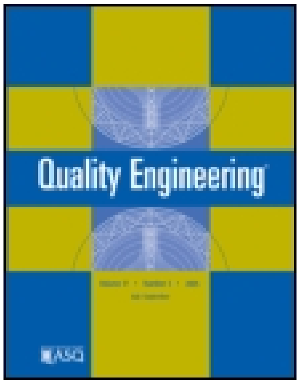

Quality Engineering
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lqen20

GAUGE CAPABILITY ANALYSIS AND DESIGNED EXPERIMENTS. PART II: EXPERIMENTAL DESIGN MODELS AND VARIANCE COMPONENT ESTIMATION

DOUGLAS C. MONTGOMERY [a] & GEORGE C. RUNGER [b]

[a] College of Engineering and Applied Sciences Arizona State University , Tempe, Arizona, 85287-5506

[b] School of Engineering Rensselaer Polytechnic Institute , Troy, New York, 12180-3590

Published online: 03 Apr 2007.

Quality Engineering, 6(2), 289-305 (1993-94)

GAUGE CAPABILITY ANALYSIS AND DESIGNED EXPERIMENTS. PART II: EXPERIMENTAL DESIGN MODELS AND VARIANCE COMPONENT ESTIMATION

DOUGLAS C. MONTGOMERY
College of Engineering and Applied Sciences
Arizona State University
Tempe, Arizona 85287-5506

GEORGE C. RUNGER
School of Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180-3590

Key Words

Repeatability; Reliability; Analysis of variance; Confidence intervals.

Introduction

In Part I of this paper (1) we reviewed the classical gauge repeatability and reproducibility (R&R) study and noted its equivalence to an experiment designed to estimate certain components of variance. We focused on several practical aspects of designing and conducting a gauge R&R study and noted that, under the condition of interaction between parts and operators, the classical method gives a downwardly biased estimator of the variance component representing gauge reproducibility. Furthermore, this interaction can be easily investigated in the experimental design–analysis of variance format. This leads us to recommend the experimental design approach as a replacement for the classical R&R analysis.

In some problems, this approach will produce variance component estimates that are negative; in fact, this will always occur whenever MS_{OP} is small relative to MS_R. In Part I, we suggested a modified analysis-of-variance estimation procedure for this situation based on hypothesis testing about the individual variance components and fitting reduced models in which terms associated with variance components that do not differ significantly from zero are deleted. In this article, we illustrate other estimation methods that can be used to produce non-negative estimates and discuss their appropriateness.

We also show how to find confidence intervals on the variance components of interest, $\sigma^2_{\text{repeatability}}$ and $\sigma^2_{\text{reproducibility}}$. We use these results to make recommendations concerning sample sizes for designing gauge capability studies. Finally, we discuss other experimental design models for gauge capability studies, such as those incorporating nested factors.

The Factorial Design Model

As noted in Part I of this paper (1) and in Ref. 2., the classical gauge R&R study can be viewed as a two-factor factorial experiment, random effects model. More specifically, if there are o operators and p parts and if each operator makes n measurements on each part, then the appropriate statistical model is

$$X_{ijk} = \mu + O_i + P_j + (OP)_{ij} + R_{k(ij)}, \quad \begin{aligned} i &= 1, 2, \ldots, o, \\ j &= 1, 2, \ldots, p, \\ k &= 1, 2, \ldots, n \end{aligned} \tag{1}$$

where X_{ijk} is the kth measurement made by operator i on part j. The parameter μ represents an overall mean, factors O_i, P_j, and $(OP)_{ij}$ are the usual random effects factors, defined such that O_i is NID(0, σ^2_O), P_j is NID(0, σ^2_P), $(OP)_{ij}$ is NID(0, σ^2_{OP}), and $R_{k(ij)}$ is the replication effect, assumed to be NID(0, σ^2_R). For more details of this model and the assumptions, see Ref. 3.

The expected values of the mean squares from the analysis of variance are

$$\begin{aligned} \mathrm{E}(\mathrm{MS}_R) &= \sigma^2_R \\ \mathrm{E}(\mathrm{MS}_{OP}) &= \sigma^2_R + n\sigma^2_{OP} \end{aligned} \tag{2}$$

$$\begin{aligned} E(MS_P) &= \sigma_R^2 + n\sigma_{OP}^2 + on\sigma_P^2 \\ E(MS_O) &= \sigma_R^2 + n\sigma_{OP}^2 + pn\sigma_O^2 \end{aligned}$$

and the corresponding analysis-of-variance estimates of the variance components are

$$\begin{aligned} \hat{\sigma}_R^2 &= MS_R \\ \hat{\sigma}_{OP}^2 &= (MS_{OP} - MS_R)/n \\ \hat{\sigma}_P^2 &= (MS_P - MS_{OP})/on \\ \hat{\sigma}_O^2 &= (MS_O - MS_{OP})/pn \end{aligned} \tag{3}$$

