

Неадекватность математических моделей объектов измерений и расчёты риска согласно ГОСТ ISO/IEC 17025-2019

С. Ф. Левин

Московский институт экспертизы и испытаний, Москва, Российская Федерация,
e-mail: AntoninaEL@rostest.ru

Рассмотрена проблема неадекватности математических моделей объектов измерений в связи с проблемой «дефиниционной неопределённости измерения» и с необходимостью управления рисками согласно ГОСТ ISO/IEC 17025-2019 «Общие требования к компетентности испытательных и калибровочных лабораторий». Описана предыстория проблемы от формирования моментного и композиционного подходов к оцениванию точности до введения специального термина – погрешности неадекватности математической модели объекта измерений. Отмечено негативное влияние безвыходности понятийно-терминологических преобразований в метрологии и критическое противоречие оценок применимости «статистических методов» ГОСТ Р ИСО/МЭК 31010-2011 «Менеджмент риска. Методы оценки риска» и Руководства по выражению неопределённости измерения. Показано, что учёт неадекватности вероятностных моделей в расчётах риска является необходимым условием достоверности результатов измерений.

Ключевые слова: математическая модель, дефиниционная неопределённость, моментный подход, композиционный подход, погрешность неадекватности модели.

С 1 сентября 2019 г. введён в действие ГОСТ ISO/IEC 17025-2019 «Общие требования к компетентности испытательных и калибровочных лабораторий». И теперь **риски статистических предположений, ложного принятия и ложного отклонения** решений о соответствии лаборатории должны оцениваться объективно, беспристрастно и постоянно. При этом отмечено, что международные или национальные стандарты и признанные технические требования, содержащие достаточную информацию о том, как осуществлять лабораторную деятельность, не требуются дополнять или переписывать в качестве внутренних процедур лаборатории.

«Риск-ориентированный подход» описан в ряде стандартов «международного происхождения»: ГОСТ Р ИСО 31000-2010 «Менеджмент риска. Принципы и руководство», ГОСТ Р 51901.7-2017/ISO/TR 31004:2013 «Менеджмент риска. Руководство по внедрению ИСО 31000» и ГОСТ Р ИСО/МЭК 31010-2011 «Менеджмент риска. Методы оценки риска» [в июне 2019 г. появилась новая версия – ISO/IEC 31010:2019 “Risk management – Risk assessment techniques”]. В этих стандартах **риск** определен как «влияние **неопределённости** на цели», а получение оценок предполагает процесс идентификации, анализа и оценивания риска.

Вместе с тем согласно Р 50.2.004-2000 «ГСИ. Определенные характеристики математических моделей зависимостей между физическими величинами. Основные положения» в общем случае **неопределённость в широком смысле** характеризует распределение вероятностей на множестве значений фактора риска, а **неопределённость в узком смысле** – параметр рассеяния этого же распределения [1].

Для лабораторий фактором риска в контрольных испытаниях является искомая величина, для которой установлена область несоответствия, в определительных испытаниях – показатели точности и (или) достоверности, а при калибровке – несоответствие диаграммы калибровки нормам доверительной вероятности.

В математической статистике **риск** характеризуют **вероятностями ошибок** 1-го и 2-го рода, а **риск** вообще – **вероятностью** наступления **нежелательного события**.

Перечень методов оценивания риска ГОСТ Р ИСО/МЭК 31010-2011, кроме FMEA (Failure Mode Effect Analysis) процесса, системы и программного обеспечения «для идентификации отказа компонентов, систем или процессов, которые могут привести к невыполнению их функции», содержит «вероятностно-статистические методы» – Марковский анализ, метод Монте-Карло, байесовский анализ и сети Байеса. Эти методы регламентированы ГОСТ Р 51901.12-2007 «Менеджмент риска. Метод анализа вида и последствий отказов» (FMEA), ГОСТ Р 50901.15-2005 «Менеджмент риска. Применение Марковских методов» и ГОСТ Р 54500.3.1-2011 / Руководство ИСО/МЭК 98-3:2008 / Дополнение 1:2008 «Неопределённость измерения. Ч. 3. Руководство по выражению неопределённости измерения. Дополнение 1. Трансформирование распределений с использованием метода Монте-Карло». При этом отмечено следующее:

работа с Марковским анализом невозможна без знания операций с матрицами, а полученные результаты трудны для понимания персоналом, не имеющим соответствующих технических знаний, навыков и опыта;

метод Монте-Карло используют для трансформирования распределений и расчёта вероятностей, а исходные данные должны быть *известным распределением*, но метод не может адекватно моделировать события с очень высокой и очень низкой вероятностью появления;

байесовский анализ является *экспертным методом*. Метод «матрицы последствий и вероятностей», как и метод Монте-Карло, требует, чтобы вероятности были известными.

Проблема состоит в том, что в Таблице А.1 «Характеристики методов оценки риска» ГОСТ Р ИСО/МЭК 31010-2011 в графе «Анализ риска – Вероятностные характеристики» метод FMEA, публично высмеянный членом президентской комиссии по расследованию обстоятельств катастрофы Space Shuttle Challenger Нобелевским лауреатом Р. Фейнманом за оценку вероятности катастрофы 0,0001 [2], отмечен знаком **SA** – «строго применим», а Марковский анализ, метод Монте-Карло, байесовский анализ и «сети Байеса» отмечены знаком **NA**, который означает, что для анализа вероятностных характеристик риска эти методы *не применимы*.

Опыт астрономии за две с половиной тысячи лет построения математических моделей «мира» сформировал понятие *точности* решения измерительных задач как совпадение *расчётных* предсказаний с *измеренными* положениями небесных светил (до *точки*), а их разность (число) называли ошибкой. В XIX веке в работах П. Лапласа, К. Гаусса и А. Лежандра возникло представление о расхождении результатов вычислений по моделям движения небесных тел и измерений их положений с *точностью до распределения вероятностей*. Простота вычислений выделила линейные модели, для них основным распределением из-за универсальности при суммировании стал нормальный закон и его *моменты*. Для нелинейных моделей К. Якоби в 1834 г. получил *композицию* как функциональное преобразование распределений вероятностей входных переменных, что ориентировало на применение непараметрических методов. Так возникли *моментный* и *композиционный* подходы к оцениванию точности математических моделей объектов измерений.

