Tutorial 1:
Analysis and Graphical Representation of Data

Overview

People are good at interpreting visual information, and poor at interpreting text and numbers. Although you will learn how to do statistical calculations to get more meaning from data, very often the right graphical representation of the data is more useful in finding important patterns than all of the t-tests and ANOVAs covered later. The instincts of most good experimentalists are to plot up their measurements (in several ways) before doing any calculations. Hopefully this tutorial will encourage those instincts in you!

Objectives

In this tutorial you will explore different ways to represent the data from the in-class bolt measurements you did in Week 1. This exploration will have two purposes:

a) to check your data for self-consistency and freedom from gross errors
b) to determine how to get meaning from your data

Tutorial Assignment

You should also start the tutorial assignment before the tutorial period, so that you can identify the difficult points before class and therefore have more focused questions during the tutorial. The tutorials are intended as a place to discuss the questions that you found challenging. The in-class questions and the assigned textbook questions are intended to exercise needed skills. The assigned textbook questions are not directly marked, but the topics covered may be expected to appear in the course exams.

1. Read all of **Chapter 1** and **Sections 2.1 and 2.2** of the course textbook (Taylor)
2. Answer textbook questions **2.1, 2.2, 2.3 and 2.4** (on pages 35-36)
3. Download Matlab files **Tutorial_1_Bolt_Histogram.m** and **Tutorial_1_Bolt_Data.m** from Canvas.

4. The Bolt_Histogram assignment ultimately involves drawing histograms of the bolt length and diameter measurements made in class in Week 1. Unfortunately, the low quality of the data makes this a challenging task. The casual nature of the in-class experiment created numerous gross errors, far beyond anything that would occur in a serious experiment; no gross error is the typical experimental target. But even in the most well-controlled experiments it is important to check measured data for gross errors. Among the numerous gross errors in the in-class experiment, some were repairable, such as dimensions reported in cm rather than mm, and some were false without doubt, e.g., bolt diameter = 10000mm. Data review is a very sensitive process. Certainly, very definite and specific reasons have to be identified before correcting or rejecting suspicious data, else the overall data set will become seriously falsified.
5. The Bolt_Data assignment is a simple exercise in the use of several of the specialized statisitical and presentation functions available in Matlab.

Tutorial Report Requirements

Please write a report according to the following requirements and submit it as a PDF version on Canvas by noon on Monday January 14. Refer to **Tutorial_1_slides.pdf** for some hints.

1. Complete **Tutorial_1_Histogram.m.** When you first run the Matlab script, the errors in the raw data set will produce obviously faulty results. Correct or delete faulty data to enable your analysis to work well. In your report, point how you have cleaned the data with examples in bullet points. After cleaning the data, attach the good version of the two histograms. At the end, report the two average diameters and two average lengths of the various bolt sizes.
2. Complete **Tutorial_1_Bolt_Data.m**. Report the two Box Plots of the length measurements of the five example bolts and for the all bolts considered together. Compare the results to check whether they are consistent with those from **Tutorial_1_Histogram.m.** Report **means**, **medians, quartiles, standard deviations**, and **variances** for the **individual bolts** and for **all the bolts.**
3. Write a few sentences to comment on what you learnt from the data analysis about the sizes of the bolts and the measurement approach.