As noted in Ref. 4, the normality assumption is not required to produce the estimates in Eq. (3). Furthermore, these are unbiased estimates of the corresponding variance components. Because $\sigma^2_{\text{repeatability}} = \sigma_R^2$, MS_R is an unbiased estimator of gauge repeatability; that is, $\hat{\sigma}^2_{\text{repeatability}} = MS_R$. The variance component corresponding to gauge reproducibility is

$$\sigma^2_{\text{reproducibility}} = \sigma_O^2 + \sigma_{OP}^2$$

There is no direct estimator of $\sigma^2_{\text{reproducibility}}$, in that no mean square from the analysis of variance has expected value $\sigma_O^2 + \sigma_{OP}^2$. However, an unbiased estimator of $\sigma^2_{\text{reproducibility}}$ is

$$\begin{aligned} \hat{\sigma}^2_{\text{reproducibility}} &= \hat{\sigma}_O^2 + \hat{\sigma}_{OP}^2 \\ &= (MS_O + (p-1)MS_{OP} - pMS_R)/pn \end{aligned} \tag{4}$$

An unbiased estimator of σ^2_{gauge} is

$$\begin{aligned} \hat{\sigma}^2_{\text{gauge}} &= \hat{\sigma}^2_{\text{repeatability}} + \hat{\sigma}^2_{\text{reproducibility}} \\ &= (MS_O + (p-1)MS_{OP} + p(n-1)MS_R)/pn \end{aligned} \tag{5}$$

Confidence Intervals

A major advantage of the experimental design–variance components approach to a gauge R&R study is that if the normality assumptions made in reference to the model of Eq. (1) are satisfied, then confidence intervals on $\sigma^2_{\text{repeatability}}$, $\sigma^2_{\text{reproducibility}}$, σ^2_{gauge}, and other functions of the variance components can be easily determined. These confidence intervals convey much more information than the usual point estimates, and they offer some insight into the sample size determination problem as well as other aspects of gauge capability analysis.

One always obtains a $100(1 - \alpha)\%$ confidence interval on $\sigma^2_{\text{repeatability}}$ from

$$\frac{df_E\,\hat{\sigma}^2_{\text{repeatability}}}{\chi^2_{\alpha/2,\,df_E}} \leqslant \sigma^2_{\text{repeatability}} \leqslant \frac{df_E\,\hat{\sigma}^2_{\text{repeatability}}}{\chi^2_{1-\alpha/2,\,df_E}} \tag{6}$$

where df_E is the error degrees of freedom in the design and $\chi^2_{\alpha,k}$ is the percentage point of the chi-square distribution with k degrees of freedom that has $\alpha\%$ of the probability to the right of $\chi^2_{\alpha,k}$. Unfortunately, there may not be exact confidence intervals for $\sigma^2_{\text{reproducibility}}$ or σ^2_{gauge}. This will always be the case when these variances are estimated from the full model in Eq. (1), using Eqs. (4) and (5). However, using results from Graybill (5) and Searle (4), approximate $100(1-\alpha)\%$ confidence intervals can be easily obtained. An approximate $100(1-\alpha)\%$ confidence interval on $\sigma^2_{\text{reproducibility}}$ is

$$\frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{\alpha/2,\,\nu}} \leqslant \sigma^2_{\text{reproducibility}} \leqslant \frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{1-\alpha/2,\,\nu}} \tag{7}$$

where

$$\nu = (\hat{\sigma}^2_{\text{reproducibility}})^2 \left[\frac{(1/np)^2\text{MS}_O^2}{o-1} + \frac{[(p-1)/np]^2\text{MS}_{OP}^2}{(o-1)(p-1)} + \frac{(1/n)^2\text{MS}_R^2}{op(n-1)} \right]^{-1} \tag{8}$$

In practice, ν will seldom be an integer, and the constant for the confidence interval will have to be found either by interpolation in the chi-square tables or by using the nearest integer to ν. The approximate $100(1-\alpha)\%$ confidence interval on σ^2_{gauge} is

$$\frac{u\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{\alpha/2,\,u}} \leqslant \sigma^2_{\text{gauge}} \leqslant \frac{u\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{1-\alpha/2,\,u}} \tag{9}$$

where

$$u = (\hat{\sigma}^2_{\text{gauge}})^2 \left[\frac{(1/np)^2\text{MS}_O^2}{o-1} + \frac{[(p-1)/np]^2\text{MS}_{OP}^2}{(o-1)(p-1)} + \frac{[(n-1)/n]^2\text{MS}_R^2}{op(n-1)} \right]^{-1} \tag{10}$$

Obviously, the results of Eqs. (9) and (10) can be used to define an approximate $100(1-\alpha)\%$ confidence interval on the P/T ratio because $P/T = 6\sigma_{\text{gauge}}/(\text{USL}-\text{LSL})$, where USL and LSL are the upper and lower specification limits, respectively. The confidence interval for P/T is

$$6\left[\frac{u\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{\alpha/2,\,u}}\right]^{1/2}(\text{USL}-\text{LSL})^{-1} \leqslant \frac{\hat{\sigma}_{\text{gauge}}}{\text{USL}-\text{LSL}} \leqslant 6\left[\frac{u\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{1-\alpha/2,\,u}}\right]^{1/2}(\text{USL}-\text{LSL})^{-1} \tag{11}$$

The intervals in Eqs. (7), (9), and (11) are appropriate whenever the variance component estimates are estimated from the full factorial model and the estimates are non-negative. If a negative estimate of $\sigma^2_{\text{reproducibility}}$ is obtained from the full model, fitting a submodel may result in changes to the confidence intervals.

