Extra Bonus Assignment

STAT 4705, Fall 2017

Option 1 (max 1 pt), due Dec 13

Analyze the famous (Fisher's or Anderson's) iris data, which is uploaded on scholar with the name "irisdata.csv". This data set gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of 3 species of iris. The species are Iris setosa, versicolor, and virginica.

- (a) Report the descriptive statistics for the sepal width of the whole 150 iris data. By descriptive statistics, I mean things like sample mean, sample variance, sample median, Q1, Q3 etc.
- (b) Draw histogram, boxplot for the sepal width of the whole 150 iris data. Comment on each plot.
- (c) Check the normality of sepal width. That is, draw the normal Q-Q plot for the sepal width. Comment on the plot.
- (d) Software option: R or Python

Option 2 (max 3 pts), signup due Dec 9

- group study (1 - 3 people)
- pick and read one chapter (1 - 22) of the book: E.T. Jaynes, Probability Theory, The Logic of Science
- 10 to 15 mins presentations in class
- presentation time: probably 12/11 Monday (depending the status of signup), email notice
- book link: http://www.med.mcgill.ca/epidemiology/hanley/bios601/GaussianModel/JaynesProbabilityTheory.pdf
- sign up link: https://docs.google.com/a/vt.edu/spreadsheets/d/1Wj5DzjD2nUUJT5PLsd1TUBAWfSzRYZ_SP_Mg9G_BsOc/edit?usp=sharing