

## Читайте и узнаете:

- о том, какой критерий для оценки работы лаборатории в новом стандарте предлагается использовать в качестве одного из основных;
- причины, по которым в новом стандарте большое внимание уделяется анализу графического представления результатов участников МСИ;
- как стандарт помогает предотвратить ошибочные выводы провайдера МСИ о неудовлетворительном участии лаборатории.

## Ключевые слова:

ГОСТ Р 50779.60-2017, проверки квалификации, критерии оценки, термины, обозначения, представление результатов

# Новая версия *ISO 13528* — шаг к надежным методам оценки качества работы лабораторий

**Е.Н. Тупицын**

начальник отдела ОАО «Лига», г. Саратов, канд. физ.-мат. наук

Участие испытательной лаборатории в проверках квалификации посредством межлабораторных сличений является одним из требований ГОСТ ИСО/МЭК 17025–2009<sup>1</sup>, международно-го стандарта *ISO/IEC 17025:2017*<sup>2</sup>, а также важным подтверждением ее компетентности, объективным свидетельством возможности получения в лаборатории достоверных результатов испытаний.

В декабре 2018 г. введен в действие стандарт ГОСТ Р 50779.60–2017 (ИСО 13528:2015)<sup>3</sup>. На сегодняшний день *ISO 13528* является основным международ-

Недавнее введение в действие ГОСТ Р 50779.60–2017 (ИСО 13528:2015) «Статистические методы. Применение при проверке квалификации посредством межлабораторных испытаний», отражающего подходы и требования соответствующего международного стандарта, чрезвычайно важно для корректной оценки деятельности испытательной лаборатории по итогам участия в межлабораторных сличениях. Представлено описание значимых отличий нового документа от ГОСТ Р ИСО 13528-2010, взамен которого он был принят.

ным сводом статистических методов, которые применяются для обработки данных, полученных от лабораторий-участников провайдером проверки квалификации, выполняющим требования стандарта ГОСТ *ISO/IEC 17043-2013*<sup>4</sup>. Применение нового стандарта крайне важно для корректной оценки испытательной лаборатории по итогам участия в межлабораторных сличениях.

Из названия нового стандарта исчез термин «проверка компе-

тентности», являющийся синонимом «проверки квалификации» и ранее вносивший путаницу. Это положительный момент, так как тем самым устанавливается единообразие в понятиях. Одновременно можно посетовать на то, что термин «межлабораторные сличительные испытания» остался без изменений, потому что, если следовать единообразию, его следовало бы обозначить как «межлабораторные сличения». Объем стандарта увеличился более чем в полтора раза и пополнился новыми разделами. Даны рекомендации по решению ряда проблем и задач, ответы на которые в прежней редакции стандарта отсутствовали.

<sup>1</sup> ГОСТ ИСО/МЭК 17025–2009 «Общие требования к компетентности испытательных и калибровочных лабораторий» Приказом Федерального агентства по техническому регулированию и метрологии № 41-ст от 04.04.2011 г. введен в действие в качестве национального стандарта с 01.01.2012 г.

<sup>2</sup> *ISO/IEC 17025:2017. General requirements for the competence of testing and calibration Laboratories.*

<sup>3</sup> ГОСТ Р 50779.60–2017 (ИСО 13528:2015) «Статистические методы. Применение при проверке квалификации посредством межлабораторных испытаний» введен в действие Приказом Росстандарта № 1061-ст от 12.09.2017 г.

<sup>4</sup> ГОСТ *ISO/IEC 17043-2013* «Оценка соответствия. Основные требования к проведению проверки квалификации» Приказом Росстандарта № 1941-ст от 22.11.2013 г. введен в действие в Российской Федерации для применения в качестве национального стандарта с 01.03.2015 г.

Начнем с обзора основных изменений, которые коснулись новых обозначений характеристик, используемых при проверке квалификации (см. таблицу).