The construction of these confidence intervals may be easily illustrated using the data from Ref. 3, analyzed in Tables 3 and 4 of Part I of this paper (1). The 95% confidence interval on $\sigma^2_{\text{repeatability}}$ is computed, using Table 4 of Part I, to be

$$\frac{98(0.88)}{127.28} = 0.68 \leqslant \sigma^2_{\text{repeatability}} \leqslant 1.19 = \frac{98(0.88)}{72.50}$$

However, the full model estimate of $\sigma^2_{\text{reproducibility}}$ is negative because MS_{OP} is less than MS_R, so Eq. (7) cannot be used. Table 4 of Part I of this paper shows that in the reduced model without the interaction term one may estimate $\sigma^2_{\text{reproducibility}}$ as

$$\hat{\sigma}^2_{\text{reproducibility}} = \hat{\sigma}^2_O$$

Therefore, finding a confidence interval on $\sigma^2_{\text{reproducibility}}$ is equivalent to finding a confidence interval on σ^2_O. Using the results in Milliken and Johnson (6), this confidence interval is

$$c \leqslant \sigma^2_{\text{reproducibility}} \leqslant d \tag{12}$$

where

$$c = \max\left\{\frac{(o-1)\text{MS}_O/\chi^2_{\alpha/2,o-1} - df_E\text{MS}_R/\chi^2_{1-\alpha/2,df_E}}{pn}, 0\right\} \tag{13}$$

and

$$d = \frac{(o-1)\text{MS}_O/\chi^2_{1-\alpha/2,o-1} - df_E\text{MS}_R/\chi^2_{\alpha/2,df_E}}{pn} \tag{14}$$

Equation (12) leads to an interval for which the confidence is at least $100(1-\alpha)\%$. Therefore, using Eqs. (12)-(14), a 95% confidence interval on $\sigma^2_{\text{reproducibility}}$ is

$$0 \leqslant \sigma^2_{\text{reproducibility}} \leqslant 1.00$$

The interval is relatively wide because only three operators were used in the study. Furthermore, a confidence interval for σ^2_{gauge} can be obtained for the reduced model by using an obvious modification of Eqs. (9) and (10). Now,

$$\frac{w\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{\alpha/2,w}} \leqslant \sigma^2_{\text{gauge}} \leqslant \frac{w\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{1-\alpha/2,w}} \tag{15}$$

where $\hat{\sigma}^2_{\text{gauge}} = \hat{\sigma}^2_{\text{repeatability}} + \hat{\sigma}^2_{\text{reproducibility}}$ and

$$w = (\hat{\sigma}^2_{\text{gauge}})^2 \left[\frac{(1/np)^2 \text{MS}_O^2}{o-1} + \frac{[(pn-1)/pn]^2 \text{MS}_R^2}{df_E} \right]^{-1} \tag{16}$$

From Table 4 in Part I of this series, $\hat{\sigma}^2_{\text{gauge}} = 0.88 + 0.01 = 0.89$, $w = 98.4$, and, from Eqs. (15) and (16),

$$0.69 \leqslant \sigma^2_{\text{gauge}} \leqslant 1.21.$$

Because $\hat{\sigma}^2_{\text{reproducibility}}$ is very small in this example, the resulting confidence interval is very similar to the interval for $\sigma^2_{\text{repeatability}}$.

The sensitivity of these confidence intervals to the number of levels of the factors and the amount of replication in the gauge capability study can be investigated by considering selected cases of the full model. For example, the precision of the confidence interval for $\sigma^2_{\text{repeatability}}$ in Eq. (6) is determined by the ratios

$$\frac{m}{\chi^2_{\alpha/2,m}} \quad \text{and} \quad \frac{m}{\chi^2_{1-\alpha/2,m}} \tag{17}$$

where $m = op(n-1)$ for the full model. These ratios are tabulated for selected values of m in Table 1. From Table 1, the improvement to the width of the confidence interval is small for m larger than approximately 50. For m equal to 50, the confidence interval is $0.70\hat{\sigma}^2_{\text{repeatability}}$ to $1.55\hat{\sigma}^2_{\text{repeatability}}$. By dividing by $\hat{\sigma}^2_{\text{repeatability}}$, the width can be expressed as 85% of $\hat{\sigma}^2_{\text{repeatability}}$. For $m = 100$, the width of the confidence interval is 58% of $\hat{\sigma}^2_{\text{repeatability}}$.