Согласно ГОСТ 24026-80 «Планирование эксперимента. Термины и определения» адекватность математической модели определяется как «соответствие математической модели экспериментальным данным по выбранному критерию». «Статистический рецепт Фишера» по спецификации параметрических моделей критериями Стюдента и Фишера (F-критерия) для гипотезы нормальности ведёт к парадоксу Эльясберга–Хампеля [3, 4]. В «рецепте» нет ни слова об отсутствии независимости и катастрофической неробастности F-критерия, о том, что «различные критерии согласия оценивают часто несущественные аспекты подгонки и что статистически “плавную” нулевую гипотезу доказать нельзя, её можно только отклонить, хотя и при очень большом объёме выборки» [4].

Однако математический эквивалент «погрешности неадекватности математической модели» – погрешность предсказания моделью значений случайной функции, не использованных при её идентификации, как распределение погрешностей экстраполяции (функции компактности)

в методе максимума компактности (ММК) [5, 6] был. Прототипом ММК стал метод группового учёта аргументов [7] и схема перекрёстного наблюдения [8]. В 1980-е гг. ММК позволил оценивать точность и риск в задачах контроля объектов, к надёжности которых предъявлялись повышенные требования [6, 9]. Термин «погрешности неадекватности математической модели объекта измерений» и средний модуль погрешности неадекватности (СМПН) модели объекта были введены в Рекомендациях по метрологии Р 50.2.004-2000. Кроме того, анализ в рамках проекта «Прогноз–ММК» [9–11] выявил неизвестный ранее фактор катастроф – *погрешность неадекватности* математических моделей объектов. В список причин катастроф попали статистическая обработка результатов измерений и статистическое моделирование, погрешности измерений и погрешности неадекватности математических моделей.

В МИ 1967-89 «ГСИ. Выбор методов и средств измерений при разработке методик выполнения измерений. Общие положения» *погрешность неадекватности* названа проблемой «наименее развитых областей метрологии» и как методическая составляющая погрешности измерений вошла в условия проверки правильности постановки измерительных задач МИ 2091-90 «ГСИ. Измерения физических величин. Общие требования». Также в МИ 1967-89 *погрешность неадекватности* названа фактором «№ 1» без каких-либо указаний на методы идентификации и в этом качестве фигурирует в ГОСТ Р 22.2.04-94 «Безопасность в чрезвычайных ситуациях. Техногенные аварии и катастрофы. Метрологическое обеспечение контроля состояния сложных технических систем. Основные положения и правила», ГОСТ Р 8.563-96 «ГСИ. Методики выполнения измерений» и его редакции ГОСТ Р 8.563-2009 «ГСИ. Методики (методы) измерений».

В 1986–1987 гг. в СССР по рекомендации рабочей группы ВНИИ стандартизации были отменены 24 стандарта по статистическим методам из 31, позже – и по прикладной статистике. Результаты анализа рабочей группы не были полностью опубликованы, а «традиционные предубеждения» продолжали действовать. Конечно, свято место пусто не бывает. В науке это – терминология, но не столько как список названий и определений, а как *подсистема* в *системе* ассоциаций естественного языка. Ассоциации индивидуумов обладают устойчивостью, нарушение которой представляет собой весьма «болезненный» процесс. Поэтому дискуссии о терминологии отличаются особой напряжённостью. В частности, *метрологический менталитет* опирается на ассоциативно-семантическое поле терминов метрологии с ассоциациями естественного языка, что влияет на выбор математических моделей объектов измерений и на логику статистического вывода при построении моделей. Другими словами, роль искусственных искажений *метрологического менталитета* в проблеме неадекватности чрезвычайно высока.

Терминологическая чехарда вместо работы над «наименее развитыми областями метрологии» привела к беспрецедентному «потоку» переводов с «новой» терминологией: Guide to the Expression of Uncertainty in Measurement (далее – GUM), ISO 3534.1-93 «Statistical methods. Probability

and general statistical terms. Terms and definitions» [ГОСТ Р 50779.10-2000 «Статистические методы. Вероятность и основы статистики. Термины и определения»]; ISO 3534.2-93 "Statistical methods. Statistical quality control. Terms and definitions" [ГОСТ Р 50779.11-2000 «Статистические методы. Статистическое управление качеством. Термины и определения»]; ISO 5725:1994 "Accuracy (trueness and precision) of measurement methods and results" [ГОСТ Р ИСО 5725-2002 «Точность (правильность и прецизионность) методов и результатов измерений»]; ISO 16269-6:2003 "Statistical interpretation of data – Pt. 6. Determination of statistical tolerance interval" [ГОСТ Р ИСО 16269-6-2005 «Статистические методы. Статистическое представление данных. Определение статистических толерантных интервалов»]; словарь VIM-3 [12].

Перевод GUM с художественной точки зрения можно считать почти безупречным, но отсутствие профессионального комментария насторожило [13]. Дело в том, что ряд терминов англоязычной научной литературы переводится как «неопределённость» [14], но имеет иной, чем «рассеяние», смысл. В первую очередь это относится к uncertainty как *неадекватности* математической модели физической величины измеряемому свойству объекта или *неадекватности* прибора цели измерения. В GUM «врождённой» *неопределённости* «на любом уровне детализации определения измеряемой величины» уделено много внимания (подробнее см. [15]) и заявлено о том, что «эта» неадекватность «должна быть» пренебрежимо мала. При этом использованы названия терминов математической статистики, не соответствующие в контексте GUM их определениям, например, *уровень доверия*, или термины *точечная оценка* и *интервальная оценка* подменялись словами «реалистическая оценка» и «безопасная оценка» с явным выражением им предпочтения. А моментом истины стал п. С.3.2 GUM: «Надлежащей мерой неопределённости результата измерения является не дисперсия наблюдаемой величины, а дисперсия среднего арифметического по выборке наблюдений».

При некотором сходстве описаний на уровне понятий явления неадекватности математических моделей в GUM и рекомендациях по метрологии Р 50.2.004-2000 в тексте GUM математическая сторона вопроса не раскрыта. Однако при представлении формул Тейлора как «новых законов» подчёркивается роль теории планирования эксперимента, тогда как в этой теории даже в ситуации, когда неизвестную модель «можно представить в виде отрезка ряда Тейлора, проверка адекватности оказывается единственным критерием годности модели. Все стройное здание построения математической модели может рухнуть, если не выполняются исходные предпосылки, лежащие в основе статистического анализа, такие, как предпосылки о виде законов распределения, однородности дисперсий и т. д.» [16].

