Homework for 1/27 Due 2/5

1. [§8-13] In Example D of Section 8.4, the pdf of the population distribution is

\[
f(x|\alpha) = \begin{cases} 
\frac{1 + \alpha x}{2} & -1 \leq x \leq 1, \\
0 & \text{otherwise,}
\end{cases} 
\]

and the method of moments estimate was found to be \( \hat{\alpha} = 3\bar{X} \) (where \( \bar{X} \) is the sample mean of the random sample \( X_1, \ldots, X_n \)). In this problem, you will consider the sampling distribution of \( \hat{\alpha} \).

(a) Show that the estimate \( \hat{\alpha} \) is unbiased.

(b) Find \( \text{Var}[\hat{\alpha}] \). \( \text{Hint: What is } \text{Var}[\bar{X}]? \)

(b) Use the central limit theorem to deduce a normal approximation to the sampling distribution of \( \hat{\alpha} \). According to this approximation, if \( n = 25 \) and \( \alpha = 0 \), what is the \( P(|\hat{\alpha}| > 0.5) \)?

(a) We notice that

\[
\mathbb{E}[\bar{X}] = \mathbb{E}[X_1] = \int_{-1}^{1} x \cdot \frac{1 + \alpha x}{2} \, dx = \int_{-1}^{1} \left( \frac{x^2 + \alpha x^2}{2} \right) \\
= \left( \frac{x^3}{4} + \frac{\alpha x^4}{6} \right) \bigg|_{-1}^{1} = \frac{\alpha}{3}.
\]

Therefore,

\[
\mathbb{E}[\hat{\alpha}] = \mathbb{E}[3\bar{X}] = 3\mathbb{E}[\bar{X}] = 3 \cdot \frac{\alpha}{3} = \alpha.
\]

(b) First, we have

\[
\mathbb{E}[X_1^2] = \int_{-1}^{1} x^2 \cdot \frac{1 + \alpha x}{2} \, dx = \int_{-1}^{1} \left( \frac{x^2}{2} + \frac{\alpha x^3}{2} \right) \\
= \left( \frac{x^4}{6} + \frac{\alpha x^5}{8} \right) \bigg|_{-1}^{1} = \frac{1}{3},
\]

and

\[
\text{Var}[X_1] = \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = \frac{1}{3} \cdot \left( \frac{\alpha}{3} \right)^2 = \frac{3 - \alpha^2}{9}.
\]

Thus,

\[
\text{Var}[\bar{X}] = \text{Var} \left[ \frac{1}{n} \sum_{i=1}^{n} X_i \right] = \frac{1}{n^2} \text{Var} \left[ \sum_{i=1}^{n} X_i \right] = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}[X_i] \\
= \frac{1}{n} \text{Var}[X_1] = \frac{3 - \alpha^2}{9n}.
\]
Therefore, we have

\[ \text{Var}[\hat{\alpha}] = \text{Var}[3\bar{X}] = 9\text{Var}[\bar{X}] = \frac{3-\alpha^2}{n}. \]

(c) According to the central limit theorem, we have \( \bar{X} \sim N\left(\frac{\alpha}{3}, \frac{3-\alpha^2}{9n}\right) \), approximately. Therefore, \( \hat{\alpha} = 3\bar{X} \) implies that

\[ \hat{\alpha} \sim N\left(\alpha, \frac{3-\alpha^2}{n}\right), \quad \text{approximately}. \]

In the case \( \alpha = 0 \) and \( n = 25 \), we have \( \hat{\alpha} \sim N(0, 0.12) \), approximately. Thus

\[
\mathbb{P}(|\hat{\alpha}| > .5) = \mathbb{P}(\hat{\alpha} > .5) + \mathbb{P}(\hat{\alpha} < -.5) \\
= \mathbb{P}\left(\frac{\hat{\alpha} - 0}{\sqrt{0.12}} > \frac{0.5 - 0}{\sqrt{0.12}}\right) + \mathbb{P}\left(\frac{\hat{\alpha} - 0}{\sqrt{0.12}} < \frac{-0.5 - 0}{\sqrt{0.12}}\right) \\
\approx \mathbb{P}(Z > 1.44) + \mathbb{P}(Z < -1.44) \\
= 0.0749 + 0.0749 = 0.1498.
\]

2. [§8-53] Let \( X_1, \ldots, X_n \) be i.i.d. uniform on \([0, \theta]\).

