ISSN 2782-2001
Системы анализа
и обработки данных
том 97, № 1, 2025, с. 49–72

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 97, No. 1, 2025, pp. 49–72

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.23

DOI: 10.17212/2782-2001-2025-1-49-72

Об исследовании свойств оценок и влияния методов оценивания на распределения статистик критериев проверки гипотез*

Б.Ю. ЛЕМЕШКО[a], С.Б. ЛЕМЕШКО[b]

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

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

Выражения статистик многих критериев проверки статистических гипотез содержат оценки параметров или числовых характеристик, таких как математическое ожидание или дисперсия, предполагаемых законов распределения вероятностей. Оценки можно находить различными методами. Исследование асимптотических свойств некоторой оценки аналитическими методами, как правило, представляет собой нетривиальную задачу. Еще сложнее аналитическими методами исследовать свойства оценок при ограниченных объемах выборок, когда реальные свойства оценок могут существенно отличаться от асимптотических. Компьютерные технологии позволяют исследовать асимптотические свойства оценок и уточнять, при каких объемах выборок можно опираться на асимптотические свойства, а также показывают, как реальные свойства отличаются от асимптотических при ограниченных объемах выборок.

Как следствие, распределения статистик критериев при справедливости проверяемой гипотезы зависят от свойств используемых оценок, то есть от применяемого метода оценивания. Исследование аналитическими методами изменившихся свойств критериев вследствие выбора другого метода оценивания также оказывается нетривиальной задачей. В то же время компьютерные технологии позволяют не только решать такие задачи, но и обеспечивать корректное применение соответствующих критериев в изменившихся условиях.

В рамках развиваемой программной системы продемонстрированы возможности компьютерных технологий моделирования и анализа данных при исследовании свойств различных оценок и влияния вида оценок на распределения статистик критериев. Исследованы свойства *MD*-оценок, минимизирующих статистики различных непараметрических критериев согласия. Показано, что эти оценки, в отличие от оценок максимального правдоподобия, не являются асимптотически эффективными, и их асимптотические распределения различаются. Показано, что при проверке сложных гипотез, когда параметры закона находятся по анализируемой выборке, распределения статистик всех непараметрических критериев согласия и критерия χ^2 Пирсона, имеющие место при справедливости проверяемой гипотезы H_0, зависят от метода оце-

* *Статья получена 30 декабря 2024 г.*

нивания параметров. При этом распределения статистики одного и того же критерия различаются при разных методах оценивания. Полученные оценки мощности критериев, знание асимптотических свойств оценок и характера влияния этих оценок на распределения статистик критериев позволяют судить о предпочтительности комбинаций применяемых критериев с методами оценивания.

Программная система позволяет проводить аналогичные исследования относительно любых параметров более чем 30 параметрических моделей законов распределения, включенных в систему, и относительно любых из более чем десятка критериев согласия.

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

ВВЕДЕНИЕ

В настоящей работе преследуются две цели: 1) показать, как используемые методы оценивания параметров предполагаемых законов распределения влияют на свойства различных критериев проверки статистических гипотез, в статистиках которых предусмотрено использование тех или иных оценок параметров; 2) продемонстрировать, как компьютерные технологии моделирования и анализа данных позволяют при современном уровне вычислительных средств закрывать «белые пятна» в математической статистике, способствуя изучению статистических свойств оценок и критериев, что, в общем, расширяет прикладные возможности соответствующих методов. Компьютерные технологии позволяют исследовать асимптотические свойства оценок и уточнять, при каких объемах выборок можно опираться на асимптотические свойства, а также показывают, как при ограниченных объемах выборок реальные свойства могут отличаться от асимптотических. Более того, эти технологии дают возможность корректного применения множества статистических критериев в тех приложениях, где нарушаются стандартные «классические» предположения.

Среди множества критериев, ориентированных на проверку гипотезы о принадлежности выборок нормальному закону, можно указать ряд критериев, при вычислении статистик которых предусматривается применение смещенных оценок дисперсии (например, критерий Фросини [1], критерий Гири [2] и др.), и критерии, предусматривающие использование несмещенных оценок дисперсии (например, критерии Хегази – Грина [3] и др. [4]). Замена одних оценок на другие приводит к определенным *изменениям* распределений статистик соответствующих критериев. Если этого не учитывать, выводы могут оказаться некорректными.

Иногда в руководящих документах, касающихся анализа результатов измерений, встречаются рекомендации, например, такого типа: оценку дисперсии, используемую далее при вычислении статистики некоторого критерия, заменить, как крайне неустойчивую, ее робастным аналогом на базе размаха выборки. При этом к сожалению, никак не учитываются последствия такой замены, связанные с возможным *изменением* распределения статистики этого критерия.

В критерии согласия χ^2 Пирсона, наиболее часто используемом в приложениях, предусматривается разбиение области определения случайной величины на k непересекающихся интервалов, подсчет числа наблюдений n_i, попавших в *i*-й интервал, и вычисление статистики:

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

где $P_i(\theta)$ – вероятность попадания в интервал в соответствии с функцией распределения $F(x,\theta)$ предполагаемого закона, $n = \sum_{i=1}^{k} n_i$ – объем выборки. При проверке простой гипотезы асимптотическим распределением статистики является χ^2_{k-1}-распределение. При проверке сложной гипотезы и вычислении r компонент вектора параметров θ в результате минимизации статистики критерия асимптотическим распределением статистики является χ^2_{k-r-1}-распределение. Если же оценки будут находиться каким-то методом по исходным негруппированным данным, то распределение статистики критерия уже не будет χ^2_{k-r-1}-распределением и будет зависеть от используемого метода оценивания. Этим фактом на практике часто пренебрегают, что приводит, как правило, к увеличению вероятности ошибок 1-го рода (отклонению справедливой проверяемой гипотезы H_0).

Эта проблема для ситуации использования оценок максимального правдоподобия (ОМП) по негруппированным данным получила решение в работах Никулина [5, 6] и Рао – Робсона [7] введением поправки к статистике Пирсона, учитывающей (для законов распределения, определяемых только параметрами сдвига и масштаба) разность информационных матриц Фишера по точечным и сгруппированным выборкам [4]. Этот критерий получил название критерия Никулина – Рао – Робсона [8]. Никулиным были рассмотрены и другие модификации критерия [9]. Однако, несмотря на хорошие свойства, критерий Никулина – Рао – Робсона не получил широкого распространения на практике по банальной причине: вид поправки для статистики критерия зависит от закона, относительно которого проверяется гипотеза, и ее нахождение не всегда тривиально и требует определенных усилий. Устранение такой причины решается наличием соответствующего программного обеспечения.

Эта же проблема, ограничивающая применение критерия χ^2 Пирсона в случае использования ОМП по негруппированным данным, в современных условиях может решаться и другим образом. Для принятия решения о результатах проверки гипотезы нам необходимо знать распределение статистики $G(\chi^2 | H_0)$-критерия, имеющее место при справедливости проверяемой гипотезы H_0, чтобы найти по этому распределению достигнутый уровень значимости P_v. Распределение $G(\chi^2 | H_0)$ в данном случае зависит от метода оценивания, числа интервалов, способа разбиения области определения случай-

ной величины на интервалы и от объема выборки n. Поэтому, используя метод Монте-Карло, *несложно* либо *смоделировать* распределение статистики критерия в указанных конкретных условиях, либо найти оценку P_v. Этим подходом открывается возможность применения критерия χ^2 в нестандартных условиях приложений с формированием корректных выводов по результатам проверки гипотезы.

Большинство ошибок использования непараметрических критериев согласия [10, 11] также связано с некорректным применением в условиях проверки сложных гипотез. В ситуации проверки сложной гипотезы на распределения статистик $G(S|H_0)$-критериев воздействует следующий ряд факторов [12]:

(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 [13];

(e) – в некоторых ситуациях распределения статистик зависят от конкретных значений параметров закона $F(x,\theta)$ (например, от значений параметров формы обобщенного нормального закона [14], гамма-распределения, бета-распределений и т. п.).

Очевидно, что в присутствии факторов (d) и (e) корректность выводов по непараметрическим критериям согласия можно обеспечить только *за счет* использования методов *статистического моделирования*.

