IOP PUBLISHING METROLOGIA
Metrologia **44** (2007) L57–L61 doi:10.1088/0026-1394/44/6/N03

SHORT COMMUNICATION

Comments on 'Bayesian evaluation of comparison data'

Raghu Kacker

National Institute of Standards and Technology, Gaithersburg, MD 20899-8910, USA

E-mail: raghu.kacker@nist.gov

Received 2 February 2007
Published 11 October 2007
Online at stacks.iop.org/Met/44/L57

Abstract
A recent paper by Ignacio Lira in this journal (2006 *Metrologia* **43** S231–4) proposes a new procedure to evaluate the data from a simple key comparison where a travelling standard of a stable value during the comparison is independently measured by the participants. In particular, Lira presents an expression, which he claims to be the kernel of a Bayesian posterior probability density function for the value of the measurand. Lira claims that his expression encodes the collective state of knowledge whether or not the results are consistent. Thus Lira offers an alternative to the method recommended by an Advisory Group on Uncertainties commissioned by the BIPM. We discuss Lira's procedure from the viewpoint of what we believe to be a good practice of Bayesian statistics.

1. Introduction

Statistical analysis of the results from a key comparison is an important problem of interest to the national metrology institutes. The outputs of a key comparison include a reference value, unilateral and bilateral degrees of equivalence and their associated uncertainties. Lira [1] addresses a simple key comparison where a travelling standard of a stable value during the comparison is independently measured by the participants. This problem has previously been addressed in Cox [2] by an Advisory Group on Uncertainties commissioned by the BIPM. Cox [2] recommends two procedures: (a) for consistent results and (b) for inconsistent results. Lira [1, section 1, paragraph 2] states that his paper develops a Bayesian procedure, which yields an analytic probability density function (pdf) that encodes the collective state of knowledge, and it applies irrespective of whether or not the data are consistent; then from his pdf, an estimated value for the measurand together with its associated uncertainty and the degrees of equivalence may be determined. Thus Lira [1] offers an alternative to Cox [2]. We discuss the procedure of Lira [1] from the viewpoint of what we believe to be a good practice of Bayesian statistics.

1.1. Notation

In this paper we use symbols different from those used by Lira [1]. The correspondence between the two sets of symbols is given in table 1. The symbols we use make it easier to distinguish quantities that have constant values from those that are random variables. We use upright lower-case Greek symbols for statistical and metrological parameters of supposedly constant values (that is, essentially unique and stable values), such as $\mathrm{\mu}_1, \ldots, \mathrm{\mu}_n$, and ξ. The symbol ξ is a metrological parameter; it represents the value of the measurand. The symbols $\mathrm{\mu}_1, \ldots, \mathrm{\mu}_n$ are statistical parameters; they represent expected values of the sampling pdfs of the results of measurement from the individual laboratories. The metrological and statistical parameters have essentially identical interpretations in conventional statistics, Bayesian statistics [3] and the *Guide to the Expression of Uncertainty in Measurement (GUM)* [4]. We use upright lower-case Latin symbols for the actual results (data), $\mathrm{x}_1, \ldots, \mathrm{x}_n$, submitted by the individual laboratories. In both conventional statistics and Bayesian statistics, the data $\mathrm{x}_1, \ldots, \mathrm{x}_n$ are regarded as known realizations of random variables having sampling pdfs. The random variables of which the data $\mathrm{x}_1, \ldots, \mathrm{x}_n$ are regarded as realization are denoted by the italic lower-case Latin symbols $x_1, \ldots, x_n$. The random variables $x_1, \ldots, x_n$ are also referred to as results. The expected values of the sampling pdfs of $x_1, \ldots, x_n$ are $\mathrm{\mu}_1, \ldots, \mathrm{\mu}_n$, respectively. The variances of the sampling pdfs of $x_1, \ldots, x_n$ are denoted by $\sigma_1^2, \ldots, \sigma_n^2$, respectively. To our understanding, Lira's paper [1] is based on the assumption that the variances $\sigma_1^2, \ldots, \sigma_n^2$ of the sampling

Table 1. Correspondence between the symbols used by Lira in [1] and in this paper.