В GUM термины *свёртка* явно и *композиция* неявно упомянуты, но «в силу известных трудностей» их применение признано «нецелесообразным», поскольку идеи GUM сосредоточились на ключевом слове «неопределённость». В аутентичном переводе оно ассоциируется с «отсутствием

определённости», а в сочетании со словами «стандартная», «суммарная» и «расширенная» – превращается в идиомы.

ГОСТ Р 50779.10-2000 и ГОСТ Р 50779.11-2000 внесли свою лепту в неадекватность терминологии: доверительная вероятность и уровень доверия – синонимы, нет терминов *композиция*, *свёртка* и *вероятность согласия*, а *толерантный интервал* = statistical coverage interval (статистический интервал охвата). Перевод tolerance interval = «**поле [область] допуска**» и формула толерантного интервала ГОСТ Р ИСО 16269-6-2005, исправленная лишь в одноимённом ГОСТ Р 50779.29-2017, стали фоном нетолерантной критики *толерантных интервалов* специалистами по внедрению «неопределённости» в отечественную поверку [17]. «Критика» закончилась казусом: GUM стал одноимённым ГОСТ Р 54500.3-2011 с правильным определением *толерантного интервала*, но со «старыми» формулами, а в РМГ 29-2013 исчезли все «статистические» термины РМГ 29-99, кроме СКО как синонима стандартного отклонения.

Неадекватность ISO 5725:1994 проявилась в названии, так как «термин precision не следует использовать для термина accuracy» [18]. Математический аппарат ISO 5725:1994 – дисперсионный анализ. Условиями его применимости являются сбалансированность плана измерений, а также гауссовость, некоррелированность и гомоскедастичность данных [19]. В ISO 5725:1994 этим условиям не соответствуют все практические примеры [20–25], раскритикованные прямо в тексте стандарта. Даже установка ГОСТ Р ИСО 5725-2002 на «борьбу с выбросами» превратилась в подгонку под «нормальность», от которой пытались избавиться путём отмены стандартов по статистическим методам в 1986–1987 гг.

В итоге искажение понятий *неопределённость* в GUM и *прецизионность* в ISO 5725:1994 вызвало бесплодную дискуссию.

В 2004 г. на семинаре во ВНИИМ имени Д. И. Менделеева ведущие метрологи Национальной физической лаборатории Великобритании и Физико-технического института Германии М. Сох, W. Wöger, P. Harris и B. Siebert познакомились со статистической идентификацией погрешности неадекватности согласно Р 50.2.004-2000. Тогда было высказано общее мнение, что математическая часть теории безупречна, но можно поспорить по философским вопросам. А в 2007 г. в третьей редакции VIM [12] появился термин: «2.27 **дефинициальная неопределённость** – составляющая неопределённости измерений, являющаяся результатом ограниченной *детализации* в определении измеряемой величины. Примечания: 1. Дефинициальная неопределённость есть практический минимум неопределённости измерений при *любом* измерении данной величины. 2. *Любое* изменение *детализации* в определении величины ведет к другой дефинициальной неопределённости». Здесь *детализация* – аналог *кода структуры модели* из Р 50.2.004-2000.

«Спор» продолжился [1, 26–33] и закончился признанием: «задача, стоящая перед составителями словаря (VIM-3 – *авт.*) – гармонизировать терминологию, используемую при разных подходах, таким образом, чтобы одному понятию соответствовал только один термин, оказалась трудновыполнимой» [28].

Учёный совет ВНИИМ им. Д. И. Менделеева ещё в 2002 г. осудил «англоязычную гармонизацию», но, когда дискуссия о GUM дошла до поверки, один теоретик заявил, что этот вопрос в настоящее время стал тривиальным, и необходимость применения GUM, фактически являющегося международным регламентом, не обсуждается, так как дискуссии по неопределённости измерений – пройденный этап. Но в VIM-3 сохранены оба термина – «неопределённость» и «погрешность».

Если в GUM проблема неадекватности обозначена как «в принципе» решаемый вопрос, то в VIM-3 этот «вопрос» назван *дефинициальной неопределённостью* без какого-либо намёка на принцип решения.

Да, термина «*дефинициальная неопределённость*» в GUM нет. Есть «дефиниция» [34, п. 3.1.3] и «неадекватность» [34, п. 4.1.2], но без математического определения. «Неадекватность» описана в [34, п. D.1.1] как «неполнота модели» на понятийном уровне, а в плане идентификации – в Р 50.2.004-2000. В целом же проблема *дефинициальной неопределённости* с самого начала была центральной для GUM, но об отражении этого факта в международных стандартах говорить пока рано, а критика GUM [13] сразу назвала эту ситуацию «драмой неадекватности».

Нормативно подходы к оцениванию точности «триединством измерений» были разделены в ГОСТ Р 8.820-2013 «ГСИ. Метрологическое обеспечение. Основные положения»: «... следует использовать показатели точности: неопределённость по ГОСТ Р 54500.3, характеристики погрешности по МИ 1317 и РМГ 83, показатели правильности и прецизионности по ГОСТ Р ИСО 5725-1». Композиционный подход здесь представлен свёрткой в Приложении Д МИ 1317-2004 «ГСИ. Результаты и характеристики погрешности измерений. Формы представления. Способы использования при испытаниях образцов продукции и контроле их параметров». В списке нет ГОСТ 8.009-84 «ГСИ. Нормируемые метрологические характеристики средств измерений» и МИ 2916-2005 «ГСИ. Идентификация распределений вероятностей при решении измерительных задач» со свёртками (композициями для линейных моделей) распределений, ГОСТ Р 50779.21-2004 «Статистические методы. Правила определения и методы расчета статистических характеристик по выборочным данным. Ч. 1. Нормальное распределение», ГОСТ Р ИСО 16269-6-2005 с интервальными оценками и Р 50.2.004-2000 с проверкой неадекватности.

ГОСТ Р 50779.21-2004 соответствует ISO 2854:76 “Statistical interpretation of data – Techniques of estimation and tests relating to means and variance” [ИСО 2854-1976 «Статистическое представление данных. Методы оценки и проверки гипотез о средних значениях и дисперсиях»]. Вот только в GUM статистическая проверка гипотез не предусмотрена.

