Math 541b, Fall 2025, 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 5 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.

Problem	Points	Score
1	6	
2	8	
3	10	
4	8	
5	10	
Total:	42	

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

1. (6 points) Let $X_1, X_2, \ldots : \Omega \to \mathbf{R}$ be i.i.d random variables. Let $Y_1, Y_2, \ldots$ be a sequence of estimators so that for any $n \geq 1$, $Y_n = t_n(X_1, \ldots, X_n)$ for some $t_n \colon \mathbf{R}^n \to \mathbf{R}$. For any $n \geq 1$, define the **jackknife estimator** of Y_n to be

$$Z_n := nY_n - \frac{n-1}{n} \sum_{i=1}^{n} t_{n-1}(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n). \qquad (*)$$

   - Assume that there exists $\theta, a, b \in \mathbf{R}$ such that

     $$\mathbb{E}Y_n = \theta + a/n, \quad \forall n \geq 1.$$

     Show that

     $$\mathbb{E}Z_n = \theta.$$

   - The jackknife described above involves summing over all ways to delete one of the samples from $X_1, \ldots, X_n$. Write a formula for a term to add to $(*)$ that also sums over all ways to delete exactly two of the samples from $X_1, \ldots, X_n$ in t_{n-2}.

   [this was a repeated homework exercise]

2. (8 points) Write down the generalized likelihood ratio estimate for the following alpha particle data, as we did in class for a slightly different data set. The corresponding test treats individual counts of alpha particles as independent Poisson random variables, versus the alternative that the probability of a count appearing in each box of data is a sequence of nonnegative numbers that sum to one.

m	0,1,2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	≥ 17
# Ints	16	26	58	102	125	146	163	164	120	100	72	54	20	12	10	4

Suppose we wanted to plot the MLE for the Poisson statistic (i.e. plot the denominator of the generalized likelihood ratio test statistic $\frac{\sup_{\theta\in\Theta} f_\theta(X)}{\sup_{\theta\in\Theta_0} f_\theta(X)}$) as a function of λ.

Describe in detail how you would plot this MLE on a computer, with particular detail on how to avoid outputting zeros or infinities that should not occur.

[this was a repeated homework exercise]

3. (10 points) Let $X_1, \ldots, X_n$ be a random sample from a Gaussian distribution with known variance $\sigma^2 > 0$ and unknown mean $\mu \in \mathbb{R}$. Fix $\mu_0 \in \mathbb{R}$. Suppose we want to test the hypothesis H_0 that $\mu = \mu_0$ versus the alternative H_1 that $\mu \neq \mu_0$. That is, $\Theta = \mathbb{R}$, $\Theta_0 = \{\mu_0\}$ and $\Theta_0^c = \{\mu \in \mathbb{R} \colon \mu \neq \mu_0\}$. Also, for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$f_\mu(x) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}.$$

- Explicitly describe the rejection region of the generalized likelihood ratio test.
- Denote $X = (X_1, \ldots, X_n)$. If H_0 is true, describe the distribution of

$$2 \log \frac{\sup_{\theta \in \Theta} f_\theta(X)}{\sup_{\theta \in \Theta_0} f_\theta(X)}$$

(Hint: you can freely use the identity $\sum_{i=1}^{n}[(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j)^2 - (x_i - \mu_0)^2] = n(\mu_0 - \frac{1}{n}\sum_{j=1}^{n} x_j)^2 - 2n(\frac{1}{n}\sum_{i=1}^{n} x_i - \mu_0)(\frac{1}{n}\sum_{j=1}^{n} x_j - \mu_0))$
[this was a repeated example from class]

4. (8 points) Let $X_1, \ldots, X_{16}$ denote real valued random variables with $\sum_{j=1}^{16} X_j = 1207$. Denote $X = (X_1, \ldots, X_{16})$. Suppose we know that

$$Z := 2 \log \frac{\sup_{\theta \in \Theta} f_\theta(X)}{\sup_{\theta \in \Theta_0} f_\theta(X)} = 2 \cdot 1207 \sum_{j=1}^{16} \frac{X_j}{1207} \log\left(\frac{X_j/1207}{p_j}\right),$$

for some constants $p_1, \ldots, p_{16} > 0$ with $\sum_{j=1}^{16} p_j = 1$. Suppose we also know that $X_j/1207 \approx p_j$ for all $1 \leq j \leq 16$.

Using a Taylor expansion of the function $h(a) = a \log(a/b)$, show that Z is approximately equal to Pearson's chi-squared statistic

$$\sum_{j=1}^{16} \frac{(X_j - 1207 p_j)^2}{1207 p_j}$$

[this was a repeated example from class]

5. (10 points) Let X_1, X_2, X_3 be i.i.d. continuous random variables. Let W_1, W_2, W_3 be a bootstrap sample from X_1, X_2, X_3. Let Y denote the sample median of X_1, X_2, X_3. (That is, Y is the middle value among X_1, X_2, X_3, which is unique with probability one since the random variables are continuous.)

   - Describe the distribution of $(W_{(1)}, W_{(2)}, W_{(3)})$.
   - Describe the bootstrap estimator of Y.

   [this was a repeated homework exercise]

(Scratch paper)