	Lira [1]	This paper
Constant value of the common measurand	X	ξ
Constant realized result from laboratory i	x_{ei}	x_i
Constant uncertainty submitted by laboratory i	u_i	$\mathrm{u}(\mathrm{x}_i)$
Constant expected value of result from laboratory i	X_i	μ_i
Variable with state-of-knowledge pdf concerning μ_i	x_i	X_i
Variable with state-of-knowledge pdf concerning ξ	x	Y

pdfs of $x_1, \ldots, x_n$ are *known* and that $\sigma_1^2, \ldots, \sigma_n^2$ are equal to the squares $\mathrm{u}^2(\mathrm{x}_1), \ldots, \mathrm{u}^2(\mathrm{x}_n)$ of the standard uncertainties $\mathrm{u}(\mathrm{x}_1), \ldots, \mathrm{u}(\mathrm{x}_n)$ submitted by the individual laboratories. Following the notation used in the GUM, we use the italic upper-case Latin symbol X_i for a random variable having a state-of-knowledge pdf concerning the expected value μ_i based on the result x_i and uncertainty $\mathrm{u}(\mathrm{x}_i)$ from the laboratory labelled i, for $i = 1, 2, \ldots, n$. Similarly we use the symbol Y for a random variable having a state-of-knowledge pdf concerning the value ξ of the common measurand. To sum up, we use upright symbols for constants and italic symbols for random variables.

1.2. Assumptions

To our understanding, Lira's paper [1] is based on the following assumptions.

Assumption 1. A measurand of an unknown essentially unique and stable value ξ is measured by n laboratories. The laboratories submit the following paired results and standard uncertainties $(\mathrm{x}_1, \mathrm{u}(\mathrm{x}_1)), \ldots, (\mathrm{x}_n, \mathrm{u}(\mathrm{x}_n))$ for the common value ξ of the measurand [1, section 1, line 3].

Assumption 2. For $i = 1, 2, \ldots, n$, the submitted result x_i is a realization of a random variable x_i. The sampling pdf of x_i is normal (Gaussian) with *unknown* expected value μ_i and *known* variance σ_i^2. The known variance σ_i^2 is equal to $\mathrm{u}^2(\mathrm{x}_i)$, for $i = 1, 2, \ldots, n$. The sampling pdfs of $x_1, \ldots, x_n$ are mutually independent. (This assumption is not explicitly stated in Lira [1, section 2] but it is implied and corroborated by [5, 6] cited in that section.)

Assumption 3. The difference $\mu_i - \xi$ is the bias (offset, systematic error) in the result x_i, for $i = 1, \ldots, n$. The random error in x_i is $\mathrm{x}_i - \mu_i$ and the total error is $\mathrm{x}_i - \xi = (\mathrm{x}_i - \mu_i) + (\mu_i - \xi)$. All biases (offsets) $\mu_i - \xi$ are assumed to be zero; that is, $\mu_1 = \ldots = \mu_n = \xi$ [1, section 2.1, line 1].

Assumption 4. The state-of-knowledge pdf of each X_i is normal with expected value x_i and standard deviation $\mathrm{u}(\mathrm{x}_i)$, for $i = 1, 2, \ldots, n$ [1, section 1, line 8]. The state-of-knowledge pdfs of $X_1, \ldots, X_n$ are mutually independent [1, section 1, line 6].

Assumption 5. The number n of laboratories is three or more [1, p S232, column 1, line 6].

Note 1. To avoid overload of symbols, Bayesian statisticians declare the expected values $\mu_1, \ldots, \mu_n$ as random variables with state-of-knowledge pdfs rather than introduce new symbols for the variables. For clarity, in this paper, we have used the symbols $\mu_1, \ldots, \mu_n$ for constants and $X_1, \ldots, X_n$ for random variables with state-of-knowledge pdfs concerning $\mu_1, \ldots, \mu_n$.

