Instrumentation and Data Processing*. 40(3). pp. 3–18.
22. Lilliefors, H.W. (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*. 62. pp. 399–402. DOI: 10.1080/01621459.1967.10482916
23. Lilliefors, H.W. (1969) On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*. 64. pp. 387–389. DOI: 10.1080/01621459.1969.10500983
24. R 50.1.037–2002. *Rekomendatsii po standartizatsii. Prikladnaya statistika. Pravila proverki soglasiya opytnogo raspredeleniya s teoreticheskimi* [Recommendations for standardization. Applied statistics. Rules of check of experimental and theoretical distribution of the consent]. Part II. Moscow: Izd-vo standartov.
25. Lemesenko, B.Yu. & Lemesenko, S.B. (2009a) Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. Part I. *Izmeritel'naya tekhnika – Measurement Techniques*. 52(6). pp. 555–565. DOI: 10.1007/s11018-009-9330-3
26. Lemesenko, B.Yu. & Lemesenko, S.B. (2009) Models for statistical distributions in nonparametric fitting tests on composite hypotheses based on maximum-likelihood estimators. Part II. *Izmeritel'naya tekhnika – Measurement Techniques*. 52(8). pp. 799–812. DOI: 10.1007/s11018-009-9353-9
27. Lemesenko, B.Yu., Lemesenko, S.B. & Postovalov, S.N. (2010) Statistic Distribution Models for Some Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses. *Communications in Statistics – Theory and Methods*. 39(3). pp. 460–471. DOI: 10.1080/03610920903140148
28. Lemesenko, B.Yu. (2014) *Neparametricheskie kriterii soglasiya: rukovodstvo po primeneniyu* [Nonparametric goodness-of-fit tests. Guide on the application]. Moscow: INFRA-M. DOI: 10.12737/11873

