

Module 3

Using some intermediate-level statistical methods to quantify the importance of contributing sources of variability in measurements

Prof. Stephen B. Vardeman
Statistics and IMSE
Iowa State University

March 4, 2008

Separating Sources of Variation

Through reference to familiar elementary one- and two-sample methods of statistical inference, Modules 2 illustrated the basic insight that:

How sources of physical variation interact with a data collection plan governs what of practical importance can be learned from a data set, and in particular, how measurement error is reflected in the data set.

We begin to consider some computationally more complicated statistical methods and what they provide in terms of quantification of the impact of measurement variation on quality assurance data.

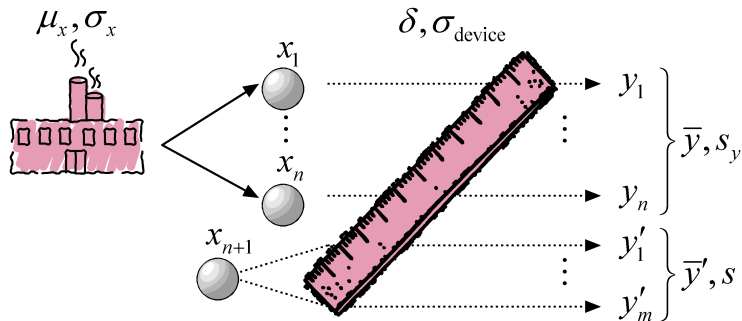
A Simple Method for Separating Process and Measurement Variation

In Module 1 we essentially observed that

- repeated measurement of a single measurand with a single device allows one to estimate device variability, and
- single measurements made on multiple measurands from a stable process using a single linear device allow one to estimate a combination of process and measurement variability,

and remarked that these facts suggest a way to estimate a process standard deviation alone (by removing variation of the first type from a measure of variation of the second type). Our first objective in this module is to elaborate on this.

A Simple Method for Separating Process and Measurement Variation (cont.)



$$y_i \text{'s} \sim \text{ind} \left(\mu_x + \delta, \sqrt{\sigma_x^2 + \sigma_{\text{device}}^2} \right) \text{ independent of}$$

$$y_i \text{'s} \sim \text{ind} \left(x_{n+1} + \delta, \sigma_{\text{device}} \right)$$

A Simple Method for Separating Process and Measurement Variation (cont.)

The sample standard deviation of the y 's, s_y , is a natural empirical approximation for $\sigma_y = \sqrt{\sigma_x^2 + \sigma_{\text{device}}^2}$ and the sample standard deviation of the y' 's, s , is a natural empirical approximation for σ_{device} . That suggests that one estimate the process standard deviation with

$$\hat{\sigma}_x = \sqrt{\max(0, s_y^2 - s^2)} .$$

It is further possible to use this estimate to make approximate confidence limits for σ_x .

A Simple Method for Separating Process and Measurement Variation (cont.)

Another Satterthwaite approximation gives approximate degrees of freedom

$$\hat{\nu} = \frac{\hat{\sigma}_x^4}{\frac{s_y^4}{n-1} + \frac{s^4}{m-1}}$$

for $\hat{\sigma}_x$ and corresponding approximate confidence limits for σ_x

$$\hat{\sigma}_x \sqrt{\frac{\hat{\nu}}{\chi_{\text{upper}}^2}} \quad \text{and} \quad \hat{\sigma}_x \sqrt{\frac{\hat{\nu}}{\chi_{\text{lower}}^2}} .$$

Example 3-1

In Example 2A-1, we considered $m = 5$ measurements made by a single analyst on a single physical specimen of material using a particular assay machine that produced $s = .0120$ mol/l. Subsequently, specimens from $n = 20$ different batches were analyzed and $s_y = .0300$ mol/l. An estimate of real process standard deviation uninflated by measurement variation is

$$\hat{\sigma}_x = \sqrt{\max\left(0, (.0300)^2 - (.0120)^2\right)} = .0275 \text{ mol/l}$$

and this value can be used to make confidence limits. Approximate degrees of freedom for $\hat{\sigma}_x$ are

$$\hat{\nu} = \frac{(.0275)^4}{\frac{(.0300)^4}{19} + \frac{(.0120)^4}{4}} = 11.96 .$$

Example 3-1 (cont.)