Note 2. Assumptions 2 and 4 are not equivalent. Assumption 2 is about the *sampling probability distributions* of the data $x_1, \ldots, x_n$ conditional on the fixed (constant) values of the statistical parameters $(\mu_1, \sigma_1^2), \ldots, (\mu_n, \sigma_n^2)$, where $\sigma_i^2 = \mathrm{u}^2(\mathrm{x}_i)$, for $i = 1, 2, \ldots, n$. A sampling distribution is a property of the measurement procedure (i.e. the data generation process). A sampling distribution is required in both conventional statistics and Bayesian statistics to define the likelihood function of the parameters conditional on the realized data. A sampling distribution has at least one unknown parameter which is to be estimated. For example, under assumption 2 and the subsequent assumption 3, the expected value of each x_i, for $i = 1, 2, \ldots, n$, is ξ which is to be estimated.

Assumption 4 is about the *state-of-knowledge probability distributions* of $X_1, \ldots, X_n$ concerning $\mu_1, \ldots, \mu_n$ conditional on the fixed (constant) values of the realized data $\mathrm{x}_1, \ldots, \mathrm{x}_n$. A state-of-knowledge distribution represents belief probabilities about the possible values of a parameter based on all available information. In Bayesian statistics, a prior distribution represents the state of knowledge about the value of a parameter before measurement data are seen and a posterior distribution represents an updated state of knowledge about the value of that parameter in view of the measurement data. All parameters of a state-of-knowledge probability distribution are fully specified, possibly through a hierarchy of probability distributions. For example, under assumption 4, the probability distribution of each X_i, for $i = 1, 2, \ldots, n$, is a fully specified normal distribution with expected value x_i and variance $\sigma_i^2 = \mathrm{u}^2(\mathrm{x}_i)$.

Assumption 4 does not imply assumption 2. Assumption 2 provides a likelihood function of $X_1, \ldots, X_n$ concerning $\mu_1, \ldots, \mu_n$ conditional on the realized data $\mathrm{x}_1, \ldots, \mathrm{x}_n$. Then with independent non-informative improper prior distributions for $X_1, \ldots, X_n$, it can be shown, using Bayes's theorem, that the Bayesian posterior distributions for $X_1, \ldots, X_n$ conditional on the realized data $\mathrm{x}_1, \ldots, \mathrm{x}_n$ are independent and normal with expected values $\mathrm{x}_1, \ldots, \mathrm{x}_n$ and variances $\sigma_1^2, \ldots, \sigma_n^2$, respectively [3]. Since Bayesian posterior distributions are state-of-knowledge pdfs, one may say that assumption 4 follows from assumption 2. In this sense assumption 4 is redundant. If the prior distributions for $X_1, \ldots, X_n$ were proper probability distributions, then assumption 4 would not follow from assumption 2.

1.3. Review of Lira's procedure

Lira's [1] procedure for determining a state-of-knowledge pdf for Y, based on assumptions 1 through 5, is as follows.

Step 1. Lira [1, section 2.1, line 7] states that for $i = 1, \ldots, n$, 'With non-informative prior pdfs for all quantities involved, Bayes's theorem [5] gives their joint pdf after all measurements have been carried out as

$$f(X_1, \ldots, X_n, Y|\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}) \propto \prod_{i=1}^{n} L_i(Y, X_i|\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i))\delta(Y - X_i), \tag{1}$$

where δ is Dirac's delta function and $L_i(Y, X_i|\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i))$ is the likelihood of obtaining the data x_i, $\mathrm{u}(\mathrm{x}_i)$ given that $\xi = Y$ and $\mu_i = X_i$.' The symbol $\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}$ represents the sequence $(\mathrm{x}_1, \mathrm{u}(\mathrm{x}_1)), \ldots, (\mathrm{x}_n, \mathrm{u}(\mathrm{x}_n))$.