Because the width of the confidence interval depends on o, p, and n only through the product $op(n-1)$, increasing any one of these parameters reduces the width of the confidence interval. However, when n is small, the impact of

Table 1. Values Useful in Evaluation of Confidence Intervals for Variance Components

m	$\chi^2_{0.975,m}$	$m/\chi^2_{0.975,m}$	$\chi^2_{0.025,m}$	$m/\chi^2_{0.025,m}$
4	0.48	8.333	11.14	0.359
20	9.59	2.086	34.17	0.585
30	16.79	1.787	46.98	0.639
40	24.43	1.637	59.34	0.674
50	32.37	1.545	71.42	0.700
60	40.48	1.482	83.30	0.720
70	48.76	1.436	95.02	0.737
80	57.15	1.400	106.63	0.750
90	65.65	1.371	118.14	0.762
100	74.22	1.347	129.56	0.772
200	162.3	1.232	240	0.833

increasing n on the width of the confidence interval is greater than for p or o. For example, for the gauge study from Ref. 3, $p=20$, $o=3$, and $n=2$. Therefore, if we fit the full factorial model, $df_E=op(n-1)=60$. Doubling the number of observations by doubling o or p increases $op(n-1)$ to 120. However, doubling n to 4 increases $df_E = op(n-1)$ to 180. In this example, the effect of these changes on the width of the confidence interval is small, but for smaller experiments, the choice of which parameter to increase can be important.

From Eqs. (7) and (8), the precision of the confidence interval for $\sigma^2_{\text{reproducibility}}$ in the full model similarly depends on the ratio of the degrees of freedom of a chi-square distribution to the tabulated values of the chi-square distribution. Therefore, Table 1 is also useful in this case. However, the impact of selections for o, p, and n depends on the calculated value of ν. It is useful to consider some special cases. Consider the full factorial model case, and replace each mean square in Eq. (8) by the combination of variance components that it estimates. Then, we can show that the limit of ν as p tends to infinity is

$$\lim_{p\to\infty} \nu = (o-1)\left[1+\frac{\sigma^2_{OP}}{\sigma^2_O}\right] \tag{18}$$

Therefore, whenever the ratio of σ^2_{OP} to σ^2_P is near zero, the number of degrees of freedom in the confidence interval in Eq. (7) is approximately $o-1$. Traditionally, gauge capability studies use values for o of 5 or less, and often the ratio of σ^2_{OP} to σ^2_P is small. Therefore, traditional practice results in wide confidence intervals for $\sigma^2_{\text{reproducibility}}$ even for experiments with many parts. Furthermore, whenever the ratio of σ^2_{OP} to σ^2_O is small, the limit as n tends to infinity of ν is approximately $o-1$. Therefore, neither increasing the number of parts nor the number of replications will be very effective in improving the estimate of $\sigma^2_{\text{reproducibility}}$ when σ^2_{OP} is small relative to σ^2_O. For example, suppose that five operators are used in a gauge capability study. Assume that σ^2_{OP} is negligible relative to σ^2_O and the number of parts are large enough that the asymptotic result in Eq. (18) is appropriate. Then, from Table 1 the approximate width of the confidence interval for $\sigma^2_{\text{reproducibility}}$ is nearly eight times the value of $\hat{\sigma}^2_{\text{reproducibility}}$.

This leads us to conclude that gauge capability studies should use a sufficient number of operators whenever estimates of $\sigma^2_{\text{reproducibility}}$ are important. Often there is no preliminary information on the variance components, so that a preliminary estimate of ν is not available. In this case, the number of parts can be chosen large enough (say $p>40$) such that ν is conservatively approximated by $o-1$. Then, o can be selected to provide the desired confidence width according to Table 1. This method can provide a useful estimate of $\sigma^2_{\text{reproducibility}}$; however, the approach might result in substantially larger experiments than have traditionally been used in gauge capability studies. When a sufficiently large number of operators are not available, the analyst must realize that relatively imprecise estimates of $\sigma^2_{\text{reproducibility}}$ will result.

Other Parameter Estimation Methods

In addition to the analysis-of-variance method for estimating variance components, there are two other useful approaches that can be adapted to a gauge R&R experiment. They are the maximum likelihood procedure and the MINQUE procedure. The maximum likelihood estimators maximize the likelihood function of the sample over a constrained parameter space where each variance component is required to be non-negative. For more details of the maximum likelihood procedure, see the work of Hemmerle and Hartley (7) and Searle (4). The MINQUE method produces estimates that are best quadratic unbiased. For details of the procedure, see Refs. 6, 8, and 9. Like maximum likelihood, the MINQUE procedure guarantees non-negative estimates of the variance components. Both procedures are iterative schemes and require computer software for their implementation.

SAS PROC VARCOMP provides an excellent way to estimate variance components by these methods.* Table 2 shows the results of applying PROC VARCOMP to the data from our gauge capability study (Table 2 of Part I of this paper). Panel A of Table 2 gives the maximum likelihood estimates and Panel B shows the MINQUE estimates. Note that the point estimates from both methods do not differ dramatically from the estimates obtained previously from our modified analysis-of-variance method (refer to Table 4, Part I of this paper). In addition to the point estimates of the variance components, the asymptotic (large-sample) covariance matrix of the estimates is also given.

