Nonparametric Statistics Midterm

Spring 2022

Problem 1 (6pts)

Suppose you have random number generator that can generate iid samples from exp(1). Briefly explain how would you use it to generate iid samples from $N(0,1)$.

Problem 2 (6pts)

Let Z be a random variable. What are the first order approximations of $\mathbb{E}\{g(Z)\}$ and $Var\{g(Z)\}$?

Problem 3 (10pts)

Suppose we have a sequence consists of n_1 type 1 objects and n_2 type 2 objects.
(a) Write down the definition of "run". (2pts)
(b) Let R denote the number of runs in the sequence. What is the variance of R under the assumption that the sequence is completely random? (8pts)

Problem 4 (26pts)

Suppose you observe $X_1, ..., X_n \overset{iid}{\sim} F_X$ and you want to test the null hypothesis $H_0 : F_X = F_0$ against the alternative $H_1 : F_X \neq F_0$.

1. Give two test statistics that you think are appropriate. (4pts)
2. What are the intuitions behind the two test statistics? (10pts)
3. Write pseudo-codes to explain how would you estimate the p-values for the two test statistics. (12pts)

Problem 5 (26pts)

Suppose you observe $X_1, ..., X_n \overset{iid}{\sim} F_X$ and $Y_1, ..., Y_n \overset{iid}{\sim} F_Y$ and you know $F_X(x) = F_Y(\theta + x)$. You want to test the null hypothesis $H_0 : \theta = 0$ against the alternative $H_1 : \theta \neq 0$.

1. Give two test statistics that you think are appropriate. (4pts)

2. What are the intuitions behind the two test statistics? (10pts)

3. Write pseudo-codes to explain how would you estimate the p-values for the two test statistics. (12pts)

Problem 6 (22pts)

(a)Please list at least 5 conditions that a reasonable non-parametric measure of association should satisfy. (5pts)
(b) Write down the definition of Kendall's τ. (2pts)
(c) How does Kendall's τ relate to Pearson correlation coefficient? (5pts)
(d) Suppose the data you collected are of the form (X_i, Y_i, Z_i), $i = 1, \ldots n$. What is Kendall's partial correlation between X and Y when Z is hold constant? (5pts)
(e) What does a partial correlation measure? Give an intuitive explanation (5pts)