Fiducial approach to uncertainty assessment accounting for error due to instrument resolution

2007 Metrologia 44 476

(http://iopscience.iop.org/0026-1394/44/6/006)

Metrologia 44 (2007) 476–483 doi:10.1088/0026-1394/44/6/006

Fiducial approach to uncertainty assessment accounting for error due to instrument resolution

Jan Hannig[1,2], Hari K Iyer[1,2] and C M Wang[1]

[1] Statistical Engineering Division, National Institute of Standards and Technology, Boulder, CO 80305, USA
[2] Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA

E-mail: jwang@boulder.nist.gov

Received 16 July 2007
Published 7 November 2007
Online at stacks.iop.org/Met/44/476

Abstract
This paper presents an approach for making inference on the parameters μ and σ of a Gaussian distribution in the presence of resolution errors. The approach is based on the principle of fiducial inference and requires a Monte Carlo method for computing uncertainty intervals. A small simulation study is carried out to evaluate the performance of the proposed procedure and compare it with some of the existing procedures. The results indicate that the fiducial procedure is comparable to the best of the competing procedures for inference on μ. However, unlike some of the competing procedures, the same Monte Carlo calculations also provide inference for σ and many other related quantities of interest.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

One common source of uncertainty in metrological applications is the one due to limited resolution of instruments. The error in the *result* that is attributable to limited instrument resolution or quantization during analog-to-digital conversion is sometimes referred to as resolution error or quantization error. We say that an instrument has resolution d (in appropriate units) if the absolute difference between the actual value and the instrument reading could be as high as $d/2$. That is, if the instrument reading is y then the actual value is known to be between $y - d/2$ and $y + d/2$. No finer information about the value is available. Early attempts to account for resolution error took the approach of treating resolution error as a random quantity having a uniform distribution within the interval $(-d/2, d/2)$. This led to a model of the form

$$X_i = \mu + \sigma Z_i + dU_i, \qquad i = 1, \ldots, n, \qquad (1)$$

where μ is the measurand, σ is the standard deviation of the measurement process under repeatability conditions, Z_i are independent standard Gaussian errors, d is instrument resolution and U_i are independent uniform $(-1/2, 1/2)$ variates. An implicit assumption is that Z_i and U_i are independent. Using such a model Lira and Wöger [1] proposed the use of the formula

$$s_{\bar{x}} = \sqrt{\frac{s_x^2}{n} + \frac{d^2}{12}}$$

for the standard uncertainty of $\bar{x}$. Here s_x^2 is the sample variance of the measured values.

However, Lira [2] correctly pointed out that Z_i and U_i cannot be independent and provided a method, using a Bayesian approach, for calculating the standard uncertainty of a quantity using discretized measurement data. Prior to Lira's work, Elster [3] had proposed a Bayesian treatment of this resolution error problem and provided a solution that did not make any distributional assumptions about the resolution errors. Lira [2] remarked that the Elster solution is computationally unsuitable for routine work and provided an approximation that is easy to use. Frenkel and Kirkup [4] considered this resolution error problem and provided charts that can be used to estimate the unknown variance σ^2 from the observed variance s_x^2. Taraldsen [5] and Willink [6] have also presented simple alternative approaches for the problem.

In this paper we propose a different approach for making inference on the parameters μ and σ of a Gaussian distribution in the presence of resolution errors. Our approach is based on generalizations of *fiducial* inference proposed by Fisher [7] and *structural inference* introduced by Fraser [8]. The paper is organized as follows. In section 2 we review the measurement model in the presence of resolution or quantization errors. The main problem is that of estimating the measurand μ and obtaining an appropriate uncertainty interval. A related problem of importance is the estimation of σ along with an uncertainty interval. This latter estimate provides information about the variability of the measurement process in the absence of resolution errors. Section 3 contains an outline of the fiducial method and gives simple illustrations of how the method is implemented. We then briefly describe a technical formulation of a generalized fiducial recipe for deriving uncertainty intervals for parameters in general problems. Section 4 uses two examples to illustrate the fiducial uncertainty intervals for μ and compares them with the uncertainty intervals obtained based on the existing procedures mentioned in this section. Section 5 uses a Monte Carlo study to examine the coverage properties of the fiducial uncertainty intervals under a variety of scenarios involving different values for the measurand, the experimental standard deviation and the instrument resolution, namely, μ, σ and d, respectively. Section 6 discusses two other related problems of interest in metrology that can also be solved using the fiducial approach. Finally, we conclude with some summary remarks in section 7.

2. Measurement model with resolution error

If an instrument has a resolution of d units, the measured value can only be reported as the closest multiple of d. If x is the measured value, then there is a unique integer k such that $kd - d/2 < x \leqslant kd + d/2$ and we define $\psi(x) = kd$. It follows that

$$x - d/2 < \psi(x) \leqslant x + d/2.$$

As in [5], we let the independent random variables $X_1, \ldots, X_n$ represent the measurement process with perfect resolution. The reported values $Y_1, \ldots, Y_n$ are then given by $Y_i = \psi(X_i)$. The measurement model is

$$Y_i = \psi(X_i) = \psi(\mu + \sigma Z_i), \qquad i = 1, \ldots, n. \tag{2}$$

In this paper we use an extension of the fiducial argument of Fisher, described by Hannig [9], to obtain the fiducial distribution of (μ, σ) for the measurement model in (2). This approach is described in the next section. Earlier works on *generalized confidence intervals* and *generalized inference* [10–13] are special cases of generalized fiducial inference. See Hannig [9] for a detailed discussion of generalized fiducial inference.