It is very easy to construct confidence intervals on these variance components. If $\hat{\sigma}^2$ is an estimator of the variance component and if $V(\hat{\sigma}^2)$ is the variance of $\hat{\sigma}^2$, which is obtained as the corresponding diagonal element of the covariance matrix, then an approximate $100(1-\alpha)\%$ confidence interval on σ^2 is

*Version 6.0 of SAS PROC VARCOMP actually provides four different methods of variance component estimation. Method 1 is the moment or analysis of variance estimator. Method 2 is called the MINQUE method. It is a noniterative procedure that may produce negative estimates of variance components, and for this reason we do not recommend it. Method 3 is the maximum likelihood method, and Method 4, called the restricted maximum likelihood method, is actually an iterated MINQUE procedure. We recommend one of these last two procedures.

Table 2. Maximum Likelihood and MINQUE Estimates of Variance Components for the Gauge Capability Study

POINT ESTIMATES	ASYMPTOTIC COVARIANCE MATRIX OF ESTIMATORS				
	A. Maximum Likelihood Estimates				
$\hat{\sigma}_R^2 = 0.88$		σ_O^2	σ_P^2	σ_{OP}^2	σ_R^2
$\hat{\sigma}_{OP}^2 = 0$	σ_O^2	0.0010	0.0004	0	−0.0004
$\hat{\sigma}_P^2 = 9.73$	σ_P^2	0.0004	9.7658	0	−0.0028
$\hat{\sigma}_O^2 = 0.01$	σ_{OP}^2	0	0	0	0
	σ_R^2	−0.0004	−0.0028	0	0.0159
	B. MINQUE Estimates				
$\hat{\sigma}_R^2 = 0.88$		σ_O^2	σ_P^2	σ_{OP}^2	σ_R^2
$\hat{\sigma}_{OP}^2 = 0$	σ_O^2	0.0011	0.0001	0	−0.0004
$\hat{\sigma}_P^2 = 10.25$	σ_P^2	0.0001	10.3822	0	−0.0027
$\hat{\sigma}_O^2 = 0.011$	σ_{OP}^2	0	0	0	0
	σ_R^2	−0.0004	−0.0027	0	0.0159

$$\hat{\sigma}^2 - z_{\alpha/2}\sqrt{V(\hat{\sigma}^2)} \leqslant \sigma^2 \leqslant \hat{\sigma}^2 + z_{\alpha/2}\sqrt{V(\hat{\sigma}^2)} \tag{19}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution. Because this interval is based on asymptotic results, large sample sizes are required. To illustrate the calculations, because $\hat{\sigma}^2_{\text{repeatability}} = \hat{\sigma}_R^2 = 0.88$ and $V(\hat{\sigma}_R^2) = 0.0159$ from the maximum likelihood procedure, the approximate 95% confidence interval on $\sigma^2_{\text{repeatability}}$ is

$$0.88 - 1.96\sqrt{0.0159} \leqslant \sigma^2_{\text{repeatability}} \leqslant 0.88 + 1.96\sqrt{0.0159}$$

or

$$0.63 \leqslant \sigma^2_{\text{repeatability}} \leqslant 1.13$$

Note that this confidence interval is about the same length as the confidence interval on $\sigma^2_{\text{repeatability}}$ from the same data using Eq. (6) based on the analysis-of-variance method. Maximum likelihood and MINQUE estimates will often produce shorter confidence intervals than the moment estimator does, and so this is another reason to recommend these methods.

Because $\hat{\sigma}_{OP}^2 = 0$, we would estimate $\hat{\sigma}^2_{\text{reproducibility}} = \hat{\sigma}_O^2 = 0.01$, by the maximum likelihood method. Panel A of Table 2 indicates that $V(\hat{\sigma}_O^2) = 0.0010$. The approximate 95% confidence interval on $\sigma^2_{\text{reproducibility}}$ is

$$0.01 - 1.96\sqrt{0.0010} \leqslant \sigma^2_{\text{reproducibility}} \leqslant 0.01 + 1.96\sqrt{0.0010}$$

or

$$0 \leqslant \sigma^2_{\text{reproducibility}} \leqslant 0.07$$

This interval is considerably shorter that the confidence interval on $\sigma^2_{\text{reproducibility}}$ computed from Eq. (12).

Using the asymptotic distribution of the maximum likelihood estimates, a confidence interval can be obtained for σ^2_{gauge}. The point estimate of σ^2_{gauge} is

$$\hat{\sigma}^2_{\text{gauge}} = \hat{\sigma}^2_R + \hat{\sigma}^2_{OP} + \hat{\sigma}^2_O = [\hat{\sigma}^2_R, \hat{\sigma}^2_{OP}, \hat{\sigma}^2_O] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

For the example in Panel A of Table 2, $\hat{\sigma}^2_{\text{gauge}} = 0.89$. An asymptotic estimate of the variance of $\hat{\sigma}^2_{\text{gauge}}$ is

$$V(\hat{\sigma}^2_{\text{gauge}}) = [111]\, \Sigma \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

where Σ is the asymptotic covariance matrix of the three maximum likelihood estimators of the variance components used in the point estimate. For this example,

$$V(\hat{\sigma}^2_{\text{gauge}}) = 0.016$$

Therefore, an asymptotic 95% confidence interval for σ^2_{gauge} using the maximum likelihood estimators is $0.89 \pm 1.96\sqrt{0.016}$, or

$$0.64 \leqslant \sigma^2_{\text{gauge}} \leqslant 1.14$$

This interval is very similar to the result obtained from the reduced model using Eqs. (15) and (16).

