Measurement 45 (2012) 2194–2202

Contents lists available at SciVerse ScienceDirect

Measurement

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

A Bayesian paradox and its impact on the GUM approach to uncertainty

Filippo Attivissimo, Nicola Giaquinto *, Mario Savino

Department of Electrics and Electronics, Politecnico di Bari, Via E. Orabona, 70125 Bari, Italy

ARTICLE INFO

Article history:
Available online 3 February 2012

Keywords:
GUM
Uncertainty
Estimation theory
Bayesian estimation
Stein's paradox
Noise cancellation

ABSTRACT

The paper presents ideas and observations about the use of the frequentist and the Bayesian approach to estimation and uncertainty. The merits and the pitfalls of the Bayesian approach, compared with the frequentist one, are illustrated using a simple example, which gives rise to an instructive paradox. The impact of the paradox on the GUM approach to uncertainty prescribed in Supplement 1 is highlighted and discussed.

1. Introduction

The "Guide to the Expression of Uncertainty in Measurement" (GUM), maintained by the Joint Committee for Guides in Metrology (JCGM) of BIPM, is currently made of three separate documents: an Introduction [1], the core document [2], and the Supplement 1 [3] (other Supplements are under preparation).

The GUM core document, denoted hereafter JCGM-100, dates back to 1993 and has remained substantially unchanged through years. JCGM-100 does not prescribe explicitly a Bayesian approach to measurement uncertainty, even if the very same definition of uncertainty ("dispersion of the values that could reasonably be attributed to the measurand") seems to suggest a Bayesian view. Other parts of the document are implicitly frequentist, e.g. when an "uncertainty of the uncertainty" is admitted and, consequently, effective degrees of freedom are attached to the uncertainty. Some statisticians have explicitly judged the GUM a frequentist–Bayesian hybrid [4]. In particular, having an uncertainty on the knowledge of a probability density function (pdf), which is a model of incomplete knowledge in itself, has been judged a serious internal inconsistency [5], and has caused various suggestions to take the Bayesian viewpoint [6–8]. Also the latest edition of the International Vocabulary of Metrology (VIM) [9] has pushed in this direction, since it is clearly Bayesian in many points, for example in its definition of expanded uncertainty (in terms of "coverage intervals", and *not* confidence intervals).

The Bayesian revision of the GUM is in Supplement 1, denoted hereafter JCGM-101. This document is very clear-cut in prescribing Bayesian definitions and procedures for measurement uncertainty. The incomplete knowledge of quantities is always described by exactly known pdfs, and rules are given to establish the proper pdf for typical cases of incomplete information. Another feature of JCGM-101 is the accurate evaluation, via Monte Carlo (MC) method, of both the best estimate and the uncertainty of a quantity, even when first-order approximations are not acceptable, and when Central Limit Theorem is not applicable (the usual assumptions in JCGM-100).

The theoretical framework of JCGM-101 has been analysed in many papers, and some issues have been raised. For example, in [10] and in [11], JCGM-101 computations are compared with alternative ones made under the Bayesian paradigm, and discrepancies between the results are analysed. Without going deep into the content of these papers, it must be said that the discrepancies are due – as

* Corresponding author.
E-mail addresses: attivissimo@misure.poliba.it (F. Attivissimo), giaquinto@misure.poliba.it (N. Giaquinto), savino@misure.poliba.it (M. Savino).

0263-2241/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.measurement.2012.01.022

usual with Bayesian inference – to explicit or implicit differences in the choice of the prior distribution. In particular, [10] mentions the marginalization paradox and the possible use of techniques for assigning priors that, although well-known among statisticians, are different from the set of rules stated in JCGM-101.

Paradoxes in Bayesian statistics are notorious, and hardly considered a fatal problem; on the contrary, historically they have stimulated new findings and refinements of the theory [12]. The Bayesian approach of the GUM has stimulated researchers in the measurement field to extend the scope of their analyses, developing Bayesian estimators, and comparing them with previously developed frequentist estimators [13,14]. Therefore, in the authors' opinion, there is no point in criticising Bayesianism on general terms (unless one wants to take part in the endless Bayesian–frequentist dispute).

A more practical question is whether some of the known Bayesian paradoxes can actually lead to unacceptable results, in a concrete measurement situation, following the normal JCGM-101 methodology. A related, and more general question, is whether it is really convenient to exclude completely frequentist methodologies, which have obviously a very long tradition, and are customarily used in countless practical estimation problems.

