Designing a Measurement Precision Experiment considering Distribution of Estimated Precision Measures

T Suzuki¹,³, J Takeshita², J Ono¹ and X Lu¹

¹ Tokyo University of Science, 2641 Yamazaki, Noda, Chiba, 278-8510, Japan
² National Institute of Advanced Industrial Science and Technology, 16-1, Onogawa, Tsukuba, Ibaraki, 305-8569, Japan

E-mail: szk@rs.tus.ac.jp

Abstract. Evaluating performance of a measurement method is essential in metrology. Concepts of repeatability and reproducibility are introduced in ISO 5725 series including experiments (usually collaborative studies) to obtain these precision measures. ISO 5725 series mainly describes managerial part on how to run an experiment, but there is only a little guidance on statistical aspects of the experimental design. In such an experiment, accuracy of these precision measures is important, and the experimenter should be aware of biases and precisions of these precision measures. The objective of such an experiment is not limited to simply estimate the precision measures but to obtain these measures with intended accuracy. This paper proposed how to design a measurement precision experiment from statistical point of view. This paper first investigated the bias and precision of the precision measures based on their distributions and then proposed how to design an experiment with the desired accuracy.

1. Introduction
Evaluating performance of a measurement method is essential in metrology. Concepts of repeatability and reproducibility are introduced in ISO 5725 series including experiments (usually collaborative studies) to obtain these precision measures. ISO 5725 series mainly describes managerial part on how to run an experiment, but there is only a little guidance on statistical aspects of the experimental design. The objective of such an experiment is not limited to simply estimate the precision measures but to obtain these measures with intended accuracy. This paper proposed how to design a measurement precision experiment from statistical point of view. This paper first investigated the bias and precision of the precision measures based on their distributions and then proposed how to design an experiment with the desired accuracy.

2. Evaluation of Measurement Performance

2.1. Basic Concept and Statistical Model
The basic model adopted in ISO 5725 series can be expressed as
where $y_{jk}$ denotes the measurement result for $k$th repetition in $j$th laboratory, $\mu$ the general mean, $\alpha$ the laboratory effect, $\epsilon$ the error term. The subscripts are $j = 1, 2, \ldots, b$ where $b$ is the number of laboratories and $k = 1, 2, \ldots, n$ where $n$ is the number of repetitions within a laboratory. Variance of $\alpha$ is called between-laboratory variance and is denoted by $\sigma^2_\alpha$. Variance of $\epsilon$ is denoted $\sigma^2_\epsilon$ and is within-laboratory variance. More importantly this is called repeatability variance and also written as $\sigma^2_\epsilon$. $\sigma^2_\epsilon$ is the variance under repeatability conditions, that is, 'conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short interval of time'. Repeatability standard deviation is the square root of the repeatability variance. Reproducibility variance is defined by

$$
\sigma^2_R = \sigma^2_\alpha + \sigma^2_\epsilon = \sigma^2_L + \sigma^2_E,
$$

and is the variance under reproducibility conditions, that is, 'conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipments'. Reproducibility standard deviation is the square root of the reproducibility variance.

2.2. Estimation of Precision Measures

Repeatability variance and between-laboratory variance are estimated using formulae in calculating one way ANOVA (analysis of variance) table shown in table 1.

**Table 1. One way ANOVA Table.**

| Source | Sum of Squares | Degree of Freedom | Mean Square | $F_0$ | Expected Mean Square |
|--------|----------------|-------------------|-------------|------|----------------------|
| $B$ (Laboratory) | $SSB = n \sum_{j=1}^{b} (\bar{y}_j - \bar{y})^2$ | $b - 1$ | $MSB = \frac{SSB}{b - 1}$ | $SSB$ | $\frac{SSB}{MSE}$ | $\sigma^2_\alpha + n\sigma^2_L$ |
| $E$ (Repetition) | $SSE = \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{jk} - \bar{y}_j)^2$ | $b(n - 1)$ | $MSE = \frac{SSE}{b(n - 1)}$ | $SSE$ | $\frac{SSE}{MSE}$ | $\sigma^2_\epsilon$ |
| $T$ (Total) | $SST = \sum_{j=1}^{b} \sum_{k=1}^{n} (\bar{y}_{jk} - \bar{y})^2$ | $bn - 1$ | | | |

In table 1, $\bar{y}_j$ stands for the average of all the results of $j$th laboratory and $\bar{y}$ stand for the overall average of all the measurement results. The estimate of the repeatability variance is denoted by $\hat{\sigma}^2_\epsilon$ and its estimate is given by

$$
\hat{\sigma}^2_\epsilon = \hat{\sigma}^2_\epsilon = MSE.
$$

The estimate of the between-laboratory variance is denoted by $\hat{\sigma}^2_L$ and its estimate is given by

$$
\hat{\sigma}^2_L = \frac{MSB - MSE}{n}.
$$

Using the two equations above, the estimate of the reproducibility variance, which is denoted by $\hat{\sigma}^2_R$ is given by

$$
\hat{\sigma}^2_R = \hat{\sigma}^2_L + \hat{\sigma}^2_\epsilon.
$$

These three precision measures are used to evaluate performance of measurement methods.
3. Distribution of Estimates of Precision Measures

3.1. Negative Estimates of Between-laboratory Variance

In precision experiments, negative values of the estimate of between-laboratory variance are sometimes obtained due to taking differences between sums of squares as in equation (4). In ISO 5725-2, it is recommended to use value zero instead of the estimated negative values, which leads to bias in the estimated between-laboratory variance. In other words, the expected value of the between-laboratory variance becomes larger than the true between-laboratory variance. The same can be said for the reproducibility variance because it is estimated as the sum of the estimated repeatability variance and the estimated between-laboratory variance as in equation (5). However, it is not discussed in ISO 5725-2 how large the bias and its effect will be. We focused on the bias.

