408 Midterm 2 Solutions[1]

1. Question 1

(a) The negation of the statement
"There exists an integer j such that $j^2 - j < 8$" is:
"For every integer j, we have $j^2 - j \geq 8$."

TRUE. By the rules of negation, "There exists" is negated to"For every," and the inequality $<$ is negated to $\geq$.

(b)We have

$$\sup_{\lambda \in (0,1)} \lambda^2 = 1.$$

TRUE. Each $\lambda \in (0,1)$ satisfies $\lambda^2 \leq 1$. And for any number $t < 1$, there is some $\lambda \in (0,1)$ such that $\lambda^2 > t$ (e.g. we can choose $\lambda = \sqrt{t} + (1 - \sqrt{t})/2$). So, 1 is the smallest number satisfying $\lambda^2 \leq 1$ for all $\lambda \in (0,1)$.

(c) Let Z be a sufficient statistic for $\{f_\theta \colon \theta \in \Theta\}$ and let Y be an unbiased estimator for θ. Define $W := \mathbf{E}_\theta(Y|Z)$. Let $\theta \in \Theta$ with $\mathrm{Var}_\theta(Y) < \infty$. Then

$$\mathrm{Var}_\theta(W) \leq \mathrm{Var}_\theta(Y).$$

TRUE. This is the Rao-Blackwell Theorem.

(d) There can be at most one maximum likelihood estimator. That is, if a maximum likelihood estimator exists, it is unique.

FALSE. Example 4.32 in class (using uniform random variables on $[\theta - 1/2, \theta + 1/2]$) shows that there can be infinitely many MLEs. That is, the MLE might not be unique.

(e) Suppose X is a UMVU for $\theta \in \Theta$, and Y is an estimator for θ. Then

$$\mathbf{E}_\theta(X - \mathbf{E}_\theta X)^2 \leq \mathbf{E}_\theta(Y - \mathbf{E}_\theta Y)^2, \qquad \forall\, \theta \in \Theta.$$

(In answering this question, you can freely use a result from the homework.)

FALSE. On the homework (quiz 4), we found an estimator Y such that $\mathbf{E}_\theta(X - \mathbf{E}_\theta X)^2 > \mathbf{E}_\theta(Y - \mathbf{E}_\theta Y)^2$ for some θ. This estimator Y was biased, so it does not contradict the definition of UMVU.

(f) Let $X \colon \Omega \to \mathbf{R}^n$ be a random variable with distribution f_θ. Let $I_X(\theta)$ denotes the Fisher Information of X. Let Y be an unbiased estimator for $\theta \in \Theta$. Suppose

$$\mathrm{Var}_\theta(Y) = \frac{1}{I_X(\theta)}, \qquad \forall\, \theta \in \Theta,$$

Then Y is UMVU for θ.

TRUE. By the Cramér-Rao Theorem, any estimator Z that is unbiased for θ must satisfy $\mathrm{Var}_\theta(Z) \geq \frac{1}{I_X(\theta)}$ $\forall\, \theta \in \Theta$. Since the variance of Y is equal to the last quantity, we get $\mathrm{Var}_\theta(Z) \geq \mathrm{Var}_\theta(Y)$ $\forall\, \theta \in \Theta$, for any estimator Z of θ.

2. Question 2

Let $X_1, \ldots, X_n$ be i.i.d continuous random variables with $\mathbf{E}\,|X_1| < \infty$.

In each question below, simplify your answer to the best of your ability.

Unlike other questions, in this question you can freely use a result from a previous homework concerning general properties of conditional expectation.

- Compute $\mathbf{E}(X_1 \mid X_1)$.
- Compute $\mathbf{E}(X_1 \mid X_2)$.
- Compute $\mathbf{E}(X_1 \mid X_1 + \cdots + X_n)$.

Solution. From Exercise 2(iv) on HW3, $\mathbf{E}(X_1|X_1) = X_1$.

From Exercise 2(v) on HW3, $\mathbf{E}(X_1|X_2) = \mathbf{E}X_1$, since X_1, X_2 are independent. (Since X_2 is independent of X_1, $\mathbf{E}(X_1|X_2 = x_2) = \mathbf{E}X_1$, so $\mathbf{E}(X_1|X_2) = \mathbf{E}X_1$.)

Since the random variables are i.i.d., for any $1 \leq k < \ell \leq n$, the joint distribution of $(X_k, \sum_{i=1}^n X_i)$ is equal to the joint distribution of $(X_\ell, \sum_{i=1}^n X_i)$. So, by the definition of conditional expectation,

$$\mathbf{E}(X_k| \sum_{i=1}^n X_i) = \mathbf{E}(X_\ell| \sum_{i=1}^n X_i).$$

Therefore, by Exercise 2(iv) on HW3,

$$W := \mathbf{E}(X_1| \sum_{i=1}^n X_i) = \frac{1}{n}\sum_{j=1}^n \mathbf{E}(X_j| \sum_{i=1}^n X_i) = \frac{1}{n}\mathbf{E}(\sum_{j=1}^n X_j| \sum_{i=1}^n X_i) = \frac{1}{n}\sum_{i=1}^n X_i.$$

3. Question 3

Let $\theta \in \mathbf{R}$ be an unknown parameter. Consider the density

$$f_\theta(x) := \begin{cases} e^{-(x-\theta)}, & \text{if } x \geq \theta \\ 0, & \text{if } x < \theta. \end{cases}$$

Suppose $X_1, \ldots, X_n$ is a random sample of size n, such that X_i has density f_θ for all $1 \leq i \leq n$.