(a) Find the method of moments estimate of \( \theta \), and the mean, variance, bias, and MSE of the MME.

(b) The mle of \( \theta \) is \( \hat{\theta} = \max_{1 \leq i \leq n} X_i \). The pdf of \( \max_{1 \leq i \leq n} X_i \) (How do we find this?) is

\[
f(x|\theta) = \begin{cases} 
\frac{n x^{n-1}}{\theta^n} & 0 < x < \theta \\
0 & \text{otherwise}
\end{cases}
\]

Calculate the mean and variance of the mle. Compare the variance, the bias, and the mean squared error to those of the method of moments estimate.

(c) Find a modification of the mle that renders it unbiased.
(a) Since
\[ \mu_1 = \mathbb{E}[X_1] = \int_0^\theta x^1 \frac{1}{\theta} \, dx = \frac{x^2}{2\theta} \bigg|_0^\theta = \frac{\theta}{2}, \]
we have
\[ \theta = 2\mu_1, \]
and the MME for \( \theta \) is
\[ \hat{\theta} = 2\bar{X}, \]
where \( \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \). Furthermore, we have
\[ \mu_2 = \mathbb{E}[X_1^2] = \int_0^\theta x^2 \frac{1}{\theta} \, dx = \frac{x^3}{3\theta} \bigg|_0^\theta = \frac{\theta^2}{3}, \]
and
\[ \text{Var}[X_1] = \mu_2 - \mu_1^2 = \frac{\theta^2}{3} - \left( \frac{\theta}{2} \right)^2 = \frac{\theta^2}{12}. \]
Thus
\[ \mathbb{E}[\bar{X}] = \mathbb{E}[X_1] = \frac{\theta}{2}, \quad \text{and} \quad \text{Var}[\bar{X}] = \frac{1}{n} \text{Var}[X_1] = \frac{\theta^2}{12n}. \]
It follows that
\[ \mathbb{E}[\hat{\theta}] = \mathbb{E}[2\bar{X}] = 2\mathbb{E}[\bar{X}] = \theta, \]
and
\[ \text{Var}[\hat{\theta}] = \text{Var}[2\bar{X}] = 4\text{Var}[\bar{X}] = \frac{\theta^2}{3n}. \]
In particular, the MME \( \hat{\theta} \) is unbiased. The bias and MSE of \( \hat{\theta} \) are
\[ b(\hat{\theta}) = 0, \]
and
\[ \text{MSE}(\hat{\theta}) = \text{Var}[\hat{\theta}] + b(\hat{\theta})^2 = \text{Var}[\hat{\theta}] = \frac{\theta^2}{3n}. \]

(b) The mean of \( \hat{\theta} \) is
\[ \mathbb{E}[\hat{\theta}] = \int_0^\theta x \cdot \frac{n x^{n-1}}{\theta^n} \, dx = \left( \frac{n x^{n+1}}{(n+1)\theta^n} \right) \bigg|_0^\theta = \frac{n}{n+1} \theta. \]
Since
\[ \mathbb{E}[\hat{\theta}^2] = \int_0^\theta x^2 \cdot \frac{n x^{n-1}}{\theta^n} \, dx = \left( \frac{n x^{n+2}}{(n+2)\theta^n} \right) \bigg|_0^\theta = \frac{n}{n+2} \theta^2, \]
the variance of \( \hat{\theta} \) is

\[
\text{Var}[\hat{\theta}] = \mathbb{E}[\hat{\theta}^2] - \left( \mathbb{E}[\hat{\theta}] \right)^2 = \frac{n\theta^2}{n+2} - \left( \frac{n\theta}{n+1} \right)^2 = \frac{n\theta^2}{(n+1)^2(n+2)}.
\]