Выбирая определенные методы оценивания, мы должны ориентироваться на свойства соответствующих оценок. Хорошо, если оценки являются асимптотически эффективными. Тогда возникает вопрос: начиная с каких объемов выборок можно пользоваться свойством асимптотической эффективности? Или чем отличаются свойства оценок при ограниченных объемах выборок от асимптотических? Например, ОМП параметров законов распределения, вычисляемые по цензурированным выборкам, в основном являются асимптотически эффективными. Однако при ограниченных объемах выборок и увеличении степени цензурирования распределения оценок по цензурированным данным, как показано в [15, 16], оказываются асимметричными, а оценки – смещенными.

Степень уверенности в корректности статистических выводов по используемым критериям и в адекватности построенных вероятностных моделей базируется на знании реальных свойств применяемых методов и статистических критериев, которыми они обладают при использовании в конкретном приложении. Эти свойства могут существенно отличаться от тех, которые они имеют в условиях стандартных предположений. Как правило, во многих ситуациях единственный выход для получения требуемых знаний о реальных свойствах методов и критериев видится в использовании компьютерных технологий и статистического моделирования.

1. МЕТОДЫ ОЦЕНИВАНИЯ ПАРАМЕТРОВ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ

При идентификации закона распределения наблюдаемой случайной величины и выборе из множества возможных претендентов наиболее подходящей параметрической модели $F(x, \theta)$, где θ – скалярный или векторный параметр закона, возникает необходимость в оценивании параметров этих моделей на основе имеющейся выборки $x_1, x_2, \ldots, x_n$ объемом n или, может быть, совокупности выборок.

Для оценки неизвестных параметров законов могут использоваться различные методы. В ситуациях, когда вид предполагаемой модели закона достаточно близок к неизвестному истинному закону, оценки параметров, получаемые разными методами, могут оказываться очень близкими. Если же наши предположения о виде закона далеки от истины или вследствие некоторых причин данные «засорены» результатами посторонних экспериментов или наличием грубых ошибок измерений, различные методы могут приводить к резко отличающимся оценкам. Получаемые результаты в обеих ситуациях объясняются различием статистических свойств оценок, вычисляемых разными методами.

Существует множество критериев проверки различных статистических гипотез, в которых предусмотрено использование оценок, получаемых определенным методом, так как именно в такой ситуации известно распределение статистики критерия $G(S|H_0)$ при справедливости проверяемой гипотезы H_0 и, следовательно, гарантирована корректность статистического вывода по результатам проверки гипотезы по этому критерию. Замена метода оценивания параметров может изменить свойства критерия (распределение его статистики). Если этого не учитывать, то выводы о результатах проверки могут оказаться некорректными.

В частности, при проверке сложных гипотез от метода оценивания параметров существенно зависят распределения статистик непараметрических критериев согласия: каждому типу оценок при конкретной сложной проверяемой гипотезе соответствует свое предельное распределение $G(S|H_0)$ статистики критерия [12]. По-видимому, впервые о такой зависимости было указано в работе [13].

В приложениях обычно отдают предпочтение методу максимального правдоподобия, так как оценки максимального правдоподобия (ОМП) обладают лучшими асимптотическими свойствами [17, 18]. Как правило, они оказываются асимптотически эффективными.

Оценки максимального правдоподобия $\hat{\theta}$ вычисляют в результате максимизации по θ-функции правдоподобия

$$L(\theta) = \gamma \prod_{i=1}^{n} f(x_i, \theta),$$

где $f(x, \theta)$ – плотность закона распределения, или ее логарифма

$$\ln L(\theta) = \ln \gamma + \sum_{i=1}^{n} \ln f(x_i, \theta).$$

Чаще всего в случае скалярного параметра θ ОМП определяют как решение уравнения, а в случае векторного параметра – как решение системы уравнений правдоподобия вида

$$\frac{\partial \ln L(\theta)}{\partial \theta_l} = \sum_{i=1}^{n} \frac{\partial \ln f(x_i, \theta)}{\partial \theta_l} = 0,\ l = \overline{1, m},$$

где m – размерность вектора параметров θ. В общем случае эта система является нелинейной и, за редким исключением, решается только численно.

Оценки максимального правдоподобия обладают лучшими статистическими свойствами, но, к сожалению, в большинстве своем не являются робастными, они чувствительны к наличию в выборках аномальных измерений, к отклонению реально наблюдаемого закона от предполагаемого, к засорению выборок данными, принадлежащими другому закону. Всё это подтверждается опытом эксплуатации программной системы [19] и многочисленными результатами модельных экспериментов.

В работах [20, 21] робастность ОМП исследована с позиций *функции влияния*, предложенной Хэмпелем [22]. Анализ функций влияния ОМП параметров различных распределений, в том числе того множества законов, которое включено в программную систему [19], показал, что за редкими исключениями функции влияния ОМП по точечным (негруппированным) данным оказываются неограниченными. Это является свидетельством неробастности соответствующих оценок. В то же время функции влияния ОМП по группированным данным всегда ограничены, а такие оценки оказываются робастными [23, 24]. Робастными являются оптимальные L-оценки параметров сдвига и масштаба распределений по выборочным квантилям [25, 26]. Эти оценки строятся в виде линейных комбинаций выборочных квантилей, соответствующих асимптотически оптимальному группированию, при котором минимизируются потери в информации Фишера, связанные с группированием. Функции влияния этих оценок также ограничены по абсолютной величине.

На практике часто используют так называемые оценки по методу моментов. Для некоторых законов оценки параметров по методу моментов совпадают с ОМП, как, например, в случае нормального закона. Но в общем случае свойства оценок по методу моментов отличаются от свойств ОМП.

Вторым часто используемым видом оценок являются различные *MD*-оценки, которые получаются при минимизации некоторого расстояния между эмпирической и теоретической функцией распределения. Например, в качестве такого расстояния могут рассматриваться статистики различных критериев согласия. В этом случае по параметру θ будет минимизироваться статистика соответствующего критерия. По сравнению с ОМП, *MD*-оценки более устойчивы к отклонениям наблюдаемого закона от предполагаемого.

Статистика критерия Колмогорова с поправкой Большева [27] имеет вид

$$S_K = \sqrt{n} D_n + \frac{1}{6\sqrt{n}},$$

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

При использовании статистики Колмогорова *MD*-оценка $\tilde{\theta}$ находится в результате минимизации S_K:

$$\tilde{\theta} = \arg \min_{\theta} S_K . \quad (1)$$

При минимизации статистики критерия Крамера – Мизеса – Смирнова [27], имеющей вид

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

получаем *MD*-оценку $\hat{\theta}$:

$$\hat{\theta} = \arg \min_{\theta} S_{CMS} . \quad (2)$$

При минимизации статистики критерия Андерсона – Дарлинга [28, 29]

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

получаем *MD*-оценку $\breve{\theta}$:

$$\breve{\theta} = \arg \min_{\theta} S_{AD} . \quad (3)$$

В качестве основы для *MD*-оценок могут быть взяты также статистики критериев Купера [30], Ватсона [31, 32] и критериев Жанга [33].

Об исследованиях статистических свойств *MD*-оценок, в отличие от ОМП, известно мало. По крайней мере, авторам этой работы они не известны. Но встроенные в программную систему [19] средства моделирования позволяют численными методами восполнить недостающие знания.

2. ИССЛЕДОВАНИЕ АСИМПТОТИЧЕСКИХ СВОЙСТВ ОЦЕНОК

В данном случае на примере нормального закона с плотностью

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

проведем сравнительный анализ статистических свойств оценок его параметров, получаемых методом максимального правдоподобия и с использованием *MD*-оценок (1)–(3).

Известно, что ОМП параметров сдвига θ_0 и масштаба θ_1 нормального закона представляют собой асимптотически несмещенные, состоятельные и асимптотически эффективные оценки. Это означает, что при оценивании только параметра сдвига по выборке объемом n распределение его оценки должно быть нормальным с математическим ожиданием $E[\hat{\theta}_{0,n}] = \theta_0$ и дисперсией $D[\hat{\theta}_{0,n}] = n^{-1}J^{-1}(\theta_0)$, где $J(\theta_0)$ – количество информации Фишера о параметре θ_0 в одном наблюдении. Аналогично, при оценивании только параметра масштаба θ_1 оценка должна подчиняться нормальному закону с математическим ожиданием $E[\hat{\theta}_{1,n}] = \theta_1$ и дисперсией $D[\hat{\theta}_{1,n}] = n^{-1}J^{-1}(\theta_1)$, где $J(\theta_1)$ – количество информации Фишера о параметре θ_1. В случае нормального закона $J(\theta_0) = 1/\theta_1^2$, а $J(\theta_1) = 2/\theta_1^2$.

