Statistics 571 Midterm 2

Hanlon/Larget, Fall 2010

Name: _____

Please circle the lecture section *in which you are registered*: Hanlon Larget

Please circle the lecture section *you attend*: Hanlon Larget

---

Instructions:

1. You may use a calculator, but you may not use a laptop computer or phone.
2. The examination is open book, open notes, but not open neighbor. You may use any course handouts including lecture notes and homework solutions.
3. Do all of your work in the space provided. Use the backs of pages if necessary, indicating clearly that you have done so (so the grader can easily find your complete answer).

For Graders' Use:

Question	Possible Score	Score
1	20	
2	20	
3	20	
4	25	
5	15	
Total	100	

1. (20 points) As part of a study of the treatment of anemia in cattle, researchers measured the concentration of selenium in the blood of 36 cows who had been given a dietary supplement of selenium (2 mg/day) for one year. The cows were all the same breed (*Santa Gertrudis*) and all had borne their first calf during the year. Various graphs of the sampled data show a fairly symmetric distribution with no extreme outliers. In this study, the sample mean selenium concentration was 6.21 μg/dLi and the standard deviation was 1.84 μg/dLi.

   (a) Conduct a two-sided t-test of $H_0 : \mu = 6$. State hypotheses, calculate a test statistic, declare the sampling distribution, and use the table to give a range for the p-value.

   (b) Interpret this test within the context of the problem

   (c) The researchers want to conduct a second experiment, using the data from the first experiment as a pilot study. In this case, they want to compute a 95% confidence interval for the mean selenium concentration. How many cows should be used in the new study so that the margin of error for this confidence interval is less than 0.25 μg/dLi?

2. (20 points) Certain types of nerve cells have the ability to regenerate a part of the cell that has been amputated. In an experiment, measurements were made on the nerves in the spinal cord of eight rhesus monkeys. Nerves emanating from the left side of the cord were cut, while nerves from the right side were kept intact. During the regeneration process, the content of creatine phosphate (CP) was measured in the left and the right portion of the spinal cord. The following table provides summary statistics from the experiment. The units of measurement are mg CP per 100 g tissue.

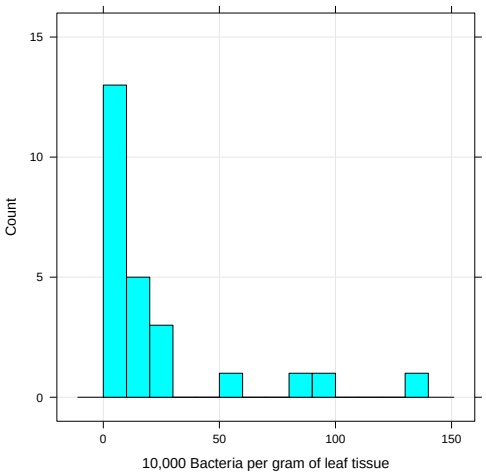
	Right	Left	Difference
n	8	8	
mean	15.50	10.86	
SD	7.61	4.49	4.89

(a) Construct a 90% confidence interval for the difference in mean CP between the right and the left sides.

(b) Interpret this confidence within the context of the problem.

3. (25 points)

The surfaces of leaves on corn plants harbor large populations of bacteria, some of which can affect the health of the plant. A researcher wishes to estimate the abundance of bacteria per gram of leaf tissue in a corn field. The researcher collected leaf samples from 25 separate plants and measured the number of bacteria from each sample. A histogram of the bacteria densities (in units 10^4 individuals per gram of leaf tissue) is shown at the right—a value of 10 corresponds to 100,000 bacteria per gram of leaf tissue. The following table summarizes the sample data both before and after a natural log transformation.

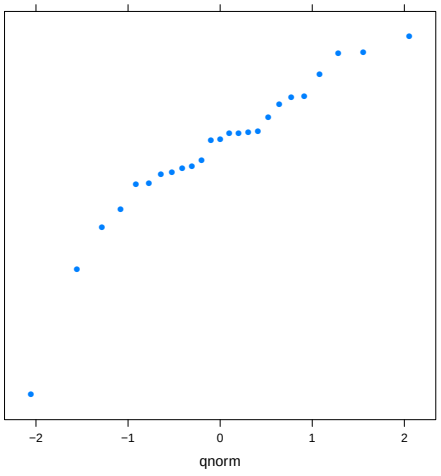


	Mean	SD
Original Data	22.4	34.8
Log-transformed Data	1.94	1.98

(a) Use an appropriate method to find a 99% confidence interval for the mean number of bacteria per gram of leaf tissue, expressing the interval in the units of the problem.

(b) Interpret this interval in the context of the problem.

(c) The second plot to the right shows a plot of the sorted data versus quantiles of a standard normal distribution. Does the plot show the original or log-transfomed data? Briefly explain.



(d) A statistical consultant suggested a different way to calculate a confidence interval. She used the computer to sample at random 25 observations with replacement from the original untransformed sample, computed and recorded the mean, and repeated this 10,000 times. Briefly explain how she can calculate a 99% confidence interval from these 10,000 numbers.

4. (20 points)

Consider the following sets of hypotheses, significance levels, and sample sizes for inference about μ, the mean of a population where it is assumed that $\sigma = 20$.

