MEEG 215: Statistics and Calibration

Lab Objectives

1) Calibrate the workstation instrumentation using standards or other pre-calibrated instruments.
2) Use measurement results, engineering judgment and deduction to determine the uncertainty of each measurement device.
3) Measure the force to break a steel wire as a function of wire length and diameter.
4) Apply statistical analyses and measurement uncertainties to help draw conclusions about the functional relationships between break force and the geometric properties of the wire.

Background

Statistics

Statistics is the mathematical science pertaining to the collection, analysis, interpretation and presentation of data. Statistical analyses are used to establish relationships between variables, develop models and forecast outcomes. Statistics is an invaluable tool to the design engineer for determining the probabilities of critical events (e.g. failure).

The following example will demonstrate how engineers apply statistics. Let’s consider the design of a doorway. If it is too low, people will be uncomfortable; too high and the space and materials are wasted
making for an inefficient design. A common strategy is to design for failure for an acceptably small fraction of the time; here we will specify a door such that 99.9% of American males (since they are statistically taller than females) may pass unobstructed. Engineers use statistical methods to accomplish such goals.

A large fraction of natural variables, including American male height, follow the Normal or Gaussian distribution shown below; the characteristic 'bell' curve depicts relative probability versus variable value. The most probable value is the mean or average (the mean is zero in the figure) and the curve width is defined in terms of the standard deviation. This curve is used to extract probabilities of events. For exactly one sample, the probability that the value will be less than the mean is 50%. If the doorway height was designed for the population mean, 50% would pass through unobstructed. According to the graph, the design needs to be 3σ, or 3 standard deviations greater than the mean, so that 99.9% may pass unobstructed. If the population were known to be normal with a mean of 69.4" and a standard deviation of 2.8", it would be trivial to design the door; three times the standard deviations is 8.4", so the doorway would need to be 69.4" + 8.4" = 77.8" tall to meet the performance criterion.

The population is the complete group under study; in this case, the population is every American. It is impossible to measure every American, so the population cannot be known exactly. The experimenter must design an experiment and measure a randomly selected group. The individual height measurements taken from this group are called random samples and the result of all of the measurements is a sample distribution. The sample distribution cannot have exactly the same statistics as the population. The mean and standard deviation are the most common statistical descriptors of a distribution of samples.

Mathematically, the mean is:

\[ \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \]

If the first three sampled measurements are 72", 75" and 70". The mean is \( \frac{72"+75"+70"}{3} = 72.3" \). The standard deviation is:

\[ \sigma_n = \sqrt{\frac{1}{N-1} \sum_{i=1}^{n} (X - \bar{X}_i)^2} \]

For the same measurements, the standard deviation would be

\[ \sqrt{\frac{1}{2} \left( (72.3"-72")^2 + (72.3"-75")^2 + (72.3"-70")^2 \right)} = 2.5" \]

It is intuitive that the quality of the statistics will be a function of the sample size. For example, if only one person is measured, little is known of the population mean. If ten people are measured one would have more confidence in the population mean and so on. The central limit theorem can be used to make inferences about population means; it basically asserts that the confidence in the knowledge of the population mean increases with the square root of the number of samples. Following 10,000 measurements in the U.S., the engineer found that the sample group had an average of 69.4" with a standard deviation of 2.8". This is more compactly written as 69.4"±2.8" or N(69.4",2.8").
The engineer must now make an inference about the population based on a limited sample. Imagine the engineer were to measure 10,000 random American males 1,000 times. He would obtain 1,000 average heights, \( \bar{Y} \), which are normally distributed about a mean, \( \mu \). According to Central Limit Theorem (CLT), the standard deviation of the \( \bar{Y} \) distribution from N=10,000 random samples is the standard deviation in Y divided by the square root of N: \( \sigma_{\bar{Y}} = \frac{\sigma_Y}{\sqrt{N}} \). Given N=10,000 and \( \sigma_Y = 2.8" \) gives \( \sigma_{\bar{Y}} = 0.028" \). Therefore, it can be concluded that the average American male height is \( 69.4 \pm 0.028" \) with 68.2% confidence. Equivalently, average American male height is \( 69.4 \pm 0.056" \) with 95.4% confidence. It is customary to use \( 2\sigma \) for a 95% (rounded) confidence interval.

