Finger Exercises — Due Thursday, April 9, 1998

Please prepare answers for these exercises, to be turned in. Your grade will be based on how many questions you made a reasonably strong effort to answer, **not** *how many questions you get right or partially right. There is no partial credit, but the grader may give you additional written feedback on each question you attempt.*

Feel free to discuss these exercises with each other and with me. You will get the most benefit from the grader's remarks, however, if the work you turn in is your own.

This hw, like most material for this course, is posted on the World Wide Web at URL `http://www.stat.cmu.edu/~brian/402/`; you can cut and paste data directly out of the TEX file if you wish. Look for something under week11. The files you need are: `ratcancer.dat` and `motorettes.dat`.

There are two problems. Both problems have four parts.

1. Please refer to the rats cancer data at the beginning of the notes from Tue Mar 31, 1998. The data set is also in `ratcancer.dat`.

   (a) Fit and plot Kaplan-Meier PL estimates of the survival curves for the rats in the two groups. Indicate in your plot which survival curve goes with which group.

   (b) Is there evidence that these two groups have different survival curves? Use the Mantel-Haenszel test (a.k.a. the log-rank test) as well as a modified Wilcoxon test, to test this hypothesis. (see documentation for `survdiff` if you forget how to do this).

   (c) Now refit the data using a Cox proportional hazards model, and plot the resulting survival curves, indicating which curve goes with whih group. Qualitatively compare the Cox PH curves with the Kaplan-Meier PL curves from part (a).

   (d) Use an appropriate test based on the Cox model to decide whether the two survival curves are the same. What test(s) did you use (i.e. give the statistical name(s) of the test(s))? Compare your answer with part (b) above.

2. Please refer to the motorettes testing data from the beginning of the notes from Tue Mar 31, 1998. The data set is also on line in `motorettes.dat`. Let us suppose that a normal operating temperature for these motorettes is 25^oC.

   (a) Fit the Cox proportional hazards model to this data, regressing survival time on temperature. Should temp be kept in the model or can it be dropped? Justify.

   Prepare and examine appropriate residual plots; does the model look like a reasonable fit?

   (b) Plot (on the same plot) the survival curves for temp=25, 150, 170, 190 and 220^oC. What conclusions can you draw about survival at the normal operating temperature, from this or other ways of looking at the Cox model?

   (c) Use `survreg` to fit a Weibull model to these data (note: you can indicate censored observations just as you would with `survfit` or `coxph`).

   Prepare and examine appropriate residual plots; does the model look more or less acceptable than the Cox model?

(d) Find a way to summarize (probably graphically but maybe not) the survival function under the Weibull model at 25^oC.

In particular, find a value of t such that $P[T > T] \approx 95\%$, where T is the time until failure for a motorette under the Weibull model.