ISSN 2782-2001
Системы анализа
и обработки данных
том 86, № 2, 2022, с. 21–38

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 86, No. 2, 2022, pp. 21–38

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.23

DOI: 10.17212/2782-2001-2022-2-21-38

Непараметрические критерии согласия при проверке нормальности в условиях округления измерений*

Б.Ю. ЛЕМЕШКО [a], С.Б. ЛЕМЕШКО [b]

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

[a] *lemeshko@ami.nstu.ru* [b] *skyer@mail.ru*

При анализе рядов измерений в различных приложениях в качестве обязательной процедуры рассматривается проверка принадлежности ошибок измерений нормальному закону. В этих целях могут использоваться различные специальные критерии проверки гипотез о нормальности, могут применяться непараметрические критерии согласия или критерии типа хи-квадрат.

При использовании для проверки нормальности непараметрических критериев согласия необходимо учитывать, что проверяется сложная гипотеза. При проверке сложной гипотезы распределения статистик критериев согласия существенно отличаются от классических, имеющих место при проверке простых гипотез.

Известно, что наличие ошибок округления может существенно изменять распределения статистик критериев. В таких ситуациях игнорирование факта влияния может приводить к некорректным выводам о результатах проверки нормальности.

Метрологи при проведении высокоточных измерений, как правило, не допускают и мысли о возможном влиянии ошибок округления Δ на результаты статистического анализа. Этим самым допускается возможность некорректности выводов, так как влияние округления отсутствует не просто при малых Δ, а при значениях Δ, много меньших среднеквадратического отклонения σ закона распределения ошибок измерения, и объемах выборок n, не превышающих некоторых максимальных значений. При объемах выборок больших, чем эти максимальные значения, реальные распределения статистик критериев отклоняются от асимптотических в сторону больших значений статистик.

В работе на реальных и хорошо знакомых данных с использованием методов статистического моделирования демонстрируется зависимость распределений статистик непараметрических критериев согласия при проверке нормальности от соотношения Δ и σ при конкретных значениях n. Показывается и реализуется возможность корректного применения критериев в условиях влияния на выводы ошибок округления.

* *Статья получена 17 марта 2022 г.*

Ключевые слова: проверка нормальности, непараметрические критерии согласия, сложная гипотеза, распределение статистики, ошибки округления, достигнутый уровень значимости, ошибка 1-го рода, ошибка 2-го рода, имитационное моделирование

ВВЕДЕНИЕ

При анализе рядов измерений в различных приложениях в качестве обязательной процедуры рассматривается проверка принадлежности ошибок измерений нормальному закону. В этих целях могут использоваться различные специальные критерии проверки гипотез о нормальности, могут применяться непараметрические критерии согласия или критерии типа хи-квадрат [1].

При использовании для проверки нормальности непараметрических критериев согласия необходимо учитывать, что проверяется сложная гипотеза. При проверке сложной гипотезы распределения статистик критериев согласия существенно отличаются от классических, имеющих место при проверке простых гипотез. Игнорирование этого факта является одной из наиболее распространенных ошибок при применении непараметрических критериев согласия [2].

Известно, что наличие ошибок округления может существенно изменять распределения статистик критериев [3–5]. В таких ситуациях пренебрежение фактом влияния может приводить к некорректным выводам о результатах проверки нормальности.

В метрологии при проведении высокоточных измерений, как правило, не допускают и мысли о возможном влиянии ошибок округления Δ на результаты статистического анализа. В результате возрастает риск некорректности выводов, так как влияние округления отсутствует не просто при малых Δ, а при значениях Δ, много меньших среднеквадратического отклонения σ закона распределения ошибок измерения, и объемах выборок n, не превышающих некоторых максимальных значений. При объемах выборок бóльших, чем эти максимальные значения, реальные распределения статистик критериев отклоняются от асимптотических в сторону бóльших значений статистик [6].

В настоящей статье на реальных и хорошо знакомых данных [7] с использованием методов статистического моделирования демонстрируется зависимость распределений статистик непараметрических критериев согласия от соотношения между Δ и σ и от объемов выборок n при проверке гипотезы о нормальности.