The door was designed for 77.86" and was such a success that a market developed in Norway. The engineer measured 10,000 random Norwegian males to determine if they are from the same population. The Norwegian sample had a distribution of 70.8"±2.64". Accordingly, \( \sigma_Y = 0.026" \). CLT enables us to make a defensible statement about our confidence in the assessments of population means. Based on our knowledge of normally distributed variable, we can say with 99.9% confidence that the American male mean is no more than \( 69.49" (\mu \pm 3\sigma) \) and the Norwegian male mean is no less than 70.72". We can therefore say with better than 99.9% confidence that Norwegian males are taller than American males. The analysis is represented graphically below.

**Measurement and Data Collection**

It is impossible to make an exact measurement which means that while the population has an inherent distribution, imperfect measurements provide an altered representation of the distribution of interest. The
difference between the two is related to the **accuracy** of the measurement. The engineer’s job is to understand the shortcomings of the measurement so that the probable dispersion of any measurement from its true value may be evaluated. A measurement must always be accompanied by its **uncertainty** if it is to be of value to a reader. A working measurement vocabulary is provided below.

**Accuracy:** The deviation from the true value of a known calibration standard (not truly known, but to within some uncertainty).

**Bias Error:** An artificial offset of the measurement. Bias errors are often results of thermal or electrical influences.

**Calibration:** The use of standard or otherwise known input to determine the output scale on an instrument.

**Random Error:** Random fluctuation about the mean measurement which may be biased. Systematic fluctuations that are not random often occur at 60 Hz due to electrical noise.

**Range:** The maximum value an instrument can measure; 12” for a 12” ruler

**Repeatability:** The variation in measurements of the same artifact.

**Resolution:** The minimum detectable change in value; 1/32” for many rulers, or 0.001” for lab calipers

**Standard:** An artifact with a known input to the measurement device. The freezing point of pure water at ambient pressure is the standard for 0°C.

**Uncertainty:** The degree to which a measurement is uncertain. Engineering judgment or calibration datasheets from the manufacturer are needed to assess uncertainty. The uncertainty can be no less than the greater of resolution, repeatability and bias. Without better information, the uncertainty can be assumed equal to five times this value. Uncertainty, \( U_c \), is often treated as a standard deviation. By definition, the *true value deviates less than 1 \( U_c \) from the measurement with 68.2% confidence* (refer to the first figure). As an example, with a nominal measurement of 5 inches and an uncertainty of 0.1 in, one can claim that the true value is less than 5+3*0.1 = 5.3” with 99.9% confidence.

### Procedure

Each workstation is equipped with several measurement capabilities. A pressure gauge and a load cell are available for redundant force measurements. Calipers, measuring tape and a Linear Variable Differential Transformer (LVDT) are available for dimensional measurements. The pressure gauge, calipers and measuring tape are pre-calibrated mechanical measurements (note: the pressure measurement is at the regulator and not the cylinder). There is an uncertainty in this pressure, there are pressure losses from the regulator to the piston and there is friction in the cylinder all of which contribute to measurement uncertainty. A load cell and LVDT provide voltage responses that are ideally linear functions of the applied load and displacement, respectively. Each is sensitive to inputs along the instrument axis only.

In this lab you will calibrate the electronic instrumentation and determine the measurement limitations of each. You will use this knowledge to determine the break force variability of a material and the dependence of break force on length and diameter.
1) **Load cell calibration**

   a) Make sure the load cell is in the upmost position and loosen the safety collar if necessary (7/64)
   b) Slowly increase the tension pressure (left regulator) to 90 psi
   c) Mount a steel hanger wire in the upper wire grip
   d) Mount the hooked weight hanger
   e) Open MEEG215 Labview Software (MEEG215lab.vi)
   f) Click arrow in upper left corner to run
   g) Click ‘set files to start measurement’ button
   h) Zero force under ‘voltages’ tab. Under calibrated force tab confirm zero voltage and a force calibration of 1.
   i) Place a calibrated weight (5kg±5g) on the hanger and record the voltage output. Remove and repeat with each of the weights. What can be said about the weights? About the load cell? Can you add these instead of stacking the weights to obtain a voltage for the sum total?
   j) **Taking care to avoid tipping,** stack the weights (misalign slot) and measure voltage at each increment. Do these deviate from sums in i)? If so, why?
   k) Pass the weights and hanger to the next group
   l) In Excel, convert kg to lb (2.2046lb/kg) and plot force versus voltage
      To plot, click insert then scatter on the toolbar. Right click on the blank plot and click ‘select data’. Click add. Select all voltage values for X and all force values for Y. Hit ‘OK’.
   m) Right click on data to add a linear trendline and check add equation and $R^2$ value. Right click on label and format trendline label. Highlight scientific and use 5 decimal places.
   n) Record (using trendline options) the slope and the $R^2$ value (indication of linearity).
   o) Enter the slope as the calibration constant under the calibrated force tab in Labview
   p) Place voltmeter on DC 20V and measure the voltage from ‘V’ to ‘G’ on the black conditioner. If the voltage is non-zero, you will need to subtract the zero-load condition from subsequent measurements to get the change in voltage needed for use with your calibration constant.
   q) Use the voltmeter and software to measure the unknown weight 3 times each (remember to zero voltage for zero load). Do they differ? Why? Which can be trusted?
   r) Compare the average Labview measurement with the TA’s measurements. The measurements should differ by less than 0.01 lb
   s) Comment on the resolution, repeatability, bias and accuracy of each technique

