Homework 1

BIOST 544

Please submit your homework on Canvas, in a compiled R-markdown file (to pdf or html).

Your writeup should be clearly written, and should integrate images, analysis and text. Think of this as a writeup that you are sending to a collaborator who provided the data.

Analysis Question

There is a belief that the effectiveness of the anti-angiogenesis agent TFD725 (evaluated in the nsclc dataset) may be different for older vs younger patients. We will use the nsclc dataset to attempt to evaluate this.

1) As a first pass, we will consider a few subgroups of patients: Those 50 and older (50+), 55+, 60+, 65+, and 70+. Please estimate/evaluate the probability a patient on TFD725+docetaxel will survive past 400 days in each of those subgroups. Please also give an interval estimate for each of those probabilities.

2) Now, in each of those subgroups evaluate whether TFD725+docetaxel is more effective than docetaxel alone (and the magnitude of any potential treatment effect). In addition, evaluate if the treatment effect appears to substantively and/or systematically differ across age (or if the data doesn't give a clear answer to this).

In answering these questions please use appropriate visuals! (please do not submit a zillion ugly tables!). Also attempt to be efficient in code-reuse (try not to copy/paste a lot of code! Use functions and loops where possible/appropriate)