При проверке квалификации посредством межлабораторных сличений основным подходом является сопоставление результатов измерений, выполненных на некотором образце  $x_i$ , с приписанным значением  $x_{pt}$ . Для оценки работы лаборатории разность между  $x_i$  и  $x_{pt}$  сравнивают с допустимой ошибкой (погрешностью) измерений и делят на допустимое стандартное отклонение  $\sigma_{pt}$ , получая безразмерный стандартизованный индекс  $Z, Z', \xi, E_n$ . Далее устанавливают границы зон сигнала предупреждения, действия и сообщают участникам соответствующие им результаты индексов или оценки отклонения. Результат считают приемлемым, если значение индекса  $|Z| \leq 2$ , сомнительным (в зоне предупреждения), если  $2 < |Z| < 3$ , и неприемлемым, если  $|Z| \geq 3$ . В целом подход остался прежним за исключением того, что теперь провайдер в некоторых случаях может устанавливать границу зоны действия, равную двум, то есть отнести результат участника к неудовлетворительным при  $|Z| \geq 2$ .

### Максимально допустимая погрешность

В новом стандарте в качестве одного из основных критериев оценки работы лаборатории предлагается использовать максимально допустимую погрешность измерений (*Error*). Применение данного критерия представляется предпочтительнее остальных, присутствующих в стандарте, так

как в Российской Федерации действуют нормативные документы, устанавливающие точность измерений различных объектов анализа, например, ГОСТ 27384–2002<sup>5</sup>,

### Обозначения в ГОСТ Р ИСО 13528–2010 и ГОСТ Р 50779.60–2017

[таблица]

ГОСТ Р ИСО 13528–2010	ГОСТ Р 50779.60–2017
Приписанное значение $x$	Приписанное значение $x_{pt}$
Стандартное отклонение оценки компетентности $\hat{\sigma}$	Стандартное отклонение для оценки квалификации $\sigma_{pt}$

ГОСТ 17.2.4.02–81<sup>6</sup>, ГОСТ 28473–90<sup>7</sup>, ГОСТ Р 54569–2011<sup>8</sup>. Поэтому вопрос о критериях оценки при использовании участниками различных методик измерений решается автоматически. Кроме того, многие методики измерений в РФ содержат показатель качества — характеристику погрешности, которая, будучи представлена интервалом, определяет границы точности измерений (как правило, для доверительной вероятности  $P = 95\%$ ) и является по сути той же максимально допустимой погрешностью измерений. Поэтому, если в раунде межлабораторных сличений все участники используют одинаковую методику измерений, вопрос

теряется вводится новая характеристика, связанная с максимально допустимой погрешностью, — оценка отклонения  $D_i = x_i - x_{pt}$ , которую выражают или в абсолютных единицах, или в относительных —  $D_i\%$  (если разделить  $D_i$  на величину  $x_{pt}$ ).

### Обнаружение выбросов

Одной из основных задач при обработке провайдером результатов лабораторий-участников является обнаружение выбросов — некорректных, ошибочных значений. Иногда их идентификация происходит по очевидным причинам: представление результатов в других единицах измерений, или ошибка на несколько порядков из-за неправильного вычисления по формуле. Но зачастую определить выбросы в выборке бывает непросто, особенно когда они не сильно отличаются от основной массы результатов. В прежней редакции стандарта присутствовал всего один абзац, посвященный этому процессу и тема далее не развивалась. Вопросы возникали в основном, когда обнаруживались квазивыбросы (потенциальные выбросы с уровнем значимости от 0,05 до 0,01). Если автоматически отнести квазивыбросы к выбросам, то мож-

<sup>5</sup> ГОСТ 27384–2002 «Вода. Нормы погрешности измерений показателей состава и свойств» Постановлением Росстандарта № 408-ст от 12.11.2002 г. введен в действие непосредственно в качестве государственного стандарта с 01.01.2004 г.

<sup>6</sup> ГОСТ 17.2.4.02–81 «Охрана природы (ССОП). Атмосфера. Общие требования к методам определения загрязняющих веществ» Постановлением Госстандарта СССР № 4837 от 09.11.1981 г. дата введения установлена 01.07.1982 г.

<sup>7</sup> ГОСТ 28473–90 «Чугун, сталь, ферросплавы, хром, марганец металлические. Общие требования к методам анализа» введен в действие Постановлением Госстандарта СССР № 526 от 23.03.1990 г.

<sup>8</sup> ГОСТ Р 54569–2011 «Чугун, сталь, ферросплавы, хром и марганец металлические. Нормы точности количественного химического анализа» введен в действие Приказом Росстандарта № 656-ст от 30.11.2011 г.