Step 2. Lira [1, section 2.1] determines the likelihood function $L_i(Y, X_i|\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i))$ used in (1) from assumption 4, which states that the state-of-knowledge pdfs for $X_1, \ldots, X_n$ are independent and the pdf of X_i is normal with expected value x_i and variance $\mathrm{u}^2(\mathrm{x}_i)$. That is,

$$f_i(X_i|\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)) \propto \exp\left[-\frac{1}{2}\frac{(X_i - \mathrm{x}_i)^2}{\mathrm{u}^2(\mathrm{x}_i)}\right] \times [\mathrm{u}^2(\mathrm{x}_i)]^{-1/2}, \tag{2}$$

for $i = 1, \ldots, n$. The term $[\mathrm{u}^2(\mathrm{x}_i)]^{-1/2}$ is included in (2) because Lira [1, section 2.1] replaces the known constant variance $\mathrm{u}^2(\mathrm{x}_i)$ in (2) by the random variable $\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2$ of unknown pdf to obtain the following expression for the likelihood function:

$$L_i(Y, X_i|\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)) \propto \exp\left[-\frac{1}{2}\frac{(X_i - \mathrm{x}_i)^2}{\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2}\right] \times [\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2]^{-1/2}, \tag{3}$$

for $i = 1, \ldots, n$.

Step 3. Lira [1, section 2.1] substitutes (3) in (1), which amounts to multiplying the n expressions given in (3) for $i = 1, \ldots, n$ and then replacing each of the random variables $X_1, \ldots, X_n$ with the same random variable Y to obtain the following expression:

$$f(Y|\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\frac{(Y - \mathrm{x}_i)^2}{\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2}\right] \times \prod_{i=1}^{n} [\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2]^{-1/2}. \tag{4}$$

Lira [1] claims that when the number n is three or more, expression (4) can be normalized to form a pdf for Y. Expression (4) cannot be normalized when n is less than three. Further, Lira refers to expression (4) as the kernel of a posterior pdf for Y. Lira [1, abstract] refers to this procedure for obtaining (4) as a Bayesian procedure.

In section 2, we show that if assumptions 1, 2 and 3 could reasonably be attributed to the data $\mathrm{x}_1, \ldots, \mathrm{x}_n$ then a Bayesian posterior pdf for Y is normal with expected value $\mathrm{x_W}$ and variance $\mathrm{u}^2(\mathrm{x_W})$, where $\mathrm{x_W} = \Sigma \mathrm{w}_i \mathrm{x}_i / \Sigma \mathrm{w}_i$ is the weighted mean, $\mathrm{w}_i = 1/\mathrm{u}^2(\mathrm{x}_i)$ for $i = 1, 2, \ldots, n$, and $\mathrm{u}(\mathrm{x_W}) = 1/\sqrt{\Sigma \mathrm{w}_i}$. In section 3, we comment on Lira's procedure. Concluding remarks are given in section 4.

2. A Bayesian posterior pdf for Y conditional on $\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}$

Assumptions 1, 2 and 3 imply that the only unknown parameter is the value ξ of the measurand. The problem is to determine a pdf for the random variable Y representing the state of knowledge concerning ξ. The sampling pdf of x_i conditional on its unknown expected value ξ is

$$f(x_i|\xi) \propto \exp\left[-\frac{1}{2}\frac{(x_i - \xi)^2}{\mathrm{u}^2(\mathrm{x}_i)}\right], \tag{5}$$

for $i = 1, \ldots, n$, and the joint pdf of $x_1, \ldots, x_n$ conditional on their common expected value ξ is

$$f(x_1, \ldots, x_n|\xi) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i - \xi)^2}{\mathrm{u}^2(\mathrm{x}_i)}\right]. \tag{6}$$

