STAT 796: Homework 10

Due Friday, April 19 at 11:59pm on Canvas. Please append your code at the end your assignment.

1. Figure 1 shows the time until failure in a sample of 70 diesel engine fans. Using this figure, answer the following questions:
   a. What is the approximate estimated probability of the fan lasting at least 6,000 hours?
   b. Approximately at what time did the second fan failure occur?
   c. Approximately at what time did the first censoring occur?

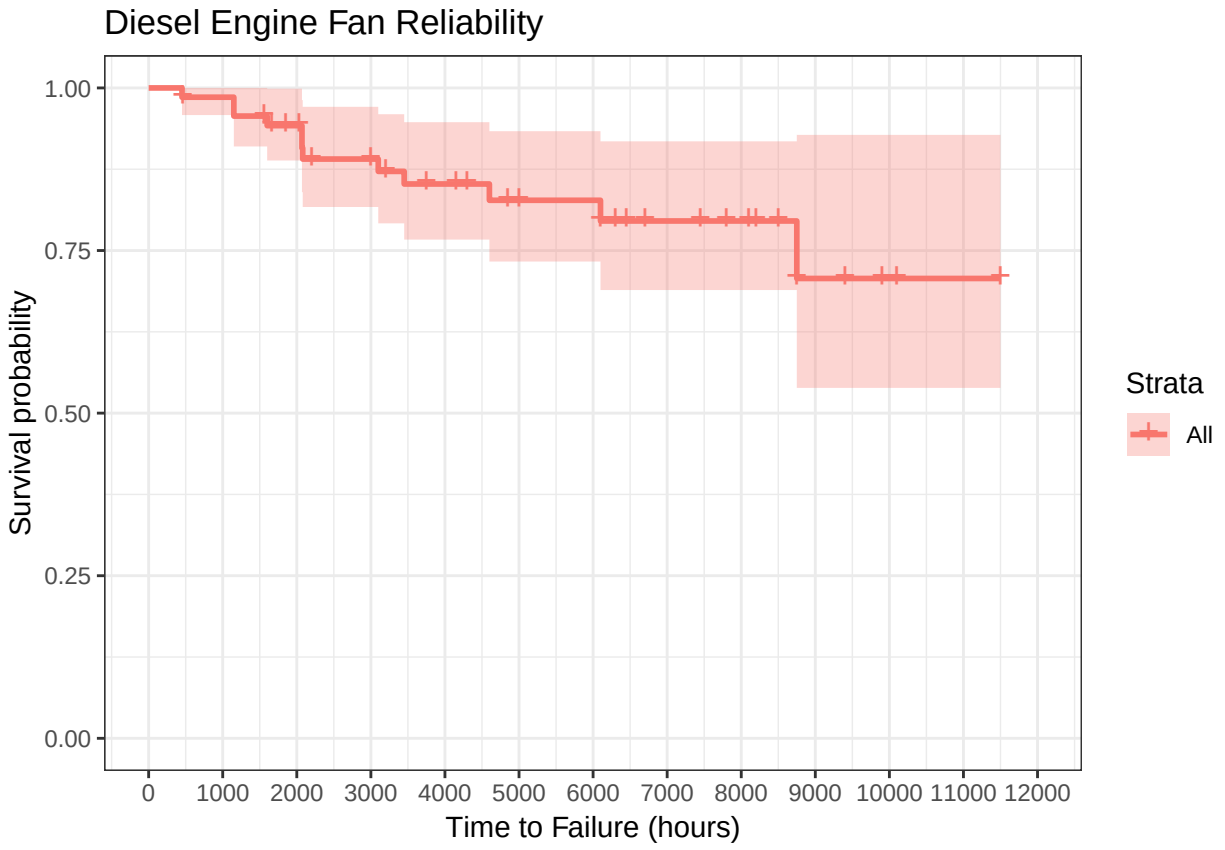

Figure 1: Failure time for diesel engine fans.

The following questions are about survival time in the larger dataset from the Worcester Heart Attack Study (WHAS). These data are described in HLM and are available on Canvas in the file `whas500.csv` (note that this is different from the `whas100.csv` dataset!). The variable `died` provides an indicator of an event. Follow-up time is in the variable `l_of_fol`.

2. For this question, consider the whole cohort together.
   a. Plot the estimated survival (i.e. K-M curves) over time.
   b. What is the median survival time after a heart attack?
   c. What is the estimated survival at 1 year (365 days)? Provide a point estimate and confidence interval.

3. For this question, consider the cohort stratified by whether or not this is the first heart attack (`mi_recurr=0`) or a recurrence (`mi_recurr=1`).
   a. Plot the estimated survival (i.e. K-M curves) over time for each group in the same figure
   b. What is the median survival time after a heart attack for each group?

c. What is the estimated survival at 1 year (365 days) for each group? Provide point estimates and confidence intervals.
d. Is there a significant difference in the survival curves comparing the two groups?