В данном случае наша задача заключается в проверке асимптотических свойств различных видов оценок. Воспользовавшись возможностями программной системы [19], будем моделировать выборки объемом $n = 10^3$ в соответствии с нормальным законом с параметрами $\theta_0 = 1,\ \theta_1 = 1$. В первом случае будем оценивать только параметр θ_1, вычисляя оценки четырьмя различными методами. Задав количество экспериментов $N = 16\,600$, получим 4 выборки такого объема с оценками $\hat{\theta}_1$, $\tilde{\theta}_1$, $\widehat{\theta}_1$, $\breve{\theta}_1$ параметра. Если какая-то из оценок является асимптотически эффективной, то соответствующая ей выборка оценок должна принадлежать асимптотически нормальному закону с параметром сдвига 1 и масштабом $\sqrt{10^{-3}/2} \approx 0.02236068$.

Результаты моделирования распределений оценок представлены на рис. 1 и 2, на которых слева показаны эмпирические распределения оценок $F_N(\hat{\theta}_1)$, $F_N(\tilde{\theta}_1)$, $F_N(\widehat{\theta}_1)$, $F_N(\breve{\theta}_1)$ и функция распределения нормального закона, соответствующего асимптотически эффективной оценке, а справа – гистограмма распределения соответствующей оценки и плотность асимптотически эффективной оценки.

Как можно видеть на рис. 1, *а*, эмпирическое распределение $F_N(\hat{\theta}_1)$ «совпадает» с функцией распределения нормального закона $N(1,\ 0.02236068)$, соответствующего асимптотически эффективной оценке. Вид гистограммы на рис. 1, *б* демонстрирует ее близость к соответствующей функции плотности. Это же подтверждают представленные в табл. 1 результаты проверки простой гипотезы о принадлежности распределения оценок $\hat{\theta}_1$ закону $N(1,\ 0.02236068)$ по критериям Колмогорова, Крамера – Мизеса – Смирнова, Андерсона – Дарлинга и χ^2 Пирсона.

Заметим, что при использовании критерия χ^2 Пирсона выборка оценок $\hat{\theta}_1$ разбивалась на 10 асимптотически оптимальных интервалов [14]. В этом случае данный критерий имеет максимальную мощность относительно близких конкурирующих гипотез [34, 35]. Таким образом, все результаты подтверждают факт асимптотической эффективности ОМП.

Напротив, *MD*-оценки $\tilde{\theta}_1$, $\hat{\theta}_1$ и $\breve{\theta}_1$ не являются асимптотически эффективными. Это подтверждает картина, представленная на рис. 1, *в*, *г* и на рис. 2, где эмпирические распределения оценок $F_N(\tilde{\theta}_1)$, $F_N(\hat{\theta}_1)$, $F_N(\breve{\theta}_1)$ явно отличаются от распределения асимптотически эффективной оценки. А проверки простых гипотез о принадлежности $F_N(\tilde{\theta}_1)$, $F_N(\hat{\theta}_1)$, $F_N(\breve{\theta}_1)$ нормальному закону асимптотически эффективной оценки по всем применяемым критериям отклоняются (табл. 1). Более того, по виду $F_N(\tilde{\theta}_1)$, $F_N(\hat{\theta}_1)$, $F_N(\breve{\theta}_1)$ можно обратить внимание на факт смещенности оценок (рис. 1, *в*, 2, *а*, *г*).

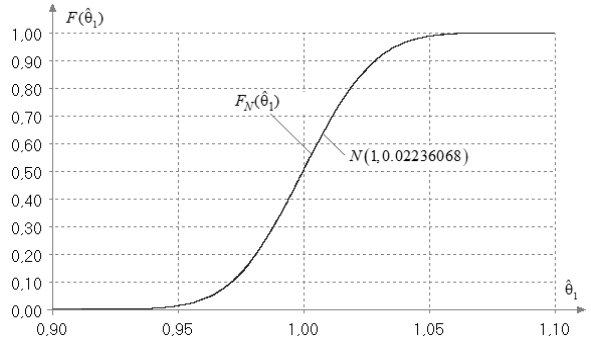
Таблица 1

Table 1

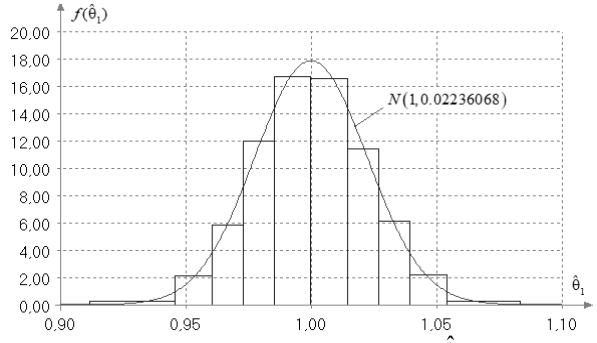
Результаты проверки простых и сложных гипотез о принадлежности выборок оценок $\hat{\theta}_1$, $\tilde{\theta}_1$, $\hat{\theta}_1$ и $\breve{\theta}_1$ нормальному закону

The results of testing simple and composite hypotheses about the belonging of samples of estimates $\hat{\theta}_1$, $\tilde{\theta}_1$, $\hat{\theta}_1$, and $\breve{\theta}_1$ to the normal law

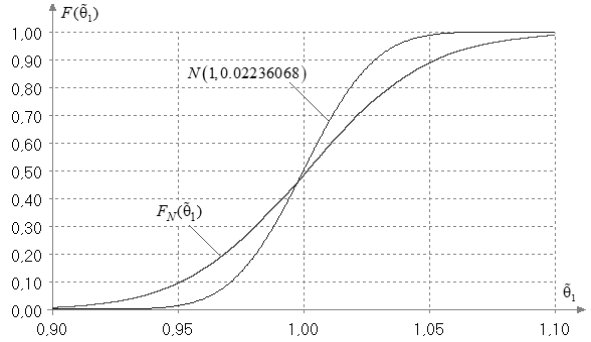
Критерий	Простая гипотеза		Сложная гипотеза	
	Статистика	P_v	Статистика	P_v
ОМП $\hat{\theta}_1$	$N(1, 0.02236068)$		$N(0.999973, 0.022366)$	
Пирсона	5.8946	0.7504	6.030986	0.5361
Колмогорова	0.68451	0.7368	0.623803	0.4846
Крамера – Мизеса – Смирнова	0.08452	0.6662	0.069179	0.2879
Андерсона – Дарлинга	0.48672	0.7604	0.438981	0.2958
MD-оценка $\tilde{\theta}_1$	$N(1, 0.02236068)$		$N(1.002029, 0.040524)$	
Пирсона	30273.99	0.0000	89.8896	0.000
Колмогорова	18.5146	0.0000	1.99452	0.000
Крамера – Мизеса – Смирнова	147.9105	0.0000	0.87782	0.000
Андерсона – Дарлинга	564.000	0.0000	6.08439	0.000
MD-оценка $\hat{\theta}_1$	$N(1, 0.02236068)$		$N(1.001813, 0.0282137)$	
Пирсона	2867.007	0.0000	5.90033	0.5514
Колмогорова	9.8156	0.0000	0.65592	0.3995
Крамера – Мизеса – Смирнова	33.7463	0.0000	0.07672	0.2280
Андерсона – Дарлинга	326.8707	0.0000	0.48643	0.2272
MD-оценка $\breve{\theta}_1$	$N(1, 0.02236068)$		$N(1.002151, 0.024538)$	
Пирсона	505.623	0.0000	8.90452	0.2596
Колмогорова	6.4200	0.0000	0.73944	0.2218
Крамера – Мизеса – Смирнова	15.768	0.0000	0.10029	0.1103
Андерсона – Дарлинга	108.9496	0.0000	0.59371	0.1239



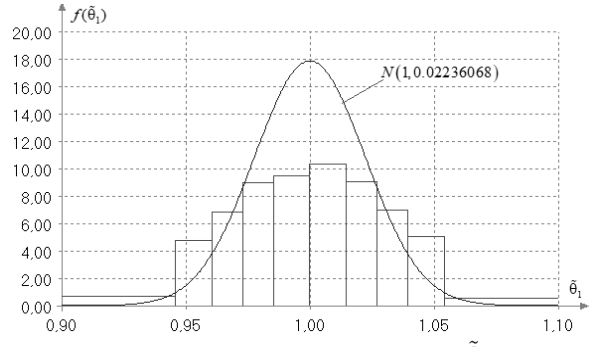
а – эмпирическое распределение оценок $\hat{\theta}_1$ и нормальный закон асимптотически эффективной оценки