Rounding down to $\hat{v} = 11$, since the upper 2.5% point of the χ_{11}^2 distribution is 21.920 and the lower 2.5% point is 3.816, approximate 95% confidence limits for the real process standard deviation (σ_x) are

$$.0275\sqrt{\frac{11}{21.920}} \quad \text{and} \quad .0275\sqrt{\frac{11}{3.816}},$$

i.e.

$$.0195 \text{ mol/l} \quad \text{and} \quad .0467 \text{ mol/l}.$$

The One-Way Random Effects Model

One of the basic models of intermediate statistical methods is the so-called "one-way random effects model" for I samples of observations

$$\begin{aligned} & y_{11}, y_{12}, \dots, y_{1n_1} \\ & y_{21}, y_{22}, \dots, y_{2n_2} \\ & \vdots \\ & y_{I1}, y_{I2}, \dots, y_{In_I} \end{aligned}$$

This model says that the observations may be thought of as

$$y_{ij} = \mu_i + \epsilon_{ij}$$

where the ϵ_{ij} are independent normal random variables with mean 0 and standard deviation σ , while the I values μ_i are independent normal random variables with mean μ and standard deviation σ_μ (independent of the ϵ 's).

The One-Way Random Effects Model (cont.)

One can think of I means μ_i drawn at random from a normal distribution of μ_i 's, and subsequently observations y generated from I different normal populations with those means and a common standard deviation. In this model, the three parameters are σ (the "within group" standard deviation), σ_μ (the "between group" standard deviation), and μ (the overall mean). The squares of the standard deviations are called "variance components" since for any particular observation, the laws of expectation and variance imply that

$$\mu_y = \mu + 0 = \mu \quad \text{and} \quad \sigma_y^2 = \sigma_\mu^2 + \sigma^2$$

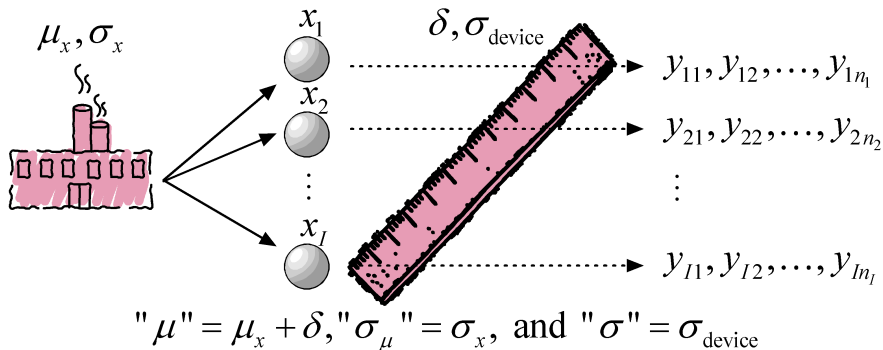
(i.e. σ_μ^2 and σ^2 are components of the variance of y).

Two QC Applications of the One-Way Random Effects Model

Two quality assurance contexts where this model can be helpful are where

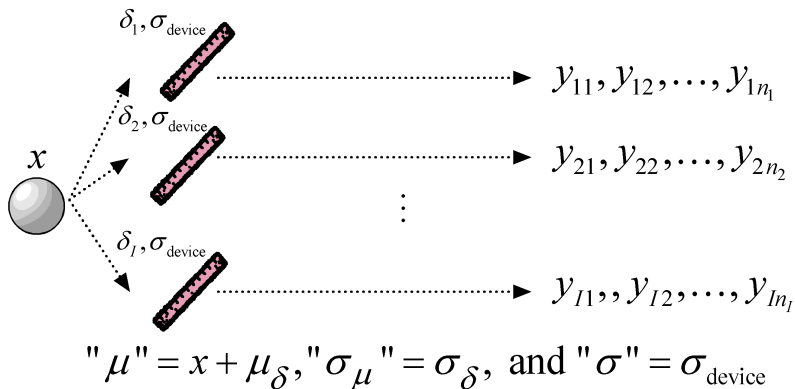
- multiple measurands from a stable process are each measured multiple times using the same device, and
- a single measurand is measured multiple times using multiple devices.

Multiple Measurands from a Stable Process Each Measured Multiple Times Using the Same Device



Here separating σ_x and σ_{device} could be central.

