ASEN2012 - Coding Challenge 1 (25 pts) - Due Wednesday 9/27 by 11:59 pm
Experimental and Computational Methods in Aerospace Engineering Sciences

---

Key Deliverables:

- 5 "Check your answer" questions in Canvas (**10 pts**)
  - **You only get 3 attempts** (correct answers carry over to next attempt) at the quiz.
  - **There is a 5 minute "cool-down" between attempts.** Use this time to debug and reevaluate your code. Once you compute a new answer, make sure to critically think about it before submitting again.
  - **No additional attempts will be provided for any reason.**
- 1 .zip file containing the following (naming convention lastname_firstname_CC1.zip) (**15pts**):
  - Functioning Matlab code (.m file(s) - if using multiple files, please name driver code main.m)
  - Published code (.pdf file)
  - Flowchart sketch (.pdf file)

Background

We have began to study statistical concepts such as mean, standard deviation, probability density function (PDF), and standard error of the mean. In this coding challenge you will analyze a single data set with N samples, that does not contain any outliers, and quantify some of these values. Recall that for a measurement (x), the mean ($\overline{x}$) is defined by Eqn. 1, the standard deviation of x (σ_x) is defined by Eqn. 2, and the standard error of the mean ($\sigma_{\overline{x}}$) is defined by Eqn. 3. For a normal distribution, the Gaussian (PDF) is appropriate for comparison which is shown in Eqn. 4.

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{2}$$

$$\sigma_{\overline{x}} = \frac{\sigma_x}{\sqrt{N}} \tag{3}$$

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\overline{x}}{\sigma_x}\right)^2\right] \tag{4}$$

Note that x_i is not part of $f(x)$ since the PDF does not rely on being calculated at individual data points but can instead be calculated for all valid values of x.

The Data

We are going on a hike! The data provided are times, in minutes, for the Sanitas Climb (only going up), publicly available by Strava (`https://www.strava.com/segments/633433`). A topological map, elevation map and various distances are provided in Fig. 1. Our goal is to understand what we might expect, in terms of time, for our first hike together.

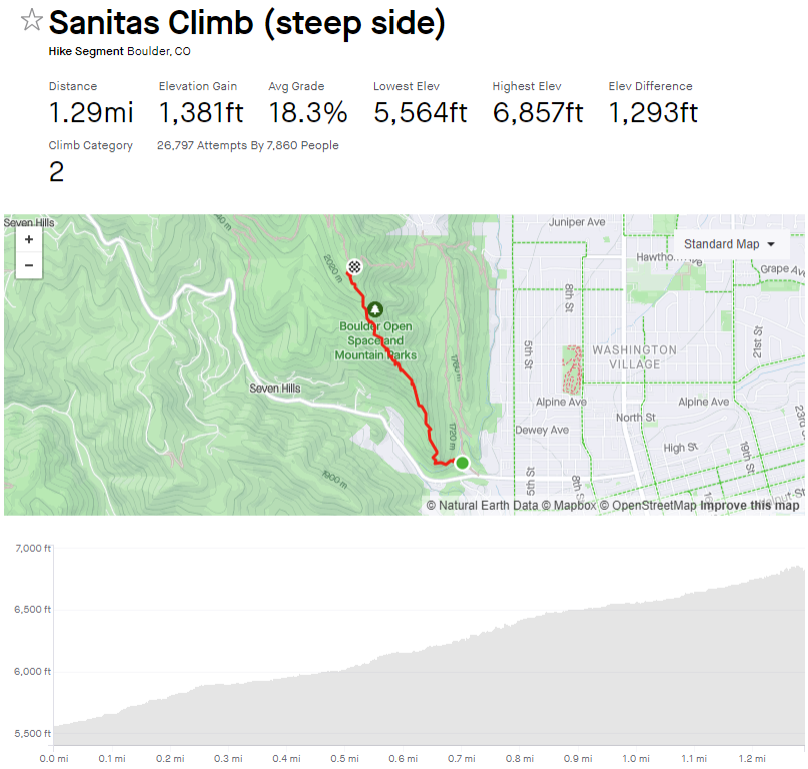

Figure 1: Informational graphic from Strava of the Sanitas Climb.