1. ПРИМЕНЯЕМЫЕ ПРИ ПРОВЕРКЕ НОРМАЛЬНОСТИ НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ СОГЛАСИЯ

В работе рассматривается применение восьми непараметрических критериев согласия.

В критерии **Колмогорова** (K) [8] (в отечественной практике) используется статистика с поправкой Большева [9] в виде

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

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

Статистика критерия **Крамера – Мизеса – Смирнова** (CMS) имеет вид [9]

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

Статистика критерия **Андерсона – Дарлинга** (AD) [10, 11], как правило, применяется в виде

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

В критерии **Купера** (Ku) [12] можно использовать статистику в виде [13]

$$V_n^{mod} = \sqrt{n}\left(D_n^+ + D_n^-\right) + \frac{1}{3\sqrt{n}}. \tag{4}$$

Статистика критерия **Ватсона** (W) [14, 15] используется в форме

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

В критериях **Жанга** [16, 17] статистики задаются выражениями:

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

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

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

Критерии Жанга являются развитием критериев со статистиками (1)–(3).

Модели распределений статистик (1)–(5), имеющие место при проверке нормальности (в отсутствие влияния ошибок округления) с оцениванием параметров нормального закона методом максимального правдоподобия, получены в [18–20] и представлены в табл. 1.

Таблица 1

Table 1

Модели распределений статистик (1)–(5) при проверке нормальности

Models of statistics distributions (1)–(5) when testing normality

№ п/п	Критерий	Модель распределения статистики
1	Колмогорова	γ(6.4721; 0.0580; 0.2620)
2	Крамера – Мизеса – Смирнова	B_{III}(4.1153; 4.1748; 11.035; 0.5116; 0.009)
3	Андерсона – Дарлинга	B_{III}(4.7262; 4.6575; 9.4958; 2.717; 0.0775)
4	Купера	B_{III}(7.4917; 8.0016; 2.4595; 2.1431; 0.4937)
5	Ватсона	B_{III}(3.5230; 4.4077; 9.2281; 0.4785; 0.0104)

В табл. 1 через $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ обозначено бета-распределение III рода с плотностью

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

через $\gamma(\theta_0, \theta_1, \theta_2)$ – гамма-распределение с функцией плотности

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

Распределения статистик (6)–(8), в том числе при проверке нормальности, зависят от объемов выборок n. Поэтому, применяя соответствующие критерии, для вычисления достигнутого уровня значимости p_{value} в программном обеспечении (как в [21]) приходится предусматривать использование интерактивного режима моделирования распределений статистик.

2. ВЛИЯНИЕ ОКРУГЛЕНИЯ НА РАСПРЕДЕЛЕНИЯ СТАТИСТИК

Измерения всегда фиксируются с некоторой погрешностью округления. Возможность влияния ошибок округления на статистические выводы давно признавалась многими авторами [22–25]. Вопрос заключался лишь в том, как реально наличие ошибок округления сказывается на распределениях статистик критериев.

В ситуациях, когда ошибки округления $\Delta << \sigma$, где σ – среднеквадратичное отклонение ошибки измерения, и n меньше некоторого $n_{\max}$, зависящего от n и σ, влиянием Δ на распределения статистик можно пренебречь [6] и пользоваться асимптотическими результатами, имеющими место при проверке простых и сложных гипотез.

При соизмеримости Δ и σ, когда значения Δ и σ величины одного порядка, всё меняется. Влияние ошибок округления на распределения статистик критериев проверки различных гипотез рассматривалось нами в [2–5].

Результаты, полученные в [2], касающиеся критериев со статистиками (1)–(5) и имеющие отношение к проверке нормальности, кратко можно охарактеризовать следующим образом:

– наличие ошибок округления приводит к появлению зависимости $G(S|H_0)$ от n;

– факт наличия округлений в анализируемых данных исключает возможность использования асимптотических распределений статистик в условиях больших выборок;

– в условиях соизмеримости Δ и σ распределения $G(S_n|H_0)$ статистик могут значительно отличаться от асимптотических и при относительно небольших объемах выборок;

– вследствие наличия округлений потеря свойства «свободы от распределения» может происходить и в условиях проверки простых гипотез;

– на распределения статистик (6)–(8), которые зависят от n и в отсутствие ошибок округления, все вышеперечисленные факторы, связанные с Δ, воздействуют аналогичным образом.

