ECON 6190

Problem Set 4

Gabe Sekeres

October 9, 2024

1. We have that the marginal density for a uniform random variable is $f_X(x) = \mathbb{1}_{x\in(\theta,\theta+1)}$. Thus, the joint pdf is

$$f(x \mid \theta) = \begin{cases} 1 & X_i \in (\theta, \theta+1)\ \forall\, i = 1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$$

Consider the statistic $T(X) = (\min_i X_i, \max_i X_i) \coloneqq (x^{(1)}, x^{(n)})$. This is a sufficient statistic by the factorization theorem, as taking $g(T(X) \mid \theta) = \mathbb{1}_{\theta \le x^{(1)}, x^{(n)} \le \theta+1} = \mathbb{1}_{\theta \in (x^{(n)}-1, x^{(1)})}$ and $h(x) = 1$, we get that $f(x \mid \theta) = g(T(X) \mid \theta)h(x)$. Take some $x, y \in X$. Then we have that

$$\frac{f(x \mid \theta)}{f(y \mid \theta)} = \frac{g(T(x) \mid \theta)}{g(T(y) \mid \theta)} = \begin{cases} 0 & \theta \notin (x^{(n)}-1, x^{(1)}), \theta \in (y^{(n)}-1, y^{(1)}) \\ 1 & \theta \in (x^{(n)}-1, x^{(1)}), \theta \in (y^{(n)}-1, y^{(1)}) \\ \infty & \theta \notin (x^{(n)}-1, x^{(1)}), \theta \notin (y^{(n)}-1, y^{(1)}) \end{cases}$$

This ratio is not dependent on θ if and only if $(x^{(n)}-1, x^{(1)}) = (y^{(n)}-1, y^{(1)})$, so T is a minimal sufficient statistic for θ.

2. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$ for unknown μ and known σ^2. We are interested in estimating μ.

   (a) Consider the statistic $T(X) = \bar{X}$, which we showed in class was sufficient because taking $g(\bar{X} \mid \mu) = \exp\left(-\frac{n(\bar{X}-\mu)^2}{2\sigma^2}\right)$ and $h(x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{2\sigma^2}\right)$, we get that

$$\begin{aligned} g(\bar{X} \mid \mu)h(x) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\sum_{i=1}^n (X_i-\bar{X})^2}{2\sigma^2}\right) \exp\left(-\frac{n(\bar{X}-\mu)^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\sum_{i=1}^n (X_i-\bar{X})^2 + n(\bar{X}-\mu)^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\sum_{i=1}^n (X_i-\mu)^2}{2\sigma^2}\right) \\ &= f(x \mid \mu) \end{aligned}$$

   To show that it is minimal, consider samples $X \sim \{X_1, \dots, X_n\}$ and $Y \sim \{Y_1, \dots, Y_n\}$. We get that

$$\begin{aligned} \frac{f(X \mid \mu)}{f(Y \mid \mu)} &= \frac{(2\pi\sigma)^{-n/2} \exp\left(-\frac{(n-1)s_X^2 + n(\bar{X}-\mu)^2}{2\sigma^2}\right)}{(2\pi\sigma)^{-n/2} \exp\left(-\frac{(n-1)s_Y^2 + n(\bar{Y}-\mu)^2}{2\sigma^2}\right)} \\ &= \exp\left(\frac{(n-1)(s_X^2 - s_Y^2) + n(\bar{Y}-\bar{X}) + 2n\mu(\bar{X}-\bar{Y})}{2\sigma^2}\right) \end{aligned}$$

   Which does not depend on μ if and only if $\bar{X} = \bar{Y}$, so $T(X) = \bar{X}$ is minimal.

   (b) Suppose $\sigma^2 = 1$ and $n = 1$. Consider the estimator $\hat{\theta} = \frac{c^2}{c^2+1} X_1$ for some $c > 0$.

      i. The MSE of $\hat{\theta}$ is

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

We have that

$$\text{bias}(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta = \frac{c^2}{c^2+1}\,\mathbb{E}[X_1] - \mu = \mu\left(\frac{c^2}{c^2+1} - 1\right) = -\frac{\mu}{c^2+1}$$

So the estimator is not unbiased. We also have that

$$\text{Var}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2] = \mathbb{E}\left[\left(\frac{c^2}{c^2+1}(X_1 - \mu)\right)^2\right] = \frac{c^4}{(c^2+1)^2}\,\mathbb{E}[(X_1-\mu)^2] = \frac{c^4}{(c^2+1)^2}$$

Thus, we have that

$$\text{MSE}(\hat{\theta}) = \frac{\mu^2 + c^4}{(c^2+1)^2}$$

ii. We will first find

$$\text{MSE}(\tilde{\theta}) = \text{bias}(\tilde{\theta})^2 + \text{Var}(\tilde{\theta})$$

