Math 541b, Fall 2026, 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	5	
2	10	
3	10	
4	10	
5	8	
Total:	43	

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

[a] November 2, 2026, © 2026 Steven Heilman, All Rights Reserved.

1. (5 points) Let X be a Gaussian random variable with unknown mean $\mu_X \in \mathbf{R}$ and variance 1, so that X has the following density (with $\sigma = 1$)

$$f_{\mu_X}(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-(z-\mu_X)^2/(2\sigma^2)\Big), \qquad \forall\, z \in \mathbf{R}.$$

Let Y be a Gaussian random variable with unknown mean $\mu_Y \in \mathbf{R}$ and variance 1.

Assume that X and Y are independent. Let $\alpha \in (0,1)$.

Find a $100(1-\alpha)\%$ confidence interval for $\mu_X - \mu_Y$.

Simplify the expression to a reasonable extent.

[this is a repeated example from class/homework]

2. (a) (3 points) Complete the statement of the following Theorem from the notes about the limiting distribution of the GLR statistic:

Let $X := (X_1, ..., X_n)$ be a random sample of size n from a family of distributions $\{f_\theta : \theta \in \Theta\}$. Fix $\theta_0 \in \Theta \subseteq \mathbf{R}$. Suppose we test the hypothesis H_0 that $\{\theta = \theta_0\}$ versus the alternative $\{\theta \neq \theta_0\}$. Suppose the assumptions of Theorem 2.69 from the notes hold. Let $\lambda(X) := \frac{\sup_{\theta\in\Theta} f_\theta(X)}{\sup_{\theta\in\Theta_0} f_\theta(X)}$ denote the generalized likelihood ratio statistic. If H_0 is true, then as $n \to \infty$ [complete this sentence].

(b) (3 points) In the alpha particle examples from class and the homework, we had a table with 16 entries of the following form, corresponding to the results of 1207 independent experiments

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

What did we conclude about the distribution of the random variable

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

(Here $X = (X_1, \ldots, X_{16})$, X_i is the i^{th} entry in the table, so that $\sum_{j=1}^{16} X_j = 1207$.)

(c) (4 points) Are the conclusions of parts (a) and (b) above different? Explain **in detail** why the conclusions of parts (a) and (b) are the same, or not the same.

3. (10 points) Let $(x_1, y_1), \ldots, (x_n, y_n)$ in $\mathbf{R} \times \mathbf{R}$.

For any $m, b \in \mathbf{R}$, define

$$f(m, b) := \sum_{i=1}^{n} (y_i - mx_i - b)^2.$$

Using the second derivative test, show that the minimum value of f is achieved when

$$m = \frac{\left(\sum_{i=1}^n x_i\right)\left(\sum_{j=1}^n y_j\right) - n\left(\sum_{k=1}^n x_k y_k\right)}{\left(\sum_{i=1}^n x_i\right)^2 - n\left(\sum_{j=1}^n x_j^2\right)} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{j=1}^n (x_j - \overline{x})^2}.$$

$$b = \frac{1}{n}\left(\sum_{i=1}^n y_i - m \sum_{j=1}^n x_j\right) = \overline{y} - m\overline{x}.$$

Briefly explain why this is actually the minimum value of $f(m, b)$. (You are allowed to use the inequality $(\sum_{i=1}^n x_i)^2 \leq n(\sum_{i=1}^n x_i^2)$.)

[this is a repeated homework question]

4. Let

$$h(t) = \frac{1}{1+e^{-t}}, \qquad \forall\, t \in \mathbf{R}.$$

In logistic regression, we are given independent random variables $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ and we assume there exists $a, b \in \mathbf{R}$ such that

$$P(Y_i = 1) = h(aX_i + b), \qquad \mathbf{P}(Y_i = 0) = 1 - \mathbf{P}(Y_i = 1), \qquad \forall\, i \geq 1.$$

Suppose $n = 2$ and

$$(X_1, Y_1) = (-1, 0), \qquad (X_2, Y_2) = (1, 1).$$

(a) (2 points) Write an expression for the log-likelihood $\ell(a, b)$. Simplify it to a reasonable extent.

(b) (3 points) Verify that $\frac{d^2}{dt^2} \log h(t) < 0$ for all $t \in \mathbf{R}$. (You don't have to show it, but this implies that ℓ is strictly concave on $\mathbb{R}^2$, That is, for any $s \in \mathbf{R}$, and for any $u, w \in \mathbf{R}$ $\frac{d^2}{ds^2}|_{s=0} \ell(a + su, b + sw) < 0$.)

(c) (5 points) Does ℓ have a maximum value in the plane? Explain why your conclusion does not contradict with part (b).

5. (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/ practice exams]

(Scratch paper)