В Руководстве Guide OIML G 19:2017 “The role of measurement uncertainty in conformity assessment in legal metrology” (Роль неопределённости измерений в решениях об оценке соответствия в законодательной метрологии) была заявлена «новая» идея. «Включение понятия неопределённости измерений в принятие решений о соответствии испытаний

в рамках законодательной метрологии требует иного мышления и других рассуждений о таких решениях. В таком вероятностном подходе понятно, что следует учитывать определённые риски, указанные в ISO/IEC 17025:2005. Одним из способов рассмотрения этого является построение функции, известной как функция плотности вероятностей (PDF), которая представляет степень веры (уровень доверия) о знании истинного значения измеряемой величины». Там же была предложена «новая» концепция вероятности: «...PDF кодирует всю известную информацию об измеряемой величине, включая как систематические, так и случайные воздействия. В то время как кривая, подходящая для гистограммы только случайных колебаний, часто имеет гауссову форму, PDF не является подходящей для гистограммы, она скорее содержит дополнительную информацию, поступающую от систематических воздействий при измерении». Но «несмотря на то, что вместо понятий *погрешность* и *анализ погрешности* применительно к измерениям теперь используется понятие *неопределённость*, термин и понятие *погрешность* по-прежнему применимы при поверке средств и систем измерений».

И тут же Росаккредитация обязала лаборатории при определении измерительных возможностей «брать пример с GUM» и учитывать требования поверочных схем к эталонам. Но определить измерительные возможности по «косвенным и совместным измерениям» как «измерениям» значит определить их *дефинициальную неопределённость*.

Только в лабораториях сообразили, что «пересчитать погрешность в неопределённость» просто делением на «корень из трех» нельзя [35], как в ГОСТ ISO/IEC 17025-2019 в соответствии с риск-ориентированным мышлением фактически уже декларируется необходимость «значительного переходного периода для наработки методического аппарата и обучения специалистов» [36].

Согласно Р 50.2.004-2000 «*погрешность неадекватности математической модели* – расчётная величина, разность расчётного значения выходной переменной модели объекта по данным измерений входных переменных и результата её измерения в соответствующих расчёту условиях». Ранее это были погрешности аппроксимации.

Погрешность неадекватности моделей функционального типа как случайная функция имеет характеристики положения и рассеяния. Первая – собственно функциональная зависимость, которую можно рассматривать как систематическую составляющую, а вторая – распределение отклонений от характеристики положения, её можно представлять зоной толерантности. Погрешность неадекватности для характеристики положения определяют остатки экстраполяционного функционала, а для характеристики рассеяния – контурная оценка по статистикам Смирнова или квантилям отклонений эмпирической функции распределения от теоретической. Контурные оценки погрешности неадекватности принятого распределения позволяют не рассматривать альтернативные гипотезы и последовательно строить свёртку функции распределения $F_X(x)$ наблюдаемой составляющей и ненаблюдаемых составляющих в виде

эквивалентных равномерных распределений на интервалах $[a, b]$:

$$f_X(x) * f_R(x) = f_{XR}(x) = [F_X(x-a) - F_X(x-b)] / (b-a).$$

Объединение «*» составляющих погрешности в ГОСТ 8.009-84, МИ 1317-2004 и свёртка распределений вероятностей в ГОСТ Р 54500.3-2011 (п. G.1.4) – одно и то же. Свёртка – это композиция при единичном якобиане функционального преобразования. По ней «могут быть рассчитаны интервалы с заданным уровнем доверия p », а «трапецеидальное распределение можно рассматривать как свёртку двух прямоугольных распределений» [34, пп. 4.3.9, G.1.4]. Но в [34] нет примеров такого расчёта, хотя в ГОСТ Р ИСО 16269-6-2005 есть таблицы толерантных интервалов.

Испытания в целях утверждения типа и поверка средств измерений являются контрольными. Их качественный результат – решение о степени соответствия количественных результатов установленным требованиям, требованиям государственных поверочных схем по доверительной вероятности или ГОСТ 27883-88 «Средства измерения и управления технологическими процессами. Общие требования и методы испытаний». ГОСТ 16504-81 «Система государственных испытаний продукции. Испытания и контроль качества. Основные термины и определения» также характеризует качественные результаты достоверностью, ошибками контроля 1-го и 2-го рода. В МИ 187-86 оценкой вероятности ошибки 2-го рода принята « $P_{\text{бам}}$ – наибольшая вероятность ошибочного признания годным любого в действительности дефектного экземпляра средства измерений», а «при малом числе (или одном) поверяемых СИ вместо $(P_{\text{гр}})_{\text{мг}}$ в качестве критерия может быть установлена наибольшая вероятность $P_{\text{грм}}$ ошибочного признания дефектным любого в действительности годного экземпляра СИ». Риск *принятия ложного* решения связан с вероятностью нахождения контролируемой характеристики объекта вне границ допуска. Трудность лишь в том, как оценить эту вероятность.

Согласно ГОСТ 16504-81 достоверность имеет методическую и инструментальную составляющие. Методическая составляющая достоверности результата контрольных испытаний определяется *планом измерений* при испытаниях и браковочным условием с учётом нормы доверительной вероятности. Инструментальная составляющая определяется алгоритмом контроля и вероятностными свойствами контура контроля, включающего эталоны и средства вычислений.

Различают априорную и апостериорную достоверности. Априорную достоверность результата поверки устанавливают согласно МИ 188-86 «ГСИ. Средства измерений. Установление значений параметров методик поверки» по распределению Иордана. Априорной достоверности фактически посвящен и ГОСТ Р 8.731-2010 «ГСИ. Системы допускового контроля. Основные положения». Апостериорную достоверность поверки согласно МИ 187-86 может

охарактеризовать оперативная характеристика по нормализованной плотности распределения вероятностей:

$$L(\alpha) = \int_{-\gamma}^{\gamma} \varphi(\tilde{\alpha}/\alpha) d\tilde{\alpha},$$

где α , γ – нормализованные значение погрешности и её границы соответственно; $\varphi(\tilde{\alpha}/\alpha)$ – условная (при $\alpha = \tilde{\alpha}$) плотность распределения вероятностей.

В обоих случаях допустима оценка PDF вида $\varphi = f_X(x) * f_R(x)$.

В ГОСТ ISO/IEC 17025-2019 подразумевается, что риски должны оцениваться апостериорно с учётом параметров методик испытаний и калибровки. Это напрямую связано со статистическими предположениями относительно выбора структуры и методов оценивания параметров для моделей метрологических характеристики функционального вида.

