Problem Set 0

EN 600.438/638

January 30, 2019

Due date: February 4, 2019 by midnight
Submission: Please sign up and submit your assignments on https://www.gradescope.com. The code for this course is 9PYWKP. Please read the instructions on the latex template about how to write up the answer file. Problem sets of written questions should be submitted in PDF.
Reminders:
For this problem set only, please work independently: This problem set is intended to test your knowledge of the prerequisite material for the course and you should use it as a tool to honestly assess your own background. Therefore, we ask that you do not discuss any of the problems with others and conduct all of your work independently. You may not copy any of your work or code from others, including but not limited to any resources you may find on the Internet.
Late days: There is no late day for this assignment. You have 5 total late days for the semester. Days will be rounded up; for example, if you submit your assignment 3 hours late, you will use 1 full late day. If you submit your assignment 25 hours after the deadline, you will use 2 full late days.
Programming Language: We will be using Python 3.6.5. We are *not* using Python 2, or other programming languages, and will not accept assignments written in other languages. We recommend using a recent release of Python 3.6.x, but anything in this line (e.g. 3.x) should be fine. For each assignment, we will tell you which Python libraries you may use. We *strongly* recommend using a *virtual environment* to ensure compliance with the permitted libraries.
Virtual Environments Virtual environments are easy to set up and let you work within an isolated Python environment. In short, you can create a directory that corresponds to a specific Python version with specific packages, and once you activate that environment, you are shielded from the various Python / package versions that may already reside elsewhere on your system. Here is an example:

```
# Create a new virtual environment.
python3 -m venv python3-hw1
# Activate the virtual environment.
source python3-hw1/bin/activate
# Install packages as specified in requirements.txt.
pip3 install -r requirements.txt
# Optional: Deactivate the virtual environment, returning to your system's setup.
deactivate
```

Exercise 1. Expected value and likelihood (1 point)
You are given a coin which is loaded so that probability of heads is $\theta = 0.6$. You toss the coin 30 times.

1. What is the probability that all 30 tosses yield heads? (You should round the answer to the third decimal digit).
2. If we assign a value of 0 to heads, and 1 to tails, what is the expected value of a single coin toss?
3. What is the expected value of summing the values of 4 coin tosses?
4. For a new coin, you are given the following recorded sequence of coin tosses:

   $$H, H, H, T, T, H, H, T, H, H$$

   Assume a model where every toss is independent with $p_H = \theta$, but we don't know θ yet.

   (a) What is the likelihood of this data according to our model if $\theta = 0.5$?
   (b) What is the likelihood of this data according to our model if $\theta = 0.8$?
   (c) Which model do you prefer? Why?
   (d) Is there an even better setting of θ?

Exercise 2. Probability and independence (0.5 points)
Suppose you are randomly picking balls from a black box. You pick one ball at a time, record two features of the ball - color and texture, and put it back. You record the feature of each ball during the sampling process, and fill in the table below:

	Steel	Wooden
Red		
White		

1. Suppose A is the event of getting a red ball, and B is the event of getting a wooden ball. If we know $P(A|B) > P(A)$. Prove it implies $P(B|A) > P(B)$.
2. After 25 runs, you get the table below.

	Steel	Wooden
Red	5	8
White	10	2

   (a) From the data, are A and B independent? and Why?
   (b) What is the probability of getting a steel ball if you already know that the ball is red.
   (c) Given that you get a red ball, would it be more likely to be steel or wooden?

Exercise 3. Bayes' Theorem (0.5 points)
There is a blood test for a rare disease that affects 1 of every $250,000$ people in the population. If a patient has the disease, the test will come up positive (correctly) with probability 0.96. If a patient does not have the disease, the test will (incorrectly) come up positive with probability 0.005. If you take this test and it comes up positive, what is the probability you actually have the disease?

Exercise 4. Gaussian data and likelihood (1.5 points)
The standard normal distribution refers to a Gaussian distribution with mean 0, variance 1. Generate 10 values sampled from a standard normal distribution using the numpy.random.normal module in python. We refer to your data as $x_1, x_2...x_{10}$.

1. What is the expected value of the sum of ten numbers sampled from a standard normal distribution?
2. What is the expected value of the sum of the square of ten numbers sampled from a standard normal distribution $E[\sum_{i=1}^{10} x_i^2]$?
3. What were your actual ten values (3 signif digits) here, their sum, and sum of squares?
4. Write out the probablity for the first three values: $p(x_1), p(x_2), p(x_3)$.
5. What is the value of the full likelihood $l(\mu, \sigma^2|X) = p(x_1, x_2, ...x_{10})$ assuming $\mu = 0$, $\sigma^2 = 1$?
6. What is the **log** likelihood of your data, and why might people prefer to work in log space rather than raw space?

Exercise 5. Linear algebra. (.5 points)

1. If A is an orthonormal matrix, what is $A^T A$?
2. Compute the inverse of the following matrices if you can. If not, please justify why.

   (a)
   $$A = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

   (b)
   $$A = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & 0 \\ -6 & 2 & 2 \end{bmatrix}$$

3. Write a function called **inverseMatrix.py** that can take in a matrix of an arbitrary size. It should be able to decide whether a matrix is invertible. If so, return the inverse of the matrix. If not, print out why. You can (and only can) use all the functions from **numpy.linalg** module. Submit the *.py* file to the programming assignment in gradescope.

Exercise 6. Matrix derivatives. (1 point)

1. Let $Y \in \mathbb{R}^{N\text{x}1}$, $X \in \mathbb{R}^{N\text{x}K}$, $\beta \in \mathbb{R}^{K\text{x}1}$, and $N \geq K$
   Let $Q = (Y - X\beta)^T(Y - X\beta)$. This is the standard objective function that is minimized in linear regression. Compute the derivative of Q with respect to β

2. Find the β that minimizes Q (ie. find the value of β that makes $\frac{\partial Q}{\partial \beta} = 0$)

3. Why are we only able to get a unique solution for β iff $N \geq K$?