Module 2A

Elementary statistical methods and measurement: simple one-sample methods

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

March 4, 2008

Basic "Stat 101" Formulas for a Single Sample

"Ordinary" confidence interval formulas based on a model that says that $y_1, y_2, \ldots, y_n$ are a sample from a normal distribution with mean μ and standard deviation σ are

$$\overline{y} \pm t \frac{s}{\sqrt{n}} \quad \text{for estimating } \mu$$

and

$$\left(s\sqrt{\frac{n-1}{\chi^2_{\text{upper}}}}, s\sqrt{\frac{n-1}{\chi^2_{\text{lower}}}} \right) \quad \text{for estimating } \sigma \quad .$$

What we usually don't point out in Stat 101 is that the data collection plan governs the practical meaning of μ and σ. Here we proceed to think carefully about this, *paying attention to the impact of measurement error.*

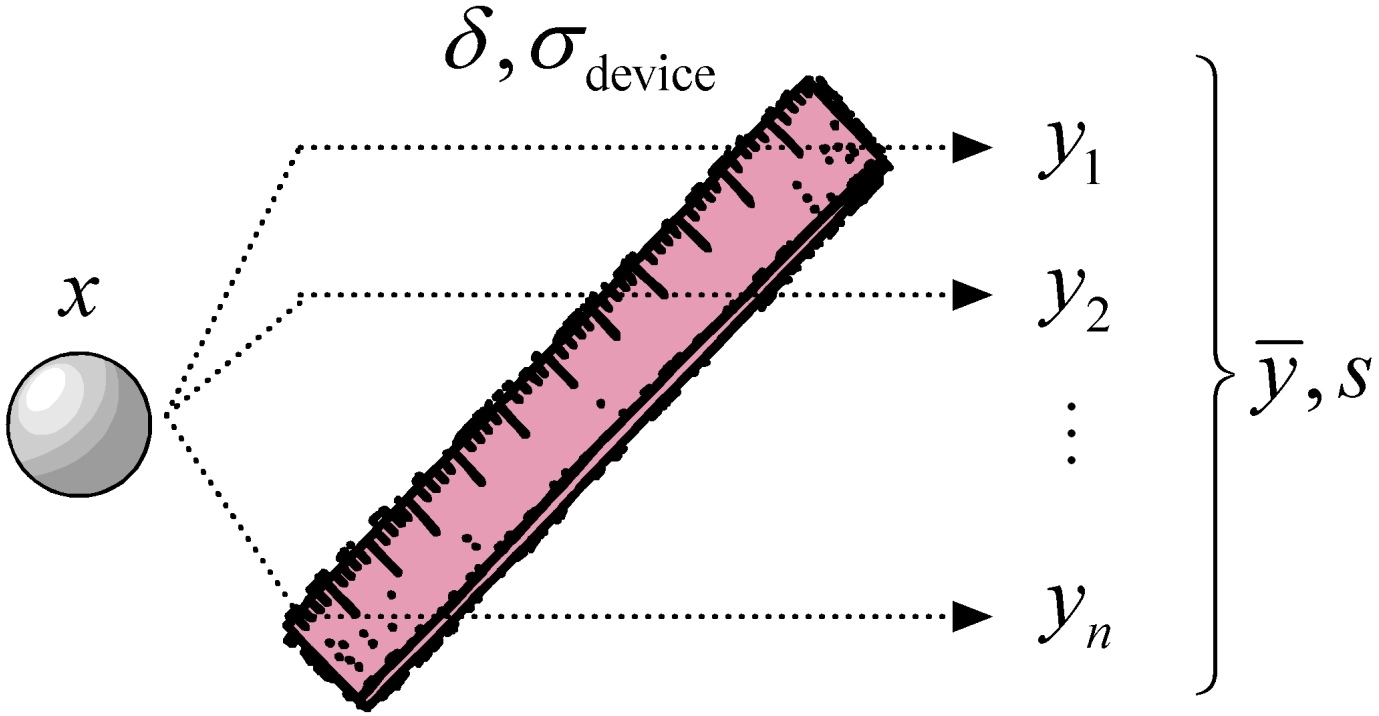
Two "Obvious" Applications

Two different ways that a single "sample" of n observed values y might arise are as

- repeat measurements on a single measurand made using the same device, and
- single measurements made on multiple measurands coming from a stable process made using the same device.

Henceforth we will use the language "device" as shorthand for a fixed combination of physical measurement equipment, operator identity, measurement procedure, and surrounding physical circumstances (like time of day, temperature, etc.).

