

УДК 661.1

Н.Ю. Ефремова

Белорусский государственный институт метрологии, Минск, Республика Беларусь

ОСОБЕННОСТИ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ ПРОВЕРКИ КВАЛИФИКАЦИИ В СООТВЕТСТВИИ С МЕЖДУНАРОДНЫМ ГАРМОНИЗИРОВАННЫМ ПРОТОКОЛОМ IUPAC

В статье представлены основные особенности и преимущества статистического анализа данных проверок квалификации и процедуры анализа данных при проверке на однородность испытываемых образцов, предлагаемых Международным Гармонизированным протоколом IUPAC, по отношению к представленным в международном стандарте ISO 13528. Рассматривается пример с реальными данными проверки квалификации и определение по ним приписанного значения и его неопределенности с помощью рекомендованного Протоколом IUPAC алгоритма, основанного на построении функции ядерной плотности и процедуры «бутстреппинга». Рассматривается точка зрения, высказываемая в Протоколе IUPAC по отношению к неопределенности измерения, представляемой участниками проверки квалификации, и возможность учета ее значений при обработке данных.

Ключевые слова: проверка квалификации, статистический анализ, статистики функционирования, бутстреппинг, ядерная плотность, приписанное значение, неопределенность измерений.

Проверки квалификации, их регламентация и выполнение

Чтобы постоянно производить надежные данные, лаборатория должна внедрить соответствующую систему обеспечения качества и процедуры контроля выполняемых работ (измерений, испытаний, анализа). Всесторонняя программа обеспечения

качества в испытательных лабораториях включает следующие элементы: валидация аналитических методов; использование сертифицированных стандартных образцов (CRM) (когда возможно), в том числе для внутреннего контроля качества. В теории, валидация методов и внутренний контроль качества сами по себе достаточны для обеспечения точности. Но на практике, они часто менее совершенны: при

валидации метода неизвестные факторы влияния могут достаточно сильно воздействовать на процесс измерения, а CRM во многих секторах являются недоступными. Без внешнего эталона лаборатории могут работать в течение длительного периода со смещением или значительными случайными вариациями результатов измерений. Проверка квалификации является средством определения и инициирования исправления таких проблем, а также средством гарантии того, что эти две внутрилабораторные процедуры работают успешно. Проверка квалификации является инструментом, с помощью которого участники могут получать внешнюю независимую оценку точности своих результатов. Именно это свойство проверки квалификации, а также современные требования к участию в программах проверки квалификации аккредитованных лабораторий привели к необычайному развитию проверки квалификации в последнее десятилетие.

Проверки квалификации реализуются на основании программ, и проводится независимыми организациями – Провайдерами. Программы проверки квалификации основаны на распределении образцов испытываемого однородного материала участникам проверки квалификации. Участники анализируют образец без имеющихся сведений о правильном результате и возвращают результат измерений Провайдеру. Провайдер преобразует результаты в соответствующие количественные показатели, называемые статистики функционирования и отражающие качество функционирования участвующей лаборатории при выполнении определенного вида измерений (испытаний, анализа).

Международная организация по стандартизации (ISO) опубликовала стандарт ISO/IEC 17043 [1], предъявляющий требования к компетентности провайдеров проверки квалификации, а также к разработке и реализации программ проверки квалификации и стандарт ISO 13528 [2] по использованию статистических методов при проверке квалификации. Данные стандарты внедрены в действие на территории РБ в качестве предварительного СТБ П ISO/IEC 17043 [3] и государственного стандарта СТБ ISO 13528 [4]. Стандарт ISO 13528 основан на первоначальной версии Гармонизированного протокола IUPAC (далее Протокола) [5], опубликованного в 1993 году. В 2006 году была опубликована обновленная версия Протокола [6], которая представляет оптимальный набор методов, основанных на детальном практическом опыте реализации программ проверки квалификации, интерпретированном специально для аналитических лабораторий, а также объединением новейших идей. Обновленная версия Протокола в некоторых случаях расходится с рекомендациями стандарта ISO 13528, и об этом более подробно будет говориться далее в статье. Однако,

следует отметить, что в настоящее время стандарт ISO 13528 уже пересматривается.