3. Generalized fiducial inference

The following three simple examples serve to illustrate the basic ideas of fiducial inference.

3.1. Example 1

Suppose $X \sim N(\mu, 1)$ where μ is the measurand, the measurement process has a known variance equal to 1, and X is the random variable representing values that may be observed. One might express the relationship between the measured values and the underlying random experimental error process by the following equation:

$$X = \mu + E, \tag{3}$$

where E is a random error with $N(0, 1)$ distribution. Each measured value is associated with a particular random experimental error. Suppose a single measurement is made and its value is 10. The associated measurement error is denoted by e. So

$$10 = \mu + e.$$

Hence $\mu = 10-e$. If the value of e were known, then we would know the measurand exactly, but the value of e is not known. Nevertheless, the fact that we know the distribution from which e was generated helps us determine a set of values of μ that we consider plausible. For instance, how plausible is the value $\mu = 2$ for the measurand? For this to be true we need $e = 8$. A value of 8 is highly unlikely to have come from a $N(0, 1)$ distribution. So we conclude that the value $\mu = 2$ is highly unlikely. How likely is it that μ is between 10 and 12? For μ to be between 10 and 12, e needs to be between 0 and 2 and we can calculate the probability for this to be $\Phi(2) - \Phi(0)$, where $\Phi(z)$ is the value of the cumulative standard normal distribution at z. Thus, probabilities associated with E can be transferred to probabilities for μ. Our knowledge about μ, based on the measured value of 10, can be described by the distribution of the random variable $\tilde{\mu}$ whose distribution is given by that of $10 - E$. That is, $\tilde{\mu} \sim N(10, 1)$. We say that the fiducial distribution of μ (that is, the distribution of $\tilde{\mu}$) is $N(10, 1)$.

The fiducial distribution of μ may be used to obtain a coverage interval for μ. An equal-tails 95% probability interval for $\tilde{\mu}$ is

$$P(10 - z_{0.975} \leqslant \tilde{\mu} \leqslant 10 + z_{0.975}) = 0.95,$$

where z_γ is the 100γ percentile of a standard normal distribution. This probability interval for $\tilde{\mu}$ may be interpreted as a coverage interval for μ and we say that $(10 - z_{0.975}, 10 + z_{0.975})$ is a 95% fiducial coverage interval for μ. Note that this is exactly the interval one would get using either the classical frequentist approach or a Bayesian approach with a *non-informative* improper prior for μ.

3.2. Example 2

In the above example, suppose we consider making two measurements. Let X_1 and X_2 be the random variables denoting the possible values one might obtain for the two measurements. We can write

$$X_1 = \mu + E_1,$$

$$X_2 = \mu + E_2. \tag{4}$$

Suppose the actual measurements are 10 and 8. We then have the following equations relating the measured values, the measurand and the realized values of experimental errors, say e_1 and e_2.

$$10 = \mu + e_1,$$

$$8 = \mu + e_2.$$

Plausible values for μ are related to plausible values of (e_1, e_2). What makes this example different from the first example is that here we *know* $e_1 - e_2$ equals 2. So the universe of possible values for (e_1, e_2) is now limited by this requirement. We know (e_1, e_2) is from a standard bivariate Gaussian distribution but is constrained to lie on the line $e_1 - e_2 = 2$. So the probabilities one would associate with μ are the probabilities one would associate with either $10 - e_1$ (since $\mu = 10 - e_1$) or $8 - e_2$ (since μ is also equal to $8 - e_2$), knowing that (e_1, e_2) is a realization from a bivariate standard Gaussian distribution subject to the additional condition that $e_1 - e_2 = 2$. Hence we define the random variable $\tilde{\mu}$ to have a distribution that is equal to the conditional distribution of $10 - E_1$ given that $E_1 - E_2 = 2$. This is the same distribution as the conditional distribution of $8 - E_2$ given that $E_1 - E_2 = 2$. A simple calculation tells us that the distribution of $\tilde{\mu}$ is $N(\bar{x}, 1/2)$ where $\bar{x} = (x_1 + x_2)/2 = (10 + 8)/2 = 9$. Thus, a 95% fiducial coverage interval for μ is $(9 - z_{0.975}/\sqrt{2}, 9 + z_{0.975}/\sqrt{2})$. Again, this is the same interval that one would obtain by the classical frequentist method or a Bayesian method with a non-informative improper prior.

Interestingly, this argument is fully generalizable and one can develop fiducial distributions for model parameters in very general problems. The starting point for this process is what we call a *structural equation*. In example 1, equation (3) constitutes the structural equation. In example 2, equations (4) constitute the structural equations. The structural equations relate the observations with model parameters and error processes whose distributions are fully known. For instance, in example 1 we know the distribution of E completely.