In Bayesian statistics, the likelihood function of the random variable Y, representing the state of knowledge concerning ξ, given the data $\mathrm{x}_1, \ldots, \mathrm{x}_n$ is the joint pdf (6) interpreted as a function of ξ [3]. Thus

$$L(Y|\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\frac{(Y - \mathrm{x}_i)^2}{\mathrm{u}^2(\mathrm{x}_i)}\right]. \tag{7}$$

In Bayesian statistics, the likelihood function (7) is used to update via Bayes's theorem [3] a prior pdf for Y to determine a posterior pdf for Y. The likelihood function can be multiplied by any function of the data (which are known quantities) because it gets cancelled out when the constant of proportionality is determined. Let us define $\mathrm{x_W} = \Sigma \mathrm{w}_i \mathrm{x}_i / \Sigma \mathrm{w}_i$, where $\mathrm{w}_i = 1/\mathrm{u}^2(\mathrm{x}_i)$ for $i = 1, 2, \ldots, n$, and $\mathrm{u}(\mathrm{x_W}) = 1/\sqrt{\Sigma \mathrm{w}_i}$. If we multiply (7) by the function $\exp[-\frac{1}{2}(\Sigma \mathrm{x}_i^2/\mathrm{u}^2(\mathrm{x}_i) - \mathrm{x_W^2}/\mathrm{u}^2(\mathrm{x_W}))]$ of the data $\mathrm{x}_1, \ldots, \mathrm{x}_n$, we get

$$L(Y|\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}) \propto \exp\left[-\frac{1}{2}\frac{(Y - \mathrm{x_W})^2}{\mathrm{u}^2(\mathrm{x_W})}\right]. \tag{8}$$

If for a prior distribution for Y, we take the non-informative improper prior distribution

$$g(Y) \propto 1, \tag{9}$$

then by Bayes's theorem [3], the posterior distribution for Y is

$$h(Y|\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}) = g(Y) \times L(Y|\{\mathrm{x}_i, \mathrm{u}(\mathrm{x}_i)\}) \propto \exp\left[-\frac{1}{2}\frac{(Y - \mathrm{x_W})^2}{\mathrm{u}^2(\mathrm{x_W})}\right]. \tag{10}$$

From (10), it follows that a Bayesian posterior pdf for Y is normal with expected value $\mathrm{x_W}$ and variance $\mathrm{u}^2(\mathrm{x_W})$. This pdf represents the collective state of knowledge concerning ξ implied by assumptions 1, 2 and 3 attributed to the data $\mathrm{x}_1, \ldots, \mathrm{x}_n$. Whether these assumptions could reasonably be attributed to the given data $\mathrm{x}_1, \ldots, \mathrm{x}_n$ is a matter beyond the scope of this paper.

3. Comments on Lira's procedure

Lira [1] does not explain how expression (1) comes out of Bayes's theorem. Also, what does this expression mean? In particular, what does the delta function $\delta(Y - X_i)$ mean? According to assumption 4, the pdf of X_i is normal with expected value x_i and variance $\sigma_i^2 = \mathrm{u}^2(\mathrm{x}_i)$, for $i = 1, \ldots, n$. Thus the random variables $X_1, \ldots, X_n$ have different fully specified distributions. A pdf for the random variable Y is to be determined. As discussed in section 2, according to Bayes's theorem, assumptions 1, 2 and 3 imply that a Bayesian posterior pdf for Y is normal with expected value $\mathrm{x_W}$ and variance $\mathrm{u}^2(\mathrm{x_W})$. Therefore, for each $i = 1, \ldots, n$, the distributions of X_i and Y are different, unless $n = 1$.

3.1. Definition of a likelihood function

In Bayesian statistics, the likelihood function of an unknown parameter (conditional on the realized data) is the sampling pdf of the data generation process (conditional on the value of the unknown parameter) regarded as a function of the value of the parameter [3]. Thus, the likelihood function (7) is the sampling pdf (6) regarded as a function of Y representing the state of knowledge concerning the unknown value ξ of the measurand. Unlike Bayesian statistics [3], Lira [1] develops his likelihood function (3) from the fully specified state-of-knowledge distribution (2) rather than a sampling distribution. Also, Lira's [1] method of developing the likelihood function (3) from the fully specified state-of-knowledge distribution (2) does not agree with Bayesian statistics [3].