Белорусский Государственный институт метрологии (БелГИМ) назначен Госстандартом Республики Беларусь Провайдером проверки квалификации калибровочных, испытательных и поверочных лабораторий. Для осуществления деятельности по проведению проверок квалификации БелГИМ разрабатывает и внедряет систему менеджмента в соответствии с СТБ П ISO/IEC 17043. Статистический анализ данных проверок квалификации выполняется в соответствии с СТБ ISO 13528. Однако, в процессе использования данного стандарта при обработке данных автор статьи, являющийся экспертом по статистике Провайдера-БелГИМ, столкнулся с излишней краткостью и сухостью изложения основных положений документа, а также отсутствием рекомендаций по использованию тех или иных способов обработки данных проверок квалификации. Для улучшения понимания «философии» принципов (в том числе статистических), на которой основаны проверки квалификации, Провайдером-БелГИМ было принято решение использовать обновленный Протокол [6] и оптимизировать основные статистические критерии обработки данных проверок квалификации на его основе.

Протокол IUPAC и область его применения

Хотя документ [6] носит название "Протокол", он не пропагандирует философию о том, что существует только один хороший способ проведения проверки квалификации. В Протоколе говорится, что при определенных обстоятельствах могут потребоваться альтернативные процедуры, и что выбор такой процедуры является ответственностью провайдера, консультирующегося с техническим экспертом.

Собственно сам Протокол идет первым и включает серию относительно коротких разделов, излагающих основные организационные действия, необходимые для проведения программ проверки квалификации. Далее следует ряд более длинных разделов и приложений, имеющих практическую направленность, касающуюся особенностей расчета статистик функционирования, определения приписанных значений и их неопределенностей, установления стандартного отклонения для проверки квалификации, выполнения испытаний на однородность и стабильность образцов для проверки квалификации.

Протокол в большей степени ограничивается научными и техническими аспектами проверки квалификации, используемой как средство для улучшения выполнения измерений, поэтому в нем не затрагиваются такие вопросы, как квалификация или дисквалификация лабораторий или персонала для

конкретных целей, или аккредитация провайдеров проверки квалификации. Документ также не предназначен для оценивания калибровочных служб.

Преобладающая черта Протокола состоит в том, что проверка квалификации должна обеспечить информацией *о соответствии назначению* результатов измерений, представляемых участниками, с тем, чтобы помочь им выполнить определенные требования.

Этот протокол применим, когда:

- основная цель состоит в оценивании функционирования лаборатории по отношению к установленным критериям, основанным на соответствии общей (единой) цели;

- о соответствии этим критериям можно судить на основании отклонения результатов измерений от приписанного значения;

- результаты участников представляются в шкале интервалов или отношений.

Особенности подхода протокола IUPAC к анализу данных проверок квалификации

Далее в статье представлены основные особенности Протокола, касающиеся статистического анализа данных проверок квалификации, которые используются Провайдером-БелГИМ с целью оптимизации статистических методов, представленных в ISO 13528. Основные положения Протокола используются Провайдером-БелГИМ в первую очередь при статистическом анализе данных проверок квалификации, проводимых в аналитических химических лабораториях. Но также были предприняты попытки по использованию Протокола для данных, полученных при проверках квалификации испытательных лабораторий, выполняющих другие виды испытаний, такие как, строительные и лаборатории радиационного контроля, как будет показано далее на примере.

1. Статистики функционирования

В отличие от ISO 13528, который предлагает 8 статистик функционирования, такие как оценка лабораторного смещения, процентные разности, ранги и процентные ранги, числа E_n , количественные показатели z , z' , $ж$ и E_z , Протокол рекомендует преобразование результатов участников только в z -показатели, так как опыт последнего десятилетия продемонстрировал широкую применимость и признание z -показателя в проверках квалификации. Кроме того, основная идея z -показателя состоит в том, чтобы сделать все показатели проверки квалификации сравнимыми, так чтобы смысл показателя был сразу же очевиден для любого провайдера, участника, или конечного пользователя, связанного с проверкой квалификации вне зависимости от концентрации или подлинности аналита, природы контрольного материала или физического принципа,

лежащего в основе измерения. Скажем показатель $z = -3,5$ вне зависимости от условий его получения должен сразу же иметь следующий смысл: функционирование участника в данном туре проверки квалификации при измерении данной величины является неудовлетворительным.