In examples 1 and 2 we have assumed that the variance of the measurement error process is known to be 1. This is for simplicity only. One can assume that the measurement error process has an unknown variance σ^2. In this case, for example 2, we will start with the following structural equations:

$$X_1 = \mu + \sigma E_1,$$

$$X_2 = \mu + \sigma E_2.$$

More generally, if we have n measurements of μ, then we can write

$$X_1 = \mu + \sigma E_1,$$

$$X_2 = \mu + \sigma E_2,$$

$$\dots,$$

$$X_n = \mu + \sigma E_n,$$

where $E_1, \dots, E_n$ are independent, standard Gaussian random variables. It is well known that the fiducial argument will yield the following fiducial distribution for μ:

$$\tilde{\mu} \sim \bar{x} - \frac{s}{\sqrt{n}} T_{n-1}, \tag{5}$$

namely, a shifted and scaled Student's t distribution with $n-1$ degrees of freedom. Here $\bar{x}$ is the mean of the n measured values and s is the sample standard deviation of these n measurements. The coverage interval for μ turns out to be the same as the standard Student's t intervals.

The fiducial distribution for μ in (5) can also be obtained with a simpler approach based on minimal sufficient statistics $\bar{X}$ and S^2. Since $\bar{X} \sim N(\mu, \sigma^2/n)$ and $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$, where $\chi^2(\nu)$ stands for the chi-squared distribution with ν degrees of freedom, we have the following two structural equations:

$$\bar{X} = \mu + \frac{\sigma}{\sqrt{n}} E,$$

$$S^2 = \frac{\sigma^2}{n-1} V,$$

where E is standard normal and V is chi-squared with $n-1$ degrees of freedom random variables. By solving the structural equations, we obtain the fiducial distribution for μ in (5). This is the approach used in [10] and [11].

One may wonder if the fiducial method will lead to anything different from standard results. The answer is yes. In fact, except in very simple situations such as those discussed above, the fiducial method will differ from other methods. The theoretical properties of coverage intervals derived from fiducial distributions are discussed in detail in [9, 13]. Many articles have also examined, via statistical simulation, the coverage properties of generalized confidence intervals in a variety of applications, and as shown in [13], these generalized confidence intervals are all, in fact, fiducial coverage intervals.

The next example develops the fiducial reasoning for the model that is the topic of this paper, namely, the measurement model with resolution errors in addition to random experimental errors.

3.3. Example 3

The goal of this example is to illustrate how the fiducial argument is applied to the measurement error model with resolution error and so we consider a simple situation with only three measurements of μ, the measurand. Suppose $d = 1$ is the resolution (by appropriate choice of units, d can always be taken to be 1). Let Y_1, Y_2, Y_3 be the random variables representing the three measurements. Suppose the observed values of these random variables are $y_1 = 4$, $y_2 = 5$ and $y_3 = 6$. The structural equations are

$$\begin{aligned} Y_1 &= \psi(\mu + \sigma E_1), \\ Y_2 &= \psi(\mu + \sigma E_2), \\ Y_3 &= \psi(\mu + \sigma E_3), \end{aligned} \tag{6}$$

where $\psi(\cdot)$ is the function defined in section 2.

From the observed values we know

$$\begin{aligned} 3.5 &\leqslant \mu + \sigma e_1 < 4.5, \\ 4.5 &\leqslant \mu + \sigma e_2 < 5.5, \\ 5.5 &\leqslant \mu + \sigma e_3 < 6.5, \end{aligned} \tag{7}$$

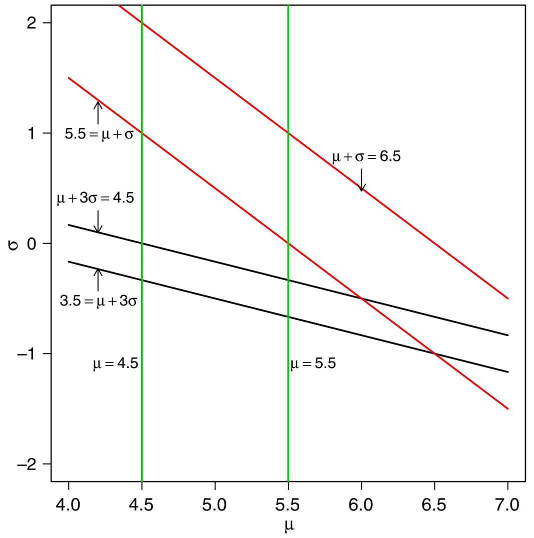

Figure 1. Empty region for (μ, σ) when $(y_1, y_2, y_3) = (4, 5, 6)$ and $(e_1, e_2, e_3) = (3, 0, 1)$, demonstrating that not all values for (e_1, e_2, e_3) are allowable.

