

IE 361 Module 5

Gauge R&R Studies Part 1: Motivation, Data, Model and Range-Based Estimates

Reading: Section 2.2 *Statistical Quality Assurance for Engineers*
(Section 2.4 of Revised *SQAME*)

Prof. Steve Vardeman and Prof. Max Morris

Iowa State University

August 2008

Standard R&R Data and Descriptive Statistics (Based on Ranges) for Partitioning Measurement Variation

A very common type of industrial measurement study is one where a single gauge or piece of measuring equipment is used (according to a standard protocol) by multiple operators to measure multiple parts, with the primary end goal of quantifying repeatability and reproducibility measurement variation, and comparing measurement imprecision to the basic engineering requirements that a part must satisfy in order to be functional. Remember from what we have already said in Modules 3 and 4, that

- "Repeatability" variation is variation characteristic of one operator/analyst remeasuring one specimen
- "Reproducibility" variation is variation in operator biases, i.e. variation characteristic of many operators measuring a single specimen after accounting for (or somehow mathematically eliminating) repeatability variation

R&R Data and Descriptive Statistics

In a typical (balanced data) industrial Gauge R&R study, each of I items is measured m times by each of J operators. For example, a typical data layout for $I = 2$ parts, $J = 3$ operators, and $m = 2$ repeats per "cell" might be represented as in this figure.

		Operator		
		1	2	3
Part	1	y_{111}	y_{121}	y_{131}
		y_{112}	y_{122}	y_{132}
2	2	y_{211}	y_{221}	y_{231}
		y_{212}	y_{222}	y_{232}

Figure: Hypothetical Gauge R&R Data

If only *one* part/measurand were involved, the one-way model and analyses of Module 4 could be used to do inference for what we have been calling σ_{δ} and σ_{device} . (See again Example 4-3 of Module 4.) But (presumably in order to have some check on how measurement performs across a spectrum of parts) it is common in Gauge R&R studies to use multiple parts. The next figure shows some real R&R data collected in-class in IE 361 on $I = 4$ parts, by $J = 3$ operators, making $m = 2$ repeats per cell, and some summary statistics based on ranges (rather than standard deviations). (The data are measurements of the sizes of some Styrofoam peanuts, and range-based methods are presented here because of their connection to fairly standard industry practice.)

R&R Data and Descriptive Statistics

	Operator #1		Operator #2		Operator #3		Δ_i
Part # 1	.52	$\bar{y}_{11} = .52$.54	$\bar{y}_{12} = .535$.55	$\bar{y}_{13} = .55$	$\Delta_1 = .03$
	.52	$R_{11} = 0$.53	$R_{12} = .01$.55	$R_{13} = 0$	
Part # 2	.56	$\bar{y}_{21} = .555$.54	$\bar{y}_{22} = .54$.55	$\bar{y}_{23} = .555$	$\Delta_2 = .015$
	.55	$R_{21} = .01$.54	$R_{22} = 0$.56	$R_{23} = .01$	
Part # 3	.57	$\bar{y}_{31} = .565$.55	$\bar{y}_{32} = .555$.57	$\bar{y}_{33} = .57$	$\Delta_3 = .015$
	.56	$R_{31} = .01$.56	$R_{32} = .01$.57	$R_{33} = 0$	
Part # 4	.55	$\bar{y}_{41} = .55$.54	$\bar{y}_{42} = .545$.56	$\bar{y}_{43} = .555$	$\Delta_4 = .01$
	.55	$R_{41} = 0$.55	$R_{42} = .01$.55	$R_{43} = .01$	

$$\bar{R} = \frac{.07}{12} = .00583 \quad \text{and} \quad \bar{\Delta} = \frac{.07}{4} = .0175$$

Figure: Styrofoam Peanut Size Measurements and Summary Statistics (Inches)

\bar{R} is a very simple descriptive measure of within cell variability and is related to repeatability variation. Similarly, $\bar{\Delta}$ is a measure of between-operator variation and is related to reproducibility variation. In order to be more more precise about these (and to do statistical inference) one must adopt a probability model for the data collected in an R&R study.

The "Two-Way Random Effects" Model for Gauge R&R Data

Typical analyses of Gauge R&R studies are based on the so-called "two-way random effects" model. With

y_{ijk} = the k th measurement made by operator j on specimen i

this model is that y_{ijk} is made up as a sum of independent contributions,

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

where

- μ is an (unknown) constant, an average (over all possible operators and all possible parts/specimens) measurement
- the α_i are normal with mean 0 and variance σ_{α}^2 , (random) effects of different parts/specimens
- the β_j are normal with mean 0 and variance σ_{β}^2 , (random) effects of different operators

The Two-Way Random Effects Model

- the $\alpha\beta_{ij}$ are normal with mean 0 and variance $\sigma_{\alpha\beta}^2$, (random) joint effects peculiar to particular part/operator combinations
- the ϵ_{ijk} are normal with mean 0 and variance σ^2 , (random) errors that are peculiar to a particular attempt to make a measurement (they change measurement-to-measurement, even if the part and operator remain the same)

σ_{α}^2 , σ_{β}^2 , $\sigma_{\alpha\beta}^2$, and σ^2 are called "variance components" and their sizes govern how much variability is seen in the measurements y_{ijk} .

The Two-Way Random Effects Model

Example 5-1

The reader should conduct a "Thought Experiment" generating a Gauge R&R data set, and fill in formulas for the 12 measurements in the table below. (For example, $y_{111} = \mu + \alpha_1 + \beta_1 + \alpha\beta_{11} + \epsilon_{111}$.)

