Announcement: There are solutions to some extra problems

(-1) **Problem 19, (a) and (b):** This is a special case of Problem 8(b) of the More Problems Section, with $\lambda = 1$.

(0) **Problem 21:** For a more general version, see Problem 2 in Practice Midterm Solutions downloadable off the course website.

(0)' **Problem 29:** We have,

$$\text{Prob} \left( \frac{\sqrt{n}(\bar{X} - \mu)}{s} \geq -t_{n-1}^{1-\alpha} \right) = 1 - \alpha.$$ 

Here $t_{n-1}^{1-\alpha}$ is the $1 - \alpha$'th quantile of the $t_{n-1}$ distribution. This is equivalent to the statement that

$$\text{Prob} \left( \mu \leq \bar{X} + \frac{s}{\sqrt{n}} t_{n-1}^{1-\alpha} \right) = 1 - \alpha.$$ 

For this particular problem, choose $\alpha = 0.05$.

(1) **Problem 39, Chapter 8:** (a) Note that $x \geq x_0$ and $\theta > 1$ and

$$f(x, \theta) = \theta x_0^\theta x^{-\theta - 1}.$$ 

$$E(X_1) = \int_{x_0}^{\infty} \theta x_0^\theta x^{-\theta} dx$$

$$= \theta x_0^\theta \left[ \frac{-1}{-\theta + 1} x^{-\theta+1} \right]_{x_0}^{\infty}$$

$$= \theta x_0^\theta \frac{x_1^{1-\theta}}{-\theta + 1}$$

$$= \frac{\theta x_0}{\theta - 1}$$

$$= \mu_1.$$
Solving for $\theta$ we get, 

$$\theta = \frac{\mu_1}{\mu_1 - x_0}.$$ 

Thus, 

$$\hat{\theta}_{MOM} = \frac{\bar{X}}{\bar{X} - x_0}.$$ 

(b) The joint density, 

$$f(X_1, X_2, \ldots, X_n, \theta) = \theta^n x_0^n \prod_{i=1}^{n} X_i^{-\theta - 1}.$$ 

Setting, 

$$\frac{\partial}{\partial \theta} \log f(X_1, X_2, \ldots, X_n, \theta) = 0$$ 

gives 

$$\frac{n}{\theta} + n \log x_0 - \sum_{i=1}^{n} \log X_i = 0$$ 

and solving for $\theta$ we get, 

$$\hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^{n} (\log X_i - \log x_0)}.$$ 

(c) 

$$\sqrt{n} (\hat{\theta} - \theta) \rightarrow_d N(0, I_1(\theta)^{-1}).$$ 

Here, 

$$I_1(\theta) = -E_\theta \left( \frac{\partial^2}{\partial \theta^2} \log f(X_1, \theta) \right).$$ 

A trite computation shows that, 

$$\frac{\partial^2}{\partial \theta^2} \log f(X_1, \theta) = -\frac{1}{\theta^2}$$ 

showing that $I_1(\theta)^{-1} = \theta^2$. 

(d) It is easy to see that $\prod_{i=1}^{n} X_i$ is sufficient; equivalently, $\sum_{i=1}^{n} \log X_i$ is sufficient. 

(2) **Problem 42, Chap.8:** (a) Use the Jacobian Theorem to show that $X_1^2$ follows an Exponential$(1/2 \theta^2)$ distribution. Thus, 

$$\mu_2 = E(X_1^2) = 2 \theta^2$$ 

and hence 

$$\theta = \sqrt{\frac{\mu_2}{2}}.$$ 

It follows that 

$$\hat{\theta}_{MOM} = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{2n}}.$$ 

2
(b) Denoting the joint density by \( p(X, \theta) \) we have,

\[
\log p(X, \theta) = -2n \log \theta + \sum_{i=1}^{n} \log X_i - \frac{\sum_{i=1}^{n} X_i^2}{2 \theta^2}.
\]

Now,

\[
\frac{\partial}{\partial \theta} \log p(X, \theta) = 0
\]

gives

\[-2n \frac{\theta}{\theta} + \frac{\sum_{i=1}^{n} X_i^2}{\theta^3} = 0,
\]

and on solving for \( \theta \) this yields,

\[
\theta^2 = \frac{\sum_{i=1}^{n} X_i^2}{2n}.
\]

Thus,

\[
\hat{\theta}_{MLE} = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{2n}}.
\]

(c)

\[
\log f(x, \theta) = \log x - 2 \log \theta - \frac{x^2}{2 \theta^2}.
\]

From this we get,

\[
\frac{\partial^2}{\partial \theta^2} \log f(x, \theta) = \frac{2}{\theta^2} - \frac{3x^2}{\theta^4}.
\]

Thus,

\[
I_1(\theta) = -E_\theta \left( \frac{\partial^2}{\partial \theta^2} \log f(X_1, \theta) \right) = -\left( \frac{2}{\theta^2} - \frac{3}{\theta^4} E(X_1^2) \right) = + \frac{4}{\theta^2}
\]

on using the fact that \( E(X_1^2) = 2 \theta^2 \). Thus,

\[
\sqrt{n} (\hat{\theta} - \theta) \rightarrow_d N(0, \theta^2/4).
\]