Repeat Measurements on a Single Measurand Made With the Same Device



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

Repeat Measurements on a Single Measurand Made With the Same Device (cont.)

The t confidence interval for a mean estimates

$$x + \delta = \text{measurand plus bias .}$$

So

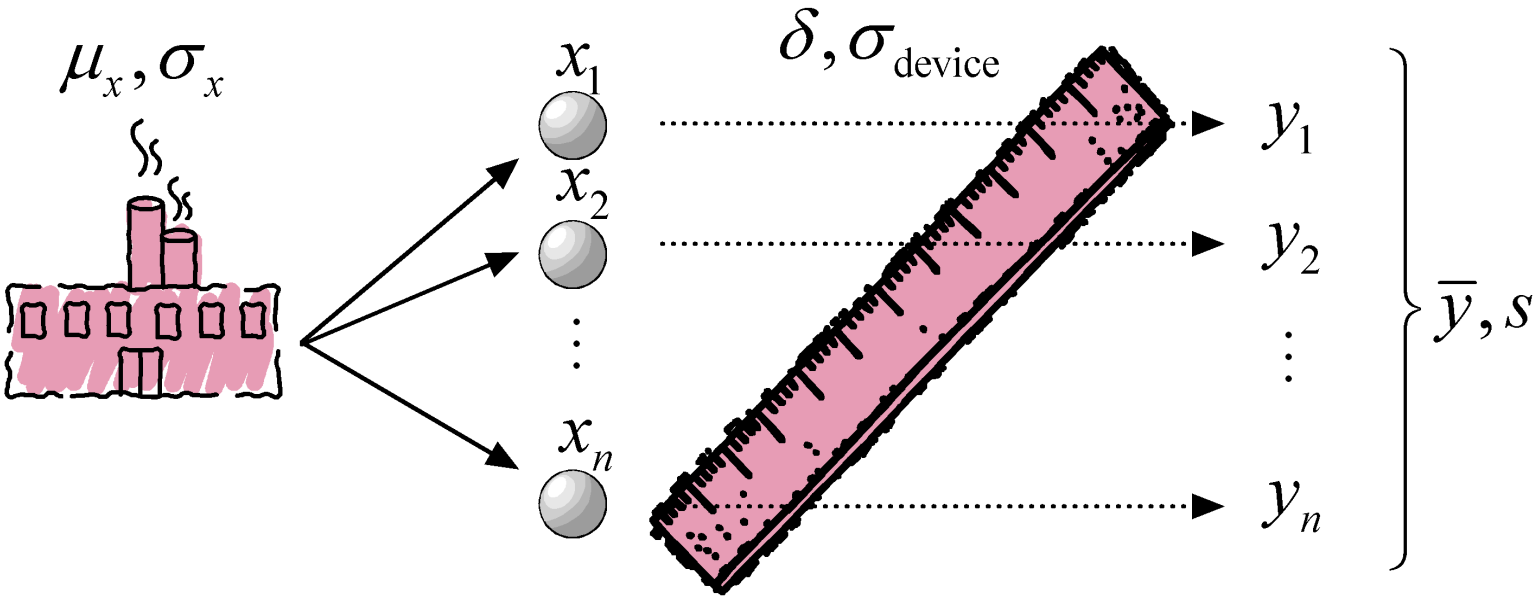
- if the device is known to be well-calibrated (one is sure that $\delta = 0$), the limits $\overline{y} \pm ts/\sqrt{n}$ based on $\nu = n - 1$ df are limits for x, and
- if what is being measured is a *standard* **for which** x **is known**, the limits

$$(\overline{y} - x) \pm t\frac{s}{\sqrt{n}}$$

  (once again based on $\nu = n - 1$ df) estimate the device bias, δ.

Further, applying the χ^2 confidence interval for a standard deviation, one has an inference for the size of the device "noise," σ_{device}.

Single Measurements Made on Multiple Measurands Coming from a Stable Process Made With the Same (Linear) Device



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

Single Measurements Made on Multiple Measurands Coming from a Stable Process Made With the Same (Linear) Device (cont.)

Here

- the limits $\overline{y} \pm ts/\sqrt{n}$ (for t based on $\nu = n - 1$ df) are limits for

$$\mu_x + \delta = \text{the mean of the distribution of true values plus bias ,}$$

  and
- the quantity s estimates $\sigma_y = \sqrt{\sigma_x^2 + \sigma_{\text{device}}^2}$, that really isn't of fundamental interest. So there is little point in direct application of the χ^2 confidence limits for a standard deviation in this context.

