PH C240B: Assignment 3

Rachael Phillips

11/9/2017

Problem 1

Show the CAR condition, $x \to Pr(O = o|X = x)$ for $x \in \mathcal{C}(o)$ is constant implies $Pr(X = x|O = o) = Pr(X = x|x \in \mathcal{C}(o))$. You may assume all random variables here are discrete for simplicity.

Problem 2

Let $P_{X,\epsilon}$ be a path through P_X, the distribution of the full data, X, and having score $S_1(X)$. This then defines a path $P_{P_{X,\epsilon},G}$ through the observed data distribution, $P_{P_X G}$. Show that the scores generated by these paths are $E[S_1(X)|O = o]$.

Problem 3

Let G_ϵ be a path through G, the distribution of the censoring time, C, given X, having score $S_2(C, X)$. This then defines a path $P_{P_X G_\epsilon}$ through the observed data distribution, $P_{P_X G}$. Show that the scores generated by these paths are $E[S_2(C, X)|O = o]$.

Problem 4

This problem involves simulating data under a general Cox model. Let's make the assumption we have a conditional hazard of death at time, t, given by $\lambda(t|X) = \lambda_0(t)exp(f_\beta(X))$ where X is a set of covariates and f_β is a function indexed by β, say finite dimensional. Assume the baseline hazard is $\lambda_0(t) = exp(rt)$ for positive r. Given X, what is the distribution of death times? Prove your answer.

Problem 5

Complete the first problem from LabCox in the lab section of the files on bCourses.

Bonus

Assume a CAR model for full data consisting of survival time, censoring time, the continuous baseline covariates and randomly assigned treatment indicator. We have observed data $min(T, C), \Delta$ along with the covariates and treatment indicator. Someome receives a data set of 1000 independent subjects drawn from this model from an RCT and runs a Cox Proportional hazards regression with treatment as the only covariate, showing a significantly negative coefficient. Can you convince this person he may be wrong via simulation? Explain how you set up your simulation and turn in your code to show the results.