Measurement 88 (2016) 194–201

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Removing divergence of JCGM documents from the GUM (1993) and repairing other defects

Raghu N. Kacker [a,*], Rüdiger Kessel [b], James F. Lawrence [a,c]

[a] National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
[b] Physikalisch-Technische Bundesanstalt, Braunschweig D-38116, Germany
[c] George Mason University, Fairfax, VA 22030, USA

ARTICLE INFO

Article history:
Received 11 December 2015
Received in revised form 21 March 2016
Accepted 25 March 2016
Available online 29 March 2016

Keywords:
Bayesian statistics
Coverage interval
GUM-S1
Uncertainty in measurement
VIM3

ABSTRACT

The mission of the Joint Committee for Guides in Metrology (JCGM) is to maintain and promote the use of the Guide to the Expression of Uncertainty in Measurement (GUM) and the International Vocabulary of Metrology (VIM, second edition). The JCGM has produced the third edition of the VIM (referred to as VIM3) and a number of documents; some of which are referred to as supplements to the GUM. We are concerned with the Supplement 1 (GUM-S1) and the document JCGM 104. The signal contribution of the GUM is its operational view of the uncertainty in measurement (as a parameter that characterizes the dispersion of the values that could be attributed to an unknown quantity). The operational view promulgated by the GUM had disconnected the uncertainty in measurement from the unknowable quantities true value and error. The GUM-S1 has diverged from the operational view of the uncertainty in measurement. Either the disparities should be removed or the GUM-S1 should not be referred to as a supplement to the GUM. Also, the GUM-S1 has misinterpreted the Bayesian concept of a statistical parameter and the VIM3 definitions of coverage interval and coverage probability are mathematically defective. We offer practical suggestions for revising the GUM-S1 and the VIM3 to remove their divergence from the GUM and to repair their defects.

Published by Elsevier Ltd.

1. Introduction

The Joint Committee for Guides in Metrology (JCGM) was formed in 1997 [1] to maintain and promote the use of the Guide to the Expression of Uncertainty in Measurement (GUM, 1993) [2] and the International Vocabulary of Metrology (VIM, 1993, second edition, now referred to as VIM2) [3]. The JCGM has two working groups. The JCGM working group 1 (JCGM WG1) has produced a number of documents; some of which are referred to as supplements to the GUM. We are concerned with the supplement 1, the GUM-S1 published in 2008 and the document JCGM 104 published in 2009. The GUM-S1 is entitled Supplement 1 to the GUM – Propagation of distributions using a Monte Carlo (MC) method [4] and the JCGM 104 is entitled Introduction to the GUM and related documents [5]. The concept and definitions in the JCGM 104 apply to the GUM-S1. The JCGM working group 2 (JCGM WG2) published in 2008 the third edition of the VIM, identified as VIM3 or JCGM 200 [6].

The GUM is not completely consistent with either conventional or Bayesian statistical concepts [7]. However, the GUM can be made fully consistent with Bayesian concepts by using for the Type A (statistical) evaluations Bayesian statistics (with non-informative prior distributions) [8]. Then the GUM concept (from conventional statistics) of

* Corresponding author.
E-mail addresses: raghu.kacker@nist.gov (R.N. Kacker), ruediger.kessel@ptb.de (R. Kessel), lawrence@gmu.edu (J.F. Lawrence).

http://dx.doi.org/10.1016/j.measurement.2016.03.057
0263-2241/Published by Elsevier Ltd.

quantifying uncertainty of the uncertainty in measurement by 'degrees of freedom' would vanish. An introductory reference for conventional statistical inference is [9]. Bayesian inference based on the use of Bayes' rule (theorem) is described in many textbooks and articles such as the following [10–21]. In this paper, phrases displayed in the *italic* font are direct quotes from a cited reference; sometimes additional words are inserted in parentheses to clarify the intended meaning.

To a statistician, the term Bayesian implies the use of Bayes' rule. However, metrologists use the term Bayesian in a broad sense meaning only that the state of knowledge about an unknown quantity is expressed by a subjective (personal degree of belief) probability distribution. The GUM-S1 does not use the Bayes' rule, yet it considers the input and output probability distributions of its Monte Carlo method to be Bayesian. For example, the GUM-S1 makes the following statements: *(i) A coverage interval is sometimes known as a credible interval or Bayesian interval, (ii) the use of probability distributions in Type B evaluation is a feature of Bayesian inference, (iii) in the Bayesian context of this Supplement, concepts such as the reliability, or the uncertainty, of an uncertainty are not necessary* [4, Sections 3.12, 5.1.2, 6.4.9.4 Note 2]. We use the term Bayesian in the broad sense of metrologists, unless stated otherwise.

