519.24

Смещённость непараметрических критериев согласия относительно некоторых пар конкурирующих гипотез

Б. Ю. ЛЕМЕШКО, П. Ю. БЛИНОВ, С. Б. ЛЕМЕШКО

Новосибирский государственный технический университет, Новосибирск, Россия, e-mail: Lemeshko@ami.nstu.ru

Рассмотрено применение непараметрических критериев согласия Андерсона—Дарлинга, Крамера—Мизеса—Смирнова, Купера, Ватсона, Колмогорова, Жанга при проверке простых и сложных гипотез. На основании исследования мощности впервые показано, что существуют пары конкурирующих гипотез, которые указанные критерии не могут различать при малых объёмах выборок и малых вероятностях ошибок 1-го рода. Показано, что причина этого кроется в смещённости критериев в соответствующих ситуациях.

Ключевые слова: *критерии Андерсона—Дарлинга, Крамера—Мизеса—Смирнова, Купера, Ватсона, Колмогорова, Жанга, мощность критерия.*

Problems of the application of nonparametric goodness-of-fit tests (Anderson—Darling test, Cramer—Mises—Smirnov test, Kuiper test, Watson test, Kolmogorov test, Zhang tests) for testing simple and composite hypothesis are considered. Results based on studies of test power demonstrate that there are pairs of competing hypotheses, which considered tests can't distinguish for small sample size and small values of type I error. It is shown that the reason of this fact is bias of tests in these cases. It is indicated that is more preferable to use tests with some advantages for greater objectivity of statistical inference.

Key words: *Anderson—Darling test, Cramer—Mises—Smirnov test, Kuiper test, Watson test, Kolmogorov test, Zhang tests, test power.*

Используя при анализе результатов экспериментов критерии проверки статистических гипотез, исследователь порой слепо верит результатам вывода, не задумываясь о том, что критерий представляет собой математический инструмент, предназначенный для измерения (обнаружения, оценки?) в анализируемых данных некоторого отклонения от выдвигаемых предположений, и что этот инструмент также требует соответствующей «настройки и поверки». Множество статистических «инструментов», применение которых правомерно для измерения соответствующих величин, достаточно обширно, а сами инструменты принадлежат к различным «классам точности». Квалифицированное применение любого измерительного инструмента предполагает знание его возможностей и области применения.

В данном случае речь пойдет о некоторых фактах, свидетельствующих об ограниченных возможностях непараметрических критериев согласия (Колмогорова, Крамера—Мизеса—Смирнова, Андерсона—Дарлинга, Купера, Ватсона, Жанга) по различению законов распределения вероятностей. О существовании подобных фактов в связи с непараметрическими критериями согласия в литературных источниках ранее не сообщалось. Знание же этих фактов должно способствовать более осмысленному восприятию результатов применения критериев проверки статистических гипотез.

Напомним, что при использовании критериев согласия необходимо различать проверку простых и сложных гипотез. Простая проверяемая гипотеза имеет вид H_0: $F(x) = F(x, \theta)$, где $F(x, \theta)$ — функция распределения вероятностей, с которой проверяют согласие наблюдаемой выборки, а θ — известное значение параметра (скалярного или векторного).

Сложная проверяемая гипотеза имеет вид H_0: $F(x) \in \{F(x, \theta), \theta \in \Theta\}$, где Θ — область определения параметра θ. Отличие возникает, если при проверке сложной гипотезы оценку $\hat{\theta}$ параметра распределения вычисляют по той же самой выборке, по которой проверяют согласие, так как тогда распределение статистики, соответствующее справедливости гипотезы H_0, значительно отличается от существующего при простой проверяемой гипотезе [1].

При проверке гипотезы H_0 о принадлежности выборки закону с функцией распределения $F(x, \theta)$ критерий Колмогорова [2] опирается на статистику

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

где $F_n(x)$ — эмпирическая функция распределения; $F(x, \hat{\theta})$ — функция распределения нормального закона; $\hat{\theta}$ — оценка вектора параметров, найденная по этой же выборке; n — объем выборки.

При проверке гипотез с применением критерия Колмогорова рекомендуется использовать статистику с поправкой Большева в форме [2]:

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

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

$D_n^- = \max\limits_{1 \le i \le n} \left\{ F\left(x_{(i)}, \hat{\theta}\right) - \frac{i-1}{n} \right\}$; $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$ — здесь и далее, элементы вариационного ряда, построенного по исходной выборке $X_1, X_2, \ldots, X_n$.