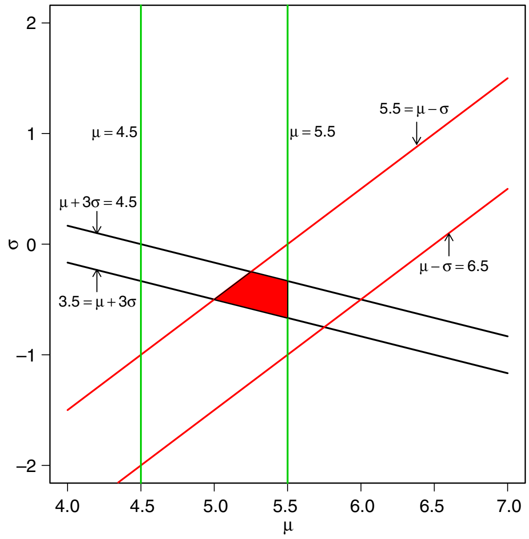

Figure 2. Solutions for (μ, σ) when $(y_1, y_2, y_3) = (4, 5, 6)$ and $(e_1, e_2, e_3) = (3, 0, -1)$, demonstrating that some values for (e_1, e_2, e_3) are allowable.

where e_1, e_2, e_3 are independently realized values from a standard Gaussian distribution. However, knowing the measured values, these equations imply that certain combinations of values for (e_1, e_2, e_3) are impossible. For instance, could the underlying errors (e_1, e_2, e_3) equal $(3, 0, 1)$? Plug in these values for the e_i into equations (7). This will imply that μ and σ must satisfy the inequalities

$$\begin{aligned} 3.5 &\leqslant \mu + 3\sigma < 4.5, \\ 4.5 &\leqslant \mu < 5.5, \\ 5.5 &\leqslant \mu + \sigma < 6.5. \end{aligned} \tag{8}$$

In particular, this implies that (μ, σ) belongs to the region in the (μ, σ) plane determined by the inequalities in (8). This region turns out to be *empty* if $(e_1, e_2, e_3) = (3, 0, 1)$. See figure 1.

So we know, on the strength of the observed values, that $(e_1, e_2, e_3) = (3, 0, 1)$ is an impossibility. On the other hand, $(e_1, e_2, e_3) = (3, 0, -1)$ leads to the following inequalities.

$$\begin{aligned} 3.5 &\leqslant \mu + 3\sigma < 4.5, \\ 4.5 &\leqslant \mu < 5.5, \\ 5.5 &\leqslant \mu - \sigma < 6.5. \end{aligned} \tag{9}$$

The points in the (μ, σ) plane that satisfy these inequalities belong to the shaded region in figure 2. In particular, $(e_1, e_2, e_3) = (3, 0, -1)$ is, in fact, a feasible realization for (E_1, E_2, E_3).

In fact, it is easy to characterize the set $\mathcal{S}$ of values of (e_1, e_2, e_3) in $\mathbf{R}^3$ which are possible realizations of the random errors given the observed data. Given that (E_1, E_2, E_3) belongs to the set $\mathcal{S}$, the inequalities in (7) define a set $\mathcal{Q} = \mathcal{Q}(E_1, E_2, E_3)$ in the (μ, σ) plane that is not empty. Let $(\tilde{\mu}(\mathcal{Q}), \tilde{\sigma}(\mathcal{Q}))$ be a randomly chosen element from the set $\mathcal{Q}$. The distribution of $(\tilde{\mu}(\mathcal{Q}), \tilde{\sigma}(\mathcal{Q}))$, conditional on the event that (E_1, E_2, E_3) belongs to the set $\mathcal{S}$, is defined to be the generalized fiducial distribution of (μ, σ).

In general, it is impossible to derive an analytic expression for the fiducial distribution of (μ, σ) in the resolution error model. However, Markov Chain Monte Carlo (MCMC) methods [14] are available such that we can obtain random samples of any size from the fiducial distribution. Hence the joint fiducial distribution of (μ, σ), as well as the marginal fiducial distributions of μ and σ, can be empirically estimated. Appropriate percentiles from these empirical distributions can be used to estimate the theoretical percentiles for computing coverage intervals. The description of an MCMC algorithm that can be used to obtain random samples from the joint fiducial distribution of (μ, σ) is given in the appendix. An R [15] program, which takes the recorded data and the minimum resolution as inputs and calculates various coverage intervals for μ and/or σ, is available from the authors on request.

3.4. Generalized fiducial recipe

We will briefly describe the generalization of the idea described above to arbitrary statistical models. Let $\boldsymbol{X}$ be a (possibly discrete) random vector with a distribution indexed by a parameter $\xi \in \Xi$. Assume that the data-generating mechanism for $\boldsymbol{X}$ could be expressed in the following form

$$\boldsymbol{X} = G(W, \xi), \tag{10}$$

where G is a function and W is a random variable or vector with a completely known distribution independent of any parameters. We call equation (10) the *structural equation.* Suppose $\boldsymbol{X}$ has been observed and the realized value is $\boldsymbol{x}$. This must correspond to some realized value of W which we denote by w. We do not know the value of w but we do know that it has to be a value such that there is a value ξ_w of ξ which, together with w, results in $\boldsymbol{x} = G(w, \xi_w)$. So, given the data, only those values of w are possible for which $\boldsymbol{x} = G(w, \xi_w)$. Denote this set of values of w by $\mathcal{S}$. For any given value of w in $\mathcal{S}$, let $Q(\boldsymbol{x}, w)$ be the set of values of ξ such that $\boldsymbol{x} = G(w, \xi)$. That is,