Show that $X_{(1)} = \min_{1\leq i\leq n} X_i$ is a sufficient statistic for θ.

Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. If it occurs that $\min_{1\leq i\leq n} x_i < \theta$, then some $1 \leq i \leq n$ satisfies $x_i < \theta$, so $f_\theta(x_i) = 0$ and the joint PDF $\prod_{i=1}^n f_\theta(x_i)$ is also zero. On the other hand, if it occurs that $\min_{1\leq i\leq n} x_i \geq \theta$, then all $1 \leq i \leq n$ satisfy $x_i \geq \theta$, and the joint PDF $\prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n e^{-(x_i-\theta)}$. That is, we can write

$$\prod_{i=1}^n f_\theta(x_i) = 1_{\{\min_{1\leq i\leq n} x_i \geq \theta\}} \cdot \prod_{i=1}^n e^{-(x_i-\theta)} = e^{n\theta} 1_{\{\min_{1\leq i\leq n} x_i \geq \theta\}} \cdot \prod_{i=1}^n e^{-(x_i)}.$$

The factorization theorem then implies that $t(x) := \min_{1\leq i\leq n} x_i$ gives our sufficient statistic $Y = t(X_1, \ldots, X_n)$, since if we define $g_\theta(z) := e^{n\theta} 1_{\{z\geq\theta\}}$ and $h(x) := \prod_{i=1}^n e^{-(x_i)}$, then we have written the joint PDF as

$$\prod_{i=1}^n f_\theta(x_i) = g_\theta(t(x))h(x), \qquad \forall\, x \in \mathbf{R}^n, \quad \forall\, \theta \in \mathbf{R}.$$

4. Question 4

Suppose $X_1, \ldots, X_n$ are i.i.d. random variables, and X_1 has distribution

$$f_\theta(x) = \theta x^{\theta-1}, \qquad \forall\, 0 < x < 1,$$

where $\theta > 0$ is unknown.

Find an MLE Y_n of θ. As usual, you should justify your answer.

Solution. The joint distribution of $X_1, \ldots, X_n$ is

$$\ell(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}.$$

Taking the log and then taking a derivative of that,

$$\frac{d}{d\theta} \log \ell(\theta) = \frac{d}{d\theta}[n \log \theta + (\theta - 1)\sum_{i=1}^{n} \log x_i] = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i.$$

Setting this quantity to zero, we find a single critical point

$$\theta = -\frac{1}{n}\sum_{i=1}^{n} \log x_i.$$

When θ is less than this value, $\frac{d}{d\theta} \log \ell(\theta) > 0$ and when θ is greater than this value, $\frac{d}{d\theta} \log \ell(\theta) < 0$. That is, $\log \ell(\theta)$ increases, then decreases. So, $\log \ell(\theta)$ has a unique maximum at the value $\theta = -\frac{1}{n}\sum_{i=1}^{n} \log x_i$.

5. Question 5

Let X be a geometric distributed random variable with unknown parameter p where $p \in \{1/3, 2/3\}$. (So, $\mathbf{P}(X = k) = (1-p)^{k-1}p$ for all integers $k \geq 1$.) Suppose we want to test the hypothesis H_0 that $p = 1/3$ versus the alternative H_1 that $p = 2/3$.

- Explicitly describe the rejection region C of the UMP (uniformly most powerful) test among all hypothesis tests with significance level less than or equal to 5/9.
- Suppose we observe that $X = 1$. Report a p-value for this observation, for the UMP test you found.

Solution. We have

$$\frac{f_{2/3}(k)}{f_{1/3}(k)} = \frac{(1-2/3)^{k-1}(2/3)}{(1-1/3)^{k-1}(1/3)} = \frac{(1/3)^{k-1}(2/3)}{(2/3)^{k-1}(1/3)} = \frac{(1/3)^{k-2}}{(2/3)^{k-2}} = (1/2)^{k-2}.$$

Evidently this quantity decreases as k increases. Plugging in a few values, we get

$$\frac{f_{2/3}(1)}{f_{1/3}(1)} = (1/2)^{-1} = 2, \qquad \frac{f_{2/3}(2)}{f_{1/3}(2)} = (1/2)^{0} = 1, \qquad \frac{f_{2/3}(3)}{f_{1/3}(3)} = (1/2)^{1} = 1/2.$$

So, e.g. the likelihood ratio test with rejection region

$$C = \{k \geq 1 \colon \frac{f_{2/3}(k)}{f_{1/3}(k)} \geq 1\}$$

has significance level

$$\alpha = \mathbf{P}_{1/3}(X \in C) = \mathbf{P}_{1/3}(X = 1 \text{ or } X = 2) = 1/3 + (1-1/3)(1/3) = (1/3)(1+2/3) = 5/9.$$

The Neyman-Pearson Lemma states that the test with rejection region C is UMP among all tests with significance level at most 5/9. A p-value for this test is

$$p(k) = \mathbf{P}_{1/3}\Big(\frac{f_{2/3}(X)}{f_{1/3}(X)} \geq \frac{f_{2/3}(k)}{f_{1/3}(k)}\Big).$$

In the case $k = 1$, we get

$$p(1) = \mathbf{P}_{1/3}\Big(\frac{f_{2/3}(X)}{f_{1/3}(X)} \geq \frac{f_{2/3}(1)}{f_{1/3}(1)}\Big) = \mathbf{P}_{1/3}\Big((1/2)^{X-2} \geq 2\Big) = \mathbf{P}_{1/3}(X = 1) = 1/3.$$