PID: ____________________

Last Name, First Name: ____________________

Section: ____________________

Homework 5
Math 183, UCSD, Fall 2017
Due on Friday 17th, November 12:50pm
Staple pages together

Exercise 1

In this exercise, T_n denotes a random variable having T-distribution with n degrees of freedom. Using the tables of the textbook only, find the best bounds you can for:

1. $P(T_{12} < 2.02)$ (give an answer of the form $a \leq P(T_{12} < 2.02) \leq b$).

2. $P(0 \leq T_{32} < 2.3)$ (answer of the form $a' \leq P(0 \leq T_{32} < 2.3) \leq b'$).

3. the value t^* such that $P(-t^* \leq T_5 < t^*) = 0.017$ (answer of the form $a'' \leq t^* \leq b''$).

Exercise 2

Coffee packages indicate a content of 500g. A fraud prevention inspector wants to check that the packages do contain at least 500g on average. Otherwise, he will prosecute the coffee brand. He runs a t-test with significance level $\alpha = 1\%$.

1. Define the hypotheses H_0 and H_A he takes.

2. (a) Describe what type I error is in general, and what it would mean in this case.

   (b) Assuming that the coffee packages do actually weigh 500g on average, what is the probability that the inspector rejects H_0?

3. On 41 packages the inspector sampled randomly, he got a sample mean of 495g and a standard deviation of 20g. What is his conclusion? Explain.

Exercise 3

A Chinese group of ornithologists studies the duck population of the Yellow River. Every season, they capture ducks, measure a certain number of parameters, and then eat all the ducks they captured to celebrate the end of the study. The following table gives the weights (in kg) of the captured ducks of two consecutive years (called 1 and 2).

Year 1	2.22	2.16	2.49	2.11	2.17	2.13	2.06	2.21	2.20	2.47	2.20	2.08
Year 2	2.52	2.31	2.11	2.73	2.32	2.38	2.30	2.41	2.40	2.15		

Duck survey (Yellow River)

On the other hand, an Italian group of ornithologists studies the goose population of the Po River. Every season, they capture geese, measure a certain number of parameters, and then set the geese free after marking them with an identification ring. From one year to another, they catch again the same birds they caught previously. The following table gives the weights (in kg) of the same geese that have been captured in two consecutive years (called I and II).

Year I	4.78	4.85	4.09	5.00	4.90	4.86	5.37	4.68	5.28	5.11
Year II	4.99	5.05	4.53	4.97	5.48	5.03	5.01	4.78	4.99	5.07

Goose survey (Po River)

1. Explain briefly why the statistical models of these two studies are deeply different.

2. **(Ducks of the Yellow River)** With a confidence level $\alpha = 5\%$, decide whether the mean weight of the ducks of the Yellow River has changed between Year 1 and Year 2 or not.

Make sure to: 1) state hypotheses, 2) draw a picture of a sampling distribution, 3) shade a p-value, 4) find the p-value, and 5) state your decision based on the p-value.

3. List the assumptions you made to answer question 2.

4. **(Geese of the Po River)** With a confidence level $\alpha = 5\%$, decide whether the mean weight of the geese of the Po River has changed between Year I and Year II or not.

Make sure to: 1) state hypotheses, 2) draw a picture of a sampling distribution, 3) shade a p-value, 4) find the p-value, and 5) state your decision based on the p-value.

5. List the assumptions you made to answer question 4.

Exercise 4

Do fancy math equations make a research paper seem higher quality, even if those equations are total nonsense? In 2012, one researcher decided to figure this out by taking a published sociology paper and clipping the results from it. He then designed a study using this results section: some participants saw the exact results section, while a different group of people saw the results sections with one additional nonsense equation written afterwards. All the participants had a postgraduate degree. The participants were asked to score the paper's quality (between 0 and 100) based on what they read.

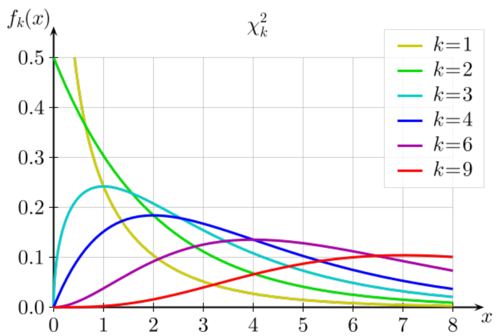
	Original Results Section	Results + Nonsense Equation
Mean	72.1	78.7
SD	16.2	17.3
n	84	75

Quality scores for the experiment

1. Using $\alpha = 0.05$, conduct a hypothesis test so see if a nonsense equation can increase people's perception of the quality of a paper. Make sure to: 1) state hypotheses, 2) draw a picture of a sampling distribution, 3) shade a p-value, 4) find the p-value, and 5) state your decision based on the p-value.

Exercise 5

As with discrete distributions, there are continuous distributions that arise all the time and answer "the common questions of the human experience". One example is known as the χ^2-distribution (pronounced ki-squared). This is actually a whole family of distributions, defined as follows by its density f_k.



$$f_k(x) = \begin{cases} \dfrac{1}{2^{k/2}\Gamma(k/2)} x^{\frac{k}{2}-1} e^{-x/2} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

Here, k is the degrees of freedom, and the Γ function is a generalization of the factorial symbol to non-integer values.

If we wanted to find the expected value and standard deviation of this distribution, we would have to do some very difficult integrals. Instead, we will simulate values from the distribution and find the mean and standard deviation of those values.

1. We generate $n = 10,000$ independent random values from the χ^2-distribution with $k = 16$ degrees of freedom, and get $\bar{x} = 16.068$ and $s_x = 5.6583$. With these data, build a 99% confidence interval for the true mean of the χ^2-distribution with 16 degrees of freedom.

2. Using advanced techniques for integals, researchers find that the standard deviation of the χ^2-distribution with k degrees of freedom is $\sigma = \sqrt{2k}$. With the method of question 1, how large should the sample size n be to get a 99% confidence interval of width smaller than 0.001?