Ma 3 - Problem Set 5

Marco Yang

2237027

RIP couldn't hit the buzzer beater gradescope submission so taking a 1 day extension on this.

1. (25 points) Find the MLE of θ when $X_1, X_2, \ldots, X_n$ are i.i.d. random samples and each X_i has a distribution given by the density function $f(x) = \frac{1}{2}e^{-|x-\theta|}$, $-\infty < x < \infty$.

Solution: The probability that every X_i came from this distribution is

$$P = \prod_{i=1}^{N} f(X_i) = \prod_{i=1}^{N} \frac{1}{2}e^{-|X_i-\theta|} = \frac{1}{2}e^{-\sum_{i=1}^{N}|X_i-\theta|}.$$

Maximizing this probability occurs at the critical point, e.g. when $\frac{\partial L}{\partial \theta} = 0$.

$$\begin{aligned}
\frac{\partial L}{\partial \theta} &= \frac{1}{2}e^{-\sum_{i=1}^{N}|X_i-\theta|} \cdot \left(-\sum_{i=1}^{N}|X_i-\theta| \cdot \operatorname{sign}(X_i-\theta)\right) \\
&= \frac{1}{2}e^{-\sum_{i=1}^{N}|X_i-\theta|} \cdot \left(-\sum_{i=1}^{N}(X_i-\theta)\right) \\
&= -\frac{1}{2}e^{-\sum_{i=1}^{N}|X_i-\theta|} \left(\sum_{i=1}^{N}(X_i-\theta)\right) \\
.
\end{aligned}$$

We know that $\sum_{i=1}^{N}(X_i-\theta) = 0$ when $\theta = \sum_{i=1}^{N} X_i/N$. Thus, θ is the mean of the X_i.

2. (25 points) The MSE of an estimator is given by $\text{MSE}(\hat{\theta}) = \mathbb{E}[(\hat{\theta}-\theta)^2]$. Show that

$$\text{MSE}(\hat{\theta}) = \text{Var}(\theta) + \text{Bias}(\theta)^2.$$

Solution: Note that $\hat{\theta}$ is independent of θ. Thus, we can take it out of any expected values w.r.t. θ as if it were a constant.

$$\begin{aligned}
\text{MSE}(\hat{\theta}) &= \mathbb{E}[(\hat{\theta}-\theta)^2] \\
&= \mathbb{E}[\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2] \\
&= \mathbb{E}[\hat{\theta}^2] - 2\mathbb{E}[\theta\hat{\theta}] + \mathbb{E}[\theta^2] \\
&= \hat{\theta}^2 - 2\hat{\theta}\mathbb{E}[\theta] + \mathbb{E}[\theta^2] \\
&= \hat{\theta}^2 - 2\hat{\theta}\mathbb{E}[\theta] + \mathbb{E}[\theta]^2 + \mathbb{E}[\theta^2] - \mathbb{E}[\theta]^2 \\
&= (\mathbb{E}[\theta^2] - \mathbb{E}[\theta]^2) + (\hat{\theta} - \mathbb{E}[\theta])^2 \\
&= \text{Var}(\theta) + \text{Bias}(\theta)^2.
\end{aligned}$$

3. (25 points) Let $X_1, X_2, \ldots, X_n$ be an i.i.d. random sample where X_i are distributed according to $U(0, \theta)$, i.e. a uniform distribution on $[0, \theta]$ where θ is unknown.

(a) Find $E(X)$ if $X \sim U(0, \theta)$.

Solution:

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[X_i] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[U(0,\theta)] = \frac{1}{N}\sum_{i=1}^{N} \frac{\theta}{2} = \frac{\theta}{2}.$$

(b) Show that the estimator $\hat{\theta} = 2 \cdot \overline{x}$ is unbiased.