б – гистограмма распределения оценок $\hat{\theta}_1$ и плотность нормального закона асимптотически эффективной оценки



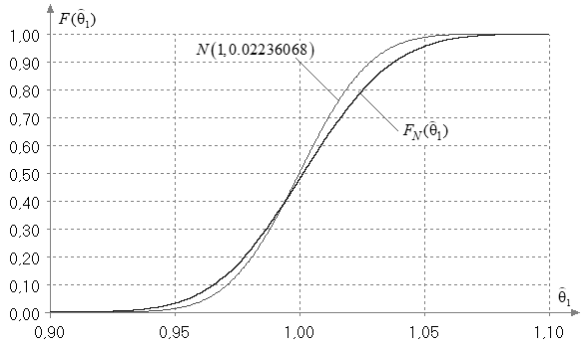
в – эмпирическое распределение *MD*-оценок $\tilde{\theta}_1$ и нормальный закон асимптотически эффективной оценки



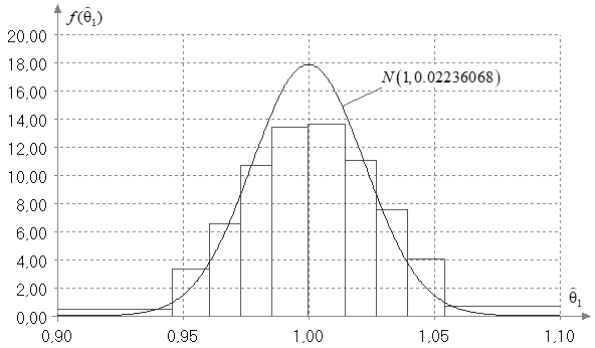
г – гистограмма распределения *MD*-оценок $\tilde{\theta}_1$ и плотность нормального закона асимптотически эффективной оценки

Рис. 1. Асимптотические свойства оценок параметра масштаба нормального закона: ОМП (*а*, *б*) и *MD*-оценки (*в*, *г*), минимизирующей статистику S_K

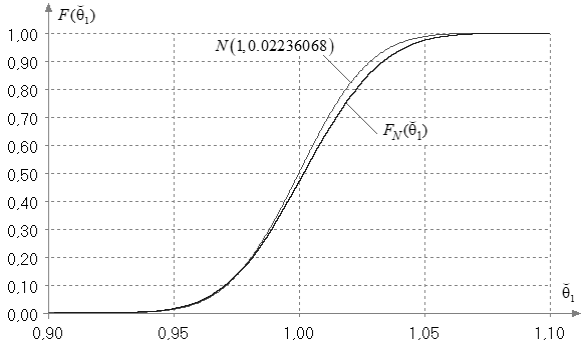
Fig. 1. Asymptotic properties of the estimates of the scale parameter of the normal law: MLE (*a*, *b*) and *MD*-estimate (*c*, *d*), minimizing the statistics S_K



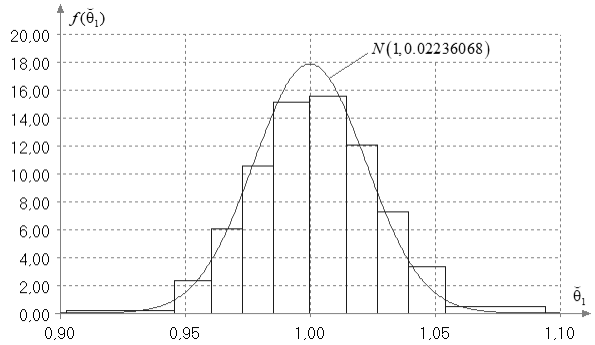
а – эмпирическое распределение *MD*-оценок $\hat{\theta}_1$ и нормальный закон асимптотически эффективной оценки



б – гистограмма распределения *MD*-оценок $\hat{\theta}_1$ и плотность нормального закона асимптотически эффективной оценки



в – эмпирическое распределение *MD*-оценок $\breve{\theta}_1$ и нормальный закон асимптотически эффективной оценки



г – гистограмма распределения *MD*-оценок $\breve{\theta}_1$ и плотность нормального закона асимптотически эффективной оценки

Рис. 2. Асимптотические свойства *MD*-оценок параметра масштаба нормального закона, минимизирующих статистики S_{CMS} (*а*, *б*) и S_{AD} (*в*, *г*)

Fig. 2. Asymptotic properties of *MD*-estimates of the scale parameter of the normal law, minimizing the statistics S_{CMS} (a, b) and S_{AD} (c, d)

В то же время распределения всех оценок $\hat{\theta}_1$, $\tilde{\theta}_1$, $\widehat{\theta}_1$, $\breve{\theta}_1$ хорошо описываются нормальными законами, найденными при проверке сложных гипотез, как можно видеть в табл. 1. О величине смещения оценок можно судить, например, по отклонению $E[\hat{\theta}_1]$ от единицы, а о том, насколько распределения оценок отклоняются от распределения асимптотически эффективной оценки, – по разности $\sqrt{D[\hat{\theta}_1]} - 0.02236068$.

Можно заметить, что для распределения *MD*-оценки $\tilde{\theta}_1$ гипотеза о нормальности отклоняется, причем по всем критериям достигнутый уровень значимости $P_v = 0$. Хотя визуально отклонение $F_N(\tilde{\theta}_1)$ от нормального закона $N(1.002029, 0.040524)$ едва заметно, но мощность критериев при таких больших выборках (N = 16 600) очень высокая и применяемые критерии однозначно «отклоняют» гипотезу.

Аналогичным образом при тех же значениях параметров нормального закона $\theta_0 = 1$, $\theta_1 = 1$ (при $n = 1000$, $N = 16\,600$) рассмотрим ситуацию оценивания четырьмя различными методами только параметра сдвига θ_0 нормального закона. В этом случае получим 4 выборки оценок $\hat{\theta}_0$, $\tilde{\theta}_0$, $\widehat{\theta}_0$, $\breve{\theta}_0$ параметра θ_0. Если какая-то из оценок является асимптотически эффективной, то соответствующая ей выборка оценок должна принадлежать нормальному закону с параметром сдвига 1 и масштабом $\sqrt{10^{-3}} \approx 0.03162278$.

В данном случае эмпирические распределения рассматриваемых оценок оказались слишком близкими, чтобы можно было содержательно проиллюстрировать их различие на рисунках, поэтому приводим только табл. 2 с результатами проверки простых гипотез о принадлежности выборок нормальному закону, соответствующему асимптотически эффективной оценке, и сложных гипотез о их принадлежности нормальному закону. Эти результаты позволяют судить о реальных свойствах оценок. Очевидно, что асимптотически эффективной оценкой является только ОМП $\hat{\theta}_0$.

Как можно видеть, наиболее близкими по асимптотическим свойствам к ОМП являются *MD*-оценки $\breve{\theta}_0$, которые получаются в результате минимизации статистики критерия Андерсона – Дарлинга. И надо полагать, что все рассматриваемые оценки параметра сдвига являются (асимптотически) несмещенными.

В табл. 2 указан нормальный закон, соответствующий асимптотически эффективной оценке параметра сдвига (при $n = 10^3$), о принадлежности которому проверяются простые гипотезы относительно анализируемых выборок оценок. Приведены нормальные законы (с указанием параметров сдвига и масштаба), описывающие распределения соответствующих оценок и полученные в процессе проверки сложных гипотез. Приведены значения статистик и достигнутые уровни значимости P_v по применяемым критериям согласия.

3. ВЛИЯНИЕ МЕТОДОВ ОЦЕНИВАНИЯ НА СТАТИСТИЧЕСКИЕ СВОЙСТВА КРИТЕРИЕВ СОГЛАСИЯ

Вид используемой оценки существенно влияет на распределения статистик критериев согласия. Степень влияния метода оценивания на распределение статистики показывает рис. 3, на котором изображены распределения статистик непараметрических критериев согласия Колмогорова $G(S_K|H_0)$ (рис. 3, *а*), Крамера – Мизеса – Смирнова $G(S_{CMS}|H_0)$ (рис. 3, *б*), Андерсона – Дарлинга $G(S_{AD}|H_0)$ (рис. 3, *в*) и χ^2 Пирсона $G(\chi^2|H_0)$ (рис. 3, *г*) при проверке сложной гипотезы о принадлежности выборок нормальному закону в случае применения для вычисления оценок вектора параметров, рассмотренных в работе методов.

