UNIVERSITY OF TORONTO
Faculty of Arts and Science

STA261H1: Probability and Statistics II
Final Examination
August 21, 2024

Instructor: Robert Zimmerman

- This test is worth 50% of the final grade.
- Do not open this test until you are told to begin.
- You are may use a one-sided handwritten (not printed) cheat sheet. No other aids are permitted.
- The main questions on the exam are worth a total of 60 points, and there are bonus questions worth an additional 13 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. Be sure to clearly distinguish between random variables and constants, and between vectors and scalars (you can write $\mathbf{X}_n$ as $\vec{X}_n$ and $\mathbf{x}_n$ as $\vec{x}_n$).
- 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. (30 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Beta}(\theta, 1)$, where $\theta > 0$.[1] Let $Y_i = -\log(X_i)$, and let $\tau(\theta) = 1/\theta$.

   (a) (2 points) Find a complete sufficient statistic $T(\mathbf{X})$ for θ.

   (b) (2 points) Suppose that I did part (a) myself and my complete sufficient statistic is $\tilde{T}(\mathbf{X})$. Explain why it must be that $T(\mathbf{X}) = r(\tilde{T}(\mathbf{X}))$ for some bijection r.

   (c) (4 points) Show that $Q(\mathbf{X}, \theta) = X_1^\theta$ is a pivotal quantity for θ and use it to construct a $(1-\alpha)$-confidence interval for $\tau(\theta)$. Don't forget that $X_1 \in (0, 1)$.

   (d) (2 points) In general, is a pivotal quantity always an ancillary statistic? Or is it the other way around? Or neither? Justify your answer.

[1] The Beta (α, β) distribution has pdf $f_{\alpha,\beta}(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, where $x \in (0,1)$ and $\alpha, \beta > 0$. When $\alpha = \theta$ and $\beta = 1$, that becomes $f_\theta(x) = \theta x^{\theta-1}$.

(e) (2 points) Show that $Y_1, \ldots, Y_n \overset{iid}{\sim} \text{Exp}(\theta)$.

(f) (2 points) Explain why no matter what your complete sufficient statistic $T(\mathbf{X})$ from part (a) is, it's independent of Y_1/Y_n.

(g) (3 points) Show that the MLE of $\tau(\theta)$ based on $\mathbf{Y}$ is $\hat{\theta}_n(\mathbf{Y}) = \bar{Y}_n$. Don't bother with any second derivative tests.

(h) (3 points) Show that $\hat{\theta}_n(\mathbf{Y})$ is efficient for $\tau(\theta)$.

(i) (2 points) Is $\hat{\theta}_n(\mathbf{Y})$ also the UMVUE of $\tau(\theta)$? Answer YES or NO and briefly justify your answer. You don't really need any calculations here.

(j) (2 points) Find the asymptotic distribution of $\hat{\theta}_n(\mathbf{Y})$.

(k) (3 points) Construct an approximate size-α score test for testing $H_0 : \theta = \theta_0$ versus $H_A : \theta \neq \theta_0$, simplifying your rejection region as much as possible. It's probably better not to expand the square.

(l) (3 points) Invert your test into an approximate $(1 - \alpha)$-confidence interval for $\tau(\theta)$.

2. (a) (5 points) Suppose $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_{\theta_0}$. Let $S_n = \frac{1}{n} S(\theta_0 \mid \mathbf{X}_n)$, where $S(\theta \mid \mathbf{X}_n)$ is the usual score function. Show that S_n is a sample mean, and use that to determine its asymptotic distribution under the regularity conditions of the Cramér-Rao Lower Bound.

(b) (5 points) Prove that if some estimator T_n is asymptotically efficient for $\tau(\theta)$, then it must be consistent for $\tau(\theta)$.

3. (10 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{NegBin}(r, \theta)$ where $\theta \in (0, 1)$ and $r > 0$ is known.[2]

(a) (5 points) Show that Jeffreys' prior for θ is $\pi^J(\theta) \propto \dfrac{1}{\theta\sqrt{1-\theta}}$, and show that it's improper.

(b) (3 points) Nevertheless, determine the exact posterior distribution of θ (i.e., write down its name and parameters) when using this prior. The footnote on Page 2 might help.

(c) (2 points) Why would anyone consider using an improper prior? Explain.

[2]The $\text{NegBin}(r, \theta)$ distribution has pmf $f_\theta(x) = \binom{x+r-1}{x}(1-\theta)^x\theta^r$, where $x \in \{0, 1, 2, \ldots\}$, and its expectation is $\frac{r(1-\theta)}{\theta}$. In other news, if $h(x)$ is non-negative, then $\int_0^1 h(x) \cdot (1-x)^\gamma \,\mathrm{d}x \geq \int_0^1 h(x)\,\mathrm{d}x$ whenever $\gamma \leq 0$.

4. (a) (5 points) Suppose we choose a prior for θ which is itself a finite mixture of k priors; that is, $\pi(\theta) = \sum_{j=1}^{k} p_j \cdot \pi_j(\theta)$, where $p_j > 0$ and $\sum_{j=1}^{k} p_j = 1$ and each π_j is itself a pdf/pmf on Θ. Prove that the posterior is a finite mixture of k posteriors. You can make life easier by writing $f_j(\mathbf{x}) = \int \pi_j(\theta) \cdot f_\theta(\mathbf{x}) \, \mathrm{d}\theta$.

   For 3 bonus points, prove the statement for continuous mixtures instead of just finite mixtures.[3]

   (b) (5 points) Let $X_1, X_2, \ldots, X_n \stackrel{iid}{\sim} f_\theta$ and suppose we place a proper prior π on θ. Prove that $\pi(\theta \mid \mathbf{x}) = \pi(\theta)$ if and only if f_θ is free of θ.

[3] A continuous mixture distribution is a pdf of the form $f(x) = \int_{\mathcal{A}} q(a) \cdot f_a(x) \, \mathrm{d}a$, where $q(a)$ is itself a pdf on some continuum $\mathcal{A} \subseteq \mathbb{R}$, and each f_a is a pdf/pmf, for every $a \in \mathcal{A}$.

5. (a) (BONUS: 5 points) Suppose that $\theta \in \Theta \subseteq \mathbb{R}$ and $\sqrt{n}(T_n - \theta) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right)$. If some function g is k-times continuously differentiable and $g'(\theta) = g''(\theta) = \cdots = g^{(k-1)}(\theta) = 0$ but $g^{(k)}(\theta) \neq 0$, find the asymptotic distribution of $g(T_n)$ by developing a k'th order delta method.

(b) (BONUS: 5 points) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Beta}(1, \beta)$, where $\beta > 0$. Find deterministic sequences $a_n \in \mathbb{R}$ and $b_n > 0$ so that

$$\frac{X_{(n)} - a_n}{b_n}$$

converges in distribution to a non-degenerate limit.[4]

[4]In fact, the Fisher-Tippett-Gnedenko theorem (yes, *that* Fisher) says that *regardless* of the distribution of the X_i's, if there exist deterministic sequences $a_n \in \mathbb{R}$ and $b_n > 0$ such that $(X_{(n)} - a_n)/b_n$ converges in distribution to a non-denerate limit, then that limit follows a particular kind of distribution called a "GEV distribution". This result is essentially a central limit theorem for extreme values (i.e., sample maxima), and it is the basis for the field of *extreme value theory*.