Solution: Plugging in $X_i \sim U(0, \hat{\theta}) = U(0, 2 \cdot \overline{x})$,

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[X_i] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[U(0, 2 \cdot \overline{x})] = \frac{1}{N}\sum_{i=1}^{N} \overline{x} = \overline{x}.$$

Since $E(X) = \overline{x}$ when $\hat{\theta} = 2 \cdot \overline{x}$, the estimator is unbiased.

(c) Find the MLE for θ.

Solution: The probability density for $U(0, \theta)$ is $\frac{1}{\theta}$. Thus, the likelihood of getting our sample X_i (or any valid sample) is

$$p = \prod_{i=1}^{n} \frac{1}{\theta} = \theta^{-n}.$$

Taking the derivative w.r.t. θ,

$$\frac{dp}{dt} = -n\theta^{-n-1}.$$

Setting this to 0,

$$\frac{dp}{dt} = -n\theta^{-n-1} = 0 \implies \theta = 0.$$

However, this doesn't make since since if $\theta = 0$, then we know that all values of X must be 0, but we are constrained by the value we sampled. What we can do is notice that p is maximized for values of θ closer to 0. Thus, we need to minimize θ given the sample, which would make $\theta = \max(X_i)$.

(d) Compare the MSE for the two estimators, which one is lower?

Solution: The MSE for $\hat{\theta} = 2 \cdot \overline{x}$ is

$$\sum_{i=1}^{n} \left(X_i - \frac{2}{n} \sum_{i=1}^{n} X_i \right)^2.$$

Let $\max(X_i) = m$. The MSE for $\theta = \max(X_i)$ is

$$\sum_{i=1}^{n} (X_i - m)^2.$$

Subtracting the first from the second,

$$\begin{aligned}
\Delta &= \sum_{i=1}^{n}\left(2X_i - \frac{2}{n}\sum_{i=1}^{n} X_i - m\right)\cdot\left(\frac{2}{n}\sum_{i=1}^{n} X_i + m\right) \\
&= (2\overline{x}+m)\cdot\left(\left(2\sum_{i=1}^{n}(X_i-\overline{x})\right) - mn\right) \\
&= (2\overline{x}+m)(-mn) \\
&= -mn(2\overline{x}+m) < 0.
\end{aligned}$$

Thus, $\theta = \max(X_i)$ has a lower MSE.

4. (25 points) **(Hardy-Weinberg.)** Suppose that a particular gene occurs as one of two alleles (A and a), where allele A has a frequency $\theta \in (0,1)$ in the population. That is, a random copy of the gene is A with probability θ and a with probability $1-\theta$. Since a diploid genotype consists of two genes, the probability of each genotype is

$$P(AA) = \theta^2 \quad P(Aa) = 2\theta(1-\theta) \quad P(aa) = (1-\theta)^2.$$

Suppose we test a random sample of people and find that n_1 are AA, n_2 are Aa, and n_3 are aa. Find the MLE $\hat{\theta}_n$. Make sure to verify that is indeed maximizing, by computing the second derivative of the function you are maximizing.

Solution: The probability of getting the samples n_i is

$$p = \theta^2 \cdot 2\theta(1-\theta)\cdot(1-\theta)^2 = 2\theta^3(1-\theta)^3 = 2(\theta(1-\theta))^3.$$

To maximize this likelihood, we take the derivative w.r.t. θ:

$$\frac{dp}{d\theta} = 6(\theta(1-\theta))^2\cdot(1-2\theta)$$

and find where it is zero:

$$\frac{dp}{d\theta} = 0 \implies \theta = 0, 1, \frac{1}{2}.$$

Out of these candidates, only $\theta = \frac{1}{2}$ makes sense since the other are the bounds for what a probability can be.

The second derivative is

$$\frac{d^2p}{d\theta^2} = 6\left(2\theta(1-\theta)(1-2\theta)^2 - 2(\theta(1-\theta))^2\right).$$

At $\theta = \frac{1}{2}$, the second derivative is negative, meaning it is a local maximum. At $\theta = 0, 1$, the second derivative is zero, which means we