$$Q(\boldsymbol{x}, w) = \{\xi : \boldsymbol{x} = G(w, \xi)\}. \tag{11}$$

Finally, let $\tilde{\xi}$ denote a randomly chosen value from $Q(\boldsymbol{x}, W)$. We define a generalized fiducial distribution of ξ as the conditional distribution of $\tilde{\xi}$ given $W \in \mathcal{S}$.

The choice of a particular form of the structural equation (10) could influence the generalized fiducial distribution. This situation is well recognized in the fiducial literature. However, it is important to remark that, in many practical applications, the physical process by which the data were generated is known. In this case we can and should choose the structural equation to reflect this process, thus eliminating the problem of non-uniqueness due to the choice of structural equation. In the field of metrology where an unknown measurand is measured using some known processes, one typically knows that random errors influence the measurement in some pre-specified known fashion. The resulting measured values are expressed as an equation combining some unknown measured quantities and errors. This equation can be taken as the structural equation.

4. Examples

We use the examples in [1] to illustrate the proposed fiducial intervals for μ and compare them with the intervals obtained from other procedures. In the first example, ten measurements of some length μ are obtained using a micrometer with a resolution of 0.001 mm. The measurements (in millimetres) are 7.489, 7.503, 7.433, 7.549, 7.526, 7.396, 7.543, 7.509, 7.504 and 7.383. Table 1 displays the point estimates and the 95% uncertainty intervals for μ produced by the method that ignores the resolution information (called RI method), Lira and Wöger [1], Lira [2], Taraldsen [5], Willink [6], Bayesian and fiducial procedures. The intervals that can be analytically expressed are

$$\bar{y} \pm t_{0.975,n-1}\, s_y/\sqrt{n} \text{ (RI)},$$

$$\bar{y} \pm t_{0.975,n-1}\sqrt{s_y^2/n + d^2/12} \text{ (Lira and Wöger)},$$

$$\bar{y} \pm (d/2 + t_{0.975,n-1}(s_y/\sqrt{n} + d/\sqrt{n-1})) \text{ (Taraldsen)}$$

$$\bar{y} \pm t_{0.975,n-1}\, u(\bar{y}) \text{ (Willink)},$$

where

$$u^2(\bar{y}) = \begin{cases} d^2/12 & \text{if } y_{\max} = y_{\min}, \\ \max(s_y^2/n, [(y_{\max} + y_{\min})/2 - \bar{y}]^2/3) & \\ & \text{if } y_{\max} - y_{\min} = d, \\ s_y^2/n & \text{otherwise} \end{cases}$$

and where $y_{\max} = \max(y_1, \ldots, y_n)$ and $y_{\min} = \min(y_1, \ldots, y_n)$.

Table 1 shows that the intervals from the various methods are almost identical. This is because, when the resolution is small relative to the spread of the data, the resolution error can be safely ignored.

In the next example, ten measurements of the same length μ are taken with a caliper. The measured values (in millimetres) are 7.5, 7.5, 7.4, 7.5, 7.5, 7.4, 7.5, 7.5 7.5 and 7.4. The caliper has a resolution of 0.1 mm. Table 2 displays the point estimates and the 95% uncertainty intervals for μ produced by the same methods listed in table 1.

The simulation study described in the next section shows that the Lira and Wöger interval is moderately

Table 1. Estimates and 95% uncertainty intervals for μ with measurements obtained using a micrometer.

Method	Estimate	Conf. limits	
RI	7.484	7.441	7.526
Lira and Wöger	7.484	7.441	7.526
Lira	7.484	7.440	7.531
Taraldsen	7.484	7.440	7.527
Willink	7.484	7.441	7.526
Bayesian[1]	7.48	7.44	7.52
Fiducial[2]	7.483	7.441	7.525

[1] Taken from [5].
[2] Based on 10 000 Monte Carlo samples.

Table 2. Estimates and 95% uncertainty intervals for μ with measurements obtained using a caliper.

Method	Estimate	Conf. limits	
RI	7.47	7.44	7.50
Lira and Wöger	7.47	7.40	7.54
Lira	7.47	7.45	7.52
Taraldsen	7.47	7.31	7.63
Willink	7.47	7.44	7.50
Bayesian[1]	7.46	7.43	7.49
Fiducial[2]	7.47	7.44	7.51

[1] Taken from [5].
[2] Based on 10 000 Monte Carlo samples.

conservative and the Taraldsen interval is very conservative when the measurement resolution is not small relative to the measurement uncertainty, which is the case in this example. This explains why the intervals in table 2 can be roughly classified into three groups according to their interval widths: Taraldsen, Lira and Wöger and the rest. In the extreme case of no spread in the data, that is, all the ten measurements were 7.5 mm, the RI interval degenerates to a single point. With a resolution of 0.1 mm, the 95% Lira and Wöger, Taraldsen and fiducial intervals are (7.43, 7.57), (7.37, 7.63) and (7.45, 7.55), respectively. Also, the Lira and Willink intervals are identical to the Lira and Wöger interval.

