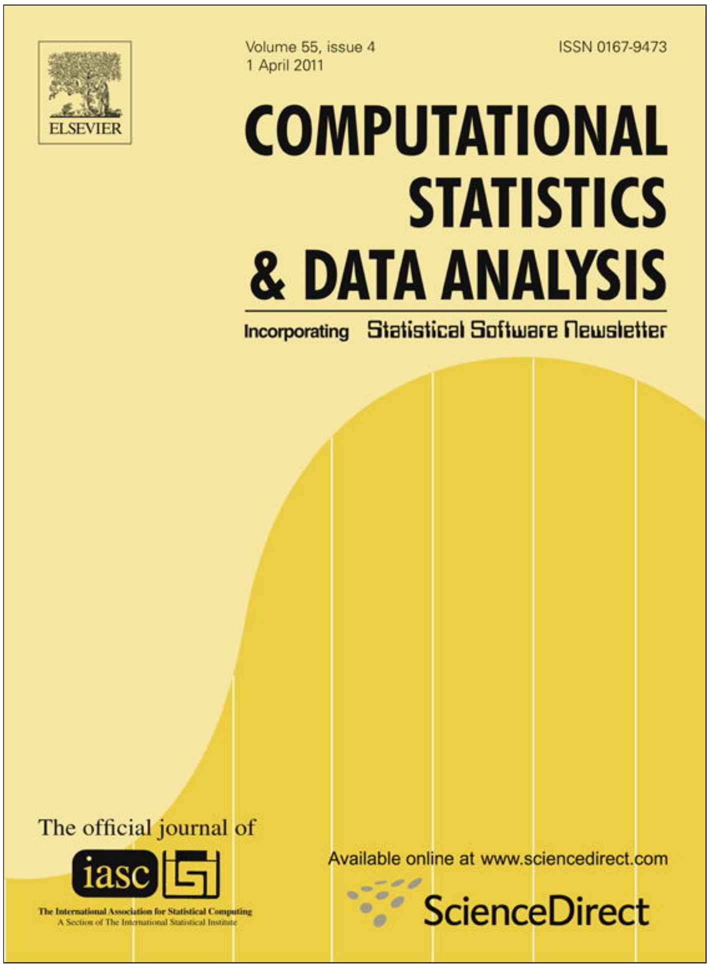
Volume 55, issue 4
1 April 2011
ISSN 0167-9473
ELSEVIER
COMPUTATIONAL STATISTICS & DATA ANALYSIS
Incorporating Statistical Software Newsletter
The official journal of
iasc
The International Association for Statistical Computing
A Section of The International Statistical Institute
Available online at www.sciencedirect.com
ScienceDirect

Computational Statistics and Data Analysis 55 (2011) 1815–1827

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Laplace random effects models for interlaboratory studies

Andrew L. Rukhin *, Antonio Possolo

Statistical Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

ARTICLE INFO

Article history:
Received 15 June 2010
Received in revised form 15 November 2010
Accepted 16 November 2010
Available online 25 November 2010

Keywords:
Bayes estimators
Degree of equivalence
Heteroscedasticity
Interlaboratory study
Key comparisons
Laplace distribution
Maximum likelihood estimation
Meta-analysis
Random effects model

ABSTRACT

A model is introduced for measurements obtained in collaborative interlaboratory studies, comprising measurement errors and random laboratory effects that have Laplace distributions, possibly with heterogeneous, laboratory-specific variances. Estimators are suggested for the common median and for its standard deviation. We provide predictors of the laboratory effects, and of their pairwise differences, along with the standard errors of these predictors. Explicit formulas are given for all estimators, whose sampling performance is assessed in a Monte Carlo simulation study.

Published by Elsevier B.V.

1. Interlaboratory studies and key comparisons

The international agreement, the so-called "Mutual Recognition Arrangement" (MRA) (CIPM, 1999) on mutual recognition of national measurement standards, calibration and measurement certificates issued by national metrology institutes (NMIs) calls for the execution of interlaboratory studies aimed at testing principal techniques and measurement methods in a particular field of science. These studies are organized by the Consultative Committees (CCs) of the *Comité International des Poids et Mesures* (CIPM), there are CCs for length, mass, amount of substance, etc. Such interlaboratory studies are called *Key Comparisons* (KCs), and one of their principal goals is to establish the degree of equivalence of national measurement standards which characterize the extent to which each institute may have confidence in the results reported by other NMIs. Typically, a KC produces a *key comparison reference value* (KCRV): for example, in a KC focusing on the length of a gauge block, this should be the block's true length (ISO/IEC, 2007, 5.18) although in actuality it is the best estimate of this length. In a KC focusing on the mass fraction of a particular substance in a certified reference material of which aliquots are distributed to the participating NMIs for analysis, this could be the mass fraction of one or more selected compounds.

The MRA defines the *Degree of Equivalence* of a national measurement standard (unilateral DoE) as comprising its deviation from the key comparison reference value and the uncertainty of this deviation. According to the MRA the degree of equivalence between a pair of national measurement standards (bilateral DoE) is formed by the difference of their deviations from the reference value and the uncertainty of this difference. If a reference value cannot be meaningfully defined (for example, when the KC involves multiple circulating artifacts and not all NMIs measure all of them), the KC results might be expressed directly in terms of the degrees of equivalence between pairs of standards.

* Corresponding author. Fax: +1 301 975 3144.
E-mail addresses: andrew.rukhin@nist.gov (A.L. Rukhin), antonio.possolo@nist.gov (A. Possolo).

0167-9473/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.csda.2010.11.016

The *International vocabulary of metrology* (VIM) (ISO/IEC, 2007, 2.26) defines *measurement uncertainty* as a "non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used", and adds that this "parameter may be, for example, a standard deviation". For this reason we follow the customary usage in statistics and use either "standard deviation" or "standard error" throughout, where metrologists might use "measurement uncertainty" instead.

The ith of n NMI participating in a KC is supposed to produce a measured value x_i and an assessment of its standard error u_i. Thus we start with a set of n (scalar) measurement values $x_1, \ldots, x_n$, and the corresponding standard errors $u_1, \ldots, u_n$, which we assume are known. The precise meaning of these uncertainties often is debated heatedly. Could they be regarded as known quantities, or instead are they merely estimates of unknown quantities? Some metrologists insist that uncertainties are *computed* (with assuredness and certainty of arithmetic), while others concede that they are only *estimated*. When x_i and u_i are modeled as in Bayesian inference, one can convincingly argue that conditionally upon the data, u_i indeed are computed as standard deviations of suitable posterior distributions, and therefore are known with certainty which is our assumption.

The next section introduces the model. We find estimators of the KCRV (8) and of its standard error (10), as well as of a scale parameter (9) in the distribution of the random interlaboratory effects. Assuming these parameters to be known, in Section 3 the estimators of the degrees of equivalence (value and standard error) are derived. Section 4 contains results of Monte Carlo simulation. Most of the formulas needed in Section 3 are collected in the Appendix.

2. Random effects and common median model

2.1. Mixed effects Laplace model

In many practical cases, all confidence intervals based on measured values x_i and their standard errors u_i do not overlap, which suggests that the dispersion of the measured values x_i is greater than what their standard errors might lead one to expect. It is also fairly common that a few of the measurements deviate markedly from the bulk of the rest.

The first kind of situation can be dealt with by modeling the measurements as outcomes of independent random variables $X_i = \mu + B_i + E_i$ for $i = 1, \ldots, n$, where μ is the unknown KCRV, B_i is a lab-specific random effect, and E_i represents measurement error. Similarly to a common practice in robust estimation (Wilcox, 2005) the second eventuality can be addressed by modeling the distributions of B_i and E_i as suitably heavy-tailed. The results are then analyzed using either *ad hoc* robust statistical methods, or likelihood methods, conventional or Bayesian, that guarantee similar robustness within a parametric framework.

There is a precedent to this general approach. For example, Pinheiro et al. (2001) describe a model that is based on Student's t-distribution. In the same spirit, the median has been suggested as a possible consensus KCRV estimator e.g., Cox (2002). However, this method does not use the standard errors u_i at all; neither do other robust estimators that have been suggested to address the same problem, e.g., Analytical Methods Committee (1989a,b) and Thompson et al. (2006). The departures from the Gaussian random effects linear model, that appear most detrimental to the performance of the estimators, while staying within the realm of symmetric distributions, are heaviness of the tails of the distribution of B_i, and incomplete knowledge of the variances of the measurement errors E_i.

