Math 8452, Test, November 19, 2018 **Name:**

Instructions: Please put your answers on the separate sheets provided. You must show your work to receive full credit. In testing problems, please be sure to specify your hypotheses clearly and to state your conclusion in the context of the problem. In confidence interval problems, please be sure to give an interpretation of your interval in the context of the problem.

1. (8 points) Students A and B each took the same eight-question true/false statistics test. Their answers are tabled below, together with the correct answer for each question. Using level 0.20, test for evidence that student A is more likely to answer correctly than is student B.

Question	A	B	Correct
1	T	T	T
2	T	F	T
3	F	F	F
4	F	T	F
5	T	T	T
6	T	T	T
7	T	F	T
8	F	F	F

2. (8 points) The values tabled below are years of experience and annual salaries for statisticians employed at a certain company. Using an appropriate level-0.20 nonparametric test, test for a positive association between years of experience and annual salary.

注意理解association 指的是什么

Statistician	Years	Salary
1	4	$78,000
2	8	$87,000
3	10	$85,000
4	23	$102,000

3. (10 points) Consider the data given in problem 2. Assume that there is a linear relationship between salary and years of experience. Using Theil's method, find a 90% confidence interval for the slope of the regression line. What is the exact coverage probability of your interval?

4. (14 points) Given below are the lengths of 12 randomly selected issues of a statistics journal. (a) Find a nonparametric 80% confidence interval for the median length of an issue. What is the exact coverage probability of your interval? (b) Find a nonparametric 80% prediction interval for the length of the next issue of the journal. What is the exact coverage probability of your interval?

```
69  92  92  99  96 115  95  96 103  92  77 117
```

5. (10 points) A statistician has proposed a new "sum-of-squared-ranks" two-sample test that proceeds as follows. One pools the two samples and ranks all units from smallest to largest as one would in doing a rank sum test. The test statistic is the sum of the squares of the ranks for the values in the first sample. For example, if the first sample consists of the units with ranks 1, 2, and 5, then the test statistic is $1^2 + 2^2 + 5^2 = 30$. Find the null distribution of the test statistic when the first sample is of size 3 and the second sample is of size 2.

6. (16 points) Researchers studying two new corn varieties assigned each variety to seven plots at random. The yields (in bushels per acre) for the plots are tabled below. (a) Using a level-0.10 rank sum test or an equivalent test, test for a difference in the median corn yield for the two varieties. (b) Find the value of the two-sample Kolmogorov-Smirnov test statistic for these data.

Group	Yield in Bushels per Acre					
A	212	178	209	226	203	194
B	174	184	165	210	150	153

7. (10 points) Suppose that our two samples are $\{2, 4\}$ (first sample) and $\{6, 8\}$ (second sample). Using listing, find the bootstrap distribution of the difference between the two sample means. Please assume that we draw one independent bootstrap sample from each sample.

8. (10 points) A study was run to compare three different formulations of a drug. Thirty healthy subjects were available, and ten were randomly assigned to each of the formulations. Each formulation was administered to the appropriate subjects, and a single measurement was made on each subject one hour later. Suppose that we wish to test the theory that there is no difference between the three formulations. State the normal-theory approach to this problem, and also list two nonparametric approaches. For each of the three approaches that you list, indicate the assumptions that are required for the Type I error rate to be properly controlled. Be sure to highlight cases where one approach requires fewer assumptions than another.

9. (6 points) Provide an example of a small ($n \leq 5$) bivariate data set such that Spearman's rank correlation r_s is positive, but Pearson's product moment correlation r is negative.

10. (8 points) Independent simple random samples were drawn from two populations. A level-0.05 Ansari-Bradley test rejected the hypothesis of equal distributions ($p = 0.01$), but a level-0.05 rank sum test did not ($p = 0.57$). Without doing any calculations, explain how this is possible and what, if anything, it suggests about the two populations.