A Single Measurand Measured Multiple Times Using Multiple Devices



Here, if different "devices" are different operators, $\sigma_{\delta} = \sigma_{\text{reproducibility}}$ and $\sigma_{\text{device}} = \sigma_{\text{repeatability}}$ (and this is a 1 part R&R study).

Inference for the One-Way Model

There are at least two possible sources of inferences based on one-way data:

- some "by hand" formulas based on one-way ANOVA and more Satterthwaite approximations *could be used*, or (more conveniently)
- high quality statistical software prints out inferences for the parameters directly.

(If you are interested in the "hand formulas" e-mail me and I'll send you details from the revision of Vardeman and Jobe.)

Example 3-2

Part Hardness. Below are $m = 2$ hardness values (in mm) measured on each of $l = 9$ steel parts by a single operator at a farm implement manufacturer.

Part	1	2	3	4	5	6	7	8	9
	3.30	3.20	3.20	3.25	3.25	3.30	3.15	3.25	3.25
	3.30	3.25	3.30	3.30	3.30	3.30	3.20	3.20	3.30

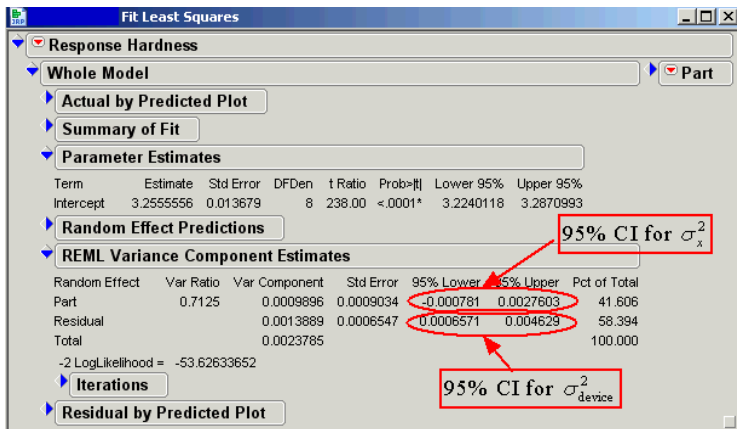
This is an instance of the first type of application of the one-way model. The TMJMP package produces limits for σ_x

$$0 \text{ mm and } \sqrt{.0027603} = .053 \text{ mm}$$

and limits for σ_{device}

$$\sqrt{.0006571} = .0256 \text{ mm and } \sqrt{.004629} = .0680 \text{ mm}$$

Example 3-2 (cont.)



What is clear from this analysis is that this is a case where part-to-part variation in hardness (measured by σ_x) is small enough and poorly determined enough in comparison to basic measurement noise (measured by σ_{device}) that it is impossible to really tell its size.

Example 3-3

Paper Weighing Below are $m = 3$ measurements of the weight (in g) of a single $20\text{ cm} \times 20\text{ cm}$ piece of 20 lb bond paper made by each of $l = 5$ different technicians using a single balance.

Operator	1	2	3	4	5
	3.481	3.448	3.485	3.475	3.472
	3.477	3.472	3.464	3.472	3.470
	3.470	3.470	3.477	3.473	3.474

This is an instance of the second type of application of the one-way model and illustrates the concepts of repeatability (fixed device) variation and reproducibility (here, device-to-device, i.e. operator-to-operator) variation. The TMJMP statistical package produces 95% confidence limits for the two standard deviations σ_{δ} ($= \sigma_{\mu}$ here) and σ_{device} ($= \sigma$ here).

Example 3-3 (cont.)

These intervals are

$$0 < \sigma_{\delta} < \sqrt{4.5 \times 10^{-5}} = .0067 \text{ g}$$

and

$$.0057 \text{ g} = \sqrt{3.2 \times 10^{-5}} < \sigma_{\text{device}} < \sqrt{.0002014} = .0142 \text{ g} .$$

Here repeatability variation is clearly larger than reproducibility (operator-to-operator) variation in weight measuring. If one doesn't like the overall size of measurement variation, it appears that some fundamental change in equipment or how it is used will be required. Simple training of the operators aimed at making how they use the equipment more uniform (and reduction of differences between their biases) has far less potential to improve measurement precision.

Example 3-3 (cont.)

