MSCI 718

Assignment 2

Akash Kadiri (akadiri.uwaterloo.ca, 20925765**)**

Solution 1.b

p-value associated with the beta 1 is 0.0056 which is less than the level of significance 0.05 so we are rejecting the null hypothesis.

Solution 1.c

p-value of beta 1 is 0.285 which is greater than the level of significance 0.05 so we are accepting the null hypothesis

Solution 1.d

1. Statistical significance doesn't mean experimental significance. Statistical significance simply means that the results are unlikely to have been due to sampling error. Experimental significance means we can trust the results are real and will likely hold up in future studies. As mentioned in the article, it appears that statistical significance was achieved in this case, as the number of participants is very small given the effect size expected. The classical hypothesis testing is aimed at finding out if a sample data has features that make it possible to conclude that a hypothesis tested was true or false.
2. The classical hypothesis testing assumes that hypotheses are true, then it tests them against reality. As described above, the classical hypothesis testing isn't always sufficient for achieving scientific conclusions, and other possibilities need to be considered when there is controversy.
3. The classical hypothesis testing only allows publication of results that are "statistically significant" The classical hypothesis testing is designed to avoid publishing findings that aren't true. It is better than not publishing anything, but it doesn't provide information about the probability for the truth of our findings.
4. After reporting p-values of 0.05 or less, it is almost impossible to report results with p-value higher than 0.05. This means that researchers have a vested interest in reporting results with p-value lower than 0.05. It is true that most of the classical hypothesis testing results are p-values lower than 0.05, but it doesn't mean that all are.
5. The classical hypothesis testing does not provide information about the statistical significance of the result (i.e., almost all p-values are lower than 0.05).
6. The classical hypothesis testing doesn't provide information about the actual probability for the result being true or false. Also, it doesn't always consider the "p-value inaccuracy" that often occurs in behavioural sciences.
7. The classical hypothesis testing provides little information about whether the difference between groups X and Y is due to chance or other errors.

Solution 1.e

1. One type of bias, known as “p-hacking,” occurs when researchers collect or select data or statistical analyses until nonsignificant results become significant, and quantifying p-hacking is important because publication of false positives hinders scientific progress. When false-positive results enter the literature, they can be very persistent.
2. Text-mining is used to obtain reported p-values in papers drawn from a broad range of scientific disciplines, we then looked for evidence of p-hacking based on the shape of the p-curves which are produced from primary data used in published meta-analyses. This allows us to test the evidence for

p-hacking when looking at specific hypotheses which researchers have clearly identified as being of general interest.

3. Tests for evidential value and p-hacking can readily be used to detect biases in datasets used in meta-analyses, if researchers p-hack when there is no true effect, the p-curve will shift from being flat to left-skewed hence we employ binomial tests to look for evidence of evidential value and p-hacking in both our text-mined and meta-analyses datasets.

Solution 2.a

The scope of multiple linear regression model is limited due to normality assumptions of errors. The normal distribution is defined for continuous variables. However, for the current exercise the outcome variable is bounded between 0 and 1, it does not make sense to assume the error of the outcome is distributed normally.

Solution 2.b

The key characteristic of the data is that the outcome variable x is either 0 or 1 or the sum of repeated binary outcomes given n. To model this, we need to specify an outcome distribution that gives either 0 or 1 (Outcome distribution of binomial logistic function is given as – $\text{logit}(p) = \alpha + \beta 1x1 + \beta 2x2$)

Solution 2.c

The following are some interpretations we may draw based on the output:

1. For every one-unit change in marital_id i.e., x1, the log odds of vote intention, i.e., y, will decrease by -1.014. Here the marital status has negative impact on voting intention. If a person is not married, then there is a chance that the individual is democrat leaning
2. By observing the p value of marital_id it is quite evident that it is not significant in determining the vote intention as the p value is way too less than 1%
3. Also, by above step, the number 0.956 indicates that the individual being a republican leaning is decreasing by 4.4% if we do not have any predictor variables.
4. There is not much difference between Null deviance and Residual deviance through which we can conclude that the model isn't a good fit. The Residual deviance indicates the response by the model when all the variables are included. As the difference between null deviance and residual deviance should be significantly large to consider the model to be good fit.
5. It is always advisable to prefer a model that has less AIC value because it is the measure of fit which penalizes model for the number of independent variables. In this case it is way too high to consider the model to be good fit.