При проверке простой гипотезы H_0 статистика подчиняется распределению Колмогорова $K(s)$ [2].

В критерии Купера [3] в качестве меры расстояния между эмпирическим и теоретическим законом рассматривается величина

$$V_n = D_n^+ + D_n^-,$$

где D_n^+, D_n^- определены выше.

В качестве статистик критерия можно использовать статистику [4]:

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

или статистику [5]:

$$V_n^{\text{mod}} = \sqrt{n}\,V_n + \left(3\sqrt{n}\right)^{-1}.$$

Предельное распределение этих статистик при справедливости простой гипотезы H_0 получено в [3] и приведено в [1].

Статистика критерия ω^2 Крамера—Мизеса—Смирнова имеет вид

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

При простой проверяемой гипотезе, когда параметры теоретического закона $F(x, \theta)$ известны, при справедливости H_0 данная статистика в пределе подчиняется закону с функцией распределения $a1(s)$ [2].

Статистика критерия Ватсона [6, 7] используется в следующей удобной для расчётов форме:

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

Предельное распределение этой статистики при справедливости простой проверяемой гипотезы получено в [6, 7].

В критерии согласия Ω^2 Андерсона—Дарлинга [8, 9] используется статистика

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

При справедливости простой проверяемой гипотезы H_0, когда параметры теоретического закона $F(x, \theta)$ известны, эта статистика в пределе подчиняется закону с функцией распределения $a2(s)$ [2].

В [10, 11] были предложены непараметрические критерии согласия, статистики которых имеют вид:

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\ln\left\{F(x_{(i)}, \hat{\theta})\right\}}{n-i+1/2} + \frac{\ln\left\{1 - F(x_{(i)}, \hat{\theta})\right\}}{i-1/2}\right];$$

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

$$Z_K = \max_{1\le i\le n}\left(\left(i - \frac{1}{2}\right)\ln\left\{\frac{i-1/2}{nF\left(x_i, \hat{\theta}\right)}\right\} + \right.$$

$$\left. + \left(n - i + \frac{1}{2}\right)\ln\left[\frac{n-i+1/2}{n\left\{1 - F\left(x_{(i)}, \hat{\theta}\right)\right\}}\right]\right).$$

Применение критериев с этими статистиками осложняет сильная зависимость распределений статистик от объёма выборки n.

При проверке сложных гипотез, когда по той же выборке оценивают параметры наблюдаемого закона распределения вероятностей, все рассматриваемые непараметрические критерии согласия теряют свойство «свободы от распределения» [12]. Более того, предельные распределения статистик непараметрических критериев согласия зависят от ряда факторов, определяющих «сложность» гипотезы.

На закон распределения статистики $G(S|H_0)$ влияют следующие факторы [1]:

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

тип оцениваемого параметра и число оцениваемых параметров;

в некоторых ситуациях конкретное значение параметра, например, в случае гамма- и бета-распределений;

используемый метод оценивания параметров.

Вопросам проверки сложных гипотез по отношению к различным законам распределения и построению моделей предельных распределений статистик перечисленных выше критериев посвящены работы [13—18]. Рекомендации по применению критериев при проверке простых и сложных гипотез сконцентрированы в руководстве по применению [1].

Используя критерий согласия, экспериментатор предполагает, что применяемый критерий представляет собой инструмент, (в принципе) позволяющий ему отличать закон $F(x, \theta)$, соответствующий проверяемой гипотезе H_0, от (близких) конкурирующих законов распределения. Напомним, что с проверкой статистических гипотез связаны ошибки двух типов. Ошибка 1-го рода заключается в отклонении справедливой проверяемой гипотезы H_0, и её вероятность обозначают α. Ошибка 2-го рода состоит в неотклонении в случае справедливости некоторой конкурирующей гипотезы H_1, а её вероятность обозначают β. Предположение экспериментатора о желаемых свойствах критерия включает требование его несмещённости, т. е. при любой заданной вероятности ошибки 1-го рода α и любой конкурирующей гипотезе H_1 (любом конкурирующем законе) мощность критерия $1 - \beta$ относительно H_1 должна подчиняться неравенству $\alpha \le 1 - \beta$.