3.2. Replacing the variances $\mathrm{u}^2(\mathrm{x}_i)$ by the random variables $\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2$

The *GUM* [4] recommends that a result which is subject to a recognized non-negligible bias (offset) should be corrected and the uncertainty associated with the correction should be included in the combined standard uncertainty associated with the corrected result. Suppose q is an uncorrected result (which is subject to a recognized bias/offset) with uncertainty u(q). If c is the correction with uncertainty u(c), then the corrected result, based on the measurement equation $M = Q + C$, is $\mathrm{m} = \mathrm{q} + \mathrm{c}$. The combined uncertainty associated with m is $\mathrm{u(m)} = \sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c})}$ [6]. It has been suggested that in some situations it is more practical to enlarge the uncertainty $\mathrm{u(m)} = \sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c})}$ in lieu of correcting the result q [7]. Lira and Wöger [6] suggest that if the result q is not corrected then its uncertainty may be enlarged from $\sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c})}$ to $\sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c}) + \mathrm{c}^2}$. The approach of Lira and Wöger [6] is an alternative[1] to the approach proposed in [7]. Note that the pairs (q, u(q)), (c, u(c)) and (m, u(m)) are known quantities (constants) representing the expected values and standard deviations of the state-of-knowledge distributions for Q, C and M, respectively. In particular, the uncertainties $\sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c})}$ and $\sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c}) + \mathrm{c}^2}$ are known quantities (constants).

In Lira [1], all n biases (offsets) $\mu_i - \xi$ are assumed to be zero; therefore, Lira and Wöger [6], which addresses biased results, does not apply to the problem addressed in Lira [1]. Nevertheless, Lira [1] uses the random variable $Y - \mathrm{x}_i$ as a correction c_i to enlarge the uncertainty associated with the result x_i from $\mathrm{u}(\mathrm{x}_i)$ to $\sqrt{\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2}$. Note that, unlike Lira and Wöger [6], Lira [1] equates the random variable $Y - \mathrm{x}_i$ to the correction c_i, a supposedly known constant representing the expected value of a correction variable and that $\sqrt{\mathrm{u}^2(\mathrm{x}_i) + (Y - \mathrm{x}_i)^2}$ is a random variable rather than a constant such as $\sqrt{\mathrm{u}^2(\mathrm{q}) + \mathrm{u}^2(\mathrm{c}) + \mathrm{c}^2}$.

3.3. Replacing the random variables X_i with the random variable Y

Lira [1] replaces each of the n variables X_i with the same variable Y in (3) to obtain expression (4). Then Lira [1] claims expression (4) to be the kernel of a pdf for Y when n is at least three. Lira [1] does not provide a mathematical basis for replacing the n random variables X_i with the same random variable Y. In addition, it is not clear why expression (4) should be a Bayesian posterior pdf. The seemingly arbitrary requirement that n must be at least three for expression (4) to yield a pdf must raise a flag that perhaps something is wrong in Lira's procedure.

3.4. Applicability of expression (4) to inconsistent results

Lira [1, section 1] states that expression (4) yields a state-of-knowledge pdf for Y that is valid whether or not the results $\mathrm{x}_1, \ldots, \mathrm{x}_n$ are consistent. A chi-square test of consistency, widely accepted by metrologists [2], implies that if the results $\mathrm{x}_1, \ldots, \mathrm{x}_n$ are judged to be inconsistent then the expected values $\mu_1, \ldots, \mu_n$ may not be regarded as equal. If $\mu_1, \ldots, \mu_n$ were not all equal, then not all of $\mu_1, \ldots, \mu_n$ could equal the value ξ of the measurand. Thus assumption 3 is refuted. Consequently, contrary to the claim of Lira [1], expression (4) determined from assumptions 1 through 5 does not apply when the results $\mathrm{x}_1, \ldots, \mathrm{x}_n$ are inconsistent.

