Homework 6
(Practice Final Exam)
STAT W4413: Nonparametric Statistics

DUE: Friday, April 29, 8:00 pm

(1) Please sign your home work with your name and UNI number.

(2) Homework must be submitted into the Statistics Homework Boxes room 904 on the 9th floor of SSW building.

(3) Homework is due Friday, April 29, 8:00 pm.

(4) No late homework, under any circumstances, will be accepted.

(5) At the end of semester, one of your lowest homework scores will be dropped before the final grade is calculated.

(6) Your submitted solutions should consist of (i) the hand written (or printout) of the results with all the details, (ii) printout of the relevant figures (if applicable), and (iii) the printout of the source code (if applicable).

(7) This homework serves as a practice final exam.

1. (15 points) Let $X \sim F$ and $Y \sim G$. Suppose that $F(x) \geq G(x)$ for all x. Prove that $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

2. Basic results on order statistics:

   (a) (10 points) $X_1, X_2, \ldots, X_n \sim \text{Unif}(-2, 2)$ and $Y_1, Y_2, \ldots, Y_m \sim \text{Unif}(0, 2)$. Calculate the expected value of Wilcoxon rank-sum statistic.

   (b) (10 points) Let $X_1, X_2, \ldots, X_n \sim f(x)$. Derive a formula for the pdf of $X_{(1)}$.

   (c) (15 points) Let $X_1, X_2, \ldots, X_n \sim F(x)$, where F is a continuous CDF and satisfies $F(-x) = 1 - F(x)$. Calculate

$$\mathbb{E}(rank(X_1)rank(X_1^2)),$$

where for a function $g : \mathbb{R} \to \mathbb{R}$, $\text{rank}(g(X_1))$ denotes the rank of $g(X_1)$ between $g(X_1), \ldots, g(X_n)$.

3. (25 points) A Poisson distribution with intensity parameter $\lambda = 3$ was proposed to model the number of arrivals per minute at a bank in New York City. Suppose that the actual arrivals per minute were observed in 200 one-minute periods over the course of a week. The results are summarized in the following table.

Arrivals	0	1	2	3	4	5	6	7	8
Observed frequency	14	31	47	41	29	21	10	5	2

Use the Pearson's χ^2 test to determine if the proposed distribution can be used to model the arrivals.

4. Explain in a few words:

   (a) (10 points) What is the main advantage of all the permutation, Monte-Carlo, and bootstrap tests and what is the main limitation that we could not resolve?

   (b) (5 points) What are the advantages and disadvantages of using Monte-Carlo (or bootstrap) method in the permutation test?

   (c) (10 points) Describe the difference between parametric and nonparametric bootstrap. What are the advantages and limitation of the two methods.