Таблица 2

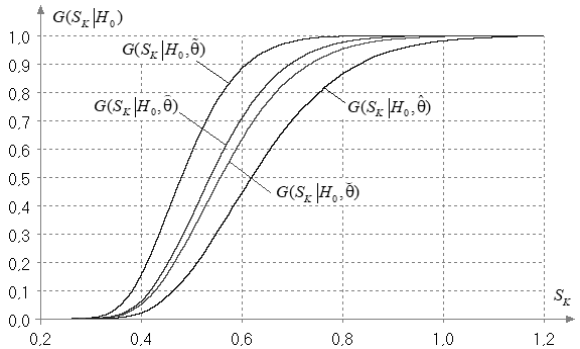
Table 2

Результаты проверки простых и сложных гипотез о принадлежности выборок оценок $\hat{\theta}_0$, $\tilde{\theta}_0$, $\widehat{\theta}_0$ и $\breve{\theta}_0$ нормальному закону

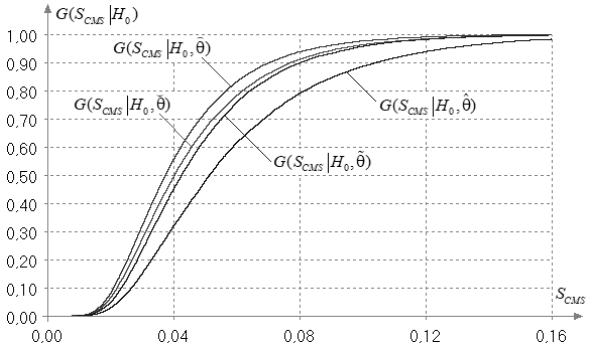
The results of testing simple and composite hypotheses about the belonging of samples of estimates $\hat{\theta}_0$, $\tilde{\theta}_0$, $\widehat{\theta}_0$, and $\breve{\theta}_0$ to the normal law

Критерий	Простая гипотеза		Сложная гипотеза	
	Статистика	P_v	Статистика	P_v
ОМП $\hat{\theta}_0$	$N(1, 0.03162278)$		$N(1.00018, 0.031787)$	
Пирсона	2.89460	0.9683	1.94355	0.9629
Колмогорова	0.74326	0.6385	0.52818	0.7556
Крамера – Мизеса – Смирнова	0.07989	0.6925	0.03042	0.8379
Андерсона – Дарлинга	0.50227	0.7445	0.19771	0.8913
MD-оценка $\tilde{\theta}_0$	$N(1, 0.03162278)$		$N(1.000117, 0.033436)$	
Пирсона	127.450	0.0000	11.6858	0.1114
Колмогорова	1.92859	0.0012	0.81864	0.1152
Крамера – Мизеса – Смирнова	1.17910	0.0009	0.09920	0.1140
Андерсона – Дарлинга	12.2006	0.0000	0.55896	0.1509
MD-оценка $\widehat{\theta}_0$	$N(1, 0.03162278)$		$N(1.000121, 0.033269)$	
Пирсона	85.8216	0.0000	9.02520	0.2509
Колмогорова	1.74374	0.0046	0.68479	0.3304
Крамера – Мизеса – Смирнова	1.01045	0.0023	0.07285	0.2570
Андерсона – Дарлинга	10.2669	0.0000	0.43703	0.2990
MD-оценка $\breve{\theta}_0$	$N(1, 0.03162278)$		$N(1.000142, 0.032342)$	
Пирсона	26.2124	0.0019	10.7229	0.1512
Колмогорова	0.9148	0.3726	0.70371	0.2894
Крамера – Мизеса – Смирнова	0.2174	0.2362	0.06994	0.2813
Андерсона – Дарлинга	2.0551	0.0857	0.40216	0.3616

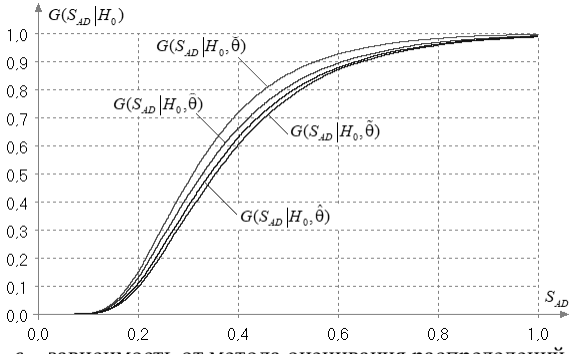
По существу, на рис. 3, *а–г* показаны асимптотические распределения статистик критериев, так как они моделировались при объемах выборок $n = 10^3$.



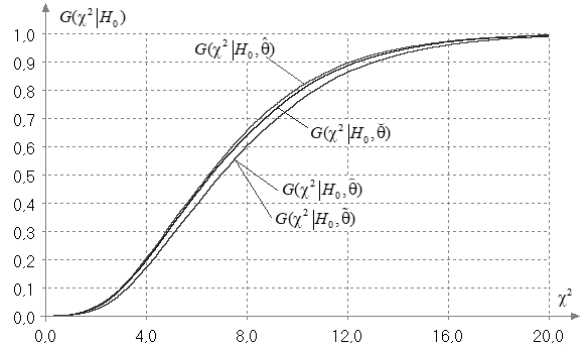
а – зависимость от метода оценивания распределений статистики критерия Колмогорова



б – зависимость от метода оценивания распределений статистики критерия Крамера – Мизеса – Смирнова



в – зависимость от метода оценивания распределений статистики критерия Андерсона – Дарлинга



г – зависимость от метода оценивания распределений статистики критерия χ^2 Пирсона

Рис. 3. Зависимость от методов оценивания распределений статистик критериев согласия при проверке нормальности с оцениванием двух параметров закона

Fig. 3. Dependence on methods of estimating distributions of statistics of goodness-of-fit tests when checking normality with the estimation of 2 parameters of the law

Все рассматриваемые критерии согласия правосторонние, т. е. проверяемая гипотеза H_0 отклоняется при больших значениях статистик. Распределения статистик всех этих критериев при проверке сложных гипотез смещаются в сторону меньших значений статистик относительно тех распределений, которым статистики подчиняются при проверке простых гипотез: относительно $K(S)$ в случае критерия Колмогорова, $a1(S)$ – критерия Крамера – Мизеса – Смирнова, $a2(S)$ – критерия Андерсона – Дарлинга [27, 12].

Если вид используемых оценок очень сильно отражается на распределениях статистик непараметрических критериев согласия, то точность оценивания параметров на этих распределениях сказывается в существенно меньшей степени [34].

В случае критерия χ^2 Пирсона моделируемые выборки разбивались на $k = 10$ асимптотически оптимальных интервалов [35, 36]. Все распределения на рис. 3, *г* расположены между распределениями χ^2_9 и χ^2_7, которым статистика подчиняется при проверке простой и сложной гипотез, если при проверке сложной гипотезы вектор оценок находится в результате минимизации статистики критерия. На рис. 3, *г* распределения $G(\chi^2 \left| H_0, \tilde{\theta}\right.)$ и $G(\chi^2 \left| H_0, \widehat{\theta}\right.)$ практически сливаются, но отличие между ними всё-таки есть.

В целом же из картины, представленной на рис. 3, очевидно, что при формировании статистического вывода о результатах проверки сложной гипотезы H_0 нельзя пренебрегать влиянием на распределения статистик критериев используемого метода оценивания параметров закона, относительно которого проверяется гипотеза.

Корректность применения различных непараметрических критериев согласия в рамках программной системы [19] в случае использования различных *MD*-оценок обеспечивается интерактивным моделированием распределений статистик критериев в этих условиях с последующим оцениванием по этим распределениям достигнутых уровней значимости P_v для применяемых критериев.

Следует напомнить, что распределения статистик критериев согласия в случае больших [37] и не только больших выборок дополнительно могут сильно изменяться под влиянием ошибок округления в анализируемых данных [11, 38–41]. В таких выборках оказывается много повторяющихся значений. Возможность такой ситуации предусмотрена в системе [19]. В таких случаях распределения статистик критериев могут моделироваться с учетом той величины ошибки округления Δ, которая соответствует анализируемой выборке.

Таким образом, к указанным выше факторам (a)–(e), влияющим на распределения статистик непараметрических критериев согласия при проверке сложных гипотез, необходимо добавить фактор (f), связанный с возможным влиянием ошибок округления. Если влияние такого фактора проявляется, то распределения статистик становятся зависящими от величины Δ и от объема выборки n.

