Homework 5

Stat 345 - Spring 2020

Name: _____

Problem 1

X and Y are jointly continuous with joint pdf

$$f(x,y) = 2, \quad x > 0, y > 0, x + y \leq 1$$

and 0 otherwise.

a) Find marginal pdf's of X and of Y.

b) Find covariance Cov(X,Y).

c) Find correlation Corr(X,Y). What you can say about the relationship between X and Y?

Problem 2

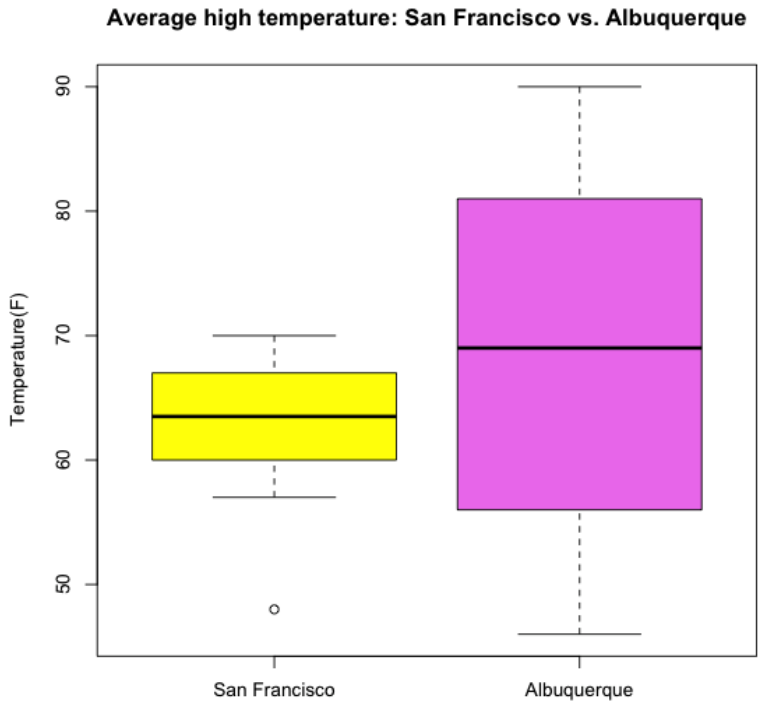
The number, Y, of spam messages sent to a server in a day has a Poisson distribution with parameter $\lambda = 24$. Each spam message independently has a probability $p = 0.25$ of not being detected by the spam filter. Let X denote the number of spam massages getting through the filter.

a) How are you going to find the distribution of X?

b) Calculate the expected daily number of spam messages which get into the server through the filter.

Problem 3

These boxplots compare the average high temperatures of San Francisco (x) to those in Albuquerque (y) during one year.



a) Which of the following statements are true? Select all.

- Half of SF temperatures are below 70 F
- The ABQ IQR is around 23 F
- About 10% of temperatures are below 65 F
- A quarter of SF temperatures are above 67 F
- Half of ABQ temperatures are below 69 F
- A quarter of ABQ temperatures are above 55 F
- A quarter of ABQ temperatures are below 69 F
- Mean and median of ABQ temperatures are very close to each other
- The spread of SF temperatures is greater than the spread of ABQ temperatures
- The median of SF temperatures is greater than the median of ABQ temperatures
- The third quartile of SF temperatures is less than the first quartile of ABQ temperatures
- The first quartile of SF temperatures is less than the third quartile of ABQ temperatures

a) Can you identify any outliers from the plot? Identify quartiles from the plot and use the $1.5IQR$ rule to check suspected values.

Problem 4

Online R: https://rdrr.io/snippets/. If you will use R attach your histograms at the end of this homework.

The scores of two exams are given:

Exam 1: 42, 41, 76, 48, 48, 59, 58, 49, 56, 88, 83, 49, 78, 47, 61, 64, 54, 76, 91, 63, 51, 95, 64, 88, 47

Exam 2: 58, 100, 50, 62, 68, 67, 69, 78, 79, 72, 75, 83, 81, 75, 89, 87, 91, 97, 59

a) Make a histogram (by hand or use *hist(x)* in R) of the test scores on Exam 1 and Exam 2.

b) Calculate mean and median for both exams. You can calculate this by hand or use R. For Exam 2 scores, also calculate three quartiles by hand.

c) Was one test more reasonable than the other? Which measure of center is better to use for each exam?