Math 8452, Test, Due by July 10, 2020. **Name:**

Instructions: Please write your answers on separate sheets. 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. Submit your R code as a separate text file.

1. (15 points) Using listing, find the null distribution of the rank sum statistic when $n_1 = n_2 = 3$ and there are no ties. What is the rejection region for a level-0.20 two-tailed test?

2. (10 points) A statistician has proposed a new two-sample test where the test statistic is the number of values from the first sample that are larger than any value from the second sample. Using listing, find the null distribution of the test statistic when $n_1 = 3$, $n_2 = 2$, and there are no ties.

22页第二章的笔记

3. (15 points) A small experiment was run to test whether listening to statistics talks affects milk yield for cows. Before and after milk yield data for three cows appears below. **(a)** Using listing, do a level-0.25 nonparametric test of the theory that listening to statistics talks has no effect on milk yield. Specify the p-value. **(b)** Without actually doing the test, list one other nonparametric test that you could have used instead of the test that you used in part (a).

样本不是独立的

Cow	Yield before	Yield after
1	10	12
2	15	19
3	12	15

4. (5 points) Provide an example of a small bivariate data set ($n \leq 5$) such that Spearman's rank correlation coefficient r_s is -1, but Pearson's correlation r is bigger than -1. You need not do any computations, but please briefly explain why your answer works.

5. (15 points) An agronomist studied the effect of mowing height on the phosphorous content of a certain species of prairie grass. Data from a completely randomized experiment are given below. **(a)** Using R, do a level-0.10 nonparametric test of the theory that mowing height has no impact on phosphorous content. Specify the p-value. **(b)** Without actually doing the test, list one other nonparametric test that you could have used instead of the test that you used in part (a).

Mowing height	Observations					
20 cm	42.0	18.0	14.0	36.0	11.6	19.0
10 cm	15.6	23.8	24.4	24.0	21.0	21.2
5 cm	35.3	22.5	16.9	25.0	23.1	26.0

6. (10 points) Your boss has just commented that one should never use the rank sum test because using ranks instead of the raw data amounts to throwing away information. Using 60 words or less, but being as specific as you can, please respond to this comment.

7. (15 points) Consider a two-sample setting in which one population is normal with mean 0 and standard deviation 1 and the other population is normal with mean 1 and standard deviation 2. Assume that the data will be two independent samples of size 10, and suppose that you are choosing between (i) the rank sum test, (ii) the Ansari-Bradley test, and (iii) the Kolmogorov-Smirnov test. By doing an appropriate simulation study in R, determine which of the three tests has the best power to detect the difference between the two populations. Use $\alpha = 0.05$.

8. (15 points) The values given below are a simple random sample of size 8 from an unknown continuous distribution. **(a)** Do a level-0.05 nonparametric test for evidence that the population median exceeds 0.5. Specify the p-value. **(b)** Find a 75% prediction interval for the next value that we draw from this distribution. What is the exact coverage probability for your interval?

```
0.980 0.381 0.842 0.045 0.505 0.697 0.773 0.529
```