To account for these facts we suggest a mixed effects model in which both B_i and E_i have suitably scaled Laplace (double-exponential) distributions. This model, as we shall show, is far more robust than the traditional Gaussian model, while incurring only moderate loss of efficiency in this traditional case. Similarly to what other laboratory effects do, it also overcomes the problem of "inconsistency" (between the x_i). See Toman and Possolo (2009) for critique of consistency testing proposed by Decker et al. (2006) and Cox (2007). In our model both lab-specific random effects and within-lab measurement errors can be interpreted as Gaussian but with variances u_i^2 that are like random draws from exponential distributions.

Our estimation method for μ on the basis of heterogeneous data is based on the statistic,

$$\widetilde{\mu} = \arg\min_{\mu} \sum_{i=1}^{n} \frac{|x_i - \mu|}{u_i}, \tag{1}$$

which is a weighted median. This procedure has a maximum likelihood interpretation to be discussed in the next section, and there are efficient numerical algorithms for its evaluation; see Bloomfield and Steiger (1983). Indeed, medians weighed by their standard errors have already been suggested as KCRV estimators; see Müller (2000), Ratel (2006). Besides those mentioned above, Rocke (1983), Davies (1991), Lischer (1996) and Duewer (2006, 2008) advocate the use of robust statistics (including the median) in interlaboratory studies. The same robustness issues arise in the more general context of meta-analysis; see Hedges and Olkin (1985). To address meta-analysis problems robustly, Demidenko (2004) uses a setting which is somewhat similar to the following model by assuming Gaussian errors and Laplace between-lab effects. Wilcox (2006) points out difficulties with homogeneity testing for the medians.

The model we propose for KCRV estimation, and for the assessment of its standard error is this: the measured values, $x_1, \ldots, x_n$, are outcomes of random variables

$$X_i = \mu + B_i + E_i, \tag{2}$$

where B_i and E_i, $i = 1, \ldots, n$, are independent random variables with Laplace distributions whose densities are f_β and f_{u_i} such that,

$$f_\beta(b_i) = \frac{1}{2\beta} e^{-|b_i|/\beta}, \qquad f_{u_i}(e_i) = \frac{1}{2u_i} e^{-|e_i|/u_i}. \tag{3}$$

The unknown parameters are the unknown common median μ (the deterministic common effect, consensus value, or KCRV), the scale parameter β, and the realized values $b_1, \ldots, b_n$ of random, between-lab effects $B_1, \ldots, B_n$, whose density is given in (3).

The parameter β plays an important role in the estimation of μ. In Section 2 β is initially restricted to a finite set, namely to $\{u_1, \ldots, u_n\}$ with the initial value $\beta = \max_i u_i$. This parameter can be interpreted as a penalty which discourages large absolute values of b_i; see Koenker (2005).

An alternative but equivalent formulation is that the conditional distribution of X_i for given $B_i = b_i$ is the Laplace distribution with the mean $\mu + b_i$ and the scale parameter u_i. The distribution of the random effect B_i is also a Laplace distribution with zero mean and the scale parameter β. Model (2) can be motivated in the same way as the random effects model with Gaussian errors and between-lab effects. However, now the distributions of B_i and E_i are taken to be Laplace, that is, as a compound Gaussian distribution with common zero mean and exponentially distributed variance; see Kotz et al. (2001).

This fact allows for an extension of our model to accommodate KCs where u_i is qualified with the number d_i of "degrees of freedom". In this case the uncertainties u_i are not known with certainty, but instead have their degrees of freedom. Then for suitable constants κ_i represent measurement errors as $E_i = \kappa_i u_i Z_i \sqrt{\chi^2_{\nu_i}/\nu_i}$, where Z_i are independent standard normal variables, and $\chi^2_{\nu_i}$ are mutually independent random variables, also independent of the Z_i, with χ^2-distributions on ν_i degrees of freedom. Although the subsequent explicit formulas are not applicable in this case, direct maximization of the likelihood or Bayesian estimation via MCMC is viable in this case, as is illustrated in Section 4.

We present now the strategy to obtain the desirable estimators. By employing the maximum profile likelihood method (or by using the posterior mode when μ has the uniform prior) we derive the estimator (predictor) of random effects in (7) as well as the estimator (8) of the common median. This is done for a given value of β using an empirical Bayes method. Then in Section 3 we give the conditional distribution of B_i for given X_i which is used to get estimators of the unilateral and bilateral degrees of equivalence: $\widetilde{b}_i$ from (13) and $\widetilde{b}_i - \widetilde{b}_j$, and their standard errors $u(\widetilde{b}_i)$ from (15) and $V(i, j)$ from (14).

2.2. Estimators of the common median and random effects

Assume that u_i are known, and β is fixed. We start with the maximum profile likelihood estimator or by assuming that μ has the uniform (non-informative) prior, with the posterior mode which can be regarded as the Bayes estimator of μ. This estimator maximizes

$$\prod_{i=1}^{n} \left[\frac{1}{2u_i} \exp\left\{ -\frac{|x_i - b_i - \mu|}{u_i} \right\} \frac{1}{2\beta} \exp\left\{ -\frac{|b_i|}{\beta} \right\} \right]. \tag{4}$$

The function (4) can be maximized directly, for example using the Nelder–Mead algorithm as implemented in R function *optim*. Gurwitz (1990) had proposed three different algorithms (modified quicksort, modified heapsort and linear-time) for similar optimization problems, and the referee has suggested another iterative algorithm which maximizes (4) using the idea of reweighted least squares.

To find the posterior mode notice that

$$\arg\min_{\mu, b_1, \ldots, b_n} \sum_{i=1}^{n} \left(\frac{|x_i - b_i - \mu|}{u_i} + \frac{|b_i|}{\beta} \right) = \arg\min_{\mu, h_1, \ldots, h_n} \sum_{i=1}^{n} \left(\frac{|y_i - h_i - \mu|}{\alpha_i} + \frac{|h_i|}{\beta} \right). \tag{5}$$

Here $y_1 \le y_2 \le \cdots \le y_n$ are the order statistics of the x_i, that is $y_i = x_{\sigma(i)}$, where the permutation σ defines the anti-ranks of x's, and $\alpha_i = u_{\sigma(i)}$, $h_i = b_{\sigma(i)}$. One has

$$\begin{aligned} \min_{\mu, h_1, \ldots, h_n} \sum_{i=1}^{n} \left(\frac{|y_i - h_i - \mu|}{\alpha_i} + \frac{|h_i|}{\beta} \right) &= \min_{\mu, h_1, \ldots, h_n, c} \sum_{i=1}^{n} \left(\frac{|y_i - h_i - \mu|}{\alpha_i} + \frac{|h_i - c|}{\beta} \right) \\ &= \min_{\mu, h_1, \ldots, h_n} \sum_{i=1}^{n} \left(\frac{|y_i - h_i - \mu|}{\alpha_i} + \frac{|h_i - \mathrm{median}(h_i)|}{\beta} \right) \\ &= \min_{\substack{\mu, h_1, \ldots, h_n \\ \mathrm{median}(h_i) = 0}} \sum_{i=1}^{n} \left(\frac{|y_i - h_i - \mu|}{\alpha_i} + \frac{|h_i|}{\beta} \right). \end{aligned} \tag{6}$$

In other words, if $\widehat{b}_i$ denotes a minimizer of (5), $\mathrm{median}(\widehat{b}_i)$ can be taken to be zero, i.e., only differences $b_i - \mathrm{median}(b_i)$ are estimable from (5).

The objective function in (5) is a convex (continuous), piecewise linear function of μ and h_i (or b_i). To find its minimum notice that for fixed i and μ, when $y_i > \mu$, $\widehat{h}_i = 0$, if $\alpha_i \ge \beta$; $= y_i - \mu$, if $\alpha_i < \beta$, and the same formulas hold when $\mu \le y_i$. Thus,

$$\widehat{h}_i = \widehat{h}_i(\mu) = \begin{cases} 0 & \alpha_i \ge \beta, \\ y_i - \mu & \alpha_i < \beta, \end{cases} \tag{7}$$

i.e., $\widehat{h}_i = \widehat{b}_{\sigma(i)}$ vanishes if the reported standard error of lab i exceeds β. The posterior mode (7) can assume only two values, 0 and $x_i - \mu$, and is unable to predict all linear combinations of b_i, so (7) should not be used as the final estimator.

