IEDA 5270 (Fall 2023)

Marking Notes for Midterm Exam

Question 1

(a) (5 points) The joint density of $\mathbf{X}$ can be written as

$$\begin{aligned} f(\mathbf{x}|\boldsymbol{\theta}) &= \exp\left\{-\sum_{i=1}^n \left(\frac{x_i-\mu}{\sigma}\right)^4 - n\xi(\boldsymbol{\theta})\right\} \\ &= \exp\left\{-\frac{1}{\sigma^4}\sum_{i=1}^n x_i^4 + \frac{4\mu}{\sigma^4}\sum_{i=1}^n x_i^3 - \frac{6\mu^2}{\sigma^4}\sum_{i=1}^n x_i^2 + \frac{4\mu^3}{\sigma^4}\sum_{i=1}^n x_i - \frac{n\mu^4}{\sigma^4} - n\xi(\boldsymbol{\theta})\right\} \\ &=: h(\mathbf{x})c(\boldsymbol{\theta})\exp\{\mathbf{w}(\boldsymbol{\theta})^\top \mathbf{T}(\mathbf{x})\}, \end{aligned}$$

with $h(\mathbf{x}) := 1$, $c(\boldsymbol{\theta}) := \exp\{-\frac{n\mu^4}{\sigma^4} - n\xi(\boldsymbol{\theta})\}$, $\mathbf{w}(\boldsymbol{\theta}) := (-\frac{1}{\sigma^4}, \frac{4\mu}{\sigma^4}, -\frac{6\mu^2}{\sigma^4}, \frac{4\mu^3}{\sigma^4})^\top$ and $\mathbf{T}(\mathbf{x}) := (\sum_{i=1}^n x_i^4, \sum_{i=1}^n x_i^3, \sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i)^\top$. By definition, it belongs to an exponential family.

> **Marking Notes:**
> - 2 points for correct joint pdf.
> - 1 point for correctly expanding the 4th power.
> - 2 points for correctly identifying $w_i(\boldsymbol{\theta})$ and $T_i(\mathbf{x})$ etc.
>   ** 1 point if wrong sign in $w_i(\boldsymbol{\theta})T_i(\mathbf{x})$.

(b) (5 points) Let $\mathbf{x}$, $\mathbf{y}$ be any realizations of the joint distribution. Then

$$\begin{aligned} &\frac{f(\mathbf{x}|\boldsymbol{\theta})}{f(\mathbf{y}|\boldsymbol{\theta})} \text{ free of } \boldsymbol{\theta} \\ \iff &\exp\Bigg\{-\frac{1}{\sigma^4}\left(\sum_{i=1}^n x_i^4 - \sum_{i=1}^n y_i^4\right) + \frac{4\mu}{\sigma^4}\left(\sum_{i=1}^n x_i^3 - \sum_{i=1}^n y_i^3\right) \\ &\quad - \frac{6\mu^2}{\sigma^4}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right) + \frac{4\mu^3}{\sigma^4}\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\Bigg\} \text{ free of } \boldsymbol{\theta} \\ \iff &\sum_{i=1}^n x_i^k = \sum_{i=1}^n y_i^k \text{ for } k=1,2,3,4 \iff \mathbf{T}(\mathbf{x}) = \mathbf{T}(\mathbf{y}). \end{aligned}$$

From the Checking Rule, $\mathbf{T}(\mathbf{X}) = (\sum_{i=1}^n X_i^4, \sum_{i=1}^n X_i^3, \sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i)^\top$ is minimal sufficient.

> **Marking Notes:**
> - 2 points for correctly specifying $\mathbf{T}(\mathbf{X})$.
> - 3 points for correctly arguing minimal sufficiency.
>   ** 1 point if not correctly showing minimality, this includes claiming full-rank EF (this EF is NOT of full-rank, which can only imply sufficiency).
>   ** 2 points if using Checking Rule without "iff".

Question 2

(a) (2 points) The likelihood function of the sample is

$$L(p|\mathbf{x}) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}, \quad p \in \{0, 0.5, 1\},$$

so the MLE is

$$\hat{p}(\mathbf{X}) = \arg\max_{p\in\{0,0.5,1\}} L(p|\mathbf{X}) = \begin{cases} 0, & \text{if } \sum_{i=1}^n X_i = 0, \\ 0.5, & \text{if } 0 < \sum_{i=1}^n X_i < n, \\ 1, & \text{if } \sum_{i=1}^n X_i = n, \end{cases}$$
$$= 0.5\,\mathbb{1}\left\{\sum_{i=1}^n X_i > 0\right\} + 0.5\,\mathbb{1}\left\{\sum_{i=1}^n X_i = n\right\}.$$

Marking Notes:
- 1 point for correct joint likelihood function.
- 1 point for correct $\hat{p}$.