The bias of \( \hat{\theta} \) is

\[
b(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta = -\frac{\theta}{n+1},
\]

and the MSE of \( \hat{\theta} \) is

\[
\text{MSE}(\hat{\theta}) = \text{Var}[\hat{\theta}] + b(\hat{\theta})^2 = \frac{n\theta^2}{(n+1)^2(n+2)} + \frac{\theta^2}{(n+1)^2} = \frac{2\theta^2}{(n+1)(n+2)}.
\]

By comparison, although the MLE \( \hat{\theta} \) is biased while the MME \( \tilde{\theta} \) is unbiased, we see that \( \text{MSE}(\hat{\theta}) \) < \( \text{MSE}(\tilde{\theta}) \) when \( n \) is large. In fact, \( \text{MSE}(\hat{\theta}) \) decreases much faster than \( \text{MSE}(\tilde{\theta}) \).

(c) Let

\[
\bar{\theta} = \frac{n+1}{n} \hat{\theta} = \frac{n+1}{n} \max_{1 \leq i \leq n} X_i,
\]

it follows that

\[
\mathbb{E}[\bar{\theta}] = \mathbb{E}\left[ \frac{n+1}{n} \hat{\theta} \right] = \frac{n+1}{n} \mathbb{E}[\hat{\theta}] = \theta.
\]

Thus \( \bar{\theta} \) is unbiased.
This problem is concerned with the estimation of the variance of a normal distribution with unknown mean from a sample $X_1, \ldots , X_n$ of i.i.d. normal random variables $N(\mu , \sigma^2)$. In answering the following questions, use the fact that (from Theorem B of Section 6.3)

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

and that the mean and variance of a chi-square random variable with $r$ df are $r$ and $2r$, respectively.

(a) Which of the following estimates is unbiased?

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

(We discussed this in class. However, we do not assume normality. When the distribution is not normal, the argument is much more complicated, as seen in class. A technical detail is provided at the end of this homework.)

(b) Which of the estimates given in part (a) has the smaller MSE?

(c) For what value of $\rho$ does $\rho \sum_{i=1}^{n} (X_i - \overline{X})^2$ have the minimal MSE (as an estimate for $\sigma^2$)?

(a) Since

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1},$$

and

$$\mathbb{E}[U] = r \quad \text{and} \quad \text{Var}[U] = 2r,$$

where $U \sim \chi^2_r$, we have

$$\mathbb{E} \left[ \frac{(n-1)s^2}{\sigma^2} \right] = n-1 \quad \text{and} \quad \text{Var} \left[ \frac{(n-1)s^2}{\sigma^2} \right] = 2(n-1).$$

Therefore,

$$\mathbb{E} [s^2] = (n-1) \cdot \frac{\sigma^2}{(n-1)} = \sigma^2,$$

and

$$\text{Var} [s^2] = 2(n-1) \cdot \left( \frac{\sigma^2}{(n-1)} \right)^2 = \frac{2\sigma^4}{n-1}.$$ 

Furthermore, since $\hat{\sigma}^2 = \frac{n-1}{n}s^2$, we have

$$\mathbb{E}[\hat{\sigma}^2] = \frac{n-1}{n} \mathbb{E}[s^2] = \frac{n-1}{n} \sigma^2.$$
and
\[ \text{Var}[\hat{\sigma}^2] = \left( \frac{n-1}{n} \right)^2 \text{Var}[s^2] = \frac{2(n-1)}{n^2} \sigma^4. \]

Thus \( s^2 \) is an unbiased estimate for \( \sigma^2 \) while \( \hat{\sigma}^2 \) is not.