(3) Problem 43, Rice’s Book. The joint density of the observations \( X = (X_1, X_2, \ldots, X_n) \) is,

\[
p(X, \theta) = \frac{1}{2^n} \exp \left( -\sum_{i=1}^{n} |X_i - \theta| \right).
\]

To maximise the above with respect to \( \theta \) one just needs to minimise,

\[
\sum_{i=1}^{n} |X_i - \theta|.
\]

The minimum is attained when \( \theta = \text{med}(X) \); i.e. \( \theta \) is the median of the \( X \)s. If \( n \) is odd and therefore of the form \( 2m + 1 \), then \( \text{med}(X) = X_{(m+1)} \), the \( m+1 \)'st ordered statistic. When \( n \)
is of the form $2m$, any value between $X_{(m)}$ and $X_{(m+1)}$ serves as a median. A rigorous proof of the above fact can be given but will be skipped for the moment. Let’s focus on the case of 2 observations. To minimise,

$$\Psi(X_1, X_2, \theta) = |X_1 - \theta| + |X_2 - \theta|$$

over $\theta$ we proceed as follows. Without loss of generality let $X_1 < X_2$. Then for any $\theta$ between $X_1$ and $X_2$,

$$\Psi(X_1, X_2, \theta) = |X_2 - X_1| = X_2 - X_1 .$$

If $\theta < X_1$ then

$$\Psi(X_1, X_2, \theta) = X_1 - \theta + X_2 - \theta = (X_2 - X_1) + 2(X_1 - \theta) > X_2 - X_1 .$$

If $\theta > X_2$ then

$$\Psi(X_1, X_2, \theta) = \theta - X_2 + \theta - X_1 = (X_2 - X_1) + 2(\theta - X_2) > X_2 - X_1 .$$

Thus, the median works as the minimizer for $n = 2$. An extension of this idea applies to the case of general $n$.

(4) Problem 44, page 294: $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with density function,

$$f(x \mid \theta) = (\theta + 1)x^\theta, \quad 0 \leq x \leq 1 .$$

For more on these problem, refer to Problem 4 of the More Problems Section.

(a) Method of Moments Estimate of $\theta$: Let’s compute $E(X_1) = \mu_1$. We have,

$$\mu_1 = (\theta + 1) \int_0^1 x^{\theta+1} dx = (\theta + 1)/((\theta + 2) .$$

Thus we can express $\theta$ as a function of $\mu_1$. Some algebra yields,

$$\theta = \frac{1 - 2\mu_1}{\mu_1 - 1} .$$

It follows easily that,

$$\hat{\theta}_{MOM} = \frac{1 - 2\bar{X}}{\bar{X} - 1} .$$

(b) This is worked out in the More Problems Section, Problem 4.

To answer part(c) we just need to compute $I_1(\theta)$.

$$I_1(\theta) = -E_\theta \left( \frac{\partial^2}{\partial \theta^2} \log f(X_1, \theta) \right)$$

$$= -E_\theta \left( \frac{\partial^2}{\partial \theta^2} (\log(\theta + 1) + \theta \log X_1) \right)$$

$$= -E_\theta \left( -\frac{1}{(\theta + 1)^2} \right)$$

$$= \frac{1}{(\theta + 1)^2} .$$
It follows that
\[ \sqrt{n} (\hat{\theta} - \theta) \xrightarrow{d} N(0, (\theta + 1)^2). \]

(5) **Problems 47 (a), (b) and 48** I will not work with the exact numbers, but will derive the mathematical expressions. Letting \( X_1, X_2, X_3, X_4 \) denote the counts in the four cells, we have a multinomial model with \( N = X_1 + X_2 + X_3 + X_4 \). The likelihood function is,

\[ f(X, \theta) = C(X) \left( \frac{\theta + 2}{4} \right)^{X_1} \left( \frac{1 - \theta}{4} \right)^{X_2 + X_3} \left( \frac{\theta}{4} \right)^{X_4}, \]

for some \( C(X) \) not involving \( \theta \). Thus,

\[ \log f(X, \theta) = X_1 \log (2 + \theta) + (X_2 + X_3) \log (1 - \theta) + X_4 \log \theta + D(X), \]

for some \( D(X) \) not involving \( \theta \). Setting

\[ \frac{\partial}{\partial \theta} \log f(X, \theta) = 0 \]

gives,

\[ \frac{X_1}{2 + \theta} - \frac{X_2 + X_3}{1 - \theta} + \frac{X_4}{\theta} = 0, \]

and this can be rewritten as the following quadratic:

\[ X_1 (1 - \theta) \theta - (X_2 + X_3) \theta (2 + \theta) + X_4 (2 + \theta)(1 - \theta) = 0, \]

and this in turn is equivalent to

\[ \theta^2 (X_1 + X_2 + X_3 + X_4) - \theta [X_1 - 2(X_2 + X_3) - X_4] - 2X_4 = 0. \]

We obtain the MLE of \( \theta \) by solving the above quadratic and taking the admissible root.

