Midterm 1

Graded

Student

Sampad Mohanty

Total Points

52 / 60 pts

Question 1

Problem 1

18 / 20 pts

- ✔ **+ 8 pts** (a) is Correct
- ✔ **+ 7 pts** (b) is correct
- **+ 5 pts** (c) is correct
- **+ 5 pts** (a) The factorization theorem was applied, but the joint density or the final answer was incorrect.
- **+ 5 pts** (b) Most of the steps are correct. Conclusion is incorrect.
- ✔ **+ 3 pts** (c) Used the correct definition, but there is a gap in the proof.
- **+ 4 pts** (c) checked $f(x|\theta)$ is an exponential family, but did not mention the open set condition.
- **+ 1 pt** (c) state the definition correctly
- **+ 4 pts** Checked exponential family and open set condition. Only the conclusion is incorrect

1 this needs justification. g(t) can take both positive and negative values

Question 2

Problem 2

14 / 20 pts

- ✔ **+ 5 pts** (a) is correct
- **+ 4 pts** (a) partially correct
- **+ 2 pts** (a) Computed the expectation correctly
- ✔ **+ 7 pts** (b) is correct
- **+ 2 pts** (b) Use the definition of MLE correctly
- **+ 8 pts** (c) is correct
- **+ 1 pt** (c) Only write down the formula for MSE
- ✔ **+ 2 pts** (c) computed the MSE but did not find the correct distribution of $X_{(n)}$.
- **+ 4 pts** c) Show correctly that MLE is biased
- **+ 6 pts** c) Show correctly that MLE is biased but the variance computation is only partially correct
- **+ 7 pts** minor error in (c)

2 you need to know the pdf of $X_{(n)}$ first

Question 3

Problem 3

20 / 20 pts

- ✔ **+ 8 pts** (a) is Correct
- ✔ **+ 7 pts** (b) is correct
- ✔ **+ 5 pts** (c) is correct
- **+ 2 pts** (a) partial credits for $L(\theta)$.
- **+ 6 pts** (a) Write the MLE as a solution to an optimization question correctly. but didn't write down the explicit answer, or the final answer is incorrect
- **+ 4 pts** (b) mostly correct for the posterior distribution
- **+ 6 pts** (b) minor error
- **+ 2 pts** (b) partial credits for the posterior density calculation
- **+ 2 pts** (c) Write down the equation for MAP and posterior mean correctly
- **+ 4 pts** (c) Partially correct for the final expression,
- **+ 0 pts** Incorrect

Math 541 A - Spring 2025
Midterm 1
Feb 26, 2025
11:00-11:50 AM

Name: SAMPAD MOHANTY

Student ID Number: 567952831?

- There are 3 problems in total. Make sure your exam contains all these questions.
- You are allowed to use one page of double-sided **hand-written** 8.5 by 11 inch notes.
- You must show your work on all problems. The correct answer with no supporting work may result in no credit.
- If you need more room, use the backs of the pages and indicate to the grader that you have done so.

Problem 1	20	
Problem 2	20	
Problem 3	20	
Total	60	

Name: ID:

1. (20 pts) Suppose $X_1, \ldots, X_n$ are i.i.d. random variables with common pdf

$$f(x|\theta) = \begin{cases} (\theta+1)x^\theta, & x \in (0,1), \\ 0 & \text{otherwise,} \end{cases} \qquad \text{where} \quad \theta > -1.$$

(a) (8 pts) Use the Factorization Theorem to find a 1-dimensional sufficient statistic $T(X_1, \ldots, X_n)$ for θ.

(b) (7 pts) Find a minimal sufficient statistic for θ. Justify your answer.

(c) (5 pts) Find a complete statistic for θ. Justify your answer.

$$f(x_1, x_2 \ldots x_n|\theta) = \prod_{i=1}^{n} (\theta+1)x_i^\theta \, \mathbb{I}_{[x_i \in (0,1)]} = (\theta+1)^n \left[\prod(x_i)\right]^\theta \, \mathbb{I}_{\{\max_i x_i < 1,\ \min_i x_i > 0\}}$$

define $T(X) = \prod_i x_i \rightarrow f(\vec{x}|\theta) = \underbrace{(\theta+1)^n T^\theta}_{g(T|\theta)} \underbrace{\mathbb{I}_{\{\max x_i < 1\}} \cdot \mathbb{I}_{\{\min x_i > 0\}}}_{h(x)}$

By factorization Thm; $T(X) = \prod_i x_i$ is sufficient for θ.

I claim $T(X) = \prod_i x_i$ is sufficient & minimal too.

