UNIVERSITY OF TORONTO
Faculty of Arts and Science

STA261H1: Probability and Statistics II
Midterm 1
July 16, 2024

Instructor: Robert Zimmerman

- Do not open this test until you are told to begin.
- Midterm 1 is closed-book; no aids are allowed.
- Whatever you write on the backs of these pages will not be graded, so you can use them for scratch work.
- 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.
- Show all of your work for full marks, and ensure your notation is legible, correct, and consistent with that used in the course. Use $\vec{X}$ and $\vec{x}$ to denote vectors.
- 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 **complete statistic**.

   (b) (2 points) State what it means for a distribution to be in an **exponential family**.

   (c) (1 point) State the theorem that we called the **bias-variance tradeoff**.

   (d) (1 point) State the **Lehmann-Scheffé theorem**.

2. (5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Unif}\,(-\theta, \theta)$, where $\theta \in (0, \infty)$.[1]

(a) (2.5 points) Prove that $T(\mathbf{X}) = \max_{1 \le i \le n} |X_i|$ is a minimal sufficient statistic for θ.

(b) (2.5 points) Compute the MOM estimator for θ.

[1]Recall that when $a < b$, the general Unif (a, b) distribution has pdf given by $f_{a,b}(x) = \frac{1}{b-a} \cdot \mathbb{1}_{x \in (a,b)}$.

3. (5 points) Let $X_1, X_2, \ldots, X_n$ be a random sample from a discrete distribution with pmf

$$f_\theta(x) = (\theta - 1) \cdot \theta^{-(x+1)}, \quad \theta > 1, \quad x \in \{0, 1, 2, 3, \ldots\}.$$

Let $\tau(\theta) = \theta/(\theta - 1)$.

(a) (2.5 points) Show that the MLE of $\tau(\theta)$ is $\tilde{\tau}(\mathbf{X}) = \bar{X}_n + 1$. You can skip the second derivative test.

(b) (2.5 points) Assuming that the MLE above is unbiased for $\tau(\theta)$, show that it is best unbiased.

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) Suppose $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_\theta$ and $T = T(\mathbf{X})$ is sufficient for θ. If no unbiased estimator of θ based on T exists, can we conclude that no unbiased estimator based on $\mathbf{X}$ exists either?

   (b) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_\theta$. Assuming it exists, is the MLE of θ always an unbiased estimator of θ?

5. (5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_{\mu,\sigma}$ be a random sample from a location-scale family, where μ is the location parameter and σ is the scale parameter (note that $f_{\mu,\sigma}$ is *not necessarily* the normal distribution). Prove that

$$D(\mathbf{X}) = \frac{X_1 - X_n}{X_{n-1} - X_n}$$

is ancillary for $\theta = (\mu, \sigma)$.

6. (5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_\theta$ and suppose that $T(\mathbf{X})$ is the UMVUE of θ. Prove that for any random variable U satisfying $\mathbb{E}_\theta[U] = 0$ and $\mathbb{E}\left[U^2\right] < \infty$, we must have $\text{Cov}_\theta(T(\mathbf{X}), U) = 0$.

Hint: define $S_\lambda = T + \lambda U$ for $\lambda \in \mathbb{R}$ and derive a contradiction by choosing λ to minimize $\text{Var}_\theta(S_\lambda)$.

7. (BONUS: 5 points) Without assuming unbiasedness, calculate the MSE of the MLE in Question 3.[2] It might help to use mgfs.

[2] Just so that you don't have to keep flipping pages, here it is again: the pmf is $f_\theta(x) = (\theta - 1) \cdot \theta^{-(x+1)}$ for $\theta > 1$ and $x \in \{0, 1, 2, \ldots\}$, and the MLE of $\tau(\theta) = \theta/(\theta + 1)$ is $\tilde{\tau}(\mathbf{X}) = \bar{X}_n + 1$.