Now

\[ \hat{\theta}_N \sim_{\text{approximately}} N(0, I_N(\theta)^{-1}) \]

where

\[ I_N(\theta) = -E_\theta \left( \frac{\partial^2}{\partial \theta^2} \log f(X, \theta) \right). \]

But this is just,

\[ -E_\theta \left( \frac{\partial}{\partial \theta} \left( \frac{X_1}{2 + \theta} - \frac{X_2 + X_3}{1 - \theta} + \frac{X_4}{\theta} \right) \right). \]

This is just

\[ E_\theta \left( \frac{X_1}{(2 + \theta)^2} + \frac{X_2 + X_3}{(1 - \theta)^2} + \frac{X_4}{\theta^2} \right). \]
Using the fact that $X_1 \sim Bin(N, (2 + \theta)/4)$, that $X_2 + X_3 \sim Bin(N, (1 - \theta)/2)$ and that $X_4 \sim Bin(N, \theta/4)$ we get,

$$I_N(\theta) = \frac{N}{4} \left[ \frac{1}{2 + \theta} + \frac{2}{1 - \theta} + \frac{1}{\theta} \right].$$

This simplifies to,

$$I_N(\theta) = \frac{N}{2} \left[ \frac{\theta + 1}{\theta (2 + \theta)(1 - \theta)} \right].$$

Thus the asymptotic variance of the MLE is,

$$I_N(\theta)^{-1} = \frac{2 \theta (2 + \theta)(1 - \theta)}{N (\theta + 1)}.$$

An approximate 95% confidence interval for $\theta$ is

$$\left[ \hat{\theta} - z_{\alpha/2} \sqrt{I_N(\hat{\theta})^{-1}}, \hat{\theta} + z_{\alpha/2} \sqrt{I_N(\hat{\theta})^{-1}} \right],$$

where $I_N(\theta)$ is estimated by $I_N(\hat{\theta})$, since $\theta$ is unknown.

**Problem 48:** Using the facts that $X_1 \sim Bin(N, (\theta + 2)/4)$ and that $X_4 \sim Bin(N, \theta/4)$, it can be easily checked that $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased (these are MOM estimators). Also it can be checked that,

$$\text{Var}(\hat{\theta}_1) = \frac{16}{N^2} \text{Var}(X_1) = \frac{4 - \theta^2}{N}$$

and

$$\text{Var}(\hat{\theta}_2) = \frac{\theta (4 - \theta)}{N}.$$

The standard errors are computed by plugging in $\hat{\theta}$ for $\theta$ in the expressions for the variances and then taking the square root.

(6) **Problem 49, page 295:** I will work out part (c). This essentially encompasses parts (a) and (b). Now,

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

Now,

$$E \left( \rho \sum_{i=1}^{n} (X_i - \bar{X})^2 \right) = \rho \sigma^2 E \left( (n - 1) \frac{s^2}{\sigma^2} \right) = \rho (n - 1) \sigma^2.$$

We use the fact that $(n - 1) s^2/\sigma^2$ is $\chi^2_{n-1}$ and hence has expectation $n - 1$. Also using the fact that $\text{Var}(\chi^2_{n-1})$ is $2(n - 1)$ we readily get,

$$\text{Var} \left( \rho \sum_{i=1}^{n} (X_i - \bar{X})^2 \right) = 2(n - 1) \rho^2 \sigma^4.$$
The M.S.E. of $\rho \sum_{i=1}^{n} (X_i - \bar{X})^2$ as an estimate of $\sigma^2$ is then,

$$M.S.E = 2(n-1)\rho^2 \sigma^4 + (\rho (n-1) \sigma^2 - \sigma^2)^2$$
$$= \sigma^4 \left[ \rho^2 2(n-1) + (\rho (n-1) - 1)^2 \right].$$

To find the $\rho$ that minimizes the MSE, differentiate the expression within square brackets above, with respect to $\rho$ and set the derivative equal to 0 to get,

$$2 \rho \times 2(n - 1) + 2 (\rho (n-1) - 1) (n - 1) = 0,$$

which yields,

$$2 \rho + \rho (n - 1) - 1 = 0,$$

which yields,

$$\rho = 1/(n + 1).$$

It can be checked that this is indeed a minimum (compute the sign of the second derivative). The estimators in (a) are special cases of the more general estimator in (c) with $\rho = 1/n$ and $\rho = 1/(n - 1)$. Note that the MSE as a function of $\rho$ is bowl-shaped hitting a minimum at $\rho = 1/(n + 1)$. To the right of $1/(n + 1)$, the MSE keeps on increasing. Therefore the MSE with $\rho = 1/n$ will be smaller than the MSE with $\rho = 1/(n - 1)$. In other words, $\hat{\sigma}^2$ has smaller MSE than $s^2$. 