1. Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a uniform distribution on the interval $(0, \theta)$.

a. Obtain the method of moment estimator of $\theta$. [5]

$$
\mathbb{E}(X) = \frac{\theta}{2} \quad \Rightarrow \quad \bar{X} = \frac{\theta}{2}
$$

$$
\Rightarrow \quad \hat{\theta} = 2 \bar{X}
$$

b. Is the moment estimator for $\theta$ unbiased? [5]

$$
\mathbb{E}(\bar{X}) = \mathbb{E}(X) = \frac{\theta}{2} \quad \Rightarrow \quad \mathbb{E}(\hat{\theta}) = \mathbb{E}(2 \bar{X}) = 2 \mathbb{E}(\bar{X}) = \frac{2\theta}{2} = \theta
$$

$$
\Rightarrow \quad \mathbb{E}(\hat{\theta}) = \theta, \quad \text{so} \quad \hat{\theta} \text{ is an unbiased estimator}
$$

2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from the distribution with probability density function

$$
f(x; \theta) = \begin{cases} 
\frac{1}{\theta} x^\frac{1}{\theta - 1} & 0 \leq x \leq 1 \\
0 & \text{otherwise}
\end{cases}
$$

a. Obtain the method of moments estimator of $\theta$. [10]

$$
\mathbb{E}(X) = \int_0^\infty x \frac{1}{\theta} x^\frac{1}{\theta - 1} \, dx = \int_0^\infty \frac{x^\frac{1}{\theta}}{\theta} \, dx
$$

$$
= \frac{1}{\theta} \left[ \frac{x^\frac{1}{\theta} + 1}{\frac{1}{\theta} + 1} \right]_0^\infty
$$

$$
= \frac{1}{\theta} \left( \frac{1}{\frac{1}{\theta} + 1} \right) = \frac{1}{\theta + 1}
$$

$$
\Rightarrow \quad \bar{X} = \frac{1}{\theta + 1} \quad \Rightarrow \quad \hat{\theta} = \frac{1}{\bar{X}}
$$
b. Obtain the maximum likelihood estimator of $\theta$.\[10\]

$$L(x; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \frac{1}{\bar{x}^n (e^{\theta})^{\sum_{i=1}^{n} x_i}}$$

$$\ln(L) = -n \ln(\theta) + \frac{1-\theta}{\theta} \sum_{i=1}^{n} \ln(x_i)$$

$$d \ln(L) \over d\theta = 0 \Rightarrow -n - \frac{1}{\theta} \sum_{i=1}^{n} \ln(x_i) = 0$$

$$\Rightarrow \hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln(x_i)}$$

3. The superintendent of a large school district, having once had a course in probability and statistics, believes that the number of teachers absent on any given day has a Poisson distribution with parameter $\lambda$. Use the accompanying data on absences for 50 days to derive a 95% CI for $\lambda$. [Hint: The mean and variance of a Poisson variable both equal $\lambda$, so $Z=(\bar{X} - \lambda)/\sqrt{\lambda/n}$ has approximately a standard normal distribution. $\theta = \lambda, \hat{\theta} = \bar{X}$ and $\sigma_{\hat{\theta}} = \sqrt{\frac{\lambda}{n}} \Rightarrow \sigma_{\theta} = \sqrt{\frac{\bar{X}}{n}}$ [10]

<table>
<thead>
<tr>
<th># Absences</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

$$\sum x_i = 205$$, so $\bar{x} = \frac{205}{50} = 3.9423$

95% CI for $\lambda$ is

$$\left[ 3.9423 \pm 1.96 \sqrt{\frac{3.9423}{50}}, 4.482 \right]$$
b. How large must $n$ be if the width of the 99% interval for $\mu$ is to be 1.0?[5]

$$
\begin{align*}
 n &= \left[ \frac{2 \times Z_{\alpha/2}}{\sqrt{\frac{\sigma}{n}}} \right]^2 \\
 &= \left[ \frac{2 \times 2.575}{\sqrt{\frac{3.9423}{10}}} \right]^2 \\
 &= 10.456
\end{align*}
$$