На основании таких аргументов Провайдер-БелГИМ для оценивания характеристик функционирования испытательных лабораторий использует по возможности только z -показатель. Количественный показатель z рассчитывается в соответствии с выражением:

$$z = (x - x_a)/y_p \quad (1)$$

где x_a – "приписанное значение", наилучшая оценка провайдера для значения измеряемой величины; y_p – "стандартное отклонение для оценки квалификации", основанное на соответствии назначению.

В ISO 13528 для стандартного отклонения для оценки квалификации используется символ $\hat{\sigma}$. В Протоколе обозначение y_p используется для подчеркивания важности приписывания диапазона, соответствующего конкретной цели.

В формуле (1) разность $(x - x_a)$ является, по сути, погрешностью измерения. Стандартное отклонение для оценки квалификации y_p – это параметр, который используется для масштабирования лабораторных отклонений $(x - x_a)$ от приписанного значения.

2. Определение стандартного отклонения для оценки квалификации

Существует несколько способов определения значения этого параметра, но, как было сказано выше, Протокол делает упор на значении, которое отражает соответствие назначению для области использования результата измерения. В этом способе определения стандартного отклонения для оценки квалификации провайдер проверки квалификации определяет уровень неопределенности, который является «подходящим» в качестве общепринятого для участников и конечных пользователей данных в определенной области использования результата, и выражает его в терминах y_p . Такая стандартная неопределенность, "*соответствующая назначению*" будет необязательно близка к неопределенности, связанной с представленными результатами, но она не представляет основную идею о том, как лаборатории функционируют, а лишь то, как они должны функционировать, чтобы выполнить обязательства перед своими клиентами.

Неопределенность, которая *соответствует назначению* результата измерения, зависит от его использования. Например, если относительная стандартная неопределенность измерений, равная 10 %, является подходящей для многих измерений в окружающей среде, то для определения содержания металла в партии металлолома, содержащего золото,

с целью определения ее коммерческого значения требуется гораздо меньшая относительная неопределенность измерений.

Решение о неопределенности для *соответствия назначению* - это компромисс между затратами на анализ и затратами на принятие некорректного решения. Получение меньшей неопределенности требует непропорционально больших расходов на анализ. Но использование методов с большей неопределенностью означает большую вероятность принятия дорогого некорректного решения, основанного на данных. *Соответствие назначению* определяется неопределенностью, которая балансирует между этими факторами, т. е. она должна быть достаточно малой, чтобы решения, основанные на данных, оказывались очень редко некорректными, но не такой малой, чтобы расходы на анализ оказались чрезмерно высоки. Специалисты по выполнению измерений (технические эксперты) и их заказчики обычно не проводят формальный математический анализ ситуации, но они должны, по крайней мере, прийти к соглашению о том, что подразумевается под *соответствием назначению* для каждого конкретного применения.

Некоторые программы проверки квалификации не оперируют положением «*соответствие назначению*» и провайдер проверки квалификации рассчитывает показатель исключительно из результатов участников, опираясь на стандартные статистические идеи, заключающиеся в использовании в качестве u_r робастного стандартного отклонения результатов участников. Последствия этой стратегии в том, что в каждом случае примерно 95 % участников получают очевидно приемлемые значения z -показателей. Это удобный результат как для участников, так и для провайдеров проверки квалификации, но, опять же, он предназначен только для определения его отличия от других результатов. Кроме того, что значение, используемое для u_r , будет изменяться от тура к туру, что будет приводить к нестабильной основе для сравнения показателей из нескольких туров.

В некоторых случаях участник может найти, что критерий *соответствия назначению*, используемый провайдером, является неподходящим для основных видов работ, выполняемых лабораторией. На самом деле такая ситуация может быть обычной для лаборатории, имеющей нескольких клиентов, желающих определить одну и ту же величину, но имеющих различные требования к неопределенности. В некоторых программах проверки квалификации участники могут рассчитывать показатели, основанные на своих личных требованиях, касающихся соответствия назначению. В таких случаях участник должен договориться с клиентом о критерии *специфического соответствия назначению* u_{jpr} для каждого специфического применения, и использо-

вать его для расчета соответствующего модифицированного z -показателя.

