A classical method for uncertainty analysis with multidimensional data

R. Willink[1] B. D. Hall[2]

[1] Applied Mathematics Centre, Industrial Research Ltd,
PO Box 31-310, Lower Hutt,
New Zealand.

[2] Measurement Standards Laboratory of New Zealand,
PO Box 31-310, Lower Hutt,
New Zealand.

Abstract

A method of uncertainty analysis based on classical statistical principles is presented for a measurand that is a linear combination of multidimensional input quantities. The method assigns the measurand a combined standard uncertainty matrix and an effective degrees of freedom, which enables the measurand to be estimated by an ellipsoidal confidence region in the multidimensional space. Simulations for a 95% nominal confidence level show the ellipsoids to contain the measurand with probability approximately 0.95, as required. The derivation of the method assumes all input uncertainties to be evaluated by the Type A method. The method is analogous to the Welch-Satterthwaite formula for one-dimensional data, a derivation of which is given in an appendix.

1 Introduction

The 'Guide to the Expression of Uncertainty in Measurement' (*Guide*), published by the ISO, is widely accepted as describing best-practice in the evaluation and reporting of measurement uncertainty [1]. The *Guide*'s approach, however, does not apply to all types of measurement. There is an on-going effort to revise and expand the scope of the *Guide* which is being coordinated by the Joint Committee for Guides in Metrology [2].

One important area not yet addressed by the *Guide* concerns measurements of multi-dimensional quantities, of which complex-valued quantities are a particular two-dimensional example. These can arise, for example, in some electrical and optical measurements. There does not appear to be an accepted method of evaluating the uncertainty in a complex-valued quantity which is estimated by a linear combination of uncertain complex-valued 'input' quantities, especially when the input uncertainties have finite degrees of freedom (i.e., they have been estimated from finite sized samples).

In one-dimensional cases the so-called Welch-Satterthwaite (WS) formula is recommended [1, Appendix G.4]. This is an approximate method based on classical normal statistics associated with Type A evaluation of uncertainty analysis [1, Clause 4.2]. It obtains a number of 'effective degrees of freedom' associated with the standard uncertainty in a linear combination of uncertain inputs, provided the input quantities are uncorrelated [3, 4, 5]. The effective degrees of freedom is used in obtaining an approximate confidence interval, or 'expanded uncertainty', for the quantity of interest using tabulated values of the Student's t-distribution.

This paper describes a new procedure, analogous to the WS formula, that can be applied in two-dimensional, and higher-dimensional, problems. In the case of complex-valued data, the procedure can be implemented on a programmable calculator, or computer, in a small number of simple equations given in Appendix 1. The more general multidimensional procedure can be implemented using a few algorithms from standard linear algebra.

Specifically, we give a method for the evaluation of the uncertainty in the weighted sum of independent multidimensional inputs. The derivation is directly analogous to the derivation of the WS formula for the sum of independent scalar inputs, which is given in Appendix 2. Readers unfamiliar with the derivation of the WS formula may wish to read this appendix first. The development of the method is based on the principles of classical statistics, so initially the material is presented in a corresponding style and notation. The familiar notation of the *Guide* is introduced at a later stage for the presentation of the results.

Section 2 reviews the construction of a confidence region in p-dimensional space for a single mean with unknown variance-covariance matrix – this is the multidimensional analogue of a confidence interval for a single normal mean with unknown variance constructed using a Student's t-distribution. Section 3 develops the approximate confidence region for a linear combination of several p-dimensional means – this is a multidimensional analogue of the WS formula. Section 3.1 presents alternative ways in which to calculate the effective degrees of freedom, and Appendix 1 gives explicit formulae for complex-valued inputs, i.e., $p = 2$. The corresponding procedures have been tested with an extensive series of Monte Carlo simulations. These results are discussed in Section 4.

Scalar quantities are indicated in italic type, e.g. Z, x_i. Vector quantities are indicated in bold italic type, e.g. $\boldsymbol{V}, \boldsymbol{x}_i$, and matrix quantities in bold upright type, e.g. $\mathbf{M}, \mathbf{s}$. Writing $\mathbf{M} \equiv (M_{jk})$ indicates that M_{jk} is the element in the jth row and kth column of $\mathbf{M}$. Following conventional statistical practice, random variables are indicated in upper case, e.g. Z, $\mathbf{S}$, S_{jk}, while particular values taken by these variables are indicated in lower case, e.g. z, $\mathbf{s}$, s_{jk}. Greek symbols represent fixed quantities also known as statistical parameters. A prime, $'$, indicates the transpose. In general, the running index i is applied to vectors and sets, (i.e. one-dimensional arrays), while the indices j and k are used to identify rows and columns of matrices. Exceptions to these rules are the familiar use of t_ν and χ^2_ν to denote variables with Student's t and chi-square distributions with ν degrees of freedom, and the addition of a decimal subscript to indicate a fixed point in the tail of a distribution, e.g. $t_{\nu,0.975}$.

2 The single-input confidence region

First it is helpful to review the theory describing the uncertainty of a single multidimensional quantity evaluated by the Type A method. This is the statistical theory of the estimation of the mean of a p-dimensional multinormal variable. Let $\boldsymbol{Z} \equiv (Z_1, \ldots, Z_p)'$ be a random column vector having the p-dimensional multinormal distribution[1] with unknown mean vector $\boldsymbol{\mu} \equiv (\mu_1, \ldots, \mu_p)'$ and unknown covariance matrix $\boldsymbol{\Sigma} \equiv (\sigma_{jk})$, i.e.

$$\begin{aligned} \mu_j &= E[Z_j] \\ \sigma_{jk} &= E[(Z_j - \mu_j)(Z_k - \mu_k)]. \end{aligned}$$

(Note that σ_{jk} has the dimensions of units-squared, and σ_{jj} is the notation for a variance.) We write $\boldsymbol{Z} \sim \boldsymbol{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\sim$ means '(is) distributed as'. Suppose that we wish to find a p-dimensional confidence region for $\boldsymbol{\mu}$ from a random sample of n such vectors, denoted $\boldsymbol{Z}_{[1]}, \ldots, \boldsymbol{Z}_{[n]}$, with sample-mean vector and sample covariance matrix

$$\begin{aligned} \bar{\boldsymbol{Z}} &\equiv \tfrac{1}{n} \textstyle\sum_{k=1}^{n} \boldsymbol{Z}_{[k]} \\ \mathbf{S} &\equiv \tfrac{1}{n-1} \textstyle\sum_{k=1}^{n} \left(\boldsymbol{Z}_{[k]} - \bar{\boldsymbol{Z}}\right) \left(\boldsymbol{Z}_{[k]} - \bar{\boldsymbol{Z}}\right)'. \end{aligned}$$

The sample-mean vector and the sample-mean covariance matrix, $\mathbf{S}/n$, obey independently

$$\bar{\boldsymbol{Z}} - \boldsymbol{\mu} \sim \boldsymbol{N}_p(\mathbf{0}, \boldsymbol{\Sigma}/n) \tag{1}$$

$$\mathbf{S}/n \sim \tfrac{1}{n-1} \mathbf{W}_p(n-1, \boldsymbol{\Sigma}/n), \tag{2}$$

where $\mathbf{W}_p(\nu, \boldsymbol{\Sigma}/n)$ is a $p \times p$ matrix variable with the Wishart distribution with ν degrees of freedom and scale matrix $\boldsymbol{\Sigma}/n$ [6]. These equations serve as a definition of the Wishart distribution.