Step 1: Read in and clean the data set

Download the data set from Coding Challenge 1 Assignment in Canvas, named **sanitas_times.csv**. Use any Matlab function of your choosing to read in the data set (you may **not** use the Import Data Tool). As previewed in class, clean the data set of any missing (or NaN) values and create a histogram plot to take note of any trends.

In Canvas quiz, enter the number of elements in the data set.

Step 2: Calculate statistical quantities 3 different ways

Calculate the mean and standard deviation using 3 methods and time how long each method takes:

1. `for` loop
2. `sum()` function
3. `mean()` and `std()` functions

In order to time how long Matlab takes to run snippets of code, use the `tic` and `toc` commands around code. For example,

```
tic
i=1;
while i < 10
    i=i+1;
end
% toc % just typing toc will print the time to the command window
whileRunTime = toc; % setting a variable to toc will save this for
    comparison later
```

In Canvas quiz, enter the mean and standard deviation; and answer if all 3 methods were equal for both quantities. (3 separate questions)
Note that run times are hardware dependent and don't always scale as expected (take a look at the 3 times).

Step 3: Calculate the SEM

Compute the standard error in the mean using any method you'd like.
In Canvas quiz, enter the SEM.

Step 4: Compare data with the PDF

Create a vector using `linspace()` that spans the entire data range with sufficient number of elements, and compute the PDF given the mean and standard deviation computed earlier. Compare the histogram and PDF on one plot (you may find that the `histogram` function options are helpful). Include lines indicating the mean (solid black line) and ± 1 standard deviation from the mean (using dashed black lines). Be sure to include axis labels and a legend.

Reflection Questions

Does the distribution of times look to be normally distributed? Does the Gaussian (perfectly normal distribution) visually compare well with the provided data? Why do you think that is or is not the case?
Please write out the answers to these questions in the comments of your Matlab script at the end of your .m file. This should be about 1 paragraph in length.

Coding Rubric

	Excellent (100%)	Above Average (80%)	Average (70%)	Below Average (50%)
Requirements and Delivery (2pts)	• Completed 90-100% of the requirements • Delivered on time, and in correct format.	• Completed 80-90% of the requirements • Delivered on time, and in correct format.	• Completed 70-80% of the requirements • Delivered on time, and in correct format.	• Completed <70% of the requirements • Delivered on time, but not in correct format.
Coding Standards (2pts)	• Includes full header* • Excellent use of variables (no global or unambiguous variable naming)	• Includes full header* • Good use of variables (1-2 global or ambiguous variable naming)	• Includes incomp. header* • Fine use of variables (3-5 global or ambiguous variable naming)	• No header • Poor use of variables (many global or ambiguous variable naming)
Documentation (2pts) Comment your code	• Clearly documented • Specific purpose noted for each function and/or section	• Well documented • Specific purpose noted for each function and/or section	• Some documentation • Purpose noted for each function and/or section	• Limited to no documentation
Runtime (1pt)	• Executes quickly, without errors	• Executes without errors, over 1 min runtime	• Executes with warnings/errors	• Does not execute
Efficiency (2pts)	• Easy to understand, and maintain	• Logical, without sacrificing readability and understanding	• Difficult to follow	• Difficult to follow, huge and appears patched together
Figure Quality (2pts)	• Easy to understand, labels and legend present	• Easy to understand, lacks labels and legend	• Difficult to understand, labels and legend present	• Difficult to understand, lacks labels and legend
Reflection Questions (4pts)	• Easy to understand, fully thought out	• Easy to understand, mostly thought out	• Hard to understand, effort made	• Hard to understand, lacking effort

* Header includes author name(s), assignment title, purpose, creation date, revisions (applicable to **group** projects only)