(b) (3 points) Since $\sum_{i=1}^n X_i \sim \text{Bin}(n,p)$, we know that

$$\mathbb{E}[\hat{p}] = 0.5\left[1 - \mathbb{P}\left(\sum_{i=1}^n X_i = 0\right)\right] + 0.5\,\mathbb{P}\left(\sum_{i=1}^n X_i = n\right)$$
$$= 0.5(1-(1-p)^n) + 0.5p^n = p$$

for all $p \in \{0, 0.5, 1\}$ and for all n. Thus $\text{bias}(\hat{p}) = \mathbb{E}[\hat{p} - p] = 0 = \lim_{n\to\infty} \text{bias}(\hat{p})$.

Marking Notes:
- 2 points for correctly finding the $\mathbb{E}[\hat{p}]$ for the $\hat{p}$ in (a).
  ** 1 point if $\hat{p}$ in (a) is incorrect.
- 1 point for both correct bias and asymptotic bias.

(c) (3 points) Note that $\hat{p}^2 = 0.25\,\mathbb{1}\{\sum_{i=1}^n X_i > 0\} + 0.75\,\mathbb{1}\{\sum_{i=1}^n X_i = n\}$. Hence

$$\text{MSE}(\hat{p}) = \text{Var}(\hat{p}) + \text{bias}(\hat{p})^2 = \mathbb{E}[\hat{p}^2] - p^2 + 0$$
$$= 0.25(1-(1-p)^n) + 0.75p^n - p^2$$
$$= \begin{cases} 0, & \text{if } p \in \{0,1\} \\ 0.5^{n+1} & \text{if } p = 0.5 \end{cases}.$$

Marking Notes:
- 2 points for correctly finding the $\text{Var}(\hat{p})$ for the $\hat{p}$ in (a).
  ** 1 point if $\hat{p}$ in (a) is incorrect.
- 1 point for correct MSE.

(d) (2 points) Since $\lim_{n\to\infty} \text{MSE}(\hat{p}) = 0$ for all $p \in \{0, 0.5, 1\}$, the MLE $\hat{p}$ is consistent.

Marking Notes:
- 1 point for claiming consistency for all $p \in \{0, 0.5, 1\}$.
- 1 point for correct justification.
  ** 0 points if the justification uses an incorrect MSE in (c).

Question 3

(a) (4 points) Direct integration gives $\mathbb{E}[X_1] = \int_0^1 ax^a\,\mathrm{d}x = \frac{a}{a+1}$, $\mathbb{E}[X_1^2] = \int_0^1 ax^{a+1}\,\mathrm{d}x = \frac{a}{a+2}$, and hence $\text{Var}(X_1) = \frac{a}{a+2} - (\frac{a}{a+1})^2 = \frac{a}{(a+2)(a+1)^2}$.

Marking Notes:
- 2 points for correct $\mathbb{E}[X_1]$.
- 2 points for correct $\text{Var}(X_1)$.

(b) (6 points) The log-likelihood function of a single observation x and its derivatives with respect to a are, respectively,

$$l(a|x) = \log(a) + (a-1)\log(x),$$
$$\frac{\partial l(a|x)}{\partial a} = \frac{1}{a} + \log(x), \qquad \frac{\partial^2 l(a|x)}{\partial a^2} = -\frac{1}{a^2}.$$

Hence, the Fisher Information for a in a single observation is $\mathcal{I}(a) = -\mathbb{E}[-\frac{1}{a^2}] = \frac{1}{a^2}$.

Let $T(\mathbf{X})$ be any unbiased estimator of $m(a) = \frac{a}{a+1} = \mathbb{E}[X_1]$. Then Cramer-Rao Inequality states that its variance has a lower bound

$$\mathrm{Var}\left(T(\mathbf{X})\right) \geq \frac{[m'(a)]^2}{n\mathcal{I}(a)} = \frac{a^2}{n(a+1)^4} = \text{CRLB}.$$

Marking Notes:
There are several acceptable alternatives of finding CRLB, e.g. via $\mathcal{I}_n(a)$, $\mathcal{I}(\mu)$, $\mathcal{I}_n(\mu)$ etc.
For those using $\mathcal{I}(a)$ or $\mathcal{I}_n(a)$:
- 1 point for correct likelihood or log-likelihood.
- 2 points for correct Fisher Information.
  ** 1 point if minor calculation mistake.
- 1 point for $m(a) = \frac{a}{a+1}$ and $m'(a) = (a+1)^{-2}$.
- 2 points for correct CRLB.
  ** 1 point if $m(a)$ is completely omitted and all other components are correct.
  ** 1 point if minor calculation mistake.

