1. The following data show the number of A’s obtained in a population of 5 students enrolled in an Accounting PhD program at the Micanopy Institute of Technology.

<table>
<thead>
<tr>
<th>Student</th>
<th>Number of A’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alma</td>
<td>3</td>
</tr>
<tr>
<td>Bud</td>
<td>0</td>
</tr>
<tr>
<td>Carmella</td>
<td>1</td>
</tr>
<tr>
<td>Dora</td>
<td>3</td>
</tr>
<tr>
<td>Elmer</td>
<td>2</td>
</tr>
</tbody>
</table>

(i) Define the random variable to be the number of A’s that a student obtains. Write down the probability distribution of $X$ and then calculate $E(X)$ and $Var(X)$.

(ii) List the ten possible samples of size 3 which we could take from this population if we were to sample without replacement (i.e., name the students who would be in each of these ten samples). Then calculate the ten sample means associated with these samples.

(iii) Use your answers to (ii) to derive the distribution of the random variable $X$.

(iv) Calculate the mean of $X$ (a) using your answers from part (i), (b) directly from the probability distribution of $X$ from part (iii).

(v) Explain why it is not appropriate to use the formula $Var(X) = \sigma^2/n$ in this particular example, and then calculate the $Var(X)$ from the probability distribution found in part (iii).

(i) Probability distribution - just add up the number of occurrences of each number of A's and divide by the total (5).

<table>
<thead>
<tr>
<th># A's</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$1/5$</td>
</tr>
<tr>
<td>1</td>
<td>$1/5$</td>
</tr>
<tr>
<td>2</td>
<td>$1/5$</td>
</tr>
<tr>
<td>3</td>
<td>$2/5$</td>
</tr>
</tbody>
</table>

\[ E(X) = \sum_{x} xf(x) = 0\left(\frac{1}{5}\right) + 1\left(\frac{1}{5}\right) + 2\left(\frac{1}{5}\right) + 3\left(\frac{2}{5}\right) = \frac{1}{5} + \frac{2}{5} + \frac{6}{5} \Rightarrow E(X) = \frac{9}{5} = 1.8 \]

\[ Var(X) = \sum_{x} (x - \mu)^2 f(x) = \\
(0 - 1.8)^2\left(\frac{1}{5}\right) + (1 - 1.8)^2\left(\frac{1}{5}\right) + (2 - 1.8)^2\left(\frac{1}{5}\right) + (3 - 1.8)^2\left(\frac{2}{5}\right) \Rightarrow Var(X) = \frac{34}{25} = 1.36 \]
(ii) Just using the first letter for the students.

<table>
<thead>
<tr>
<th>Samples</th>
<th># A's</th>
<th>Sample Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>4</td>
<td>$\frac{4}{3} = 1.333$</td>
</tr>
<tr>
<td>ABD</td>
<td>6</td>
<td>$\frac{6}{3} = 2$</td>
</tr>
<tr>
<td>ABE</td>
<td>5</td>
<td>$\frac{5}{3} = 1.667$</td>
</tr>
<tr>
<td>ACD</td>
<td>7</td>
<td>$\frac{7}{3} = 2.333$</td>
</tr>
<tr>
<td>ACE</td>
<td>6</td>
<td>$\frac{6}{3} = 2$</td>
</tr>
<tr>
<td>ADE</td>
<td>8</td>
<td>$\frac{8}{3} = 2.667$</td>
</tr>
<tr>
<td>BCD</td>
<td>4</td>
<td>$\frac{4}{3} = 1.333$</td>
</tr>
<tr>
<td>BCE</td>
<td>3</td>
<td>$\frac{3}{3} = 1$</td>
</tr>
<tr>
<td>BDE</td>
<td>5</td>
<td>$\frac{5}{3} = 1.667$</td>
</tr>
<tr>
<td>CDE</td>
<td>6</td>
<td>$\frac{6}{3} = 2$</td>
</tr>
</tbody>
</table>

(iii) Probability distribution - just add up the number of occurrences of each sample mean and divide by the total (10).