2) **LVDT calibration**

The LVDT should be mounted in the lower left corner of the base plate. Alternating current in the primary coil of an LVDT induces voltages across two secondary coils. When the magnetic core (probe) splits the secondary coils, the coils cancel each other and the output is zero. As the core moves toward one of the coils, a proportional differential voltage develops.

In this case, we do not have a dimensional standard. Instead, we have a calibration block with four unique measurement lengths, L1, L2, L3 and L4. For calibration purposes, only the change in length of the LVDT is important. As illustrated above, we can infer these length changes if L stays constant and if we can accurately measure L1, L2, L3 and L4 with a precalibrated device.
a) Do not ‘zero’ the LVDT channel. Insert the calibration block with side 4 contacting the LVDT. Load the leading edge so the LVDT does not push it up and load the back edge flush against the baseplate. The voltage needs to lie between 7 and 10. Reposition LVDT if necessary and lock it securely. Hint: slow insertion and removal will give best results.

b) Record the voltage given by the software

c) Repeat for each side

d) Repeat 4 times to examine repeatability and establish a statistically robust calibration

e) Measure L1, L2, L3 and L4 with the calipers. Take care to minimize the measurement by minimizing the misalignment angle.

f) Repeat 4 times to examine repeatability and establish a statistically robust calibration.

g) Plot length versus voltage and compute calibration trendline

h) Comment on accuracy, resolution and repeatability of the LVDT displacement measurement.

i) Measure the thickness of the unknown sample and compare with caliper results. Hint: use the same concept illustrated in the above figure. The measurements should be within 0.003”.

3) **Measuring mechanical properties of a material**

a) Section an appropriate length of the thinnest steel wire (ø approximately 0.009”)

b) Measure the first wire diameter four times in the same location and then in multiple locations with the calipers.

c) Release the collar if necessary, lower the piston and zero the load cell

d) Mount the wire in the grips

e) Use measuring tape to measure length and use engineering judgment to determine one standard uncertainty; by definition, there is a 68.2% probability that the true value lies less than 1Uc away from the measurement.

f) After properly calibrating and ‘zeroing’ the software, reduce the average save interval to provide an effective refresh rate on the average data plots. If the refresh rate is too slow, the data will lag the force and provide an inaccurate measure (0.5s is a good starting point)
g) Gradually increase the tension pressure, monitoring pressure and force until the sample breaks. **Set safety collar if force exceeds 32 lb.** Note that the plotted points are averages of the sampling window.

h) Record the break force from software and break pressure from the gage. Zoom in on the graph by right-clicking, unchecking autoscale and adjusting the axis limits. Recheck autoscale when finished.

i) Repeat three more times to obtain statistics taking care to reproduce the previous conditions

j) Add the extender rod and repeat the experiments for shorter samples

k) Compute averages and standard deviations

l) Repeat break test for the two thicker sample using the short test length

m) Quickly re-weigh your unknown to assess calibration stability with time.

**Analysis Guidelines**

Reports should follow a technical format which includes a brief introduction, a description of methods, results, discussion and conclusions. Technical writing is a learned skill that is best taught through study of others’ works and practice. Google Scholar is a tremendous tool for locating journal articles. This will be your best resource for studying high quality works of technical writing.

The introduction is often the motivating background for the report. The Methods section describes what is being done. This is in effect a procedure, but it should be in technical writing format and not in a bulleted list as given here. You will need to exercise judgment on which details are appropriate. For example, you should describe the measurement of precalibrated weights; you should not describe the use of excel to calculate trendlines and plot data. Assume that the reader is comfortable with data analysis. The results may be presented as the groups see fit. Include Tables for raw data and **select** graphs to demonstrate trends. It is important that you carefully consider the graphs you include. Graphical presentation is arguably more important in engineering that written presentation. To include a thoughtless graph is akin to rambling incoherent paragraph. Effective graphics optimize the transfer of results and trends to the reader. Again, see published works for examples. You should discuss the meaning of the results in the Discussion and offer closing remarks or conclusions in the Conclusions.

