1. (15 pts.) The following histogram contains 62 data values—namely 62 intervals, in days, between "serious" earthquakes, across the globe, from December 16, 1902 through March 4, 1977.

![Histogram of earthquake intervals](image)

**a.** The data are best modeled by a(n) (choose one):

- i. Exponential
- ii. Gamma
- iii. Normal
- iv. Poisson

**B. METHOD OF MOMENTS**

Set \( \lambda = \bar{x} \)

i.e. \( \frac{1}{\lambda} = \frac{44}{\lambda} \Rightarrow \hat{\lambda}_{\text{MM}} = \frac{1}{44} \)

**b.** The average of the 62 values is approximately 440 days. Suppose a "serious" earthquake has just occurred. Estimate the chance that the next earthquake occurs within 44 days.

\[
P(\text{Interval} \leq 44) = \int_{0}^{44} \frac{1}{440} e^{-\frac{x}{440}} \, dx
\]

\[
= e^{-\frac{44}{440}} \left|^{44}_{0}
\right.
\]

\[
= 1 - e^{-\frac{44}{440}} \approx 0.952
\]

---

1 The data are from Hand, D.J. (1994), *A Handbook of Small Datasets*, Chapman & Hall, p. 203 where a "serious" earthquake is carefully defined.
2. (25 pts.) On the evening of 4/20/04 I generated 250 pseudo-random numbers between 0 and 1 on my TI-83 calculator. As a partial test on the randomness of my calculator, test the hypothesis that numbers are equally likely to fall in any one of the 5 intervals [0.00, 0.20), [0.20, 0.40), [0.40, 0.60), [0.60, 0.80), and [0.80, 1.00) by answering the questions below.

Data:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number of Data Points Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0.00, 0.20)</td>
<td>45</td>
</tr>
<tr>
<td>[0.20, 0.40)</td>
<td>51</td>
</tr>
<tr>
<td>[0.40, 0.60)</td>
<td>57</td>
</tr>
<tr>
<td>[0.60, 0.80)</td>
<td>43</td>
</tr>
<tr>
<td>[0.80, 1.00)</td>
<td>54</td>
</tr>
<tr>
<td>Total = 250</td>
<td></td>
</tr>
</tbody>
</table>

(a) Compute (Pearson's) $\chi^2$ statistic.

(b) Give a rough sketch of the $\chi^2$ curve that the $\chi^2$ statistic follows if the null hypothesis of equally likely intervals is true.

(c) What is the degrees of freedom number for the $\chi^2$ curve?

(d) On your graph (part b), label the numeric value along the horizontal axis for which there is area 0.05 to the right.

(e) On your graph (parts b, d), indicate the region of $\chi^2$ statistic values for which you accept the null hypothesis. Likewise, indicate the region of $\chi^2$ statistic values for which you reject the null hypothesis.

(f) Come to a conclusion about the null hypothesis. That is, do you accept or reject the null hypothesis (use $\alpha = 0.05$)?

\[ \chi^2_{\text{statistic}} = \frac{(45-50)^2}{50} + \frac{(51-50)^2}{50} + \frac{(57-50)^2}{50} + \frac{(43-50)^2}{50} + \frac{(54-50)^2}{50} \]

\[ = \frac{140}{50} = 2.80 \]

\[ \text{df} = 4 \]

Accept $H_0$
3. (10 pts.) A random variable is a Beta random variable if it has density

\[ f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \]

for some parameters \( \alpha \) and \( \beta \). Random variables having such a density have the following mean and variance:

\[ \mu = \frac{\alpha}{\alpha + \beta}, \quad \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} \]

Suppose you have a random variable which you believe to follow some Beta (e.g. by checking the histogram). If \( \bar{x} = 0.60 \) is the average of the data values and \( s^2 = 0.28 \) is the sample variance of the data values state, but do not try to solve, the equation(s) you would need to solve to find the method of moments estimates for \( \alpha \) and \( \beta \).

\[ \begin{align*}
\mu &= \bar{x} \\
\sigma^2 &= s^2
\end{align*} \]

That is, \( \bar{x} \) and \( s^2 \)

Solve to find \( \alpha \) and \( \beta \).

\[ \begin{align*}
\frac{\alpha}{\alpha + \beta} &= 0.60 \\
\frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} &= 0.28
\end{align*} \]
4. (15 pts.) Here is a histogram of the waist circumferences of 507 physically active (several hours of exercise per week) individuals:

