MECH 305/306 Statistics

Tutorial 1

Athena (Xiaohe) Liu

Jan-2019

About the TAs

- Athena Liu

- Email: sdathena@mail.ubc.ca

- Hooman Esfandiari

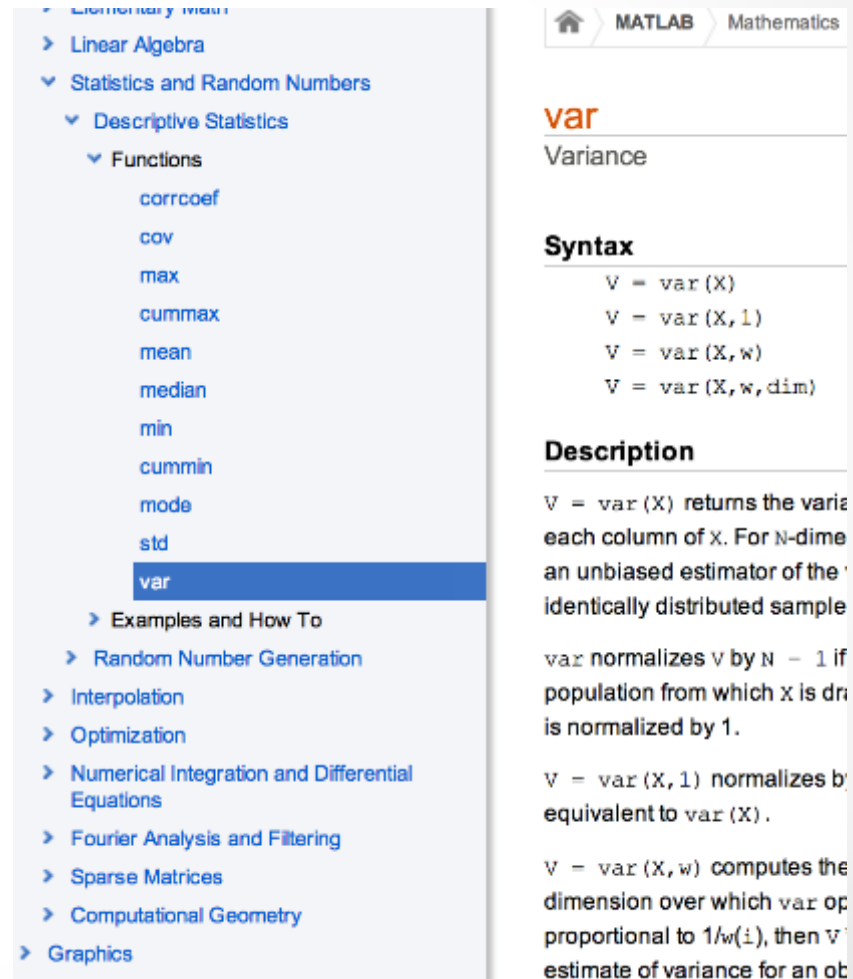- Email: hooman.esfandiari@gmail.com

What the tutorials are about?

- Before the tutorial:
  - Read through the tutorial instructions
  - Do the assigned textbook problems and bring any questions
- During the tutorial:
  - Hands-on instruction on useful MATLAB functions & data analyzing techniques (25 min) **(Bring your laptop!)**
  - **You may use other data analysis tools (R, python, excel…), but it may be difficult for me to provide help.**
  - Work on the tutorial problems by yourself with help of the TA   + Q&A on the textbook assignments (20 min)

- Tutorial Outcome:
  - Turn in a **PDF version** of your report **on Canvas by the next Monday**
  - 5 tutorials *1% each = 5%
  - No attendance checks
  - Textbook problems are not marked but we highly recommend you to do them!

Statistics MATLAB functions and where to find them

- GOOGLE
  - e.g., Matlab quartile
  - e.g., Matlab load excel

- Ask MATLAB :
  - **help** mean
  - **doc** mean

- Read the documentation

Use the **full** version of
MATLAB
(available to all UBC students)
https://it.ubc.ca/services/deskto
p-print-services/software-
licensing/matlab

Step 1: Data Checking

The first task when analyzing data is to check that all items are correct. Errors can occur by faulty measurements or recording. Obvious errors may be corrected if their cause is clearly understood and the required correction is definitely known. Grossly faulty data should be deleted. However, be careful, do not be heavy-handed in deleting imperfect measurements just because they are noisy or unexpected, else you will end up falsifying your data.

The in-class bolt length and diameter measurements were done under non-ideal conditions and contain much noise and several gross errors. The gross errors will need to be identified and either corrected or removed before the data analysis can proceed.

Step 1: Data Checking

- **Repair or Delete faulty data**

- Things to check:
  - o **Type error:** e.g., Bolt ID: expecting a letter, get a number "5.5" (Delete)
  - o **Data out of range:** e.g., get an "M" bolt ID (Delete)
  - o **Data logged in the wrong Unit:** a "7.4 mm" length → 74 mm (Repairable)
  - o **Measurement error:** Bolt length should be for shank only, but some measurements include the 10mm thick head (your decision)
  - o **Completely unphysical data:** A bolt with 999 mm length and 999 mm diameter!! (Delete)

Step 2: Box Plot

The lengths in mm of five of the bolts measured in the in-class exercise are:
A = [112.5 113 113 113.4 112 112.5 113.5]
D = [112 112.5 113.5 112.5 113 113 112.5]
E = [112 112 112.5 113 113.5 113.5 114.3]
G = [113 111 112 112.5 112.5 110 112.5]
H = [111.5 112 114 112 112.5 113 112.5]

Draw boxplots for the **individual bolts** and for **all the bolts**.