<table>
<thead>
<tr>
<th>Sample Mean</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>$\frac{4}{3}$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\frac{5}{3}$</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.3</td>
</tr>
<tr>
<td>$\frac{7}{3}$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\frac{8}{3}$</td>
<td>0.1</td>
</tr>
</tbody>
</table>

(iv) (a) $E(\bar{X}) = E(X) = \boxed{1.8}$

(b) $E(\bar{X}) = \sum \bar{X}_i f(\bar{X}_i) = 1(0.1) + (\frac{4}{3})(0.2) + (\frac{5}{3})(0.2) + 2(0.3) + (\frac{7}{3})(0.1) + (\frac{8}{3})(0.1) = \boxed{1.8}$

(v) The distribution in part (iii) doesn't come from a random sample because it was generated by looking at every combination of the five students rather than random draws.

$$Var(\bar{X}) = \sum (\bar{X}_i - E(\bar{X}))^2 f(\bar{X}_i) = (1 - 1.8)^2(0.1) + (\frac{4}{3} - 1.8)^2(0.2) + (\frac{5}{3} - 1.8)^2(0.2) + (2 - 1.8)^2(0.3) + (\frac{7}{3} - 1.8)^2(0.1) + (\frac{8}{3} - 1.8)^2(0.1) = \frac{17}{75} = 0.22667$$
2. (a) Generate 100 samples of 10 observations drawn from a standard normal distribution.
   (i) For each sample calculate sample mean and sample variance.
   (ii) Plot a histogram of the sample means and sample variances, indicating on your plot the true
        mean or variance of the random variables. Comment on the relationship between the sample
        mean and the true mean, the sample variance and the true variance.
   (b) Repeat the above exercise for exponentially distributed random variables with parameter
       \( \theta = 2 \).
   (c) Repeat the above exercise for the Weibull distribution with distribution function

\[
F(x) = 1 - \exp \left[ -\left(\frac{x}{\theta}\right)^{1.5} \right]
\]

   (i) Standard Normal: don't need \( f(x) \) of \( F(x) \) because we can use \texttt{invnorm} command in
       Stata, \( E(X) = 0 \), \( Var(X) = 1 \)

Sample program for Stata (from Scott). I added the red line to get the program to work.

```
program define nsim /*arguments are obs iterations*/
version 6
args n it
tempname sim
postfile `sim' mean var using results, replace
quietly {
    local i=1
    while `i' <= `it' {
        drop _all
        set obs `n'
        gen z = invnorm(uniform())
        summarize z
        post `sim' r(mean) r(Var)
        local i = `i' + 1
    }
}
postclose `sim'
use results,clear
end
```
Results:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>100</td>
<td>.0005718</td>
<td>.3383416</td>
<td>-1.118136</td>
<td>1.017259</td>
</tr>
<tr>
<td>var</td>
<td>100</td>
<td>.9885811</td>
<td>.4143226</td>
<td>.0791966</td>
<td>2.038543</td>
</tr>
</tbody>
</table>

(ii) Exponential Distribution: 
\[ f(x) = \theta e^{-\theta x}, \quad \frac{1}{\theta} = E(X) = 0.5, \quad \frac{1}{\theta^2} = Var(X) = 0.25 \]

CDF: 
\[ F(x) = \int_0^x \theta e^{-\theta y} \, dy = -e^{-\theta x} \bigg|_0^x = -e^{-\theta x} - (-e^0) = 1 - e^{-\theta x} \]

Solve CDF for \( x \):
\[ e^{-\theta x} = 1 - F(x) \Rightarrow -\theta x = \ln(1 - F(x)) \Rightarrow x = -\frac{\ln(1 - F(x))}{\theta} \]

Generate Uniform(0,1) and plug it in for \( F(x) \) to generate values for \( x \)

