# Module 2B

## Elementary statistical methods and measurement: simple two-sample methods

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

March 4, 2008

# Basic CI Formulas for Two (Independent) Samples

These are based on a model that says that

$$y_{11}, y_{12}, \ldots, y_{1n_1} \quad \text{and} \quad y_{21}, y_{22}, \ldots, y_{2n_2}$$

are samples from normal distributions with means $\mu_1$ and $\mu_2$ and standard deviations $\sigma_1$ and $\sigma_2$. The "Satterthwaite approximation" gives limits

$$\bar{y}_1 - \bar{y}_2 \pm \hat{t}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \text{ for estimating } \mu_1 - \mu_2 \ ,$$

where "approximate degrees of freedom for $\hat{t}$ are

$$\hat{\nu} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{(n_1 - 1)n_1^2} + \frac{s_2^4}{(n_2 - 1)n_2^2}}$$

A simple conservative version of this method uses

$$\hat{\nu}^* = \min\left((n_1 - 1), (n_2 - 1)\right) \ .$$

# Basic CI Formulas for Two (Independent) Samples (cont.)

Further, in the two-sample context, there are elementary confidence limits

$$\frac{s_1}{s_2} \cdot \frac{1}{\sqrt{F_{(n_1-1),(n_2-1),\text{upper}}}} \quad \text{and} \quad \frac{s_1}{s_2} \cdot \frac{1}{\sqrt{F_{(n_1-1),(n_2-1),\text{lower}}}} \qquad \text{for} \quad \frac{\sigma_1}{\sigma_2}.$$

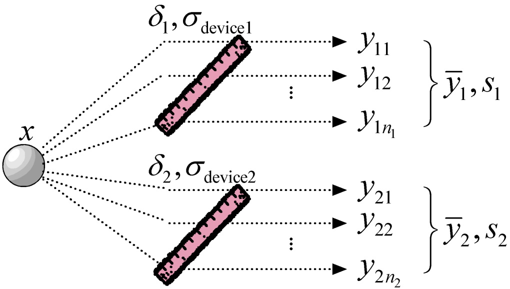
($F_{(n_1-1),(n_2-1),\text{lower}} = 1/F_{(n_2-1),(n_1-1),\text{upper}}$ so that standard $F$ tables giving only upper percentage points can be employed.)

# Four Applications

Four different ways that two "samples" of $n_1$ observed values $y_1$ and $n_2$ values $y_2$ might arise are as

- repeat measurements of a single measurand made using two different devices,
- single measurements made with two devices on multiple measurands from a stable process (only one device being used for a given measurand),
- repeat measurements made with one (linear) device on two measurands, and
- single measurements made using a single (linear) device on multiple measurands produced by two stable processes.

# Repeat Measurements of a Single Measurand Made Using Two Different Devices




$$y_{1i}\text{'s} \sim \text{ind}\left(x+\delta_1, \sigma_{\text{device1}}\right) \text{ independent of}$$
$$y_{2i}\text{'s} \sim \text{ind}\left(x+\delta_2, \sigma_{\text{device2}}\right)$$

Here the $t$ interval is for $\delta_1 - \delta_2$ and the $F$ interval is for $\sigma_{\text{device1}}/\sigma_{\text{device2}}$.

**Measuring Styrofoam "Packing Peanut" Size**. In an in-class measurement exercise, two students used the same caliper to measure the "size" of a single Styrofoam "packing peanut" according to a class-standard measurement protocol. Summary statistics are

| Student 1 | Student 2 |
|---|---|
| $n_1 = 4$ | $n_2 = 6$ |
| $\bar{y}_1 = 1.42\,\mathrm{cm}$ | $\bar{y}_2 = 1.44\,\mathrm{cm}$ |
| $s_1 = .20\,\mathrm{cm}$ | $s_2 = .40\,\mathrm{cm}$ |

Here, the difference in "devices" is the difference in "operators" making the measurements. Let's compare"devices."

To begin,

$$\hat{\nu} = \frac{\left(\frac{(.20)^2}{4} + \frac{(.40)^2}{6}\right)^2}{\frac{(.20)^4}{(4-1)(4)^2} + \frac{(.40)^4}{(6-1)(6)^2}} \approx 7.7$$

# Example 2B-1 (cont.)

Or being more conservative, $\hat{\nu}^* = \min\left((4-1),(6-1)\right) = 3$. So (rounding the first down to 7) one should use either 7 or 3 degrees of freedom in the $t$ formula.

For sake of example, using $\hat{\nu}^* = 3$ degrees of freedom, the upper 2.5% point of the $t$ distribution with 3 df is 3.182. So 95% confidence limits for the difference in biases for the two operators using this caliper are

$$1.42 - 1.44 \pm 3.182\sqrt{\frac{(.20)^2}{4} + \frac{(.40)^2}{6}}$$

i.e.

$$-.02\,\mathrm{cm} \pm .61\,\mathrm{cm}$$

The apparent difference in biases is small in comparison to the imprecision associated with that difference.