3. Определение приписанного значения и его неопределенности

Методы определения приписанного значения, представленные в Протоколе, практически совпадают с представленными в ISO 13528, и являются следующими:

- измерение в референтной лаборатории;
- использование сертифицированного стандартного образца (SRM);
- непосредственное сравнение материала для проверки квалификации с SRM;
- согласованное значение от экспертных лабораторий;
- по процедуре приготовления;
- согласованное значение от участников.

Тем не менее, в связи с тем, что в секторах измерений, где стандартные образцы находятся в дефиците или не существуют совсем, в Протоколе отдельный упор делается на получение приписанного значения, как согласованного на основании результатов участников. Преимуществом согласованного значения участников является низкая стоимость его получения.

В ISO 13528 согласованное значение от участников определяется как робастное среднее на основании «алгоритма А». В Протоколе представлен другой алгоритм для оценивания согласованного значения и его неопределенности, который включает:

- исследование визуального представления данных с помощью столбиковой или точечной диаграммы или гистограммы,
- сравнение робастного стандартного отклонения со стандартным отклонением для оценки квалификации,
- использование оценок ядерных плотностей (kernel density) для распределения результатов,
- нахождение моды и медианы для распределения результатов измерений и сравнение их друг с другом.

Оценивание совокупности с помощью построения ядерной плотности является достаточно новым и малоизвестным в отечественной метрологической практике способом анализа данных. Построение ядерной плотности - это хороший метод восстановления неизвестного распределения генеральной совокупности на основании случайной выборки данных, который заключается в сглаживании данных при сохранении их общей структуры и требует большого количества вычислений, реализовать которые возможно только с помощью вычислений на компьютере [7].

Рассмотрим более подробно представленный в Протоколе алгоритм получения приписанного значения и его неопределенности на основании резуль-

татов участников на конкретном примере и покажем отличие результатов использования этого алгоритма по сравнению с использованием робастного значения в соответствии с «алгоритмом А» по ISO 13528.

Результаты измерений для иллюстрации примера взяты из программы проверки квалификации по определению качественных параметров строительных материалов и представлены в таблице 1. Анализ данных и расчет согласованного значения в соответствии с Протоколом будет следующий:

а) Проанализируем весь набор результатов измерений с целью выявления экстремальных выбросов (например, тех которые находятся вне диапазона $\pm 50\%$ от медианы). Экстремальных результатов измерений в данной совокупности обнаружено не было.

б) Исследуем визуальное представление результатов измерений с помощью точечного графика,

так как количество результатов измерений в совокупности данных менее 50, для больших наборов данных (более 50) Протокол рекомендует выполнять визуальное исследование с помощью построения столбиковой диаграммы или гистограммы. Визуальный анализ точечного графика (см. рисунок 1) показывает, что распределение совокупности результатов измерений является унимодальным, но обладает явной асимметрией. В этом случае можно поставить под сомнение использование робастного значения в качестве приписанного значения. В таких ситуациях Протокол рекомендует сделать оценку ядерной плотности (kernel density) распределения результатов измерений, и затем использовать моду в качестве приписанного значения, а ее стандартную ошибку в качестве неопределенности приписанного значения.

Таблица 1

Результаты измерений предела прочности цементного раствора при сжатии при пропаривании

шифр лаборатории	A	B	C	D	E	G	H	I	J	K
результат измерения, МПа	34,03	39,55	40,20	35,84	33,90	37,36	39,68	28,70	36,45	42,05
шифр лаборатории	L	M	N	O	P	Q	R	S	T	U
результат измерения, МПа	39,83	39,90	40,20	34,80	28,45	37,56	39,83	40,10	37,78	42,78

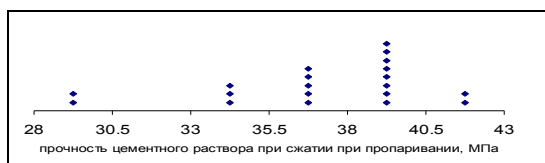


Рис. 1. Точечный график распределения результатов измерений

в) Выполним оценку ядерной плотности. Для построения графика ядерной плотности используется специально разработанное для Excel программное обеспечение [www.rcc.org/ame]. График полученной ядерной плотности представлен на рис. 2. График ядерной плотности показывает асимметрию и наличие второстепенной моды.