The variable $\mathbf{W}_p(\nu, \boldsymbol{\Sigma})$, for general p, ν and $\boldsymbol{\Sigma}$, is the multidimensional analogue of $\sigma^2\chi^2_\nu$, where σ^2 is a variance and χ^2_ν is a variable with the chi-square distribution with

[1] Also simply called the multivariate normal distribution. In statistics the adjective 'multivariate' implies a variable with many components rather than a set of many single variables.

ν degrees of freedom. In both cases the scalar ν acts as a shape parameter. The matrix parameter $\mathbf{\Sigma}$ is a scale parameter incorporating the scale factor σ^2.

The following result acts as a definition of the distribution of Hotelling's T^2, which is the multivariate analogue of the square of a Student's t variable.

- *If* $\boldsymbol{V} \sim \boldsymbol{N}_p(\mathbf{0}, \mathbf{\Sigma})$ *and independently* $\mathbf{M} \sim \frac{1}{\nu}\mathbf{W}_p(\nu, \mathbf{\Sigma})$ *then the scalar* $\boldsymbol{V}'\mathbf{M}^{-1}\boldsymbol{V}$ *has the distribution of Hotelling's* T^2 *with parameters* ν *and* p*. This can be written*

$$\boldsymbol{V}'\mathbf{M}^{-1}\boldsymbol{V} \sim T^2_{\nu,p}. \tag{3}$$

It turns out that

$$T^2_{\nu,p} \sim \frac{\nu p}{\nu + 1 - p} F_{p,\nu+1-p} \tag{4}$$

where the variable $F_{p,\nu+1-p}$ *has the better-known* F*-distribution with* p *and* $\nu+1-p$ *degrees of freedom.*

Therefore, from (1) and (2),

$$(\bar{\boldsymbol{Z}} - \boldsymbol{\mu})'(\mathbf{S}/n)^{-1}(\bar{\boldsymbol{Z}} - \boldsymbol{\mu}) \sim T^2_{n-1,p}.$$

We now introduce notation similar to that of the *Guide*. Let the observed estimate of the required quantity $\boldsymbol{\mu}$ be the vector $\boldsymbol{x}$, and let the uncertainty data for $\boldsymbol{x}$ be in the form of a $p \times p$ matrix $\mathbf{u} \equiv (u_{jk})$ of *squared* standard uncertainties. We can associate $\boldsymbol{x}$ with $\bar{\boldsymbol{z}}$ and $\mathbf{u}$ with $\mathbf{s}/n$, where $\bar{\boldsymbol{z}}$ and $\mathbf{s} \equiv (s_{jk})$ are the observed values of $\bar{\boldsymbol{Z}}$ and $\mathbf{S}$. Therefore the 95% confidence region quoted for $\boldsymbol{\mu}$ is the set of all p-dimensional vectors $\boldsymbol{\mu}_0$ for which

$$(\boldsymbol{x} - \boldsymbol{\mu}_0)'\mathbf{u}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_0) \leq \frac{(n-1)p}{n-p} F_{p,n-p,0.95} \tag{5}$$

where $F_{p,n-p,0.95}$ is the 0.95 quantile of the distribution of $F_{p,n-p}$. This region is an ellipsoid in p-dimensional space.

For example, suppose that

$$(4.61,\ 3.13)',\ (5.00,\ 3.37)',\ (4.00,\ 2.47)',\ (2.64,\ 4.38)',\ (5.03,\ 2.72)'$$

is a sample of size $n = 5$ that is drawn from a bivariate distribution (i.e. $p = 2$) with $\boldsymbol{\mu}= (4, 3)$ and some $\mathbf{\Sigma}$. This gives

$$\boldsymbol{x} = \begin{pmatrix} 4.26 \\ 3.21 \end{pmatrix} \qquad \mathbf{u} = \begin{bmatrix} 0.16 & -0.08 \\ -0.08 & 0.09 \end{bmatrix} \qquad \mathbf{u}^{-1} = \begin{bmatrix} 11.18 & 10.17 \\ 10.17 & 20.21 \end{bmatrix}.$$

The critical value on the right-hand side of (5) is 25.47 and the confidence region is the interior of the ellipse shown in Figure 2, where the hollow marker indicates $\boldsymbol{x}$, the solid marker indicates $\boldsymbol{\mu}$ and the crosses indicate the data points. We expect the ellipse, (which is random with respect to the sampling), to enclose the fixed point $\boldsymbol{\mu}$ for 95% of samples, as it does on this occasion.[2,3]

[2] This is the proper understanding of the statement that the region is a 95% confidence region for $\boldsymbol{\mu}$. To avoid confusion, a region derived under a statistical methodology that does not treat $\boldsymbol{\mu}$ as fixed but assigns it a distribution should not be simply called a *confidence* region.

[3] The above is data generated with $\sigma_{11} = \sigma_{22} = 2$ and $\sigma_{12} = \sigma_{21} = 0.6\surd(\sigma_{11}\sigma_{22})$, which implies a

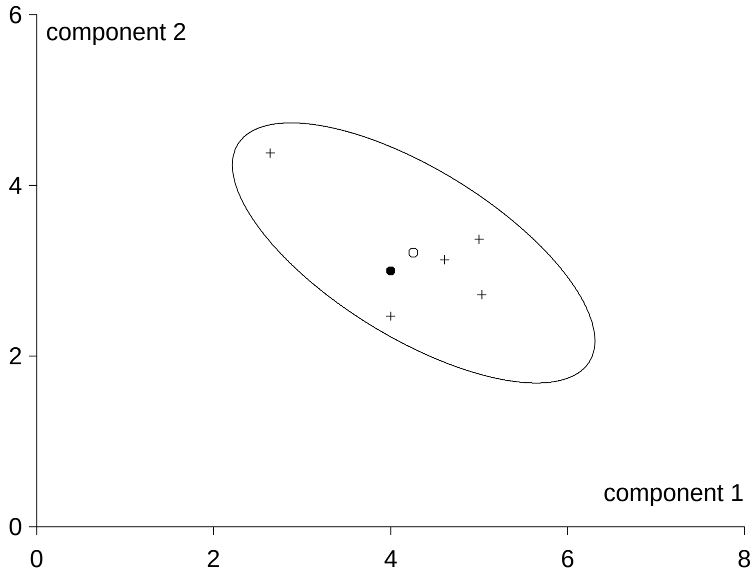

Figure 1: - 95% confidence region for a complex-valued quantity. Hollow marker - estimate. Solid marker - true value. Crosses - data points.

