Statistics 533 Final Examination Name _____
Spring 2000 **You must show all of your work**

When asked to explain something, provide an explanation that could be understood by someone who does not have formal training in statistical methods. Your explanations should be clear, but concise. Each question or part of a question is worth 5 points for a total of 100 points.

1. Consider the failure-time distribution with cdf

$$F(t) = 1 - \exp\left[-\left(\frac{t-\mu}{\alpha}\right)^3\right], \quad t > \mu. \tag{1}$$

(a) Derive an expression for the hazard function corresponding to the cdf in equation (1). How would you describe the shape of this hazard function? Be as specific as possible.

(b) Write down an expression for the log likelihood for a sample of n observations that are reported as exact failures (i.e., use the density approximation), based on the model expressed in equation (1).

(c) Derive a simple expression for the ML estimator of α in equation (1), assuming that the value of μ is given.

2. The Arrhenius-lognormal accelerated test model is

$$\Pr[T \leq t; \texttt{temp}] = \Phi_{\text{nor}}\left[\frac{\log(t) - \mu}{\sigma}\right]$$

where $\mu = \beta_0 + \beta_1 x$, $x = 11605/(^\circ\text{C} + 273.15)$, and σ is constant.

(a) What is the interpretation of β_1 in this model?

(b) Derive an expression for t_p, the p quantile of the lognormal distribution as a function of temperature in degrees C.

(c) Engineers often think in terms of acceleration factors. Derive an expression for the acceleration factor that one would have for testing at $\texttt{temp}_1$ versus $\texttt{temp}_2$ under the Arrhenius-lognormal accelerated test model described above.

3. Describe some of the potential pitfalls (dangers) of using accelerated testing to make inferences about product reliability.

4. Bayesian methods are becoming increasingly popular in many areas of application of statistics. Part of the reason for this is that modern computing capabilities have made it possible to solve real problems that could not be solved in the past. One difficult issue is how to obtain useful prior information.

   (a) What is the most important advantage or potential advantage for the use of Bayesian methods in reliability data analysis and decision making?

   (b) Briefly describe appropriate and inappropriate sources of prior "informative" information.

   (c) Some analysts advocate the use of Bayesian methods with "noninformative" or "diffuse" prior distributions when informative prior information is not available. Discuss some of the potential difficulties associated with the use of such distributions.

5. Suppose that a reliability engineer believes that the failure-time distribution of a particular kind of switch is lognormal. She also has available prior information on the lognormal parameters as well as Type I censored life test data from a sample of 100 switches. A total of 50 switches failed in the life test. A customer for one of these switches has asked for a prediction interval that will, with probability .95, contain the life time of the particular switch that it will purchase.

   (a) Given a Monte Carlo sample of 2000 pairs of $\mu^\star, \sigma^\star$ values from the posterior pdf $f(\mu, \sigma|\text{DATA})$, list the steps needed to obtain a Bayesian prediction interval for the time of the failure.

   (b) Is there any extrapolation involved in making the above prediction? *Explain* and describe the possible difficulties involved with the extrapolation(s).

6. A company that owns a fleet of 130 automobiles keeps track of all repairs and component replacements. Part of the data base provides the life times of components (e.g., light bulbs, hoses, and belts) that are replaced upon failure or through preventive maintenance in an attempt to reduce unexpected failures. For the purposes of this question, let us consider the bulbs used in head lights. These bulbs, at present, are not replaced until there is a failure.

   (a) Suggest how you could look at available data to see if the failure times of light bulbs within a car over time are identically and independently distributed (so that a renewal process would provide an appropriate model for the data).

   (b) If light bulbs failure times are identically and independently distributed, what shape would you expect for the hazard function of the light bulbs? Draw a picture and explain.

   (c) It has been suggested that after the first headlight in an automobile fails, both of the headlights should be replaced. Assume that the failure record makes known whether the replaced bulb is on the left or on the right of the car. Explain how you would analyze the available data to see if this is a good idea or not. Suggest more than one way of looking at the data, if appropriate.

   (d) Explain why fitting a NHPP data to the light bulb replacement recurrence data could be useful.

7. A system contains four components. The failure items for the components are independently distributed with cdf's F_1, F_2, F_3, and F_4. The system fails after the second of these four units fails.

   (a) Draw a system-structure diagram for this system.

   (b) Find an expression for the failure time distribution for the system as a function of F_1, F_2, F_3, and F_4.

8. An life test experiment has been conducted on a random sample of devices that have a failure-time distribution that can be approximated with an exponential distribution. An estimate the exponential mean θ has been reported along with its estimated standard error $\widehat{\text{se}}_{\widehat{\theta}}$. The exponential hazard is $\lambda = 1/\theta$.

   (a) Use the delta method to derive an expression for $\widehat{\text{se}}_{\widehat{\lambda}}$ as a function of $\widehat{\text{se}}_{\widehat{\theta}}$ where $\widehat{\lambda} = 1/\widehat{\theta}$.

   (b) Provide an expression for a normal approximation confidence interval for λ, the lower endpoint of which will always be positive.