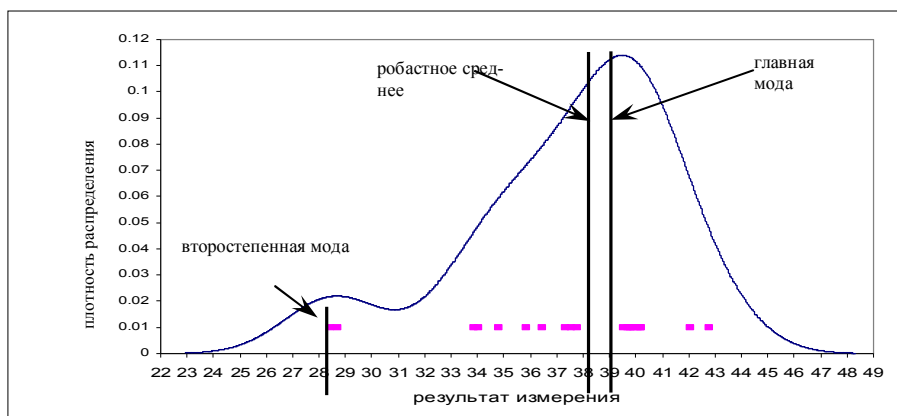


Рис. 2. График ядерной плотности

г) Рассчитаем итоговые статистики по результатам построения ядерной плотности. Стандартная ошибка моды рассчитывается с помощью процедуры бутстреппинга (bootstrapping) в соответствии с [8]. Основные числовые характеристики (итоговые статистики), рассчитанные по результатам измерений, в том числе на основании построенной ядерной плотности, являются следующими:

- среднее арифметическое значение составляет 37,45 МПа;

- медиана составляет 38,67 МПа;
- робастное среднее составляет 37,88 МПа;
- робастное стандартное отклонение – 3,45 МПа;
- мода составляет 39,50 МПа;
- стандартная ошибка моды – 1,07 МПа.

Из этих итоговых статистик видно, что значения робастного среднего и главной моды отличаются друг от друга. При этом значение медианы оказалось значительно ближе к значению главной моды, чем значение робастного среднего.

д) За приписанное значение принимается значение главной моды, равное 39,50 МПа, а за стандартную неопределенность приписанного значения – стандартная ошибка моды, равная 1,07 МПа.

Таким образом, из рассмотренного примера видно, что процедура получения приписанного значения в соответствии с рекомендациями Протокола приводит к приписанным значениям, отличным от тех, которые получают в соответствии с алгоритмом А по ISO 13528. Процедура установления приписанного значения является очень важной, так как от его величины зависят значения статистик функционирования, а, сле-

довательно, и оценки функционирования, которые получают участники проверки квалификации по результатам участия в проверке квалификации. Даже в данном примере, не смотря на, казалось бы, незначительное отличие между значениями робастного среднего и моды (равное 1,62 МПа) различие в статистиках функционирования, получаемым при принятии этих значений в качестве приписанных, оказалось значительным, и будет приводить к различным оценкам функционирования участников в данном туре проверки квалификации. Эти отличия показаны на примере расчета z-показателей и представлены в табл. 2.

Таблица 2

Сравнение z-показателей, рассчитанных на основании разных приписанных значений

шифр лаборатории	A	B	C	D	E	G	H	I	J	K
z (приписанное значение=робастное среднее)	-1,1	0,5	0,7	-0,6	-1,1	-0,1	0,5	-2,6	-0,4	1,2
z (приписанное значение=мода)	-1,6	0,0	0,2	-0,1	-1,6	-0,6	0,1	-3,1	-0,9	0,7
шифр лаборатории	L	M	N	O	P	Q	R	S	T	U
z (приписанное значение=робастное среднее)	0,6	0,6	0,7	-0,9	-2,7	-0,1	0,6	0,6	0,0	1,4
z (приписанное значение=мода)	0,1	0,1	0,2	-1,3	-3,2	-0,6	0,1	0,2	-0,5	0,9

Из табл. 2 видно, что для лабораторий I и P использование в качестве приписанного значения моды, полученной на основании оценки ядерной плотности распределения вероятностей, приводит к неудовлетворительным значениям z-показателей, в то время как использование робастных оценок приводит только к сомнительным значениям.