В литературных источниках не встречается упоминаний о примерах смещённости непараметрических критериев согласия относительно каких-либо конкурирующих гипотез. По-видимому, большинство применяемых на практике критериев являются асимптотически несмещёнными, а «огрехи» проявляются при ограниченных объёмах выборок и относительно близких конкурирующих гипотезах.

В [19, 20] при исследовании свойств множества критериев, ориентированных на проверку нормальности, была обнаружена смещённость (при ограниченных объёмах выборок n и малых α) целого ряда специальных критериев (Шапиро—Уилка, Эппса—Палли, Хегази—Грина, Шпигельхалтера, Ройстона). При этом в качестве конкурирующей гипотезы H_1 рассматривался обобщённый нормальный закон с плотностью

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

и параметром формы $\theta_2 = 4$. Соответствующие плотности нормального и обобщённого нормального законов пред-

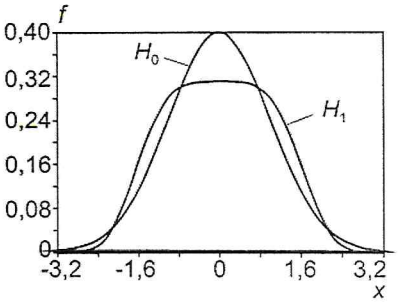

Рис. 1. Плотности, соответствующие *1* — нормальному (H_0) и *2* — обобщённому нормальному (H_1) законам

ставлены на рис. 1. Конкурирующая гипотеза H_1, которой соответствует обобщённый нормальный закон с параметром формы $\theta_2 = 4$, представляет собой «лакмусовую бумагу», на которой проявились ранее скрытые недостатки множества критериев.

Непараметрические критерии согласия (Колмогорова, Купера, Крамера—Мизеса—Смирнова, Ватсона и Андерсона—Дарлинга) относительно данной пары гипотез продемонстрировали несмещённость [21]. Однако критерии Жанга со статистиками Z_C и Z_A при проверке нормальности относительно этой же конкурирующей гипотезы показали существенную смещённость [21]. Оценки мощности критерия Жанга со статистикой Z_A по отношению к конкурирующей гипотезе H_1 приведены в табл. 1, а со статистикой Z_C — в табл. 2.

Таблица 1

Мощность критерия z_A Жанга относительно гипотезы H_1

n	α				
	0,15	0,1	0,05	0,025	0,01
10	0,127	0,078	0,033	0,014	0,005
20	0,148	0,090	0,036	0,014	0,004
30	0,199	0,128	0,056	0,023	0,006
40	0,263	0,180	0,087	0,039	0,012
50	0,333	0,239	0,127	0,063	0,022
100	0,650	0,548	0,389	0,259	0,139
150	0,844	0,775	0,641	0,503	0,335
200	0,939	0,901	0,815	0,706	0,545
300	0,992	0,985	0,962	0,923	0,841

Таблица 2

Мощность критерия z_C Жанга относительно гипотезы H_1

n	α				
	0,15	0,1	0,05	0,025	0,01
10	0,163	0,101	0,041	0,017	0,004
20	0,211	0,130	0,049	0,014	0,002
30	0,277	0,179	0,071	0,020	0,002
40	0,348	0,238	0,104	0,033	0,003
50	0,421	0,300	0,142	0,049	0,005
100	0,715	0,599	0,390	0,201	0,045
150	0,879	0,806	0,635	0,420	0,150
200	0,955	0,917	0,808	0,634	0,322
300	0,995	0,988	0,961	0,895	0,688

Серым цветом в табл. 1, 2 выделены оценки мощности, меньшие α. Образно говоря, «с позиции критерия» это означает, что закон, соответствующий H_1, является «более нормальным, чем нормальный». Таким образом критерий со статистикой Z_A при n = 10, 20 не способен отличать закон, соответствующий гипотезе H_1, от нормального.

Заметим, что в общем случае критерии Андерсона—Дарлинга, Ватсона и Крамера—Мизес—Смирнова (да и критерии Жанга со статистиками Z_C и Z_A) не так уж сильно (и не всегда) уступают в мощности специальным критериям нормальности.

