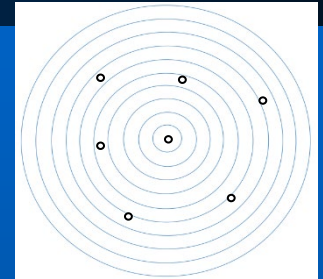


ME103:: Experimentation and Measurements

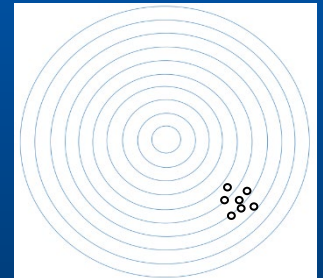
Lecture #4

Bias and Random Error

- Precision error = "random scatter" in the experimental results that CAN be found from repeated measurements.
 - Different for each successive measurement but have an average value of zero
- Bias error = **Systematic** errors that CANNOT be found from repeated measurements
 - Occur the same way each time a measurement is made
 - e.g. if a scale consistently reads 5% high, the entire set of measurements will be biased +5% above the true value
 - Can be estimated by comparison of the instrument to a more accurate standard, from knowledge of how the instrument was calibrated, or from experience with that specific instrument (or type of instrument)



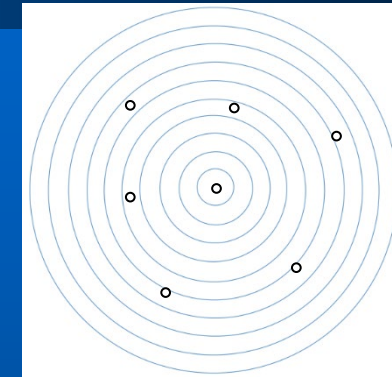
Random or
Precision Error



Systematic
or
Bias Error

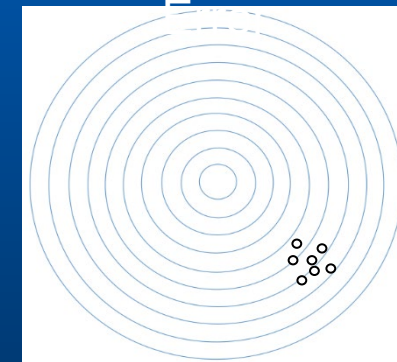
Some Causes of Bias and Random Errors

- Precision error from:
 - disturbances (uncompensated vibrations, temperature, pressure, etc.),
 - observation errors,
 - friction,
 - not repeating experimental procedure properly,
 - etc.



Random or Precision

- Bias error from:
 - calibration
 - reoccurring human error,
 - defective equipment,
 - theoretical (e.g. incorrect system model),
 - etc.



Systematic or Bias Error

- Blunders
 - recording data incorrectly,
 - reading scale incorrectly,
 - etc.



Blunders →

Terminology: Uncertainty

- **Uncertainty:** A measure of the *expected error* in measurement that combines both systematic and random errors.
- For sample x , if B represents an estimate of bias (systematic) errors and P represents an estimate of precision (random) errors, then the uncertainty U is usually estimated as:

$$U_x = \left(B_x^2 + P_x^2 \right)^{1/2}$$

Terminology: Uncertainty

- Most of the time, you will be dealing with **random** errors
- The best way to **quantify random errors** is to take a lot of measurements and perform **statistics**

How Do We Go about Quantifying Error?

- ▶ When you make measurements and take data, you will get a **population** of data
- ▶ Within this **population**, you will have a **distribution about a mean value**

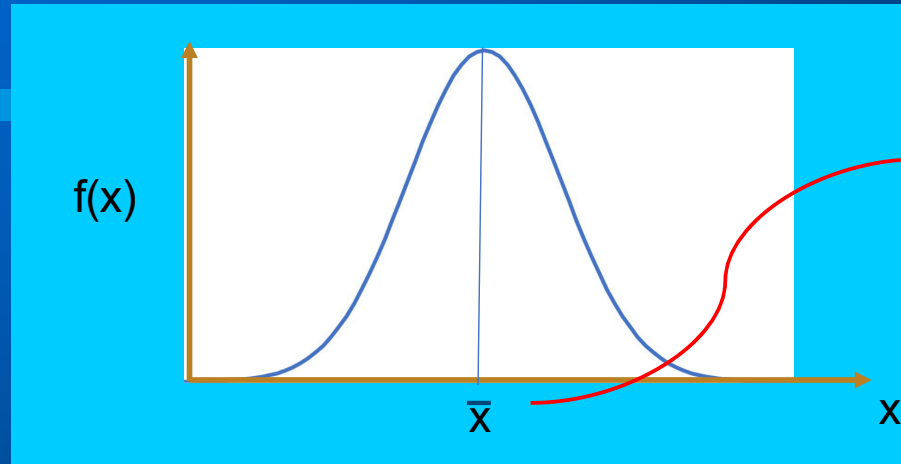
Average or mean

$$\bar{X} = \sum_{i=1}^n \frac{x_i}{n}$$

x_i = data point

n = number of data points

If your data are independent of one another, then the distribution about the mean is a Gaussian



Centered around the mean value

$$PDF = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

x = individual data point

μ = TRUE value

σ = standard of deviation (gives the "spread" of the Gaussian)

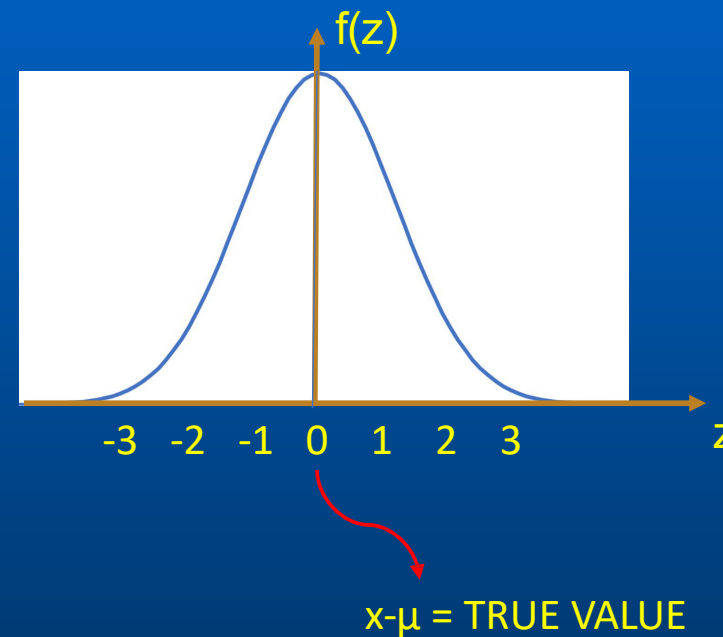
$$\sigma = \sqrt{\frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n}}$$

To make life easier for us, especially when we are comparing distributions, we can **re-write the Gaussian distribution**:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$z = \left| \frac{x - \mu}{\sigma} \right|$$

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$



By re-writing your Gaussian in terms of z , you have the following...

$$z = \left| \frac{x - \mu}{\sigma} \right|$$

$$Z\sigma = |x - \mu|$$

$$x = \mu \pm Z\sigma \quad \text{or} \quad \mu = x \pm Z\sigma$$

$\mu = x \pm z\sigma$ is an important statement: If we do not know the true value μ , we have at least a range of what the true value is

$x = \mu \pm z\sigma$ tells us how good our data is and tells us if our data point is an outlier. It helps us decide whether our data can be "tossed" if it is σ , 2σ , 3σ away from μ

