STA 312s19 Assignment Five[1]

The paper and pencil part of this assignment is not to be handed in. It is practice for Quiz 5 on February 11th. The R part may be handed in as part of the quiz. **Bring hard copy of your printout to the quiz**. Do not write anything on your printout in advance except possibly your name and student number.

1. Consider a random sample of right censored data from an exponential distribution with parameter $\lambda > 0$. In addition to the data $T_1, \ldots, T_n$, we have indicators $\delta_1, \ldots, \delta_n$, where $\delta_j = 1$ if T_j is a failure time, and $\delta_j = 0$ if T_j is a censoring time.

   (a) Derive the maximum likelihood estimate of λ. Show your work. Include the second derivative test. Circle your final answer. It is on the lecture slides. What happens if *all* the observations are censored?

   (b) Give a formula for the estimated asymptotic variance of $\widehat{\lambda}$. Show your work. Circle your answer.

   (c) Give a 95% confidence interval for λ. Your answer is two formulas, one for the lower confidence limit and one for the upper confidence limit.

   (d) Give a formula for the MLE of the survival function $S(t)$. Your answer is a formula that could be computed from the sample data for any given t.

   (e) Using the one-variable delta method (see formula sheet), derive a 95% confidence interval for the survival function $S(t)$. As above, your answer is two formulas, one for the lower confidence limit and one for the upper confidence limit.

   (f) Now do the same thing a different way. Writing your confidence interval for λ as $0.95 \approx P(A < \lambda < B)$, multiply by $-t$ and take the exponential function to obtain a different confidence interval for $S(t)$, one that is also valid. Is it possible for these confidence limits to be outside the range zero to one?

2. The file `http://www.utstat.toronto.edu/~brunner/data/legal/expo.data2.txt` contains a random sample from an exponential distribution with right censoring. Using R as a high-powered calculator (no numerical search yet),

   (a) Give the maximum likelihood estimate of λ. The answer is a number. You are calculating your answer to Question 1a.

   (b) What is the estimated asymptotic variance of $\widehat{\lambda}$? The answer is a number. You are calculating your answer to Question 1b.

(c) Give a 95% confidence interval for λ. Your answer is two numbers, the lower confidence limit and the upper confidence limit. You are calculating your answer to Question 1c.

(d) Based on your answers to Questions 1d and 1e, calculate your estimated $S(t)$ and the confidence interval for the following values of t: `seq(from=0,to=3,by=0.1)`. Bind t, $\widehat{S}(t)$ the lower confidence limit and the upper confidence limit into a matrix with 4 columns.

(e) Now do the same thing, using your confidence limits from Question 1f. The result is another 4×31 matrix. Which one do you like more and why?

3. Using the data from Question 2,

(a) Obtain the maximum likelihood estimate of λ by a numerical search. Compare your answer to Question 2a.

(b) Obtain the estimated variance of $\widehat{\lambda}$ from the output of your search. If you think about it, you will realize that $-\ell''(\hat{\lambda})$ is a number on your printout. Display the estimated variance on your printout. Compare your answer to Question 2b.

After that, what you'd do would be the same as in Question 2. You don't have to do it twice.

Bring your printout for Questions 2 and 3 to the quiz. All the requested numbers and the code that produced them should appear on your printout. Do not write anything on your printout in advance except possibly your name and student number.