В качестве демонстрационного примера на рис. 1 в условиях соизмеримости Δ и σ иллюстрируется влияние Δ на распределение $G(S_K|H_0)$ статистики (1) критерия Колмогорова в следующей ситуации: проверяется гипотеза H_0 о принадлежности выборок нормальному закону распределения при $n = 50$ с вычислением оценок максимального правдоподобия (ОМП) параметров μ и σ закона. Приведенные на рисунке зависимости построены на основе результатов статистического моделирования распределений статистик в условиях округления данных с использованием возможностей системы ISW [21]. Аналогично на рис. 2 показано, как меняются распределения статистики (1) в зависимости от объема выборки n при справедливости гипотезы H_0 о нормальности при фиксированной ошибке округления $\Delta = 0.1\sigma$ и использовании ОМП. Мы можем видеть, что и при росте Δ, и при росте n распределения $G(S_K|H_0)$ всё дальше отклоняются от асимптотического распределения статистики, имеющего место в отсутствие ошибок округления, модель которого приведена в табл. 1.

Таким же образом в этих условиях меняются распределения статистик (2)–(8) других непараметрических критериев согласия.

3. ПРОВЕРКА НОРМАЛЬНОСТИ В УСЛОВИЯХ ОКРУГЛЕНИЯ ИЗМЕРЕНИЙ

В работе [7] результаты измерений характеристик ирисов были использованы для решения задачи таксономии. Заимствованные в [7] результаты измерений в сантиметрах представлены в табл. 2.

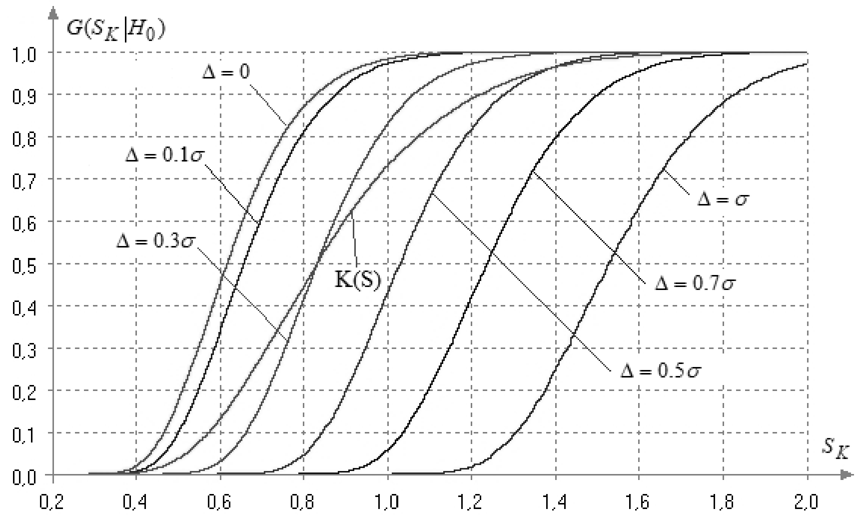

Рис. 1. Зависимость распределения статистики (1) критерия Колмогорова от Δ при справедливости сложной гипотезы H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $n = 50$

Fig. 1. Dependence of the distribution of statistic (1) of the Kolmogorov test on Δ under the validity of the composite hypothesis H_0 that the sample belongs to the normal law (in the case of MLE) for $n = 50$