Just change green line in program above to this:
\[ gen \ z = -\ln(1 - \text{uniform}())/2 \]

Results:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>100</td>
<td>.4997093</td>
<td>.1652599</td>
<td>.149499</td>
<td>1.164545</td>
</tr>
<tr>
<td>var</td>
<td>100</td>
<td>.2295738</td>
<td>.1771686</td>
<td>.002916</td>
<td>.963869</td>
</tr>
</tbody>
</table>
(iii) Weibull Distribution: 

\[
\frac{dF(x)}{dx} = -\exp\left[-\left(\frac{x}{2}\right)^{1.5}\right] \cdot \left(-1.5\left(\frac{x}{2}\right)^{0.5}\right) \cdot \frac{1}{2} = \\
1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]
\]

\[
E(X) = \int_{0}^{\infty} xf(x)\,dx = \int_{0}^{\infty} x \cdot 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dx = \\
\int_{0}^{\infty} 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot x \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dx
\]

Let \( y = \left(\frac{x}{2}\right)^{1.5} \) \( \Rightarrow \frac{x}{2} = y^{2/3} \Rightarrow x = 2y^{2/3} \)

Note: \( \frac{dy}{dx} = 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \) so multiply \( E(X) \) by 1 = \( \frac{dy}{dx} \cdot \frac{dy}{dx} = \frac{dy}{dx} \cdot 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} 
\]

\[
E(X) = \int_{0}^{\infty} 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot x \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dx \cdot \frac{dy}{dx} \cdot \frac{dy}{dx} = \int_{0}^{\infty} \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dy
\]

Now substitute \( x = 2y^{2/3} \) for the first term and \( \left(\frac{x}{2}\right)^{1.5} = y \) for the second

\[
E(X) = \int_{0}^{\infty} 2y^{2/3} e^{-y}\,dy = 2\int_{0}^{\infty} y^{5/3-1} e^{-y}\,dy = 2\Gamma(5/3) = 2(0.9027) = E(X) = 1.8054
\]

\[
E(X^2) = \int_{0}^{\infty} x^2 f(x)\,dx = \int_{0}^{\infty} x^2 \cdot 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dx = \\
\int_{0}^{\infty} 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot x^2 \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dx
\]

Use the same \( y \) from before \( \Rightarrow \ x^2 = 4y^{4/3} \)

\[
E(X^2) = \int_{0}^{\infty} 1.5\left(\frac{x}{2}\right)^{0.5} \cdot \frac{1}{2} \cdot x^2 \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dx \cdot \frac{dy}{dx} \cdot \frac{dy}{dx} = \\
\int_{0}^{\infty} x^2 \exp\left[-\left(\frac{x}{2}\right)^{1.5}\right]\,dy
\]

Now substitute \( x^2 = 4y^{4/3} \) for the first term and \( \left(\frac{x}{2}\right)^{1.5} = y \) for the second

\[
E(X^2) = \int_{0}^{\infty} 4y^{4/3} e^{-y}\,dy = 4\int_{0}^{\infty} y^{7/3-1} e^{-y}\,dy = 4\Gamma(7/3) = 4(1.1906) = 4.7624
\]

\[
Var(X) = E(X^2) - [E(X)]^2 = 4.7624 - (1.8054)^2 = Var(X) = 1.5029
\]

Solve CDF for \( x \):
\[
\exp\left[-\left(\frac{x}{2}\right)^{1.5}\right] = 1 - F(x) \quad \Rightarrow -\left(\frac{x}{2}\right)^{1.5} = \ln(1 - F(x)) \quad \Rightarrow \\
x/2 = \left[-\ln(1 - F(x))\right]^{\frac{1}{1.5}} \quad \Rightarrow \quad x = 2\left[-\ln(1 - F(x))\right]^{\frac{2}{3}}
\]

Generate Uniform(0,1) and plug it in for \( F(x) \) to generate values for \( x \)

