

Name _____

Lab Assignment 2

(40 points)

DIRECTIONS

- *An uploaded or hard copy is due at the start of next Thursday class by 11:00 AM to Bioinformatics 104.*
- *In a typed word processing document, or on this sheet clearly answer the following questions.*
- *Be sure to show all of your work by providing your R commands, solutions/output, and interpretation.*
- *You may work with a partner to complete this assignment, but you each must hand-in your own independent work.*
- *Please circle or highlight your final answers.*

A. Bayes Theorem (20 points)

1) 3% of people have a certain genetic defect. 85% of tests for the gene detect the defect (true positives). 7.8% of the tests are false positives. If a person gets a positive test result, what are the odds they actually have the genetic defect (10 points).

$P(A)$ _____

$P(B)$ _____

$P(B|A)$ _____

Final Probability _____

2) Given the following statistics, what is the probability that a woman over 45 has cancer if she has a positive mammogram result (10 points).

Three percent of women over 45 have breast cancer.

Ninety-Two percent of women who have breast cancer test positive on mammograms.

Ten percent of women will have false positives.

$P(A)$ _____

$P(B)$ _____

$P(B|A)$ _____

Final Probability _____

B. Descriptive statistics (10 points)

1) Give an example of a Population and Sample. (2 pts)

Population: _____

Sample: _____

2) Fill in the blank. A sample is considered random if every different sample has

_____ probability to be selected. (2 pts)

3) What is the variable for population parameter for mean. What is the variable for sample statistic for mean? (2 pts)

Population: _____ Sample: _____

4) Give an example of a stratified random sample of a population. (2 pts)

Example: _____

5) What is the K^{th} element for a population of 30 and a systematic sample size of 6? (2 pts)

$K =$ _____

C. Draw a Box Plot in R (10 points)

Using the “area” data from the “rock” dataset within in R, draw a boxplot in R, print out the boxplot. Be sure to add correct axis labels and a title using R commands. List the R commands in the following lines.

Using R determine the median, IQR, first quartile, third quartile, list both the command and the value (10 points).