The maximum likelihood procedure illustrated by the results in Table 2, Panel A, provides non-negative estimates of the variance components without preliminary evaluation of tentative models. However, a software package capable of providing the estimates is required for routine application of this procedure. Based on our examples, an effective alternative is to use a reduced model to estimate the components whenever the analysis of a full model results in negative estimates. The reduced model procedure is simple, very similar to the traditional analysis-of-variance method, provides non-negative estimates, and does not require specialized software. The basic results can be obtained from ANOVA tables produced by standard statistical software.

The Nested Design Model

We now consider a variation of the factorial design model studied previously. Suppose that the p parts measured n times by operator 1 are physically different than the p parts measured by operators $2, 3, \ldots, o$. This situation occasionally occurs in practice; for example, when only a limited number of measurements can be made on each part, or when the operators are in different locations. This situation leads to a two-stage nested design model, with parts nested within operators, say

$$X_{ijk} = \mu + O_i + P(O)_{j(i)} + R_{k(ij)}, \quad \begin{aligned} i &= 1, 2, \ldots, o, \\ j &= 1, 2, \ldots, p, \\ k &= 1, 2, \ldots, n \end{aligned} \tag{20}$$

The parameters O_i, $P(O)_{j(i)}$, and $R_{k(ij)}$ are typically random effects, each normally and independently distributed with mean zero and variances σ_O^2, $\sigma_{P(O)}^2$, and σ_R^2, respectively. As is shown by Montgomery (10), the expected mean squares from the two-stage nested analysis of variance are

$$\begin{aligned} E(\mathrm{MS}_R) &= \sigma_R^2 \\ E(\mathrm{MS}_{P(O)}) &= \sigma_R^2 + n\sigma_{P(O)}^2 \\ E(\mathrm{MS}_O) &= \sigma_R^2 + n\sigma_{P(O)}^2 + pn\sigma_O^2 \end{aligned} \tag{21}$$

and the estimates of the variance components are

$$\begin{aligned} \hat{\sigma}_R^2 &= \mathrm{MS}_R \\ \hat{\sigma}_{P(O)}^2 &= \frac{\mathrm{MS}_{P(O)} - \mathrm{MS}_R}{n} \\ \hat{\sigma}_O^2 &= \frac{\mathrm{MS}_O - \mathrm{MS}_{P(O)}}{pn} \end{aligned} \tag{22}$$

Because there is no interaction term in this model, the point estimates of $\sigma^2_{\text{repeatability}}$, $\sigma^2_{\text{reproducibility}}$, and σ^2_{gauge} are

$$\hat{\sigma}^2_{\text{repeatability}} = \hat{\sigma}_R^2 \tag{23}$$

$$\hat{\sigma}^2_{\text{reproducibility}} = \hat{\sigma}_O^2 \tag{24}$$

and

$$\begin{aligned} \hat{\sigma}^2_{\text{gauge}} &= \hat{\sigma}^2_{\text{repeatability}} + \hat{\sigma}^2_{\text{reproducibility}} \\ &= \frac{\mathrm{MS}_O - \mathrm{MS}_{P(O)} + pn\mathrm{MS}_R}{pn} \end{aligned} \tag{25}$$

To illustrate construction of these point estimates, consider the data from the gauge R&R study of Montgomery (1). We have previously observed that the correct experimental design model for this example is the two-factor factorial. However, suppose that each of the three operators made their measurements on *different* samples of $p=20$ parts; consequently, the two-stage nested model now applies. Table 3 shows the analysis of variance for this case. The estimates from Eqs. (23)–(25) are

$$\hat{\sigma}^2_{\text{repeatability}} = \text{MS}_R = 0.99$$

$$\hat{\sigma}^2_{\text{reproducibility}} = \frac{\text{MS}_O - \text{MS}_{P(O)}}{pn} = \frac{1.31 - 21.27}{20(2)} = -0.50$$

and

$$\hat{\sigma}^2_{\text{gauge}} = \hat{\sigma}^2_{\text{repeatability}} + \hat{\sigma}^2_{\text{reproducibility}} = 0.99 - 0.50 = 0.49$$

The analysis-of-variance method estimators of $\sigma^2_{\text{repeatability}}$ and $\sigma^2_{\text{reproducibility}}$ are not very satisfactory because of the relatively large negative value of $\hat{\sigma}^2_{\text{reproducibility}}$. Unfortunately, this will occur frequently in the nested model when variability among operators is relatively small and variability among parts is relatively large. One solution to this problem is to fit a reduced model and employ the modified analysis-of-variance estimation procedure discussed earlier. If the factor Operators is deleted, the model becomes a single-factor experiment in the factor Parts, which has $p = 20$ levels, and each part is measured $n = 6$ times. Table 4 shows the analysis of variance. We would now estimate

Table 3. Analysis of Variance for the Gauge Capability Study in Ref. 3, Treated As a Two-Stage Nested Design

SOURCE OF VARIABILITY	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F_0	P-value
Operators	2.62	2	1.31	0.06	0.9418
Parts (within operators)	1212.48	57	21.27	21.48	0.0001
Repeatability	59.50	60	0.99		
Total	1274.60	119			