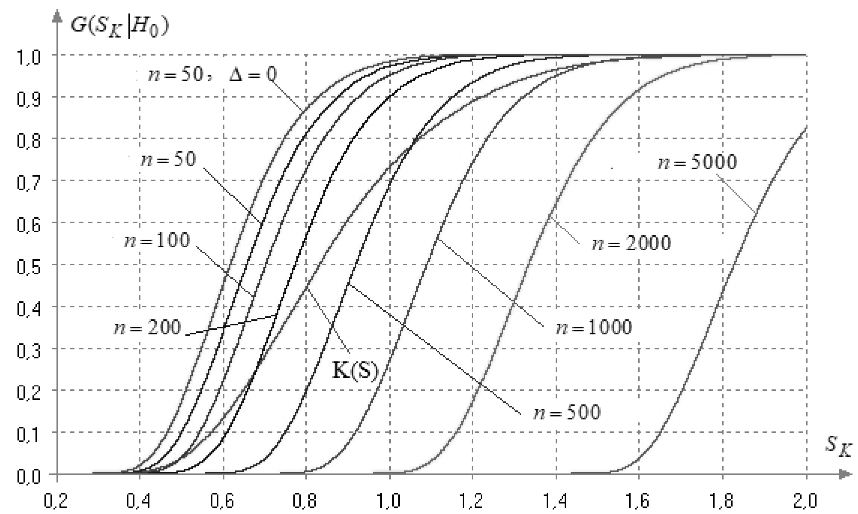

Рис. 2. Зависимость распределения статистики (1) критерия Колмогорова от объема выборки n при справедливости сложной гипотезы H_0 о принадлежности выборки нормальному закону (в случае ОМП) при $\Delta = 0.1\sigma$

Fig. 2. Dependence of the distribution of statistic (1) of the Kolmogorov test on the sample size n with the correctness of the composite hypothesis H_0 that the sample belongs to the normal law (in the case of MLE) at $\Delta = 0.1\sigma$

В табл. 2 для каждого из трех видов ириса (Iris setosa, Iris versicolor, Iris virginica) представлены измерения четырех характеристик для 50 представителей каждого вида: Sepal length – длина чашелистика, Sepal width – ширина чашелистика, Petal length – длина лепестка, Petal width – ширина лепестка. Погрешность округления $\Delta = 0.1$ одна и та же для всех измерений.

В выборках непрерывных одинаково распределенных случайных величин вероятность появления двух одинаковых значений равна нулю. В рядах реальных измерений вследствие естественного округления бывают повторяющиеся значения, и иногда их много. В столбцах табл. 2 также наблюдаются повторяющиеся значения. Посмотрим, как это отражается на результатах проверки гипотезы о принадлежности измерений характеристик ирисов нормальному закону.

Результаты проверки принадлежности всех 12 выборок нормальным законам по восьми рассматриваемым критериям для каждого из трех видов ириса представлены в табл. 3–5.

Для каждой выборки приводятся ОМП параметров μ и σ нормального закона, вычисленные значения S статистик применяемых критериев согласия, значения достигнутого уровня значимости p_{value}, рассчитанного по асимптотическому $G(S|H_0)$ распределению статистики в предположении об отсутствии ошибок округления (при $\Delta = 0$) и полученного по реальному распределению $G(S_n|H_0)$ (при $\Delta = 0.1$ и при соответствующей ОМП для σ). В последнем случае оценка p_{value} находилась по смоделированному в интерактивном режиме реальному распределению $G(S_n|H_0)$ (при $\Delta = 0.1$ и конкретном σ). Такая возможность реализована в [21].

Как можно видеть, оценки p_{value}, вычисленные по реальному распределению $G(S_n|H_0)$, кардинально отличаются от значений p_{value}, вычисленных по асимптотическому распределению статистики $G(S|H_0)$. И если пренебречь влиянием ошибок округления на распределения статистик критериев, то во многих случаях гипотеза о нормальности будет несправедливо отклоняться.

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

В отсутствие влияния округлений распределения $G(S_n|H_0)$ статистик (1)–(5) быстро сходятся к асимптотическим $G(S|H_0)$ распределениям этих статистик: отклонением $G(S_n|H_0)$ от $G(S|H_0)$ можно пренебречь, как правило, при $n \geq 25\ldots30$. При наличии влияния ошибок округления (как в данном случае) распределения $G(S_n|H_0)$ не сходятся к асимптотическим $G(S|H_0)$, а с ростом n всё дальше отклоняются от них.

Таблица 2

Table 2

Результаты измерений характеристик ирисов

Results of measurements of the irises characteristics

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

Таблица 3

Table 3

Результаты проверки гипотез о принадлежности нормальному закону характеристик для Iris setosa

Results of testing the hypotheses about the belonging to the normal law of characteristics for Iris setosa

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

Таблица 4

Table 4

Результаты проверки гипотез о принадлежности нормальному закону характеристик для Iris versicolor

Results of testing hypotheses about belonging to the normal law of characteristics for Iris versicolor

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

Таблица 5

Table 5

Результаты проверки гипотез о принадлежности нормальному закону характеристик для Iris virginica

Results of testing hypotheses about belonging to the normal law of characteristics for Iris virginica

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

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

Отметим также, что в общем случае проверки сложной гипотезы о принадлежности выборки некоторому закону $F(x,\theta)$ к факторам, влияющим на распределения статистик $G(S | H_0)$ при сложной гипотезе [13], добавляется зависимость от n, Δ и от значений оценок параметров формы и масштаба закона $F(x,\theta)$.