For those using $\mathcal{I}(\mu)$ or $\mathcal{I}_n(\mu)$:
- 1 point for $l(\mu|x) = \log(\frac{\mu}{1-\mu}) + \frac{2\mu-1}{1-\mu}\log(x)$.
- 2 points for $l''(\mu|x) = -\frac{1}{\mu^2} + \frac{1}{(1-\mu)^2} + \frac{2}{(1-\mu)^3}\log(x)$.
  ** 2 points if minor calculation mistake.
  ** 1 point if only $l'(\mu|x) = \frac{1}{\mu} + \frac{1}{1-\mu} + \frac{1}{(1-\mu)^2}\log(x)$ is correct.
- 1 point for $\mathbb{E}[\log(X)] = -\frac{1}{a} = \frac{\mu-1}{\mu}$.
- 2 points for CRLB $= \frac{1}{n\mathcal{I}(\mu)} = \frac{\mu^2(1-\mu)^2}{n}$.
  ** 1 point if mistake occurs in simplifying $\mathbb{E}[l''(\mu|X)] = \frac{1}{\mu^2(1-\mu)^2}$.

(c) (4 points) $\mathrm{Var}(\bar{X}) = \frac{1}{n}\mathrm{Var}(X_1) = \frac{a^2}{na(a+2)(a+1)^2}$. Since $a(a+2) < (a+1)^2$, we have $\mathrm{Var}(\bar{X}) >$ CRLB, i.e. it does not attain the CRLB.

Marking Notes:
- 1 point for showing $\mathrm{Var}(\bar{X}) = \frac{1}{n}\mathrm{Var}(X_1)$.
- 2 points for proving strict inequality for correct $\mathrm{Var}(\bar{X})$ and CRLB.
  ** 1 point if either $\mathrm{Var}(X_1)$ or CRLB is incorrect due to minor calculation mistake.
  ** 0 point if either $\mathrm{Var}(X_1)$ or CRLB is incorrect.
- 1 point for concluding "does not attain CRLB".

(d) (6 points) The parameter space is $a \in (0, \infty)$. Note that the joint density of $\mathbf{X}$ has the form $f(\mathbf{x}|a) = a^n(\prod_{i=1}^n x_i)^{a-1} = a^n \exp\{(a-1)\sum_{i=1}^n \log(x_i)\}$ for $x_i \in [0,1]$, which belongs to a 1-parameter exponential family of full-rank, thus $-\sum_{i=1}^n \log(X_i)$ is complete and sufficient for a.

By Lehmann-Scheffe/the conditioning method, the UMVUE for $\mathbb{E}[X_1]$ is $\mathbb{E}[X_1|-\sum_{i=1}^n \log(X_i)]$. Now consider the distribution of $-\log(X_i)$. Notice that for $y > 0$, the cdf of $-\log(X_i)$ is

$\mathbb{P}(-\log(X_i) \le y) = 1 - \mathbb{P}(X_i < e^{-y}) = 1 - e^{-ay}$, so $-\log(X_i)$ are iid and follows exponential distribution with mean $1/a$. As such, the sum $-\sum_{i=2}^n \log(X_i) \sim \text{Gamma}(n-1, a)$ is independent with $-\log(X_1) \sim \text{Gamma}(1, a)$.

Now our task boils down to find the conditional distribution of U given $U+V$, where $U \sim \text{Gamma}(1, a)$ and $V \sim \text{Gamma}(n-1, a)$ are independent. The joint density of (U, V) is

$$\begin{aligned} f_{U,V}(u, v) = f_U(u) f_V(v) &= a e^{-au} \frac{a^{n-1}}{\Gamma(n-1)} v^{n-2} e^{-av} \mathbb{1}\{u > 0, v > 0\} \\ &= \frac{a^n}{\Gamma(n-1)} v^{n-2} e^{-a(u+v)} \mathbb{1}\{u > 0, v > 0\}. \end{aligned}$$

