


Introduction to Biostatistics (171:161)
Breheny

Assignment 7
Due: Tuesday, March 25

1. We saw in class that the heights of adult women in the United States are roughly normally distributed with a mean of 63.5 inches and a standard deviation of 2.75 inches. What is the probability that a randomly selected adult woman in the United States will be exactly 66 inches tall?
2.  This question is based on the NHANES sample of adult males available on the course website. For each of the following, **provide two answers**: the answer based on the normal approximation and the answer calculated directly from the data.
 - (a) What percent of men weigh between 150 and 200 pounds?
 - (b) What is the 25th percentile of weight?
 - (c) How many men in the sample weigh above 300 pounds?
 - (d) These approximations do not seem to be as good as those we saw in class. Why?
3. A study at Boston University found that for men who will develop coronary artery disease, cholesterol levels are normally distributed with a mean of 244 mg/dl and a standard deviation of 51 mg/dl. They also found that for men who do not develop the disease, cholesterol levels are normally distributed with a mean of 219 mg/dl and standard deviation 41 mg/dl. Consider the following “test” for coronary artery disease: if a man has cholesterol level above 240, we predict that he will develop coronary artery disease (*i.e.*, he tests positive).
 - (a) What is the sensitivity of this screening tool?
 - (b) What is the specificity of this screening tool?
 - (c) If ten individuals who will not develop coronary artery disease take this test, what is the probability of obtaining at least 2 false positives?
4.
 - (a) A coin is tossed several times. You win a dollar if the number of heads is equal to the number of tails. Which is better for you: 10 tosses or 100 tosses?
 - (b) A coin is tossed several times. You win a dollar if the percentage of heads is between 45% and 55%. Which is better for you: 10 tosses or 100 tosses?
5. A hundred draws will be made at random with replacement from one of the following urns, each of which contain two numbered balls. Urn A contains a 1 and a 9. Urn B contains a 4 and a 6. The value of each draw will be recorded and the mean of the 100 draws calculated. Someone offers you a dollar if you can predict the value of the sample mean to within 0.1. Which urn offers you the best chance of winning the dollar?
6. According to the Massachusetts Department of Health, 224 women who gave birth in the state of Massachusetts in 1988 were HIV positive. At the time, a child born to an HIV positive woman had about a 25% chance of also being HIV positive. Fill in the blanks: the number of HIV positive babies born in Massachusetts in 1988 was around _____, give or take _____ or so (*i.e.* mean \pm standard deviation).

7. A researcher wants her sample mean to be twice as accurate; how much does she have to increase her sample size by?
8. An article in the *New England Journal of Medicine* reported that among adults living in the United States, the average level of albumin in cerebrospinal fluid is 29.5 mg/dl, with a standard deviation of 9.25 mg/dl. We are going to select a sample of size 20 from this population.
 - (a) How does the variability of our sample mean compare with the variability of albumin levels in the population?
 - (b) What is the probability that our sample mean will be greater than 33 mg/dl?
 - (c) What is the probability that our sample mean will lie between 29 and 31 mg/dl?
 - (d) What two values will contain the middle 50% of our sample means?