The GUM-S1 is aligned with Bayesian concepts. The GUM-S1 identifies the true value of an unknown quantity (measurand) as a statistical parameter and regards that parameter as a random variable [4, Section 5.1.1, c], [5, Section 3.17]. A probability distribution for this random variable describes the probabilities of the unknown true value lying in different intervals [5, Section 3.17]. The GUM-S1 introduced an expression of uncertainty in measurement called a coverage interval. A coverage interval is defined to be an *interval containing the value of a quantity with a stated probability, based on the information available* [4, Section 3.12], and its associated coverage probability is the *probability that the value of a quantity is contained within a specified coverage interval.* In these GUM-S1 definitions, 'the value of a quantity' refers to 'the true value' even though the adjective 'true' is suppressed to create an appearance of fealty to the GUM [4, Section 3.12 Note 4]. In the GUM-S1, the true value of a quantity is assumed to be essentially unique [4, Section 1], [6, Section 2.11 Note 3], [22, Section 2.5]. A real physical quantity involved in measurement has a set of multiple true values rather than a single true value, because of the inherently incomplete amount of detail to which the quantity can be specified [2, Annex D including Figures D.1 and D.2], [6, Section 2.11 Note 1], [22, Section 2.4]. Fundamental constants of nature are exceptions. So the VIM3 gives the following general definitions: a coverage interval is an *interval containing the set (range) of true quantity values of a measurand with a stated probability, based on the information available* [6, Section 2.36], and the coverage probability is the *probability that the set (range) of true quantity values of a measurand is contained within a specified coverage interval* [6, Section 2.37]. Uncertainty arising from the finite amount of detail in the definition of a quantity is called definitional uncertainty [6, Section 2.27]. A quantity is considered to have an essentially unique true value when the definitional uncertainty is believed to be negligible [6, Section 2.11 Note 3].

The GUM-S1 has misinterpreted the Bayesian concept of a statistical parameter and the VIM3 definitions of coverage interval and coverage probability are mathematically defective. These defects should be repaired. The GUM-S1 concepts of coverage interval and coverage probability do not agree with the operational view of the uncertainty in measurement promulgated by the GUM. Since the GUM-S1 is a supplement to the GUM it should be revised to agree with the GUM.

In Section 2, we describe the Bayesian concepts of a statistical parameter, a probability distribution (expressing the state of knowledge about the value of that parameter), and an interval estimate for that value. In Section 3, we review various concepts and terms introduced by the GUM and identify the operational view of the uncertainty in measurement as the signal contribution of the GUM. In Section 4, we show that the GUM-S1 misinterprets the Bayesian concept of a statistical parameter, and in Section 5, we show that the VIM3 definitions of coverage interval and coverage probability are mathematically defective. In Section 6, we demonstrate that the GUM-S1 concepts of coverage interval and coverage probability do not agree with the GUM. In Section 7 we offer suggestions for revising the GUM-S1 to remove its divergence from the GUM. Summary and concluding remarks appear in Section 8.

2. Bayesian concepts of parameter, probability distribution, and interval estimate

In this section we are concerned with statistical inference for the unknown value of a statistical parameter based on the Bayes' rule. It is often said that in conventional statistical inference a statistical parameter has a fixed value but in Bayesian statistical inference the statistical parameter is treated as a random variable with a probability distribution which describes the possible variation of that parameter [9, Section 7.2.3]. This is a widespread misinterpretation of the Bayesian concept of a statistical parameter. In Bayesian inference also, the value of a parameter is fixed. That fixed value is the target of statistical inference. What changes is a probability distribution (over the possible values for that parameter) expressing the state of knowledge about that fixed value.

Statistician Dennis Lindley was a leading expert and advocate of Bayesian inference. We quote Lindley from the Ref. [15, p. 301]: "*The parameter is also uncertain. Indeed, it is that uncertainty that is the statistician's main concern. The recipe says that it also should be described by a probability .... In so doing we depart from the conventional attitude. It is often said that the parameters are assumed to be random quantities. This is not so. It is the axioms that are assumed, from which the randomness property is deduced.*"

Physicist E.T. Jaynes was also a leading expert and advocate of Bayesian inference. We quote Jaynes from the Ref. [12, p. 11]: "*For decades Bayesians have been accused of "supposing that an unknown parameter is a random variable"; and we have denied hundreds of times,*