но, во-первых, зацензурировать выборку — получить малые значения стандартного отклонения, во-вторых, — гарантированно оценить участие лаборатории, предоставивший такой «квазивыбросный» результат, как неудовлетворительный. Новый стандарт обращает на это внимание и рекомендует не рассматривать квазивыбросы как выбросы, а оценку участников проводить с использованием методов робастной статистики.

Стандарт был существенно дополнен методами робастной статистики, которые используются при оценке параметра положения и стандартного отклонения выборки, «загрязненной» выбросами. Простейшей робастной оценкой положения является медиана — находящаяся в середине ряда значение выборки, которая упорядочена по возрастанию.

требует значительных вычислительных ресурсов. Но и надежность сложных методов выше, они работают даже тогда, когда большинство результатов участников представляют собой одинаковые значения и использование простых робастных методов дает некорректные оценки. Применение пакетов специальных программ необходимо и для большинства вычислений по другим разделам стандарта, например, для вычисления неопределенности приписанного значения  $x_{pt}$ , для которого используются алгоритмы на языке программирования R бутстреп-методами [3].

## Графическое представление результатов

В математической статистике большую роль играет анализ графического представления данных, по этой причине в новом

ваемых для отображения на гистограмме. Как пример на *рисунке* представлены гистограммы разного вида для одной выборки при варьировании числа интервалов для раунда межлабораторных сличений по определению массовой концентрации общего железа в воде (28 лабораторий-участниц).

Для решения этой проблемы стандарт предлагает помимо гистограммы использовать график ядерной плотности, зависящий только от робастного стандартного отклонения, который представляет собой сглаженную кривую, описывающую общую форму плотности распределения набора данных. Из *рисунка* видно, что на графике ядерной плотности вид распределения может быть явно отнесен к бимодальному.

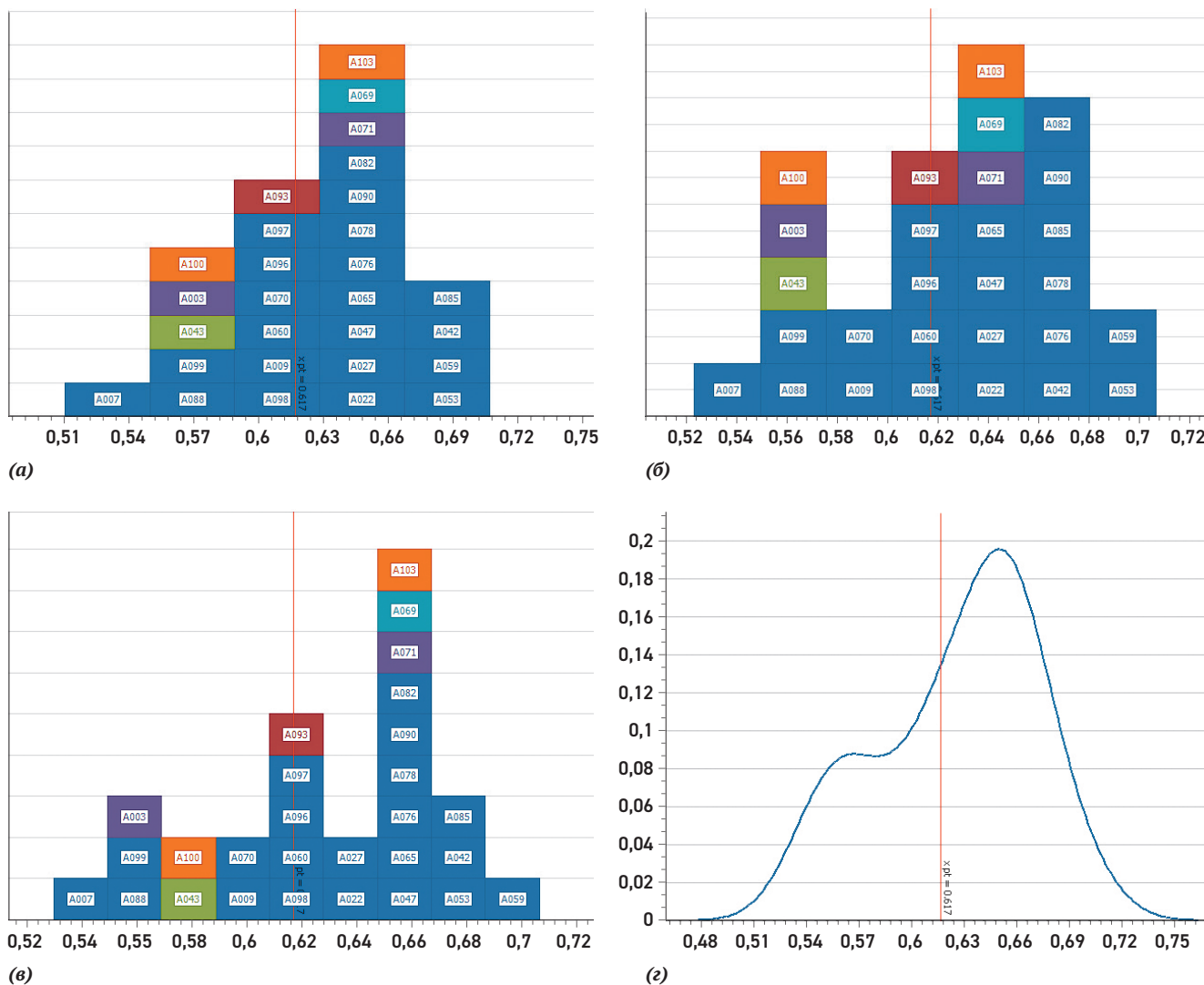
Отдельно остановимся на практических примерах расчетов, приведенных в стандарте. Они занимают примерно треть его объема и оказывают существенную помощь при разборе теоретических алгоритмов. Приводятся примеры для обработки как количественных, так и качественных результатов испытаний. С ними можно свериться при программировании и отладке процедур и лучше понять реализацию тех или иных методов.