4. СВЯЗЬ МОЩНОСТИ КРИТЕРИЕВ С МЕТОДАМИ ОЦЕНИВАНИЯ

При заданной вероятности α ошибки первого рода (при уровне значимости) мощность определяется величиной $1-\beta$, где β – вероятность ошибки второго рода (не отклонить проверяемую гипотезу H_0 при справедливости H_1). Мощность представляет собой вероятность отклонения проверяемой гипотезы H_0 при справедливости гипотезы H_1.

Для того чтобы оценить мощность критерия при заданном объеме выборки *n*, необходимо знать распределения статистики $G(S_n|H_0)$ и $G(S_n|H_1)$ соответствующего критерия при данном *n*. Знание распределения $G(S_n|H_1)$ – событие крайне редкое. Поэтому оценки мощности критериев практически всегда находят с использованием методов статистического моделирования.

В работах [42, 43] был проведен сравнительный анализ мощности непараметрических критериев согласия Колмогорова, Крамера – Мизеса – Смирнова, Андерсона – Дарлинга и критериев типа χ^2. В качестве конкурирующих гипотез рассматривались пары «нормальный закон – логистический закон» и «гамма-распределение – распределение Вейбулла». В данном случае ограничимся первой парой и акцентируем внимание на том, как на мощности критериев отражается применяемый метод оценивания параметров закона.

Логистическое распределение с плотностью

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

– это распределение, очень близкое к нормальному, а при значениях параметров $\theta_0=0$, $\theta_1=\sqrt{3}/\pi\approx 0.551328895$ – близкое к стандартному нормальному закону.

В табл. 3 представлены полученные в рамках программной системы [19] оценки мощности критериев согласия χ^2 Пирсона, Колмогорова (K), Крамера – Мизеса – Смирнова (CMS) и Андерсона – Дарлинга (AD) при проверке сложной гипотезы о принадлежности выборок нормальному закону при конкурирующей гипотезе в виде логистического распределения (в зависимости от используемого метода оценивания параметров).

В случае критерия χ^2 Пирсона использовалось асимптотически оптимальное группирование [35, 36] и выборки разбивались на $k = 10$ интервалов.

Как можно видеть по приводимым в табл. 3 оценкам мощности, непараметрические критерии согласия более высокую мощность показывают в случае применения метода максимального правдоподобия.

Если рассматривать различные *MD*-оценки, то увидим, что мощность критериев Колмогорова и Крамера – Мизеса – Смирнова оказывается выше при *MD*-оценках $\breve{\theta}$, минимизирующих статистику Андерсона – Дарлинга. Возможно, это объясняется тем, что свойства этих оценок ближе к свойствам ОМП $\hat{\theta}$. Однако мощность критерия Андерсона – Дарлинга принимает

бо̀льшие значения при *MD*-оценках, получаемых при минимизации статистики Крамера – Мизеса – Смирнова.

Таблица 3

Table 3

Оценки мощности критериев при проверке сложных гипотез о принадлежности выборок нормальному закону при конкурирующем логистическом законе и заданном уровне значимости $\alpha = 0.1$

Estimates of the power of criteria when testing composite hypotheses about the belonging of samples to the normal law with a competing logistic law and a given significance level $\alpha = 0.1$

Критерий	$n = 20$	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 1000$
При ОМП $\hat{\theta}$						
Пирсона	0.151	0.133	0.233	0.340	0.639	0.906
K	0.143	0.181	0.240	0.354	0.651	0.906
CMS	0.152	0.209	0.294	0.449	0.783	0.967
A–D	0.163	0.231	0.328	0.498	0.830	0.981
При *MD*-оценке $\tilde{\theta}$						
Пирсона	0.166	0.230	0.312	0.448	0.739	0.944
K	0.108	0.113	0.125	0.146	0.235	0.417
CMS	0.108	0.117	0.133	0.167	0.277	0.490
A–D	0.175	0.216	0.274	0.381	0.666	0.908
При *MD*-оценке $\widehat{\theta}$						
Пирсона	0.173	0.243	0.330	0.474	0.761	0.968
K	0.105	0.113	0.123	0.145	0.231	0.401
CMS	0.100	0.108	0.123	0.151	0.247	0.435
A–D	0.185	0.232	0.297	0.412	0.707	0.931
При *MD*-оценке $\breve{\theta}$						
Пирсона	0.164	0.217	0.277	0.400	0.691	0.924
K	0.129	0.140	0.157	0.190	0.299	0.491
CMS	0.132	0.147	0.170	0.224	0.391	0.645
A–D	0.169	0.209	0.265	0.373	0.670	0.917

Что касается критерия χ^2 Пирсона, то следует заметить, что при *MD*-оценках мощность его не уменьшилась, а зависимость мощности от вида *MD*-оценок такая же, как в случае критерия Андерсона – Дарлинга.

В целом же, полученные результаты демонстрируют, как вид используемых методов оценивания параметров предполагаемых законов отражается на мощности применяемых критериев согласия.

ЗАКЛЮЧЕНИЕ

Таким образом, в рамках программной системы [19] продемонстрированы возможности компьютерных технологий моделирования и анализа данных при исследовании статистических свойств различных оценок параметров вероятностных моделей и влияния видов оценок на распределения статистик критериев, в которых эти оценки используются.

В частности, на примере нормального закона и его параметров исследованы свойства *MD*-оценок, минимизирующих статистики различных непараметрических критериев согласия. Показано, что *MD*-оценки, минимизирующие статистики различных непараметрических критериев согласия, в отличие от ОМП, не являются асимптотически эффективными оценками. При этом асимптотические распределения *MD*-оценок $\tilde{\theta}_0$, $\widehat{\theta}_0$, $\breve{\theta}_0$ различаются.

Показано, что при проверке сложных гипотез, когда параметры закона находятся по анализируемой выборке, распределения статистик всех непараметрических критериев согласия и критерия χ^2 Пирсона, имеющие место при справедливости проверяемой гипотезы H_0, зависят от метода оценивания параметров. При этом распределения статистики одного и того же критерия различаются при разных методах оценивания.

Программная система [19] позволяет проводить аналогичные исследования относительно любых параметров более чем 30 параметрических моделей законов распределения, включенных в систему, и относительно любых из более чем десятка критериев согласия.

Полученные оценки мощности критериев, знание асимптотических свойств оценок и характера влияния этих оценок на распределения статистик критериев позволяют судить о предпочтительности комбинаций применяемых критериев с методами оценивания.

Всё это в совокупности способствует повышению корректности выводов, формируемых по результатам статистического анализа.

СПИСОК ЛИТЕРАТУРЫ

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

2. *Geary R.C.* The ratio of the mean deviation to the standard deviation as a test of normality // Biometrika. – 1935. – Vol. 27. – P. 310–322.

3. *Hegazy Y.A.S.*, *Green J.R.* Some new goodness-of-fit tests using order statistics // Applied Statistics. – 1975. – Vol. 24 (3). – P. 299–308.

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

5. *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.

6. *Никулин М.С.* О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. – 1973. – Т. 18, № 3. – С. 675–676.

7. *Rao K.C., Robson D.S.* A chi-squared statistic for goodness-of-fit tests within the exponential family // Communications in Statistics. – 1974. – Vol. 3 (12). – P. 1139–1153.

8. *Vaart A.W. van der*. Asymptotic statistics. – Cambridge University Press, 1998. – DOI: 10.1017/CBO9780511802256.

9. *Greenwood P.E., Nikulin M.S.* A guide to chi-squared testing. – New York: John Wiley & Sons, 1996. – 280 p.

10. *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.

11. *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.

12. *Лемешко Б.Ю.* Непараметрические критерии согласия. Руководство по применению: монография. – 2-е изд., перераб. и доп. – М.: Инфра-М, 2024. – 201 с. – (Научная мысль). – DOI: 10.12737/2058731.

13. *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.

14. *Лемешко Б.Ю., Маклаков А.А.* Непараметрические критерии при проверке сложных гипотез о согласии с распределениями экспоненциального семейства // Автометрия. – 2004. – № 3. – С. 3–20.

15. *Лемешко Б.Ю., Гильдебрант С.Я., Постовалов С.Н.* К оцениванию параметров надежности по цензурированным выборкам // Заводская лаборатория. Диагностика материалов. – 2001. – Т. 67, № 1. – С. 52–64.

16. *Лемешко Б.Ю.* Об оценивании параметров распределений и проверке гипотез по цензурированным выборкам // Методы менеджмента качества. – 2001. – № 4. – С. 32–38.

