

Comment on ‘Empirical *versus* modelling approaches to the estimation of measurement uncertainty caused by primary sampling’ by J. A. Lyn, M. H. Ramsey, A. P. Damant and R. Wood

B. Geelhoed*

First published as an Advance Article on the web 9th July 2009

DOI: 10.1039/b812422a

Recently, Lyn *et al.* (*Analyst*, 2007, **132**, 1231) compared two ways of estimating the standard uncertainty of sampling pistachio nuts for aflatoxins – a modelling method and an empirical method. Their case study used robust analysis of variance (RANOVA) to derive the uncertainty estimates, highlighting a substantial difference between the two: the estimate of sampling uncertainty derived from the modelling method was six-fold greater than that using the empirical approach (*cf.* 136% and 22.5%, respectively, when expressed as relative standard deviations (RSDs) at 68% confidence). A further analysis of this case study is reported here and suggests that the estimation uncertainty during RANOVA in the empirical approach could account for this difference.

Lyn *et al.*¹ addressed the question of what the most appropriate method for estimating measurement uncertainty is. The authors expressed their preference for the empirical approach by stating that the empirical approach to uncertainty estimation is recognized as generally the most practical procedure providing the more reliable estimates.¹ In response to this, a further analysis of the case study presented by Lyn *et al.*¹ is given here.

In the above-mentioned case study, robust analysis of variance (RANOVA) was performed on the 32 analytical results obtained after duplicate analysis of each sample in eight pairs of samples (where each pair constitutes a sample duplet), estimating the Relative Standard Deviation (RSD) caused by sampling as 22.5% in the empirical approach. This value was then compared with the much higher RSD of 136% predicted by the modelling approach by Gy's model for the fundamental sampling error in the absence of grouping and segregation. To explain the difference, the authors stated in their article that the difficulty in establishing reliable estimates for the input variable for the modelling approach is thought to be the main cause of the discrepancy. They mentioned that empirical measurements are subject to measurement uncertainty,¹ but they did not further explore estimation uncertainty associated with the result of RANOVA as one of the possible causes of the difference. Therefore, this is studied here, based on the information that was provided in the article by Lyn *et al.*:¹ the average aflatoxin concentration in the sample population is estimated by $a = 0.86 \mu\text{g kg}^{-1}$; the concentration of aflatoxin in a contaminated nut is assumed to be $\alpha = 250 \mu\text{g kg}^{-1}$; and the mass of a single nut is given by the product of volume and density and is estimated by $1.42 \text{ cm}^3 \times 0.9951 \text{ g cm}^{-3} = 1.41 \text{ g}$. Given a sample mass of 250 g, this results in samples containing $N = 177$ pistachio nuts. The fraction of contaminated nuts (p) can be calculated using the ratio of the concentration in the population and the

concentration in a contaminated nut. The average number of contaminated particles in a sample is then calculated as $Np = 0.61$.

Modelling sampling using a binomial process, with N selections, where during each selection there is a probability (p) of selecting a contaminated nut and a probability ($1 - p$) of selecting an uncontaminated nut, the relative standard deviation of the concentration in a sample is given by $\sqrt{((1 - p)/(Np))}$. For the population under study the RSD is thus evaluated as 128%. This is slightly smaller than the value obtained with Gy's theory (136%), mainly due to ignoring the small variations in particle masses, which are taken into account in the (more general) model of Gy.

In order to assess estimation uncertainty associated with the estimate of RSD obtained by RANOVA in the empirical approach, computer simulations are performed that mimic the experimental design described in the section “Empirical estimation of sampling uncertainty” by Lyn *et al.*¹ Independent series of eight random independent duplicate samples are repeatedly simulated using the above-mentioned binomial process. For simplicity, it is assumed that analytical errors are absent, so that the replicate analysis results for each sample will be equal. For each data set obtained in this way (each consisting of 32 values), RANOVA is performed to estimate the RSD caused by sampling. In order to match the computational procedure used in the empirical approach, the program ROBCOOP4.exe (which was downloaded from www.rsc.org/suppdata/JA/1998/97/index.sht) was selected, because the use of ROBCOOP4.exe is suggested in ref. 26 of Lyn *et al.*¹ The entire process of generating 32 “analytical results” and performing RANOVA on these data is repeated 10^4 times and each time (if ROBCOOP4.exe did not report an error) the value for RSD as estimated with RANOVA is recorded. This allows the scatter of RSD estimates to be assessed, which is indicative of the uncertainty associated with the estimate of RSD obtained using the empirical method as implemented by Lyn *et al.*¹ under the conditions of the case

Delft University of Technology, Mekelweg 15, 2629 JB Delft, The Netherlands. E-mail: B.Geelhoed@tudelft.nl