Still the form of (7) provides a useful/estimate of μ. Let

$$\begin{aligned}\widehat{\mu} &= \arg\min_{\mu}\left[\sum_{i=1}^{n}\frac{|y_i - \widehat{h}_i - \mu|}{\alpha_i} + \frac{|\widehat{h}_i|}{\beta}\right] \\ &= \arg\min_{\mu}\left[\sum_{i:\alpha_i \ge \beta}\frac{|y_i - \mu|}{\alpha_i} + \sum_{i:\alpha_i < \beta}\frac{|y_i - \mu|}{\beta}\right] \\ &= \arg\min_{\mu}\sum_{i=1}^{n}\min\left(\frac{1}{\alpha_i}, \frac{1}{\beta}\right)|y_i - \mu| = \arg\min_{\mu}\sum_{i} w_i|y_i - \mu|,\end{aligned} \tag{8}$$

i.e., $\widehat{\mu}$ is a weighted median of y_i with weights $w_i = 1/\max(\alpha_i, \beta)$. For its numerical evaluation find a, $1 \le a \le n$, such that $\sum_1^{a-1} w_k < 2^{-1}\sum_1^n w_k \le \sum_1^a w_k$, and put $\widehat{\mu} = y_a$. Thus, although the minimum in (8) is possibly attained on an interval, as the classical median, the weighted median can be restricted to the set of observed values.

If $\beta \ge \max_i u_i$, $\widehat{\mu}$ is the classical median, and Cox's (2002) proposal mentioned in Section 2 is a particular case of our method (in which the standard errors u_i do play a role!) In this situation $\widehat{b}_i \ne 0$ except possibly for exactly one value of i. If $\beta \le \min_i u_i$, then $\widehat{b}_i \equiv 0$, and $\widehat{\mu}$ coincides with (1).

If the set $\mathcal{C} = \{i : \widehat{b}_i \ne 0\}$ is not empty, the maximization of the likelihood function (4) gives the empirical Bayes estimator of β,

$$\widehat{\beta} = \frac{1}{\operatorname{card}(\mathcal{C})}\sum_{i\in\mathcal{C}}|\widehat{b}_i|.$$

Otherwise the initial value of β is to be increased. If this (recommended by us) initial value is $\max_i u_i$, then

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}|x_i - \text{median}(x_i)|}{n-1}. \tag{9}$$

Simulation studies, including those whose results are reported in Section 4, suggest that the estimator (9) be used to determine $\widehat{\mu}$ in (8).

The variability of $\widehat{\mu}$ needs to be assessed. The asymptotic theory, Koenker (2005) suggests that $\widehat{\mu} - \mu$ is approximately normal with zero mean and variance $\sigma^2 = (\sum_k w_k^2)[\sum_k w_k/(u_k + \beta)]^{-2}$, which can be estimated by

$$\widehat{\sigma}^2 = \frac{\sum_k \hat{w}_k^2}{\left[\sum_k \frac{\hat{w}_k}{u_k + \widehat{\beta}}\right]^2}, \tag{10}$$

$\hat{w}_k = 1/\max(\alpha_k, \widehat{\beta})$, $k = 1, \ldots, n$. If $\widehat{\beta} = \max_i u_i$, then $\hat{w}_k \equiv \widehat{\beta}^{-1}$, and $\widehat{\sigma}^2 = n\left[\sum_k (u_k + \widehat{\beta})^{-1}\right]^{-2}$. According to simulation results, $\widehat{\sigma}^2$ underestimates the variance for small to medium sample sizes. In such cases the distribution of $(\widehat{\mu} - \mu)/\widehat{\sigma}$ is better approximated by a t-distribution with $n - 1$ degrees of freedom, and $\widehat{\mu} \pm t_{\alpha/2,n-1}\widehat{\sigma}$, can be suggested as an approximate $100(1 - \alpha)\%$ confidence interval for μ.

To conclude this section notice that when β is unknown, the joint profile likelihood for μ and β is not bounded. Indeed formulas (5), (7) and (8) show that the negative logarithm of this function is $\sum |x_i - \mu|/\max(u_i, \beta) + n\log\beta$, which tends to $-\infty$ as $\beta \to 0$. Thus, formally the maximum likelihood estimator of β is 0, which is not useful, and this is the reason for employing the maximum profile likelihood estimator.

In Section 4 we report some simulation results for the maximum likelihood estimator of μ based on the marginal likelihood $\prod_i p_i(x_i - \mu)$ where

$$p_i(x) = \frac{1}{4\beta u_i}\int \exp\left\{-\frac{|x - t|}{u_i} - \frac{|t|}{\beta}\right\}dt = \frac{\beta e^{-|x|/\beta} - u_i e^{-|x|/u_i}}{2(\beta^2 - u_i^2)}, \tag{11}$$

denotes the density of the sum of two independent Laplace distributed random variables with parameters $u_i \neq \beta$. When $u_i = \beta$,

$$p_i(x) = \frac{e^{-|x|/u_i}}{4u_i}\left(1 + \frac{|x|}{u_i}\right).$$

Notice that in the formula (2.3.23) for p_i in Kotz et al. (2001) s_i should be replaced by $1/s_i$.

3. Degrees of equivalence and their standard errors

A clear benefit of model (2) is an easy interpretation and estimation of DoE. Indeed as in Toman and Possolo (2009), let the unilateral DoE for lab i be b_i, and let the bilateral DoE for two labs, say, i and i', be $b_i - b_{i'}$. The standard deviation of the Laplace distribution (3) is measured by the parameter β which is the absolute moment of order one, or by the square root of one-half of the second moment,

$$\beta = E|B_i| = \left(\frac{EB_i^2}{2}\right)^{1/2}.$$

Since β has been estimated by (9), and μ was determined from (8), we will take these parameters, as well as the u_i, to be known when using the notation,

$$g_i(t) = \frac{\exp\{-|x_i - \mu - t|/u_i - |t|/\beta\}}{4u_i\beta p_i(x_i - \mu)}$$

for the density of the conditional distribution of B_i for given $x_i - \mu$. The same holds for the corresponding distribution function,

$$G_i(t) = P(B_i \le t|x_i) = \int_{-\infty}^{t} g_i(s)\mathrm{d}s,$$

for which

$$G_i(t) = \begin{cases} \dfrac{e^{-(x_i-\mu)/u_i + t(1/u_i+1/\beta)}}{4(u_i+\beta)p_i(x_i-\mu)} & t \le 0.5(x_i - \mu - |x_i - \mu|), \\ 1 - \dfrac{e^{(x_i-\mu)/u_i - t(1/u_i+1/\beta)}}{4(u_i+\beta)p_i(x_i-\mu)} & t > 0.5(x_i - \mu + |x_i - \mu|), \end{cases}$$

so that G_i has exponential tails. In the interval $\{t : |t - 0.5(x_i - \mu)| \le 0.5|x_i - \mu|\}$, the function $G_i(t)$ is an integral of the exponential function $\exp\{\mathrm{sgn}(x_i - \mu)t(1/u_i - 1/\beta)\}$.

The particular cases are used in the following sections: (i) $x_i = \mu$, G_i the distribution function of the Laplace distribution with the parameter $\gamma_i = u_i\beta/(u_i + \beta)$, i.e.,

$$\frac{1}{\gamma_i} = \frac{1}{u_i} + \frac{1}{\beta},$$

(ii) $u_i = 0$, G_i the point mass at $x_i - \mu$, (iii) $\beta = 0$, G_i the point mass at zero.

3.1. Unilateral degrees of equivalence

We examine two (empirical) Bayes estimators of the random effect (or unilateral DoE) namely, the posterior mean $\bar{b}_i = E(B_i|x_i)$ (which is the classical unbiased predictor) and the posterior median $\tilde{b}_i$ (which is the median-unbiased predictor.) The posterior mode $\hat{b}_i$ is less attractive than these as was discussed in Section 2.

The posterior mean is

$$\bar{b}_i = \int_{-\infty}^{\infty} t g_i(t)\mathrm{d}t = \frac{\mathrm{sgn}(x_i-\mu)\beta}{p_i(x_i-\mu)(\beta^2 - u_i^2)}\left[e^{-|x_i-\mu|/\beta}\left(\frac{|x_i-\mu|}{2} + \gamma_i\right) - 2u_i\beta p_i(x_i-\mu)\right]. \tag{12}$$

The median $\tilde{b}_i$ of the posterior distribution is the solution of the equation $G_i(\tilde{b}_i) = 0.5$,

$$\tilde{b}_i = \frac{\beta(x_i-\mu)}{\beta - u_i} + \frac{\beta u_i \mathrm{sgn}(x_i-\mu)}{\beta - u_i}\log\left(\frac{\beta e^{-|x_i-\mu|/\beta} + u_i e^{-|x_i-\mu|/u_i}}{\beta + u_i}\right). \tag{13}$$