17. *Rao C.R.* Criteria of estimation in large samples // Sankhyā: The Indian Journal of Statistics, Series A. – 1963. – Vol. 25 (2). – P. 189–206.

18. *Рао С.Р.* Линейные статистические методы и их применение. – М.: Наука, 1968. – 548 с.

19. 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: 25.02.2025).

20. *Лемешко Б.Ю., Постовалов С.Н.* К вопросу о робастности оценок по группированным данным // Сборник научных трудов НГТУ. – 1996. – № 2 (4). – С. 9–18.

21. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск: Изд-во НГТУ, 2011. – 888 с.

22. *Hampel F.R.* The influence curve and its role in robust estimation // Journal of the American Statistical Association. – 1974. – Vol. 69 (346). – P. 383–393. – DOI: 10.2307/2285666.

23. *Лемешко Б.Ю.* Робастные методы оценивания и отбраковка аномальных измерений // Заводская лаборатория. Диагностика материалов. – 1997. – Т. 63, № 5. – С. 43–49.

24. *Лемешко Б.Ю.* Группирование наблюдений как способ получения робастных оценок // Надежность и контроль качества. – 1997. – № 5. – С. 26–35.

25. *Лемешко Б.Ю., Чимитова Е.В.* Построение оптимальных L-оценок параметров сдвига и масштаба распределений по выборочным квантилям // Сибирский журнал индустриальной математики. – 2001. – Т. 4, № 2. – С. 166–183.

26. *Лемешко Б.Ю., Чимитова Е.В.* Оптимальные L-оценки параметров сдвига и масштаба распределений по выборочным квантилям // Заводская лаборатория. Диагностика материалов. – 2004. – Т. 70, № 1. – С. 54–66.

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

28. *Anderson T.W., Darling D.A.* A test of goodness of fit // Journal of the American Statistical Association. – 1954. – Vol. 49 (268). – P. 765–769. – DOI: 10.1080/01621459.1954.10501232.

29. *Anderson T.W., Darling D.A.* Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes // The Annals of Mathematical Statistics. – 1952. – Vol. 23 (2). – P. 193–212. – DOI: 10.1214/aoms/1177729437.

30. *Kuiper N.H.* Tests concerning random points on a circle // Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Proceedings. Series A. – 1960. – Vol. 63. – P. 38–47.

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

32. *Watson G.S.* Goodness-of-fit tests on a circle. II // Biometrika. – 1962. – Vol. 49 (1–2). – P. 57–63.

33. *Zhang J.* Powerful goodness-of-fit and multi-sample tests: PhD Thesis. – Toronto, York University, 2001. – 113 p. – URL: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (accessed: 28.01.2013).

34. *Лемешко Б.Ю., Чимитова Е.В.* Численное сравнение оценок максимального правдоподобия с одношаговыми и влияние точности оценивания на распределения статистик критериев согласия // Заводская лаборатория. Диагностика материалов. – 2003. – Т. 69, № 5. – С. 62–68.

35. *Лемешко Б.Ю.* Асимптотически оптимальное группирование наблюдений – это обеспечение максимальной мощности критериев // Надежность и контроль качества. – 1997. – № 8. – С. 3–14.

36. *Лемешко Б.Ю.* Асимптотически оптимальное группирование наблюдений в критериях согласия // Заводская лаборатория. Диагностика материалов. – 1998. – Т. 64, № 1. – С. 56–64.

37. *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.

38. *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.

39. *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.

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

41. *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.

42. *Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н.* Сравнительный анализ мощности критериев согласия при близких конкурирующих гипотезах. I. Проверка простых гипотез // Сибирский журнал индустриальной математики. – 2008. – Т. 11, № 2 (34). – С. 96–111.

43. *Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н.* Сравнительный анализ мощности критериев согласия при близких альтернативах. II. Проверка сложных гипотез // Сибирский журнал индустриальной математики. – 2008. – Т. 11, № 4 (36). – С. 78–93.

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

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

Lemeshko Boris Yu., D.Sc. (Eng,), Professor, Professor at the Department of Theoretical and Applied Informatics, Novosibirsk State Technical University. The main direction of his scientific research is computer technologies for studying 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 in the Center for Statistical Technology, Novosibirsk State Technical University. The main direction of his scientific research is computer modeling technologies and research of statistical laws. He has 69 publications, including 1 monograph. E-mail: skyer@mail.ru

DOI: 10.17212/2782-2001-2025-1-49-72

On the study of the properties of estimates and the influence of estimation methods on the distribution of statistics of hypothesis testing criteria [*]

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

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

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

Abstract

Statistical expressions of many criteria for of testing statistical hypotheses contain estimates of parameters or numerical characteristics such as mathematical expectation or dispersion, of assumed laws of probability distribution. Estimates can be found by different methods and differ in properties. The study of asymptotic properties of some estimate by analytical methods is, as a rule, a non-trivial task. It is even more difficult to study the properties of estimates by analytical methods with limited sample sizes, when the real properties of estimates may differ significantly from the asymptotic ones. Computer technologies make it possible to study the asymptotic properties of estimates, to clarify at what sample sizes one can rely on asymptotic properties, and to show how real properties differ from asymptotic ones at limited sample sizes.

As a consequence, the distributions of the statistics of the tests for the validity of the hypothesis being tested depend on the properties of the estimates used, that is, on the evaluation method used. The study of changed properties of tests due to the choice of another evaluation method using analytical methods also turns out to be a non-trivial task. At the same time, computer technologies make it possible not only to solve such problems, but also to ensure the correct application of relevant tests in changing conditions.

Within the framework of the developed software system, the capabilities of computer technologies for modeling and data analysis in studying the properties of various assessments and the influence of the type of assessment on the distribution of statistic of criteria are demonstrated. The properties of MD estimates that minimize the statistics of various nonparametric goodness-of-fit tests are investigated. It is shown that, unlike maximum likelihood estimators, these estimators are not asymptotically efficient and their asymptotic distributions differ. It is shown that when testing complex hypotheses, when the parameters of the law are found from the analyzed sample, the distributions of the statistics of all nonparametric goodness-of-fit tests and the Pearson test, which occur when the hypothesis being tested is valid, depend on the method of estimating the parameters. Moreover, the distributions of statistics of the same test differ for different evaluation methods. The obtained estimates of the power of the criteria, knowledge of the asymptotic properties of the estimates and the nature of the influence of these estimates on the distribution of the statistics of the test allow us to judge the preference of the combinations of the applied tests with the evaluation methods.

The software system allows conducting similar studies with respect to any parameters of more than 30 parametric models of distribution laws included in the system, and with respect to any of more than a dozen goodness-of-fit tests.

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

[*] *Received 30 December 2024.*

REFERENCES

1. Frosini B.V. A survey of a class of goodness-of-fit statistics. *Metron*, 1978, vol. 3 (1–2), pp. 3–49.

2. Geary R.C. The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika*, 1935, vol. 27, pp. 310–322.

3. Hegazy Y.A.S., Green J.R. Some new goodness-of-fit tests using order statistics. *Applied Statistics*, 1975, vol. 24 (3), pp. 299–308.

4. Lemeshko B.Yu. *Kriterii proverki otkloneniya raspredeleniya ot normal'nogo zakona. Rukovodstvo po primeneniyu* [Criteria for checking the deviation of the distribution from the normal law. Application Guide]. 2nd ed., rev. Moscow, Infra-M Publ., 2023. 353 p. DOI: 10.12737/1896110.

5. Nikulin M.S. Kriterii khi-kvadrat dlya nepreryvnykh raspredelenii s parametrami sdviga i masshtaba [Chi-square test for continuous distributions with shift and scale parameters]. *Teoriya veroyatnostei i ee primenenie = Probability Theory and its Applications*, 1973, vol. 18, no. 3, pp. 583–591.

6. Nikulin M.S. O kriterii khi-kvadrat dlya nepreryvnykh raspredelenii [On the chi-square test for continuous distributions]. *Teoriya veroyatnostei i ee primenenie = Probability Theory and its Applications*, 1973, vol. 18, no. 3, pp. 675–676.

7. *Rao K.C., Robson D.S.* A chi-squared statistic for goodness-of-fit tests within the exponential family. *Communications in Statistics*, 1974, vol. 3 (12), pp. 1139–1153.