Для проверки нормальности кроме непараметрических критериев согласия может применяться более трех десятков специальных критериев [26], ориентированных только на проверку принадлежности выборок нормальному закону, в том числе описанных в [27–36]. Часть из этих критериев обладает определенными преимуществами по сравнению с непараметрическими критериями согласия, в частности, выигрывают в мощности относительно определенных конкурирующих гипотез. Однако и на распределениях статистик специальных критериев нормальности в аналогичных ситуациях сказывается влияние ошибок округления [4].

ЗАКЛЮЧЕНИЕ

Таким образом, применяя непараметрические критерии согласия для проверки гипотезы о принадлежности выборок нормальному закону распределения, следует учитывать возможное влияние ошибок округления на распределения статистик критериев.

Ошибки округления есть всегда. В ситуации, когда $\Delta << \sigma$ и n меньше некоторого $n_{\max}$, зависящего от n и σ, влиянием Δ на распределения статистик можно пренебречь. Но при $n > n_{\max}$ реальные распределения статистик отклоняются от асимптотических, использование которых при проверке увеличивает вероятности ошибок 1-го рода – отклонения верной гипотезы H_0.

В условиях соизмеримости Δ и σ отклонение реальных распределений статистик от асимптотических (или от имеющих место в отсутствие округлений при зависимости распределений статистик от n) может проявляться при относительно малых объемах выборок. Это подтверждает рассмотренный пример с характеристиками ирисов. А с ростом n эта проблема только усугубляется.

С выборками такого рода сталкиваются не только в биологии, зоологии, медицине и других науках, где в связи со спецификой измеряемых величин ошибки округления Δ всегда достаточно велики.

С такого же рода выборками сталкиваются при высокоточных измерениях в технических приложениях, когда измерения осуществляются на пределе точности измерительных систем. Общим и в том и в другом случае является проведение измерений на пределе точности.

Не следует считать, что ошибки округления могут изменять свойства только критериев согласия. Ошибки округления влияют на распределения статистик специальных критериев, ориентированных только на проверку нормальности. Под их влиянием изменяются свойства критериев равномерности, экспоненциальности и других десятков и сотен критериев, предназначенных для проверки различных гипотез.

Решать обозначенную проблему применения критериев проверки различных гипотез в условиях влияния ошибок округления можно единственным способом: разработать программное обеспечение, позволяющее методами статистического моделирования исследовать распределения статистик критериев (или находить оценки p_{value}) в конкретных условиях приложения и при конкретном значении Δ. Таким примером является система [21], в рамках которой проведены настоящие исследования.

СПИСОК ЛИТЕРАТУРЫ

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

2. *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.

3. *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.

4. *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.

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

6. *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.

7. *Fisher R.A.* The use of multiple measurements in taxonomic problems // Annals of Eugenics. – 1936. – Vol. 7. – P. 179–188.

8. *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.

9. *Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. – М.: Наука, 1983. – 416 с.

10. *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. – P. 193–212.

11. *Anderson T.W., Darling D.A.* A test of goodness of fit // Journal of the American Statistical Association. – 1954. – Vol. 29. – P. 765–769.

12. *Kuiper N.H.* Tests concerning random points on a circle // Indagationes Mathematicae (Proceedings). – 1960. – Vol. 63. – P. 38–47.

13. *Лемешко Б.Ю.* Непараметрические критерии согласия: руководство по применению. – М.: Инфра-М, 2014. – 163 с. – DOI: 10.12737/11873.

14. *Watson G.S.* Goodness-of-fit tests on a circle. 1 // Biometrika. – 1961. – Vol. 48, N 1–2. – P. 109–114.