Алгоритм получения приписанного значения, представленный в Протоколе, является хоть и более сложным, чем просто расчет робастного значения в соответствии с ISO 13528, но зато более правильным с точки зрения статистического анализа данных, так как он не просто позволяет «бездумно» рассчитать числовую характеристику и принять ее в качестве приписанного значения, а предусматривает анализ распределения данных с помощью оценки ядерной плотности, позволяющий получить более точное и обоснованное для данной совокупности приписанное значение.

Еще одним преимуществом Протокола является возможность предъявления менее жесткого требования к неопределенности приписанного значения, чем в ISO 13528, определяющего соотношение $u(x_a) \leq 0,3 \hat{\sigma}$, которое для ряда программ проверки квалификации в области испытаний является практически недостижимым условием. Протокол в зависимости от требований программы проверки квалификации и особенностей видов испытаний, на усмотрение и ответственность Провайдера, допускает выбор соотношения из интервала: $0,3 < u(x_a)/y_p < 0,7$. Для ряда программ проверки квалификации Провайдер-БелГИМ устанавливает соотношение $u(x_a)/y_p < 0,5$, что является более реалистичным с точки зрения достижения на практике.

4. Данные от участников, представленные с неопределенностью

Как ни странно, Протокол не рекомендует участникам представлять вместе с результатом неопределенность измерения, так как предполагается, что после тщательного рассмотрения техническим экспертом, программа, как правило, устанавливает значение y_p , которое представляет *соответствие назначению* для всей области использования, как было подробно объяснено выше. То есть такое оптимальное требование неопределенности уже подразумевается в программах. От участников ожидается, что они будут функционировать в соответствии с этим требованием, и поэтому (в контексте участия в проверке квалификации) от них не требуется сообщать неопределенности в явном виде. Корректно оцененные неопределенности будут ожидаться главным образом похожими на значение y_p , а недооценки будут приводить к неудовлетворительным z-показателям. В таких обстоятельствах представление неопределенности не добавляет ценности программе. Однако эти аргументы не могут быть универсальными, так как данные по неопределенности все чаще требуются клиентами лабораторий, и лаборатории должны соответственно проверять свои процедуры для выполнения этого требования. Но даже если участники представляют свои результаты с неопределенностью, Протокол рекомендует для получения согласованного значения и его неопределенности использовать невзвешенные методы (т. е. методы, не принимающие во внимание индивидуальные неопределенности измерений).

5. Испытания для определения достаточной однородности

Материалы, подготавливаемые для проверок квалификации и других межлабораторных исследований,

обычно являются в некоторой степени неоднородными, и получаемые образцы будут незначительно отличаться по составу друг от друга. Протокол требует, чтобы эта вариация была достаточно малой для цели проверки квалификации и, так как стоимость испытаний на однородность часто является высокой, предлагает сделать основной упор на то, чтобы избежать «ошибки 1 рода» (т. е., ошибочный отказ от удовлетворительного материала). Такое решение отразилось в создании новой статистической процедуры для проверки достаточной однородности. Представленная в Протоколе процедура предлагает проверку результатов, полученных при испытании материалов на однородность, на наличие выбросов с помощью статистического критерия Кохрена и основана на использовании методов дисперсионного анализа и F-критерия. Протокол также предлагает детальные инструкции, касающиеся испытаний на однородность, которые он рекомендует выдавать лабораториям, выполняющим такие испытания.