Compute **means**, **medians, quartiles, standard deviations**, and **variances** for the **individual bolts** and for **all the bolts**

MATLAB Functions

e.g., X=[0 1 4];

mu = **mean**(X);
sigma = **std**(X);
variance = **var**(X);

qtls = **quantile**(X,[.25 .50 .75]);

Mdn = **median**(X);

boxplot(X);

Important : every plot must have a title and labels. Unit of the label should be displayed, e.g., 'Bolt length, mm' !!

Summary: MATLAB Functions

help <some_MATLAB_function>
doc <some_MATLAB_function>

histogram(X)
histogram(X, nbins);

mu = **mean**(X);
sigma = **std**(X);
variance = **var**(X);
qtls = **quantile**(X,[.25 .50 .75]);
Mdn = **median**(X);

figure(n)
boxplot(X);
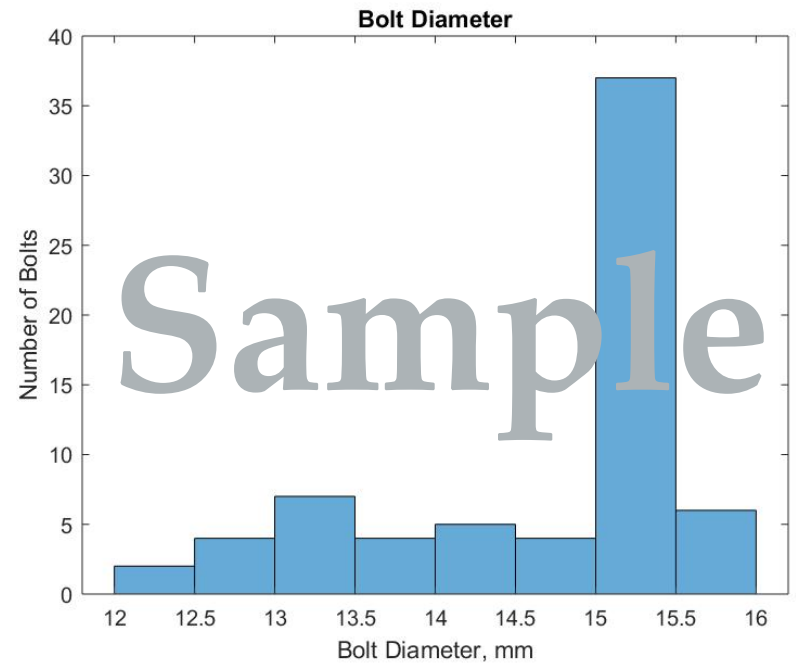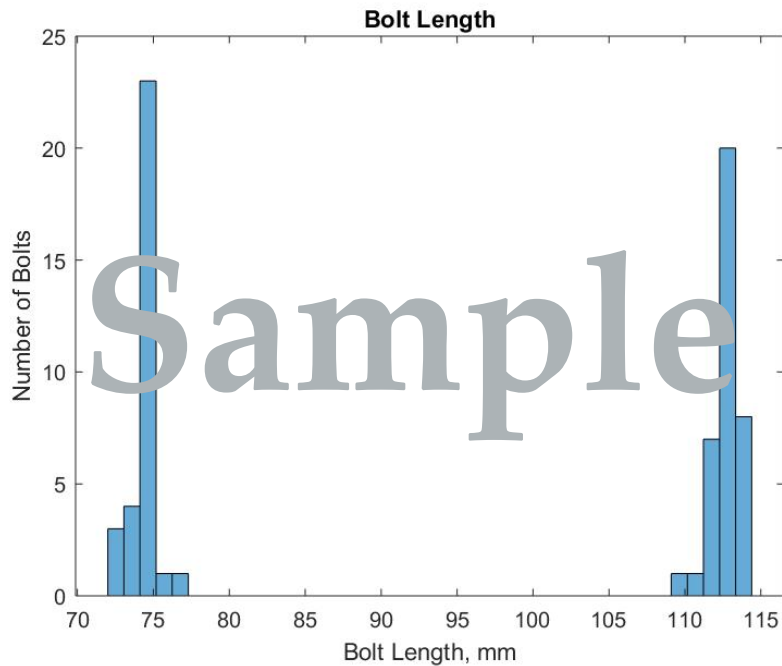**xlabel**('All bolts')
ylabel('Bolt Length, mm')
title('All Bolt Lengths')

Tutorial Report Requirements

1. **Download** Matlab **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 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. **Download** Matlab **Tutorial_1_Bolt_Data.m**. Draw 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 learn from the data analysis about either the sizes of the bolts or the measurement approach.

Histogram Solutions

Box Plot Solutions