8. Vaart A.W. van der. *Asymptotic statistics*. Cambridge University Press, 1998. DOI: 10.1017/CBO9780511802256.

9. Greenwood P.E., Nikulin M.S. *A guide to chi-squared testing*. New York, John Wiley & Sons, 1996. 280 p.

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

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

12. Lemeshko B.Yu. *Neparametricheskie kriterii soglasiya. Rukovodstvo po primeneniyu* [Nonparametric Goodness-of-Fit Tests. Guide on the application]. 2nd ed., rev. Moscow, Infra-M Publ., 2024. 201 p. DOI 10.12737/2058731.

13. Lemeshko B.Yu., Postovalov S.N. O zavisimosti raspredelenii statistik neparametricheskikh kriteriev i ikh moshchnosti ot metoda otsenivaniya parametrov [On the dependence of the distributions of statistics of nonparametric tests and their power on the method of parameter estimation]. *Zavodskaya laboratoriya. Diagnostika materialov = Industrial laboratory. Diagnostics of materials*, 2001, vol. 67, no. 7, pp. 62–71.

14. Lemeshko B.Yu., Maklakov A.A. Nonparametric test in testing composite hypotheses on goodness of fit exponential family distributions. *Optoelectronics, Instrumentation and Data Processin*g, 2004, vol. 40 (3), pp. 3–18. Translated from *Avtometriya*, 2004, no. 3, pp. 3–20.

15. Lemeshko B.Yu., Gil'debrant S.Ya., Postovalov S.N. K otsenivaniyu parametrov nadezhnosti po tsenzurirovannym vyborkam [On the estimation of reliability parameters from censored samples]. *Zavodskaya laboratoriya. Diagnostika materialov = Industrial laboratory. Diagnostics of materials*, 2001, vol. 67, no. 1, pp. 52–64.

16. Lemeshko B.Yu. Ob otsenivanii parametrov raspredelenii i proverke gipotez po tsenzurirovannym vyborkam [On estimating distribution parameters and testing hypotheses using censored samples]. *Metody menedzhmenta kachestva = Methods of Quality Management*, 2001, no. 4, pp. 32–38.

17. Rao C.R. Criteria of estimation in large samples. *Sankhyā: The Indian Journal of Statistics, Series A*, 1962, vol. 25 (2), pp. 189–206.

18. Rao C.R. *Lineinye statisticheskie metody i ikh primenenie* [Linear statistical inference and its applications]. Moscow, Nauka Publ., 1968. 548 p. (In Russian).

19. 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 25.02.2025).

20. Lemeshko B.Yu., Postovalov S.N. K voprosu o robastnosti otsenok po gruppirovannym dannym [On the issue of robustness of estimates based on grouped data]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta = Transaction of scientific papers of the Novosibirsk state technical university*, 1996, no. 2 (4), pp. 9–18.

21. Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N., Chimitova E.V *Statisticheskii analiz dannykh, modelirovanie i issledovanie veroyatnostnykh zakonomernostei. Komp'yuternyi podkhod* [Statistical data analysis, modeling and study of probabilistic patterns. Computer approach]. Novosibirsk, NSTU Publ., 2011. 888 p.

22. Hampel F.R. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 1974, vol. 69 (346), pp. 383–393. DOI: 10.2307/2285666.

23. Lemeshko B.Yu. Robust methods for the estimation and rejection of anomalous measurements. *Industrial Laboratory*, 1997, vol. 63 (5), pp. 297–302. Translated from *Zavodskaya laboratoriya. Diagnostika materialov*, 1997, vol. 63, no. 5, pp. 43–49.

24. Lemeshko B.Yu. Gruppirovanie nablyudenii kak sposob polucheniya robastnykh otsenok [Grouping observations as a way to obtain robust estimates]. *Nadezhnost' i kontrol' kachestva = Reliability and Quality Control*, 1997, no. 5, pp. 26–35.

25. Lemeshko B.Yu., Chimitova E.V. Postroenie optimal'nykh L-otsenok parametrov sdviga i masshtaba raspredelenii po vyborochnym kvantilyam [Construction of optimal L-estimates of the shift and scale parameters of distributions by sample quantiles]. *Sibirskii zhurnal industrial'noi matematiki = Siberian Journal of Industrial Mathematics*, 2001, vol. 4, no. 2, pp. 166–183. (In Russian).

26. Lemeshko B.Yu., Chimitova E.V. Optimal'nye L-otsenki parametrov sdviga i masshtaba raspredelenii po vyborochnym kvantilyam [Optimal L-estimates of the shift and scale parameters of distributions over sample quantiles]. *Zavodskaya laboratoriya. Diagnostika materialov = Industrial laboratory. Diagnostics of materials*, 2004, vol. 70, no. 1, pp. 54–66.

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

28 Anderson T.W., Darling D.A. A test of goodness of fit. *Journal of the American Statistical Association*, 1954, vol. 49 (268), pp. 765–769. DOI: 10.1080/01621459.1954.10501232.

29. Anderson T.W., Darling D.A. Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 1952, vol. 23 (2), pp. 193–212. DOI: 10.1214/aoms/1177729437.

30. Kuiper N.H. Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Proceedings. Series A*, 1960, vol. 63, pp. 38–47.

31. Watson G.S. Goodness-of-fit tests on a circle. *Biometrika*, 1961, vol. 48 (1–2), pp. 109–114.

32. Watson G.S. Goodness-of-fit tests on a circle. II. *Biometrika*, 1962, vol. 49 (1–2), pp. 57–63.

33. Zhang J. *Powerful goodness-of-fit and multi-sample tests*. PhD Thesis. Toronto, York University, 2001. 113 p. Available at: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (accessed 28.01.2013).

34. Lemeshko B.Yu., Chimitova E.V. Chislennoe sravnenie otsenok maksimal'nogo pravdopodobiya s odnoshagovymi i vliyanie tochnosti otsenivaniya na raspredeleniya statistik kriteriev soglasiya [Numerical comparison of maximum likelihood estimates with one-step ones and the influence of estimation accuracy on the distributions of goodness-of-fit statistics]. *Zavodskaya laboratoriya. Diagnostika materialov = Industrial laboratory. Diagnostics of materials*, 2003, vol. 69, no. 5, pp. 62–68.

35. Lemeshko B.Yu. Asimptoticheski optimal'noe gruppirovanie nablyudenii – eto obespechenie maksimal'noi moshchnosti kriteriev [Asymptotically optimal grouping of observations is to ensure the maximum power of the criteria]. *Nadezhnost' i kontrol' kachestva = Reliability and Quality Control*, 1997, no. 8, pp. 3–14.

36. Lemeshko B.Yu. Asymptotically optimum grouping of observations in goodness-of-fit tests. *Industrial laboratory*, 1998, vol. 64 (1), pp. 59–67. Translated from *Zavodskaya laboratoriya. Diagnostika materialov*, 1998, vol. 64, no. 1, pp. 56–64.

37. 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, vol. 44, pp. 40–49. DOI: 10.17223/19988605/44/5.

38. 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 (3),

pp. 242–250. DOI: 10.3103/S8756699020030103. Translated from *Avtometriya*, 2020, vol. 56, no. 3, pp. 35–45. DOI: 10.15372/AUT20200305.

39. 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, vol. 53, pp. 47–60. DOI: 10.17223/19988605/53/5.

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

41. 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.

42. Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N. Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses. I. The verification of simple hypotheses. *Journal of Applied and Industrial Mathematics*, 2009, vol. 3 (4), pp. 462–475. DOI: 10.1134/S199047890904005X. Translated from *Sibirskii zhurnal industrial'noi matematiki*, 2008, vol. 11, no. 2 (34), pp. 96–111.

43. Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N. Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses. II. Verification of complex hypotheses. *Journal of Applied and Industrial Mathematics*, 2010, vol. 4 (1), pp. 79–93. DOI: 10.1134/S1990478910010114. Translated from *Sibirskii zhurnal industrial'noi matematiki*, 2008, vol. 11, no. 4 (36), pp. 78–93.

Для цитирования:

Лемешко Б.Ю., Лемешко С.Б. Об исследовании свойств оценок и влияния методов оценивания на распределения статистик критериев проверки гипотез // Системы анализа и обработки данных. – 2025. – № 1 (97). – С. 49–72. – DOI: 10.17212/2782-2001-2025-1-49-72.

For citation:

Lemeshko B.Yu., Lemeshko S.B. Ob issledovanii svoistv otsenok i vliyaniya metodov otsenivaniya na raspredeleniya statistik kriteriev proverki gipotez [On the study of the properties of estimates and the influence of estimation methods on the distribution of statistics of hypothesis testing criteria]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2025, no. 1 (97), pp. 49–72. DOI: 10.17212/2782-2001-2025-1-49-72.