5. Performance evaluation

We conducted a simulation study to evaluate the coverage properties of the fiducial interval for μ discussed above. The simulation parameters for this problem were μ, σ, n and d. The value of μ was fixed at 10. The parameters that were varied are n (5, 10 or 30), σ (0.01, 0.1, 0.2 or 1) and d (0.001, 0.01 or 0.1).

For each combination of σ, n and d, x_i, $i = 1, \ldots, n$, were generated from $N(10, \sigma^2)$ and rounded according to the value of d. That is, $y_i = [x_i/d]\,d$, where $[\cdot]$ indicates the nearest integer, were taken to be the rounded value of x_i. Using these y_i, 95% fiducial intervals for μ based on 10 000 fiducial samples were computed. Also, with the same generated data, 95% confidence intervals for μ were computed using competing methods. The competing methods we considered here are RI, Lira and Wöger, Taraldsen and Willink. This process was repeated 10 000 times. The percentage of times that the intervals contained $\mu = 10$ was recorded. The average lengths for the coverage intervals were also recorded.

Table 3. Coverage probabilities of the fiducial and other intervals for μ corresponding to sample size $n = 10$.

σ	d	RI	Lira and Wöger	Taraldsen	Willink	Fiducial
1.0	0.001	0.9503[a]	0.9503	0.9503	0.9503	0.9504
		1.3918[b]	1.3918	1.3943	1.3918	1.3922
1.0	0.01	0.9508	0.9508	0.9544	0.9508	0.9508
		1.3928	1.3928	1.4178	1.3928	1.3931
1.0	0.1	0.9513	0.9523	0.9788	0.9513	0.9509
		1.3949	1.4014	1.6457	1.3949	1.3950
0.2	0.1	0.9503	0.9730	0.9998	0.9503	0.9486
		0.2802	0.3105	0.5311	0.2802	0.2801
0.1	0.1	0.9574	0.9953	1.0000	0.9574	0.9479
		0.1442	0.1959	0.3950	0.1443	0.1436
0.01	0.1		1.0000	1.0000	1.0000	1.0000
			0.1306	0.2508	0.1306	0.1000

The simulated coverages of the intervals were found to be mainly dependent on the ratio d/σ. To simplify the presentation, we only report the results corresponding to $n = 10$ and some selected values of $(\sigma,\ d)$ to cover a range of practical values of d/σ. Table 3 contains the results.

Each cell in table 3 consists of two entries. The first entry ([a]) is the simulated coverage probabilities of the intervals. The second entry ([b]) is the average width of the intervals. Table 3 indicates that when d is small relative to σ, the coverage probabilities of all the five intervals are close to the stated value of 95%. However, as d/σ increases, both the Lira and Wöger and Taraldsen intervals become more conservative, especially the Taraldsen interval. In the extreme case of $d/\sigma = 0.1/0.01 = 10$, the variation of the distribution cannot be estimated from the data and the RI interval degenerates to a single point and is not included in the study. All the other intervals are conservative in this case.

Both Willink and fiducial intervals perform well under a wide range of values of d/σ. The Willink interval is simpler to compute, so if only μ is of interest, then this interval can be recommended in practice. If both μ and σ are of interest, or, if a tolerance interval or a conformance interval (see the next section) is required for the application, then the fiducial method is recommended. The strength of the fiducial approach is its ability to obtain the joint fiducial distribution of $(\mu,\ \sigma)$ and use it to make inference on many related problems, not just the construction of uncertainty intervals for μ.

6. Related problems

In the previous section we were concerned with the estimation of μ and σ and constructing coverage intervals for them. We proposed the fiducial method as a way to accomplish this. One of the strengths of the fiducial approach is that once a joint fiducial distribution for $(\mu,\ \sigma)$ has been developed, a number of related problems can also be solved without much additional work. One such related problem is the estimation of specified percentiles of the distribution of interest and another is the estimation of the probability content associated with a specified interval. We explain the significance of these problems below.

6.1. Estimation of percentiles

Suppose it is desired to characterize the distribution of a population of standard reference materials (SRMs). One way of doing this is by estimating the mean and the standard deviation of the population. Another, somewhat more useful, approach is to provide tolerance intervals for the distribution. With a tolerance interval, one makes a statement of the following type: on the basis of the data, we can claim with $100\gamma\%$ confidence that $100\beta\%$ of the SRMs will have a response value between $L(X)$ and $H(X)$ units. The interval $[L(X),\ H(X)]$ is called a β content, γ confidence tolerance interval for the distribution of interest. We apply the fiducial method for constructing such tolerance intervals.