![Histogram of waist circumferences](image)

a. Waist girth is best modeled by a(n) (circle one):

   i. Exponential
   ii. Gamma
   iii. Normal
   iv. Poisson

b. Suppose that the average of the 507 data values is about 93.3 cm and the standard variance, \( s^2 \), is about 100.6 cm². Write down, but don't evaluate, an integral which estimates the proportion of such physically active people who have waist circumference of at least 95 cm.

\[
\hat{\beta}_{mm} = \frac{s^2}{\bar{x}} \approx 86.530
\]

\[
\hat{\alpha}_{mm} = \frac{\bar{x}^2}{s^2} \approx 1.078
\]

\[
P(\omega \geq 95) = \int_{95}^{\infty} \frac{1}{\beta^x \Gamma(x)} x^{\alpha-1} e^{-x/\beta} \, dx
\]

where \( \alpha \) is replaced by \( \hat{\alpha}_{mm} = 86.530 \)

\[
\beta = \hat{\beta}_{mm} = 1.078
\]

---

5. (15 pts.) Suppose\(^3\) that cars arrive mid-afternoon to the McDonalds on Mt. Rushmore at a rate of roughly 1 per minute. Also suppose that cars arrive independently.

a. Estimate the chance that no cars arrive in 2 minutes. \( (\lambda = 2) \)

\[ P(\bar{X} = 0) = \frac{2^0}{0!} e^{-2} = e^{-2} \approx 0.1353 \]

b. Estimate the chance that at most 3 cars arrive in 5 minutes. \( (\lambda = 5) \)

\[ P(\bar{X} \leq 3) = P(\bar{X} = 0) + P(\bar{X} = 1) + P(\bar{X} = 2) + P(\bar{X} = 3) \]

\[ = \frac{5^0}{0!} e^{-5} + \frac{5^1}{1!} e^{-5} + \frac{5^2}{2!} e^{-5} + \frac{5^3}{3!} e^{-5} \]

\[ = (1 + 5 + \frac{25}{2} + \frac{125}{6}) e^{-5} \approx 0.2650 \]

\(^3\) Based upon data collected from 2:13-2:58 pm on October 26, 2000.
6. (10 pts.) Let $X_1, X_2, X_3$ be i.i.d. random variables having mean $\mu$ and variance $\sigma^2$.

Compute the mean squared error of $\hat{\mu}$ (an estimate of $\mu$) where

$$\hat{\mu} = \frac{2X_1 - X_2 + 3X_3}{4}$$

$$E(\hat{\mu}) = \frac{2}{4} E(X_1) + \frac{-1}{4} E(X_2) + \frac{3}{4} E(X_3) = \frac{2}{4} \mu + \frac{-1}{4} \mu + \frac{3}{4} \mu = \mu$$

$$\text{Var}(\hat{\mu}) = \left(\frac{2}{4}\right)^2 \text{Var}(X_1) + \left(\frac{-1}{4}\right)^2 \text{Var}(X_2) + \left(\frac{3}{4}\right)^2 \text{Var}(X_3)$$

$$= \frac{1}{4} \sigma^2 + \frac{1}{16} \sigma^2 + \frac{9}{16} \sigma^2$$

$$= \frac{7}{8} \sigma^2$$

$$\text{M.S.E.}(\hat{\mu}) = \text{Var}(\hat{\mu}) + [E(\hat{\mu}) - \mu]^2 = \frac{7}{8} \sigma^2 + \sigma^2 = \frac{7}{8} \sigma^2$$

7. (10 pts.) Multiple Choice:

a. If $m.s.e.(\hat{\mu}_1) < m.s.e.(\hat{\mu}_2)$ where $\hat{\mu}_1$ and $\hat{\mu}_2$ are estimate of $\mu$, then which estimate is to be preferred for estimating $\mu$ (circle one):

i. $\hat{\mu}_1$

ii. $\hat{\mu}_2$

b. If we decide to restrict our attention to unbiased estimates of a parameter using the m.s.e. measure, then this is equivalent to choosing the unbiased estimate having the smallest variance. True or False (circle one).

True

c. To check whether a random quantity can reasonably be modeled by some exponential, especially in the small sample case, we can check an exponential probability plot. For us to reasonably trust an exponential model, the points in this plot should follow what kind of pattern (circle one)?

i. Exponential

ii. Linear

iii. Normal

iv. None of the above