3 The multiple-input confidence region

This section extends the above method to be applicable in the situation of main interest, where there are many independent multidimensional input quantities combining in a known way to define a multidimensional measurand.

By taking the first-order terms of a multidimensional Taylor expansion of the measurement function, the measurand can be approximated by a weighted combination of each input, where the weights are known as sensitivity coefficients. This is consistent with the approach for one-dimensional variables in the *Guide*. So, if there are m inputs, the measurand is approximately equal to a sum, $\sum_{i=1}^{m} \boldsymbol{\mu}_i$, where the vector $\boldsymbol{\mu}_i$ is the weighted ith input quantity.

In this situation the following theory can be applied. Let $\boldsymbol{Z}_i$, $i = 1, \ldots, m$, be a variable with the p-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i \equiv (\sigma_{i \cdot jk})$; that is, $\sigma_{i \cdot jk}$ is the element in the jth row and kth column of $\boldsymbol{\Sigma}_i$. Suppose a confidence region for $\sum_{i=1}^{m} \boldsymbol{\mu}_i$ is to be found from independent random samples of size n_i. Let the sample of $\boldsymbol{Z}_i$ have random sample-mean vector $\bar{\boldsymbol{Z}}_i$ and random sample covariance matrix $\mathbf{S}_i \equiv (S_{i \cdot jk})$. The sample-mean covariance matrix is

correlation of +0.6 between the two components. In this case the positive correlation is not reflected in the shape of the confidence region.

$\mathbf{S}_i/n_i$.

We seek a product having the form of (3), in order to obtain a region defined by an equation such as (5). The variable

$$\textstyle\sum_{i=1}^m \bar{\boldsymbol{Z}}_i - \sum_{i=1}^m \boldsymbol{\mu}_i \sim \boldsymbol{N}_p(\mathbf{0}, \sum_{i=1}^m \mathbf{\Sigma}_i/n_i) \tag{6}$$

can take the part of $\boldsymbol{V}$ in (3). However, in general no function of all the $\mathbf{S}_i/n_i$ matrices has a Wishart distribution with scale matrix proportional to $\sum_{i=1}^m \mathbf{\Sigma}_i/n_i$ to play the part of $\mathbf{M}$. So we cannot obtain a product with an exact T^2-distribution. Therefore we look for an approximation in which a multiple of a Wishart variable with matrix parameter $\sum_{i=1}^m \mathbf{\Sigma}_i/n_i$ is constructed to match first and second moments of the variable $\sum_{i=1}^m \mathbf{S}_i/n_i$. The approximation is

$$\textstyle\sum_{i=1}^m \mathbf{S}_i/n_i \overset{\text{approx}}{\sim} \frac{c}{\nu}\mathbf{W}_p(\nu, \sum_{i=1}^m \mathbf{\Sigma}_i/n_i), \tag{7}$$

where c and ν are scalar constants to be found by matching means and covariances of corresponding matrix elements on each side of the equation. This approximation gives

$$\left(\sum_{i=1}^m \bar{\boldsymbol{Z}}_i - \sum_{i=1}^m \boldsymbol{\mu}_i\right)' \left(\sum_{i=1}^m \mathbf{S}_i/n_i\right)^{-1} \left(\sum_{i=1}^m \bar{\boldsymbol{Z}}_i - \sum_{i=1}^m \boldsymbol{\mu}_i\right) \overset{\text{approx}}{\sim} \frac{T^2_{\nu,p}}{c}, \tag{8}$$

which is then used to obtain an ellipsoidal confidence region, as in (5).

There are two unknown constants, c and ν, to be determined but there are $q \equiv p(p+1)/2$ elements in the upper-triangular parts defining the symmetric random matrices on each side of (7). So there are q equations involved in matching the means and $q(q+1)/2$ equations in matching the covariances. Therefore, depending on one's viewpoint, there is either flexibility or ambiguity in the method of choosing c and ν. This contrasts with the WS situation of $p = 1$ in Appendix 2, where there is only one solution because there are just two equations for the two unknowns.

To proceed we require the following formulae which give the means and covariances of the elements of a (scaled) Wishart matrix for general p, ν and $\mathbf{\Sigma}$.

- *If* $\mathbf{M} \sim \frac{1}{\nu}\mathbf{W}_p(\nu, \mathbf{\Sigma})$, *where* $\mathbf{M} \equiv (M_{jk})$ *and* $\mathbf{\Sigma} \equiv (\sigma_{jk})$, *then*

$$E[M_{jk}] = \sigma_{jk} \tag{9}$$

$$\text{cov}(M_{jk}, M_{rt}) = (\sigma_{jr}\sigma_{kt} + \sigma_{jt}\sigma_{kr})/\nu. \tag{10}$$

- *Also, for any positive scalar* a,

$$a\mathbf{W}_p(\nu, \mathbf{\Sigma}) \sim \mathbf{W}_p(\nu, a\mathbf{\Sigma}). \tag{11}$$

Applying (2) and (11) and then summing gives

$$\textstyle\sum_{i=1}^m \mathbf{S}_i/n_i \sim \sum_{i=1}^m \frac{1}{n_i-1}\mathbf{W}_p(n_i - 1, \mathbf{\Sigma}_i/n_i).$$

Consequently, (9) implies that the means of the q upper-triangular elements on each side specifying (7) can be matched by choosing $c = 1$. However it is unclear how best to combine the $q(q+1)/2$ covariances to define ν. Some possibilities are considered below. For the moment, let us assume that we make such a definition, and that we estimate this quantity, ν, by an effective degrees of freedom, ν_{eff}, obtainable from the data.

In the preferred notation, let $\boldsymbol{x}_i$ be the observed estimate of $\boldsymbol{\mu}_i$, and let $\mathbf{u}_i \equiv (u_{i \cdot jk})$ be its matrix of squared standard uncertainties. So, from (4) and (8) with $c = 1$, an approximate 95% confidence region for the measurand $\sum_{i=1}^m \boldsymbol{\mu}_i$ is the set of all vectors $\boldsymbol{\mu}_0$ for which

$$\left(\sum_{i=1}^m \boldsymbol{x}_i - \boldsymbol{\mu}_0\right)' \left(\sum_{i=1}^m \mathbf{u}_i\right)^{-1} \left(\sum_{i=1}^m \boldsymbol{x}_i - \boldsymbol{\mu}_0\right) \leq \frac{p\nu_{\text{eff}}\, F_{p,\nu_{\text{eff}}+1-p,0.95}}{\nu_{\text{eff}}+1-p}.$$

As in (5), this is an ellipsoid in p-dimensional space. We can define the matrix $\mathbf{u}_{\text{c}} \equiv \sum_{i=1}^m \mathbf{u}_i$ for use in this equation.