The paper develops ideas illustrated in [15] and presents a simple example of measurement made according to the Bayesian and the frequentist paradigm, leading to an instructive paradox. Although the paradox is already known [16,17], its actual importance for a GUM user has never been detected, and is not easy to recognise, as demonstrated by an actual example in JCGM-101, in which the paradox affects the result. The paper provides a detailed analysis in order to clarify the issue, and to show the merits and the limitations of both the Bayesian approach (hereafter, BA) and the frequentist approach (hereafter, FA) to estimation and uncertainty.

The paper is organised as follows. Section 2 outlines the basic measurement situation examined, and the computation methodologies employed throughout the paper. Section 3 illustrates the first part of the example, in which the BA obtains the same results of the FA, but using a far more convenient, and even elegant, theoretical framework. Section 4 illustrates the second part of the example, in which BA and FA leads to opposite results. In Section 5 the different conclusions of the Bayesian and the frequentist methodology are analysed and commented. The paper is closed by final considerations (Section 6) and conclusions.

2. Basics: proposed example and methods of computations

Fig. 1 depicts a simple and idealised, yet not unrealistic, measurement situation.

The voltmeter yields the single measurement $Y = X + N$, where X is the voltage of a battery and N is additive Gaussian thermal noise, with zero mean and known variance σ^2 (the assumption of known variance is common and not unrealistic, as it can be computed from the resistor value, the absolute temperature, and the equivalent noise bandwidth of the system). The voltage X is known to be in the interval $[-a; a]$, which is "large" with respect to the noise perturbing the measurement, i.e. $a \gg \sigma$. No other prior information is available about X. The measured voltage is $Y = y_0$. The aim is to give the best estimate, and the associated uncertainty, of two quantities:

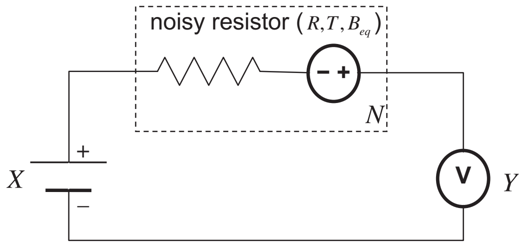

Fig. 1. Measurement of a voltage with additive Gaussian noise.

(1) the voltage X;
(2) the power $W = X^2$, transferred by the voltage source on a unit-value resistor.

Although apparently simple and straightforward, a full analysis of the problem with the Bayesian and the frequentist approach requires the computation of a number of pdfs, together with the associated expected values and variances. The complete list is:

(1) $f_X(x)$, (prior) pdf of the variable X;
(2) $f_W(w)$, (prior) pdf of the variable $W = X^2$;
(3) $f_{Y|X=x_0}(y)$, pdf of the variable Y for a given value of the measurand, $X = x_0$;
(4) $f_{Z|X=x_0}(z)$, pdf of the variable $Z = Y^2$ for a given value of the measurand, $X = x_0$;
(5) $f_{X|Y=y_0}(x)$, (posterior) pdf of the variable X for a given value of the measurement, $Y = y_0$;
(6) $f_{W|Y=y_0}(w)$, (posterior) pdf of the variable $W = X^2$ for a given value of the measurement, $Y = y_0$.

In the following, all the listed pdfs, and the related expected values and standard deviations, are computed in three ways:

(a) analytically, whenever possible, by making proper assumptions and approximations if needed;
(b) numerically, by evaluation of exact formulae (e.g. involving integrals, etc.);
(c) by MC simulations.

For the numeric and MC evaluations the following values are used: $[-a; a] = [-10; +10]$ mV, $\sigma = 1$ mV, $y_0 = 2.5$ mV. The numeric evaluations of distributions are made in $N = 10^5$ points, equispaced in a suitable interval, i.e. $[-a - 4\sigma; a + 4\sigma] = [-14; +14]$ for the variables X and Y, and $[0; a^2 + 4\sigma] = [0; 104]$ for the variables $W = X^2$ and $Z = Y^2$. Numeric values of expectations and standard deviations, as derived on the basis of the numerically evaluated distributions (point b), are reported with seven significant digits, in order to show the very small differences with