University of California, Davis
Department of Public Health Sciences

Fall 2023	**Survival Analysis**	**BST 222**
David M. Rocke		**September 28, 2023**

Homework Assignment 1
Due October 10, 2023

1. The log-logistic distribution has survival function

$$S(t) \quad = \quad \frac{1}{1+\lambda t^{\alpha}}$$

   where $\lambda, \alpha > 0$ and $t \geq 0$. Here, the distribution of the logarithm of the failure times is logistic with mean μ and scale parameter σ, and $\alpha = 1/\sigma$, $\lambda = \exp(-\mu/\sigma)$

   (a) Find the hazard rate for this distribution. For what values of α is the hazard function monotone. When it is monotone, is it increasing or decreasing?

   (b) When the hazard function is not monotone, where does it change slope? Is it decreasing to increasing or increasing to decreasing?

   (c) Describe a practical situation where the hazard function might change slopes in this manner.

2. Do Exercise 2.18 from Klein and Moeschberger.

3. Using the data from Exercise 3.6 in Klein and Moeschberger,

   (a) Suppose that time from the start of the study to relapse is assumed to be exponential with parameter λ_1 and time from start of the study to death is assumed to be exponential with parameter λ_2. Find the likelihoods for λ_1 and λ_2 and find the MLE for each of these parameters.

   (b) Suppose that we could not observe the death times of patients who had not relapsed. Find the estimate of λ_2 from the truncated sample and discuss the difference.

4. Suppose that the failure distribution of a particular type of disk drive is exponential with a mean time between failures (MTBF) of 4 years (so that $\lambda = 0.25$). In this model we suppose that after each failure, the time to the next failure is still exponential with the same parameter λ. This means that the failures form a Poisson process with the mean number of failures per year of $\lambda = 0.25$. If we have n disks on test, then the failure process is still Poisson with MTBF $1/(n\lambda)$ and mean number of failures per year of $n\lambda$.

   Suppose that we put n disks on test and need to have a 95% chance of at least 30 failures in the first year to have sufficient accuracy in estimating λ. How many disks need to be put on test? We can attack this either using the Poisson distribution (`qpoiss`, `ppoiss` in R) or by using the gamma distribution (`qgamma`, `pgamma`). For the Poisson calculation, you need to compute the mean number of failures in the first year as a function of n, and then you can use `qpoiss` with `lower=F` and trial and error in choosing n until you find the minimum value of n that results in 30 as the point with upper tail equal to 0.95. For the gamma calculation, the waiting time to the kth failure with exponential waiting times with parameter λ_0 has rate λ_0 and shape k. You can use `pgamma` to find the least value of n that results in a time to 30 failures of one year that has a probability of 0.95 or greater. Do this calculation both ways until you get the same answer from both methods.