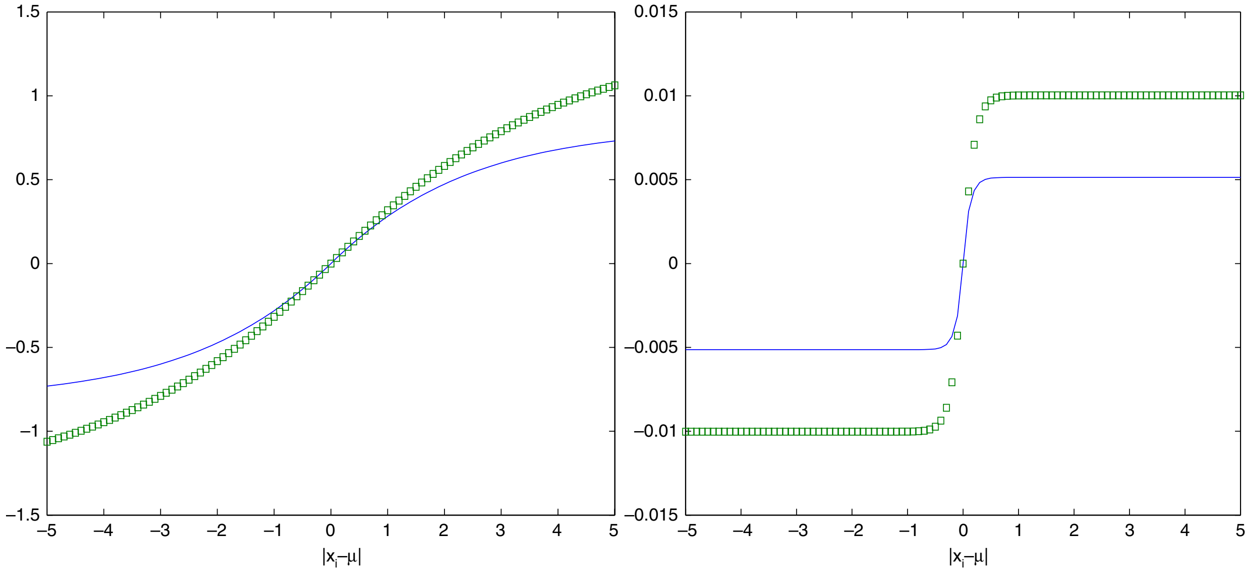

Fig. 1. Graphs of $\bar{b}_i$ (line marked by squares) and $\widetilde{b}_i$ (continuous line) when $\beta = 1 < 2 = u_i$ (left panel), and when $\beta = 0.1 < 2 = u_i$ (right panel).

Favoring $\widetilde{b}_i$ (over $\bar{b}_i$) are its fairly simple algebraic form (which does not require evaluation of $p_i(x_i - \mu)$) and the mean–median–mode inequality according to which $\widetilde{b}_i$ is always between $\bar{b}_i$ and $\widehat{b}_i$. Fig. 1 depicts the graphs of these estimators for several values of β and u_i when $\beta < u_i$. When $\beta > u_i$, these graphs look like straight lines.

Two alternative estimators of DoE's standard error are, $u(\widetilde{b}_i) = E(|B_i||x_i)$, $Eu(\widetilde{b}_i) = E|B_i|$, and $u(\bar{b}_i) = [2^{-1}E(B_i^2|x_i)]^{1/2}$, $Eu(\bar{b}_i)^2 = 2^{-1}EB_i^2$. Their form is given in the Appendix. Fig. 2 show the plots of $u(\widetilde{b}_i)$ and $u^2(\bar{b}_i)$ for several values of β and u_i when $\beta < u_i$.

These formulas reveal that if $\beta < u_i$, $\bar{b}_i$ and $\widetilde{b}_i$, as well as $u(\widetilde{b}_i)$ and $u(\bar{b}_i)$, are bounded functions of $x_i - \mu$. Interpreting $\mu + b_i$ in (2) as a signal, β/u_i becomes the signal-to-noise ratio. Our results imply that if this ratio is less than one, then the possible values of the signal's estimators cannot exceed a certain threshold: $\beta u_i(\log(u_i + \beta) - \log u_i)/(u_i - \beta)$ for the median, and $2\beta^2 u_i/(u_i^2 - \beta^2)$ for the mean. (For the mode, the estimator is zero.) The same is true for their standard errors. This fact is in stark contrast with the Gaussian model. Further discussion of "soft" detection based on Laplace distributions is in Poor (1994).

3.2. *Bilateral degrees of equivalence*

To estimate the bilateral DoE for labs i and j, $1 \le i \ne j \le n$ one can use $\widetilde{b}_i - \widetilde{b}_j$ with standard error estimator

$$\begin{aligned} U(i,j) &= E\left(|B_i - B_j|\Big|x_i, x_j\right) = \int\int |t-s| g_i(t) g_j(s) \mathrm{d}t\mathrm{d}s \\ &= \int_{-\infty}^{\infty} G_i(t)[1 - G_j(t)]\mathrm{d}t + \int_{-\infty}^{\infty} [1 - G_i(t)] G_j(t)\mathrm{d}t \\ &= u(\widetilde{b}_i) + u(\widetilde{b}_j) - 2\int_0^{\infty} \left[(1 - G_i(t))(1 - G_j(t)) + G_i(-t)G_j(-t)\right]\mathrm{d}t. \end{aligned}$$

An alternative bilateral DoE estimator, $\bar{b}_i - \bar{b}_j$, allows the following standard error estimator,

$$V(i,j) = \left[2^{-1}E((B_i - B_j)^2|x_i, x_j)\right]^{1/2} = [u^2(\bar{b}_i) + u^2(\bar{b}_j) - \bar{b}_i\bar{b}_j]^{1/2}, \tag{14}$$

which is much easier to evaluate. Typically the behavior of $V(i,j)$ is similar to that of $U(i,j)$. See Fig. 3 for the graphs and Appendix for some comparison results.

4. Example and simulation results

Table 1 shows the estimates of the KCRV corresponding to three different models for the mass fractions (ng/g) of several polychlorinated biphenyls (PCBs) in sediments that were the targets of KC CCQM-K25 (Schantz and Wise, 2004) and corresponding standard errors: for the weighted average (WAVE), Gaussian random effects model (GAU), and Laplace random effects model (LAP). The table's caption also lists the values of the estimates of the scale parameter β, and the medians of the corresponding u_i. Fig. 4 compares the estimates of the unilateral degrees of equivalence corresponding to the Gaussian and Laplace random effects models.

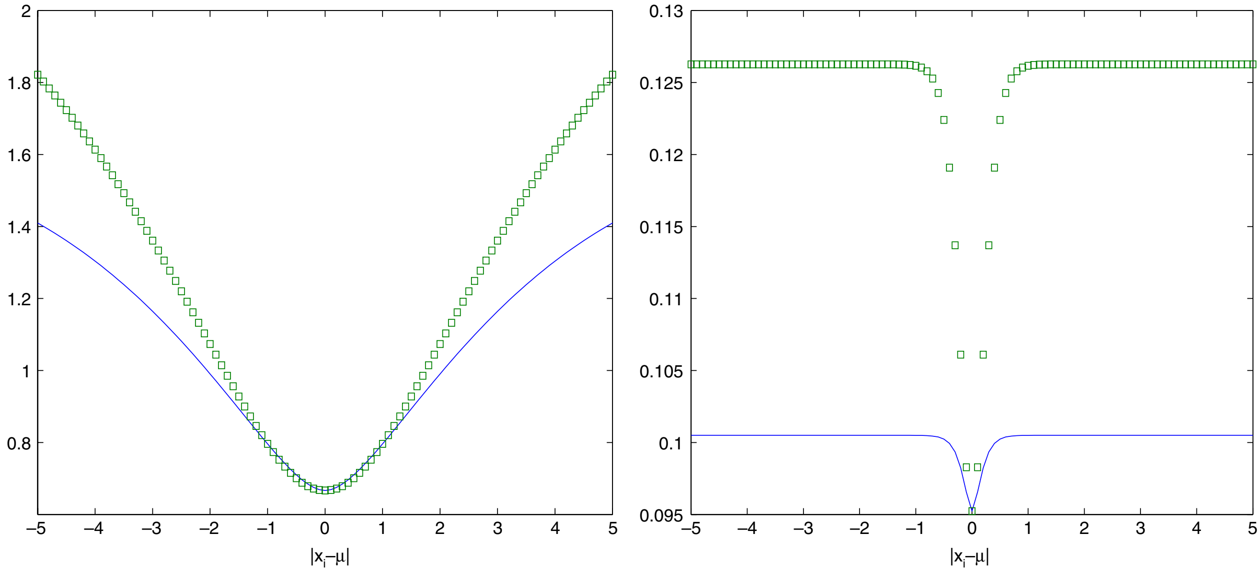

Fig. 2. Graphs of $u(\widetilde{b}_i)$ (continuous line) and $u(\bar{b}_i)$ (line marked by squares) when $\beta = 1 < 2 = u_i$ (left panel), and when $\beta = 0.1 < 2 = u_i$ (right panel).