Where

$$\text{bias}(\tilde{\theta}) = \mathbb{E}[\tilde{\theta}] - \theta = \mathbb{E}[X_1] - \mu = \mu - \mu = 0$$

and

$$\text{Var}(\tilde{\theta}) = \mathbb{E}[(\tilde{\theta} - \mathbb{E}[\tilde{\theta}])^2] = \mathbb{E}[(X_1 - \mu)^2] = 1$$

Thus, $\text{MSE}(\tilde{\theta}) = 1$. We have that

$$\text{MSE}(\hat{\theta}) > \text{MSE}(\tilde{\theta}) \iff \mu^2 > 2c^2 + 1$$

Thus, if $\mu^2 > 2c^2 + 1$, $\tilde{\theta}$ is more efficient than $\hat{\theta}$.

iii. From my answer to (ii), when $\mu = c$, then $\mu^2 + c^4 < (c^2+1)^2$, so $\hat{\theta}$ is more efficient because $\text{MSE}(\hat{\theta}) < 1 = \text{MSE}(\tilde{\theta})$.

3. We have that $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$ is an estimator for $\sigma^2 = \text{Var}(X)$. We have that

$$\begin{aligned}
\mathbb{E}[\hat{\sigma}^2] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2\right] \\
&= \left(\frac{n-1}{n}\right)\mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2\right] \\
&= \left(\frac{n-1}{n}\right)\mathbb{E}\left[s^2\right] \\
&= \left(1 - \frac{1}{n}\right)\sigma^2
\end{aligned}$$

where the last equality follows from a theorem in class that $\mathbb{E}[s^2] = \sigma^2$. We thus have that the bias of $\hat{\sigma}^2$ is

$$\text{bias}(\hat{\sigma}^2) = \mathbb{E}[\hat{\sigma}^2] - \sigma^2 = \sigma^2\left(1 - \frac{1}{n} - 1\right) = -\frac{\sigma^2}{n}$$

4. Suppose $X \sim \mathcal{N}(0, \sigma^2)$. Consider the following estimator for σ^2:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2$$

(a) We have that $n\hat{\sigma}^2/\sigma^2 = \frac{\sum_{i=1}^n X_i^2}{\sigma^2}$. Then, since $X_i \sim \mathcal{N}(0, \sigma^2)$, we have that defining $Y_i \sim \mathcal{N}(0,1)\ \forall\, i$,

$$\frac{\sum_{i=1}^n X_i^2}{\sigma^2} = \sum_{i=1}^n Y_i^2 \sim \chi_n^2$$

Thus, $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi_n^2$.

(b) We have that

$$\mathbb{E}[\hat{\sigma}^2] = \frac{1}{n}\,\mathbb{E}\left[\sum_{i=1}^n (X_i)^2 - 0^2\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[X_i^2 - \mathbb{E}[X_i]^2\right] = \frac{1}{n} n\sigma^2 = \sigma^2$$

(c) We have that

$$\begin{aligned} \mathrm{Var}(\hat{\sigma}^2) &= \mathrm{Var}\left(\frac{\sigma^2}{n}\frac{n\hat{\sigma}^2}{\sigma^2}\right) \\ &= \frac{\sigma^4}{n^2}\,\mathrm{Var}(\chi_n^2) \\ &= \frac{\sigma^4}{n^2}\cdot 2n = \frac{2\sigma^4}{n} \end{aligned}$$

(d) We have that

$$\begin{aligned} \mathrm{MSE}(\hat{\sigma}^2) &= \mathrm{bias}(\hat{\sigma}^2)^2 + \mathrm{Var}(\hat{\sigma}^2) \\ &= \left(\mathbb{E}[\hat{\sigma}^2] - \sigma^2\right)^2 + \frac{2\sigma^4}{n} \\ &= \left(\sigma^2 - \sigma^2\right)^2 + \frac{2\sigma^4}{n} \\ &= \frac{2\sigma^4}{n} \end{aligned}$$

5. Let $\{X_1, \ldots, X_n\}$ be a random sample from a Poisson distribution with parameter λ:

$$\mathbb{P}\{X_i = j\} = \frac{e^{-\lambda}\lambda^j}{j!}\ \forall\, j = 0, 1, 2, \ldots$$

(a) We have that

$$f(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & x \in \mathbb{Z}_+ \\ 0 & \text{otherwise} \end{cases}$$