3.1 Definition of ν_{eff}

We return to the question of defining ν to best match the covariances of the corresponding elements on each side of (7). In effect, the covariance matrices of the q upper-triangular elements are to be condensed into single best figures of variability for matching.

Let $\boldsymbol{V}(\mathbf{M})$ be a vector made up of the q upper-triangular elements of a $p \times p$ matrix $\mathbf{M}$, which represents the matrix on either side of (7). We look for appropriate measures figure of the 'variance' of $\boldsymbol{V}(\mathbf{M})$, which is unaffected by the re-ordering of its elements. Two such figures are its *total variance* and its *generalized variance*, which are the trace and determinant of its $q \times q$ covariance matrix $\mathbf{C}(\mathbf{M})$ respectively [7].[4] So ν can be defined by matching either the traces or determinants of the $q \times q$ covariance matrices arising from the exact and approximating matrices in (7), with $c = 1$. These covariance matrices are easily shown to be

$$\begin{aligned}\boldsymbol{\Lambda} &\equiv \textstyle\sum_{i=1}^m \mathbf{C}(\mathbf{S}_i)/n_i^2\\ \boldsymbol{\Omega} &\equiv (1/\nu)^2\, \mathbf{C}\left(\mathbf{W}_p(\nu, \textstyle\sum_{i=1}^m \boldsymbol{\Sigma}_i/n_i)\right).\end{aligned}$$

If we define $\boldsymbol{\Theta} \equiv \nu\boldsymbol{\Omega}$ then matching the traces or determinants of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$ gives $\nu = \text{tr}\,\boldsymbol{\Theta}/\text{tr}\,\boldsymbol{\Lambda}$ or $\nu = (\det \boldsymbol{\Theta}/\det \boldsymbol{\Lambda})^{1/q}$ respectively. The matrices $\boldsymbol{\Theta}$ and $\boldsymbol{\Lambda}$ are unknown but are estimable by matrices $\hat{\boldsymbol{\Theta}}$ and $\hat{\boldsymbol{\Lambda}}$ derived from the input data. So two possibilities for the effective degrees of freedom, ν_{eff}, are

$$\nu_{\text{tv}} \equiv \text{tr}\,\hat{\boldsymbol{\Theta}}/\text{tr}\,\hat{\boldsymbol{\Lambda}}, \tag{12}$$

$$\nu_{\text{gv}} \equiv \left(\det \hat{\boldsymbol{\Theta}}/\det \hat{\boldsymbol{\Lambda}}\right)^{1/q}. \tag{13}$$

The subscripts reflect the matching of the total variances and the generalized variances respectively. The method of matching generalized variances is due to Tan and Gupta [8].

We shall now find the elements of the symmetric matrices $\boldsymbol{\Theta} \equiv (\theta_{JK})$, $\boldsymbol{\Lambda} \equiv (\lambda_{JK})$, $\hat{\boldsymbol{\Theta}} \equiv (\hat{\theta}_{JK})$ and $\hat{\boldsymbol{\Lambda}} \equiv (\hat{\lambda}_{JK})$. Let $J(j,k) \equiv K(j,k)$ be the position of the upper-triangular

[4] The invariance of these to the ordering of the elements in $\boldsymbol{V}$ can be established easily.

element M_{jk}, $(j \le k)$, in the vector $\boldsymbol{V}(\mathbf{M})$[5]. Then the element in the $J(j,k)$th row and $K(r,t)$th column of $\mathbf{C}[\mathbf{M}]$ is $\text{cov}(M_{jk}, M_{rt})$. Also, note the following general result.

- *For constants $\{a_g\}$ and $\{b_h\}$ and variables $\{X_g\}$ and $\{Y_h\}$,*

$$\text{cov}\left(\sum_{g=1}^{n} a_g X_g, \sum_{h=1}^{m} b_h Y_h\right) = \sum_{g=1}^{n} \sum_{h=1}^{m} a_g b_h \text{cov}(X_g, Y_h).$$

Applying this result and (10), and defining $\nu_i \equiv n_i - 1$, gives

$$\begin{aligned}\lambda_{J(j,k)K(r,t)} &= \textstyle\sum_{i=1}^{m} (\sigma_{i\cdot jr}\sigma_{i\cdot kt} + \sigma_{i\cdot jt}\sigma_{i\cdot kr})/(n_i^2 \nu_i), \\ \theta_{J(j,k)K(r,t)} &= \textstyle\sum_{i=1}^{m} \sigma_{i\cdot jr}/n_i \sum_{i=1}^{m} \sigma_{i\cdot kt}/n_i + \sum_{i=1}^{m} \sigma_{i\cdot jt}/n_i \sum_{i=1}^{m} \sigma_{i\cdot kr}/n_i.\end{aligned}$$

As is common practice, the unknown $\sigma_{i\cdot jk}$ is replaced by its estimate $s_{i\cdot jk}$, where $\mathbf{s}_i \equiv (s_{i\cdot jk})$ is the observed value of $\mathbf{S}_i$. Also $u_{i\cdot jk} \equiv s_{i\cdot jk}/n_i$. This gives

$$\begin{aligned}\hat{\lambda}_{J(j,k)K(r,t)} &= \textstyle\sum_{i=1}^{m} (u_{i\cdot jr}u_{i\cdot kt} + u_{i\cdot jt}u_{i\cdot kr})/\nu_i \\ \hat{\theta}_{J(j,k)K(r,t)} &= \textstyle\sum_{i=1}^{m} u_{i\cdot jr} \sum_{i=1}^{m} u_{i\cdot kt} + \sum_{i=1}^{m} u_{i\cdot jt} \sum_{i=1}^{m} u_{i\cdot kr},\end{aligned}$$

from which we can evaluate (12) and (13). It follows that an alternative expression for (12) is

$$\nu_{\text{tv}} \equiv \frac{\sum_{j=1}^{p} \sum_{k=j}^{p} [\sum_{i=1}^{m} u_{i\cdot jj} \sum_{i=1}^{m} u_{i\cdot kk} + (\sum_{i=1}^{m} u_{i\cdot jk})^2]}{\sum_{j=1}^{p} \sum_{k=j}^{p} [\sum_{i=1}^{m} (u_{i\cdot jj}u_{i\cdot kk} + u_{i\cdot jk}^2)/\nu_i]}. \tag{14}$$