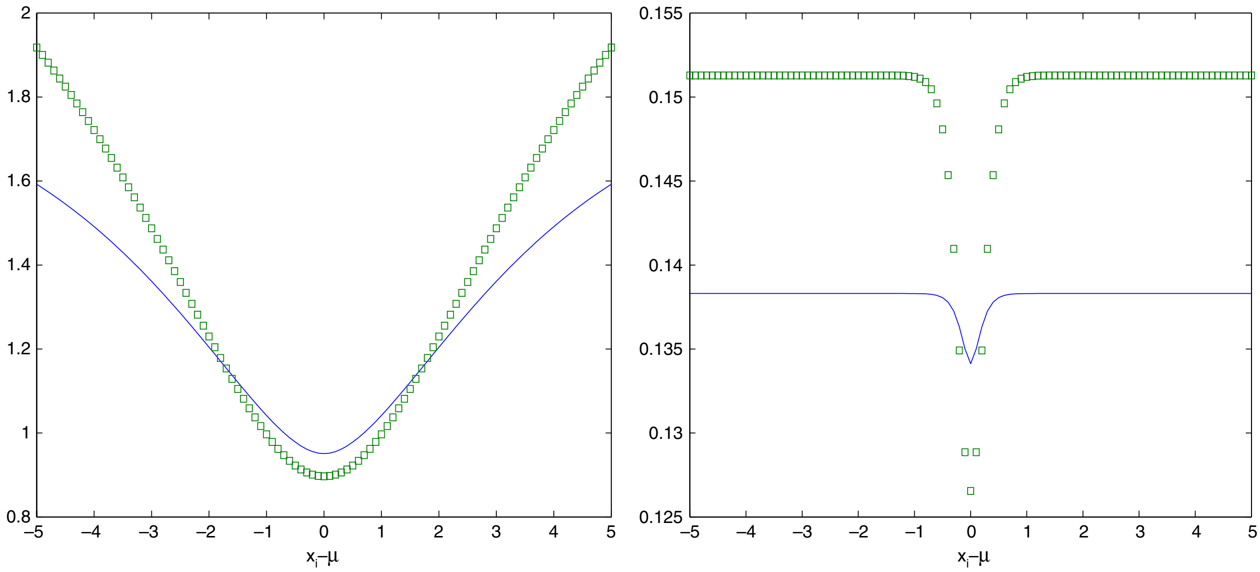

Fig. 3. Graphs of U (continuous line) and V (line marked by squares) when $\beta = 1 < 2 = u_i, x_j = \mu, u_j = 1.5$ (left panel), and when $\beta = 0.1 < 2 = u_i$, $x_j = \mu, u_j = 0.5$ (right panel).

Table 1
CCQM-KC25 estimates of the KCRV in ng/g units for five measurands corresponding to three different models, and corresponding standard errors in parentheses: LAP, Laplace random effects; GAU, Gaussian random effects; and WAVE, weighted average. For LAP, the estimates of the scale factor β were 1.23 (PCB 28), 0.34 (PCB 101), 0.32 (PCB 105), 1.00 (PCB 153), and 0.17 (PCB 170); the corresponding medians of the u_i were 0.54, 0.50, 0.18, 0.53, and 0.15.

PCB	LAP	GAU	WAVE
28	33.6 (0.74)	33.7 (0.65)	33.3 (0.18)
101	30.4 (0.35)	30.2 (0.26)	30.2 (0.18)
105	10.8 (0.19)	10.6 (0.18)	10.7 (0.05)
153	31.8 (0.53)	31.9 (0.45)	31.9 (0.14)
170	9.0 (0.11)	8.9 (0.10)	9.1 (0.04)

A Bayesian hierarchical model, where the number of degrees of freedom ν_i is taken into account by modeling $(x_i - \delta_i)/u_i$ as outcomes of Student's t_{ν_i}, and the δ_i as outcomes of Laplace random variables with mean μ and scale β, both with diffuse priors, was also fitted to the five PCB datasets via MCMC as implemented in function `metrop` of package `mcmc`, Geyer (2009) for R Development Core Team (2009). The estimates and their standard errors turned out quite similar to those listed under LAP in Table 1.

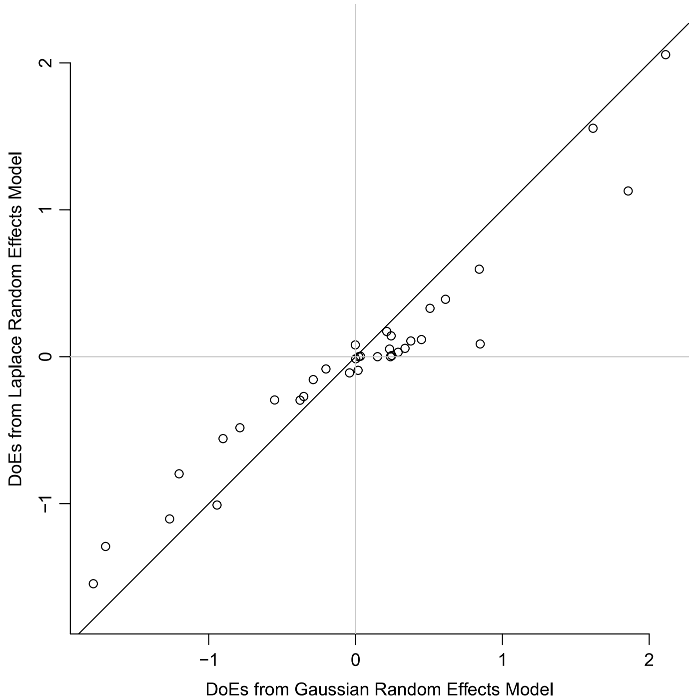

Fig. 4. Estimates of the unilateral degrees of equivalence produced by the Laplace random effects model plotted against the corresponding estimates produced by the Gaussian random effects model for the five PCBs listed in Table 1. Comparison of the locations of the points relative to the straight line with slope 1 and 0 intercept indicates that the Laplace random effects model generally "shrinks" the estimates of the DoEs toward 0 relative to its Gaussian counterpart.

Comparison of the standard errors in the last column of Table 1, under WAVE, with the standard errors in the other columns suggests that the conventional assessment for weighted averages could produce unrealistically small standard errors. The entries in the columns labeled LAP and GAU indicate that, in the absence of strikingly outlying labs, when the Gaussian random effects models may well be tenable, the Laplace random effects model reproduces its results fairly well.

We have undertaken a comparative Monte Carlo simulation study of the frequentist performance of the estimators of the KCRV μ corresponding to the Gaussian and Laplace random effects models, $X_i = \mu + B_i + E_i$, with $n = 13$, $\tau = 5$, $\beta = \tau/\sqrt{2}$, and heterogeneous variances for the E_i, under the sampling situations described next. The dispersion of the between-lab effects was about 20% of μ, and the dispersion of the measurement errors was about 5% of μ.

GAU Heterogeneous Gaussian random lab effects and measurement errors: $B_i \sim N(0, \tau^2)$, $E_i \sim N(0, \sigma_i^2)$, with different but known σ_i.

LAP Laplace random lab effects scaled to have variance τ^2, and Laplace measurement errors scaled to have variances σ_i^2 different but known.

SLA Slash random lab effects (ratio of independent Gaussian and uniform random variables scaled so that the median absolute deviation (MAD) of B_i is τ) and Gaussian measurement errors with different but known variances.

WIL Gaussian One-Wild random lab effects, $B_1 \sim N(0, 100\tau^2)$, and $B_2, \ldots, B_n \sim N(0, \tau^2)$, and Gaussian measurement errors with different but known variances.

The efficiency of the estimator of μ for the Laplace random effects model introduced in Section 2, relative to its counterpart assuming Gaussian random effects, corresponding to these four sampling situations, and based on half a million samples each, was: GAU, 66% (quite close to the median's of a homogeneous Gaussian sample); LAP, 130%; SLA 690%; and WIL, 520%. These relative efficiencies are defined as the squared ratio of MADs for corresponding sets of estimates. The efficiency loss for scenario GAU is more than offset by the vast gains for SLA and WIL, which model situations likely to be encountered in practice.

That same weighted median as an estimator of μ incurs only a small loss in efficiency relative to the MLE (for the marginal likelihood function which involves products of functions p defined in Eq. (11), shifted to be centered at μ, and was maximized using Nelder–Mead's procedure): its variance is about 1.1 larger than the MLE's for all four sampling situations considered above.

Fig. 5 shows the performance of the posterior median $\widetilde{b}_i$ as estimate of the unilateral DoE, and Fig. 6 does so for (15) the estimate $u(\widetilde{b}_i)$.

