QPM Final

Josh Klier

5/1/19

**Question 1:**

The intercept is interpreted as the thermometer rating of Clinton when both variables are set to 0. In this case, the intercept can be interpreted as the thermometer rating when a respondent is a strong republican (1) and they live in the North. Both variables Party ID and south are statistically significant. The interpretation of Party ID is that on average, keeping everything else the same, an increase of 1 in the Party ID will result in an increase of 6.09 of Clinton’s thermometer rating. The interpretation of South is that on average, keeping everything else the same, if a respondent lives in the South it will result in a decrease of 13.53 in Clinton’s thermometer rating.

In order to interpret that partisanship causes changes in attitudes towards Clinton we would need to assume that the sample size was large enough to determine a statistically significant conclusion, and that the respondents were randomly sampled and representative of the population. As it doesn’t seem that either of these things occurred in this situation, it would not seem reasonable to imply a causal relationship.

F= (r^2/p)/(1-r^2)/[n-(p+1)]

F= 15.38

We are comparing the amount of variance explained by the regression to the amount unexplained.

**Question 2:**

The intercept is interpreted as the thermometer rating of Clinton when all variables are set to 0. In this case, the intercept can be interpreted as the thermometer rating when a respondent is a strong republican (1) and when they live in the north. Both variables Party ID and North are statistically significant in the regression model. The interpretation of Party ID is that on average, keeping everything else the same and having South equal 0, an increase of 1 in the Party ID will result in an increase of 8.09 of Clinton’s thermometer rating. The interpretation of South is that on average, keeping everything else the same and having Party ID equal 1, if a respondent lives in the South it will decrease Clinton’s thermometer rating by 10.53. The interaction between South and Party ID is the effect of Party ID on thermometer rating of Clinton when the respondent lives in the South.

Y= 36 + 8.09X1 – 10.53X2 -3.57X3

Ysouth= 36 + (8.09-3.57)X1 – 10.53X2

YNorth= 36 + 8.09X1

They are different because of the interaction variable. This means that for people exclusively living in the south Party ID has less of an effect on Clinton’s Thermometer rating, meaning, on average, people in the south don’t like Clinton as much.

**Question #3:**

The causal effect of the mailers on attitudes towards Clinton is -3.57. It is represented by the interaction variable received Mailers X Wave 2. It represents the effect of mail on attitudes measured after mailers sent.

The key assumption is that the data collected was representative of the population and a random sample. This is clearly violated because they say they didn’t choose people at random to send the mailers to.

**Question #4:**

Mean= 93.87

Stdev= 9.50

N=11

Z score= (1-.1/2)

=1.644854

93.87 + (1.644854\*9.50)/sqrt(11)

93.87 - (1.644854\*9.50)/sqrt(11)

(89.158, 98.581)

**Question #5**

|  |  |  |  |
| --- | --- | --- | --- |
|  | yes | no | total |
| dem | 221  (205.58) | 225  (240.41) | 446 |
| rep | 139  (154.42) | 196  (180.58) | 335 |
| total | 360 | 421 | 781 |

* Calculate the cell component for the χ2 statistic for the lower-right cell of the table (i.e., Republicans who responded No).

(196-180.58)^2/180.58 = 1.316737

* The χ2 statistic for this table is (approximately) 5.02. Specify and conduct a hypothesis test using this number.

H0: Fobserved=Fexpected

Ha: Fobserved=/Fexpected

DF=(2-1)(2-1)=1

pchisq(5.02, df=1, lower.tail=FALSE)

=.025

P value < alpha, reject the null hypothesis. Party affiliation has an impact on attitudes towards the Affordable care act.

**Question #6**

Civics class

95% confidence interval. Point estimate= mean=3.8

Mean= 3.8, SD= 2.4, N=288, Z-score= qnorm(1-.05/2) = 1.96

3.8+(1.96\*2.4)/sqrt(288)

3.8-(1.96\*2.4)/sqrt(288)

(3.52, 4.07)

Art class

95% confidence interval. Point estimate=mean=3.5

Mean=35, SD=2.2, N=242, Zscore = 1.96

3.5 + (1.96\*2.2)/sqrt(242)

3.5 - (1.96\*2.2)/sqrt(242)

(3.22, 3.77)

Hypothesis test.

H0: Civics class has no effect on party ID

Ha: Civics class has an effect on Party ID

Standard error civics = 2.4/sqrt(288)

Standard error art= 2.2/sqrt(242)

Both equal = .1414

T stat= 2.4-2.2/(.1414) = 1.41

pnorm(1.41, lower.tail=TRUE) = .920

failure to reject the null at the .05 level. Civics has no effect on party ID.

We can treat is as causal because there is a big enough sample size and the the populations are randomly sampled.

**Question #7**

Population distribution is the whole set of values or individuals, while the sample distribution is a subset of that population. The sampling distribution statistic is used to describe the sample distribution, and depends on the size of the sample and population.

We might be concerned with autocorrelation. We can address this by using a lagged GDP variable.

**Question #8**

P-value- the probability that the test statistic equals the observed valyue or a value even more extreme in the direction predicted by Ha. Calculated by presuming that Ho is true.

Outlier- an observation is an outlier if it falls more than 1.5(IQR) above the upper quantile or more than 1.5(IQR) below the lower quantile.

Standard error- the standard deviation of the sampling distribution of ybar.

Instrument- used to account for unexpected behavior between variables in a regression.

Autocorrelation- the correlation of a signal with a delayed copy of itself as a function of delay.

Counterfactual- a comparison between what actually happened and what would have happened in the absence of the intervention.