Концептуально *дефинициальная неопределённость* – стандартное или среднее квадратическое (СКО) отклонение, но принцип её идентификации в GUM не указан. «Безусловным остается факт совпадения свёрток распределений составляющих бюджета неопределённости, оцениваемых по типу В, и равномерно распределённых компонентов бюджета погрешностей» [15]. Совпадение расчётов погрешности и неопределённости уже было в Р 50.2.038-2004 «ГСИ. Измерения прямые однократные. Оценивание погрешностей и неопределённости результата измерений», и Приложение Б с неопределённостью измерения было исключено.

Расчёт методами теории погрешностей с учётом погрешности неадекватности позволяет оценить *дефинициальную неопределённость* как компонент параметра рассеяния композиции или свёртки (т. е. PDF) выходной переменной «уравнения измерений» – *СКО композиции* [37–39, 15].

В «образцовом» примере «калибровки по GUM» термометра по статистическому ряду поправок b_k , $k=1,11$ с СКО $s=0,0035$ °C получена линейная функция поправок [34, п. Н.3]:

$$b(t) = y_1 + y_2(t - t_0).$$

При $t_0 = -20$ °C и $t = 21..27$ °C калибровка методом наименьших квадратов (МНК) дала оценки $y_1 = -0,1712$ °C, $y_2 = 0,00218$ и поправку $b(t = 30$ °C) = $-0,1494$ °C с суммарной стандартной неопределённостью $u_c = 4,1 \cdot 10^{-3}$ °C.

Однако анализ «примера GUM» выявил ряд «странностей». Во-первых, точность эталонного термометра не дана, контрольного измерения нет, а правила округления нарушены. Во-вторых, выбор модели обоснован тем, что «угловой коэффициент y_2 более чем в три раза превосходит свою стандартную неопределённость». Но согласно [34, п. 3.4.2] функция поправок как модель может быть *неполной*, и «для получения достоверных оценок неопределённости рекомендуется, по возможности, использовать эмпирические математические модели», причём возможность есть, а оценок неадекватности нет. В-третьих, хотя согласно [34, п. Н.3.6] утверждается, что «метод наименьших квадратов применим в случае, когда данные измерений известны неточно»,

для МНК это нарушение условия неконфликтности [37–39]. В-четвёртых, гипотеза «нормальности» не проверяется. В-пятых, функция поправок термометра содержит «разладки» [13], что ограничивает экстраполяцию модели.

Конечно, описанный выше пример калибровки «по GUM» – не образец для описания области аккредитации лабораторий. Калибровочная кривая [34, п. Н.3] как модель с кодом структуры $\mathcal{G}=11$ была аттестована по программе «ММК–стат 2.0» [11, 40]: $b_{11}^{[2]}(t) = -0,2148514 + 0,002182436t$, где жирным шрифтом выделены совпавшие при аттестации цифры оценок параметров модели; СМПН $\bar{\varepsilon}_{11}^{[2]} = 4,750704 \cdot 10^{-3} \text{ } ^\circ\text{C}$; $^{[2]}$ – признак МНК; «11» – код; среднее абсолютное отклонение (CAO) поправок $d_{11}^{[2]} = 2,632977476 \cdot 10^{-3} \text{ } ^\circ\text{C}$.

Контроль погрешностей неадекватности систематической составляющей функции поправок, играющий ту же роль, что и контроль параметра Ω_p согласно МИ 187-86 «Методические указания ГСИ. Средства измерений. Критерии достоверности и параметры методик поверки» [41], позволяет оценить их модуль как дополнительную составляющую погрешности эталона $\Omega_N = \text{СМПН} - \text{CAO} = 2,117726524 \cdot 10^{-3} \text{ } ^\circ\text{C}$.

Идентификация по программе «ММИ–поверка» распределения поправок b_k , $k=1,11$ относительно среднего арифметического значения поправки $b_{cp} = -0,1635 \text{ } ^\circ\text{C}$ по критерию минимума контурной оценки погрешности неадекватности для функции распределения вероятностей среди типовых усечённых распределений свёртки в качестве наиболее правдоподобного дала распределение Лапласа при $\Delta_N = 1,617 \cdot 10^{-3} \text{ } ^\circ\text{C}$. Его свёртка с эквивалентным равномерным распределением в пределах $\pm \Omega_N$ соответствует интервалу $[-0,012618; +0,012852] \text{ } ^\circ\text{C}$ и даёт оценку дефинициальной неопределённости $u_D = (\Delta_N + \Omega_N) / \sqrt{3} = 2,156245364 \cdot 10^{-3} \text{ } ^\circ\text{C}$ или 61,6 % СКО поправок s [34, п. Н.3].

Статистические предположения в методе совместных измерений [37–39] предусматривают оценивание неадекватности характеристик положения и рассеяния моделей в ходе перебора различных сочетаний их структур и методов оценивания параметров [15, 33, 41]. Эту задачу усложняет влияние разрядной сетки средства вычислений, «разладки» модели и нерегулярность зависимости критериев идентификации от кода структуры модели [41], что требует полного перебора вариантов.

Для данных [34, п. Н.3] среди пятипараметрических моделей не выше четвёртой степени первоначально оптимальной по минимуму СМПН оказалась функция поправок

$$B_{11011}(t) = -0,4409122 + 1,379538 \cdot 10^{-2}t + 5,303979 \cdot 10^{-6}t^3 - 3,728615 \cdot 10^{-7}t^4,$$

где $\bar{\varepsilon}_{11011}^{[2]} = 2,712152 \cdot 10^{-3} \text{ } ^\circ\text{C}$ при $\Delta_N = 1,5895 \cdot 10^{-3} \text{ } ^\circ\text{C}$; $d_{11011}^{[2]} = 1,843555225 \cdot 10^{-3} \text{ } ^\circ\text{C}$; $\Omega_N = 8,685967747 \cdot 10^{-4} \text{ } ^\circ\text{C}$.

Отклонения от этой модели на интервале $[-0,013088; +0,012183] \text{ } ^\circ\text{C}$ описывает свёртка усечённого распределения Гаусса с равномерным распределением. Однако в связи с выявленными ранее [13] «разладками» модели перепроверка на компьютерах с различными размерами

разрядной сетки и при различных порядках модели максимальной сложности показала, что алгоритмы оптимального перебора кода структуры (детализации) [41] из-за нерегулярности зависимости СМПН от числа параметров модели могут «пропустить» модель оптимальной сложности. Такой «неоптимальной» моделью оказалась $B_{111110000}(t)$ при $\bar{\varepsilon}_{111110000}^{[2]} = 2,8516 \cdot 10^{-3} \text{ } ^\circ\text{C}$, а при проверке по программе «ММК–стат М» – $B_{010011}(t)$ при $\bar{\varepsilon}_{010011}^{[2]} = 3,4298559 \cdot 10^{-3} \text{ } ^\circ\text{C}$. Поэтому при калибровке нужен полный перебор и контроль «разладок» моделей.

