STAT 796: Homework 11

Due Tuesday, May 7 at 11:59pm on Canvas. Please append your code at the end your assignment.

This assignment considers data from the University of Massachusetts AIDS Research Unit (UMARU) IMPACT Study (UIS). UIS was a study of residential treatment programs for drug abuse. The primary study question was to ascertain the relative effectiveness of the different programs. For this assignment, you are asked to consider that question and related questions based on the study data.

The data are available in the file `uis.csv` on Canvas and contain the following variables:

- `id` – Participant ID
- `age` – Age at enrollment in years
- `beck_score` – Beck Depression score at enrollment
- `hercoc_use` – Drug use in three months prior to enrollment. Takes values `Cocaine`, `Heroin`, `Heroin & Cocaine`, `Neither`
- `iv_use` – IV Drug use history at admission (`Recent`, `Previous`, `Neither`)
- `nprior_tx` – Number of prior drug treatments
- `white` – Indicator of white race
- `tx_long` – Indicator of long treatment (compared to short treatment)
- `site` – Treatment site (`A`, `B`)
- `time` – Time of Return to drug use or censoring
- `status` – Indicator of return to drug use (1 = yes, 0 = censored)

1. Assume treatment type (long or short) was randomly assigned. Answer the question of interest for the study by doing the following:
    a. Provide a plot of estimated survival curves for time until drug use comparing long treatment to short treatment.
    b. Estimate the hazard ratio for time until drug use, comparing long treatment to short treatment. Report the estimate in a summarizing sentence.
2. Is there a treatment effect when adjusting for age and prior treatment status? Answer this using a proportional hazards regression model with adjustment for age and whether or not an individual has more than one episode of prior treatment.
3. Is there a difference in time until return to drug use by an individual's history of IV drug use? Answer this question by doing the following:
    a. Provide a plot of estimated survival curves that correspond to this question.
    b. Fit a proportional hazards regression model with IV usage history as the predictors of interest and that adjusts for age. Report the estimated hazard ratios for IV usage categories in 1 or 2 summarizing sentences.
4. Is there a difference in time until return to drug use by an individual's use of heroin or cocaine the 3 months prior to admission? Answer by providing and interpreting estimated hazard ratios (you may choose what to adjust for).