Table 4. Analysis of Variance for the Gauge Capability Study in Ref. 3, Reduced Model

SOURCE OF VARIABILITY	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F_0	P-value
Parts	1185.43	19	62.39	70.10	0.0001
Repeatability	89.17	100	0.89		
Total	1274.60	119			

$$\hat{\sigma}^2_{\text{repeatability}} = \hat{\sigma}^2_R = 0.89$$

and

$$\hat{\sigma}^2_{\text{reproducibility}} = 0$$

Therefore,

$$\hat{\sigma}^2_{\text{gauge}} = \hat{\sigma}^2_{\text{repeatability}} + \hat{\sigma}^2_{\text{reproducibility}}$$
$$= 0.89 + 0 = 0.89$$

This is very consistent with the estimate of $\hat{\sigma}^2_{\text{gauge}}$ computed with the reduced factorial model.

An exact $100(1-\alpha)\%$ confidence interval on $\sigma^2_{\text{repeatability}}$ can be obtained by the same method [Eq. (6)] used in the factorial experiment. An approximate $100(1-\alpha)\%$ confidence interval on $\sigma^2_{\text{reproducibility}}$ is

$$\frac{\nu\hat{\sigma}^2_O}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2_{\text{reproducibility}} \leqslant \frac{\nu\hat{\sigma}^2_O}{\chi^2_{1-\alpha/2,\nu}} \tag{26}$$

where

$$\nu = (\hat{\sigma}^2_O)^2\left[\frac{(1/pn)^2\,\text{MS}^2_O}{o-1} + \frac{(1/pn)^2\,\text{MS}^2_{P(O)}}{o(p-1)}\right]^{-1} \tag{27}$$

This confidence interval is only appropriate when $\hat{\sigma}^2_{\text{reproducibility}}$ is positive. An approximate $100(1-\alpha)\%$ confidence interval on σ^2_{gauge} is

$$\frac{\nu\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2_{\text{gauge}} \leqslant \frac{\nu\hat{\sigma}^2_{\text{gauge}}}{\chi^2_{1-\alpha/2,\nu}} \tag{28}$$

where

$$\nu = (\hat{\sigma}^2_{\text{gauge}})^2\left[\frac{(1/pn)^2\,\text{MS}^2_O}{o-1} + \frac{(1/pn)^2\,\text{MS}^2_{P(O)}}{o(p-1)} + \frac{\text{MS}^2_R}{op(n-1)}\right]^{-1} \tag{29}$$

An alternative to fitting a reduced model is to estimate the variance components in the full model using either the maximum likelihood or MINQUE estimation method. These estimates, computed from SAS PROC VARCOMP, are shown in Table 5. Panel A of Table 5 shows the maximum likelihood estimates and Panel B of Table 5 shows the MINQUE estimates. Note that both of these procedures give $\hat{\sigma}_O^2 = 0$; therefore, the estimate of $\sigma^2_{\text{reproducibility}}$ is zero. The estimate of $\sigma^2_{\text{repeatability}}$ is $\hat{\sigma}_R^2 = 0.99$ by both methods. Once again, these estimates are consistent with those produced by the modified analysis-of-variance method. Confidence interval on these maximum likelihood and MINQUE estimates can be obtained by the method illustrated previously.

Other Experimental Design Models

Often gauge capability analysis requires the consideration of more than operators, parts, and replications to fully understand the measurement process. For example, more than one measurement tool might be used on the manufacturing floor and there is interest in estimating the effect of different tools as a separate source of variability. We can design an experiment with Operators, Tools, and Parts being arranged in a factorial design. The Tools can be either random or fixed effects. If we assume they are random effects, a model that can be used is

$$X_{ijkl} = \mu + O_i + P_j + (OP)_{ij} + T_k + (OT)_{ik} + (PT)_{jk} + (OPT)_{ijk} + R_{l(ijk)} \qquad (30)$$

where X_{ijkl} is the lth measurement made with tool k by operator j on part i.

Table 5. Maximum Likelihood and MINQUE Estimates of Variance Components for the Nested Design Model for a Gauge Capability Study from Ref. 3

POINT ESTIMATES	ASYMPTOTIC COVARIANCE MATRIX OF ESTIMATORS			
	A. Maximum Likelihood Estimates			
$\hat{\sigma}_O^2 = 0$	σ_O^2	σ_O^2	$\sigma_{P(O)}^2$	σ_R^2
$\hat{\sigma}_{P(O)}^2 = 9.63$	$\sigma_{P(O)}^2$	0	0	0
$\hat{\sigma}_R^2 = 0.99$	σ_R^2	0	3.4260	−0.0164
		0	−0.0164	0.0328
	B. MINQUE Estimates			
$\hat{\sigma}_O^2 = 0$	σ_O^2	σ_O^2	$\sigma_{P(O)}^2$	σ_R^2
$\hat{\sigma}_{P(O)}^2 = 9.80$	$\sigma_{P(O)}^2$	0	0	0
$\hat{\sigma}_R^2 = 0.99$	σ_R^2	0	3.6027	−0.0027
		0	−0.0027	0.0328

It is important to investigate the interaction terms between Operators and Tools, Operators and Parts, and Parts and Tools in Eq. (30). A significant interaction invalidates the simple addition of variance components for Operator, Tool, and Replication to produce an estimate of total measurement variability. If a significant interaction is present, an inspection of the raw data might be useful. For example, if there is a significant interaction between Operators and Tools, we might identify a particular operator and tool combination which led to the interaction. If there are different models of tools used in the study, or tools from different manufacturers, a conclusion might be that additional training is required for some operators on specific tools. Alternatively, the conclusion might be that unstable characteristics of all the tools have been compensated for by a few dedicated operators. In any event, the identification of significant interactions is an important component of a gauge capability study.