In simulations for $p = 2$ described below, both ν_{tv} and ν_{gv} produced confidence regions that enclosed $\sum_{i=1}^{m} \boldsymbol{\mu}_i$ on approximately 95% of occasions when the input degrees of freedom were moderate or large, as required. However, with small degrees of freedom, the regions derived using ν_{tv} and ν_{gv} enclosed $\sum_{i=1}^{m} \boldsymbol{\mu}_i$ on more and fewer than 95% of occasions respectively, i.e. the methods were *conservative* and *anti-conservative*. This suggested that $\frac{1}{2}(\nu_{\text{tv}} + \nu_{\text{gv}})$ might perform even better at low degrees of freedom, which proved to be the case. Further, we guarded against unreasonable values of the effective degrees of freedom by defining the hybrid value

$$\nu_{\text{hy}} \equiv \text{median}\{\min\{\nu_i\}, \tfrac{1}{2}(\nu_{\text{tv}} + \nu_{\text{gv}}), \textstyle\sum_{i=1}^{m} \nu_i\}.$$

That is, ν_{hy} is the mean of ν_{tv} and ν_{gv} with the proviso that it may not be smaller than the smallest number of degrees of freedom nor greater than the total number of degrees of freedom. Applying these limits improved performance by a small amount. The simulations also provided strong evidence that

$$\min\{\nu_i\} \le \nu_{\text{tv}} \le \textstyle\sum_{i=1}^{m} \nu_i$$

because no instance was observed where this did not hold.

The above limits, $\min\{\nu_i\}$ and $\sum_{i=1}^{m} \nu_i$, are applicable because each degree of freedom represents a piece of information about a spread around a mean. The effective amount of information associated with spread about the unknown sum of means must be at least as much as the minimum amount of information in its individual parts, and no

[5] If $\boldsymbol{V}(\mathbf{M})$ is formed by reading the upper-triangular elements off row by row in the natural order then $J(j,k) \equiv j + k + (j-1)p - j(j+1)/2$ for $1 \le j \le k \le p$.

greater than the sum of the information in its parts.

If $p = 1$ then $\hat{\Lambda}$ and $\hat{\Theta}$ are 1×1 matrices with the elements $\sum_{i=1}^{m} 2u_{i\cdot 11}^2/\nu_i$ and $2(\sum_{i=1}^{m} u_{i\cdot 11})^2$ respectively. Equations (12) and (13) become

$$\nu_{\text{tv}} = \nu_{\text{gv}} = \frac{(\sum_{i=1}^{m} u_{i\cdot 11})^2}{\sum_{i=1}^{m} u_{i\cdot 11}^2/\nu_i},$$

which is equal to the WS effective degrees of freedom given in Appendix 2 as (22), (where $u_i \equiv \sqrt{u_{i\cdot 11}}$). This quantity always lies between $\min\{\nu_i\}$ and $\sum_{i=1}^{m} \nu_i$ as required, and does on occasion attain these bounds. So the WS effective degrees of freedom is also just a special case of ν_{hy}.

For an example of the method, consider the addition of three complex-valued quantities, i.e. where $m = 3$ and $p = 2$. Suppose samples of $n_1 = 6$, $n_2 = 4$ and $n_3 = 7$ measurements give

$$\mathbf{u}_1 = \begin{bmatrix} 0.96 & -0.34 \\ -0.34 & 0.27 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 0.51 & 0.33 \\ 0.33 & 0.31 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0.45 & 0.28 \\ 0.28 & 1.65 \end{bmatrix}.$$

Appendix 1 shows how ν_{tv} and ν_{gv} are easily calculated in the two-dimensional case. We find $\nu_{\text{tv}} = 11.3$ and $\nu_{\text{gv}} = 12.4$. Also $\min\{\nu_i\} = 3$ and $\sum_{i=1}^{m} \nu_i = 14$, so $\nu_{\text{hy}} = 11.9$.

4 Performance evaluation results

A simulation was carried out to determine the performance of $\nu_{\text{tv}}, \nu_{\text{gv}}$ and ν_{hy} when a pair of two-dimensional weighted inputs are added, i.e. $m = 2$ and $p = 2$.

The weighted inputs $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ were both set to $(0, 0)'$ without loss of generality. Any situation could then be specified by the known sample sizes, n_1 and n_2, and the unknown covariance matrices, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. With complete generality, these matrices were defined by

$$\boldsymbol{\Sigma}_i = n_i \begin{bmatrix} \kappa_{i\cdot 11}^2 & \rho_i \kappa_{i\cdot 11} \kappa_{i\cdot 22} \\ \rho_i \kappa_{i\cdot 11} \kappa_{i\cdot 22} & \kappa_{i\cdot 22}^2 \end{bmatrix}$$

for $i = 1, 2$, where ρ_i is the correlation coefficient and $\kappa_{i\cdot 11}^2 \equiv \sigma_{i\cdot 11}^2/n_i$ and $\kappa_{i\cdot 22}^2 \equiv \sigma_{i\cdot 22}^2/n_i$ are the variances of the components of $\bar{\boldsymbol{Z}}_i$.

A Monte Carlo method was used to assess the performance of $\nu_{\text{tv}}, \nu_{\text{gv}}$ and ν_{hy} for every pair of sample sizes where

$$n_1, n_2 \in \{3, 4, 8, 10, 15, \infty\}, \quad n_1 \leq n_2$$

and every combination of the unknown parameters where

$$\begin{aligned} \rho_1 &\in \{0, 0.4, 0.8\} \\ \rho_2 &\in \{-0.8, -0.4,\ 0,\ 0.4,\ 0.8\} \\ \kappa_{1\cdot 11}^2 &= 1 \\ \kappa_{1\cdot 22}^2, \kappa_{2\cdot 11}^2, \kappa_{2\cdot 22}^2 &\in \{\tfrac{1}{256}, \tfrac{1}{64}, \tfrac{1}{16}, \tfrac{1}{4},\ 1,\ 4,\ 16,\ 64,\ 256\}. \end{aligned}$$

(Allowing infinite sample sizes provides useful limiting information.) For each parameter