By transformation, the joint density of $(U, U+V)$ is

$$f_{U,U+V}(u, w) = \frac{a^n}{\Gamma(n-1)} (w-u)^{n-2} e^{-aw} \mathbb{1}\{w > u > 0\}.$$

So it follows that the conditional density of U given $U+V = w$ is

$$\begin{aligned} f_{U|U+V}(u|w) = \frac{f_{U,U+V}(u, w)}{f_{U+V}(w)} &= \frac{a^n (w-u)^{n-2} e^{-aw}}{\Gamma(n-1)} \frac{\Gamma(n)}{a^n w^{n-1} e^{-aw}} \mathbb{1}\{w > u > 0\} \\ &= \frac{n-1}{w} \left(1 - \frac{u}{w}\right)^{n-2} \mathbb{1}_{(0,w)}(u). \end{aligned}$$

Therefore,

$$\mathbb{E}\left[X_1 \middle| -\sum_{i=1}^n \log(X_i) = w\right] = \mathbb{E}\left[e^{-U} \middle| U+V = w\right] = \frac{n-1}{w} \int_0^w e^{-u} \left(1 - \frac{u}{w}\right)^{n-2} \mathrm{d}u,$$

i.e. the UMVUE of $\mathbb{E}[X_1]$ is $\frac{n-1}{-\sum_{i=1}^n \log(X_i)} \int_0^{-\sum_{i=1}^n \log(X_i)} e^{-u} (1 - \frac{u}{-\sum_{i=1}^n \log(X_i)})^{n-2} \, \mathrm{d}u$.

Marking Notes:

- 1 point for showing full-rank EF.
  - ** 1 point for showing EF.
- 1 point for identifying complete sufficient statistics $\sum_{i=1}^n \log(X_i)$.
  - ** 1 point if did not show that $\sum_{i=1}^n \log(X_i)$ is complete (e.g. did not mention full-rank EF).
  - ** 1 point if mentioned but did not find the correct complete sufficient statistics.
- 2 points for using Lehmann-Scheffe to claim that $\mathbb{E}[X_1 | \sum_{i=1}^n \log(X_i)]$ is UMVUE.
  - ** 2 points for describing the UMVUE as "the estimator $g(\sum_{i=1}^n \log(X_i))$ that satisfies $\mathbb{E}[g(\sum_{i=1}^n \log(X_i))] = \mu$".
  - ** 1 point if did not show that $\sum_{i=1}^n \log(X_i)$ is complete.
  - ** 1 point if mentioned but did not find the correct complete sufficient statistics.
- 2 points for unstructing $\mathbb{E}[X_1 | \sum_{i=1}^n \log(X_i)]$ into integral form.

Question 4

(a) (5 points) The LRT statistic is

$$\lambda(\mathbf{x}) = \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} = \frac{e^n \prod_{i=1}^n x_i!}{2^{n+\sum_{i=1}^n x_i}}.$$

Marking Notes:
- 3 points for correct joint pdf of f_0 amd f_1.
  ** 2 points if minor calculation mistake.
  ** 1 point if only one f_i is correct.
- 2 points for correct LRT statistic.
  ** Accept using $f_0(\mathbf{x})/f_1(\mathbf{x})$ as the test statistic.

(b) (5 points) The rejection region is $\{\mathbf{x} : \frac{e^n \prod_{i=1}^n x_i!}{2^{n+\sum_{i=1}^n x_i}} > c\}$ for some constant c.

Marking Notes:
- 5 points for $\lambda(\mathbf{x}) > c$.
  ** 3 points if sign is reversed.
  ** 5 points if minor calculation mistake in calculating $\lambda(\mathbf{x})$ occurs in (a).

Question 5

(a) (5 points) The likelihood function of the random sample is

$$L(\theta|\mathbf{x}) = \prod_{i=1}^n \frac{\exp\left\{-\frac{x_i}{\theta}\right\}}{\theta} = \frac{\exp\left\{-\frac{\sum_{i=1}^n x_i}{\theta}\right\}}{\theta^n}$$

for $\theta \in \Theta = \mathbb{R}^+$. To obtain the maximizer, differentiate the log-likelihood $l(\theta|\mathbf{x}) = -n\log(\theta) - \frac{\sum_{i=1}^n x_i}{\theta}$ gives

$$l'(\theta) = -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i}{\theta^2} \begin{cases} > 0 & \text{if } \theta < \frac{\sum_{i=1}^n x_i}{n} \\ = 0 & \text{if } \theta = \frac{\sum_{i=1}^n x_i}{n} \\ < 0 & \text{if } \theta > \frac{\sum_{i=1}^n x_i}{n} \end{cases}.$$