3.5. Degrees of equivalence

Lira [1, section 2.2] suggests that the unilateral degree of equivalence of a laboratory can be obtained from an expression such as (4) by deleting the contribution of that laboratory. We note that expression (4) is determined from assumptions 1 through 5 including assumption 3 that all biases (offsets) $\mu_i - \xi$ are zero. To quantify unilateral degrees of equivalence, we need to determine state-of-knowledge distributions for $X_i - Y$ without assuming that all biases are zero [8]. Lira [1] has not determined a pdf for Y without assuming that all biases are zero; thus, he has not determined pdfs for $X_i - Y$ that are needed to quantify unilateral degrees of equivalence, for $i = 1, \ldots, n$.

4. Concluding remarks

Lira [1] offers an alternative to the method [2] recommended by a BIPM Advisory Group on Uncertainties for the evaluation

[1] Lira and Wöger [6] propose an alternative to the method of Phillips and Eberhardt [7]. However, Lira and Wöger [6] do not support the idea of enlarging the uncertainty in lieu of correcting a result for recognized bias. We quote Lira and Wöger [6, p 1011], 'Thus it is hard to think of any reason why this approach should be followed, although the decision must of course be left to the (well informed) producers and users of the results of measurements.'

of data from a simple key comparison. For the reasons discussed in section 3, we are skeptical of Lira's [1] claim that expression (4) is the kernel of a pdf for the unknown value of the common measurand that encodes the collective state of knowledge whether or not the interlaboratory results are consistent. We cannot agree with Lira's [1] claim that his procedure is Bayesian. Bayesian statistics is discussed in books such as [3, 9–13]. Lira's procedure, reviewed here in section 1.3, does not agree with the cited literature. In our view, a statistical method that disagrees with the cited literature (on Bayesian statistics) is not a Bayesian procedure.

Acknowledgment

Blaza Toman, an anonymous referee, Jim Lawrence, Jeff McFadden, Ron Boisvert, Charles Hagwood and Javier Bernal provided useful comments on earlier drafts of this paper; thank you.

References

[1] Lira I 2006 Bayesian evaluation of comparison data *Metrologia* **43** S231–4
[2] Cox M G 2002 The evaluation of key comparison data *Metrologia* **39** 589–95
[3] Box G E P and Tiao G C 1973 *Bayesian Inference in Statistical Analysis* (Reading, MA: Addison Wesley)
[4] ISO 1995 *Guide to the Expression of Uncertainty in Measurement* 2nd edn (Geneva: International Organization for Standardization) ISBN 92-67-10188-9
[5] Lira I and Wöger W 2006 Comparison between the conventional and Bayesian approaches to evaluate measurement data *Metrologia* **43** S249–59
[6] Lira I and Wöger W 1998 Evaluation of the uncertainty associated with a measurement result not corrected for systematic effects *Meas. Sci. Technol.* **9** 1010–11
[7] Phillips S D and Eberhardt K R 1997 Guidelines for expressing the uncertainty of measurement results containing uncorrected bias *J. Res. NIST* **102** 577–85
[8] Kacker R N, Datla R U and Parr A C 2004 Statistical analysis of CIPM key comparisons based on the ISO *Guide Metrologia* **41** 340–52
[9] Berger J O 1985 *Statistical Decision Theory and Bayesian Analysis* 2nd edn (Berlin: Springer)
[10] Bernardo J M and Smith A F M 1994 *Bayesian Theory* (New York: Wiley)
[11] Lee P M 1997 *Bayesian Statistics An Introduction* 2nd edn (Oxford: Oxford University Press)
[12] Press S J 2003 *Subjective and Objective Bayesian Statistics* 2nd edn (New York: Wiley)
[13] Gelman A, Carlin J B, Stern H S and Rubin D B 2004 *Bayesian Data Analysis* 2nd edn (London: Chapman and Hall)