Следует отметить, что описанная ситуация со смещённостью критериев Жанга со статистиками Z_C и Z_A при проверке нормальности не является единственным примером с таким недостатком, касающимся применения непараметрических критериев согласия. Другой пример касается критериев проверки (простых и сложных) гипотез о принадлежности анализируемых выборок равномерному закону и использования в этих целях, в частности, непараметрических критериев согласия.

В данном случае простой проверяемой гипотезе H_0 соответствует равномерный закон на интервале [0,1], а конкурирующей гипотезе H_1 — принадлежность наблюдаемой случайной величины бета-распределению 1-го рода с функцией плотности

$$f(x) = \frac{1}{\theta_2 \mathrm{B}(\theta_0, \theta_1)} \left(\frac{x - \theta_3}{\theta_2} \right)^{\theta_0 - 1} \left(1 - \frac{x - \theta_3}{\theta_2} \right)^{\theta_1 - 1},$$

где $\mathrm{B}(\theta_2, \theta_1) = \Gamma(\theta_0)\,\Gamma(\theta_1)/\Gamma(\theta_0 + \theta_1)$ — бета-функция, со значениями параметров $\theta_0 = 1{,}5$, $\theta_1 = 1{,}5$, $\theta_2 = 1$, $\theta_3 = 0$.

Функции распределения вероятностей, соответствующие рассматриваемым гипотезам, пересекаются, а плотности распределений, приведенные на рис. 2, иллюстрируют существенное различие конкурирующих законов.

Исследование распределений статистик и анализ мощности критериев относительно H_1, проведённые в [22], позволили выявить, что ряд специальных критериев, ориентированных на проверку гипотезы о равномерности (Морана, Шермана, Гринвуда, Янга, Хегази—Грина и др.), при малых объёмах выборок n и малых уровнях значимости α не способны отличать эту гипотезу от H_0.

Оказалось, что большинство рассматриваемых в [22] непараметрических критериев согласия (за исключением критериев Купера и Ватсона) также в значительной степени имеют этот недостаток, причина которого заключается в смещённости соответствующих критериев.

В качестве иллюстрации такой ситуации на конкретном примере в табл. 3 для критерия Андерсона—Дарлинга, используемого для проверки равномерности, приведены оценки мощности относительно гипотезы H_1. Серым цветом выделены ячейки таблицы с оценками мощности, меньшими соответствующего значения α. На рис. 3 показаны распределение $G(S|H_0)$ статистики этого критерия при справедливости проверяемой гипотезы H_0 и распределения $G(S_n|H_1)$ его статистики при справедливости H_1 (для объёмов выборок n = 10, 20, 100, 300). Распределения статистики $G(S_n|H_1)$ при n = 10, 20 пересекают $G(S|H_0)$, что объясняет, почему мощность 1 – β оказывается меньше заданного значения α. На рис. 3 распределение $G(S|H_0)$ показано только при n = 10.

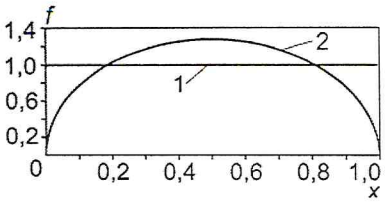

Рис. 2. Плотности, соответствующие *1* — равномерному (H_0) закону и *2* — бета-распределению 1-го рода (H_1)

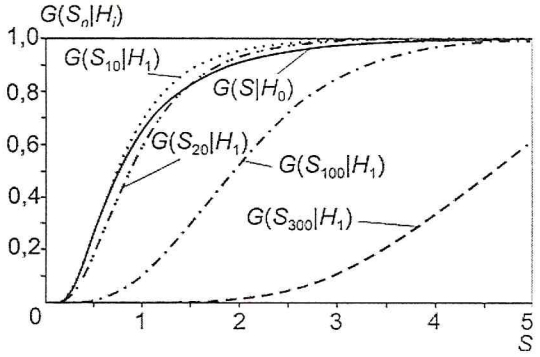

Рис. 3. Распределения $G(S|H_0)$ и $G(S_n|H_1)$ статистики критерия Андерсона—Дарлинга при проверке равномерности

При $n \geq 20$ распределения $G(S_n|H_0)$ визуально не отличаются от $G(S_{10}|H_0)$ и практически совпадают с предельным распределением статистики критерия Андерсона—Дарлинга при проверке простых гипотез $a2(s)$.

