Math 541a, Spring 2023, USC Instructor: Steven Heilman

Name: ______________________ USC ID: ______________ Date: ___________

Signature: ______________________. Discussion Section: ______________________
(By signing here, I certify that I have taken this test while refraining from cheating.)

Exam 2

This exam contains 7 pages (including this cover page) and 4 problems. Enter all requested information on the top of this page.

You may *not* use your books, notes, or any calculator on this exam.

You are required to show your work on each problem on this exam. The following rules apply:

- You have 50 minutes to complete the exam, starting at the beginning of class.
- **Organize your work**, in a reasonably neat and coherent way, in the space provided. Work scattered all over the page without a clear ordering will receive very little credit.
- **Mysterious or unsupported answers will not receive full credit**. A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.
- If you need more space, use the back of the pages; clearly indicate when you have done this. Scratch paper appears at the end of the document.

Do not write in the table to the right. Good luck![a]

Problem	Points	Score
1	10	
2	10	
3	10	
4	10	
Total:	40	

[a]April 5, 2023,

Reference sheet

Below are some definitions that may be relevant.

We say that a sequence of random variables $X_1, X_2, \ldots : \Omega \to \mathbf{R}$ **converges in probability** to a random variable $X : \Omega \to \mathbf{R}$ if: for all $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0.$$

We say that a sequence of real-valued random variables $X_1, X_2, \ldots$ **converges in distribution** to a real-valued random variable X if, for any $t \in \mathbf{R}$ such that $\mathbf{P}(X \leq t)$ is continuous at t,

$$\lim_{n\to\infty} \mathbf{P}(X_n \leq t) = \mathbf{P}(X \leq t).$$

We say that a sequence of random variables $X_1, X_2, \ldots : \Omega \to \mathbf{R}$ **converges almost surely** to a random variable $X : \Omega \to \mathbf{R}$ if

$$\mathbf{P}(\lim_{n\to\infty} X_n = X) = 1.$$

Suppose $X = (X_1, \ldots, X_n)$ is a random sample of size n from a distribution f where $f \in \{f_\theta : \theta \in \Theta\}$ is a family of densities (such as an exponential family). Let $t : \mathbf{R}^n \to \mathbf{R}^k$, so that $Y := t(X_1, \ldots, X_n)$ is a statistic.

We say that Y is a **sufficient statistic** for θ if, for every $y \in \mathbf{R}^k$ and for every $\theta \in \Theta$, the conditional distribution of $(X_1, \ldots, X_n)$ given $Y = y$ (with respect to probabilities given by f_θ) does not depend on θ.

We say Y is **minimal sufficient** for θ if Y is sufficient for θ and, for every statistic $Z : \Omega \to \mathbf{R}^m$ that is sufficient for θ, there exists a function $r : \mathbf{R}^m \to \mathbf{R}^k$ such that $Y = r(Z)$.

We say Y is **complete** for $\{f_\theta : \theta \in \Theta\}$ if the following holds:

For any $f : \mathbf{R}^m \to \mathbf{R}$ such that $\mathbf{E}_\theta f(Y) = 0 \quad \forall\, \theta \in \Theta$, it holds that $f(Y) = 0$.

We say Y is **ancillary** for θ if the distribution of Y does not depend on θ.

Let $X, Y, Z : \Omega \to \mathbf{R}$ be discrete or continuous random variables. Let A be the range of Y. Define $g : A \to \mathbf{R}$ by $g(y) := \mathbf{E}(X|Y = y)$, for any $y \in A$. We then define the **conditional expectation** of X given Y, denoted $\mathbf{E}(X|Y)$, to be the random variable $g(Y)$.

Let $\{f_\theta : \theta \in \Theta\}$ be a family of multivariable probability densities or probability mass functions. Assume $\Theta \subseteq \mathbf{R}$. Let X be a random vector with distribution f_θ. Define the **Fisher information** of the family to be

$$I(\theta) = I_X(\theta) := \mathbf{E}_\theta\left(\frac{d}{d\theta} \log f_\theta(X)\right)^2, \qquad \forall\, \theta \in \Theta,$$

if this quantity exists and is finite.

1. (10 points) Let X, Y be random variables such that (X, Y) is uniformly distributed in the region

$$\{(x, y) \in \mathbf{R}^2 \colon x^2 + y^2 \leq 1\}.$$

Compute the following quantities:

- $\mathbf{E}(X|Y)$.
- $\mathbf{E}[\mathbf{E}(X|Y)]$.

2. (10 points) Let $X := (X_1, \ldots, X_n)$ be a random sample of size n from a binomial distribution with parameters n and p. Here n is a positive (known) integer and $0 < p < 1$ is unknown. (That is, $X_1, \ldots, X_n$ are i.i.d. and X_1 is a binomial random variable with parameters n and p, so that $\mathbf{P}(X_1 = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for all integers $0 \leq k \leq n$.)

   You can freely use that $\mathbf{E}X_1 = np$ and $\mathrm{Var}X_1 = np(1-p)$.

   - Computer the Fisher information $I_X(p)$ for any $0 < p < 1$.
     (Consider n to be fixed.)
   - Let Z be an unbiased estimator of p (assume that Z is a function of $X_1, \ldots, X_n$). State the Cramér-Rao inequality for Z.
   - Let W be an unbiased estimator of p^3 (assume that W is a function of $X_1, \ldots, X_n$). State the Cramér-Rao inequality for W.

3. (10 points) Let X be a binomial random variable with parameters n and p. Here n is a positive (known) integer and $0 < p < 1$ is unknown. That is, $\mathbf{P}(X = k) = \binom{n}{k}p^k(1-p)^{n-k}$ for all integers $0 \leq k \leq n$.)

   Prove that no UMVU exists for the quantity $1/p$. (The sample size in this case is one.)

   (Recall that UMVU means: uniform minimum variance unbiased estimator, i.e. the estimator has smallest variance among all unbiased estimators, uniformly over all unknown parameters.)

4. (10 points) Let $n \geq 2$. Let $X_1, \ldots, X_n$ be a random sample from the Gaussian distribution with unknown mean $\mu \in \mathbf{R}$ and unknown variance $\sigma^2 > 0$.

   Find the UMVU for μ^3.

   (When you find the UMVU, denote it by Y_n, and you must assume that Y_n is a function of $X_1, \ldots, X_n$.)

   (In this question you can freely cite facts from the homework.)

   (You can freely use the following computations:
   $\mathbf{E}\overline{X}_n^2 = \mu^2 + \sigma^2/n$, and $\mathbf{E}\overline{X}_n^3 = \mu^3 + 3\mu\sigma^2/n$, and $\mathbf{E}S_n^2 = \sigma^2$.)

   (Recall that $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $S_n := \sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2}$.)

(Scratch paper)