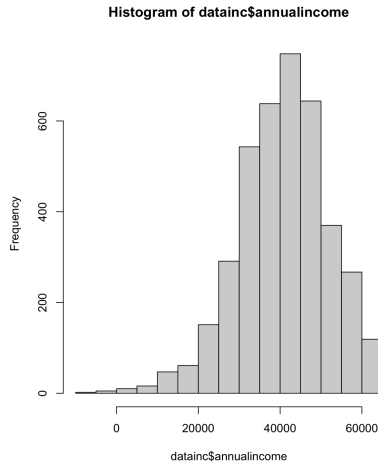


Section 1: Data Work

1) Shown in code



2) Ran merge command (thank you for including this super helpful)

3) The coefficient attached to a mother's education is 19.219. Given that the beta isn't 0 and the p-value is significant, there is a correlation between the mother's education and the baby's birth weight. However, we can not definitively say that one is causing the other. The mother's education may cause the baby's birth weight or the birth weight of the baby may impact the mother's education (for example, a woman has a premature baby and needs to take care of it and puts off school).

4) Shown in code

5) Shown in code

6) Shown in code

7) Shown in code

8) The coefficient attached to "above high school" is .027608 with a significant p-value. That means people who have above a high school education are less likely to have a baby with low birth weight.

9) There is a positive correlation between a person being white and whether they have a high school education. Given that the dependent variable is binary, it is a linear probability model meaning the coefficient attached to above high school education (.032536) says for every 1 year increase in schooling leads to a higher probability of the person being white.

10) Not quite sure what the question is asking because it references "9", but has the regression used in question "8". However, I do think the omitted variable is the same for both the regressions and the omitted variable is in the "annualized income" variable. There is a correlation between "above high school" and "annualized income". The independent variables can't be correlated with anything in the error term without creating a bias. Income will most certainly have an effect on the baby's birth weight.

11) I would prefer the multiple regression over the simple regression because of the reason I stated above. The error term should not include relevant variables and they shouldn't be correlated with the independent variables (breaking 0 conditional mean).

Section 2: Lectures

- 1)
 - a) We can say that the observed data establishes people who have insurance on average have a higher number of doctor visits. This question insinuates a causal relationship of having insurance increases doctor visits, which we can not determine solely based on the beta.
 - b) Given that the beta isn't equal to 0 means that there is an effect of insurance on the number of doctor visits. It could mean that those who have insurance are more likely to go to the doctor or those who go to the doctor are more likely to have insurance.
 - c) The demand for doctors is downward sloping for those with insurance, because as price decreases the demand for doctor visits increase.
- 2)
 - a) This suggests that there is a positive relationship between those who receive treatment and the amount of doctor visits and that the demand curve is downward sloping.
 - b) No, causality is harder to establish than correlation. The $\beta > 0$ shows that there is a positive correlation between those who received treatment (insurance) had a higher number of doctor visits. The observance of treatment and non-treatment group outcomes does not necessarily establish a causal relationship.
 - c) This is the same argument as in part b; therefore, the answer does not change. Just because the observation is that people who are insured on average visit the doctor more does not mean that insurance causes people to see the doctor. We would need to test for reverse causality to make sure the direction of causality.
 - d) The Oregon Health Insurance experiment was a randomized control experiment, which sought to answer whether a year of medicaid improved health (physical and mental). The findings of the study found "increases in hospital, outpatient, and drug utilizations; increases in compliance with recommended preventive care; and declines in exposure to substantial out-of-pocket medical expenses and medical debts" (Finkelstein et al, 2012). They also find self reported increases in physical and medical reports.

Section 3: Readings

1.
 - a) A quasi natural experiment is not like a randomized control experiment because they do not randomize people into control groups. Since medicare has a set of criteria (not making people random), it is automatically a quasi natural experiment. The eligibility is for people at least 65 years old and have worked for at least 10 years in covered employment. They run a regression discontinuity approach to measure the impacts of reaching the age 65 on the intensity of treatment.
 - b) One way to get around eliminating the people who are playing a waiting game is to

only study those individuals that need immediate care. They only use emergency department data (non-deferrable conditions) which helps limit the differences in health stock. This is similar to how people with deferrable conditions will avoid going to the hospital on weekends because of staffing issues/wait times. They find that people with non-deferrable conditions have a uniform distribution across all days of the week (including weekends) meaning that the differing health stock will not be a factor.

c) Insurance coverage: The findings show that there is an increase in the fraction of patients with medicare as their primary insurer at age 65, coupled with a decline in the fraction with no insurer.

Quality of Care: There is a significant increase of the mean length of stay between those with medicare (4.5%) and those without medicare. Certain tests and diagnosis increase at age 65 for those with medicare. Overall, the intensity of care increases for those with medicare vs those without.

Mortality Outcomes: Unsurprisingly, given the results above, once again medicare reduces mortality rates. "Medicare eligibility are relatively large: they represent a 14%-20% reduction in 7-day mortality, a 7%-9% in 28-day mortality, and a 2% reduction in 1-year mortality rates relative to those 64 years old with similar diagnosis" (Card et al., 2009)

d) One of the limitations of the study was that there was only one focus on the medicaid population which is similar to other medical insurance policies, medicare covers outpatient care, prescription drugs, etc. Another limitation is the discharge data which shows approximately 5% of 64-year-old patients in the sample have missing SSN, compared with the 4% of those over 65-year-olds.