Applied Statistics HW 17

1. Let us suppose that this past year, the average SAT score nationwide was 1400, while the standard deviation of the SAT scores was 100. We pick at random 50 students around Indiana, and ask them for their SAT scores, then compute the average of those values.

   a. What can we say about the mean and standard deviation of this sample average, thinking of it as a random variable ranging over all the possible samples? What assumptions did you have to make? Does the fact that the students we sampled are from Indiana affect the answer?

   b. Can we assume that this sample average follows a normal distribution? Explain what you had to assume along the way.

c. Suppose we didn't know the national average of 1400, and the sample average turned out to be 1359. Provide 90%, 95% and 99% confidence intervals for the national average μ.

d. The figure of 1400 was provided by the ETS, but we don't trust it. Suppose we want to test that claim of the ETS against the sample average of 1359 we got. What are the chances, if that claim of 1400 national average was true, of us seeing a sample average as far away from 1400 as our 1359? Is this strong evidence against the 1400 claim?

e. Continuing from the previous part, what if the claim was that the average SAT score is at least 1400? How does that change the computation and conclusion?

f. Continuing from the previous part, how would things have changed if our sample was consisting of only 30 students instead? What if it consisted of 200 students? Make sure to consider how the different number of students affects the various steps in the process.

2. In a scientific experiment, a scientist will run his experiment multiple times. He is measuring the weight of a certain particle. From previous experiments he knows that the samples he gets have a standard deviation of about 10micrograms, so he will assume that this will be the population standard deviation.

   a. Explain how this setup fits into the sample mean setting.

   b. Suppose the scientist runs the experiment 35 times. What would the margin of error be for a 99% confidence interval?

   c. The scientist wants to have a margin of error no more than 0.1micrograms. How many times does he need to run his experiment?

3. You are invited to play a game of chance. There is a large pool containing very many little pieces of paper. Each of these pieces has written on it the amount you win or lose in this round. For instance if the paper has the number 2 on it you win two dollars, while if it has the number -3 on it you lose three dollars. We don't know much about the numbers on the pieces of paper, except that their average value is 1, and their standard deviation is 1. For each round of this game, we pull one of these papers out. We will end up playing 60 rounds of this game. We are interested in our average earnings per round, which we'll denote by $\bar{x}$.

   a. Does the situation we are discussing fit into the sample mean setting, with $\bar{x}$, μ etc? Explain, and compute the parameters.

   b. Can we assume that the average earnings per round, $\bar{x}$, follows a normal distri- bution? Explain. Regardless of your answer, assume that it does follow a normal distribution for the next part.

c. What are the chances, that our average earnings per round would be greater than or equal to 0? Do you think this is a good game for you to play?

4. We would like to try to prove a theory that in heterosexual couples the man tends to always be taller than the woman. To that effect, we select at random 234 couples across the US, and measure the difference in the heights. We found those differences in the sample to have a mean of 5 inches and a standard deviation of 25 inches.

   a. Construct a hypothesis test to test our hypothesis above, and carry the test out. Do you think there is enough evidence to conclude that our theory is correct?

   b. Construct 90% and 95% confidence intervals for the average height difference.