Show ALL your work, along with JUSTIFICATION for the steps you take.

The (symmetric) normal distribution is without question one of the most important distributions in statistics. On the other hand, asymmetric or skewed distributions, such as the gamma and weibull, also play very important roles in modeling natural phenomenon.

Life testing, or survival analysis, or reliability, is the study of the length of “life” of some type of unit. For example, in medicine one may be interested in how long a person lives once they have been diagnosed with a particular type of cancer. In such a case one would expect lifetime to be nonnegative with some “bunching” of responses in the range one to five years. Of course, there will also be a few cases where the patient out-lives all their grandchildren. This kind of situation demands that the response, lifetime, be modeled using a skewed distribution.

In this exam you will be deriving a “new” skewed distribution (we have not discussed this distribution in class) that may be applicable in the situation mentioned above. You will also be deriving inferential procedures (point estimators, interval estimators, hypothesis tests) that are appropriate for data arising from this distribution. All four parts of the exam are related and somewhat sequential in nature, that is, your response to Part 2 may be useful in completing Part 4, etc.
Part 1 (20 points)
The lognormal distribution is related to the normal distribution in the following way. If the random variable $X$ has a normal distribution with mean $\mu$ and variance $\sigma^2$, then the transformed random variable $Y = e^X$ has a lognormal distribution with parameters $\mu$ and $\sigma^2$. [It is important for you to realize that $E(Y) \neq \mu$ and $V(Y) \neq \sigma^2$. In fact, $E(Y^k) = e^{k\mu + k^2\sigma^2/2}$. You will need this later. **Do not show this!**] Use the information above to show that the p.d.f. of $Y$ is

$$f_Y(y) = \begin{cases} 
\frac{1}{\sqrt{2\pi}\sigma} \frac{e^{-(\ln y - \mu)^2/2\sigma^2}}{y} & 0 < y < \infty \\
0 & \text{otherwise}
\end{cases}$$
Part 2 (30 points)
Suppose we have a random sample $Y_1, Y_2, \ldots, Y_n$ from a lognormal distribution with parameters $\mu$ and $\sigma^2$. We would like to be able to estimate the parameters $\mu$ and $\sigma^2$.

(a) (10 points) Find the method of moments estimators of $\mu$ and $\sigma^2$.

(b) (10 points) Show that the maximum likelihood estimators of $\mu$ and $\sigma^2$ are

$$\hat{\mu}_{mle} = \frac{1}{n} \sum_{i=1}^{n} \ln Y_i \quad \text{and} \quad \hat{\sigma}^2_{mle} = \frac{1}{n} \sum_{i=1}^{n} \left[ \ln Y_i - \frac{1}{n} \sum_{j=1}^{n} \ln Y_j \right]^2$$

(c) (10 points) Show that $\hat{\mu}_{mle}$ and $n \hat{\sigma}^2_{mle}/(n-1)$ are unbiased estimators of $\mu$ and $\sigma^2$, respectively. [Hint: you may want to exploit the relationship $Y = e^X$ (which implies $\ln Y = X$), where $X \sim N(\mu, \sigma^2)$.]
Part 3 (20 points)
As an extension of Part 2, we would like to obtain an interval estimator of $\mu$, keeping in mind that $\sigma^2$ is unknown.

(a) (10 points) Show that

$$
\frac{\hat{\mu}_{mle} - \mu}{\sqrt{\sigma^2_{mle}/(n - 1)}}
$$

has a Student’s t distribution with $n - 1$ degrees of freedom. (See the hint in Part 2.)

(b) (5 points) The random variable given above is pivotal for $\mu$. Why?

(c) (5 points) Use the pivotal random variable above to find a $(1 - \alpha)100\%$ confidence interval for $\mu$. 
Part 4 (35 points)
We would like to determine whether or not the data (once collected) will support or contradict the belief that $\mu = \mu_0$. For this purpose, we will use a hypothesis testing procedure. Let the null and alternative hypotheses be $\mu = \mu_0$ and $\mu \neq \mu_0$, respectively.

(a) (10 points) Comment on whether a Neyman-Pearson Most Powerful (MP) test would be appropriate. How about a Likelihood Ratio Test (LRT)?

(b) (20 points) Irrespective of your comments in (a), show that the resulting LRT has rejection region

$$\frac{\hat{\sigma}_0^2}{\sigma_{mle}^2} > k,$$

where $k$ is a constant and $\hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^{n} (\ln Y_i - \mu_0)^2$.

(c) (5 points) The rejection region above, after simplification, becomes

$$\frac{|\hat{\mu}_{mle} - \mu_0|}{\sqrt{\sigma_{mle}^2/(n - 1)}} > c,$$

where $c$ is a constant. [BONUS (5 points): Do this simplification. This is bonus, so no partial credit will be given.]

Find $c$ to make this an $\alpha$ level test.