

STRUCTURAL INFERENCE FOR UNCERTAINTY EVALUATION WITH APPLICATION TO CALIBRATION EXPERIMENTS

Hari K. Iyer and C. M. Jack Wang
National Institute of Standards and Technology, Boulder, CO

In many situations, the value of a measurand M is related to certain quantities denoted by m_1, m_2, \dots, m_N through a known functional relationship of the form

$$M = f(m_1, m_2, \dots, m_N).$$

The function f is generally a consequence of physical laws relating the relevant quantities. In practice, the values of m_i are not known exactly. However, results X_1, X_2, \dots, X_N are available that estimate m_1, m_2, \dots, m_N , respectively. Since m_i are estimated based on direct measurements, we have a model of the form

$$X_i = m_i + b_i + e_i$$

where b_i are errors for which type-B information is available, and e_i are errors for which type-A information is available. Let Y be the estimate of M that is obtained from

$$Y = f(X_1, X_2, \dots, X_N).$$

Thus, the same functional relationship that relates the value M of the measurand to the input variables m_1, m_2, \dots, m_N is used to calculate Y from the best estimates X_1, X_2, \dots, X_N of m_1, m_2, \dots, m_N . Consequently, the error in Y must be calculated or estimated by propagating the errors in the quantities X_1, X_2, \dots, X_N used to compute Y .

The ISO *Guide to the Expression of Uncertainty in Measurement* (GUM) recommends the use of a first-order Taylor series expansion for propagating errors and uncertainties. This generally provides adequate approximations for the actual uncertainties in the final result Y . However, when the nonlinearities in the function f are significant and/or the distribution of errors are described or modeled using a non-normal distribution, the GUM also permits the use of “other analytical or numerical methods” (Annex G.1.5). In this talk, we consider an alternative method, Fraser’s structural approach [1] [2], for the calculation of uncertainties in measurement results. The structural inference itself grew out of R. A. Fisher’s fiducial inference method. Consider a random sample X_1, \dots, X_n of size n from a $N(\mu, \sigma^2)$ distribution. The unknown parameters of this model are μ and σ . Suppose that the parameter of interest is μ . It is known that \bar{X} and S^2 are minimal sufficient and complete statistics for the parameters, where \bar{X} is the mean and S^2 is the variance of the samples. Furthermore, it is known that \bar{X} and S^2 are independent and have the following respective associated distribution

$$\begin{aligned} Z &= \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \\ U &= \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \end{aligned}$$

where $\chi^2(\nu)$ stands for the chi-squared distribution with ν degrees of freedom. We rewrite the above equations as follows:

$$\begin{aligned} \mu &= \bar{X} - \sigma Z/\sqrt{n} \\ \sigma &= \sqrt{(n-1)S^2/U}. \end{aligned}$$

In particular, substituting for σ in the first equation with the equivalent expression in the second equation, we get

$$\mu = \bar{X} - \left(\frac{S}{\sqrt{n}} \right) T_{n-1}$$

where $T_{n-1} = Z/\sqrt{U/(n-1)}$ is a student's t random variable with $n-1$ degrees of freedom. Fraser argued as follows. We have observed the value of \bar{X} and S , say, \bar{x} and s . Consequently, in the absence of any other information, all we know about μ is that it has a value that satisfies the equation

$$\mu = \bar{x} - \left(\frac{s}{\sqrt{n}} \right) T_{n-1}.$$

The only quantity that is unknown on the right hand side of the above equation is T_{n-1} . Thus, the value of μ must be related to a realization t_{n-1} of T_{n-1} via the above equation. Clearly, not all values of t_{n-1} are equally likely. Hence, not all values of μ are equally likely. The probabilities associated with the random variable T_{n-1} induce a probability structure on μ , and hence a distribution on μ . This induced distribution is what Fraser called the structural distribution of μ . We use $\tilde{\mu}$ as in

$$\tilde{\mu} = \bar{x} - \left(\frac{s}{\sqrt{n}} \right) T_{n-1}$$

to indicate the structural representation of the parameter μ . Inferences about μ can be made based on the structural distribution of μ . In this simple example, the structural procedures coincide with the classical frequentist procedures. For other complicated problems, the structural approach leads to solutions that are satisfactory and rivals other competing procedures.