Таблица 3

Мощность критерия Андерсона—Дарлинга относительно гипотезы

n	α				
	0,15	0,1	0,05	0,025	0,01
10	0,095	0,053	0,019	0,007	0,002
20	0,140	0,078	0,028	0,010	0,003
30	0,196	0,114	0,042	0,014	0,004
40	0,258	0,156	0,060	0,021	0,005
50	0,325	0,206	0,084	0,031	0,007
100	0,652	0,505	0,283	0,134	0,041
150	0,861	0,760	0,544	0,332	0,138
200	0,954	0,904	0,762	0,565	0,311
300	0,998	0,990	0,959	0,882	0,702

Рассмотренные примеры демонстрируют смещённость непараметрических критериев согласия при малых объёмах выборок n и малых уровнях значимости α относительно некоторых пар конкурирующих гипотез. По-видимому, это первое упоминание о существовании такого рода недостатков непараметрических критериев согласия. Отсюда следует, что корректного использования какого-то одного из критериев для формирования «надёжного» статистического вывода часто может оказаться недостаточным. Использование совокупности критериев, опирающихся на различные меры отклонения эмпирического распределения от теоретического, повышает качество статистических выводов.

Знание реальных свойств критериев, их недостатков, отмеченных в данной работе, и следование рекомендациям [1, 21, 22], обусловливающим корректность применения критериев, позволит специалистам, решающим задачи статистического анализа при обработке результатов измерений в конкретной прикладной области, более осознано подходить к выбору критериев, не останавливаясь на использовании какого-то одного.

В заключение, хотелось бы обратить внимание, что, на наш взгляд, использование методов статистического анализа в задачах метрологии и стандартизации заметно отстаёт от уровня развития современной прикладной математической статистики. Это замечание касается как спектра применяемых методов, так и корректности их использования, в частности, описаний применения статистических критериев в регламентирующих документах. Например, желательно сделать поправки в стандарте [23], где в описании критериев, рекомендуемых для проверки нормальности, содержатся существенные неточности. Не в последнюю очередь причины такого состояния кроются и в том, что устаревшие сведения и методы тиражируются в учебных пособиях по направлению «Стандартизация и метрология» [24].

Работа выполнена при поддержке Минобразования России в рамках проектной части государственного задания (№ 2.541.2014/K).

Литература

1. **Лемешко Б. Ю.** Непараметрические критерии согласия: Руководство по применению / М.: НИЦ ИНФРА-М, 2014. DOI: 10.12737/11873.

2. **Большев Л. Н., Смирнов Н. В.** Таблицы математической статистики. М.: Наука, 1983.

3. **Kuiper N. H.** Tests concerning random points on a circle // Proc. Konikl. Nederl. Akad. Van Wettenschappen. 1960. Ser. A. V. 63. P. 38—47.

4. **Stephens M. A.** EDF statistics for goodness of fit and some comparisons // J. Am. Statist. Assoc. 1974. V. 69. N. 347. P. 730—737.

5. **Лемешко Б. Ю., Горбунова А. А.** О применении и мощности непараметрических критериев согласия Купера, Ватсона и Жанга // Измерительная техника. 2013. № 5. С. 3—9.

6. **Watson G. S.** Goodness-of-fit tests on a circle. I // Biometrika. 1961. V. 48. N. 1—2. P. 109—114.

7. **Watson G. S.** Goodness-of-fit tests on a circle. II // Biometrika. 1962. V. 49. N. 1—2. P. 57—63.

8. **Anderson T. W., Darling D. A.** Asymptotic theory of certain «Goodness of fit» criteria based on stochastic processes // Ann. Math. Statist. 1952. V. 23. P. 193—212.

9. **Anderson T. W., Darling D. A.** A test of goodness of fit // J. Amer. Statist. Assoc. 1954. V. 29. P. 765—769.

10. **Zhang J.** Powerful goodness-of-fit and multi-sample tests. Toronto: PhD Thesis. York University, 2001.

11. **Zhang J.** Powerful goodness-of-fit tests based on the likelihood ratio // J. Roy. Statist. Soc.: Series B. 2002. V. 64. № 2. P. 281—294.

12. **Kac M. Kiefer J., Wolfowitz J.** On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Statist. 1955. V. 26. P. 189—211.