Table 1: Median and mean confidence level ($\times 10^4$)

n_1	n_2	ν_{tv} median	ν_{tv} mean	ν_{gv} median	ν_{gv} mean	ν_{hy} median	ν_{hy} mean
3	3	9819	9822	9008	8941	9476	9453
3	4	9737	9725	9076	9054	9438	9411
3	8	9518	9461	8932	9006	9169	9205
3	10	9491	9402	8889	8979	9103	9149
3	15	9450	9323	8821	8934	9007	9067
3	∞	9324	9173	8515	8722	8627	8785
4	4	9719	9717	9324	9216	9482	9471
4	8	9552	9550	9258	9247	9427	9389
4	10	9525	9502	9226	9234	9396	9355
4	15	9500	9439	9181	9212	9347	9304
4	∞	9427	9308	9012	9066	9126	9107
8	8	9602	9601	9451	9389	9492	9479
8	10	9579	9578	9452	9399	9490	9476
8	15	9525	9540	9444	9405	9483	9463
8	∞	9473	9449	9367	9351	9397	9365
10	10	9580	9580	9460	9414	9493	9483
10	15	9541	9548	9460	9425	9489	9477
10	∞	9478	9466	9404	9389	9424	9399
15	15	9552	9553	9471	9445	9493	9488
15	∞	9486	9484	9443	9432	9453	9439
∞	∞	9500	9500	9500	9500	9500	9500

combination, 10 000 sampling experiments were simulated and a recording made of the number of times the confidence region enclosed the true sum of means $(0, 0)'$. Ideally the result would lie within one or two standard deviations of 9 500. (The standard deviation is $\sqrt{[10\,000 \times 0.95 \times (1 - 0.95)]} \approx 22$.)

So for each n_1, n_2 combination there were $3 \times 5 \times 9^3 = 10\,935$ parameter combinations of $\rho_1, \rho_2, \kappa^2_{1\cdot 22}, \kappa^2_{2\cdot 11}$ and $\kappa^2_{2\cdot 22}$, whose results are summarized in Tables 1 and 2. If these 10 935 combinations are regarded as being representative of those that may occur in practice then the average confidence level, as represented by the median or mean in Table 1, is of interest. If more emphasis is placed on the most extreme levels of confidence achieved for any set of the unknown parameters then the minimum and maximum levels of Table 2 are more relevant.[6]

Table 1 shows that when n_1 and n_2 are both at least 10, the average confidence for each choice of effective degrees of freedom is reasonably close to the nominal value desired, 0.95 (i.e. 9 500). When either n_1 or n_2 is less than 10, differences of performance between the three methods are more noticeable. Confidence levels appear to be highest

[6] Owing to sampling fluctuations the actual minimum and maximum levels of confidence for parameter combinations in this set are likely to be several standard deviations closer to 0.95.

Table 2: Minimum and maximum confidence level ($\times 10^4$)

		ν_{tv}		ν_{gv}		ν_{hy}	
n_1	n_2	min	max	min	max	min	max
3	3	9554	9979	7524	9560	9061	9715
3	4	9345	9962	7820	9610	8958	9690
3	8	8812	9965	8136	9597	8695	9636
3	10	8651	9968	8148	9593	8618	9608
3	15	8484	9964	8211	9589	8522	9594
3	∞	8148	9963	8015	9573	8146	9573
4	4	9337	9932	8191	9623	9166	9663
4	8	9089	9937	8494	9583	9051	9636
4	10	8991	9932	8537	9581	8986	9609
4	15	8877	9939	8578	9574	8848	9597
4	∞	8561	9943	8564	9558	8561	9558
8	8	9343	9791	8919	9573	9281	9637
8	10	9328	9796	8951	9581	9288	9640
8	15	9255	9825	9050	9585	9273	9605
8	∞	9072	9859	9076	9596	9073	9596
10	10	9361	9757	9017	9569	9310	9636
10	15	9317	9773	9105	9571	9305	9617
10	∞	9170	9818	9157	9565	9160	9582
15	15	9376	9692	9168	9573	9335	9622
15	∞	9258	9750	9257	9570	9257	9576
∞	∞	9415	9589	9415	9589	9415	9589

using ν_{tv}. With the exception of the results for $n_1 = 3, n_2 = \infty$ and $n_1 = 3, n_2 = 15$, the average values with ν_{tv} are close to or in excess of nominal. With ν_{gv} there is a tendency for the average confidence level to fall below nominal. This is accentuated when n_1 is low and n_2 high. A similar tendency is seen in the results for ν_{hy}, although in this case the figures are much closer to 9 500 when $n_1 \approx n_2$. Nowhere in Table 1 do results for ν_{gv} or ν_{hy} exceed 9 500. Therefore these methods tend to be slightly optimistic, rather than conservative, in describing a confidence region.

The relative behaviours identified in Table 1 are reflected in the minimum and maximum confidence levels of Table 2. The tendency for ν_{tv} to be slightly conservative, on average, is in keeping with very high confidence levels for some parameter settings, in some cases levels of nearly 100%. Nevertheless, the confidence level with ν_{tv} can be significantly below nominal in some cases. The range of confidence levels for ν_{gv} and ν_{hy} does extend above the nominal. However, the tabulated values reinforce the impression that ν_{gv} and ν_{hy} tend to describe confidence regions that give a level less than 95%.

The minima for ν_{hy} appear relatively close to the minima for ν_{tr} whereas the maxima for ν_{hy} appear close to those for ν_{gv}. So the range of confidence levels has been narrowed by using the hybrid method. Minimum levels for ν_{tv} and ν_{hy} become considerably lower as the sample sizes became increasingly different.

There is no clear winner amongst the three methods according to these results. The ν_{hy} method, because it has narrower ranges, might be considered the more reliable of the three. Nonetheless, there are some practical disadvantages with ν_{hy} (and ν_{gv}) associated with the subsequent use of results by others. This will be discussed in the next section. The ν_{tv} method is attractive because it tends to generate larger than nominal levels of confidence, though consequently the confidence regions are larger than is necessary. This conservative tendency is perhaps desirable, although it is unfortunate that it is not guaranteed. In practice, the sample sizes will be known, so the experimenter may choose the method accordingly.

A further short simulation was performed to investigate the behaviour of the different methods as the dimensionality of the problem increases. The matrices Σ_1 and Σ_2 were both set equal to the p-dimensional identity matrix, and the sample sizes were fixed at $n_1 = n_2 = 8$. The range of dimensions from $p = 1$ to $p = 7$ was investigated.

The results for 10 000 simulated experiments are shown in Table 3. The performance of the hybrid method is close to nominal, which is satisfactory. The total variance method is also quite well-behaved. It displays a slight conservative tendency as the value of p approaches the sample size. In contrast, the generalized variance performs badly for p close to the sample size.

5 Discussion

Ideally the results of the procedure would be unaffected by carrying out the estimation in several stages involving arbitrary groupings. This property of transitiveness is desirable so that the derivation of combined uncertainty and effective degrees of freedom can be regarded as part of a chain of uncertainty propagation. This holds in the one-dimensional case described by the WS formula.

