SURVIVAL ANALYSIS - COMPUTER ASSIGNMENT I

DESCRIPTION

This computer assignment may be solved using any program language and/or software package that you find suitable. If you do not have any preference or do not know which to choose, the recommendation is to use R.

It is ok to do the assignment in groups, but the group size must not be greater than three.

Part 1.

- Generate n Weibull distributed random numbers from the Weibull distribution given by the density function

$$f(t; a, b) = \frac{a}{b}(\frac{t}{b})^{a-1} \exp\{-(\frac{t}{b})^a\}, \ t, a, b \geq 0,$$

   when $a = 4.5$ and $b = 22.5$, for $n = 10, 100, 200, 500$ and 1 000. Denote the simulated outcomes by $T_i, i = 1, \ldots, n$, corresponding to life times. The chosen parameter values shall correspond to that the mean and standard deviation of T is approximately 20.5 and 5.2 respectively.
- Calculate the corresponding Nelson-Aalen estimates together with 95% confidence interval of choice.
- Use that a Weibull distribution given by the density function above corresponds to a counting process with intensity function

$$\alpha(t; a, b) = \frac{a}{b}(\frac{t}{b})^{a-1}, \ t, a, b \geq 0,$$

   and compare with your estimates.

Part 2.

- Generate n uniform random numbers which have support on $[20, 30]$ for $n = 10, 100, 200, 500$ and 1 000 and denote these by $C_i, i = 1, \ldots, n$, corresponding to (independent) censoring times.
- Use that $T_i \leq C_i$ corresponds to a complete observation and that $T_i > C_i$ corresponds to a censored time.
- Repeat Part 1 given censoring.

Part 3. Repeat Part 1-2, but for the Kaplan-Meier estimator.

QUESTIONS TO ANSWER

- Argue for the performance of the estimators relative to the true theoretical model used, is any of the two preferable?
- What is the effect of censoring?
- What is the effect of n?

Report

For those of you who can not attend or have not finished the assignment during computer class you need to write a report:

- The report must be submitted as a pdf.
- Maximum report length is 3 pages.
- Base your argumentation based on a number of representative figures that *must be possible to read.*
- No code.