To illustrate the use of the structural approach in calibration experiments, we consider the simple linear calibration model given by

$$Y_i = \beta_0 + \beta_1 x_i + e_i, \quad i = 1, \dots, n$$

where Y_i are the observed responses, x_i are the standards, β_0 and β_1 are the intercept and the slope of the calibration line, and e_i are independent, normally distributed random variables with mean 0 and unknown variance σ^2 . We first assume that the standards x_i 's are known exactly. The unknown parameters for this model are β_0 , β_1 , and σ . The quantity of interest, for example, is the value of the standard x_0 when the observed response $Y = y_0$. That is,

$$y_0 = \beta_0 + \beta_1 x_0 + e_0$$

where e_0 is a normal random variable with mean 0 and standard deviation σ . The joint structural distributions of β_0 , β_1 , and σ , obtained from the sampling theory for the bivariate normal distribution, are given by

$$\begin{aligned} \tilde{\beta}_0 &= \hat{\beta}_0 - \frac{s}{\sqrt{U}} t_{11} Z_1 \\ \tilde{\beta}_1 &= \hat{\beta}_1 - \frac{s}{\sqrt{U}} (t_{21} Z_1 + t_{22} Z_2) \\ \tilde{\sigma} &= \frac{s}{\sqrt{U}} \end{aligned}$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least-squares estimates of β_0 and β_1 , s^2 is the sum of squares of residuals from the fit, t_{ij} is the (i, j) th element of a lower triangular matrix \mathbf{t} such that $(\mathbf{x}^t \mathbf{x})^{-1} = \mathbf{t} \mathbf{t}^t$ with \mathbf{x} being an $n \times 2$ matrix that the elements of the first column of the matrix all equal to 1 and the

second column contains x_i , Z_1 and Z_2 are independent standard normal random variables, and U is a χ^2 random variable with $n - 2$ degrees of freedom and is independent of Z_1 and Z_2 . Since

$$x_0 = \frac{y_0 - \beta_0 - e_0}{\beta_1} = \frac{y_0 - \beta_0 - \sigma Z_3}{\beta_1}$$

where Z_3 is a standard normal random variable independent of Z_1 , Z_2 and U . The structural distribution of x_0 can then be obtained as

$$\tilde{x}_0 = \frac{y_0 - \tilde{\beta}_0 - \tilde{\sigma} Z_3}{\tilde{\beta}_1}$$

and the statistical inference about x_0 may be made based on \tilde{x}_0 .

If the values of standards x_i 's used in a calibration are not exactly known, but an estimate of x_i and its associated uncertainty are available, then the structural inference of this so-called *errors-in-variables* calibration model is just a straightforward extension of the inference of the simple calibration model. Specifically, the structural distributions of the unknown parameters of an errors-in-variables calibration model can be obtained by replacing x_i 's with their structural representations \tilde{x}_i 's in the calculation of $\hat{\beta}_0$, $\hat{\beta}_1$, s , and t .

In the context of uncertainty calculations, we can partition our measurement model parameters into two groups – parameters belonging to type-A distributions and those belonging to type-B distributions. For type-B parameters, the distributions that describe the uncertainties associated with the reported values of these parameters are assumed known. For type-A parameters, conditional on the values of the type-B parameters, we can now use their structural distributions to summarize the uncertainties associated with the reported values of these parameters. An important assumption that is made concerning the type-B parameters is that they are independent of the data. Hence the conditional distribution of the type-B parameters given the data is the same as their unconditional distribution. The net result is that we have a joint statistical distribution associated with the totality of all model parameters. It is now a straightforward technical matter to compute the distribution associated with any function of these model parameters. Such computations can be performed analytically in simple cases, or one can use simulation methods to estimate their distributions. For example, Let Y_1, \dots, Y_n be n independent determinations of a measurand μ obtained from a single measurement experiment. They are assumed to follow the model

$$Y_i = \mu + b + e_i, \quad i = 1, \dots, n$$

where e_i and b are the type-A and type-B components of error. We also assume that e_i are independent normal random variables with zero mean and variance σ^2 , and b has a known distribution. Under the circumstances, it is shown that the structural distribution of μ is given by

$$\tilde{\mu} = \bar{y} - \frac{s}{\sqrt{n}} T_{n-1} - b$$

where \bar{y} and s are the observed values of the sample mean and the sample standard deviation. The first two terms on the right hand side of the above equation account for the type-A component of the model, while the third term is for the type-B component. Since the distribution of b is completely specified, the distribution of $\tilde{\mu}$ may be determined analytically. However, it is most conveniently generated using Monte-Carlo simulation. A single realization of the structural distribution of μ may be generated as follows.

1. Generate a student's t random variables T_{n-1} .
2. Generate b according to its type-B distribution, independent of T_{n-1} .
3. Calculate $\tilde{\mu}$.

By generating a large number of independent realizations of $\tilde{\mu}$ we can estimate the mean, the uncertainty and quantiles of the structural distribution of μ and use them to make inference about μ .

Since the type-A and type-B components e and b may themselves be sums of elemental error components of their respective types, this simple example also demonstrates that the structural approach can be used to accommodate more general models for evaluating uncertainties in measurement results. It turns out that the use of structural distributions is particularly convenient especially when repeated measurements of an elemental quantity are assumed to follow a distribution which admits *complete sufficient statistics* as is the case in most metrological applications. More general cases can also be treated using the structural reasoning.

[1] Fraser, D. A. S. *The Structure of Inference*, New York: Wiley, 1968.

[2] Fraser, D. A. S. *Foundations of Statistical inference* (eds Godambe and Sprott), 32-55, Toronto: Holt, Rinehart and Winston, 1971.