Thus, $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ is the unconstrained MLE of θ, i.e. $\sup_{\theta\in\Theta} L(\theta|\mathbf{x}) = L(\bar{x}|\mathbf{x})$. On the other hand,

$$\sup_{\theta\in\Theta_0} L(\theta|\mathbf{x}) = \begin{cases} L(\bar{x}|\mathbf{x}) & \text{if } \theta_0 > \bar{x} \\ L(\theta_0|\mathbf{x}) & \text{if } \theta_0 \le \bar{x} \end{cases}.$$

So the LR test statistic is given by

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta\in\Theta_0} L(\theta|\mathbf{x})}{\sup_{\theta\in\Theta} L(\theta|\mathbf{x})} = \begin{cases} 1 & \text{if } \theta_0 > \bar{x} \\ \left(\frac{\bar{x}}{\theta_0}\right)^n \exp\left\{-n\left(\frac{\bar{x}}{\theta_0} - 1\right)\right\} & \text{if } \theta_0 \le \bar{x} \end{cases},$$

and the rejection region at size α is

$$R_\alpha = \left\{\mathbf{x} : \left(\frac{\bar{x}}{\theta_0}\right)^n \exp\left\{-n\left(\frac{\bar{x}}{\theta_0} - 1\right)\right\} < c_\alpha\right\}.$$

Since $y^n e^{-n(y-1)}$ decreases when $y \ge 1$, R_α is equivalent to

$$R_\alpha \equiv \left\{\mathbf{x} : \frac{\bar{x}}{\theta_0} > k_\alpha\right\}$$

for some k_α. To find an explicit form of k_α, notice that $X_i \overset{\text{iid}}{\sim} \text{Exp}(\theta)$. Then $\sum_{i=1}^n X_i \sim \text{Gamma}(n,\theta)$ and we can write $\frac{2n\bar{X}}{\theta} \sim \text{Gamma}(\frac{2n}{2}, 2) \equiv \chi^2_{2n}$. Consequently, we need k_α to satisfy

$$\alpha = \sup_{\theta\le\theta_0} \mathbb{P}_\theta(\mathbf{X} \in R_\alpha) = \sup_{\theta\le\theta_0} \mathbb{P}\left(V > \frac{2n\theta_0 k_\alpha}{\theta}\right) = \mathbb{P}(V > 2nk_\alpha),$$

where $V \sim \chi^2_{2n}$ is ancillary of θ. Thus $k_\alpha = \chi^2_{2n}(\alpha)/(2n)$, where $\chi^2_{2n}(\alpha)$ is the $(1-\alpha)$-th quantile of the pdf of χ^2_{2n}.

Marking Notes:
- 1 point for joint likelihood function.
- 1 point for unconstrained MLE.
- 1 point for MLE under H_0.
- 2 points for correct LRT statistic associated with rejection regions.
  ** 1 points if using $\sup_{\Theta_1}$ in the denominator, but the direction of rejection region is correct.
- Give at most 2 points if θ_1 appears in the likelihood ratio.

(b) (5 points) X follows an exponential family with strictly increasing $\eta(\theta) = -\frac{1}{\theta}$ and the test in part (a) is equivalent to a test in the form of $\varphi(\mathbf{x}) = \mathbb{1}\{\sum_{i=1}^n x_i > c\}$. Result thus follows from page 37 of Lecture Note 5.

Marking Notes:
- 2 points for establishing equivalence in the test using $\lambda(\mathbf{X})$ and that using $T(\mathbf{X})$.
- 1 point for MLR/exponential family with increasing $\eta(\theta)$.
- 2 points for calling appropriate theorem for UMP one-sided test.

Question 6

(a) (10 points) We will prove the contraposition of the original statement (i.e. we will show that "$c(\alpha_1) < c(\alpha_2) \implies \alpha_2 \le \alpha_1$"). Suppose $c(\alpha_1) < c(\alpha_2)$. Then