(b) The bias of the two estimates are
\[ b(s^2) = 0 \quad \text{and} \quad b(\hat{\sigma}^2) = \mathbb{E}[\hat{\sigma}^2] - \sigma^2 = \frac{n-1}{n} \sigma^2 - \sigma^2 = -\frac{1}{n} \sigma^2, \]
respectively. Thus the MSE of the two estimates are
\[ \text{MSE}(s^2) = \text{Var}[s^2] + b(s^2)^2 = \frac{2\sigma^4}{n-1} + 0 = \frac{2\sigma^4}{n-1}, \]
and
\[ \text{MSE}(\hat{\sigma}^2) = \text{Var}[\hat{\sigma}^2] + b(\hat{\sigma}^2)^2 = \frac{2(n-1)}{n^2} \sigma^4 + \left(-\frac{1}{n} \sigma^2\right)^2 = \frac{2n-1}{n^2} \sigma^4, \]
respectively. Since
\[ 3n > 1 \Rightarrow (2n-1)(n-1) < 2n^2 \Rightarrow \frac{2n-1}{n^2} < \frac{2}{n-1}, \]
we have \( \text{MSE}(\hat{\sigma}^2) < \text{MSE}(s^2) \).

(c) Let \( Y := \rho \sum_{i=1}^{n} (X_i - \bar{X})^2 \). Then \( Y = \rho(n-1)s^2 \). By a similar argument as in (a), we have
\[ \mathbb{E}[Y] = \rho(n-1)\mathbb{E}[s^2] = \rho(n-1)\sigma^2, \]
and
\[ \text{Var}[Y] = (\rho(n-1))^2 \text{Var}[s^2] = \rho^2(n-1)^2 \frac{2\sigma^4}{n-1} = 2\rho^2(n-1)\sigma^4. \]
Thus
\[ \text{MSE}(Y) = \text{Var}[Y] + b(Y)^2 = 2\rho^2(n-1)\sigma^4 + (\rho(n-1)\sigma^2 - \sigma^2)^2 = \sigma^4[2\rho^2(n-1) + (\rho n - \rho - 1)^2] \equiv f(\rho). \]
Since
\[ f'(\rho) = \sigma^4[4\rho(n-1) + 2(\rho n - \rho - 1)(n-1)] \]
\[ = \sigma^4[4\rho + 2(\rho n - \rho - 1)](n-1) \]
\[ = 2(n-1)\sigma^4(\rho n + \rho - 1), \]
and
\[ f''(\rho) = 2(n-1)(n+1)\sigma^4 > 0, \]
we see that \( f(\rho) \) achieve its minimum at \( \rho = \frac{1}{n+1} \). (\( f ((n+1)^{-1}) = 0 \).)
Homework for 1/29 Due 2/5

1. [§8-7] Suppose that $X$ follows a geometric distribution,

$$P(X = k) = p (1 - p)^{k-1}$$

and assume $X_1, \ldots, X_n$ is an i.i.d. sample of size $n$. Find the asymptotic variance of the mle. (*The moments of geometric distribution can be found in P117.*)

We have

$$\log f(X|p) = \log p + (X - 1) \log (1 - p),$$

$${\partial \over \partial p} \log f(X|p) = {1 \over p} - {X - 1 \over 1 - p}, \quad \text{and}$$

$${\partial^2 \over \partial p^2} \log f(X|p) = -{1 \over p^2} - {X - 1 \over (1 - p)^2}.$$

Therefore, the Fisher information is

$$I(p) = -\mathbb{E} \left[ {\partial^2 \over \partial p^2} \log f(X|p) \right] = - \left( -{1 \over p^2} - {1 \over (1 - p)^2} (\mathbb{E}[X] - 1) \right)$$

$$= {1 \over p^2} + {1 \over (1 - p)^2} \left( {1 \over p} - 1 \right) = {1 \over (1 - p)p^2},$$

and the asymptotic variance of the mle is

$$\frac{1}{nI(p)} = \frac{(1 - p)p^2}{n}.$$

2. [§8-16] Consider an i.i.d. sample of random variables with density function

$$f(x|\sigma) = {1 \over 2\sigma} \exp \left( -{|x| \over \sigma} \right), \quad -\infty < x < \infty, \quad \sigma > 0.$$

