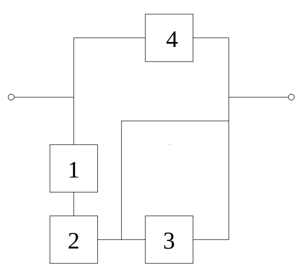
Statistics 533 | Final Examination | Name ____________________________
Spring 2002 | | **You must show all of your work**

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. A failure-time distribution for the random variable T has a hazard function $h(t) = \theta t, t > 0$. Derive an expression for the corresponding cdf for T.

2. Suppose that you have an expression $\text{Var}[\log(\widehat{\theta})]$, where $\widehat{\theta}$ is the ML estimator of the scalar parameter θ. Use the delta method to show how to get an expression for the approximate (large-sample) variance of $\widehat{\theta}$.

3. Consider the following system diagram. Time to failure for block (component) i has a cdf $F_i = F_i(t)$ and the times to failure of the individual blocks are independent. Provide an expression for the cdf of the system as a function of F_1, F_2, F_3, and F_4.



4. The Arrhenius-lognormal accelerated life test model is

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

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

(a) What is the purpose of using the constant 11605 in this model?

(b) Engineers who conduct accelerated tests often think in terms of acceleration factors. Starting with the Arrhenius-lognormal accelerated test model described above, derive an expression for the acceleration factor that one would have for testing at $\texttt{temp}_1$ versus a use condition of $\texttt{temp}_2$.

(c) Explain why the Arrhenius rate-reaction model implies that σ should be constant in the Arrhenius accelerated life test model.

(d) An accelerated life test was conducted at 60, 80, and 100°C, with at least some failure at each level of temperature. List the steps that you would use (including the model or models to be fit) in order do an approximate statistical test for the null hypothesis $\text{H}_0 : \sigma_{60} = \sigma_{80} = \sigma_{100}$, without making any assumption about the relationship between μ and temperaure..

-
-
-

5. In planning accelerated life tests with simultaneous testing and Type I (time) censoring, optimal design results suggest to test more units at lower levels of the accelerating variable. What is the intuitive explanation for this?

6. Use a simple graph to help you explain how and why having one failure mode mask another can lead the users of accelerated life tests to make incorrect conclusions.

7. 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.

8. A life test of $n = 100$ light bulbs was conducted. All light bulbs were run until failure and the number of hours (rounded to the nearest hour) was recorded. The smallest failure time was 768 hours and the largest was 1233. The likelihood might be written in either of the following two ways.

$$L_{\text{Corr}}(\theta) = \prod_{i=1}^{n} \int_{t_i-.5}^{t_i+.5} f(t)\,dt \tag{1}$$

$$L_{\text{DA}}(\theta) = \prod_{i=1}^{n} f(t_i) \tag{2}$$

We refer to (**??**) as the "correct likelihood." The expression in (**??**) is then called the "density approximation."

(a) Explain the meanings behind the names "correct likelihood" and "density approximation."

(b) Relative to the "density approximation," what are the potential advantages of using the "correct likelihood?"

(c) Relative to the "correct likelihood," what are the potential advantages of using the "density approximation?"

9. 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) Briefly describe appropriate and inappropriate sources of information to provide an "informative" prior distribution.

   (b) 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.

10. The failure-time distribution of a particular kind of switch is lognormal. Prior information is available on the lognormal parameters. A Type I censored life test was conducted on a sample of 100 switches. A total of 50 switches failed in the life test. Then a simulation was conducted to generate a Monte Carlo sample of 2000 pairs of $\mu^{\star}, \sigma^{\star}$ values from the posterior pdf $f(\mu, \sigma|\text{DATA})$.

    (a) 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. List the steps needed to obtain a Bayesian prediction interval for the time of the failure for a single switch.

    (b) Another customer is interested, instead, in estimating the time at which $100p\%$ of a population of such switches would fail. List the steps needed to obtain a Bayesian credibility interval for the p quantile of the life distribution. Also explain why this customer might have been asking for a different kind of interval.

11. 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) If light bulb failure times are identically and independently distributed, would you expect the failure-time distribution to be an exponential distribution? Explain why or why not.

    (b) 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.

12. A system consists of two identical components in series. Both components have a Weibull failure-time distribution with scale parameter η and shape parameter β and the failure times for the two components are independent. The failure time distribution of the system also has a Weibull distribution. As functions of η and β, what are the parameters of this distribution?