$$\begin{aligned}
\varphi_{\alpha_2}(X) &= \mathbb{1}\{f_1(X) > c(\alpha_2)f_0(X)\} + \gamma(\alpha_2)\mathbb{1}\{f_1(X) = c(\alpha_2)f_0(X)\} \\
&\le \mathbb{1}\{f_1(X) > c(\alpha_2)f_0(X)\} + \mathbb{1}\{f_1(X) = c(\alpha_2)f_0(X)\} \\
&= \mathbb{1}\{f_1(X) \ge c(\alpha_2)f_0(X)\} \\
&\le \mathbb{1}\{f_1(X) > c(\alpha_1)f_0(X) \text{ or } f_1(X) = f_0(X) = 0\} \\
&\le \mathbb{1}\{f_1(X) > c(\alpha_1)f_0(X)\} + \mathbb{1}\{f_0(X) = 0\} \\
&\le \mathbb{1}\{f_1(X) > c(\alpha_1)f_0(X)\} + \gamma(\alpha_1)\mathbb{1}\{f_1(X) = c(\alpha_1)f_0(X)\} + \mathbb{1}\{f_0(X) = 0\} \\
&= \varphi_{\alpha_1}(X) + \mathbb{1}\{f_0(X) = 0\}.
\end{aligned}$$

Take expectations on both sides under f_0 yields $\mathbb{E}_0[\varphi_{\alpha_2}(X)] \le \mathbb{E}_0[\varphi_{\alpha_1}(X)] + 0$, where L.H.S. and R.H.S. are α_2 and α_1 respectively, so the result follows.

Marking Notes:
- Give at most 6 points for "heuristic" solutions.
- Give at most 7 points for proofs that does not handle $\mathbb{P}_0(f_1(X) = c(\alpha)f_0(X))$ appropriately. This probability is not negligible even if both f_0 and f_1 have continuous pdf.
- -1 point for not elaborating $\varphi_{\alpha_1}(X) \ge \varphi_{\alpha_2}(X)$ under f_0.
- Accept proofs that assume $f_0(X), f_1(X) > 0$.

(b) (10 points) Notice that $\varphi_{\alpha_2}(X)$ is the UMP size α_2 test for testing $H_0 : f = f_0$ against $H_1 : f = f_1$, so its power is greater than or equal to any other test with level at most α_2. In particular, since $\varphi_{\alpha_1}(X)$ is a test of level $\alpha_1 < \alpha_2$, we have $\mathbb{E}_1[\varphi_{\alpha_2}(X)] \ge \mathbb{E}_1[\varphi_{\alpha_1}(X)]$.

Then, our goal boils down to show that the inequality is strict. Suppose $\mathbb{E}_1[\varphi_{\alpha_2}(X)] = \mathbb{E}_1[\varphi_{\alpha_1}(X)]$, then $\varphi_{\alpha_1}(X)$ is also a UMP size α_2 test. By Neyman-Pearson Lemma/uniqueness of UMP test, it follows that $\varphi_{\alpha_1}(X) = 1 \iff \varphi_{\alpha_2}(X) = 1$ for $X \sim f_i, i = 0, 1$. i.e.

$$f_1(X) > c(\alpha_1)f_0(X) \iff f_1(X) > c(\alpha_2)f_0(X) \text{ for } X \sim f_i, i = 0, 1.$$

Since from part (a) we know that $c(\alpha_1) \geq c(\alpha_2)$, the statements above can be satisfied if and only if

$$\mathbb{P}_i(c(\alpha_1)f_0(X) \geq f_1(X) \geq c(\alpha_2)f_0(X)) = 0,\ i = 0, 1.$$

Now, take expectations of $\varphi_{\alpha_1}(X)$ and $\varphi_{\alpha_2}(X)$ under f_0. Then

$$\begin{aligned}
\alpha_1 &= \mathbb{E}_0[\varphi_{\alpha_1}(X)] \\
&= \mathbb{P}_0(f_1(X) > c(\alpha_1)f_0(X)) + \gamma(\alpha_1) \cdot 0 + 0 \\
&= \mathbb{P}_0(f_1(X) > c(\alpha_2)f_0(X)) + \gamma(\alpha_2) \cdot 0 + 0 \\
&= \mathbb{E}_0[\varphi_{\alpha_2}(X)] = \alpha_2,
\end{aligned}$$

in which we will arrive at a contradiction with $\alpha_2 > \alpha_1$. Thus $\mathbb{E}_1[\varphi_{\alpha_2}(X)] > \mathbb{E}_1[\varphi_{\alpha_1}(X)]$ or $\mathbb{E}_1[1 - \varphi_{\alpha_2}(X)] < \mathbb{E}_1[1 - \varphi_{\alpha_1}(X)]$, as desired.

Marking Notes:

- Give at most 7 points for proofs that does not handle $\mathbb{P}_1(f_1(X) = c(\alpha)f_0(X))$ appropriately.
- Give at most 5 points for proofs that only show weak inequality.