PubH 7450 Homework 1 (Spring 2023) Due date: Feb 16, 2023

Note: Please include your computer programs and only relevant output; show major steps in your derivations/calculations.

1) (6 Points) The city of Saint Paul conducts a cross-sectional study on the time to retirement in Saint Paul, Minnesota. The investigators of the study take a random sample of 10,000 residents living in Saint Paul on Feb 1, 2023. Participants were asked ``How old were you when you retired?'' and ``At which age did you first enter the job market?''.

(i) When does an observation start to be at risk for the event?

(ii) Define the variable X of time to retirement, i.e. how would you measure/code it?

(iii) Provide one example of right censoring that can occur.

2) (6 Points) Provide three examples (a study, an event, a time-to-event, reason for censoring) for left censoring.

3) (6 Points) Provide three examples (a study, an event, a time-to-event, reason for censoring) for interval censoring.

4) (4 Points) Does the parametric model h(x) = a exp(-bx) for x>=0, with a,b>0, define a valid hazard function for a survival model? If, yes, please justify, if no, why?

5) (4 Points) Compute the 75-percentile, median and mean for an exponential survival model with survival function S(x)=exp(-x/12.5), x>=0.

6) (4 Points) Compute the 25-percentile and median for an Gompertz model with hazard rate h(x)=0.001 * exp(x/100).

7) (10 Points) a) Fit exponential survival models to the experimental arm(6-MP) and control arm (Placbo) in the (leukemia.dat). b) Compute a (Wald-type) 95% confidence interval for the exponential rate of the two groups, and report these intervals. c) Propose and implement a two-sided tests for the null hypothesis "*H0: the exponential rates of the 6-MP and Placbo groups are identical*" vs the alternative "*HA: the exponential rates of the 6-MP and Placbo groups differ*". d) Propose and implement a one-sided tests for the null hypothesis H0 (as in d) vs the alternative "*HA: the exponential rates of the 6-MP is small than the one for the Placbo group*".

8) (3 Points) Create an account to access the data-sharing platform projectdatasphere.org (PDS) at https://data.projectdatasphere.org/projectdatasphere/html/registration.

Note: for the fields `Reason for visiting` and `Research Description & Goals' use `Access Data for Trial Research' and `Survival and Meta-Analysis`.

9) (6 Points) Create a study-group with up to 5 students. Select a randomized clinical trial (RCT) hosted on PDS that is of interest for the group. To determine if an RCT is of interest to you by looking at the description of the study provided by PDS. Download the data, data dictionary, study protocol, and any PubMed publication linked with the data.

Notes:

(a) Make sure the study you select is (i) an RCT with (ii) data on both the experimental and control arm (iii) available for download from PDS (not via data-request at NIC) and (iv) the primary study analyses did use a time-to-even outcome (i.e., overall survival, progression free survival, time to relapse, time on treatment, etc).

(b) In subsequent assignments/work you will implement the per-protocol time-to-event analysis as planned in the study protocol of the RCT. Hence, make sure the RCT/data (outcome, treatment group, pre-treatment variables, etc.) you select are well annotated.