Just change green line in program above to this:
\[
\text{gen } z = 2\times(-\ln(1 - \text{uniform()}))^{(2/3)}
\]

Results:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>100</td>
<td>1.807677</td>
<td>.4113603</td>
<td>.8243586</td>
<td>3.266533</td>
</tr>
<tr>
<td>var</td>
<td>100</td>
<td>1.44065</td>
<td>.7496428</td>
<td>.0505821</td>
<td>3.727319</td>
</tr>
</tbody>
</table>

![Histograms of variable mean and var](histograms.png)
3. Suppose that the random variable $X$ has an unknown distribution with a mean of $\mu$ and a variance of $\sigma^2$. Given a random sample of $n$ observations, calculate $\bar{x}_{n-1}$ (the sample mean of the first $n - 1$ observations), and then form $\hat{\mu}$ defined by 

$$\hat{\mu} = 0.5(\bar{x}_{n-1} + x_n)$$

(i) Show that $\hat{\mu}$ is an unbiased estimator of $\mu$ and then explain why $\hat{\mu}$ is not consistent for $\mu$.

(ii) Find $\text{Var}(\hat{\mu})$, compare it with $\text{Var}(\bar{x}_{n-1})$, and then discuss the efficiency of $\hat{\mu}$. Does $\hat{\mu}$ attain the Cramer-Rao lower bound? Explain.

(ii)

$$\text{Var}(\bar{x}_{n-1}) = \text{Var}\left(\frac{1}{n-1}\sum_{i=1}^{n-1} x_i\right) = \frac{1}{(n-1)^2}\sum_{i=1}^{n-1} \text{Var}(x_i) = \frac{1}{(n-1)^2} \sum_{i=1}^{n-1} \sigma^2 = \frac{1}{(n-1)^2} (n-1)\sigma^2$$

$$\text{Var}(\hat{\mu}) = \text{Var}\left[0.5(\bar{x}_{n-1} + x_n)\right] = (0.5)^2 \left[\text{Var}(\bar{x}_{n-1}) + \text{Var}(x_n)\right] = 0.25 \left[\frac{\sigma^2}{n-1} + \sigma^2\right]$$

$$0.25 \left[\frac{\sigma^2}{n-1} + \frac{n\sigma^2 - \sigma^2}{n-1}\right] = 0.25n\sigma^2 > \frac{\sigma^2}{n-1} \text{ when } n > 4$$

Since $\text{Var}(\hat{\mu}) > \text{Var}(\bar{x}_{n-1})$ (when $n > 4$), $\hat{\mu}$ does not attain the Cramer-Rao lower bound.

4. If $X$ is a random variable such that $E(X) = 3$ and $E(X^2) = 13$

(a) Use Chebyshev’s inequality to determine a lower bound for the probability $\Pr(-2 < X < 8)$.

(b) Find the above probability if $X$ has a normal distribution.

(a)

Chebyshev’s Inequality: for constant $k > 0$, $\Pr[|X - E(X)| \geq k] \leq \text{Var}(X)/k^2$

Write for lower bound by flipping the inequalities: $\Pr[|X - E(X)| < k] > 1 - \text{Var}(X)/k^2$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = 13 - 9 = 4$$

For this problem: $\Pr[3-k < X < k+3] > 1 - 4/k^2$

Since $3-k = -2$ and $k+3 = 8$, $k = 5$

$\therefore \Pr[-2 < X < 8] > 1 - 4/25$

$$\frac{21}{25} = 0.84$$
(b) \( \Pr[-2 < X < 8] = \Pr\left[\frac{-2-3}{2} < \frac{X-3}{2} < \frac{8-3}{2}\right] = \Pr\left[\frac{-5}{2} < z < \frac{5}{2}\right] = \Pr[z < \frac{5}{2}] - \Pr[z < -\frac{5}{2}] = 0.9938 - 0.0062 = 0.9876 \)

5. The density function of a random variable is \( f(x; \theta) = (\ln \theta)x^{-\gamma}, \ x > 0 \text{ and } \theta > 1 \). A random sample of size \( n \) is drawn.