13. **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. V. 39. № 3. P. 460—471.

14. **Lemeshko B. Yu., Lemeshko S. B.** Models of Statistic Distributions of Nonparametric Goodness-of-Fit Tests in Composite Hypotheses Testing for Double Exponential Law Cases // Communications in Statistics — Theory and Methods. 2011. V. 40. N. 16. P. 2879—2892.

15. **Лемешко Б. Ю., Лемешко С. Б.** Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. I // Измерительная техника. 2009. № 6. С. 3—11.

16. **Лемешко Б. Ю., Лемешко С. Б.** Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. II // Измерительная техника. 2009. № 8. С. 17—26.

17. **Lemeshko B. Yu., Gorbunova A. A., Lemeshko S. B., Rogozhnikov A. P.** Solving problems of using some nonparametric goodness-of-fit tests // Optoelectronics, Instrumentation and Data Processing. 2014. V. 50. № 1. P. 21—35.

18. **Лемешко Б. Ю., Горбунова А. А.** Применение непараметрических критериев согласия Купера и Ватсона при проверке сложных гипотез // Измерительная техника. 2013. № 9. С. 14—21.

19. **Лемешко Б. Ю., Лемешко С. Б.** Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. № 2. С. 3—24.

20. **Лемешко Б. Ю., Рогожников А. П.** Исследование особенностей и мощности некоторых критериев нормальности // Метрология. 2009. № 4. С. 3—24.

21. **Лемешко Б. Ю.** Критерии проверки отклонения распределения от нормального закона: Руководство по применению. М.: ИНФРА-М, 2015. DOI: 10.12737/6086.

22. **Лемешко Б. Ю., Блинов П. Ю.** Критерии проверки отклонения распределения от равномерного закона: Руководство по применению. М.: ИНФРА-М, 2015. DOI: 10.12737/11304.

23. **ГОСТ Р 8.736—2011.** Измерения прямые многократные. Методы обработки результатов измерений. Основные положения.

24. **Анцыферов С. С., Афанасьев М. С., Русанов К. Е.** Обработка результатов измерений: Учебное пособие. М.: Изд-во ИКАР, 2014.

Дата принятия 27.07.2016 г.

519.688

Области эффективного применения статистических методов обработки результатов многократных измерений

Г. Н. СОЛОПЧЕНКО

Санкт-Петербургский политехнический университет Петра Великого, С.-Петербург, Россия, e-mail: g.n.solopchenko@mail.ru

Показано различие областей эффективного применения статистической обработки результатов многократных измерений неизменной измеряемой величины: при реальных измерениях и при определении и контроле характеристик погрешности средств измерений в процессе их метрологических испытаний. Сформулированы условия, при которых можно удовлетворительно обеспечить независимость получаемых характеристик погрешности от вида плотности распределения случайной погрешности.

Ключевые слова: *поверка, калибровка, доверительный интервал, неопределенность, толерантные пределы.*

The difference is clarified between two situations in statistical treatment of measurement results in static mode: for multiple measuring of permanent measurand and for determination and control of error characteristics of measuring instrument during its metrological testing. Recommendations are formulated on the quantity of multiple measurements that provides the independence of obtained error characteristics from the density of random error distribution.

Key words: *metrological verification, calibration test, error characteristics, confidence interval, uncertainty, tolerant bounds.*

В настоящей статье представлены две области эффективного применения статистической обработки результатов многократных измерений неизменных во времени измеряемых величин: реальные измерения и метрологические испытания (поверка и (или) калибровка) средств измерений (СИ). Многократные реальные измерения обычно проводят, когда результаты измерений содержат случайные погрешности. Цель таких измерений — усреднение результатов и снижение погрешности конечного результата. При метрологических испытаниях СИ многократные измерения величины, неизменное значение которой обеспечивается мерой с высокой точностью, выполняют в нормальных условиях, чтобы определить подлежащие нормированию характеристики инструментальной погрешности или проконтролировать погрешность испытываемого СИ на предмет удовлетворения установленным нормам. Ниже показано различие методов статистической обработки результатов измерений в указанных областях и сформулированы условия, при которых эти методы не зависят от вида плотности распределения случайных погрешностей (distribution-free methods).

В случае, когда выполняются многократные измерения величин на природном или рукотворном объекте и их ре-