By Lehman-Scheffe, $\dfrac{f(\vec{x}|\theta)}{f(\vec{y}|\theta)} = \dfrac{(\theta+1)^n T(x)^\theta \, \mathbb{I}_{\{\max_i x_i < 1\}} \mathbb{I}_{\{\min_i x_i > 0\}}}{(\theta+1)^n T(y)^\theta \, \mathbb{I}_{\{\max y_i < 1\}} \mathbb{I}_{\{\min_i y_i > 0\}}}$ is independent of θ $\Leftrightarrow$ ~~iff~~ $T(x) = T(y)$

$\therefore$ $T(X)$ is minimal & sufficient statistic.

) claim $\rightarrow$ $T(X) = \prod_i x_i$ is complete.

$\forall \theta, g \quad E_\theta[g(T)] = 0 \Rightarrow P_\theta(g(T) = 0) = 1.$

$E_\theta[g(\prod_i x_i)] = \int_0^1 \underbrace{g(\prod_i x_i)} \boxed{(\theta+1)x^\theta} dx = 0.$

+ve always

this has to be zero to make the area under curve $= 0$.

$\therefore$ $T(X) = \prod_i x_i$ is complete for θ.

Name: ID:

2. (20 pts) Suppose $X_1, \dots, X_n$ are i.i.d. from a distribution given by

$$f(x|\theta) = \frac{1}{\theta}1\{0 \le x \le \theta\}, \quad \theta > 0.$$

(a) (5 pts) Find a method of moment estimator for θ.

(b) (7 pts) Determine the Maximum Likelihood Estimator $\hat{\theta}_{\mathrm{MLE}}$. Justify your answer.

(c) (8 pts) Show that $\hat{\theta}_{\mathrm{MLE}}$ is a biased estimator for θ and find the mean squared error of $\hat{\theta}_{\mathrm{MLE}}$.

a) This is uniform $U(0,\theta)$. mean $= \theta/2$, variance $= \int_0^\theta \frac{1}{\theta}(x-\theta/2)^2 = \int_0^\theta \frac{(x^2 + \frac{\theta^2}{4} - \theta x)}{\theta} = \frac{1}{\theta}\left(\frac{\theta^3}{3} + \frac{\theta^3}{4} - \right.$
$= \theta^2\left(\frac{1}{3} + \frac{1}{4} - \frac{1}{2}\right)$
$= \theta^2\left(\frac{4+3-6}{12}\right)$
$= \theta^2/12.$

Set mean = sample mean $\Rightarrow \frac{\sum x_i}{n} = \theta/2 \Rightarrow \boxed{\theta = \frac{2\sum x_i}{n}}$

ignore this.

2nd moment $= \int_0^\theta x^2 f(x|\theta)dx = \int_0^\theta x^2 \frac{1}{\theta}dx$

$= \frac{x^3}{3}\Big|_0^\theta \times \frac{1}{\theta} = \frac{\theta^3}{3} \times \frac{1}{\theta} = \theta^2/3$

Set 2nd moment = $\underline{\text{empirical}}$ second moment.
sample.

$\Rightarrow \frac{\theta^2}{3} = \frac{\sum x_i^2}{n} \Rightarrow \theta^2 = \frac{n}{3}\frac{\sum x_i^2}{n}$ or $\boxed{\theta = \frac{n}{3}\frac{\sum x_i^2}{n}}$

b) $\underbrace{f(\vec{x}|\theta)}_{L(\theta|\vec{x})} = \prod_{i=1}^n f(x_i|\theta) = \frac{1}{\theta^n}\mathbb{I}_{\{x_i \in [0,\theta]\}}$, $\theta > 0$

We want to maximize $L(\theta|\vec{x})$. But $\frac{1}{\theta^n}$ is monotonic decreasing for $\theta > 0$
$\therefore$ It achieves maximum when θ is lowest. But $\theta > \max_i x_i$ as others we get $L(\theta|\vec{x}) = 0$ which is definitely not maximum { but a minimum as $L(\theta|\vec{x}) \geq$
$\therefore$ we set $\underline{\hat{\theta}_{MLE} = \max_i x_i}$ to maximize $L(\theta|\vec{x})$. Ans

c) $E[\hat{\theta}_{MLE}] \neq \theta/2$. Hence $\hat{\theta}_{MLE}$ is biased.

$E[\hat{\theta}_{MLE}] = \frac{n-1}{n}\theta$.