29. Lemesheko, B.Yu., Lemesheko, S.B., Blinov, P.Yu., Veretelnikova, I.V. & Novikova, A.Yu. (n.d.) *Statisticheskiy analiz interval'nykh nablyudeniy odnomernykh nepreryvnykh sluchaynykh velichin "Interval'naya statistika 5.4": sv-vo o gos. registratsii programmy dlya EVM № 2018666213* [Statistical analysis of interval observations of one-dimensional continuous random variables "Interval statistics 5.4": Certificate of state registration of the computer program No. 2018666213 / Application No. 2018663206; Dec. 11/22/2018; registered 12/13/2018]. [Online] Available from: <https://ami.nstu.ru/~headrd/ISW.htm> (Accessed: 4th March 2023).
30. Lemesheko, B.Yu., Blinov, P.Yu. & Lemesheko S.B. (2016) Smeshchennost' neparametricheskikh kriteriev soglasiya odnositel'no nekotorykh par konkuriruyushchikh gipotez [Bias of nonparametric goodness-of-fit tests relative to certain pairs of competing hypotheses]. *Izmeritel'naya tekhnika – Measurement Techniques*. 59(5). pp. 468–475. DOI: 10.1007/s11018-016-0992-3
31. Lemesheko, B.Yu., Lemesheko, S.B. & Semenova, M.A. (2018) 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*. 44. pp. 40–49. DOI: 10.17223/19988605/44/5
32. Lemesheko, B.Yu. & Lemesheko, S.B. (2020) 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*. 53. pp. 47–60. DOI: 10.17223/19988605/53/5
33. Lemesheko, B.Yu. & Lemesheko, S.B. (2021) Problemy primeneniya neparametricheskikh kriteriev soglasiya v zadachakh obrabotki rezul'tatov izmerenii [Problems of nonparametric goodness-of-fit test application in tasks of measurement results processing]. *Sistemy analiza i obrabotki dannykh – Analysis and Data Processing Systems*. 2(82). pp. 47–66. DOI: 10.17212/2782-2001-2021-2-47-66.
34. Lemesheko, B.Yu. & Lemesheko, S.B. (2022) 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*. 2(86). pp. 21–38. DOI: 10.17212/2782-2001-2022-2-21-38
35. Pearson, K. (1900) On the criterion that a given system of deviations from the probable in the case of a correlated system of variables in such that it can be reasonably to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. 50. pp. 157–175. DOI: 10.1080/14786440009463897
36. Fisher, R.A. (1922) On the interpretation of χ^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society*. 85(1). pp. 87–94. DOI: 10.2307/2340521
37. Fisher, R.A. (1924) The conditions under which χ^2 measures the discrepancy between observation and hypothesis. *Journal of the Royal Statistical Society*. 87(3). pp. 442–450.
38. Neyman, J. & Pearson, E.S. (1928) On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. *Biometrika*. 20A(1/2). pp. 175–240. DOI: 10.2307/2331945
39. Neyman, J. & Pearson, E.S. (1928) On the use and interpretation of certain test criteria for purposes of statistical inference: Part II. *Biometrika*. 20A(3/4). pp. 263–294. DOI: 10.2307/2332112
40. Mann, H.B. & Wald, A. (1942) On the choice of the number of class intervals in the application of the chi square test. *The Annals of Mathematical Statistics*. 13(3). pp. 306–317.
41. Denisov, V.I. & Lemesheko, B.Yu. (1979) *Optimal'noe gruppирование pri obrabotke eksperimental'nykh dannykh* [Optimal grouping in the processing of experimental data]. Novosibirsk: Novosibirsk Electrical Engineering Institute. pp. 5–14.
42. Lemesheko, B.Yu. (1998) Asimptoticheski optimal'noe gruppирование nablyudeniy v kriteriyakh soglasiya [Asymptotically optimum grouping of observations in goodness-of-fit tests]. *Zavodskaya laboratoriya – Industrial Laboratory*. 64(1). pp. 59–67.
43. Lemesheko, B.Y. & Chimitova, E.V. (2003) O vybore chisla intervalov v kriteriyakh soglasiya tipa χ^2 [On the choice of the number of intervals in χ^2 -type goodness-of-fit tests]. *Zavodskaya laboratoriya – Industrial Laboratory*. 69(1). pp. 61–67.
44. Chimitova E. V. & Lemesheko B. Y. (2017). Chi-Squared Goodness-of-Fit Tests: The Optimal Choice of Grouping Intervals. In: Szewczyk, R. & Kaliczynska, M. (eds) *Recent Advances in Systems, Control and Information Technology. SCIT 2016. Advances in Intelligent Systems and Computing*. Vol. 543. Springer, Cham. DOI: 10.1007/978-3-319-48923-0_82
45. Lemesheko, B.Yu. & Postovalov, S.N. (1998) O zavisimosti predel'nykh raspredeleniy statistik Pirona i otnosheniya pravdopodobiya ot sposoba gruppированиа dannykh [Limit distributions of the Pearson chi 2 and likelihood ratio statistics and their dependence on the mode of data grouping]. *Zavodskaya laboratoriya – Industrial Laboratory*. 64(5). pp. 344–351.
46. Chernoff, H. & Lehmann, E.L. (1954) The use of maximum likelihood estimates in χ^2 test for goodness of fit. *The Annals of Mathematical Statistics*. 25. pp. 579–586. DOI: 10.1007/978-1-4614-1412-4_47
47. Nikulin, M.S. (1973a) Kriteriy khi-kvadrat dlya nepreryvnykh raspredeleniy s parametrami sdviga i masshtaba [Chi-square test for continuous distributions with location and scale parameters]. *Teoriya veroyatnostey i ee primeneniye*. 18(3). pp. 583–591.
48. Nikulin, M.S. (1973b) O kriterii khi-kvadrat dlya nepreryvnykh raspredeleniy [About the Chi-square test for continuous distributions]. *Teoriya veroyatnostey i ee primeneniye*. 18(3). pp. 675–676.
49. Rao, K.C. & Robson, B.S. (1974) A chi-squared statistic for goodness-of-fit tests within the exponential family. *Communications in Statistics – Theory and Methods*. 3. pp. 1139–1153. DOI: 10.1080/03610927408827216
50. Greenwood, P.E. & Nikulin, M.S. (1996) *A guide to chi-squared testing*. New York: John Wiley & Sons.
51. Lemesheko, B., Postovalov, S. & Chimitova, E. (2001) O raspredeleniyakh statistiki i moshchnosti kriteriya tipa χ^2 Nikulina [On statistic distributions and the power of the Nikulin χ^2 test]. *Zavodskaya laboratoriya – Industrial Laboratory*. 67(3). pp. 52–58.

Информация об авторе:

Лемешко Борис Юрьевич – профессор, доктор технических наук, профессор кафедры теоретической и прикладной информатики Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: lemeshko@amil.nstu.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Lemeshko Boris Yu. (Professor, Doctor of Technical Sciences, Professor, Department of Theoretical and Applied Informatics, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: lemeshko@amil.nstu.ru

The author declares no conflicts of interests.

Поступила в редакцию 26.04.2023; принята к публикации 04.09.2023

Received 26.04.2023; accepted for publication 04.09.2023