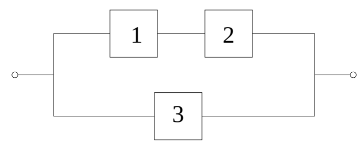
Statistics 533 Final Examination Name ______________________

Spring 1998 **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 following system diagram. Assuming that the three components fail independently, derive an expression for the cdf of the system as a function of F_1, F_2, and F_3, the cdfs of the individual components.



2. 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. In the area of reliability, however, Bayesian methods are often viewed with some skepticism.

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

   (b) What is the most important danger involved with the use of Bayesian methods in reliability data analysis and decision making?

   (c) Suppose that analysts have available prior information as well as Type I censored data from a sample of 100 switches. Given a Monte Carlo sample of 2000 pairs of $\mu^\star, \sigma^\star$ values from the posterior pdf $f(\mu, \sigma|\text{DATA})$, explain how to approximate the marginal posterior distribution for $t_{.1}$, the .1 quantile of the life distribution of the switches and how to find a 95% Bayesian confidence (or credible) interval for $t_{.1}$.

3. 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) How would you interpret the mean cumulative function (MCF) for light bulbs failures estimated from information collected on this fleet of automobiles?

   (b) How would you interpret the pointwise confidence intervals for the MCF if they were computed from information collected on this fleet of automobiles? Hint: You have to specify the population or process being described by the confidence interval.

   (c) Suppose that the failure times of light bulbs within a car over time are identically and independently distributed. Then a renewal process would provide an appropriate model and it would make sense to try to estimate the lifetime distribution of the bulbs. Give some physical reasons why this assumption might *not* be valid.

   (d) 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.

   (e) If failure times of any particular component in the automobiles are identically and independently distributed, how could the shape of the component's hazard function be used to indicate whether the component in an automobile should be replaced before it fails or not? Draw a picture and explain.

4. Refer to Table C.20 (copy attached). A life test is planned initially to run for 1000 hours and the responsible engineers feel that life can be described by a lognormal distribution with $\sigma = .5$ and with approximately 2.28% of the units failing at the end of the 1000 hour test.

(a) Management has asked for an estimate and a 95% confidence interval for $t_{.5}$ for this distribution. In planning the test, it is thought that the sample size should be large enough such that the confidence interval upper endpoint should be approximately 40% larger than the ML estimate of $t_{.5}$. What is the approximate sample size that is needed to estimate $t_{.5}$ for this distribution, according to the specified criterion, assuming that both μ and σ will be estimated using the method of maximum likelihood?

(b) Figure 10.6 (copy attached) can be used to obtain the variance factor used in the equation to find the approximate sample size needed to estimate a particular lognormal quantile. Use this figure to find the sample size needed to estimate $t_{.1}$ such that a 95% confidence interval upper endpoint should be approximately 40% larger than the ML estimate of $t_{.1}$.

(c) Give an equation to show how the variance factor from Figure 10.6, used in part (b), could be computed from the information in Table C.20.

(d) Note that some of the lines at the bottom of Figure 10.6 tend to bunch together. What is the intuitive explanation for this behavior of these lines? What general implication does this have for test planning?

(e) Suppose that there is some flexibility in choosing the length of the test (i.e., it might be possible to test longer than 1000 hours, if needed) that is to be used to estimate $t_{.1}$. Use Figure 10.6 to make a qualitative argument as to whether this is a useful way to spend additional testing resources. How much could the needed sample size be reduced by making the test longer?

5. The inverse power relationship-lognormal model is

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

where $\mu = \beta_0 + \beta_1 x$, $x = \log(\texttt{volt})$, $\texttt{volt}$ is voltage stress that could be measured, for example, in kV/mm, and σ is constant.

   (a) Derive an expression for the quantile of the lognormal distribution as a function of voltage stress.

   (b) Engineers often think in terms of acceleration factors. Derive an expression for the acceleration factor that one would have for testing at $\texttt{volt}_1$ versus $\texttt{volt}_2$.

6. Practical application of accelerated life testing methods requires some knowledge of the underlying physics/chemistry of failure. Explain why.

7. In general, planning values are needed to do test planning and to determine the sample size needed to provide a specified degree of precision.

   (a) Explain why such planning values are needed.

   (b) Product or reliability engineers may be able to provide some useful information, but they cannot be expected to provide anything like exact parameter values (otherwise they would have no reason to run the test!). What can be done to protect against the use of potentially misspecified planning values?

8. Explain how the hazard function $h(t)$ for a continuous random variable (e.g., from a Weibull distribution) is related to a probability.