Задача калибровки и идентификации дефинициальной неопределённости по данным «образцового» примера [34, п. Н.3] решена несколькими способами с использованием:

ММК для функции преобразования и градуировочной характеристики [13];

МНК линейного регрессионного анализа [42, 43];

ММК как «СКО композиции» [15], упомянутым в книге [44]; графического метода согласно МИ 2916-2005.

Оценённая дефинициальная неопределённость u_D в зависимости от гипотезы о виде распределения поправок превышала 74 % оценки суммарной стандартной неопределённости u_c при «калибровке по GUM» [34, п. Н.3].

Интересно, что расчёт риска принятия ошибочного решения по результатам контрольных испытаний на соответствие после нормализации данных не отличается от его расчёта по данным протокола поверки средства измерений с учётом параметров методики поверки согласно Р 50.2.004-2000 и МИ 2916-2005.

Наиболее существенной информацией о результате калибровки является PDF возможных значений исправленного показания. Здесь и возникает риск указания доверительных границ диаграммы калибровки, не соответствующих нормам доверительной вероятности, толерантным или предикционным интервалам согласно ГОСТ Р ИСО 16269-6-2005 и ГОСТ Р ИСО 16269-8-2005 «Статистические методы. Статистическое представление данных. Определение предикционных интервалов».

Заключение. Отсутствие учёта неадекватности математических моделей объектов испытаний как объектов измерений приводит к занижению вероятностей ложноположительных решений о соответствии или вероятностей ошибок 2-го рода, а отсутствие статистической проверки гипотез о структуре и параметрах этих моделей с целью поиска модели оптимальной сложности – к завышению вероятностей ложноотрицательных решений или ошибок 1-го рода.

Проблему, возникшую с введением ГОСТ ISO/IEC 17025-2019, точно охарактеризовал ещё П. В. Новицкий: «Очень часто доверительные погрешности рассчитывают, вводя ничем не обоснованное предположение о том, что вид закона распределения погрешностей будто бы точно известен. Такой прием является некорректным вне зависимости от того, допускается он сознательно или неосознанно» [45].

«Триединство измерений» (введение на равных правах расчётов «неопределённости, погрешности и прецизионности») в ГОСТ Р 8.820-2013 явно свидетельствует, что проблемы метрологии «новый метрологический менталитет» не

решил, и ситуация с каждым этапом «нововведений» только усложнялась.

Недостатки байесовского анализа как экспертного метода и метода Монте-Карло как метода моделирования заданных распределений, указанные в ГОСТ Р ИСО/МЭК 31010-2011, давно известны, но в явном виде в документах, связанных с расчётами согласно «Руководству по выражению неопределённости измерения», не отмечались.

Таким образом, гораздо большая определённость в расчётах риска имеет место в действующих российских нормативных документах, оперирующих понятиями «достоверность, погрешности измерений и неадекватности». Не говорится об этом и в проекте ревизии GUM. И как можно в этой ситуации вообще доверять расчётам неопределённости измерения, используемым то ли *сознательно*, то ли *неосознанно*. А так как ГОСТ Р ИСО/МЭК 31010-2011 предполагает выбирать методы расчёта риска, то можно использовать методы, описанные в Р 50.2.004-2000, МИ 2916-2005 с моделями ГОСТ 8.009-84 и МИ 1317-2004 по критериям МИ 187-86 [46].

