STAT 611 600: Theory of Inference **Spring 2021**

Homework 7: due Thursday, April 8, 2021, 11:59 pm CDT

Professor: Tiandong Wang *Name:* *UIN:*

Instructions:

- Whether you write out the solution by hand or in a text document, be sure that they are neat, legible and in order (even if you choose to solve them in different order). We highly recommend that you write your solutions in **LaTeX** and print them to a **PDF file**.
- Write/Type your name, UIN at the top of the first page. Otherwise, your submission will not be graded.
- Either scan or print your solutions to a **PDF file** under 15MB in size. It must be in a single file, not separate files for separate pages. Do not take a photo of each page and then paste them into a document - this will make your file too big and the results will generally not be very readable anyway.
- All students should login to their eCampus account to upload your file. You must do this by **11:59 pm U.S. Central time**, on the due date. You can make multiple submissions, but only the last submission will be graded.
- Write down all of your problem-solving process and cite any resources you have used in addition to lecture notes and the textbook.
- It is prohibited to share or distribute the content in this document.

1. 8.30 in C&B.

2. 8.33 in C&B.

3. (Sequential Hypothesis Testing) So far we have been considering an experimental setting where we decide in advance how many datpoints to collect, and design our test accordingly to have the desired trade-off between power (inversely related to type II error) and level (type I error bound).

   Now we would like to ask the question, can we design a sequential test which achieves the same error bounds, but reduces the expected number of data samples we need to collect? Rather than fixing the number of data points in advance, the sequential approach to testing gathers samples until a confident decision can be made.

   A *Sequential Probability Ratio Test* is based on considering the likelihood ratio as a function of the number of observations. Let $X_1, \ldots, X_k$ be independent samples (with pdf $f(\cdot)$) drawn from p_0 under the null hypothesis and p_1 under the alternative hypothesis, define

   $$\Lambda_k(X_1, \ldots, X_k) = \frac{f(X_1, \ldots, X_k; p_1)}{f(X_1, \ldots, X_k; p_0)} = \prod_{i=1}^{k} \frac{f(X_i; p_1)}{f(X_i; p_0)}.$$

   The goal of the SPRT is to decide which hypothesis is correct as soon as possible. The SPRT involves two thresholds $\gamma_1 > \gamma_0$. It rejects H_0 as soon as $\Lambda_k \geq \gamma_1$, and it accepts H_0 as soon as $\Lambda_k \leq \gamma_0$ We would like to choose these thresholds γ_0, γ_1 such that the test yields a desired level α and power β. Show that necessary conditions are

   $$\gamma_1 \leq \frac{\beta}{\alpha}, \qquad \gamma_0 \geq \frac{1-\beta}{1-\alpha}.$$

   Hint:

   $$f(X_1, \ldots, X_k; p_1) = f(X_1, \ldots, X_k; p_0) \cdot \frac{f(X_1, \ldots, X_k; p_1)}{f(X_1, \ldots, X_k; p_0)}.$$