Problem 1. (30 points) Let $X_1, \ldots, X_n$ be an iid sample from a normal distribution with mean 0 and variance $\theta$; i.e., the common PDF is

$$f_\theta(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/2\theta}, \quad -\infty < x < \infty, \quad \theta > 0.$$ 

1. Find the maximum likelihood estimator $\hat{\theta}_n$.
2. Show that $\hat{\theta}_n$ is unbiased.
3. State the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$.

\[ L(\theta) = \left( \frac{1}{2\pi\theta} \right)^{\frac{n}{2}} e^{-\frac{1}{2\theta} \sum_{i=1}^{n} x_i^2} \Rightarrow \ell(\theta) = -\frac{n}{2} \log \theta - \frac{1}{2\theta} \sum_{i=1}^{n} x_i^2 + \text{const} \]

\[ \frac{\partial}{\partial \theta} \ell(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2 \Rightarrow n\frac{1}{\theta} \sum_{i=1}^{n} x_i^2 = \hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x_i^2 \]

\[ \frac{\partial^2}{\partial \theta^2} \ell(\theta) = \frac{n}{2\theta^2} \]

\[ \frac{\partial^2}{\partial \theta^2} \ell(\theta) = \frac{1}{2\theta^2} - \frac{1}{2\theta^3} \sum_{i=1}^{n} x_i^2 \Rightarrow I(\theta) = \frac{1}{2\theta^2} + \frac{1}{2\theta^3} = \frac{1}{2\theta^2} \]

Therefore:

\[ \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, \frac{1}{I(\theta)}) = N(0, \frac{1}{2\theta^2}) \]
Problem 2. (30 points) Let $X_1, \ldots, X_n$ be an iid sample from a Bernoulli distribution with PMF $f_\theta(x) = \theta^x (1 - \theta)^{1-x}$, for $x = 0, 1$ and $\theta \in (0, 1)$.

1. Find the Fisher information $I(\theta)$.
2. Prove that the maximum likelihood estimator $\hat{\theta} = \bar{X}_n$ is efficient.
3. Find the maximum likelihood estimator of $\theta(1 - \theta)$. Is it unbiased? Is it consistent? Prove your answers.

\[
\frac{3}{2\theta} \log f_\theta(\theta) = \frac{3}{2\theta} \left[ \theta \log \theta + (1-\theta) \log (1-\theta) \right] = \frac{2}{\theta} - \frac{1-\theta}{\theta^2}
\]

\[
\frac{3}{2\theta^2} \frac{\partial f_\theta}{\partial \theta} = -\frac{2}{\theta^2} - \frac{1-\theta}{\theta^2}
\]

\[
\mathcal{I}(\theta) = -E \left[ \frac{3}{2\theta} \log f_\theta \right] = \theta \left( \frac{2}{\theta} + \frac{1-\theta}{(1-\theta) \theta^2} \right) = \frac{1}{\theta} + \frac{1}{\theta^2} = \frac{1}{\theta^2}
\]

\[
\frac{\text{Var}(\bar{X})}{n} = \frac{\theta(1-\theta)}{n}.
\]

Since these are equal, estimator $\bar{X}$ is efficient.

\[
\mathcal{I}_0 = \frac{1}{2n\theta} = \frac{\theta(1-\theta)}{n}
\]

3. Let $g(\theta) = \bar{X}(1-\bar{X})$; a continuous function of $\theta$.

So, MLE $\hat{\theta} = g(\bar{X}) = \bar{X}(1-\bar{X})$.

\[
\mathcal{I}_0 \left[ \bar{X}(1-\bar{X}) \right] = \mathcal{I}_0 \left[ \bar{X} - \bar{X}^2 \right]
\]

\[
= \mathcal{I}_0 [\bar{X}] - [\text{Var}(\bar{X}) + \mathcal{I}_0(\theta)]
\]

\[
= \frac{1}{\theta} - \frac{\theta(1-\theta)}{n} - \theta
\]

\[
\neq \mathcal{I}_0(\theta)
\]

\[
\bar{X}(1-\bar{X}) \text{ is not} \ \text{unbiased for } \theta(1-\theta)
\]

Since $\bar{X} \overset{p}{\to} \theta$ by UN, and $g(\theta) = \theta(1-\theta)$ is continuous, the continuous mapping theorem gives $\bar{X}(1-\bar{X}) \overset{p}{\to} \theta(1-\theta)$.