Список литературы

1. Левин С. Ф. Неопределённость в узком и широком смысле результатов поверки средств измерений // Измерительная техника. 2007. № 9. С. 15–19.
2. Левин С. Ф. Погрешности измерений и вычислений как причина «катастрофического феномена 1985–1986 годов» в авиационной и ракетно-космической технике // Контрольно-измерительные приборы и системы. 2000. № 3. С. 21–24.
3. Эльясберг П. Е. Измерительная информация: Сколько её нужно? Как её обрабатывать? М.: Наука, 1983.
4. Хампель Ф. и др. Робастность в статистике: Пер. с англ. М: Мир, 1989.
5. Левин С. Ф. Оптимальная интерполяционная фильтрация статистических характеристик случайных функций в детерминированной версии метода Монте-Карло и закон красного смещения. М.: АН СССР, 1980.
6. Вопросы кибернетики ВК-94. Статистические методы в теории обеспечения эксплуатации. М.: АН СССР, 1982.
7. Ивахненко А. Г. Индуктивный метод самоорганизации моделей сложных систем. Киев: Наукова думка, 1982.
8. Quenouille M. H., *Journal of Royal Statistical Society*, 1949, vol. 11, pp. 68–84.
9. Левин С. Ф. Методические рекомендации. Гарантированность программ обеспечения эксплуатации техники. Киев: Знание, 1989.
10. Левин С. Ф. Метод максимума компактности и комплексные измерительные задачи // Измерительная техника. 1995. № 7. С. 15–21.
11. Левин С. Ф. Метрологическое аттестование и сопровождение программ статистической обработки данных // Измерительная техника. 1991. № 12. С. 16–18.
12. VIM 3. Международный словарь по метрологии. Основные и общие понятия и соответствующие термины, 3-е изд.: Пер. с англ. ВНИИМ им. Д. И. Менделеева, БелГИМ, СПб.: НПО «Профессионал», 2010. URL: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf (дата обращения: 10/01/2020).
13. Левин С. Ф. Легенда о неопределённости // Партнёры и конкуренты. 2001. № 1. С. 13–25.
14. Физическая энциклопедия в 5 т. Т. 3. М.: Большая Российская энциклопедия, 1992.
15. Левин С. Ф. Дефинициальная неопределённость и погрешность неадекватности // Измерительная техника. 2019. № 11. С. 7–17. <https://doi.org/10.32446/0368-1025it.2019-11-7-17>
16. Адлер Ю. П., Грановский Ю. В., Маркова Е. В. Теория эксперимента: прошлое, настоящее, будущее. М.: Знание, 1982.
17. Левин С. Ф. Чего на самом деле должны опасаться ведущие специалисты по внедрению неопределённости в отечественные измерения // Измерительная техника. 2008. № 12. С. 61–64.
18. Исаяев Л. К., Мардин В. В. Русско-англо-французско-испанский словарь основных и общих терминов в метрологии. М.: Изд-во стандартов, 1998.
19. Шеффе Г. Дисперсионный анализ: Пер. с англ. Б. А. Севастьянова и В. П. Чистякова. М.: Государственное изд-во физ.-мат. литературы, 1963.
20. Левин С. Ф. Нерешённые проблемы «прецизионности». Ч. 1 // Главный метролог. 2003. № 3. С. 5–7.
21. Левин С. Ф. Нерешённые проблемы «прецизионности». Ч. 2 // Главный метролог. 2003. № 4. С. 44–54.
22. Левин С. Ф. Нерешённые проблемы «прецизионности». Ч. 3 // Главный метролог. 2004. № 1. С. 44–53.
23. Левин С. Ф. Нерешённые проблемы «прецизионности». Ч. 4 // Главный метролог. 2004. № 3. С. 52–56.
24. Левин С. Ф. Нерешённые проблемы «прецизионности». Ч. 5 // Главный метролог. 2005. № 1. С. 19–26.
25. Левин С. Ф. Нерешённые проблемы «прецизионности». Ч. 6 // Главный метролог. 2005. № 3. С. 20–28.
26. Кокс М., Харрис П. Основные положения Приложения 1 к Руководству по выражению неопределённости измерений // Измерительная техника. 2005. № 4. С. 17–24.
27. Ehrlich C., Dybkaer R., Wöger W., *OIML Bulletin*, April 2007, pp. 23–35.
28. Эрлик Ч., Дибкайер Р., Вёгер В. Эволюция философии и трактовки понятия «измерение» // Главный метролог. 2016. № 1. С. 11–30.
29. Левин С. Ф. Руководство по выражению неопределённости измерения: ревизия – смена парадигмы или новая санкция? // Законодательная и прикладная метрология. 2016. № 5. С. 31–44. <https://doi.org/10.32446/0368-1025it.2016-5-31-44>
30. Левин С. Ф. Метрология: понятия и термины, фразеологизмы и катахрезы. Ч. 1 // Контрольно-измерительные приборы и системы. 2017. № 1. С. 35–38.
31. Левин С. Ф. Метрология: понятия и термины, фразеологизмы и катахрезы. Ч. 2 // Контрольно-измерительные приборы и системы. 2017. № 2. С. 35–38.
32. Левин С. Ф. Концепция неопределённости – Теория погрешностей: Философский спор и математические результаты // Контрольно-измерительные приборы и системы. 2018. № 4. С. 32–36.

33. Левин С. Ф. Руководство по выражению неопределённости измерения: проблемы, нереализованные возможности и ревизия. Ч. 3. Приведение к общему терминологическому знаменателю // Измерительная техника. 2019. № 7. С. 14–22. <https://doi.org/10.32446/0368-1025it.2019-7-14-22>

34. Руководство по выражению неопределённости измерения: Пер. с англ. ВНИИМ им. Д. И. Менделеева, БелГим / Науч. ред. проф. В. А. Слаева. СПб.: ВНИИМ им. Д. И. Менделеева, 1999.

35. Левин С. Ф. Можно ли «точно» пересчитать погрешность в неопределённость? // Законодательная и прикладная метрология. 2017. № 3. С. 18–25.

36. Развитие деятельности по калибровке средств измерений. Доклад рабочей группы. М.: РСПП, 2016.

37. Левин С. Ф. Математическая теория измерительных задач. Ч. 10. Метод совместных измерений // Контрольно-измерительные приборы и системы. 2006. № 3. С. 23–24.

38. Левин С. Ф. Математическая теория измерительных задач. Ч. 10. Метод совместных измерений // Контрольно-измерительные приборы и системы. 2006. № 4. С. 32–36.

39. Левин С. Ф. Математическая теория измерительных задач. Ч. 10. Метод совместных измерений // Контрольно-измерительные приборы и системы. 2006. № 5. С. 33–34.

40. Левин С. Ф. Измерительная задача идентификации функции погрешности // Законодательная и прикладная метрология. 2016. № 4. С. 27–33.

41. Левин С. Ф. Статистические процедуры контроля при высокоточных измерениях // Контрольно-измерительные приборы и системы. 2018. № 3. С. 8–11.

42. Левин С. Ф. Калибровка средства измерений – Три решения одной измерительной задачи // Контрольно-измерительные приборы и системы. 2018. № 1. С. 35–38.

43. Большее Л. Н., Смирнов Н. В. Таблицы математической статистики. Изд. 3-е. М.: Наука, Главная редакция физ.-мат. литературы, 1983.

44. Селиванов М. Н., Фридман А. Э., Кудряшова Ж. Ф. Качество измерений. Л.: Лениздат, 1987.

45. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. Л.: Энергоатомиздат, 1985.

46. Левин С. Ф. Качество поверки средств измерений и апостериорная достоверность контроля // Измерительная техника. 2018. № 9. С. 20–25.

<https://doi.org/10.32446/0368-1025it.2018-9-20-25>

References

1. Levin S. F., *Measurement Techniques*, 2007, vol. 50, no. 9, pp. 921–928. <https://doi.org/10.1007/s11018-007-0173-5>

2. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2000, no. 3, pp. 21–24 (in Russian).

3. Elyasberg P. E., *Izmeritelnaya informatsiya: Skolko ee nuzhno? Kak ee obrabatyvat?* Moscow, Nauka Publ., 1983 (in Russian).

4. Hampel F. R. et al., *Robust Statistics*, New York: John Wiley & Sons, 1986.

5. Levin S. F., *Optimalnaya interpolacionnaya filtratsiya statisticheskikh harakteristik sluchajnykh funktsij v determinirovannoy*

versii metoda Monte-Karlo i zakon krasnogo smeshcheniya, Moscow, AN SSSR Publ., 1980 (in Russian).

6. *Voprosy kibernetiki VK-94. Statisticheskie metody v teorii obespecheniya ekspluatatsii*, Moscow, AN SSSR Publ., 1982.

7. Ivahnenko A. G., *Induktivnyj metod samoorganizatsii modelej slozhnykh system*, Kiev, Naukova dumka Publ., 1982 (in Russian).