Find the asymptotic variance of the mle.
We have

\[ \log f(X|\sigma) = -\log 2 - \log \sigma - \frac{|X|}{\sigma}, \]

\[ \frac{\partial}{\partial \sigma} \log f(X|\sigma) = -\frac{1}{\sigma} + \frac{|X|}{\sigma^2}, \quad \frac{\partial^2}{\partial \sigma^2} \log f(X|\sigma) = \frac{1}{\sigma^2} - \frac{2|X|}{\sigma^3}, \quad \text{and} \]

\[ \mathbb{E}[|X|] = \int_{-\infty}^{\infty} |x| \cdot \frac{1}{2\sigma} \exp \left( -\frac{|x|}{\sigma} \right) \, dx = 2 \int_{0}^{\infty} \frac{x}{2\sigma} \exp \left( -\frac{x}{\sigma} \right) \, dx \]

\[ = \sigma \int_{0}^{\infty} x \exp \left( -\frac{x}{\sigma} \right) \, dx = \sigma \int_{0}^{\infty} ye^{-y} \, dy \]

\[ = \sigma. \]

Therefore, the Fisher information is

\[ I(\sigma) = -\mathbb{E} \left[ \frac{\partial^2}{\partial \sigma^2} \log f(X|\sigma) \right] = -\left( \frac{1}{\sigma^2} - \frac{2}{\sigma^3} \mathbb{E}[|X|] \right) = \frac{1}{\sigma^2}, \]

and the asymptotic variance of the mle is

\[ \frac{1}{nI(\sigma)} = \frac{\sigma^2}{n}. \]
3. [§8-47] The Pareto distribution has been used in economics as a model for a density function with a slowly decaying tail:

\[
f(x|x_0, \theta) = \theta x_0^\theta x^{-\theta - 1}, \quad x \geq x_0, \theta > 1.\]

Assume that \( x_0 > 0 \) is given and that \( X_1, X_2, \ldots, X_n \) is an i.i.d. sample.

(a) Find the method of moments estimate of \( \theta \).

(b) Find the mle of \( \theta \).

(c) Find the asymptotic variance of the mle.

(a) Since

\[
\mu_1 = \mathbb{E}[X_1] = \int_{x_0}^{\infty} x \cdot \theta x_0^\theta x^{-\theta - 1} \, dx = \theta x_0^\theta \int_{x_0}^{\infty} x^{-\theta} \, dx
\]

we have

\[
\theta = \frac{\mu_1}{\mu_1 - x_0},
\]

and the MME is

\[
\hat{\theta} = \frac{\bar{X}_1}{\bar{X} - x_0},
\]

where \( \bar{X} = \sum_{i=1}^{n} X_i \).

(b) The log likelihood function is

\[
\log f(\theta) = \sum_{i=1}^{n} (\log \theta + \theta \log x_0 - (\theta + 1) \log X_i)
\]

\[
= n \log \theta + n \theta \log x_0 - (\theta + 1) \sum_{i=1}^{n} \log X_i.
\]

Thus

\[
l'(\theta) = \frac{n}{\theta} + n \log x_0 - \sum_{i=1}^{n} \log X_i,
\]

and

\[
l''(\theta) = -\frac{n}{\theta^2} < 0.
\]

Since

\[
ll' \left( \frac{1}{n \sum_{i=1}^{n} \log X_i - \log x_0} \right) = 0,
\]

9
the mle of $\theta$ is
\[ \hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \log X_i - \log x_0. \]

(c) We have
\[ \log f(X|\theta) = \log \theta + \theta \log x_0 - (\theta + 1) \log X, \]
\[ \frac{\partial}{\partial \theta} \log f(X|\theta) = \frac{1}{\theta} + \log x_0 - \log X \quad \text{and} \quad \frac{\partial^2}{\partial \theta^2} \log f(X|\theta) = -\frac{1}{\theta^2}. \]

Therefore, the Fisher information is
\[ I(\theta) = -\mathbb{E} \left[ \frac{\partial^2}{\partial \theta^2} \log f(X|\theta) \right] = -\left( -\frac{1}{\theta^2} \right) = \frac{1}{\theta^2}, \]
and the asymptotic variance of the mle is
\[ \frac{1}{nI(\theta)} = \frac{\theta^2}{n}. \]