4. Corn soy blend (CSB) is often provided to relief agencies for distribution in refugee camps or as emergency relief. Production of CSB is carefully monitored to ensure that the product has appropriate levels of certain vitamins and minerals for good nutrition. 8 bags of CSB were randomly taken from a large shipment and the amount of vitamin C was measured. The measured concentration in milligrams of vitamin C per 100 grams of CSB were

$$26, 31, 23, 22, 11, 22, 14, 31$$

a. Construct a 95% confidence interval for the true mean amount of Vitamin C per 100 grams of CSB.[5]

Since the sample size is small and $\sigma$ unknown, we use $t$-intervals.

$$
\bar{x} = 22.5 \quad \text{and} \quad s^2 = 51.7143 \quad \Rightarrow \quad s = 7.1913,
\quad s^2 = \frac{s}{n} = 0.625 \quad \Rightarrow \quad t_{0.025, 6} = 2.365
\quad \Rightarrow \quad \left[ \bar{x} - t_{0.025, n-1} \frac{s}{\sqrt{n}}, \quad \bar{x} + t_{0.025, n-1} \frac{s}{\sqrt{n}} \right]
\begin{align*}
&= \left[ 22.5 - 2.365 \times 7.1913, \quad 22.5 + 2.365 \times 7.1913 \right] \\
&= \left[ 16.487, \quad 28.513 \right]
\end{align*}
$$

b. Construct a 95% upper bound for the true variance of vitamin C per 100 grams of CSB.[5]

$$
\begin{align*}
\frac{(n-1)s^2}{\chi^2_{1-\alpha, n-1}} &= \frac{7 \times 51.7143}{2.167} = 167.14 \\
\chi^2_{0.95, 7} &= 2.167
\end{align*}
$$

2. Construct a 95% prediction interval of vitamin C per 100 grams of CSB.[5]

$$
\begin{align*}
\bar{x} \pm t_{0.025, n-1} \frac{s}{\sqrt{n}} \\
22.5 \pm 2.365 \times 7.1913 \sqrt{1 + \frac{1}{8}} \\
[4.465, \quad 40.465]
\end{align*}
$$
5. A marketing company is planning a survey to estimate the proportion of high-school students read newspapers daily.

a. What sample size should be used if the company wants to give a 95% confidence interval which is guaranteed to be more than 4 percentage points wide by assuming equal proportion for reading and not reading newspapers daily?[5]

\[ n = 2 \times 1.96 \sqrt{\frac{p(1-p)}{\sigma}} = 2 \times 1.96 \sqrt{\frac{0.5(0.5)}{\sigma}} \]

\[ = 1.96 \sqrt{\frac{\sigma}{n}} \]

So, \[ \frac{1.96}{\sqrt{n}} \leq 0.04 \Rightarrow \frac{1.96}{0.04} = 49 \]

\[ \Rightarrow n = (49)^2 = 2401 \]

b. If the company is certain that the true proportion of high school students who read a newspaper daily is less than 25%, what sample size should they use?[5]

The company is sure that \( p < 0.25 \Rightarrow p(1-p) = (0.25)(0.75) = 0.1875 \)

So, \[ n = 2 \times 1.96 \sqrt{\frac{p(1-p)}{\sigma}} = 2 \times 1.96 \sqrt{\frac{0.25(0.75)}{\sigma}} \]

\[ = 1.6974 \]

So the interval be no wider than \( 0.04 \Rightarrow \frac{1.6974}{\sqrt{n}} \leq 0.04 \Rightarrow \frac{1.6974}{0.04} = 42.44 \Rightarrow n \geq 1800.75 \]

c. In fact the company used a sample of size \( n = 2000 \) students and found that 426 said they read a newspaper every day. Find a 95% confidence interval for the true population proportion.[5]

\[ \left[ \hat{p} - Z_{0.025} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{0.025} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right] \]

\[ \left[ 0.215 - 1.96 \sqrt{\frac{(0.215)(0.785)}{2000}}, 0.215 + 1.96 \sqrt{\frac{(0.215)(0.785)}{2000}} \right] \]

\[ = [0.1951, 0.2309] \]