15. *Watson G.S.* Goodness-of-fit tests on a circle. 2 // Biometrika. – 1962. – Vol. 49, N 1–2. – P. 57–63.

16. *Zhang J.* Powerful goodness-of-fit and multi-sample tests: PhD Thesis / York University. – Toronto, 2001. – 113 p. – URL: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (дата обращения: 27.05.2022).

17. *Zhang J.* Powerful goodness-of-fit tests based on the likelihood ratio // Journal of the Royal Statistical Society: Series B. – 2002. – Vol. 64, N 2. – P. 281–294.

18. *Lemeshko B.Yu., Lemeshko S.B.* Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. Pt. 1 // Measurement Techniques. – 2009. – Vol. 52, N 6. – P. 555–565.

19. *Lemeshko B.Yu., Lemeshko 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, N 3. – P. 460–471.

20. *Lemeshko B.Yu., Gorbunova A.A.* Application of nonparametric Kuiper and Watson tests of goodness-of-fit for composite hypotheses // Measurement Techniques. – 2013. – Vol. 56, N 9. – P. 965–973.

21. 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: 27.05.2022).

22. *Pearson E.S., D'Agostino R.B., Bowman K.O.* Tests for departure from normality: comparison of powers // Biometrika. – 1977. – Vol. 64. – P. 231–246. – DOI: 10.1093/biomet/64.2.427-a.

23. *Tricker A.R.* The effect of rounding on the significance level of certain normal test statistics // Journal of Applied Statistics. – 1990. – Vol. 17, N 1. – P. 31–38. – DOI: 10.1080/757582644.

24. *Tricker A.R.* The effect of rounding on the power level of certain normal test statistics // Journal of Applied Statistics. – 1990. – Vol. 17, N 2. – P. 219–228. – DOI: 10.1080/757582833.

25. *Deidda R., Puliga M.* Sensitivity of goodness-of-fit statistics to rainfall data rounding off // Physics and Chemistry of the Earth. – 2006. – Vol. 31. – P. 1240–1251. – DOI: 10.1016/j.pce.2006.04.041.

26. *Uyanto S.S.* An extensive comparisons of 50 univariate goodness-of-fit tests for normality // Austrian Journal of Statistics. – 2022. – Vol. 51. – P. 45–97.

27. *Frosini B.V.* A survey of a class of goodness-of-fit statistics // Metron. – 1978. – Vol. 36, N 1–2. – P. 3–49.

28. *Epps T.W., Pulley L.B.* A test for normality based on the empirical characteristic function // Biometrika. – 1983. – Vol. 70. – P. 723–726.

29. *Hegazy Y.A.S., Green J.R.* Some new goodness-of-fit tests using order statistics // Journal of the Royal Statistical Society: Series C (Applied Statistics). – 1975. – Vol. 24, N 3. – P. 299–308.

30. *David H.A., Hartley H.O., Pearson E.S.* The distribution of the ratio? In a single normal sample, of range to standard deviation // Biometrika. – 1964. – Vol. 512, N 3–4. – P. 484–487.

31. *Geary R.C.* Testing for Normality // Biometrika. – 1937. – Vol. 34. – P. 209–242.

32. *D'Agostino R.B.* Transformation to normality of the null distribution of g_1 // Biometrika. – 1970. – Vol. 57 (3). – P. 679–681.

33. *Chen L., Shapiro S.S.* An alternative test for normality based on normalized spacings // Journal of Statistical and Simulation. – 2012. – Vol. 53. – P. 269–288.

34. *Desgagne A., Micheaux P.L. de.* A powerful and interpretable alternative to the Jarque–Bera test of normality based on 2nd-power skewness and kurtosis, using the rao's score test on the APD family // Journal of Applied Statistics. – 2018. – Vol. 45, N 13. – P. 2307–2327.

35. *Gel Y.R., Gastwirth J.L.* A robust modification of the Jarque–Bera test of normality // Economics Letters. – 2008. – Vol. 99, N 1. – P. 30–32.

36. *Zamanzade E., Arghami N.R.* Testing normality based on new entropy estimators // Journal of Statistical Computation and Simulation. – 2012. – Vol. 82, N 11. – P. 1701–1713.