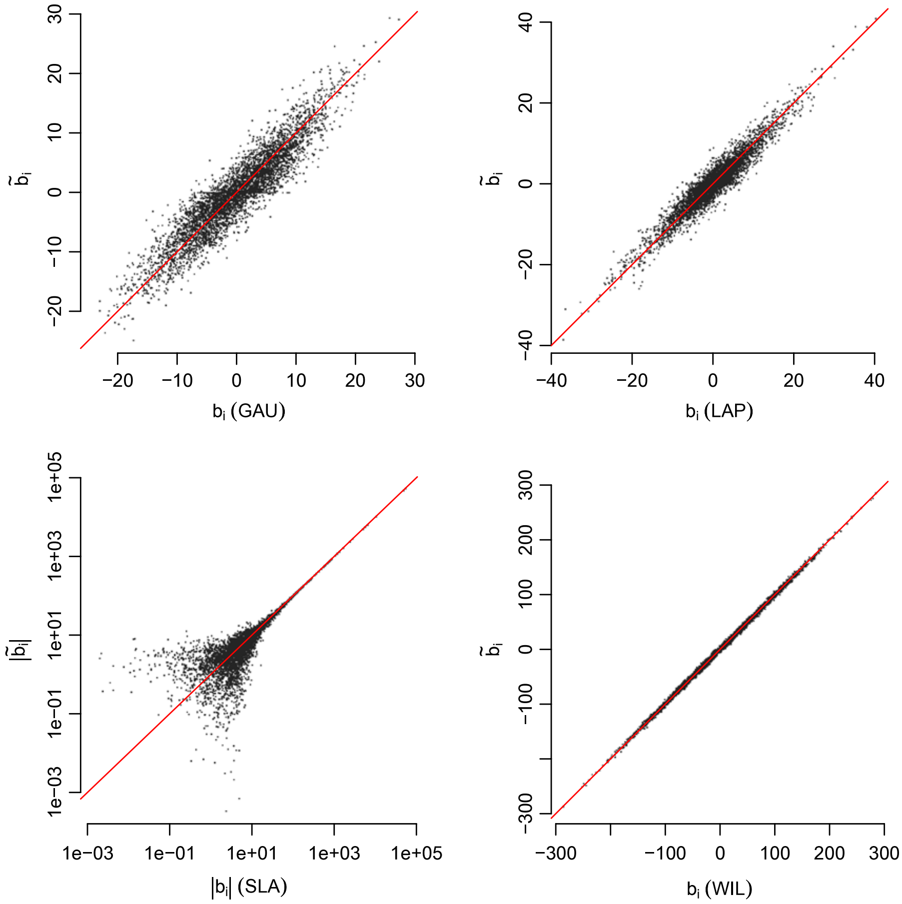

Fig. 5. The panels show plots of the estimate $\widetilde{b}_i$ in (13) of the unilateral DoE, against the corresponding, actual random effect, for the four sampling situations (GAU, LAP, SLA, WIL). The plot corresponding to SLA depicts the absolute values of the DoEs and of the random effects for both axes having logarithmic scales. The red lines have unit slope and zero intercept. The plots depict only a small sample (1%) of the simulation results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

The suggested Laplace random effects model (2) for measurements obtained in collaborative studies leads to fairly simple, robust estimators for the common median and for its standard error. An extension of this model designed to accommodate situations in which standard errors are accompanied by the number of degrees of freedom is discussed.

Predictors of the laboratory effects and of their pairwise differences interpretable as the degrees of equivalence are obtained along with their standard errors. For all these procedures we derive explicit formulas which allow direct evaluation. A Monte Carlo simulation study testifies to their good sampling performance.

Appendix

A.1. Unilateral degrees of equivalence

The formula (12) follows from the fact that

$$\bar{b}_i = \int_0^\infty [1 - G_i(t) - G_i(-t)]\mathrm{d}t = \frac{\operatorname{sgn}(x_i-\mu)}{4p_i(x_i-\mu)}\left[\frac{\mathrm{e}^{-|x_i-\mu|/\beta}}{u_i+\beta}(|x_i-\mu|+\gamma_i) - \frac{\gamma_i \mathrm{e}^{-|x_i-\mu|/u_i}}{u_i+\beta} + \frac{\mathrm{e}^{-|x_i-\mu|/u_i}}{u_i\beta}\int_0^{|x_i-\mu|} t e^{t(1/u_i-1/\beta)}\mathrm{d}t\right].$$

One has

$$u(\widetilde{b}_i) = \int_{-\infty}^{\infty} |t| g_i(t)\mathrm{d}t = \int_0^\infty [1 - G_i(t) + G_i(-t)]\mathrm{d}t$$
$$= \frac{1}{4p_i(x_i-\mu)}\left[\frac{\gamma_i \mathrm{e}^{-|x_i-\mu|/u_i}}{u_i+\beta} + \frac{\mathrm{e}^{-|x_i-\mu|/\beta}}{u_i+\beta}\left(|x_i-\mu| + \frac{u_i\beta}{u_i+\beta}\right) + \frac{\mathrm{e}^{-|x_i-\mu|/u_i}}{u_i\beta}\int_0^{|x_i-\mu|} t e^{t(1/u_i-1/\beta)}\mathrm{d}t\right]$$

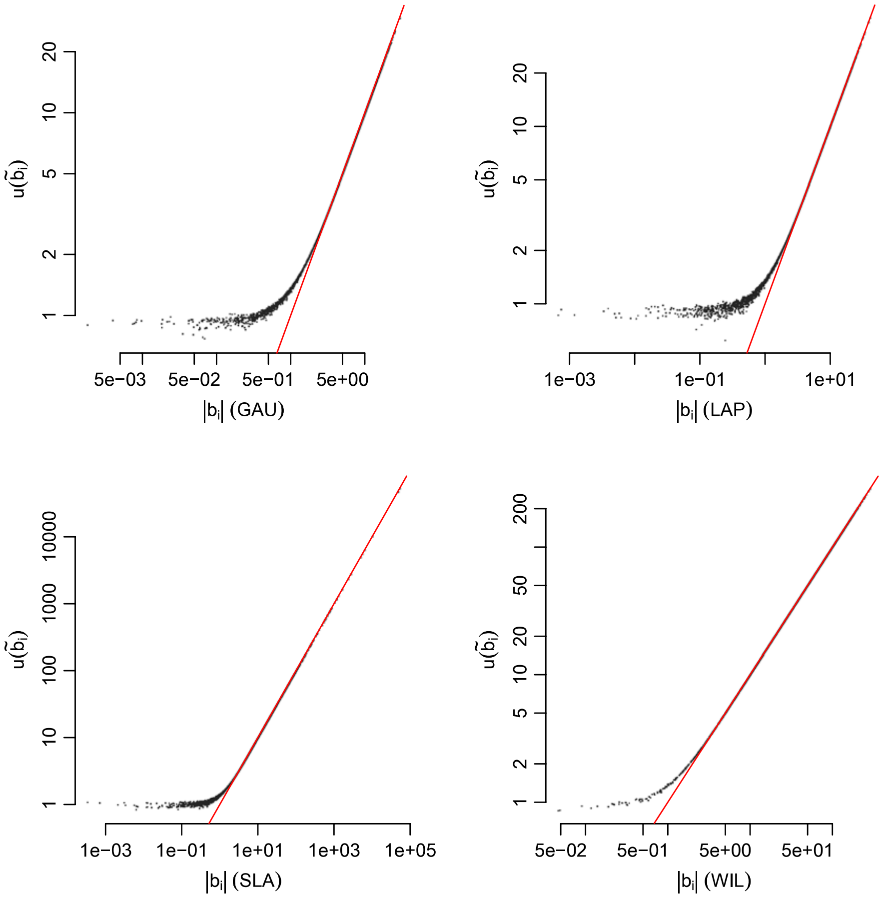

Fig. 6. The panels show plots of the estimate $u(\widetilde{b}_i)$ in (15) of the standard error of the unilateral DoE, against the absolute value of the corresponding, actual random effect, for the four sampling situations (GAU, LAP, SLA, WIL). All the axes have logarithmic scales. The red lines have unit slope and zero intercept. The plots depict only a small sample (1%) of the simulation results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$= \frac{\beta}{2p_i(x_i-\mu)(\beta^2-u_i^2)}\left[e^{-|x_i-\mu|/\beta}(|x_i-\mu|+\beta)-2(u_i^2+\beta^2)p_i(x_i-\mu)\right]$$
$$= |\widetilde{b}_i| + \frac{\gamma_i e^{-|x_i-\mu|/u_i}}{2(u_i+\beta)p_i(x_i-\mu)}, \tag{15}$$