В заключении следует отметить, что, к большому сожалению, при внимательном прочтении текста ГОСТ Р 50779.60-2017 обнаруживается ряд ошибок перевода. Некоторые из них являются принципиально важными, искажающими смысл положений стандарта. Перечень таких выявленных ошибок будет направлен автором в Технический комитет ТК 125 «Применение статистических методов».

## Из названия нового стандарта исчез термин «проверка компетентности», являющийся синонимом «проверки квалификации» и ранее вносивший путаницу

Методы стандарта содержат как простые, устойчивые к выбросам оценки для среднего и стандартного отклонения, так и сложные алгоритмы, например —  $Q_n$ ,  $Q$ -метод, оценка Хампеля [1, 2]. Простые методы можно реализовать в электронных таблицах или рассчитать на инженерном калькуляторе, а вот для реализации сложных необходимо специализированное программное обеспечение, работа которого, как указывается в стандарте,

стандарте данному виду анализа уделяется особое внимание. По частотной гистограмме распределения (которая присутствовала и в прежней редакции стандарта) проводится первичный «разведочный» анализ и выясняется, является ли распределение нормальным, одно-(би-)модальным, симметричным, есть ли в нем выбросы. Известна проблема, из-за которой картинка распределения иногда явно зависит от масштаба и интервалов, зада-



Рисунок

Зависимость вида распределения от количества интервалов, задаваемых на гистограмме: а — гистограмма для пяти интервалов (одномодальное распределение); б — гистограмма для семи интервалов (бимодальное распределение); в — гистограмма для девяти интервалов (многомодальное распределение); г — график ядерной плотности (бимодальное распределение)

### Использованная литература:

1. Uhlig S. Robust estimation of variance components with high breakdown point in the 1-way random effect model/Proceedings of the Satellite Conference to the 51st Session of the International Statistical Institute (ISI), Athens, Greece, August 16–17, 1997. — P. 65–73.

2. Maronna R.A., Martin R.D., Yohai V.J. Robust Statistics: Theory and methods. John Wiley & Sons Ltd. — Chichester, England, 2006. — 417 p.

3. Davison A.C., Hinkley D.V. Bootstrap Methods and Their Application. — Cambridge: University Press, 1997. — 582 p.



## Резюме

Подходы нового стандарта можно назвать актуальными, содержащими последние достижения математической статистики, и одновременно достаточно доступно изложенными для специалистов. Следуя требованиям и рекомендациям нового нормативного документа, провайдеры проверки квалификации получат более надежные методы для оценки, а испытательные лаборатории будут уверены в корректной обработке представленных результатов.