6.2. Estimation of conformance proportions

Often it is of interest to know what proportion of the population have values between specified numbers A and B $(A < B)$. This proportion is referred to as a *conformance proportion*. Suppose we have a lot of material and an associated distribution of a characteristic of interest (say, breaking strength). Suppose only samples having values between A and B are of acceptable quality. Hence it is of interest to estimate the proportion from the lot whose values are between A and B. We apply the fiducial method to estimate the proportion and also to provide a lower bound for this proportion with a stated confidence level.

6.3. Fiducial solutions to tolerance and conformance problems

Once we have the fiducial distribution of (μ, σ), either analytically or empirically, the derivation of tolerance intervals and bounds for conformance proportion follow in a straightforward manner. In particular, if the population of interest has the distribution $N(\mu, \sigma^2)$, then a β content, γ confidence tolerance interval is given by $(L,\ H)$, where L is the $(1-\gamma)/2$ quantile of the fiducial distribution of $\mu + z_{(1-\beta)/2}\sigma$ and H is the $(1+\gamma)/2$ quantile of the fiducial distribution of $\mu + z_{(1+\beta)/2}\sigma$. These quantiles are most conveniently estimated using a Monte Carlo approach. This involves generating a large number of realizations from the fiducial distribution of (μ, σ), denoted by $(\tilde{\mu}_1, \tilde{\sigma}_1), \ldots, (\tilde{\mu}_M, \tilde{\sigma}_M)$, and determining the empirical $(1-\gamma)/2$ quantile of $\tilde{\mu}_i + z_{(1-\beta)/2}\tilde{\sigma}_i$ and $(1+\gamma)/2$ quantile of $\tilde{\mu}_i + z_{(1+\beta)/2}\tilde{\sigma}_i$, $i = 1, \ldots, M$. An example is given in the appendix.

The conformance proportion has the theoretical value equal to $\Phi((B-\mu)/\sigma) - \Phi((A-\mu)/\sigma)$ denoted by θ. Then a $1-\alpha$ lower confidence bound for θ is the α quantile of $\Phi((B-\tilde{\mu}_i)/\tilde{\sigma}_i) - \Phi((A-\tilde{\mu}_i)/\tilde{\sigma}_i)$, $i = 1, \ldots, M$.

7. Conclusion

In this paper we have provided an approach for making inference on the parameters μ and σ of a Gaussian distribution in the presence of resolution errors. Specifically, we have provided procedures for constructing uncertainty intervals for μ and σ, tolerance intervals for the distribution $N(\mu, \sigma^2)$, and lower confidence bounds for the proportion of conformance. The approach is based on fiducial inference. Recent research results [9, 13] and many simulation studies, including the one carried out in this paper, show that fiducial inference is a valid statistical method with good operating characteristics.

A small scale simulation study showed that the fiducial interval for μ in the resolution error model performs as well as the best of the competing methods. However, unlike the other methods, the fiducial method also provides an uncertainty interval for σ and allows the calculation of tolerance intervals and confidence bounds for conformance probabilities without any additional theoretical effort and with very little additional computational effort.

We used the resolution error model to illustrate the fiducial approach in a problem where an analytic expression for the fiducial distribution is not available. We also described a generalized fiducial recipe based on structural equation(s) defining the data generation mechanism. The concept of the structural equation(s) is particularly useful in metrological applications since it is closely related to the measurement equation that describes the measurement process. Having specified the structural equation(s), the fiducial distribution of the parameters of interest can be obtained using the recipe.

Appendix

We describe an algorithm that can be used to obtain random samples from the joint fiducial distribution of (μ, σ). Given measurements $y_i, i = 1, \ldots, n$ and resolution d, the algorithm consists of the following steps:

1. Calculate $a_i = y_i - d/2$ and $b_i = y_i + d/2, i = 1, \ldots, n$, which are the bounds of the measurements, i.e., $a_i \leqslant \mu + \sigma z_i < b_i$, where z_i is the realized value from a standard Gaussian distribution.
2. Generate $v_i \sim \text{uniform}(a_i, b_i), i = 1, \ldots, n$.
3. Generate a fiducial sample of (μ, σ) of a Gaussian distribution based on $v_i, i = 1, \ldots, n$. That is,

$$\tilde{\sigma} = \sqrt{(n-1)s_v^2/w},$$

$$\tilde{\mu} = \bar{v} + \frac{\tilde{\sigma}}{\sqrt{n}}q,$$

   where $\bar{v}$ and s_v are mean and standard deviation of v_i, and q and w are random deviates from $N(0, 1)$ and $\chi^2(n-1)$.
4. Obtain the initial values of $z_i = (v_i - \tilde{\mu})/\tilde{\sigma}, i = 1, \ldots, n$.
5. Given a_i, b_i and z_i, find solutions of (μ, σ) that satisfy $a_i \leqslant \mu + \sigma z_i < b_i, i = 1, \ldots, n$. This is equivalent to finding the vertices of a polygon determined by the n pairs of two parallel lines $a_i = \mu + \sigma z_i$ and $b_i = \mu + \sigma z_i$.
6. For each $i, i = 1, \ldots, n$, update z_i by
   (a) obtaining the polygon with the ith pair of lines removed and hence the new solutions $(\mu^\star, \sigma^\star)$,
   (b) calculating

