

Assignment 2

1. Consider a very large cohort study with no covariates measured, and consider the causal risk ratio $pr(Y^1=1)/pr(Y^0=1)$. You have available the usual information on treatment received and observed outcomes. Compute/derive bounds for the causal rate ratio from the data.
2. Consider a cohort in which all subjects are exposed to some chemical. Half of the subjects die from liver cancer. Without making any further assumptions, what proportion of the study subjects might have died from cancer had they not been exposed? What are the largest and smallest values of the causal risk difference compatible with the findings? What proportion of the study subjects were Type I (treatment fatal)?
3. Consider again the cholesterol data. Which of the parameters that you defined last time are available from standard analyses of variables observable in a normal study (randomization indicator and observed cholesterol)? Compute and discuss.
4. Consider the LRC-CPPT data shown in graphical form in the notes on randomized trials. Try to formulate a structural distribution model (i.e., a parsimonious model that maps percentiles of Y^a given A into percentiles of Y^0 (given $A=a$) that is compatible with the unequal variances in the different randomized groups. Is this model compatible with a deterministic model for the effect of treatment (i.e., one in which, if you observe Y^a and parameters in the model, you can compute Y^0)? Explain.
5. Provide a numerical example to show how weak ignorability ($A \perp Y^a$ for all a) can hold even though strong ignorability ($A \perp \underline{Y}^a$) does not hold.