Graduate Statistics, 541A, Spring 2019, USC   Instructor: Steven Heilman

Name: _____   USC ID: _____   Date: _____

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

Final Exam

This exam contains 13 pages (including this cover page) and 8 problems. Check to see if any pages are missing. Enter all requested information on the top of this page.

You may *not* use your book or any calculator on this exam. You *cannot* use your homeworks. You are required to show your work on each problem on the exam. The following rules apply:

- You have 120 minutes to complete the exam.

- **If you use a theorem or proposition from class or the notes or the book you must indicate this** and explain why the theorem may be applied. It is okay to just say, "by some theorem/proposition from 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.

Problem	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
Total:	80	

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

---

[a]May 4, 2019, © 2018 Steven Heilman, All Rights Reserved.

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 $X \colon \Omega \to \mathbf{R}$ 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 $\mathbf{P}(X \le t)$ is continuous at t,

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

We say that a sequence of random variables $X_1, X_2, \ldots : \Omega \to \mathbf{R}$ **converges in** L_2 to a random variable $X \colon \Omega \to \mathbf{R}$ if

$$\lim_{n \to \infty} \mathbf{E} \, |X_n - X|^2 = 0.$$

We say that a sequence of random variables $X_1, X_2, \ldots : \Omega \to \mathbf{R}$ **converges almost surely** to a random variable $X \colon \Omega \to \mathbf{R}$ 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 $f \in \{f_\theta \colon \theta \in \Theta\}$ is a family of densities (such as an exponential family). Let $t \colon \mathbf{R}^n \to \mathbf{R}^k$, 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 $Z \colon \Omega \to \mathbf{R}^m$ that is sufficient for θ, there exists a function $r \colon \mathbf{R}^m \to \mathbf{R}^k$ such that $Y = r(Z)$.

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

For any $f \colon \mathbf{R}^m \to \mathbf{R}$ 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 $X, Y, Z \colon \Omega \to \mathbf{R}$ be discrete or continuous random variables. Let A be the range of Y. Define $g \colon A \to \mathbf{R}$ 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 $g(Y)$.

We say that Y is **uniformly minimum variance unbiased (UMVU)** for $g(\theta)$ if, for any other unbiased estimator Z for $g(\theta)$, we have $\mathrm{Var}_\theta(Y) \le \mathrm{Var}_\theta(Z) \; \forall \; \theta \in \Theta$.

1. (10 points) Let $X_1, \ldots, X_n$ be a random sample of size $n \geq 1$. Let $m \geq 1$. Recall that the **bootstrap sample** $Y_1, \ldots, Y_m$ is defined as follows. Given $X_1, \ldots, X_n$, let $Y_1, \ldots, Y_m$ be a random sample of size m from the values $\{X_1, \ldots, X_n\}$ with replacement.

   Show that the covariance of Y_1 and Y_2 is equal to $\mathrm{Var}(X_1)/n$.

   Conclude that $Y_1, \ldots, Y_m$ are not independent when $\mathrm{Var}(X_1) > 0$.

   (Hint: to compute $\mathbf{E}Y_1 Y_2$, first condition on $X_1, \ldots, X_n$).

2. (10 points) Let $X_1, \ldots, X_n$ be a random sample from the Bernoulli distribution with unknown parameter $0 < p < 1$, so that, for all $1 \le i \le n$,

$$\mathbf{P}(X_i = 1) = p, \qquad \mathbf{P}(X_i = 0) = 1 - p.$$

- Find a complete sufficient statistic for p. (As usual, justify your answer.)
- Find the UMVU for p^3. (You may assume $n \ge 3$.)
  (Hint: $X_1 X_2 X_3$ is an estimator for p^3.)

[This was part of a qual question, and a modified example from class.]

3. (10 points) Let $X_1, \ldots, X_n$ be a random sample of size n, so that X_1 is a sample from the uniform distribution on the interval $[\theta - 1/2, \theta + 1/2]$, where $\theta \in \mathbf{R}$ is unknown. Let $g \colon \mathbf{R} \to \mathbf{R}$ be a nonconstant differentiable function of $\theta \in \mathbf{R}$. Show that no UMVU of $g(\theta)$ exists when $n = 1$.

[This was a repeated homework question.]

4. (10 points) Let $f \colon \mathbf{R}^n \to \mathbf{R}$ be a convex function. Let $x \in \mathbf{R}^n$ be a local minimum of f. Show that x is in fact a global minimum of f.

Now suppose additionally that f is a C^1 function (all derivatives of f exist and are continuous), and $x \in \mathbf{R}^n$ satisfies $\nabla f(x) = 0$. Show that x is a global minimum of f.

[This was a repeated homework question.]

5. (10 points) Give an example where the maximum likelihood estimator and the method of moments estimator are the same. That is, for a random sample of size $n \geq 1$, if Y_n is the MLE for the sample of size n, and if Z_n is the Method of Moments estimator for the sample of size n, then $Y_n = Z_n$. (As usual, justify your answer.)

[This was a modified qual question.]

6. (10 points) Let $X := (X_1, \ldots, X_n)$ be a random sample of size n from a Gaussian distribution with unknown mean $\mu \in \mathbf{R}$ and unknown variance $\sigma^2 > 0$. $\forall\, n \geq 1$, define

$$Y = Y_n = Y_n(X_1, \ldots, X_n) := \frac{1}{n} \sum_{j=1}^{n} \left(X_j - \frac{1}{n} \sum_{i=1}^{n} X_i \right)^2.$$

You are required to show that Y_n has asymptotically optimal variance as $n \to \infty$ (you are not allowed to use the theorem about the limiting distribution of the MLE; you are required to argue directly.) More specifically, the Cramér-Rao inequality shows that for any unbiased estimator $Z = Z_n(X_1, \ldots, X_n)$ of σ^2, there is a constant $c_n > 0$ such that

$$\mathrm{Var}_\sigma(Z) \geq c_n.$$

You are required to compute the lower bound c_n given by the Cramér-Rao Inequality, and you should also show that

$$\lim_{n \to \infty} \frac{\mathrm{Var}_\sigma(Y_n)}{c_n} = 1.$$

[This was a calculation from class.]

7. (10 points) Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample of size n on the disc $\{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq s^2\}$ where $s > 0$ is unknown. That is, the probability density function of (X, Y) is

$$f_s(x, y) := \frac{1}{\pi s^2} 1_{[0,s]}(\sqrt{x^2 + y^2}), \qquad \forall\, (x, y) \in \mathbf{R}^2$$

Find a complete sufficient statistic Z for s and find the density function of Z.

(Hint: first try the $n = 1$ case, then try the $n > 1$ case.)

[This was part of a qual question.]

8. (10 points) Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample of size n on the disc $\{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq s^2\}$ where $s > 0$ is unknown. That is, the probability density function of (X, Y) is

$$f_s(x, y) := \frac{1}{\pi s^2} 1_{[0,s]}(\sqrt{x^2 + y^2}), \qquad \forall (x, y) \in \mathbf{R}^2$$

Find the UMVU for s.

[This was part of a qual question.]

(Scratch paper 1)

(Scratch paper 2)

(Scratch paper 3)