**PSYC234: Lecture 8 post-lecture worksheet**

This worksheet is to help you consolidate what you learned about the binary logistic regression during Lecture 8. It contains two activities.

This worksheet could be completed as part of the independent study hours for PSYC234. **It is optional but recommended**. **It is recommended that you complete this worksheet in advance of the WBA.**

Once you have finished, compare your answers to theanswer sheet provided on Moodle. You can also use this sheet and the answer sheet for revision purposes when preparing for the class test.

**Activity 1: Calculating odds ratios manually**

Activity 1 will involve work with the following data. You are a researcher interested in whether being excited (yes/no) predicts whether an individual passes their driving test (yes/no). Here is a table of frequencies:

|  |  |  |
| --- | --- | --- |
|  | Passed driving test | |
|  | Yes | No |
| Excited - Yes | 14 | 12 |
| Excited - No | 8 | 29 |

1. **What are the odds of passing the driving test in the “Excited – Yes” group?**
2. **What are the odds of passing the driving test in the “Excited – No” group?**
3. **What is the odds ratio (where “Excited – No” is the original odds)?**

1. **What does this odds ratio mean?**
2. **Is there evidence of quasi-complete separation or complete separation here? Give a reason for your answer.**

**Activity 2: Interpreting R output**

Activity 2 examines the following research question. You are a researcher interested in whether being rich (yes/no) predicts whether an individual owns a Tesla. Here is a table of frequencies:

|  |  |  |
| --- | --- | --- |
|  | Has a Tesla | |
|  | Yes | No |
| Rich - Yes | 5 | 5 |
| Rich - No | 34 | 3 |

You analyse this data in R and the output of your model is below. For the outcome, you set “0” as “Has a Tesla – No” and “1” as “Has a Tesla – Yes”

Graphical user interface, text, application

Description automatically generated

**Evaluating the model:**

**Model statistics:**

You run some code to produce the model’s chi-square statistic, the degrees of freedom and the p-value. These are displayed below:

* Chi square = 8.2
* Degrees of freedom = 1
* P-value = 0.004

1. What do these values indicate?

**Pseudo R2:**

Table

Description automatically generated with medium confidence

1. Which Pseudo R2 values might you report (based on the lecture)? What is the value of these Pseudo R2s?

**Evaluating the individual predictors:**

Looking back at the summary output:

1. What is the reference category for the predictor “Rich”?
2. What does the Intercept Estimate represent?
3. What does the RichYes Estimate represent?

Exponentiating the Estimates:

Graphical user interface, text, application

Description automatically generated

1. What does the Intercept represent?
2. What does the RichYes value represent?
3. Interpret the RichYes value

Producing confidence intervals:

Graphical user interface, text, application, email

Description automatically generated

1. What does the RichYes 95% confidence intervals represent?
2. From the p-value in the summary table for the RichYes row, what can you conclude?