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 68 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 his 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, Center for Statistical Technology, Novosibirsk State Technical University. The main field of his scientific research is computer modeling technologies and research on statistical laws. He has 68 publications, including 1 monograph. E-mail: skyer@mail.ru

DOI: 10.17212/2782-2001-2022-2-21-38

Nonparametric goodness-of-fit tests for normality testing under rounding-off measurements[*]

B.Yu. LEMESHKO[a], *S.B. LEMESHKO*[b]

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

[a] *lemeshko@ami.nstu.ru* [b] *skyer@mail.ru*

Abstract

When analyzing measurement series in various applications, the verification of whether measurement errors belong to the normal law is considered as a mandatory procedure. For this purpose, various special tests for testing hypotheses about normality can be used; non-parametric tests of goodness or chi-square tests can be used.

When using nonparametric goodness-of-fit tests to test normality, it must be taken into account that a complex hypothesis is being tested. When testing a complex hypothesis, the distributions of the statistics of the goodness-of-fit tests differ significantly from the classical ones that occur when testing simple hypotheses.

It is known that the presence of rounding errors can significantly change the distribution of test statistics. In such situations, ignoring the fact of influence can lead to incorrect conclusions about the results of the normality test.

In metrology, when carrying out high-precision measurements, as a rule, scientists do not even allow thoughts about the possible influence of Δ rounding errors on the results of statistical analysis. This allows the possibility of incorrect conclusions since there is no influence not only at small Δ, but at values of Δ much less than the standard deviation σ of the measurement error distribution law and sample sizes n not exceeding some maximum values. For sample sizes larger

[*] *Received 17 March 2022.*

than these maximum values, the real distributions of the test statistics deviate from the asymptotic ones towards larger statistics values.

In this work, based on real and well-known data, using statistical modeling methods, we demonstrate the dependence of the distributions of statistics of nonparametric goodness-of-fit tests when testing normality on the ratio of Δ and σ for specific n. The possibility of correct application of the tests under the influence of rounding errors on the conclusions is shown and implemented.

Keywords: normality check, nonparametric goodness-of-fit tests, composite hypothesis, distribution of statistics, round-off errors, achieved significance level, error of the 1st kind, error of the 2nd kind, simulation

REFERENCES

1. Lemeshko B.Yu. *Kriterii proverki otkloneniya raspredeleniya ot normal'nogo zakona: rukovodstvo po primeneniyu* [Tests for checking the deviation from normal distribution law. Guide on the application]. Moscow: INFRA–M, 2015, 160 p. DOI: 10.12737/6086.

2. Lemeshko B.Yu., Lemeshko S.B. 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*, 2021, no. 2 (82), pp. 47–66. DOI: 10.17212/2782-2001-2021-2-47-66.

3. Lemeshko B.Yu., Lemeshko S.B. Effect of the roundoff on the properties of criteria for testing statistical hypotheses. *Optoelectronics, Instrumentation and Data Processing*, 2020, vol. 56, no. 3, pp. 35–45. DOI: 10.3103/S8756699020030103. Translated from *Avtometriya*, 2020, vol. 56, no. 3, pp. 35–45. DOI: 10.15372/AUT20200305.

4. Lemeshko B.Yu., Lemeshko S.B. O vliyanii 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, no. 53. pp. 47–60. DOI: 10.17223/19988605/53/5.

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

6. Lemeshko B.Yu. Lemeshko 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, no. 44. pp. 40–49. DOI: 10.17223/19988605/44/5.

7. Fisher R.A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 1936, vol. 7, pp. 179–188.

8. 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.

9. Bol'shev L.N., Smirnov N.V. *Tablitsy matematicheskoi statistiki* [Tables for mathematical statistics]. Moscow, Nauka Publ., 1983. 416 p.

10. 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, pp. 193–212.

11. Anderson T.W., Darling D.A. A test of goodness of fit. *Journal of the American Statistical Association*, 1954, vol. 29, pp. 765–769.

12. Kuiper N.H. Tests concerning random points on a circle. *Indagationes Mathematicae (Proceedings)*, 1960, vol. 63, pp. 38–47.

13. Lemeshko B.Yu. *Neparametricheskie kriterii soglasiya: rukovodstvo po primeneniyu* [Nonparametric goodness-of-fit tests. Guide on the application]. Moscow, Infra-M Publ., 2014. 163 p. DOI: 10.12737/11873.