(i) Derive the score function and the MLE of \( \theta \).

(ii) Show that the mean and variance of \( X \) are \( E(X) = 1/(\ln \theta) \) and \( Var(X) = 1/(\ln \theta)^2 \). Use this to derive the information \( I(\theta) \) and the Cramer-Rao lower bound for the variance of \( \overline{X} \). Is the lower bound attained?

(iii) Construct a random variable \( Z_n \), which is a function of \( x \) and \( \theta \), such that its asymptotic distribution is \( N(0,1) \).

(i) \( L(x; \theta) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} (\ln \theta)^{-\gamma} \left(\theta^{-\gamma} \right)^{-1} \sum_{i=1}^{n} x_i = \prod_{i=1}^{n} (\ln \theta)^{\gamma} = (\ln \theta)^{\gamma} \) \( \theta \) \( \sum_{i=1}^{n} x_i \)

\[ \ln L(x; \theta) = n \ln(\ln \theta) - \left(\sum_{i=1}^{n} x_i \right) \ln \theta \]

Score \( = S(\hat{\theta}) = \frac{\partial \ln L(x; \theta)}{\partial \theta} \Rightarrow S(\hat{\theta}) = \frac{n \ln \ln \theta}{\theta} - \frac{1}{\theta} \sum_{i=1}^{n} x_i \)

MLE: Solve \( S(\hat{\theta}) = 0 \)

\[ \frac{n}{\theta \ln \theta} - \frac{1}{\theta} \sum_{i=1}^{n} x_i = 0 \Rightarrow \frac{n}{\theta \ln \theta} = \frac{1}{\theta} \sum_{i=1}^{n} x_i \Rightarrow \frac{n}{\ln \theta} = \sum_{i=1}^{n} x_i \Rightarrow \ln \theta = \frac{n}{\sum_{i=1}^{n} x_i} \]

\[ \hat{\theta} = \exp\left[\frac{1}{\sum_{i=1}^{n} x_i}\right] \Rightarrow \hat{\theta} = \exp\left[\frac{1}{\overline{X}}\right] \]

(ii) Substitute \( 1/\gamma = \ln(\theta) \), now \( f(x; \theta) = (\ln \theta)e^{-x/\ln(\theta)} = \frac{1}{\gamma} e^{-\frac{x}{\gamma}} \), exponential distribution, therefore, \( E(X) = \gamma = \frac{1}{\ln \theta} \) and \( Var(X) = \gamma^2 = \frac{1}{(\ln \theta)^2} \)

\[ I(\theta) = -E\left[\frac{\partial^2 \ln L(\theta; x)}{\partial \theta^2}\right] = -E\left[\frac{-n}{(\theta \ln \theta)^2} \left(\ln \theta + \frac{\theta}{\theta}\right) + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i \right] = \]

\[ \frac{n}{(\theta \ln \theta)^2} \left(\ln \theta + 1\right) - \frac{1}{\theta^2} \sum_{i=1}^{n} E(x_i) = \frac{n}{\theta^2 \ln \theta} + \frac{n}{\theta^2 (\ln \theta)^2} - \frac{1}{\theta^2} \sum_{i=1}^{n} \frac{1}{\ln \theta} = \frac{n}{\theta^2 \ln \theta} - \frac{n}{\theta^2 \ln \theta} - \frac{n}{\theta^2 (\ln \theta)^2} = I(\theta) = \frac{n}{\theta^2 (\ln \theta)^2} \]
Cramer-Rao Inequality - let \( x_i \sim f(x; \theta) \) and \( T = T(x_1, \ldots, x_n) \) be a statistic such that 
\[
E(T) = u(\theta) \text{ (some function of } \theta) \text{. Assume regularity conditions. Then}
\]
\[
\text{Var}(T) \geq \frac{[u'(\theta)]^2}{I(\theta)}
\]
Let \( T = \bar{x} \)