So the joint pmf is

$$f(x \mid \lambda) = \prod_{i=1}^n f(x_i) = \begin{cases} \prod_{i=1}^n e^{-\lambda}\frac{\lambda^{x_i}}{x_i!} & x_i \in \mathbb{Z}_+\ \forall\, i = 1, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

which can be recast as

$$f(x \mid \lambda) = \mathbb{1}_{\{x_i \in \mathbb{Z}_+\ \forall\, i\}} e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}\prod_{i=1}^n \frac{1}{x_i!}$$

Taking the statistic $T(x) = \sum_{i=1}^n x_i$, we get that by the factorization theorem, taking $g(T(x) \mid \lambda) = e^{-n\lambda}\lambda^{T(x)}$ and $h(x) = \mathbb{1}_{\{x_i \in \mathbb{Z}_+\ \forall\, i\}}\prod_{i=1}^n \frac{1}{x_i!}$ we have that

$$f(x \mid \lambda) = g(T(x) \mid \lambda)h(x) = e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}\mathbb{1}_{\{x_i \in \mathbb{Z}_+\ \forall\, i\}}\prod_{i=1}^n \frac{1}{x_i!}$$

To show that T is minimal, consider some $X \sim \{X_1, \dots, X_n\}$ and $Y \sim \{Y_1, \dots, Y_n\}$. We have that

$$\frac{f(X \mid \lambda)}{f(Y \mid \lambda)} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n X_i}\mathbb{1}_{\{X_i \in \mathbb{Z}_+ \ \forall\ i\}}\prod_{i=1}^n \frac{1}{X_i!}}{e^{-n\lambda}\lambda^{\sum_{i=1}^n Y_i}\mathbb{1}_{\{Y_i \in \mathbb{Z}_+ \ \forall\ i\}}\prod_{i=1}^n \frac{1}{Y_i!}} = \lambda^{\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}\frac{\mathbb{1}_{\{X_i \in \mathbb{Z}_+ \ \forall\ i\}}\prod_{i=1}^n \frac{1}{X_i!}}{\mathbb{1}_{\{Y_i \in \mathbb{Z}_+ \ \forall\ i\}}\prod_{i=1}^n \frac{1}{Y_i!}}$$

and since this ratio is not dependent on λ if and only if $\sum_{i=1}^n X_i = \sum_{i=1}^n Y_i$, *i.e.* when $T(X) = T(Y)$, T is minimal.

(b) Define $\hat{\theta}_1 \coloneqq \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i=0\}}$. Then we have that

$$\text{bias}(\hat{\theta}_1) = \mathbb{E}[\hat{\theta}_1] - \theta = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i=0\}} - e^{-\lambda}\right] = \frac{1}{n}n\,\mathbb{E}[X=0] - e^{-\lambda} = \mathbb{P}\{X=0\} - e^{-\lambda} = 0$$

(c) This estimator is not a function of the minimal sufficient statistic. To see why, consider the fact that taking $X = \{0, 3\}$, $\hat{\theta}_1 = \frac{1}{2}$ and $T(X) = 3$, however taking $Y = \{1, 2\}$, $\hat{\theta}_1 = 0$ but $T(Y) = T(X) = 3$.

(d) Since we have that $\hat{\theta}_2(X) = \mathbb{E}[\hat{\theta}_1(X) \mid T(X)]$, and since $\hat{\theta}_1$ is an unbiased estimator and $T(X)$ is a sufficient statistic, we have that, by Rao-Blackwell, $\text{bias}(\hat{\theta}_2) = 0$ and $\text{MSE}(\hat{\theta}_2) \leq \text{MSE}(\hat{\theta}_1)$.

(e) We have that $\hat{\theta}_2 = \mathbb{E}[\hat{\theta}_1 \mid T]$. Reformulating, we have that

$$\begin{aligned}
\hat{\theta}_2(X) &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i=0\}}\middle|\sum_{i=1}^n X_i = t\right] \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[X_i = 0\middle|\sum_{i=1}^n X_i = t\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{X_1=0\}}\middle|\sum_{i=1}^n X_i = t\right] \\
&= \mathbb{P}\left\{X_1 = 0\middle|\sum_{i=1}^n X_i = t\right\} \\
&= \frac{\mathbb{P}\left\{X_1 = 0, \sum_{i=1}^n X_i = t\right\}}{\mathbb{P}\left\{\sum_{i=1}^n X_i = t\right\}} \\
&= \frac{\mathbb{P}\left\{X_1 = 0\right\}\mathbb{P}\left\{\sum_{i=2}^n X_i = t\right\}}{\mathbb{P}\left\{\sum_{i=1}^n X_i = t\right\}}
\end{aligned}$$

Using the properties of the Poisson distribution, we can calculate this directly. We get that

$$\hat{\theta}_2(X) = \left(\frac{n-1}{n}\right)^{\sum_{i=0}^n X_i}$$