Example 2A-1

Measuring Concentration. $n = 5$ consecutive concentration measurements made by a single analyst on a single physical specimen of material using a particular assay machine (real units are not available, so call them "moles per liter," $\mathrm{mol/l}$).

$$1.0025, .9820, 1.0105, 1.0110, .9960$$

These have $\bar{y} = 1.0004\,\mathrm{mol/l}$ and $s = .0120\,\mathrm{mol/l}$.
Consulting a χ^2 table using $\nu = 5 - 1 = 4$ df, 95% confidence limits for σ_{device} (the size of basic measurement variability) are

$$.0120\sqrt{\frac{4}{11.143}} \text{ and } .0120\sqrt{\frac{4}{.484}} \quad \text{i.e.} \quad .0072\,\mathrm{mol/l} \text{ and } .0345\,\mathrm{mol/l}\,.$$

(Ordinary small sample sizes give very wide confidence limits for a standard deviation.)

Example 2A-1 (cont.)

Consulting a t table also using 4 df, 95% confidence limits for the measurand plus instrument bias $(x + \delta)$ are

$$1.0004 \pm 2.776 \frac{.0120}{\sqrt{5}} \quad \text{i.e.} \quad 1.0004\,\mathrm{mol/\,l} \pm .0149\,\mathrm{mol/\,l}\ .$$

If the measurement were done on a standard material "known" to have actual concentration $1.0000\,\mathrm{mol/\,l}$, these limits correspond to limits for device bias

$$0.0004\,\mathrm{mol/\,l} \pm .0149\,\mathrm{mol/\,l}\ .$$

Finally, suppose that subsequently samples from $n = 20$ different batches are analyzed and $\overline{y} = .9954$ and $s_y = .0300$. The 95% t confidence limits

$$.9954 \pm 2.093 \frac{.0300}{\sqrt{20}} \quad \text{i.e.} \quad .9954 \pm .0140$$

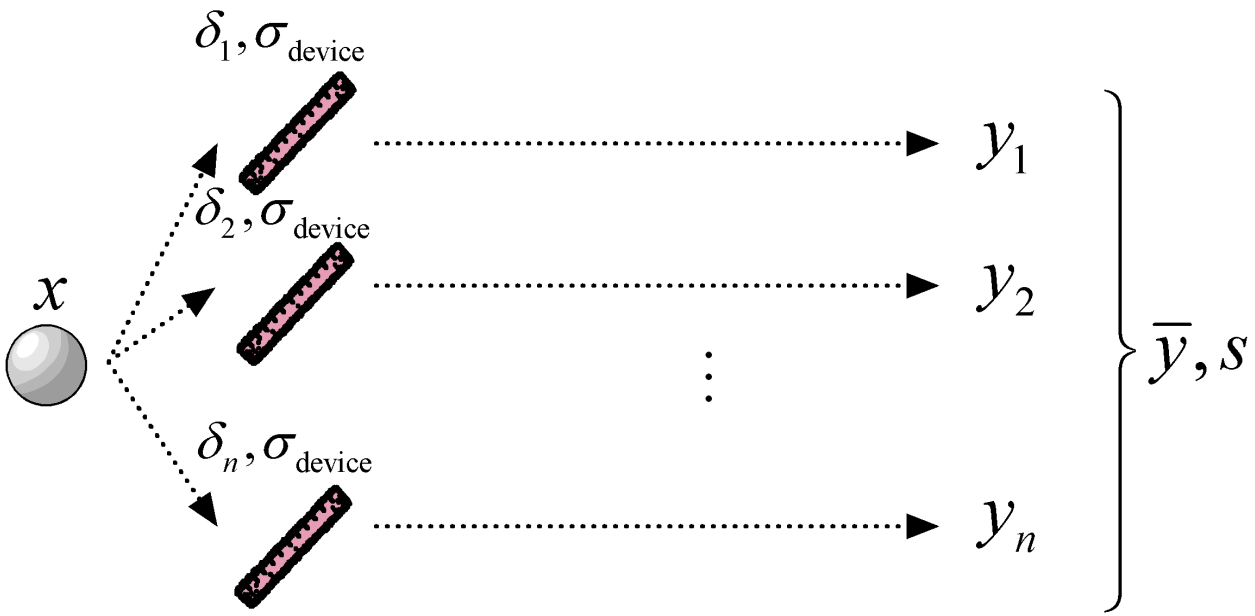
are for $\mu_x + \delta$, the process mean *plus* any device bias/systematic error.