From \( \hat{\theta} = \exp \left[ \frac{1}{\bar{x}} \right] \) (in part (i), we get \( \bar{x} = \frac{1}{\ln \theta} \)

Therefore, \( E(T) = u(\theta) = \frac{1}{\ln \theta} \)

\[
u'(\theta) = \frac{-1}{(\ln \theta)^2} \cdot \frac{1}{\theta}
\]

From Cramer-Rao inequality, \( \text{Var}(\bar{x}) \geq \frac{[u'(\theta)]^2}{I(\theta)} = \frac{\frac{-1}{(\ln \theta)^2} \cdot \frac{1}{\theta}}{\frac{n}{\theta^2 (\ln \theta)^2}} = \frac{1}{\theta^2 (\ln \theta)^2} \cdot \frac{\theta^2 (\ln \theta)^2}{n} = \frac{1}{n(\ln \theta)^2} \)

(iii) Asymptotic Normality - \( \lim_{n \to \infty} \frac{1}{\sqrt{n}}(\hat{\theta}_n - \theta) \overset{d}{\to} N(0, \Sigma^{-1}(\theta)) \); i.e., MLEs have normal distribution at the limit

In this case, we don't need matrices because there is only 1 parameter, therefore,

\[
\sqrt{n}(\hat{\theta} - \theta) \overset{d}{\to} N(0, nI(\theta))
\]

To make this \( N(0,1) \), we multiply both sides by \( \sqrt{I/n} : \)

\[
\sqrt{n}(\hat{\theta} - \theta) \cdot \sqrt{I/n} \overset{d}{\to} N(0, nI(\theta)) \cdot \frac{\sqrt{I/n}}{n} \Rightarrow \sqrt{I}(\hat{\theta} - \theta) \overset{d}{\to} N(0,1)
\]

Therefore, set random variable \( Z_n = \sqrt{I}(\hat{\theta} - \theta) = \sqrt{n} \left( \frac{1}{\theta \ln \theta} \left( e^{1/\bar{x}} - \theta \right) \right) \)

6. The Pareto distribution is frequently used as a model to study income distribution. It has the following CDF (not PDF): \( F(x) = 1 - (\theta_1 / x)^{\theta_2} \text{ for } x \geq \theta_1 \), and zero elsewhere. \( \theta_1 \) and \( \theta_2 \) are both positive.

(i) Derive the density function \( f(x) \).

(ii) Given a random sample of \( n \) observations, find the MLE of \( \theta_1 \) and \( \theta_2 \).

(iii) Derive the information matrix for \( \theta_1, \theta_2 \).

(iv) Derive the Cramer-Rao lower bounds for the variances of the estimators of \( \theta_1 \) and \( \theta_2 \). (You do not need to verify whether these lower bounds are attained.)
\(\textbf{I}(\theta) = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}, \text{ where } I_{ij} = -E \left[ \frac{\partial^2 \ln L(x; \theta)}{\partial \theta_i \partial \theta_j} \right] \)

\(\begin{align*}
I_{11} &= -E \left[ \frac{\partial^2 \ln L(x; \theta)}{\partial \theta_1^2} \right] = -E \left[ \frac{\partial}{\partial \theta_1} \left( \frac{n \theta_1}{\theta_1} \right) \right] = -E \left[ \frac{n \theta_1}{\theta_1^2} \right] = \frac{n \theta_1}{\theta_1^2} \\
I_{12} &= I_{21} = -E \left[ \frac{\partial^2 \ln L(x; \theta)}{\partial \theta_1 \partial \theta_2} \right] = -E \left[ \frac{\partial}{\partial \theta_1} \left( \frac{n \theta_2}{\theta_1} \right) \right] = -E \left[ \frac{n \theta_2}{\theta_1^2} \right] = \frac{-n \theta_2}{\theta_1^2} \\
I_{22} &= -E \left[ \frac{\partial^2 \ln L(x; \theta)}{\partial \theta_2^2} \right] = -E \left[ \frac{\partial}{\partial \theta_2} \left( \frac{n \theta_2}{\theta_1} + n \ln(\theta_1) - \sum_{i=1}^{n} \ln(x_i) \right) \right] = -E \left[ \frac{-n}{\theta_2^2} \right] = \frac{n}{\theta_2^2}
\end{align*}\)