14. Watson G.S. Goodness-of-fit tests on a circle. 1. *Biometrika*, 1961, vol. 48, no. 1–2, pp. 109–114.

15. Watson G.S. Goodness-of-fit tests on a circle. 2. *Biometrika*, 1962, vol. 49, no. 1–2, pp. 57–63.

16. Zhang J. *Powerful goodness-of-fit and multi-sample tests.* PhD Thesis. York University. Toronto, 2001. 113 p. Available at: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (accessed 27.05.2022).

17. Zhang J. Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B*, 2002, vol. 64, no. 2, pp. 281–294.

18. Lemeshko B.Yu., Lemeshko S.B. Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. Pt. 1. *Measurement Techniques*, 2009, vol. 52, no. 6, pp. 555–565.

19. Lemeshko B.Yu., Lemeshko 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, no. 3, pp. 460–471.

20. Lemeshko B.Yu., Gorbunova A.A. Application of nonparametric Kuiper and Watson tests of goodness-of-fit for composite hypotheses. *Measurement Techniques*, 2013, vol. 56, no. 9, pp. 965–973.

21. Lemeshko B.Yu., Lemeshko 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 27.05.2022).

22. Pearson E.S., D’Agostino R.B., Bowman K.O. Tests for departure from normality: comparison of powers. *Biometrika*, 1977, vol. 64, pp. 231–246. DOI: 10.1093/biomet/64.2.427-a.

23. Tricker A.R. The effect of rounding on the significance level of certain normal test statistics. *Journal of Applied Statistics*, 1990, vol. 17, no. 1, pp. 31–38. DOI: 10.1080/757582644.

24. Tricker A.R. The effect of rounding on the power level of certain normal test statistics. *Journal of Applied Statistics*, 1990, vol. 17, no. 2, pp. 219–228. DOI: 10.1080/757582833.

25. Deidda R., Puliga M. Sensitivity of goodness-of-fit statistics to rainfall data rounding off. *Physics and Chemistry of the Earth*, 2006, vol. 31, pp. 1240–1251. DOI: 10.1016/j.pce.2006.04.041.

26. Uyanto S.S. An extensive comparisons of 50 univariate goodness-of-fit tests for normality. *Austrian Journal of Statistics*, 2022, vol. 51, pp. 45–97.

27. Frosini B.V. A survey of a class of goodness-of-fit statistics. *Metron*, 1978, vol. 36, no. 1–2, pp. 3–49.

28. Epps T.W., Pulley L.B. A test for normality based on the empirical characteristic function. *Biometrika*, 1983, vol. 70, pp. 723–726.

29. Hegazy Y.A.S., Green J.R. Some new goodness-of-fit tests using order statistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 1975, vol. 24, no. 3, pp. 299–308.

30. David H.A., Hartley H.O., Pearson E.S. The distribution of the ratio? In a single normal sample, of range to standard deviation. *Biometrika*, 1964, vol. 512, no. 3–4, pp. 484–487.

31. Geary R.C. Testing for Normality. *Biometrika*, 1937, vol. 34, pp. 209–242.

32. D’Agostino R.B. Transformation to normality of the null distribution of g_1. *Biometrika*, 1970, vol. 57 (3), pp. 679–681.

33. Chen L., Shapiro S.S. An alternative test for normality based on normalized spacings. *Journal of Statistical and Simulation*, 2012, vol. 53, pp. 269–288.

34. Desgagne A., Micheaux P.L. de. A Powerful and interpretable alternative to the Jarque–Bera test of normality based on 2nd-power skewness and kurtosis, using the rao's score test on the APD family. *Journal of Applied Statistics*, 2018, vol. 45, no. 13, pp. 2307–2327.

35. Gel Y.R., Gastwirth J.L. A robust modification of the Jarque–Bera test of normality. *Economics Letters*, 2008, vol. 99, no. 1, pp. 30–32.

36. Zamanzade E., Arghami N.R. Testing normality based on new entropy estimators. *Journal of Statistical Computation and Simulation*, 2012, vol. 82, no. 11, pp. 1701–1713.

Для цитирования:

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.

For citation:

Lemeshko B.Yu., Lemeshko 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.