$$r_1 = \min_{(\mu^\star, \sigma^\star)} \{(a_i - \mu^\star)/\sigma^\star\}$$

   and

$$r_2 = \max_{(\mu^\star, \sigma^\star)} \{(b_i - \mu^\star)/\sigma^\star\},$$

   (c) obtaining the new z_i as a standard normal random deviate conditional on $z_i \in (r_1, r_2)$ or $z_i = \Phi^{-1}(v^\star)$ where $v^\star \sim \text{uniform}(\Phi(r_1), \Phi(r_2))$.
7. Based on these updated $z_i, i = 1, \ldots, n$, obtain the polygon described in step 5. Obtain a fiducial sample randomly from the vertices of the polygon.
8. Repeat steps 6 and 7 for the desired number of fiducial samples.

We implemented the above algorithm in an R function `fir`. The function has five arguments:

1. A vector of measurements.
2. Resolution value.
3. Number of fiducial samples desired.
4. Number of cycles of the Gibbs sampler for burn in. The default value is 100.
5. Number of cycles skipped between samples. The default value is 10.

The output contains the desired number of samples from the fiducial distribution of (μ, σ). With this function, the following commands may be used for the micrometer example:

```
> micrometer <- c(7.489, 7.503, 7.433,
                  7.549, 7.526, 7.396,
                  7.543, 7.509, 7.504,
                  7.383)
> d <- 0.001
> nsample <- 10000
> fs <- fir(micrometer, d, nsample)
```

The output `fs` consists of two components: `mu` and `sigma`. We can use the following command to plot the 10 000 realizations of the fiducial distribution:

```
> plot(fs$mu, fs$sigma)
```

Figure 3 displays this plot. A 95% fiducial interval for μ is obtained from

```
> quantile(fs$mu, c(0.025, 0.975))
    2.5%     97.5%
7.441340 7.524986
```

To obtain a 99% content, 95% confidence tolerance interval for this example, the following commands may be used:

```
> beta <- 0.99
> gamma <- 0.95
> z1 <- qnorm((1-beta)/2)
> z2 <- qnorm((1+beta)/2)
> low <- fs$mu + z1 * fs$sigma
> quantile(low, (1-gamma)/2)
    2.5%
7.193229
> high <- fs$mu + z2 * fs$sigma
> quantile(high, (1+gamma)/2)
  97.5%
7.76826
```

The desired tolerance interval is (7.193229, 7.76826).

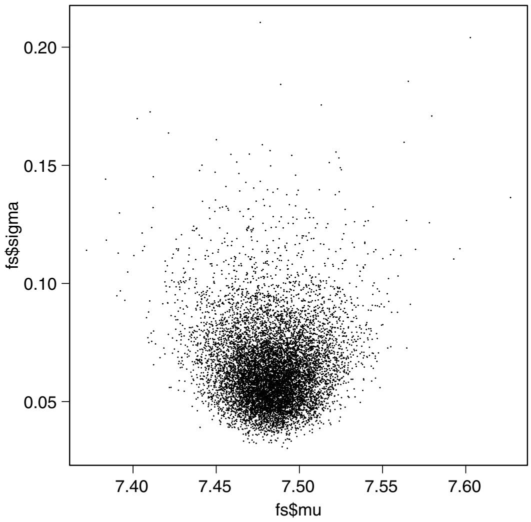

Figure 3. Samples from a fiducial distribution of (μ, σ).

Acknowledgments

The authors are grateful to the referees for their valuable suggestions. This work is a contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States.

References

[1] Lira I H and Wöger 1997 *Meas. Sci. Technol.* **8** 441–3
[2] Lira I H 2006 *Metrologia* **43** L14–7
[3] Elster C 2000 *Meas. Sci. Technol.* **11** 1359–63
[4] Frenkel R B and Kirkup L 2005 *Metrologia* **42** L27–30
[5] Taraldsen G 2006 *Metrologia* **43** 539–44
[6] Willink R 2007 *Metrologia* **44** 73–81
[7] Fisher R A 1930 *Proc. Camb. Phil. Soc.* **xxvi** 528–35
[8] Fraser D A S 1968 *The Structure of Inference* (New York: Wiley)
[9] Hannig J 2006 On fiducial inference *Technical Report* http://www.stat.colostate.edu/research/2006_3.pdf
[10] Wang C M and Iyer H K 2005 *Metrologia* **42** 145–53
[11] Wang C M and Iyer H K 2006 *Measurement* **39** 856–63
[12] Wang C M and Iyer H K 2006 *Metrologia* **43** 486–94
[13] Hannig J, Iyer H K and Patterson P D 2006 *J. Am. Stat. Assoc.* **101** 254–69
[14] Gelman A, Carlin J B, Stern H S and Rubin D B 1995 *Bayesian Data Analysis* (London: Chapman and Hill)
[15] R Development Core Team 2003 *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing) ISBN 3-900051-00-3 and URL http://www.R-project.org