$E[(\hat{\theta}-\theta)^2] = E\left(\theta\frac{n-1}{n} - \theta\right)^2 = E\left(\theta^2\left(\frac{+1}{n^2}\right)^2\right)$

$= \int \frac{\theta^2}{n^2} d\theta = \int_0^\theta \frac{\theta}{n^2}d\theta = \frac{1}{n^2}$

Name: ID:

3. (20 pts) Consider a 1-dimensional linear regression model where $y_i \in \mathbb{R}, 1 \le i \le n$ are independent random variables satisfying $y_i = \theta x_i + \varepsilon_i$, where $\varepsilon_i \sim N(0,1)$ are i.i.d. random variables, $x_i \in \mathbb{R}$, and $\theta \in \mathbb{R}$ is an unknown parameter.

   (a) (8 pts) Assume $\sum_{i=1}^n x_i^2 > 0$. Find the likelihood function $L(\theta)$ given the data $\{(x_i, y_i)\}_{i=1}^n$ and compute the maximum likelihood estimator of θ explicitly. Justify your answer.

   (b) (7 pts) Assume the prior distribution of θ is given by $\theta \sim N(\theta_0, 1)$ where $\theta_0 \in \mathbb{R}$. Show that the posterior distribution for θ is a normal distribution. Justify your answer.

   (c) (5 pts) Find the posterior mean estimator and the maximum a posteriori estimation estimator for θ. Justify your answer.

a) $L(\theta|x) = L(\theta) = \cancel{\prod_i f(y_i|\theta) = \prod_i (y_i - \theta x_i)}$

$\varepsilon_i = y_i - \theta x_i \quad \| \quad f((x_i, y_i)|\theta) = f(\varepsilon_i|\theta)$

$\therefore L(\theta) = \prod_i f(\varepsilon_i|\theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \prod_i \exp\{-\frac{1}{2}(\varepsilon_i^2)\} = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\{\sum_{i=1}^n -\frac{1}{2}(y_i - \theta x_i)^2\}$

$\arg\max_\theta L(\theta) = \arg\max_\theta \sum_{i=1}^n -\frac{1}{2}(y_i - \theta x_i)^2 = \arg\min_\theta \underbrace{\frac{1}{2}\sum_{i=1}^n (y_i - \theta x_i)^2}_{\text{Call this } Q(\theta)}$.

$\frac{dQ(\theta)}{d\theta} = -\frac{1}{2}\sum(y_i - \theta x_i)x_i = 0$

$\Rightarrow \sum_i y_i x_i - \theta \sum x_i^2 = 0 \Rightarrow \theta = \frac{\sum y_i x_i}{\sum x_i^2} = \frac{\vec{y}\cdot\vec{x}}{\vec{x}\cdot\vec{x}} = \underbrace{(\vec{y}\cdot\hat{x})\hat{x}}_{\text{projection of } \vec{y} \text{ onto } i}$

b) It has to be as the Normal $N(\theta_0, 1)$ is conjugate prior to $N(\theta, 1)$.

$f(\theta|\vec{x}) = \underbrace{f(\vec{x}|\theta)}_{\substack{L(\theta|x)\\ \text{or } L(\theta)}} \cdot \pi(\theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\{-\frac{1}{2}\sum(y_i - \theta x_i)^2\} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\theta-\theta_0)^2}$

$= \left(\frac{1}{\sqrt{2\pi}}\right)^{n+1} \exp\{-\frac{1}{2}[\underbrace{\sum_i (y_i - \theta x_i)^2 + (\theta - \theta_0)^2}_{\sum_i (y_i^2 - 2y_i\theta x_i + \theta^2 x_i^2) + \theta^2 + \theta_0^2 - 2\theta\theta_0}]\}$

This is normal dist. as we just have quadratic of θ.

$H(\theta) = \underbrace{\sum_{i=1}^n (y_i^2 - 2y_i\theta x_i + \theta^2 x_i^2) + \theta^2 + \theta_0^2 - 2\theta\theta_0}$

mean is when this function of θ peaks

The mode is the same. because mode = mean for Normal.

$\frac{\partial H}{\partial \theta} = 0 \Rightarrow \cancel{2}\sum(y_i x_i + x_i^2\theta) + \cancel{2}\theta - \cancel{2}\theta_0 = 0$

$\Rightarrow \boxed{\hat{\theta} = \theta_0 + \sum y_i x_i / 1 + \sum_i x_i^2}$

c) posterior mean = MAP because for normal, mode = mean $\therefore \hat{\theta} = \frac{\theta_0 + \sum y_i x_i}{1 + \sum x_i^2}$

$\cancel{\hat{\theta} = \theta_0 / 1 + \sum y_i x_i + x_i^2}$

Name: ID: