Homework 2: Sections 1.3, 2.1 - 2.4
KEY

STA209-04: Applied Statistics

February 1, 2019

Total Possible Points: 21

From the Book:

1.85) [3 pts]   Here, the confounding variable is the amount of ice on the roads. Other acceptable answers include temperature, road conditions, or anything else suggesting potentially more hazardous roads. [1 pts] This is a confounding variable in that it is related to both our explanatory variable (amount of salt spread on roads) and response variable (number of auto accidents). The icier the roads, the more salt that would be spread. Similarly, the icier the roads, the more likely an auto accident would occur. [2 pts]

1.89) [4 pts]

    a) [2 pts] The headline is appropriate for mice. The results cited by the headline were obtained through a randomized experiment performed on mice. Given that the experiment was randomized, causal claims may be made assuming the randomization process was such that confounding was mitigated.

    b) [2 pts] The headline is not appropriate for humans. A randomized experiment was performed only using mice and therefore can't be generalized to the human population. The only human data obtained was through an observational study whose results may be confounded by one or more lurking variables. A more appropriate headline would be "Antibiotics in infancy may be associated with obesity in adults".

Miscellaneous:

S1) [9 pts]

    a) [2 pts] Male Admission Rate: $1198/(1198 + 1493) = 0.445$. Female Admission Rate: $557/(557 + 1278) = 0.304$. Since males are admitted more often than females, there seems to be evidence of a gender bias.

    b) [2 pts] Let HS denote "High Selectivity" and MS denote "Moderate Selectivity". HS Male Admission: $333/(973 + 333) = 0.255$. HS Female Admission: $451/(451 + 1251) = 0.265$. MS Male Admission: $865/(865 + 520) = 0.625$. MS Female Admission: $106/(106 + 27) = 0.797$. When stratifying by program selectivity, the direction of the bias reverses. Females have higher admission rates than males in both highly and moderately selective programs.

    c) [5 pts] This is an example of Simpson's paradox. The marginal association between sex and admission changed after conditioning on program selectivity. This implies that selectivity is a confounding variable. We can explain this paradox by observing that substantially more females applied to more highly selective programs (this is not the case for males). As a result, the admissions rate among these programs has much greater weight when computing an overall rate. Given that the admissions rates among highly selective programs are generally lower than those for moderately selective programs, we then observe a much lower admission rate among females in the aggregate.

S2) [5 pts] Several answers possible. Check whether they wrote about three different things they found interesting [3 pts], found an example of a misleading graph, and explained why it was misleading. [2 pts]