UNIVERSITY OF TORONTO
Faculty of Arts and Science

STA261H1: Probability and Statistics II
Midterm 2
August 2, 2022

Instructor: Robert Zimmerman

- Do not open this test until you are told to begin.
- Midterm 2 is closed-book; no aids are allowed.
- There are six questions (worth a total of 30 points) on the midterm, plus one bonus question (worth five additional points). Take a quick scan through the questions first and prioritize your time accordingly. Do not attempt the bonus question until you are completely satisfied with your work on the remaining questions.
- Show all of your work for full marks, and ensure your notation is legible, correct, and consistent with that used in the course.
- If you need to use a result from lecture, either refer to it by its name (if it is a named theorem), or briefly describe it.

Good luck!

1. (5 points) Answer the following.

   (a) (1 point) State the definition of a **$(1-\alpha)$-confidence interval**.

   (b) (1 point) State what is meant by the term **type I error**.

   (c) (1 point) Write down circumstances (including the statistical model) under which we would use a **one-sided t-test**.

   (d) (2 points) Provide two examples of **p-hacking**.

2. (5 points) Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous distribution with density[1]

$$f_\theta(x) = \frac{\phi(x)}{\Phi(\theta)} \cdot \mathbb{1}_{x \leq \theta}, \quad \theta \in (-\infty, 0].$$

Suppose we want to test $H_0 : \theta = 0$ versus $H_A : \theta \neq 0$ using a likelihood ratio test, which rejects H_0 when $\lambda(\mathbf{x}) \leq c$ for some $c \in (0, 1)$.

(a) (3 points) Using the fact that the MLE of θ is $X_{(n)}$, find the LRT statistic $\lambda(\mathbf{X})$, and write down the corresponding rejection region R for your test, simplifying as much as possible. Your final $\lambda(\mathbf{X})$ shouldn't have any indicators in it.

(b) (2 points) Given that the cdf of $X_{(n)}$ is $F_{X_{(n)}}(x) = \begin{cases} \frac{\Phi(x)^n}{\Phi(\theta)^n}, & x \leq \theta \\ 1, & x \geq \theta \end{cases}$, calculate the power function $\beta(\theta)$ for the test, and find the value of c that makes the test of size-α.

[1]Here $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is the standard normal density, and $\Phi(x) = \int_{-\infty}^{x} \phi(t)\, \mathrm{d}t$ is the standard normal cdf. The inverse of this cdf is denoted by $\Phi^{-1}(\cdot)$, and it satisfies $\Phi^{-1}(\Phi(x)) = x$ for all $x \in \mathbb{R}$.

3. (5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathcal{N}\left(\mu, \sigma^2\right)$ where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Fix $\alpha \in (0, 1)$.

(a) (2 points) Recall that for this model, the sample variance S^2 is[2] sufficient for σ^2. Use the Karlin-Rubin theorem to show that for some c, the rejection region $R = \{s^2 : s^2 > c\}$ corresponds to a UMP level-α test of $H_0 : \sigma^2 \leq \sigma_0^2$ versus $H_A : \sigma^2 > \sigma_0^2$.

(b) (1 point) Recall that $\frac{(n-1)}{\sigma^2} S^2 \sim \chi^2_{(n-1)}$. Explain in one sentence why $\frac{(n-1)}{\sigma^2} S^2$ is a pivotal quantity.

(c) (2 points) Let $a < b$ satisfy $\mathbb{P}\left(a < \chi^2_{(n-1)} < b\right) = 1 - \alpha$. Derive a $(1-\alpha)$-confidence interval for σ^2.

[2] The pdf of S^2 is given by

$$g_{\sigma^2}(t) = C(\sigma^2) \cdot t^{(n-3)/2} \exp\left(-\frac{n-1}{2\sigma^2} t\right),$$

where $C(\sigma^2) > 0$ is some ugly normalizing constant.

4. (5 points) Answer each of the following questions by writing YES or NO (1 point), and justify your answer in *at most* three sentences (1.5 points).

   (a) A hypothesis test with power function $\beta(\cdot)$ is said to be **unbiased** if $\beta(\theta') \geq \beta(\theta'')$ for all $\theta' \in \Theta_0^c$ and all $\theta'' \in \Theta_0$. Suppose the true data-generating parameter is θ. Is an unbiased test always more likely to reject H_0 if $\theta \in \Theta_0^c$ than if $\theta \in \Theta_0$?

   (b) Suppose that $(L(\mathbf{X}), U(\mathbf{X}))$ is a 0.95-confidence interval for θ, and we observe $\mathbf{X} = \mathbf{x}$ for some $\mathbf{x} \in \mathcal{X}^n$. Must it be true that $\mathbb{P}_\theta\left(L(\mathbf{x}) \leq \theta \leq U(\mathbf{x})\right) \geq 0.95$?

5. (5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_\theta$ with cdf F_θ, and fix $t \in \mathbb{R}$. Recall that the ecdf $\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \leq t}$ is an estimator of the "parameter" $F_\theta(t)$. Given the concentration inequality

$$\mathbb{P}_\theta\left(\left|\hat{F}_n(t) - F_\theta(t)\right| > \epsilon\right) \leq 2e^{-2n\epsilon^2} \quad \text{for all } \epsilon > 0,$$

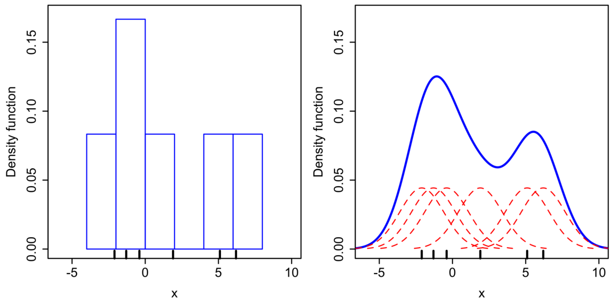
derive a $(1-\alpha)$-confidence interval for $F_\theta(t)$.

Hint: you're probably more interested in the complement of the event $\left|\hat{F}_n(t) - F_\theta(t)\right| > \epsilon$.

6. (5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_\theta$ where f_θ is continuous and supported on $\mathbb{R}$. Given any non-negative function $K(\cdot)$ such that $\int_{-\infty}^{\infty} K(x)\,\mathrm{d}x = 1$ and $K(-x) = K(x)$ for all $x \in \mathbb{R}$, the random function

$$\hat{f}_{n,h}(t) := \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h}\right)$$

is called a **kernel density estimator with kernel K and bandwidth h**. You can think of this as a kind of smoothed version of the density histogram function; here's a nice picture from Wikipedia comparing a histogram and a kernel density estimate based on six observed points:



(a) (3 points) Fix $t \in \mathbb{R}$. Prove that $\lim_{h \to 0} \mathbb{E}_\theta\left[\hat{f}_{n,h}(t)\right] = f_\theta(t)$. You can assume that limits and integrals can be swapped when it makes sense.

(b) (2 points) Assuming you can choose h to be very small, briefly explain how you'd use a kernel density estimator to visually check the hypothesis that f_θ actually generated the data.

7. (BONUS: 5 points) Let $X \sim \text{Bin}\,(k, \theta)$ with $\theta \in (0, 1)$ and k known.[3] Show that there exists no UMP level-α test of $H_0 : \theta = \theta_0$ versus $H_A : \theta \neq \theta_0$.

[3] X has pmf given by $f_\theta(x) = \binom{k}{\theta}\theta^x(1-\theta)^{k-x}$.