and

$$\begin{aligned} u^2(\bar{b}_i) &= \frac{\int_{-\infty}^{\infty} t^2 g_i(t)dt}{2} = \int_0^{\infty} t[1-G_i(t)+G_i(-t)]dt \\ &= \frac{1}{4p_i(x_i-\mu)}\left[\frac{\gamma_i^2 e^{-|x_i-\mu|/u_i}}{u_i+\beta} + \frac{e^{-|x_i-\mu|/u_i}}{2u_i\beta}\int_0^{|x_i-\mu|} t^2 e^{t(1/u_i-1/\beta)}dt \right. \\ &\quad \left. + \frac{e^{-|x_i-\mu|/\beta}}{u_i+\beta}\left(\frac{(x_i-\mu)^2}{2}+\gamma_i|x_i-\mu|+\gamma_i^2\right)\right] \\ &= \frac{1}{4p_i(x_i-\mu)}\left[e^{-|x_i-\mu|/\beta}\left((x_i-\mu)^2\left(\frac{1}{2(u_i+\beta)}+\frac{1}{\beta-u_i}\right)+|x_i-\mu|\left(\frac{\gamma_i}{u_i+\beta}-\frac{2u_i\beta}{(\beta-u_i)^2}\right)\right.\right. \\ &\quad \left.\left. + u_i\beta^2\left(\frac{1}{(u_i+\beta)^2}-\frac{2}{(\beta-u_i)^2}\right)\right) - 2u_i\beta^2\left(\frac{\beta-u_i}{(u_i+\beta)^2}-\frac{2(\beta+u_i)}{(\beta-u_i)^2}\right)p_i(x_i-\mu)\right]. \end{aligned} \tag{16}$$

If $\beta = u_i$, then

$$\bar{b}_i = \widetilde{b}_i = \frac{(x_i-\mu)e^{-|x_i-\mu|/u_i}}{8u_i^2 p_i(x_i-\mu)}(u_i+|x_i-\mu|) = \frac{x_i-\mu}{2},$$
$$u(\widetilde{b}_i) = \frac{e^{-|x_i-\mu|/u_i}}{8p_i(x_i-\mu)}\left(1+\frac{|x_i-\mu|}{u_i}+\frac{|x_i-\mu|^2}{u_i^2}\right) = \frac{|x_i-\mu|^2+u_i|x_i-\mu|+u_i^2}{2(|x_i-\mu|+u_i)},$$

$$u^2(\bar{b}_i) = \frac{2|x_i-\mu|^3 + 3|x_i-\mu|^2 u_i + 3|x_i-\mu| u_i^2 + 3u_i^3}{12(|x_i-\mu| + u_i)}.$$

If $u_i \approx 0$, then

$$p_i(x-\mu) \approx \frac{e^{-|x-\mu|/\beta}}{2\beta},$$
$$\bar{b}_i \sim \widetilde{b}_i \approx x_i - \mu,$$
$$u(\widetilde{b}_i) \approx |x_i-\mu|,$$
$$u(\bar{b}_i) \approx \frac{|x_i-\mu|}{\sqrt{2}}.$$

For small $|x_i - \mu|$

$$p_i(x_i-\mu) \approx \frac{1}{2(u_i+\beta)} - \frac{(x_i-\mu)^2}{4u_i\beta(u_i+\beta)},$$
$$\bar{b}_i \sim \widetilde{b}_i \approx \frac{\beta(x_i-\mu)}{u_i+\beta},$$
$$u(\widetilde{b}_i) \approx \frac{u_i\beta}{u_i+\beta} + \frac{(x_i-\mu)^2}{2u_i},$$
$$u^2(\bar{b}_i) \approx \frac{(u_i\beta)^2}{(u_i+\beta)^2} + \frac{\beta(x_i-\mu)^2}{2(u_i+\beta)}.$$

When $|x_i - \mu| \to \infty$,

$$\bar{b}_i \approx \begin{cases} x_i - \mu - \dfrac{\mathrm{sgn}(x_i-\mu)\beta u_i}{\beta - u_i} \log\left(\dfrac{u_i+\beta}{\beta}\right) & u_i < \beta, \\ \dfrac{x_i-\mu}{2} & u_i = \beta, \\ \dfrac{\mathrm{sgn}(x_i-\mu)\beta u_i}{u_i-\beta} \log\left(\dfrac{u_i+\beta}{u_i}\right) & u_i > \beta, \end{cases}$$

$$\widetilde{b}_i \approx \begin{cases} x_i - \mu - \dfrac{2\mathrm{sgn}(x_i-\mu)\beta u_i^2}{\beta^2 - u_i^2} & u_i < \beta, \\ \dfrac{x_i-\mu}{2} & u_i = \beta, \\ \dfrac{\mathrm{sgn}(x_i-\mu)\beta(u_i^2+\beta^2)}{u_i^2-\beta^2} & u_i > \beta, \end{cases}$$

$$u(\widetilde{b}_i) \approx \begin{cases} |x_i-\mu| - \dfrac{2\beta u_i^2}{\beta^2-u_i^2} & u_i < \beta, \\ \dfrac{|x_i-\mu|}{2} & u_i = \beta, \\ \dfrac{\beta(u_i^2+\beta^2)}{u_i^2-\beta^2} & u_i > \beta, \end{cases}$$

$$u(\bar{b}_i) \approx \begin{cases} \left[(x_i-\mu)^2 - \dfrac{4(x_i-\mu)\beta u_i^2}{\beta^2-u_i^2} + \dfrac{2\beta^2 u_i^2(\beta^2+3u_i^2)}{(\beta^2-u_i^2)^2}\right]^{1/2} & u_i < \beta, \\ \left[\dfrac{(x_i-\mu)^2}{6} - \dfrac{|x_i-\mu|\beta}{12} + \dfrac{\beta^2}{6}\right]^{1/2} & u_i = \beta, \\ \dfrac{\beta u_i\sqrt{u_i^2+3\beta^2}}{u_i^2-\beta^2} & u_i > \beta. \end{cases}$$

A.2. Bilateral degrees of equivalence

We give here the exact form of $U(i,j)$ when $x_i - \mu$ and $x_j - \mu$ have different signs, say, $x_j - \mu \le 0 \le x_i - \mu$. Integration by parts shows that

$$2\int_0^\infty \left[(1-G_i(t))(1-G_j(t))+G_i(-t)G_j(-t)\right]\mathrm{d}t$$
$$=\frac{\mathrm{e}^{-|x_j-\mu|/u_j}}{2(u_j+\beta)p_j(x_j-\mu)}\int_0^\infty \mathrm{e}^{-t/\gamma_j}[1-G_i(t)]\mathrm{d}t+\frac{\mathrm{e}^{-|x_i-\mu|/u_i}}{2(u_i+\beta)p_i(x_i-\mu)}\int_0^\infty \mathrm{e}^{-t/\gamma_i}G_j(-t)\mathrm{d}t$$
$$=\frac{\gamma_j\mathrm{e}^{-|x_j-\mu|/u_j}[1-G_i(0)]}{2(u_j+\beta)p_j(x_j-\mu)}+\frac{\gamma_i\mathrm{e}^{-|x_i-\mu|/u_i}G_j(0)}{2(u_i+\beta)p_i(x_i-\mu)}-\frac{\gamma_j\mathrm{e}^{-|x_j-\mu|/u_j}}{2(u_j+\beta)p_j(x_j-\mu)}\int_0^\infty \mathrm{e}^{-t/\gamma_j}g_i(t)\mathrm{d}t$$
$$-\frac{\gamma_i\mathrm{e}^{-|x_i-\mu|/u_i}}{2(u_i+\beta)p_i(x_i-\mu)}\int_0^\infty \mathrm{e}^{-t/\gamma_i}g_j(-t)\mathrm{d}t.$$

One gets, with $1/\eta=2/\beta+1/u_i+1/u_j=1/\gamma_i+1/\gamma_j$,

$$\int_0^\infty \mathrm{e}^{-t/\gamma_i}g_j(-t)\mathrm{d}t=\int_{-\infty}^\infty \mathrm{e}^{-|t|/\gamma_i}g_j(t)\mathrm{d}t-\frac{\eta\mathrm{e}^{-|x_j-\mu|/u_j}}{4u_j\beta p_j(x_j-\mu)}$$
$$=\frac{\delta_i\bar{p}_{ij}(x_j-\mu)}{\beta p_j(x_j-\mu)}-\frac{\eta\mathrm{e}^{-|x_j-\mu|/u_j}}{4u_j\beta p_j(x_j-\mu)}.$$