Table 3: Confidence levels ($\times 10^4$) for a standard case at various dimensions

p	ν_{tv}	ν_{gv}	ν_{hy}
1	9517	9517	9517
2	9545	9526	9532
3	9575	9531	9552
4	9598	9514	9563
5	9618	9408	9541
6	9665	7703	9538
7	9690	5776	9535

The labels $i = 1, \ldots, m$ can obviously be assigned arbitrarily. So, if the property of transitiveness also held in p-dimensions, the above confidence region for $\sum_{i=1}^{m} \boldsymbol{\mu}_i$ would be unaffected by first deriving the combined uncertainty and effective degrees of freedom for $\boldsymbol{\mu}^* \equiv \sum_{i=1}^{k-1} \boldsymbol{\mu}_i$, for any k, and then finding the confidence region for $\boldsymbol{\mu}^* + \sum_{i=k}^{m} \boldsymbol{\mu}_i$. The method giving ν_{tv} satisfies this because it involves matching the sums of the variances of elements on either side of (7), and this sum is preserved no matter how the $\mathbf{S}_i/n_i$ matrices are grouped.[7] Unfortunately the method giving ν_{gv} does not satisfy this,[8] as is easily demonstrated by counter-example. It follows that ν_{hy} also does not have this property, even discounting the effect of the median function. Therefore, if the combined uncertainty matrix and effective degrees of freedom are merely inputs to a further calculation then ν_{tv} should be quoted.

An important requirement of an estimation procedure is that is produces short confidence intervals or small confidence regions. However this requirement is secondary to achieving a confidence close to or above the stated level, 0.95. The only difference between the alternative methods is the choice of the effective degrees of freedom. Other things being equal, the size of the confidence region will be smaller with larger effective degrees of freedom. So we can regard the average sizes of the confidence regions for ν_{tv}, ν_{gv} and ν_{hy} as being inversely related to the average confidence levels.

Naturally, corresponding components of the m added vectors, $\{\boldsymbol{\mu}_i\}$, must be expressed in the same units. Less obviously, the p components of each vector are best expressed in the same units for unambiguous application of the method. Otherwise a rescaling in one or more dimensions would change the confidence ellipsoid in the p-dimensional space. For example, suppose we wished to construct a confidence region for a complex-valued impedance by combining estimates obtained from several samples. The real component, for example, would appear relatively more variable when expressed in milli-ohms than in ohms. So, it would have greater influence over the imaginary component in the matching of total variances or generalized variances. This would affect the elliptical confidence region in a way not simply equated with rescaling.

Despite this effect, the method can be used with components in different units to give proper confidence regions provided that the units are chosen before observation of the data. In this way, we do not violate the classical statement that the region to be calculated has probability (approximately) 0.95 of containing the value of the measurand.

The method is derived assuming that the estimates of the m input quantities are in-

[7] The identity $\mathrm{tr}\,(\mathbf{A}+\mathbf{B}) = \mathrm{tr}\,\mathbf{A} + \mathrm{tr}\,\mathbf{B}$ seems relevant.

[8] This seems linked to the fact that, in general, $\det(\mathbf{A}+\mathbf{B}) \neq \det \mathbf{A} + \det \mathbf{B}$.

dependent. When the only correlations between inputs are known to be perfect, i.e. $\rho = \pm 1$, the problem could be formulated in terms where the method was appropriate. However, in the case of general non-zero correlations, the method is inapplicable. This is also true in the one-dimensional case with the WS formula.

In conclusion, the method presented for the analysis of uncertainty in a multidimensional measurand gives confidence regions with actual confidence levels close to the nominal value of 95%. There are various ways of defining the effective degrees of freedom. Based on the simulations for two-dimensional data, the best performing method on average when the input estimates have low degrees of freedom is the hybrid figure ν_{hy}. However, the average performance of the simpler measure ν_{tv} is sufficiently good unless the input degrees of freedom are very low. This figure, ν_{tv}, also allows the overall procedure to be unambiguous when the analysis is viewed as part of a chain of uncertainty propagation.

6 Appendix 1 - Two-dimensional case

When $p = 2$ the formulae are easy to implement. Define

$$\Delta \equiv \sum u_{i\cdot 11} \sum u_{i\cdot 22} - \left(\sum u_{i\cdot 12}\right)^2,$$

where summation is over $i = 1$ to $i = m$. The 2×2 symmetric matrix $\mathbf{r} \equiv (r_{jk})$ where

$$\begin{aligned} r_{11} &= \sum u_{i\cdot 22}/\Delta \\ r_{12} = r_{21} &= -\sum u_{i\cdot 12}/\Delta \\ r_{22} &= \sum u_{i\cdot 11}/\Delta \end{aligned}$$

is the inverse of $\sum \mathbf{u}_i$. Suppose $\boldsymbol{x}_i \equiv (x_{i\cdot 1}, x_{i\cdot 2})'$. Then an approximate 95% confidence region for $\sum_{i=1}^m \boldsymbol{\mu}_i$ is the elliptical region of all column vectors $(\sum x_{i.1} + d_1, \sum x_{i.2} + d_2)'$ for which

$$r_{11} d_1^2 + 2 r_{12} d_1 d_2 + r_{22} d_2^2 \leq \frac{2\nu_{\text{eff}}\, F_{2,\nu_{\text{eff}}-1,0.95}}{\nu_{\text{eff}} - 1}.$$

Also $q = 3$, the elements of $\hat{\boldsymbol{\Theta}}$ are

$$\begin{aligned} A \equiv \hat{\theta}_{11} &= 2\left(\sum u_{i\cdot 11}\right)^2 \\ B \equiv \hat{\theta}_{12} &= 2\sum u_{i\cdot 11} \sum u_{i\cdot 12} \\ C \equiv \hat{\theta}_{13} &= 2\left(\sum u_{i\cdot 12}\right)^2 \\ D \equiv \hat{\theta}_{22} &= \sum u_{i\cdot 11} \sum u_{i\cdot 22} + \left(\sum u_{i\cdot 12}\right)^2 \\ E \equiv \hat{\theta}_{23} &= 2\sum u_{i\cdot 22} \sum u_{i\cdot 12} \\ F \equiv \hat{\theta}_{33} &= 2\left(\sum u_{i\cdot 22}\right)^2 \end{aligned}$$

and the elements of $\hat{\mathbf{\Lambda}}$ are

$$\begin{aligned}
a \equiv \hat{\lambda}_{11} &= 2 \sum u_{i\cdot 11}^2/\nu_i \\
b \equiv \hat{\lambda}_{12} &= 2 \sum u_{i\cdot 11} u_{i\cdot 12}/\nu_i \\
c \equiv \hat{\lambda}_{13} &= 2 \sum u_{i\cdot 12}^2/\nu_i \\
d \equiv \hat{\lambda}_{22} &= \sum \left(u_{i\cdot 11} u_{i\cdot 22} + u_{i\cdot 12}^2\right)/\nu_i \\
e \equiv \hat{\lambda}_{23} &= 2 \sum u_{i\cdot 22} u_{i\cdot 12}/\nu_i \\
f \equiv \hat{\lambda}_{33} &= 2 \sum u_{i\cdot 22}^2/\nu_i .
\end{aligned}$$

Applying (12) or (14) gives

$$\nu_{\text{tv}} = \frac{A+D+F}{a+d+f},$$

while (13) gives

$$\nu_{\text{gv}} = \left(\frac{ADF + 2BCE - AE^2 - DC^2 - FB^2}{adf + 2bce - ae^2 - dc^2 - fb^2}\right)^{1/3}.$$

7 Appendix 2 - Derivation of WS formula

Much of the derivation of the WS formula may be familiar, but the steps are presented in some detail so that analogies can be observed with the development in Sections 2 and 3. Here, also, the derivation uses a statistical notation and style, but the final steps are given in the notation of the *Guide*.