The following will guide you through the experiments. Each group needs an organized data sheet within the group notebook with room for each needed measurement. Each group will also make a copy to be handed in at the beginning of class. Some examples of effective tabular structures are given below.

<table>
<thead>
<tr>
<th>1) Force Measurement</th>
<th>Sample to sample repeatability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (#)</td>
<td>1</td>
</tr>
<tr>
<td>Voltage</td>
<td></td>
</tr>
</tbody>
</table>
Cumulative weight calibration

<table>
<thead>
<tr>
<th>Weight (kg)</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>Slope (lb/V)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (lb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voltage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unknown Measurement

<table>
<thead>
<tr>
<th>Initial voltmeter reading (V)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>measurement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>mean</th>
<th>st. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labview (lb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voltmeter (V)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibrated voltmeter (lb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) What are the resolution and repeatability of calibration weight load cell measurements?
b) If the calibration weights are known to be 5 kg±5g with 95% confidence, use your engineering judgment to calculate the uncertainty of the load cell (±1Uc = 68.2% confidence). Your result may or may not be a function of the measurement. Hint: You can simulate weight variations of ±1Uc in Excel and record the response. What would the calibration constant be if the weights were actually 4.995 kg instead of 5.000 kg?
c) Does the force uncertainty change with time? Why or why not?
d) Provide a 95% confidence interval for your measurement of a calibration weight and the unknown weight in pounds.

2) LVDT

LVDT Voltage Measurement

<table>
<thead>
<tr>
<th>Block position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>mean</th>
<th>st. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Calibration Block Caliper Measurement – not provided
**Unknown Sample Measurement** – not provided

a) What are the resolution and repeatability of LVDT measurements in inches? Describe the likely error sources (instrument noise, user error, machining form errors, etc...)
b) What are the resolution and repeatability of caliper measurements in inches? Describe the likely error sources (instrument noise, user error, machining form errors, etc...)
c) If the calipers are accurate to within 0.0005” of the measurement with 95% confidence, what is the uncertainty in LVDT measurements
d) Provide a 95% confidence interval for the thickness measurement of the unknown sample.

3) Material Properties

**Repeatability and Variability of Diameter of Wire 1**

<table>
<thead>
<tr>
<th>Measurement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>mean</th>
<th>st. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Φ (in) same place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Φ (in) along wire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Break Force Variability of Long Wire**

<table>
<thead>
<tr>
<th>Wire</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>mean</th>
<th>st. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>F_{br} load cell (lb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Break pressure (psi)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_{br} pressure gauge (lb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>μ</th>
<th>σ, U_c</th>
<th>μ</th>
<th>σ, U_c</th>
<th>μ</th>
<th>σ, U_c</th>
<th>mean</th>
<th>st. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Φ (in)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (in²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L (in)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Force Intensity (F/A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Break Force Variability of Short Wire** – not provided

**Break Force of Thick Wire** – not provided

a) What are the resolution and repeatability of sample diameter measurements? Are they sensitive to sample irregularities?
b) Calculate the effective area over which the pressure acts.
c) What is the resolution in pressure? What is the resolution in force calculated from pressure measurements? What can be said about uncertainty in force from pressure readings?
d) Estimate a 95% confidence interval for wire length for each test.
e) Calculate the averages and standard deviations in break force for each test using load cell and pressure gage results. Use sample to sample variations and measurement uncertainties to comment on the results obtained from each measurement. Are the deviations due to the materials or are they artifacts from the measurement method?
f) Plot break force versus length for the thinnest wire. Use error bars to represent $2U_c$. Can you claim with 97.7% confidence that the shorter wire has a different mean break force than the longer wire?

g) Plot break force versus wire diameter. Can you claim with 97.7% confidence that break force varies with wire diameter? Fit the data with a trendline. Is there a Y intercept? What does it tell you about the relationship? Hint: slopes on log-log graphs reveal power law relationships; what does the power tell you about the relationship and how could you test this?

h) Plot break force versus area and fit the data. What does it tell you about failure of the material? What does the proportional constant represent? What do you think this curve would look like for plastic wires of the same diameters?