Заключение

1) Основные особенности статистического анализа данных проверок квалификации в соответствии с Протоколом IUPAC позволяют при статистической обработке данных проверок квалификации:

- разрабатывать критерий анализа данных (стандартное отклонение для проверки квалификаций) исходя из концепции «соответствия назначению» результатов измерений;

- использовать для оценивания характеристик функционирования и интерпретировать в едином ключе только один количественный z-показатель;

- использовать современный и надежный алгоритм определения приписанного значения и его неопределенности на основании результатов участников, оперирующий методом оценки ядерной плот-

ности и оценивания неопределенности измерения с помощью процедуры «бутстреппинга»;

- правильно использовать и интерпретировать результаты измерений участников, представленные с неопределенностями.

2) Представленные в статье особенности подхода к анализу данных проверок квалификации, предлагаемые протоколом IUPAC, позволили оптимизировать Провайдеру – БелГИМ свои алгоритмы статистической обработки и позволяют выполнять анализ и интерпретацию данных проверок квалификации на более современном и профессиональном уровне.

Список литературы

1. ISO/IEC 17043:2010 *Conformity assessment – General requirements for proficiency testing*.

2. ISO 13528:2005 *Statistical methods for use in proficiency testing by interlaboratory comparisons*.

3. СТБ П ISO/IEC 17043 – 2010/2011 *Оценка соответствия. Основные требования к проведению проверки квалификации*.

4. СТБ ISO 13528 – 2011 *Статистические методы, применяемые при проверке квалификации лабораторий посредством межлабораторных сличений*.

5. Thompson M. and Wood R. *The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories*, AOAC, September 1993.

6. Thompson M., Ellison S.L.R., Wood R. *The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report)*, in *Pure and Applied Chemistry*, Vol. 78, No. 1, pp. 145-196, 2006.

7. Analytical Methods Committee. *Technical brief. Representing data distributions with kernel density estimates*, АМСТВ № 4, 2006.

8. Efron. B and Tibshirani R. J., *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.

Поступила в редколлегию 21.02.2013

Рецензент: д-р техн. наук, проф. Ю.П. Мачехин, Харьковский национальный университет радиоэлектроники, Харьков.

ОСОБЛИВОСТІ СТАТИСТИЧНОЇ ОБРОБКИ ДАНИХ ПЕРЕВІРОК КВАЛІФІКАЦІЇ ВІДПОВІДНО ДО МІЖНАРОДНОГО ГАРМОНІЗОВАНОГО ПРОТОКОЛУ IUPAC

Н.Ю. Ефремова

У статті представлено основні особливості та переваги статистичного аналізу даних перевірок кваліфікації та процедури аналізу даних при перевірці на однорідність зразків, що випробовуються, які запропоновано Міжнародним гармонізованим протоколом IUPAC, по відношенню до представлених у міжнародному стандарті ISO 13528. Розглядається приклад з реальними даними перевірки кваліфікації і визначення по ним приписаного значення та його невизначеності за допомогою рекомендованого Протоколом IUPAC алгоритму, заснованого на побудові функції ядерної щільності і процедури «бутстрепінга». Розглядається точка зору, висловлювана в Протоколі IUPAC по відношенню до невизначеності вимірювань, що подається учасниками перевірки кваліфікації, і можливість обліку її значень при обробці даних.

Ключові слова: перевірка кваліфікації, статистичний аналіз, статистики функціонування, бутстрепінг, ядерна щільність, приписане значення, невизначеність вимірювань.

SPECIFICITY OF STATISTICAL DESIGN FOR PROFICIENCY TESTING RESULTS IN ACCORDANCE WITH IUPAC INTERNATIONAL HARMONIZED PROTOCOL

N.Y. Efremova

The paper presents the key features and benefits of the statistical design for proficiency testing results and homogeneity checks of samples which is proposed by IUPAC International Harmonized Protocol against the statistical design provided by the international standard ISO 13528. The paper describes an example, in which the assigned value and its uncertainty for proficiency testing are determined as consensus values from real measurement results by making a kernel density estimate of the distribution and bootstrapping the data, recommended by the IUPAC Protocol. In the paper the point of view, proposed in the IUPAC Protocol to the measurement uncertainties represented by the proficiency testing participants and the possibility of taking into account of their values in statistical design for proficiency testing results is discussed.

Keywords: proficiency testing, statistical design, performance statistics, bootstrapping, kernel density, assigned value, measurement uncertainty.