Other experimental design models can be used, depending on how the experiment is actually conducted. For example, suppose that p parts are measured by each of o operators and each operator measures each part n times in random order. However, each set of operator tests are performed on n different days. A model for such a study is

$$X_{ijkl} = \mu + O_i + P_j + (OP)_{ij} + T(O)_{k(i)} + R_{l(ijk)} \tag{31}$$

where $T(O)_{k(i)}$ is a nested effect that represents the test effect within operators. The randomization restriction arising from running the n tests for each operator on different days accounts for this term in the model. Typically, all the terms in this model are treated as random effects, and the variance components can be estimated using the methods discussed in this article.

Finally, we emphasize that it is extremely important in gauge capability studies to use an experimental design model that is consistent with how the experiment is actually conducted. In our experience, using the wrong model is relatively common. This will usually produce incorrect results that can be very misleading.

Conclusion

This article has presented experimental design models for the gauge repeatability and reproducibility study and has suggested variance component estimates that are alternatives to the "classical" estimates of $\sigma^2_{\text{repeatability}}$, $\sigma^2_{\text{reproducibility}}$, and σ^2_{gauge} for both factorial and nested design models applied to gauge capability studies, assuming that either analysis-of-variance or modified analysis-of-variance estimates of the variance components are employed. Alternative estimation procedures that are presented are the maximum likelihood and MINQUE methods. We demonstrate how to obtain both point estimates and confidence intervals using these methods. We believe that the modified analysis-of-variance procedure leads to very reasonable estimates when the standard analysis-of-variance methods provide negative estimates of variance components. Alternatively, the maximum likelihood or MINQUE methods could be used in these cases.

It is very important to consider the selection of an appropriate experimental design when conducting a gauge capability study. These considerations include identification of all factors that may influence gauge performance, the proper treatment of these factors as either factorial or nested variables, and the selection of appropriate sample sizes. Our work indicates that it is frequently more difficult to obtain precise estimates of $\sigma^2_{\text{reproducibility}}$ than $\sigma^2_{\text{repeatability}}$, usually because a relatively small number of operators are used in gauge studies.

References

1. Montgomery, D. C. and Runger, G. C., Gauge Capability Analysis and Designed Experiments. Part I: Basic Methods, *Qual. Eng.*, 6, 115–135 (1993).
2. Tsai, P., Variable Gauge Repeatability and Reproducibility Study Using the Analysis of Variance Method, *Qual. Eng.*, *1*, 107–115 (1988).
3. Montgomery, D. C., *Introduction to Statistical Quality Control*, 2nd ed., John Wiley and Sons, New York, 1991.
4. Searle, S. R., *Linear Models*, John Wiley and Sons, New York, 1971.
5. Graybill, F.A., *Introduction to Linear Statistical Models*, Vol. 1, McGraw-Hill, New York, 1961.
6. Milliken, G. A. and Johnson, D. E., *Analysis of Messy Data, Volume I: Designed Experiments*, Van Nostrand Reinhold, New York, 1984.
7. Hemmerle, W. J. and Hartley, H. O., Computing Maximum Likelihood Estimates for the Mixed A.O.V. Model Using the W Transformation, *Technometrics*, 15, 819–831 (1973).
8. Rao, C. R., Minimum Variance Quadratic Unbiased Estimation of Variance Components, *J. Multivar. Anal.*, 1, 445–456 (1971).
9. Rao, C. R., Estimation of Variance and Covariance Components in Linear Models, *J. Amer. Statist. Assoc.*, 67, 112–115 (1972).
10. Montgomery, D. C., *Design and Analysis of Experiments*, 3rd ed., John Wiley and Sons, New York, 1991.

About the Authors: Douglas C. Montgomery is Professor of Industrial & Management Systems Engineering at Arizona State University. He is the author of seven books, including widely-used texts on statistical quality control and experimental design, and over 100 papers and articles. He is a Fellow of the American Society for Quality Control, a Fellow of the American Statistical Association, and a Senior Member of the Institute of Industrial Engineers, and has held several national offices in these organizations, including Director of the Statistics Division of ASQC.

George C. Runger is Assistant Professor in the Department of Decision Sciences and Engineering Systems at Rensselaer Polytechnic Institute. Formerly, Dr. Runger was a senior engineer at IBM where he implemented statistical process control in chemical and discrete manufacturing operations. His research interests are in quality and productivity improvement, integrating manufacturing data for management decisions, and experimental design strategies.