Table 1 A selection of sets of 32 “analytical results” obtained during the simulations out of the total number of 10^4 of such sets. Columns with $SXA Y$ give the Y -th analytical result ($Y = 1$ for the first and $Y = 2$ for the second) for the X -th sample ($X = 1$ for the first and $X = 2$ for the second) of a sample duplet (there are eight duplets for each set). Simulated analytical results are in units of “number of nuts in a sample”. In order to transform these units into units of $\mu\text{g kg}^{-1}$, a multiplication by 250/177 (the concentration in units of $\mu\text{g kg}^{-1}$ in a contaminated nut divided by the number of nuts in a sample) must be applied. This scaling does not affect the estimated RSD, as the estimated RSD is a ratio and, consequently, the scaling factor in the numerator and denominator of this ratio cancel against each other

Set 31				
Sample duplet	S1A1	S1A2	S2A1	S2A2
1	0	0	0	0
2	0	0	1	1
3	0	0	0	0
4	0	0	0	0
5	1	1	1	1
6	1	1	0	0
7	0	0	0	0
8	1	1	0	0
Results: robust mean = 0.2876471				
$s_{\text{sampling}} = 0.491$				
Estimated RSD = $s_{\text{sampling}}/(\text{robust mean}) \times 100 = 170.695272\%$				
Set 1375				
Sample duplet	S1A1	S1A2	S2A1	S2A2
1	0	0	0	0
2	0	0	0	0
3	1	1	1	1
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	1	1
This data set leads to a run-time error in ROBCOOP4.exe				
Set 6763				
Sample duplet	S1A1	S1A2	S2A1	S2A2
1	0	0	0	0
2	0	0	0	0
3	1	1	2	2
4	0	0	0	0
5	0	0	1	1
6	0	0	0	0
7	1	1	1	1
8	0	0	0	0
Results: robust mean = 0.3222708				
$s_{\text{sampling}} = 0.000$				
Estimated RSD = $s_{\text{sampling}}/(\text{robust mean}) \times 100 = 0.000000\%$				

study. To give a concrete impression of the here-performed simulations, in Table 1, some selected data sets out of the total number of 10^4 sets obtained are given, together with the estimated RSD for the data set (the estimated RSD is only given if ROBCOOP4.exe did not report an error). In Fig. 1, the distribution of the obtained estimates for RSD is shown in

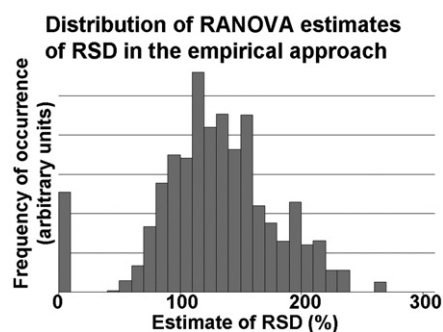


Fig. 1 A histogram of the distribution of RSDs calculated using RANOVA (using ROBCOOP4.exe) during the simulations.

a histogram. It is noted that the distribution is based on less than 10^4 results (9736 results), because there were some cases in which ROBCOOP4.exe reported a “run-time error” and these cases were therefore discarded here. Data set 1375 in Table 1 provides an example of a data set that produces a run-time error in ROBCOOP4.exe. A significant number of data sets for which the estimated RSD is zero are obtained. Data set 6763 in Table 1 provides an example of one such case.

The average of the series of 9736 estimates for RSD is 129.3% and the standard deviation is 49.9%. The uncertainty (using a coverage factor of two) in estimating RSD using RANOVA in the empirical approach by Lyn *et al.*¹ is therefore assessed to be $2 \times 49.9\% = 99.8\%$ and is likely to be a significant contributing factor to the difference between the modelling and empirical estimates of RSD ($136\% - 22.5\% = 113.5\%$). The fraction of values smaller than 22.5% (the estimate in the empirical approach obtained by Lyn *et al.*¹ using RANOVA) is also significant: 0.051 (or 5.1% when expressed as a percentage). In view of the above results, which are based on the available data combined with a binomial process as a model for the sampling, it can be concluded that the observed difference between the empirical estimate of RSD and the modelling estimate during the case study described by Lyn *et al.*¹ may have originated due to estimation uncertainty during RANOVA in the empirical approach. In general, ignoring estimation uncertainty in the result of RANOVA might lead to unreliable conclusions.

As a concluding remark, it is noted here that an exact estimator for sampling uncertainty will probably remain unachievable, whether an empirical, modelling or other approach is used. Therefore, different approaches should not be seen as competing, but rather as complementary. The co-existence of a variety of methods to estimate sampling uncertainty should be welcomed in view of the complex reality of material sampling.

References

- 1 J. A. Lyn, M. H. Ramsey, A. P. Damant and R. Wood, Empirical versus modelling approaches to the estimation of measurement uncertainty caused by primary sampling, *Analyst*, 2007, **132**, 1231–1237.