Experimental Design and Data Analysis: Assignment 6

Andrew Bedard(2566978) & Simone van Gompel(2567525)
Group 19

May 23, 2015

Exercise 1

1

2

3

4

5

6

7

8

9

10

Exercise 2

1

The data contained in *psi.txt* was read in, the following figures were obtained.

2

Fitting a logistic regression model with psi and gpa as explanatory variables for the outcome being that the student passed their assessment or not, we obtain the following table:

```
            Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)     -11.602      4.213  -2.754  0.00589 **
psi               2.338      1.041   2.246  0.02470 *
gpa               3.063      1.223   2.505  0.01224 *
---
Signif. codes:
0  ***  0.001  **  0.01  *  0.05  .  0.1      1
```

Figure 1: Parameter estimation for logistic regression model

Thus we determine that our logistic regression model should be:

$$Pr(pass = 1) = \frac{\exp(-11.602 + 2.338 * psi + 3.063 * gpa)}{1 + \exp(-11.602 + 2.338 * psi + 3.063 * gpa)} \tag{1}$$

3

Based on the p-value obtained in Figure:1, we reject the null hypothesis that there is no effect of psi on the outcomes of the students final assessment. Further based on our parameters for the logistic regression model, we see that a positive value, ie. 1, for *psi* causes an increase in probability of passing, so we conclude that *psi* does in fact work.

4

To estimate the probability that a student with a *gpa* equal to 3 who receives *psi* passes the assignment, we simply enter our values into equation 1, our logistic regression model.

$$Pr(pass = 1) = \frac{\exp(-11.602 + 2.338 * (1) + 3.063 * (3))}{1 + \exp(-11.602 + 2.338(1) + 3.063 * (3))} = 0.4813$$

So there is a 48.13 % chance of a student with *gpa* of 3 who receives *psi* of passing the final assignment.

Similarly if we wish to estimate the probability that a student with a *gpa* of 3 who does not receive *psi* passing the final assignment, we simply replace the value of *psi* with 0, and substitute this into equation 1

$$Pr(pass = 1) = \frac{\exp(-11.602 + 2.338 * (0) + 3.063 * (3))}{1 + \exp(-11.602 + 2.338(0) + 3.063 * (3))} = 0.0822$$

Thus there is a 8.22 % chance of a student with a *gpa* of 3 who did not take *psi* passing the final assignment.

5

We may investigate the relative change in odds of passing the final assignment for students with *psi* rather than those without *psi* by simply taking the exponential of the change in our linear predictor, in this case taking $psi = 1$ as opposed to 0 increases our linear predictor by 2.338 as we can see in Figure:1, thus our difference in odds is $exp(2.338) = 10.3605$.

6

Using the alternative method of analysis, and obtaining the matrix:

```
      [,1] [,2]
[1,]     3    8
[2,]    15    6
```

```
        Fisher's Exact Test for Count Data

data:  x
p-value = 0.0265
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.02016297 0.95505763
sample estimates:
odds ratio
 0.1605805
        
```

7

8

Exercise 3

1

2

3

4

5

1 R-Code

1.1 Exercise 1

1.2 Exercise 2

1.3 Exercise 3

1.4 Exercise 4