Two Other Applications

Two additional ways that a single "sample" of n observed values y might arise are

- as measurements of a single measurand made using multiple devices (from a large population of such devices), and
- as differences in measurements on multiple measurands made using two (linear) devices.

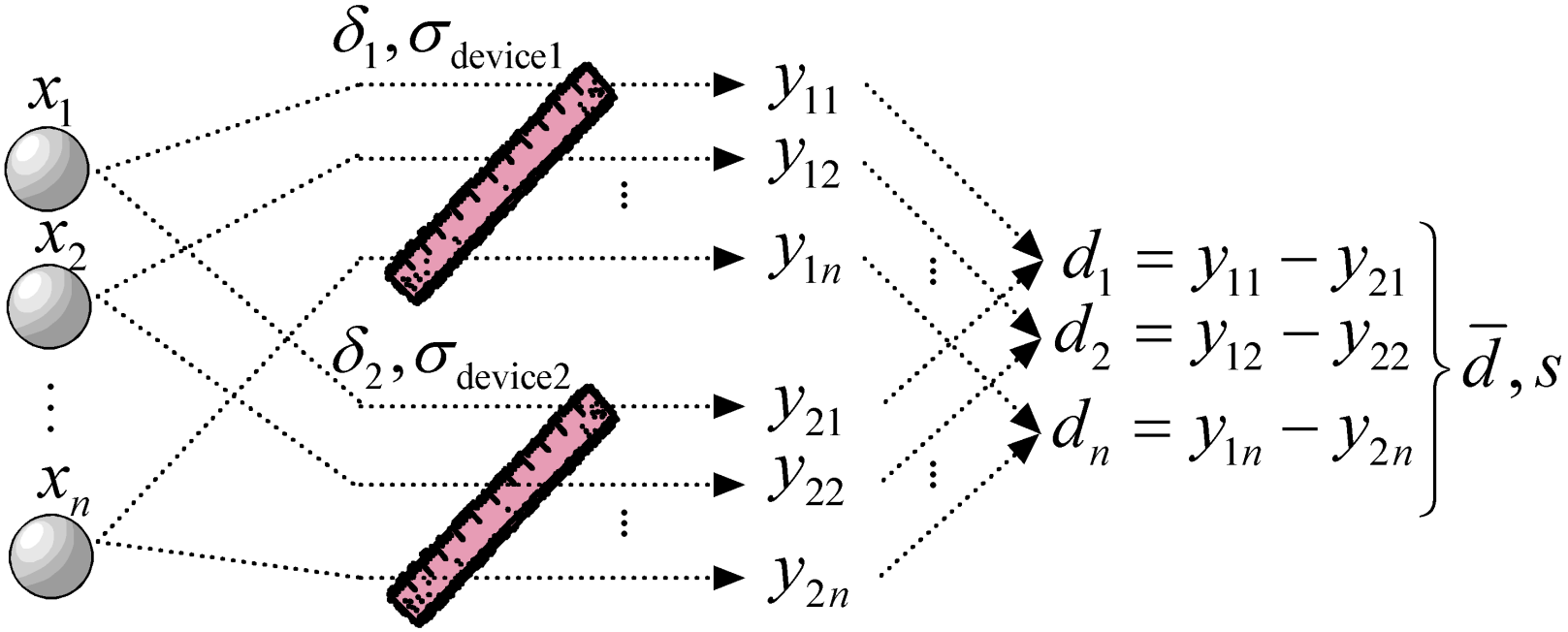
Measurements of a Single Measurand Made Using Multiple Devices



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

Here the t interval estimates $x + \mu_\delta$ and the χ^2 interval estimates $\sqrt{\sigma_\delta^2 + \sigma_{\text{device}}^2}$. (Note that the latter *can* represent $\sigma_{\text{R\&R}}$.)

Differences in Measurements on Multiple Measurands Made using Two (Linear) Devices



$$d_i\text{'s} \sim \text{ind}\left(\delta_1 - \delta_2, \sqrt{\sigma^2_{\text{device1}} + \sigma^2_{\text{device2}}}\right)$$

Here the t interval estimates $\delta_1 - \delta_2$ and the χ^2 interval estimates $\sqrt{\sigma^2_{\text{device1}} + \sigma^2_{\text{device2}}}$.