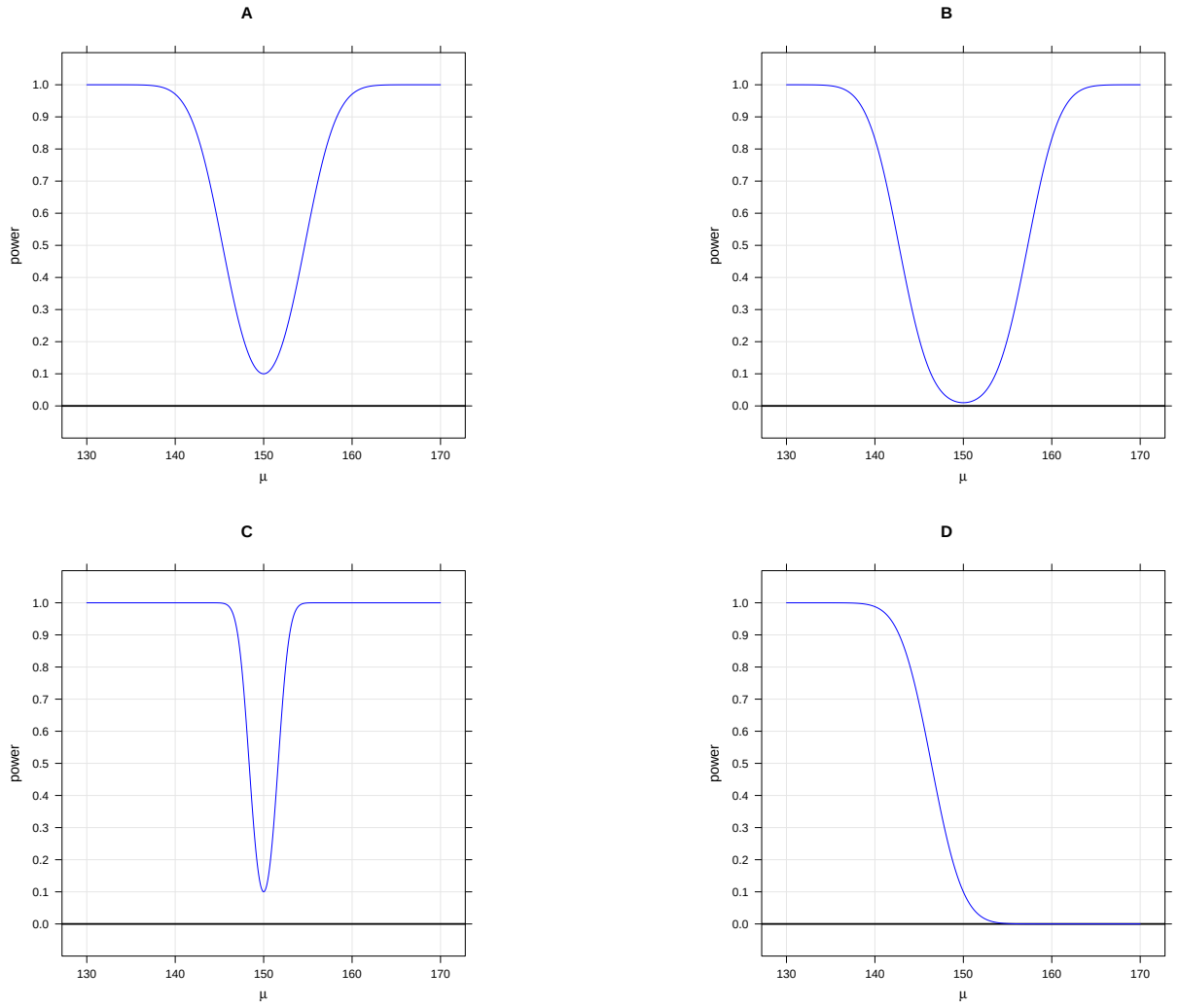
Set 1. $H_0: \mu = \mu_0$ versus $H_A: \mu < \mu_0$, $\alpha = 0.1$, $n = 50$.

Set 2. $H_0: \mu = \mu_0$ versus $H_A: \mu \neq \mu_0$, $\alpha = 0.01$, $n = 50$.

Set 3. $H_0: \mu = \mu_0$ versus $H_A: \mu \neq \mu_0$, $\alpha = 0.1$, $n = 50$.

Set 4. $H_0: \mu = \mu_0$ versus $H_A: \mu \neq \mu_0$, $\alpha = 0.1$, $n = 400$.

The power curves for these four sets of hypotheses and conditions are graphed below.



(a) What is the value of μ_0?

(b) Match each set (1, 2, 3, 4) with a power curve (A, B, C, D).

Set 1 ☐ Set 2 ☐ Set 3 ☐ Set 4 ☐

(c) For null hypothesis $H_0: \mu = \mu_0$, alternative $H_A: \mu \neq \mu_0$ (where μ_0 is the numerical value you found in part (a)), $\alpha = 0.05$, and $n = 50$, calculate the power of the test when $\mu = 145$.

5. (15 points) The following parts can be answered independently. Each question can be answered with one or two sentences that indicates comprehension of a single key concept.

   (a) A 1976 research paper reported that terminal cancer patients given vitamin C survived much longer than did historical controls. In the study, the patients treated with vitamin C were selected by surgeons from a group of cancer patients in the hospital. Briefly discuss the major flaws in this study design.

   (b) Shortly before Valentine's Day in 1999 a newspaper article was printed with the headline "Marriage makes for healthier, longer life, studies show." The headline was based on studies that showed that married people live longer and have lower rates of cancer, heart disease, and stroke than do those who never marry. Briefly discuss the validity of this headline.

   (c) For each of the following cases, state whether the study should be run blind, double-blind, or neither. If the study should be run blind or double-blind, who should be blinded.

      i. An investigation of whether taking aspirin reduces one's chances of having a heart attack.

      ii. An investigation of whether the size of the midsagittal plane of the anterior commissure (a part of the brain) of a man is related to sexual orientation of the man.