STATISTICAL RETHINKING 2023
WEEK 5

1. The data in `data(NWOGrants)` are outcomes for scientific funding applications for the Netherlands Organization for Scientific Research (NWO) from 2010–2012 (see van der Lee and Ellemers doi:10.1073/pnas.1510159112). These data have a structure similar to the UCBAdmit data discussed in Chapter 11 and in lecture. There are applications and each has an associated gender (of the lead researcher). But instead of departments, there are disciplines. Draw a DAG for this sample. Then use the backdoor criterion and a binomial GLM to estimate the TOTAL causal effect of gender on grant awards.

2. Now estimate the DIRECT causal effect of gender on grant awards. Use the same DAG as above to justify one or more binomial models. Compute the average direct causal effect of gender, weighting each discipline in proportion to the number of applications in the sample. Refer to the marginal effect example in Lecture 9 for help.

3. Considering the total effect (problem 1) and direct effect (problem 2) of gender, what causes contribute to the average difference between women and men in award rate in this sample? It is not necessary to say whether or not there is evidence of discrimination or the presence or absence of unobserved confounds (which are likely!). Simply explain how the direct effects you have estimated make sense (or not) of the total effect.

4-OPTIONAL CHALLENGE. The data in `data(UFClefties)` are the outcomes of 205 Ultimate Fighting Championship (UFC) matches (see `?UFClefties` for details). It is widely believed that left-handed fighters (aka "Southpaws") have an advantage against right-handed fighters, and left-handed men are indeed over-represented among fighters (and fencers and tennis players) compared to the general population. Estimate the average advantage, if any, that a left-handed fighter has against right-handed fighters. Based upon your estimate, why do you think left-handers are over-represented among UFC fighters?