

Homework # 5

Biostatistics 210

DUE December 5, 2017

Place in Hyeri You's mailbox at Mission Hall 2nd Floor, Suite 2653 (next to the Xerox copier)

This assignment will give you practice in applying a propensity based analysis to estimate causal effects.

1. DOWNLOAD

Download the datasets `hwk5.dta` from the course website. The variables are

- id (participant ID)
- died (1=died, 0=survived)
- hepatom (hepatomegaly: 1=present, 0=absent)
- edema (1=present, 0=absent)
- female (1=female, 0=male)
- bilirubin (in mg/dL)
- copper (copper in mg/dL)
- ast (AST in mg/dL)
- age (age at diagnosis in years)

2. CREATING PROPENSITIES

The primary interest is in estimating the causal effect of hepatomegaly on survival. Create a simple propensity model (no splines or interactions), make the quantiles, and evaluate the overlap.

```
logistic heptom female edema bilir copper ast age
predict out, p
xtile group=out, nq(5)
tab group heptom
```

*Question 1: (a) What percent of people in the lowest quintile have hepatomegaly?
(b) What percent of people in the highest quintile have hepatomegaly? Does there appear to be reasonable overlap?*

3. TESTING INTERACTION

Assess if the effect of hepatomegaly on survival appears homogeneous?

```
cc died heptom, by(group)

logistic died i.heptom##i.group
est store A
logistic died i.heptom i.group
lrtest A
```

Question 2: What do you conclude? Is there evidence of interaction?

4. CONDITIONAL ODDS RATIO

Question 3: Estimate the conditional odds ratio

```
logistic died i.hepatom i.group
```

5. CAUSAL MEAN AND AVERAGE CAUSAL EFFECT

Question 4: (Assuming no unmeasured confounders) If the entire population had hepatomegaly, what proportion would have died during follow-up? If the entire population did not have hepatomegaly, what proportion would have died during follow-up? What is the estimated risk difference?

```
logistic died i.hepatom i.group
margins hepatom
margins, dydx(hepatom)
```

6. REGRESSION MODEL

Question 5: Conduct the usual regression modeling approach adjusting for the variables in which were included in the propensity model. Have your results changed substantially? NOTE: It might be helpful to use the “margins” command here.