The probability of occurrence of negative estimates of between-laboratory variance is denoted by \( P \). The negative estimates occur when \( MSB \) is smaller than \( MSE \) in equation (5). The probability \( P \) can be expressed by

\[
P = Pr(\hat{\sigma}_L^2 < 0) = Pr \left( F < \frac{1}{1+n\Delta} \right),
\]

where

\[
F = \frac{MSB/(\sigma_r^2 + n\sigma_L^2)}{MSE/\sigma_r^2} \sim F(b - 1, b(n - 1))
\]

and \( \Delta = \sigma_L/\sigma_r \). Values of \( P \) are calculated. Figure 1 shows the results for \( \Delta = 0.0 \) and \( \Delta = 0.5 \), with the number of laboratories \( b = 2, 6, 10, 14, 18 \) and the number of repetitions \( n = 2, 3, \ldots, 40 \).

![Figure 1. Probability of negative estimates of between-laboratory variance.](image)

From figure 1(a), we can understand that \( P \) becomes larger with larger \( n \) or smaller \( b \) when \( \Delta = 0.0 \). We can also understand that \( P \) is never less than 0.5 in this case. From figure 1(b), we can understand that \( P \) becomes smaller with larger \( n \) or larger \( b \).

3.2. Bias in Estimating Between-laboratory Standard Deviation

Bias in the estimate of between-laboratory standard deviation is denoted by \( \delta_L \) and can be defined as

\[
\delta_L = E(\hat{\sigma}_L) - \sigma_L.
\]

We know that from equation (4), \( \hat{\sigma}_L \) is based on the difference of two independent chi-squared variates. Calculation of \( E(\hat{\sigma}_L) \) cannot be done readily because it includes some special function which require calculation of the sum of the infinite series.[4] Thus we performed a simulation. For each setting of parameters, a pair of two random chi-squared variates were generated one million times. The average
of one million \( \hat{\sigma}_b \)'s was used to represent \( E(\hat{\sigma}_b) \). The simulation was run for \( \Delta = 0.0, 0.1, \ldots, 1.0, b = 2, 3, \ldots, 40 \) and \( n = 2, 3, \ldots, 40 \), where \( \Delta = \sigma_b / \sigma_r \). The results for \( \Delta = 0.0, 0.5 \) and 1.0 are shown in figure 2. The biases here are compared with \( \sigma_r \) considering practical situations into account.

![Figure 2](image)

**Figure 2.** Bias (relative to \( \sigma_r^2 \)) in estimates of between-laboratory standard deviation.

From figure 2(a), we can understand that we have positive bias when \( \Delta = 0.0 \), which means replacing the negative estimates are the main cause in this case. From figure 2(b) and 2(c), we can understand that we have negative bias when \( \Delta = 0.5 \) and \( \Delta = 1.0 \), which means the main causes are that the square root of the sample variance is not the unbiased estimator of the population standard deviation. The amount of the bias is always largest at \( b = 2 \) in all cases, but on the other hand, the effect of \( n \) varies in each case. The bias was positive when \( \Delta = 0.2 \) or smaller and negative when \( \Delta = 0.3 \) or larger.

The same procedure was also applied to the estimates of the reproducibility standard deviation and we obtained the similar results.

3.3. Design of Experiments based on distribution of precision measures

From the results of the simulations performed in section 3.2, we obtained the distribution of the precision measures, namely, between-laboratory variance, between-laboratory standard deviation, reproducibility variance and reproducibility standard deviation. These distributions were then applied to design an experiment with the desired accuracy. The desired accuracy is defined by setting, say, reproducibility variance relative to repeatability variance, or maximum bias allowed relative to repeatability standard deviation, or width of confidence interval. The necessary number of participating laboratories \( b \) and/or number of repetitions within laboratory \( n \) are derived using the distribution.

4. Conclusions

This paper investigated the bias and precision of the precision measures. The analysis revealed some new interesting findings which we believe is useful in precision experiments. The results of the simulation were utilized to obtain the distribution of the precision measures, with which the design of experiment with the desired accuracy became possible.

References

[1] International Organization for Standardization 1994 *ISO 5725-1 Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions* (Geneva: ISO)

[2] International Organization for Standardization 1994 *ISO 5725-2 Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method* (Geneva: ISO)

[3] Searle S R, Casella G and McCulloch C E 1992 *Variance Components* (USA: Wiley-Interscience)

[4] Robinson J 1965 The distribution of a general quadratic form in normal variates *Australian J. Stat.* 7(3) 110-114