# Example 2B-1 (cont.)

Then, since the upper 2.5% point of the $F_{3,5}$ distribution is 7.764 and the upper 2.5% point of the $F_{5,3}$ distribution is 14.885, 95% confidence limits for the ratio of standard deviations of measurement for the two operators are

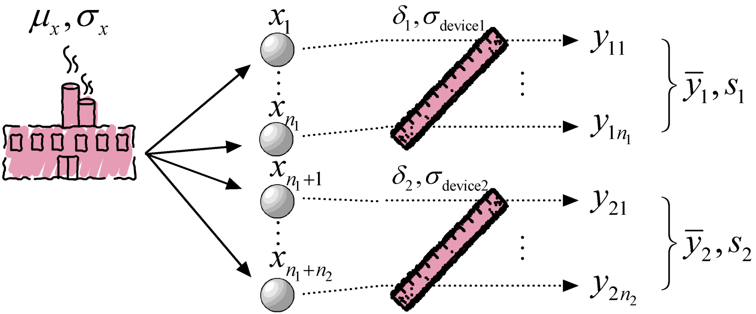
$$\frac{.20}{.40} \cdot \frac{1}{\sqrt{7.764}} \quad \text{and} \quad \frac{.20}{.40} \cdot \sqrt{14.885}$$

i.e.

$$.19 \quad \text{and} \quad 1.93$$

Since this interval covers values both smaller and larger than 1.00, there is (in the limited information available here) no clear indicator of which of these students is the most consistent in his or her use of the caliper in this measuring task.

# Single Measurements Made With Two (Linear) Devices On Multiple Measurands From a Stable Process (Only One Device Being Used for a Given Measurand)




$$y_{1i}\text{'s} \sim \text{ind}\left(\mu_x + \delta_1, \sqrt{\sigma_x^2 + \sigma_{\text{device1}}^2}\right) \text{ independent of}$$

$$y_{2i}\text{'s} \sim \text{ind}\left(\mu_x + \delta_2, \sqrt{\sigma_x^2 + \sigma_{\text{device2}}^2}\right)$$

# Single Measurements Made With Two (Linear) Devices On Multiple Measurands From a Stable Process (Only One Device Being Used for a Given Measurand) (cont.)

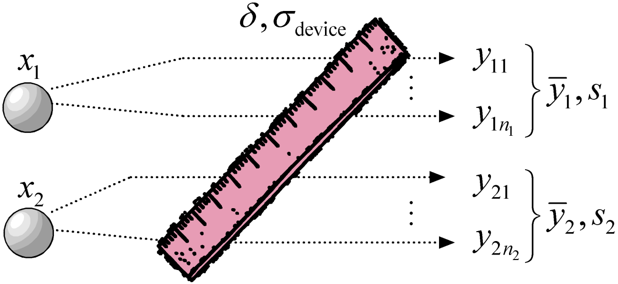
Here the $t$ interval is for $\delta_1 - \delta_2$ and the $F$ interval is for $\sqrt{\sigma_x^2 + \sigma^2_{\text{device1}}} / \sqrt{\sigma_x^2 + \sigma^2_{\text{device2}}}$.

The $t$ interval

- is important when measurement is destructive and the previous data collection plan can't be used, and
- is typically less informative (for a given pair of sample sizes) than the previous method when both can be used.

The $F$ interval provides (only) a somewhat indirect comparison of the two device precisions.

# Repeat Measurements Made With One (Linear) Device On Two Measurands

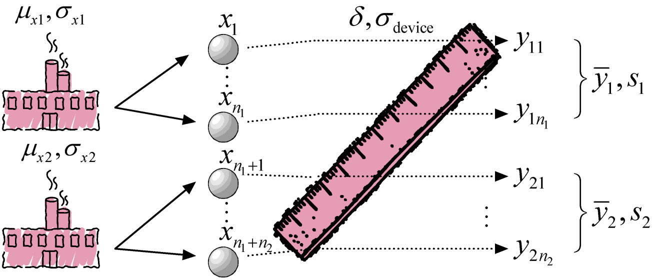



$y_{1i}$'s $\sim$ ind $\left(x_1 + \delta, \sigma_{\text{device}}\right)$ independent of

$y_{2i}$'s $\sim$ ind $\left(x_2 + \delta, \sigma_{\text{device}}\right)$

Here the $t$ interval is for $x_1 - x_2$.

# Single Measurements Made Using a Single (Linear) Device on Multiple Measurands Produced by Two Stable Processes




$$y_{1i}\text{'s} \sim \text{ind}\left(\mu_{x1}+\delta, \sqrt{\sigma_{x1}^2+\sigma_{\text{device}}^2}\right) \text{ independent of}$$

$$y_{2i}\text{'s} \sim \text{ind}\left(\mu_{x2}+\delta, \sqrt{\sigma_{x2}^2+\sigma_{\text{device}}^2}\right)$$

Here the $t$ interval is for $\mu_{x1} - \mu_{x2}$.

# Single Measurements Made Using a Single (Linear) Device on Multiple Measurands Produced by Two Stable Processes (cont.)

The $F$ interval estimates $\sqrt{\sigma_{x1}^2 + \sigma_{\text{device}}^2} / \sqrt{\sigma_{x2}^2 + \sigma_{\text{device}}^2}$ and provides an (only) indirect comparison of process precisions.