Where do gauge R&R formulas come from?

W. Robert Stephenson[1]
Department of Statistics
Iowa State University

1 Model:

The model used to derive the formulas for gauge R&R is a so called "two factor random effects model." In this model the kth measurement made by operator j on part i (denoted y_{ijk}) is described in terms of sum of several parts. Specifically,

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

Where μ is an unknown constant representing the average (over all possible operators and parts) measurement. The α's are random effects of different parts. The β's are random effects of different operators. The $\alpha\beta$'s are random joint effects of particular part/operator combinations. Finally, the ϵ's are random measurement errors.

Associated with each random effect is a so called "variance component" which quantifies the amount of variability attributable to that effect. These variance components are denoted σ^2_α, σ^2_β, $\sigma^2_{\alpha\beta}$, and σ^2. They quantify variation in parts, operators, part/operator combinations and measurement error, respectively.

The purpose of gauge R&R is to estimate these variance components, specifically σ^2 and $\sigma^2_\beta + \sigma^2_{\alpha\beta}$. The former, σ^2, quantifies the variation attributable to measurement errors for repeated measurements on a fixed part/operator combination. Estimating σ gives repeatability, $\hat{\sigma}_{Repeat}$. The latter, $\sigma^2_\beta + \sigma^2_{\alpha\beta}$, quantifies the variation experienced with many operators making a single measurement on the same part assuming that there is no error in repeated measurements. Estimating $\sqrt{\sigma^2_\beta + \sigma^2_{\alpha\beta}}$ gives reproducibility, $\hat{\sigma}_{Reprod}$.

2 Estimation:

2.1 Repeatability

For any particular part(i)/operator(j) combination, repeated measurements are subject only to the error variability, σ^2. Therefore, σ (and equivalently σ_{Repeat}) can be estimated from the repeated measurements on each ij combination. Specifically, the range of the repeated measurements on the ith part made by the jth operator, R_{ij} can be used. If measurements are normally distributed, then it is known that the center of the sampling distribution of R_{ij} is:

$$d_2(n_M)\sigma$$

[1]With help from Stephen B. Vardeman

This suggests that the estimate of σ should be:

$$R_{ij}/d_2(n_M)$$

Averaging these over all $n_P n_O$ part/operator combinations gives:

$$\hat{\sigma}_{Repeat} = \bar{R}/d_2(n_M)$$

$$\hat{\sigma}^2_{Repeat} = \left[\bar{R}/d_2(n_M)\right]^2$$

2.2 Reproducibility

Consider the mean, $\bar{y}_{ij}$, of the repeated measurements made on the ith part by the jth operator. In terms of the model:

$$\bar{y}_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \bar{\epsilon}_{ij}$$

Now, $\bar{\epsilon}_{ij}$ is an average of n_M measurements. This average has variance, σ^2/n_M. For a fixed part (i) the means for the operators differ by random quantities:

$$\beta_j + \alpha\beta_{ij} + \bar{\epsilon}_{ij}$$

which have variance:

$$\sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2/n_M$$

The range of the means for part i, R_{Mi}, has center at:

$$d_2(n_O)\sqrt{\sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2/n_M}$$

This suggests that $R_{Mi}/d_2(n_O)$, or better yet $\bar{R}_M/d_2(n_O)$, is an estimate of

$$\sqrt{\sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2/n_M}$$

Putting this together with the previous information

$$\left[\bar{R}_M/d_2(n_O)\right]^2 \text{ estimates } \sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2/n_M$$

$$\tfrac{1}{n_M}\left[\bar{R}/d_2(n_M)\right]^2 \text{ estimates } \sigma^2/n_M$$

$$\left[\bar{R}_M/d_2(n_O)\right]^2 - \tfrac{1}{n_M}\left[\bar{R}/d_2(n_M)\right]^2 \text{ estimates } \sigma^2_\beta + \sigma^2_{\alpha\beta}$$

That is

$$\hat{\sigma}^2_{Reprod} = \left(\left[\bar{R}_M/d_2(n_O)\right]^2 - \frac{1}{n_M}\left[\bar{R}/d_2(n_M)\right]^2\right)$$