Problem 1

Let $X$ be a stochastic variable with probability density (pdf)

$$f_X(x) = \frac{1}{\sqrt{2\pi}\tau x}e^{-\frac{1}{2\tau^2} (\ln x - \nu)^2}, \quad x > 0,$$

$$= 0, \quad x \leq 0,$$

(1)

where $\tau > 0$ and $\nu$ are real numbers, and $\ln x$ denotes the natural logarithm of $x$.

Consider $n$ independent soil samples taken from a given area, each having a weight of one kilogram, and let $x_1, \ldots, x_n$ denote the measured content of nickel (in milligrams) in each sample. Suppose that these measurement can be regarded as realisations of independent and identically distributed random variables $X_1, \ldots, X_n$ with probability density given by (1).

**a)** Show that if $X$ is a random variable with pdf given by equation (1), then $Y = \ln X$ is normally distributed with expectation $\nu = E[Y]$ and variance $\tau^2 = \text{Var}[Y]$. 
b) Show that $F_X(x) = \text{Prob}(X \leq x) = \Phi\left(\frac{\ln x - \nu}{\tau}\right)$, where $\Phi$ denotes the cumulative probability density of a random variable $Z \sim N(0, 1)$.

Thus, if we introduce $Y_j = \ln X_j$, $j = 1, \ldots, n$, then $Y_1, \ldots, Y_n$ are independent and identically distributed stochastic variables with $Y_j \sim N(\nu, \tau^2)$, $j = 1, \ldots, n$.

c) Assume that $\nu = 1.0$ and $\tau = 0.8$. Find $\text{Prob}(X_1 \leq 1.0)$ and $\text{Prob}(X_1 \cdot X_2 \leq 1.0)$.

d) Assume that $\nu = 1.0$, $\tau = 0.8$ and $n = 5$. Find the probability that the measured content of nickel in at least 4 out of 5 samples is less than 2.72 mg.

e) Show that

$$
\mu = \mathbb{E}[X] = e^{\nu + \tau^2/2}, \quad \sigma^2 = \text{Var}[X] = e^{2\nu}(e^{2\tau^2} - e^{\tau^2})
$$

Hint: Use the substitution $t = (\ln x - \nu)/\tau$.

f) Suppose that we have measured the content of nickel $x_1, \ldots, x_n$ in $n$ soil samples and that we want an estimate of $\mu$ based on these measurements. A possible estimator is

$$
\hat{\mu} = \frac{1}{n} \sum_{j=1}^{n} X_j = \bar{X}.
$$

Briefly discuss the properties of this estimator and compute the corresponding estimate of $\mu$ using the data given below.

g) An alternative estimator of $\mu$ can be constructed by first considering the maximum likelihood estimators (MLE) $\hat{\nu}$ and $\hat{\tau}^2$ of $\nu$ and $\tau^2$, and then, in turn, making use of the result in e). Find the estimators $\hat{\nu}$ and $\hat{\tau}^2$. What are the properties of these estimators? Find a corresponding estimator $\hat{\mu}^*$ of $\mu$ based on $\hat{\nu}$ and $\hat{\tau}^2$. Using $\hat{\mu}^*$, compute an estimate of $\mu$ based on the data given below.

h) Suppose that $\tau^2$ is known and equal to $\tau^2_0$. Based on the random $x_1, \ldots, x_n$ we wish to test

$$
H_0 : \mu \leq \mu_0
$$

against

$$
H_1 : \mu > \mu_0
$$

where $\mu_0$ is a given quantity.
Use the result in e) to express $H_0$ and $H_1$ in terms of $\nu$, and show that the above hypothesis test is equivalent to testing

\[ H'_0 : \nu \leq \nu_0 \]

against

\[ H'_1 : \nu > \nu_0 \]

where $\nu_0$ is a known quantity. Use this to construct a reasonable test of $H_0$ against $H_1$ using $\alpha$ as the level of significance.

i) Derive an expression for the power of the test in h) under the alternative hypothesis $\mu = \gamma \mu_0$ ($\gamma > 1$) when $\alpha = 0.05$ and $\tau_0^2 = 0.36$.

What is the minimum sample size $n$ if the probability of rejecting $H_0$ is 0.9 when $\mu = 1.5 \mu_0$?

j) Suppose that $\tau^2$ is known and equal to $\tau_0^2$. Derive a $100(1 - \alpha)$% confidence interval for $\nu$, and use the result to find a corresponding confidence interval for $\mu$.

Compute the confidence interval for $\mu$ when $n = 10$, $\alpha = 0.05$, $\tau_0^2 = 0.36$ using the data below.

<table>
<thead>
<tr>
<th>$x_j$</th>
<th>57</th>
<th>38</th>
<th>150</th>
<th>29</th>
<th>65</th>
<th>44</th>
<th>36</th>
<th>24</th>
<th>51</th>
<th>131</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_j = \ln x_j$</td>
<td>4.04</td>
<td>3.64</td>
<td>5.01</td>
<td>3.37</td>
<td>4.17</td>
<td>3.78</td>
<td>3.58</td>
<td>3.18</td>
<td>3.93</td>
<td>4.48</td>
</tr>
</tbody>
</table>

Table 1: Observations

\[ \sum_{j=1}^{10} y_j = 39.58 \quad \sum_{j=1}^{10} (y_j - \bar{y})^2 = 3.2360 \]