Primary Biliary Cirrhosis study (Due Date: Sept 08, 2020)

Fleming and Harrington (Counting Processes and Survival Analysis, Wiley, 1991) contain the data from the Mayo Clinic trial in primary biliary cirrhosis (PBC) of the liver conducted between 1974 and 1984. A description of the clinical background for the trial and the discussion can be found in Dickson, et al., Hepatology 10:1-7 (1989) and in Markus, et al., N Eng. J of Med 320:1709-13 (1989). A total of 424 PBC patients, referred to Mayo Clinic during that ten-year interval, met eligibility criteria for the randomized placebo controlled trial of the drug D-penicillamine.

Missing data items are denoted by "." The status is coded as 0=censored, 1=censored due to liver transplant, 2=death. The outcome of interest is time to death. For this analysis please consider status=1 is also as censored (combine 0, 1). Dataset is available at Canvas.

Variables:

case number
number of days between registration and the earlier of death, transplantation, or study analysis time in July, 1986
status
drug: 1= D-penicillamine, 2=placebo
age in days
sex: 0=male, 1=female
presence of ascites: 0=no 1=yes
presence of hepatomegaly 0=no 1=yes
presence of spiders 0=no 1=yes
presence of edema 0=no edema and no diuretic therapy for edema;
.5 = edema present without diuretics, or edema resolved by diuretics;
1 = edema despite diuretic therapy
serum bilirubin in mg/dl
serum cholesterol in mg/dl
albumin in gm/dl
urine copper in ug/day
alkaline phosphatase in U/liter
SGOT in U/ml
triglicerides in mg/dl
platelets per cubic ml / 1000
prothrombin time in seconds
histologic stage of disease

Suppose that an investigator (Dr. Blum) asks you to carry on an analysis of this data set. Dr. Blum is interested in survival time of these patients and would like to know the

impact of treatment, age and serum bilirubin as a categorical variable (<1.1, 1.1-3.3 and >3.3) on the survival.

She instructs you to do the following:

1. Dr. Blum is very familiar with linear model, sofirst, she asks you to perform a linear model after excluding all censored observations to address her question.
2. Second, she asks you to treat all censored times as death times and redo the analysis of (1).
3. Third, she asks you to perform a logistic regression analysis by defining a new outcome as dead=1 and 0, otherwise, to address her question.

In order to answer Dr. Blum's each request, please perform univariate analysis and a multivariate analysis (including all three covariates). Please interpret the effects of estimates of treatment, age and serum bilirubin regardless of their significance (In a real data analysis, you may ignore the effects of non-significant coefficients in some situations).

As a statistician you would like to comment on the appropriateness of the data analyses suggested by Dr. Blum. Provide a meaningful, convincing critique about her analysis. You can take other examples to convince her as well.

Based on the material that you have learned so far, I suggest you carry on a survival analysis assuming some parametric distributions. See the document describing SAS and R codes for carrying on parametric analyses. Report your results again for univariate and multivariate analysis. Interpret and comment on your results compared to the previous approaches of (1)-(3).

Furthermore, Dr. Blum is interested in technical details of your parametric survival method given that she has a strong background in mathematics. She is asking you to mathematically explain how you obtained the estimates and standard deviations. For example, what is the procedure? In other words, what does exactly SAS or R does in this case. You do not need to provide all details but enough so that she can understand the method. You can limit your answer only to Weibull case. Attach this technical section as an appendix to the report.

You can submit one report from the group. However, each ember in the group should submit a self-evaluation separately about the project, This should address your level of participation, and what you have learned by doing this project. Also please add **constructive comments** on how to improve your working environment within this group. Your comments are kept in confidential.