		Operator		
		1	2	3
Part	1	$y_{111} =$	$y_{121} =$	$y_{131} =$
	$y_{112} =$	$y_{122} =$	$y_{132} =$	
2	$y_{211} =$	$y_{221} =$	$y_{231} =$	
	$y_{212} =$	$y_{222} =$	$y_{232} =$	

The Two-Way Random Effects Model

In this (two-way random effects) model

- σ measures within-cell/repeatability variation
- $\sigma_{\text{reproducibility}} = \sqrt{\sigma_{\beta}^2 + \sigma_{\alpha\beta}^2}$ is the standard deviation that would be experienced by many operators measuring the same specimen once each, *in the absence of repeatability variation*
- $\sigma_{\text{R\&R}} = \sqrt{\sigma_{\text{reproducibility}}^2 + \sigma^2} = \sqrt{\sigma_{\beta}^2 + \sigma_{\alpha\beta}^2 + \sigma^2}$ is the standard deviation that would be experienced by many operators measuring the same specimen once each (this is called σ_{overall} in *SQAME*)

To make connections to what we have done earlier, consider what these two-way model parameters mean if we restrict attention to part #1. The two-way random effects model says that measurements on part #1 can be thought of as

$$y_{1jk} = \mu + \alpha_1 + \beta_j + \alpha\beta_{1j} + \epsilon_{1jk}$$

The Two-Way Random Effects Model

What then varies operator-to-operator is

$$\beta_j + \alpha\beta_{1j}$$

This quantity thus plays the role of what we before called δ_j (operator bias for operator j ... for part #1) and

$$\sigma_{\beta_j + \alpha\beta_{1j}}^2 = \sigma_{\beta}^2 + \sigma_{\alpha\beta}^2$$

plays the role of what we before called σ_{δ}^2 (the reproducibility variance). The fact that $\beta_j + \alpha\beta_{1j}$ is specific to part #1 (for example changes to $\beta_j + \alpha\beta_{2j}$ if part #2 is considered instead) has the interesting interpretation that the terms $\alpha\beta_{ij}$ play the role of "device" *nonlinearities!* That is, in the two-way random effects model where multiple parts are considered, a large variance component $\sigma_{\alpha\beta}^2$ is indicative of substantially non-constant bias on the part of the various operators in their use of the gauge. A most unpleasant circumstance indeed.

The Two-Way Random Effects Model

The most common analyses (both those based on ranges and those based on ANOVA) (e.g. following the AIAG manual) are *wrong*, in that they purport to produce estimates of $\sigma_{\text{reproducibility}}$ and $\sigma_{\text{R\&R}}$ but fail to do so. *SQAME* presents correct range-based and ANOVA-based methods. We will use primarily the generally more effective ANOVA-based estimates and confidence intervals that can be based on them (these limits are not found in *SQAME*).

If you end up doing a gauge R&R study for your project client, you will almost certainly be asked to use company standard formulas or a company spreadsheet that implements the (WRONG) AIAG formulas. If you do this you should compare those results to ones obtained using correct formulas from these modules! Failure to do so will be frowned upon when projects are graded.

Simple Range-Based Point Estimates of Repeatability and Reproducibility Standard Deviations

For motivation sake (and because of the connection to standard formulas), first briefly consider range-based estimates. Possible estimates are:

- $\hat{\sigma} = \frac{\bar{R}}{d_2(m)}$ for \bar{R} the average within-cell range and $d_2(m)$ a "control chart constant" based on "sample size" m
- $\hat{\sigma}_{\text{reproducibility}} = \sqrt{\max\left(0, \left(\frac{\bar{\Delta}}{d_2(J)}\right)^2 - \frac{1}{m} (\hat{\sigma})^2\right)}$ for $\bar{\Delta}$ the average of part ranges of cell means and $d_2(m)$ a "control chart constant" based on "sample size" J

(The second of these is NOT the AIAG estimate of reproducibility standard deviation.)

Range-Based Point Estimates

Example 5-2

Below are some simple calculations based on measurements of a geometric dimension of a machined part in a study with $I = 3$, $J = 3$, and $m = 2$.

	Operator 1	Operator 2	Operator 3	
Part 1	$\bar{y}_{11} = .34730$ $R_{11} = 0$	$\bar{y}_{12} = .34660$ $R_{12} = .0002$	$\bar{y}_{13} = .34715$ $R_{13} = .0001$	$\Delta_1 = .00070$
Part 2	$\bar{y}_{21} = .34710$ $R_{21} = 0$	$\bar{y}_{22} = .34645$ $R_{22} = .0001$	$\bar{y}_{23} = .34710$ $R_{23} = 0$	$\Delta_2 = .00065$
Part 3	$\bar{y}_{31} = .34720$ $R_{31} = 0$	$\bar{y}_{32} = .34655$ $R_{32} = .0003$	$\bar{y}_{33} = .34710$ $R_{33} = 0$	$\Delta_3 = .00065$

So $\bar{R} = .0007/9 = .000078$ and $\bar{\Delta} = .00067$ and

$$\hat{\sigma} = \frac{\bar{R}}{d_2(m)} = \frac{.000078}{1.128} = .000069 \text{ in}$$

Range-Based Point Estimates

Example 5-2

and

$$\begin{aligned}\hat{\sigma}_{\text{reproducibility}} &= \sqrt{\max\left(0, \left(\frac{\bar{\Delta}}{d_2(J)}\right)^2 - \frac{1}{m}(\hat{\sigma})^2\right)} \\ &= \sqrt{\left(\frac{.00067}{1.693}\right)^2 - \frac{1}{2}(.000069)^2} = .000391 \text{ in}\end{aligned}$$

A natural way to estimate $\sigma_{\text{R\&R}}$ is as

$$\hat{\sigma}_{\text{R\&R}} = \sqrt{(.000069)^2 + (.00039)^2} = .000396 \text{ in}$$

and the calculations here suggest that the bulk of measurement imprecision is traceable to differences between operators.