8. Quenouille M. H., *Journal of Royal Statistical Society*, 1949, B 11, pp. 68–84.

9. Levin S. F., *Metodicheskie rekomendatsii. Garantirovanost programm obespecheniya ekspluatatsii tekhniki*, Kiev, Znanie Publ., 1989 (in Russian).

10. Levin S. F., *Measurement Techniques*, 1995, vol. 38, no. 7, pp. 732–743. <https://doi.org/10.1007/BF02616256>

11. Levin S. F., *Measurement Techniques*, 1991, vol. 34, no. 12, pp. 1221–1225. <https://doi.org/10.1007/BF00982559>

12. International vocabulary of metrology – Basic and general concepts and associated terms (VIM) 3rd edition, available at: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf (accessed: 10/01/2020).

13. Levin S. F., *Partnery i konkurenty*, 2001, no. 1, pp. 13–25 (in Russian).

14. *Fizicheskaya enciklopediya v 5 tom*. Vol. 3, Moscow, Bol'shaya Rossijskaya enciklopediya Publ., 1992 (in Russian).

15. Levin S. F., *Measurement Techniques*, 2020, vol. 62, no. 11, pp. 933–944. <https://doi.org/10.1007/s11018-020-01716-0>

16. Adler Yu. P., Granovskij Yu. V., Markova E. V., *Teoriya eksperimenta: proshloe, nastoyashchee, budushchee*, Moscow, Znanie Publ., 1982 (in Russian).

17. Levin S. F., *Izmeritelnaya tekhnika*, 2008, no. 12, pp. 61–64 (in Russian).

18. Isaev L. K., Mardin V. V., *Russko-anglo-francuzsko-ispanskij slovar osnovnykh i obshchih terminov v metrologii*, Moscow, Izdatelstvo standartov Publ., 1998 (in Russian).

19. Sheffe H., *The Analysis of Variance*, New York, John Wiley & Sons, 1958.

20. Levin S. F., *Glavnyj metrolog*, 2003, no. 3, pp. 5–7 (in Russian).

21. Levin S. F., *Glavnyj metrolog*, 2003, no. 4, pp. 44–54 (in Russian).

22. Levin S. F., *Glavnyj metrolog*, 2004, no. 1, pp. 44–53 (in Russian).

23. Levin S. F., *Glavnyj metrolog*, 2004, no. 3, pp. 52–56 (in Russian).

24. Levin S. F., *Glavnyj metrolog*, 2005, no. 1, pp. 19–26 (in Russian).

25. Levin S. F., *Glavnyj metrolog*, 2005, no. 3, pp. 20–28 (in Russian).

26. Koks M., Harris P., *Measurement Techniques*, 2005, vol. 48, no. 4, pp. 336–345. <https://doi.org/10.1007/s11018-005-0146-5>

27. Ehrlich C., Dybkaer R., Wöger W., *OIML Bulletin*, April 2007, pp. 23–35.

28. Erlik Ch., Dibkajer R., Vyoger V., *Glavnyj metrolog*, 2016, no. 1, pp. 11–30 (in Russian).

29. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2017, no. 1, pp. 35–38 (in Russian).

30. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2017, no. 2, pp. 35–38 (in Russian).
31. Levin S. F., *Zakonodatelnaya i prikladnaya metrologiya*, 2016, no. 5, pp. 31–44 (in Russian).
32. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2018, no. 4, pp. 32–36 (in Russian).
33. Levin S. F., *Izmeritel'naya tekhnika*, 2019, no. 7, pp. 14–22 (in Russian).
34. Guide to the Expression of Uncertainty in Measurement (GUM). Sec. Ed. Geneva, BIMP, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 1995.
35. Levin S.F., *Zakonodatelnaya i prikladnaya metrologiya*, 2017, no. 3, pp. 18–25 (in Russian).
36. *Razvitie deyatel'nosti po kalibrovke sredstv izmerenij. Doklad rabochej gruppy*, Moscow, RSPP Publ., 2016 (in Russian)
37. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2006, no. 3, pp. 23–24 (in Russian).
38. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2006, no. 4, pp. 32–36 (in Russian).
39. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2006, no. 5, pp. 33–34 (in Russian).
40. Levin S. F., *Zakonodatelnaya i prikladnaya metrologiya*, 2016, no. 4, pp. 27–33 (in Russian).
41. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2018, no. 3, pp. 8–11 (in Russian).
42. Levin S. F., *Kontrolno-izmeritelnye pribory i sistemy*, 2018, no. 1, pp. 35–38 (in Russian).
43. Bol'shev L. N., Smirnov N. V., *Tablicy matematicheskoy statistiki. Izd. 3*, Moscow, Nauka Publ., 1983 (in Russian).
44. Selivanov M. N., Fridman A. E., Kudryashova Zh. F., *Kachestvo izmerenij*, Leningrad, Lenizdat Publ., 1987 (in Russian).
45. Novickij P. V., Zograf I. A., *Ocenka pogreshnostej rezul'tatov izmerenij*, Leningrad, Energoatomizdat Publ., 1985 (in Russian).
46. Levin S. F. *Measurement Techniques*, 2018, vol. 61, no. 9, pp. 863–871. <https://doi.org/10.1007/s11018-018-1516-0>

Дата поступления: 10.10.2019.

Дата принятия: 09.03.2020.

Inadequacy for mathematical models of measurement objects and risk calculations according to ISO/IEC 17025-2019

Sergey F. Levin

Moscow Institute for Expertise and Tests, Moscow, Russian Federation, e-mail: AntoninaEL@rostest.ru

The problem of inadequacy of mathematical models of measurement objects is considered in connection with the problem of “definitional uncertainty of measurement” and the need for risk management in accordance with GOST ISO/IEC 17025-2019 “General requirements for the competence of testing and calibration laboratories”. The prehistory of the problem is described: from the formation of moment and compositional approaches to estimating accuracy to the introduction of a special term “inadequacy error for mathematical model of measurement object”. The negative impact of hopelessness of conceptual and terminological transformations in metrology and critical contradiction of the applicability estimates for “statistical methods” of GOST R ISO/IEC 31010-2011 “Risk management – Risk assessment techniques” and “Guide to the expression of uncertainty in measurement” is noted. It is shown that taking into account the inadequacy of probabilistic models in risk calculations is a necessary condition for results reliability.

Key words: mathematical model, definitional uncertainty, moment approach, compositional approach, inadequacy of model error.