

1. True/False Problems. Write *very brief explanations*.

(a) Circle either TRUE or FALSE (and explain/correct if FALSE):

A variable consists of the numbers 1, 2, 3, 4, 5. If the number 5 is replaced by the value 20, the standard deviation will increase.

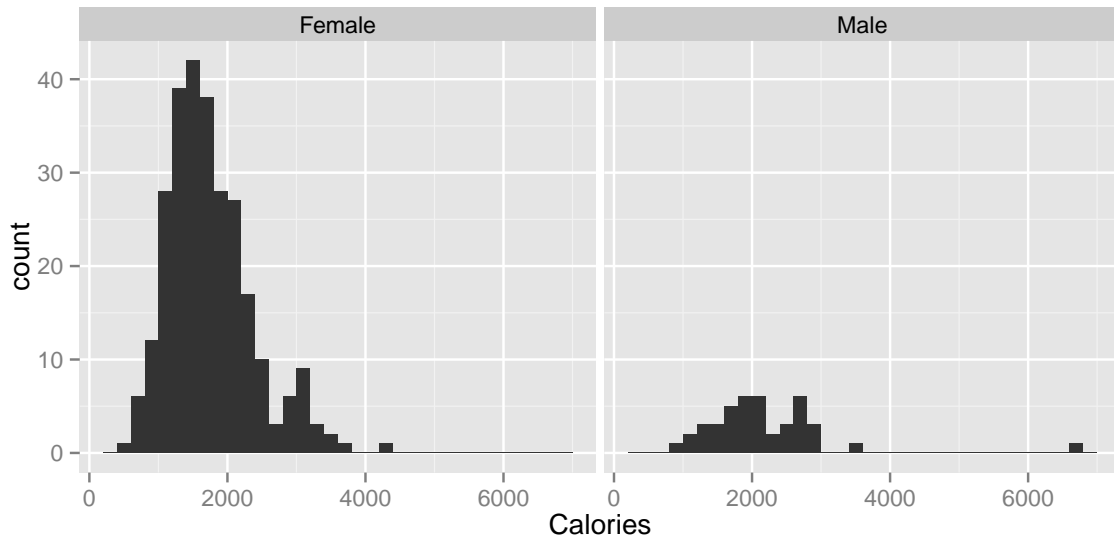
(b) Circle either TRUE or FALSE (and explain/correct if FALSE):

A 95% confidence interval for the mean birth weight of healthy babies in the US is from 7 pounds 7 ounces to 7 pounds 9 ounces. This means that 95% of all healthy babies born in the US have weights between these values.

(c) Circle either TRUE or FALSE (and explain/correct if FALSE):

When studying the association between gender (male or female) versus height, a researcher computes a correlation coefficient of  $r = 0.6$ . This is a reasonable quantitative measurement of this association.

2. A study on nutrition included a sample of 315 subjects, each of whom during a three-year period had elective surgery to remove a lesion of the lung, colon, breast, skin, ovary, or uterus that was found to be non-cancerous. The study includes many variables. Here is a graph that shows histograms of number of calories consumed per day for the men and women in the study.



- About what proportion of the men consume 1500 calories or fewer per day?
  - Guess the mean and standard deviation for calories per day for each sex.
  - Is the mean or median larger in each distribution? Why?
  - A researcher is interested in estimating the population mean number of calories consumed per day from the population of women from which the sampled women may be considered as a representative sample with a 95% confidence interval using the bootstrap standard error method. Describe the steps.
    - Find the sample mean.
    - Create a large number of samples of size (?) by sampling (?) from (?).
    - From each of these samples, calculate (?). These values are called the bootstrap statistics.
    - Compute the (?) of the bootstrap statistics.
    - Find the limits of the 95% confidence interval by adding and subtracting (?) to the (?).
3. Pictured below is the graph of a sampling distribution of the difference in sample means for male minus female daily calorie consumption.



- (a) Define parameters and state null and alternative hypotheses to test if males consume more calories per day than females in the population from which the data is sampled.
- (b) The difference in daily calories consumed (male minus female) from the sample is 415. What is an estimate of the p-value for this hypothesis test?
- (c) Write a summary of the conclusion in the context of the problem.