Therefore, $\bar{X}(1-\bar{X})$ is a consistent estimator of $\theta(1-\theta)$.
Problem 3. (20 points) Let $X_1, \ldots, X_n$ be an iid sample from a uniform distribution on the interval $[0, \theta]$, $\theta > 0$, with PDF $f_\theta(x) = \theta^{-1} I_{[0,\theta]}(x)$; here $I$ denotes an indicator function. Find the mean-square error of $\hat{\theta} = \bar{X}$.

If $X_1 \sim \text{Unif}(0, \theta)$, then $E(X) = \frac{\theta}{2}$, $V(X) = \frac{\theta^2}{12}$.

$$\text{MSE}(\bar{X}) = V(\bar{X}) + b^2(\bar{X})$$

$$= \frac{\theta^2}{12n} + \left[ \frac{\theta}{2} - \theta \right]^2$$

$$= \frac{\theta^2}{12n}$$

In fact, $\bar{X}$ is unbiased $\Rightarrow$ MSE = variance.

Problem 4. (Undergrad 10 points / Grad 0 points) Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be unbiased estimators of $\theta$. Show that, for any $w \in [0, 1]$, $\hat{\theta}_3 = w\hat{\theta}_1 + (1 - w)\hat{\theta}_2$ is also unbiased.

$$E\left[ w\hat{\theta}_1 + (1 - w)\hat{\theta}_2 \right] = wE(\hat{\theta}_1) + (1 - w)E(\hat{\theta}_2)$$

$$= w\theta + (1 - w)\theta$$

$$= \theta \checkmark$$
Problem 5. (Undergrad 10 points / Grad 20 points) Graduate students must answer four of the following five questions. Undergraduate students must answer two of the following five questions. Note that the five questions are unrelated.

1. For $X_1, \ldots, X_n$ iid from a Poisson distribution with parameter $\theta$, give three consistent estimators of $\theta$. (No proofs required.)

2. True or False: If smaller mean-square error is preferred, then only unbiased estimators should be considered. Justify your answer.

3. Give an example of a distribution $f_\theta(x)$ for which at least one of the six important regularity conditions R0-R5 fails to hold. Explain.

4. Assuming all the regularity conditions R0-R5 hold, explain what is meant by the statement “maximum likelihood estimators are asymptotically efficient.”

5. Since the Fisher information $I(\theta)$ appears in several of our variance formulas, it would be problematic if $I(\theta) < 0$. Use the definition of Fisher information to argue that $I(\theta) \geq 0$. I would be impressed enough to award extra credit if you could also characterize when $I(\theta) = 0$.

1. $\bar{X}, \frac{S^2}{n} = \frac{1}{n} \sum (X_i - \bar{X})^2, \sqrt{\bar{X} S^2}, \frac{1}{2} \bar{X} + \frac{1}{2} S^2, \ldots$

2. **False.** Often biased estimators have smaller variance than unbiased estimators, so it can happen that allowing bias will actually reduce MSE.

3. $\text{Unif}(s(x)) :$ the support $\{x : f_\theta(x) > 0\}$ depends on $\theta$.

4. The MSE is asymptotically unbiased and its variance matches the CR lower bound.

5. $I(\theta) = \mathbb{E}_{\theta} \left[ \left( \frac{2}{\theta} \frac{d}{\theta} \log f_\theta(x) \right)^2 \right]$. Since $\left( \frac{2}{\theta} \frac{d}{\theta} \log f_\theta(x) \right)^2 \geq 0$, its expected value can’t be negative. By Jensen’s inequality, $\mathbb{E}_{\theta} \left[ \left( \frac{2}{\theta} \frac{d}{\theta} \log f_\theta(x) \right)^2 \right] \geq \left( \mathbb{E}_{\theta} \left( \frac{2}{\theta} \frac{d}{\theta} \log f_\theta(x) \right) \right)^2 = 0$, with equality iff $\frac{2}{\theta} \frac{d}{\theta} \log f_\theta(x) = 0$ with $P$-probability 1. In other words, $I(\theta) = 0$ iff $f_\theta(x)$ does not depend on $\theta$. (That is, $X \sim f_\theta(x)$ has no information about $\theta$ iff its distribution does not depend on $\theta$!)