Here $1/\delta_i=1/\beta+1/\gamma_i=2/\beta+1/u_i$ and $\bar{p}_{ij}$ is the density (11) defined by the parameters u_j and δ_i. Using a similar notation $\bar{p}_{ji}$ for (11) with parameters u_i and δ_j, we obtain from (15)

$$U(i,j)=|\widetilde{b}_i|+|\widetilde{b}_j|+\frac{u_j\delta_j\bar{p}_{ji}(x_i-\mu)\mathrm{e}^{-|x_j-\mu|/u_j}}{2(u_j+\beta)^2p_i(x_i-\mu)p_j(x_j-\mu)}$$
$$+\frac{u_i\delta_i\bar{p}_{ij}(x_j-\mu)\mathrm{e}^{-|x_i-\mu|/u_i}}{2(u_i+\beta)^2p_i(x_i-\mu)p_j(x_j-\mu)}+\frac{\eta\mathrm{e}^{-|x_j-\mu|/u_j}\mathrm{e}^{-|x_i-\mu|/u_i}}{4(u_j+\beta)(u_i+\beta)p_i(x_i-\mu)p_j(x_j-\mu)}.$$

For example, when $|x_i-\mu|$ and $|x_j-\mu|$ are small,

$$U(i,j)\approx\frac{\gamma_i^2+\gamma_i\gamma_j+\gamma_j^2}{\gamma_i+\gamma_j}+\frac{(x_i-\mu)^2\gamma_i}{2u_i(\gamma_i+\gamma_j)}+\frac{(x_j-\mu)^2\gamma_j}{2u_j(\gamma_i+\gamma_j)}-\frac{(x_i-\mu)(x_j-\mu)\gamma_i\gamma_j(1+(u_iu_j)/\beta^2)}{(\gamma_i+\gamma_j)(u_iu_j)^2},$$

and

$$V(i,j)\approx\sqrt{\gamma_i^2+\gamma_j^2+\frac{(x_i-\mu)^2\beta}{2(u_i+\beta)}+\frac{(x_j-\mu)^2\beta}{2(u_j+\beta)}-\frac{(x_i-\mu)(x_j-\mu)\beta^2}{(u_i+\beta)(u_j+\beta)}}.$$

Thus, for small $|x_i-\mu|, |x_j-\mu|$,

$$1\le\frac{U(i,j)}{V(i,j)}\le\frac{1.5}{\sqrt{2}}\approx 1.0607.$$

If $|x_i-\mu|\to\infty$, $|x_j-\mu|\to\infty$, with $x_i-\mu, x_j-\mu$ having the same sign and $\beta<\min(u_i,u_j)$,

$$U(i,j)\approx\beta\left[\frac{u_i^2+\beta^2}{u_i^2-\beta^2}+\frac{u_j^2+\beta^2}{u_j^2-\beta^2}-\frac{2u_i^2u_j^2+\beta^2(u_i^2+u_j^2+4u_iu_j)}{4u_i^2u_j^2-\beta^2(u_i+u_j)^2}\right].$$

If $x_i-\mu$ and $x_j-\mu$ have different signs,

$$U(i,j)\approx\beta\left[\frac{u_i^2+\beta^2}{u_i^2-\beta^2}+\frac{u_j^2+\beta^2}{u_j^2-\beta^2}-\frac{2u_i^2u_j^2+\beta^2(u_i^2+u_j^2-4u_iu_j)}{4u_i^2u_j^2-\beta^2(u_i-u_j)^2}\right].$$

In either of these cases,

$$V(i,j)\approx\beta\times\sqrt{\frac{u_i^2(u_i^2+3\beta^2)}{(u_i^2-\beta^2)^2}+\frac{u_j^2(u_j^2+3\beta^2)}{(u_j^2-\beta^2)^2}-\frac{\mathrm{sgn}\left((x_i-\mu)(x_j-\mu)\right)(u_i^2+\beta^2)(u_j^2+\beta^2)}{(u_i^2-\beta^2)(u_j^2-\beta^2)}}.$$

Thus, if $x_i-\mu, x_j-\mu$ are of the same sign, $|x_i-\mu|\to\infty$, $|x_j-\mu|\to\infty$, and $\beta<\min(u_i,u_j)$,

$$1\le\frac{U(i,j)}{V(i,j)}\le\frac{3}{2},$$

otherwise

$$0.8660\approx\frac{\sqrt{3}}{2}\le\frac{U(i,j)}{V(i,j)}\le\frac{2}{\sqrt{3}}\approx 1.1547.$$

The largest relative discrepancy between $U(i,j)$ and $V(i,j)$ (at least $\sqrt{2}$) is when $x_i=\mu$, $x_j-\mu\to\infty$, while $u_i,u_j\to 0$, but this is not a very realistic scenario.

References

Analytical Methods Committee (Royal Society of Chemistry), 1989a. Robust statistics — how not to reject outliers. Part 1. Basic concepts. Analyst 114, pp. 1693–1697.
Analytical Methods Committee (Royal Society of Chemistry), 1989b. Robust statistics — how not to reject outliers. Part 2. Inter-laboratory trials. Analyst 114, pp. 1699–1702.
Bloomfield, P., Steiger, W.L., 1983. Least Absolute Deviations: Theory, Applications and Algorithms. Birkhauser, Boston.
Comité International des Poids et Mesures (CIPM), 1999. Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes, BIPM, Sèvres, France.
Cox, M.G., 2002. The evaluation of key comparsion data. Metrologia 39, 589–595.
Cox, M.G., 2007. The evaluation of key comparsion data: determining the largest consistent subset. Metrologia 44, 187–200.
Davies, L., 1991. A stochastic model for interlaboratory tests. Computational Statistics and Data Analysis 12, 201–209.
Decker, J.E., Brown, N., Cox, M.G., Steele, A.G., Douglas, R.J., 2006. Recent recommendations of the Consultative Committee for Length (CCL) regarding strategies for evaluating key comparison data. Metrologia 43, L51–L55.
Demidenko, E., 2004. Mixed Models. Wiley, New York.
Duewer, D.L., 2006. A robust approach for the determination of CCQM Key Comparison reference values and uncertainties, Working document CCQM/04-15. http://www.bipm.info/cc/CCQM/Allowed/10/CCQM04-15.pdf.
Duewer, D.L., 2008. A comparison of location estimators for interlaboratory data contaminated with value and uncertainty outliers. Accreditation and Quality Assurance 13, 193–216.
Geyer, C.J., 2009. mcmc: Markov Chain Monte Carlo. R package version 0.6. http://CRAN.R-project.org/package=mcmc.
Gurwitz, C., 1990. Weighted median algorithms for L_1 approximation. BIT 30, 301–310.
Hedges, L.V., Olkin, I., 1985. Statistical Methods for Meta-Analysis. Academic Press, London.
ISO/IEC, 2007. International vocabulary of metrology — basic and general concepts and associated terms (VIM), ISO/IEC Guide 99, International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC), Geneva, Switzerland.
Koenker, R., 2005. Quantile Regression. Cambridge Univeristy Press, Cambridge UK.
Kotz, S., Kozubowski, T.J., Podgorski, K., 2001. The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering and Finance. Birkhauser, Boston.
Lischer, D., 1996. Robust statistical methods in interlaboratory analytical studies. In: Rieder, H. (Ed.), Robust Statistics, Data Analysis, and Computer Intensive Methods. In: Lecture Notes in Statistics, Springer, New York, pp. 251–265.
Müller, J.W., 2000. Possible advantages of a robust evaluation of comparisons. Journal of Research of the National Institute of Standards and Technology 105, 551–555.
Pinheiro, J.C., Liu, C., Wu, Y.N., 2001. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. Journal of Computational and Graphical Statistics 10, 249–276.
Poor, H., 1994. An Introduction to Signal Detection and Estimation. Springer, New York.
R Development Core Team, 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
Ratel, G., 2006. Median and weighted median as estimators for the key comparison reference value (KCRV). Metrologia 43, S244–S248.
Rocke, D.M., 1983. Robust statistical analysis of interlaboratory studies. Biometrika 70, 421–431.
Schantz, M., Wise, S., 2004. CCQM–K25: determination of PCB congeners in sediment. Metrologia 41, Technical Supplement, 08001.
Thompson, M., Ellison, S.L.R., Wood, R., 2006. The international harmonized protocol for the proficiency testing of analytical chemistry laboratories. Pure and Applied Chemistry 78, 145–196.
Toman, B., Possolo, A., 2009. Laboratory effects models for interlaboratory comparisons. Accreditation and Quality Assurance 14, 553–563.
Wilcox, R.R., 2005. Introduction to Robust Estimation and Testing, 2nd ed. Academic Press, San Diego, CA.
Wilcox, R.R., 2006. Comparing medians. Computational Statistics & Data Analysis 51, 1934–1943.