Consider a random variable Z to have the normal distribution with unknown mean μ and unknown variance σ^2. This is denoted by $Z \sim N(\mu, \sigma^2)$, where $\sim$ can be read as "(is) distributed as". Let $\bar{Z}$ and S^2 be the random variables for the sample mean and sample variance of a random sample of size n. It is well known that, independently,

$$\bar{Z} - \mu \sim N(0, \sigma^2/n), \tag{15}$$

$$S^2/n \sim (\sigma^2/n)\chi_{n-1}^2/(n-1), \tag{16}$$

where χ_ν^2 is a variable with the chi-square distribution with ν degrees of freedom. Suppose that we require a confidence interval for the fixed parameter μ. The method generating the shortest interval involves the following result.

- *The ratio of a standard normal variable $N(0,1)$ to an independent variable distributed as $\sqrt{(\chi_\nu^2/\nu)}$ has Student's t-distribution with ν degrees of freedom. This can be written*
$$\frac{N(0,1)}{\sqrt{(\chi_\nu^2/\nu)}} \sim t_\nu. \tag{17}$$

Consequently the probability is 0.95 that the random interval

$$[\bar{Z} - t_{n-1,0.975}\, S/\sqrt{n},\ \bar{Z} + t_{n-1,0.975}\, S/\sqrt{n}]$$

encloses μ, where $t_{\nu,0.975}$ is the 0.975 quantile of the distribution of t_ν. If $\bar{z}$ and s^2 are the observed values of $\bar{Z}$ and S^2 then the quoted 95% confidence interval for μ is the interval comprising all values of μ_0 for which

$$\frac{|\bar{z}-\mu_0|}{s/\sqrt{n}} \le t_{n-1,0.975}. \tag{18}$$

Equations (15), (16) and (18) are analogous to (1), (2) and (5).

More generally, let there be m normal variables $\{Z_i\}$ with means $\{\mu_i\}$ and variances $\{\sigma_i^2\}$, with sample sizes $\{n_i\}$. Suppose that we require a confidence interval for $\sum \mu_i$, where summation is from $i = 1$ to $i = m$. In order to use a t-distribution we look for a ratio having the form of (17). Clearly

$$\sum \bar{Z}_i - \sum \mu_i \sim N(0, \sum \sigma_i^2/n_i), \tag{19}$$

and dividing this variable by $\sqrt{(\sum \sigma_i^2/n_i)}$ gives a standard normal variable. However, in general no function of all the S_i^2/n_i variables has an exact chi-square distribution, so it is not possible to obtain a ratio with an exact t-distribution, and an approximation is required. We look for a multiplier c and an effective number of degrees of freedom ν such that

$$\sum S_i^2/n_i \overset{\text{approx}}{\sim} (c \sum \sigma_i^2/n_i)\, \chi_\nu^2/\nu. \tag{20}$$

This approximation implies

$$\frac{\sum \bar{Z}_i - \sum \mu_i}{\sqrt{\sum S_i^2/n_i}} \overset{\text{approx}}{\sim} \frac{t_\nu}{\sqrt{c}}, \tag{21}$$

which can be used to find an approximate confidence interval for the value of the measurand $\sum \mu_i$. Equations (19)–(21) are analogous to (6)–(8).

The WS solution chooses c and ν so that the means and variances of the variables on each side of (20) match. The mean and variance of χ_ν^2 are ν and 2ν. It follows that $c = 1$ and

$$\nu = \frac{(\sum \sigma_i^2/n_i)^2}{\sum (\sigma_i^2/n_i)^2/(n_i - 1)}.$$

In practice, the unknowns $\{\sigma_i^2\}$ are replaced by their sample estimates $\{s_i^2\}$. Following the notation of the *Guide*, the best estimate of μ_i is x_i, its standard uncertainty is u_i, the associated degrees of freedom is ν_i and the resulting effective degrees of freedom is ν_{eff}. We associate x_i with $\bar{z}_i$, u_i with $s_i/\sqrt{n_i}$ and ν_i with $n_i - 1$, define the combined uncertainty $u_{\text{c}} \equiv \sqrt{\sum u_i^2}$ and so obtain

$$\nu_{\text{eff}} = \frac{u_{\text{c}}^4}{\sum u_i^4/\nu_i}, \tag{22}$$

which is a representation of the WS formula. Also define $k_p \equiv t_{\nu_{\text{eff}},0.975}$. So an approximate 95% confidence interval for the value of the measurand, $\sum \mu_i$, is

$$\left[\sum x_i - k_p u_{\text{c}},\ \sum x_i + k_p u_{\text{c}}\right],$$

which is a familiar result to users of the *Guide*.

References

[1] *Guide to the Expression of Uncertainty in Measurement*, Geneva, International Organisation for Standardization, 1995.

[2] `http://www.bipm.fr/enus/2_Committees/JCGM.shtml`

[3] Welch B. L., The significance of the difference between two means when the population variances are unequal. *Biometrika*, 1937, **29**, 350–362.

[4] Satterthwaite F. E., An approximate distribution of estimates of variance components. *Biometrics*, 1946, **2**, 110–114.

[5] Welch B. L., The generalization of "Student's" problem when several different population variances are involved. *Biometrika*, 1947, **34**, 28–35.

[6] Gupta A. K., Nagar D. K., *Matrix Variate Distributions*, Boca Raton, Chapman & Hall/CRC, 2000, 367 p.

[7] Johnson R. A., Wichern D. W., *Applied Multivariate Statistical Analysis*, 4th ed., Prentice-Hall, 1998, 816 p.

[8] Tan W. Y., Gupta R. P., On Approximating a Linear Combination of Central Wishart Matrices with Positive Coefficients *Comm. Statist.-Theor. Meth.*, 1983, **12**(22), 2589–2600.