\[\text{MLE: Solve } \hat{\theta} = \min(x_i)\]
(iv) CRLB is $[\mathbf{I}(\theta)]^{-1}$

Using

$$
\begin{bmatrix}
  a & b \\
  c & d
\end{bmatrix}^{-1} = \frac{1}{ad - cb}
\begin{bmatrix}
  d & -b \\
  -c & a
\end{bmatrix}
$$

$$
[I(\theta)]^{-1} = \frac{1}{n\theta_2\left(\frac{n}{\theta_1}\right) - n\left(\frac{n}{\theta_1}\right)}
\begin{bmatrix}
  n & n \\
  -\frac{n}{\theta_1} & -\frac{n}{\theta_1}
\end{bmatrix}
= \frac{1}{n^2 - n^2\theta_2^2}
\begin{bmatrix}
  \frac{n}{\theta_1} & \frac{n}{\theta_1} \\
  \frac{n}{\theta_2} & \frac{n}{\theta_2}
\end{bmatrix}
= [I(\theta)]^{-1} =
\begin{bmatrix}
  \frac{\theta_1^2}{n\theta_2(1-\theta_2)} & \frac{\theta_1\theta_2}{n(1-\theta_2)} \\
  \frac{\theta_1\theta_2}{n\theta_2(1-\theta_2)} & \frac{\theta_2^2}{n(1-\theta_2)}
\end{bmatrix}
$$

Documentation

Prof Werner practically did the entire homework assignment for me... and thankfully didn't hit me after my 50th or so question. I sat down with her on 16 Oct and went over some general points on the homework. For problem 1, Prof Werner told me part (iv)(a) just uses $E(X)$ from part (i). She also told me the answer to part (v) (i.e., not from a random sample). On problem 2, Prof Werner verified my CDF for the exponential distribution. For problem 3, Prof Werner said it was a random sample (i.e., $x$'s independent). She pretty much gave me the answer to the second part of part (i). She also showed me how to do part (ii) without computing the CRLB. For problem 4, Prof Werner confirmed that $k = 5$. For problem 5, Prof Werner reviewed the methodology for finding $L(x; \theta)$, $\ln L(x; \theta)$, the score, and the MLE. For problem 6, Prof Werner verified my answers for $L(x; \theta)$, $\ln L(x; \theta)$.

I met with Prof Werner again in and out of class on 21 Oct. She confirmed that I had to solve the CDFs for $x$ in parts (b) and (c) of problem 2. She also figured out the mean and variance of the Weibull distribution for me. On problem 4, Prof Werner showed us how to reverse Chebyshev's inequality to make it a lower bound. Josh Kneifel found a math error because I was dividing by 13 instead of by 2. On problem 5, Prof Werner told us to make the substitution to see we had an exponential distribution in order to find $E(X)$ and $\text{Var}(X)$ in part (ii). She also pointed out that $\bar{x}$ should be in the MLE for $\hat{\theta}$ and told me to use the second derivative for computing the information matrix. She gave the answer to the CRLB and part (iii) in class. For problem 6 part (ii), Prof Werner told us in class to use $\hat{\theta}_1 = \min(x_i)$ because we can't set either $\theta$ equal to zero (similar to an example she did in class). She reminded me after class that the information matrix would be $2 \times 2$ and that I had to take the inverse to get the CRLB.

Scott (TA) save us a template for generating the samples in Stata. He also gave us some basic commands to get the overall mean and variance of the samples and to plot the histograms.