

Linear and Logistic Regression Quiz

Jared Cross

2024-05-16

Spaceship Titanic

All of the questions below are based on linear and logistic regression models to predict which passengers were transported on the Spaceship Titanic.

Spaceship Titanic “Side”

```
table(train$side)
```

```
##  
##      P      S  
## 4206 4288
```

The spaceship titanic, like many ships has two sides, “Port” and “Starboard”.

As shown in the summary above, in the training set, there are 4206 passengers who were on the port “P” side of the ship and, 4288 passengers on the starboard “S” side.

Linear Regression

```
m1 = lm(Transported ~ side,  
        data=train)
```

```
summary(m1)
```

```
##  
## Call:  
## lm(formula = Transported ~ side, data = train)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -0.5550 -0.4513  0.4450  0.4450  0.5487   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept) 0.451260   0.007669  58.844  <2e-16 ***  
## sideS       0.103777   0.010793   9.615  <2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.4973 on 8492 degrees of freedom  
## (199 observations deleted due to missingness)  
## Multiple R-squared:  0.01077,    Adjusted R-squared:  0.01065   
## F-statistic: 92.45 on 1 and 8492 DF,  p-value: < 2.2e-16
```

1. The regression above uses the port side of the ship as the default and shows the effect of being on the starboard side (“S”). Please use it to predict the probability of being transported for someone on the starboard side of the ship.
2. Can you reject the null hypothesis that passengers on the two sides of the ship were equally likely to be transported (use $\alpha = 0.05$? Why or why not?

Age and P(Transported)

```
m2 = lm(Transported ~ Age,  
        data=train)
```

```
coef(m2)
```

```
## (Intercept)      Age  
## 0.578280401 -0.002589133
```

- Based on the results above, please write an equation describing the relationship between passenger age and the chance of being transported.
- Based on your equation, predict the chance of being transported for someone who is 10 years old.
- Predict the chance of being transported for someone who is 80 years old.

Multiple Regression

```
m3 = lm(Transported ~  
        side + Age,  
        data=train)
```

```
coef(m3)
```

```
## (Intercept)      sideS      Age  
## 0.528101223 0.106041148 -0.002707394
```

- Using the multiple regression above, predict the chance of being transported for someone who is 10 years old and was on the starboard side of the ship.
- Predict the chance of being transported for someone who was 80 years old and on the port side of the ship.

Logistic Regression

Note: $\text{VRDeck_log} = \log_e(\text{VRDeck} + 1)$

```
m4 = glm(Transported ~ VRDeck_log,
         data=train,
         family="binomial")

coef(m4)
```

```
## (Intercept) VRDeck_log
##    0.4788564  -0.2806875
```

8. Using the coefficients above, write an equation to predict the *log odds* of being transported based on VRDeck_log .

```
exp(coef(m4))
```

```
## (Intercept) VRDeck_log
##    1.6142273    0.7552643
```

9. Using the exponentiated coefficients above, write an equation to predict the odds of being transported based on VRDeck_log .
10. Based on this logistic regression, predict the odds of being transported for someone who spent 1000 at the VR deck (Please note that “ VRDeck_log ” is $\log_e(\text{VRDeck} + 1)$).
11. Based on this logistic regression, predict the probability of being transported for someone who spent 0 at the VR Deck.

Bonus

```
m5 = glm(Transported ~ VRDeck_log +  
        Age,  
        data=train,  
        family="binomial")
```

```
coef(m5)
```

```
## (Intercept)  VRDeck_log      Age  
##  0.559210235 -0.277149526 -0.002919541
```

```
exp(coef(m5))
```

```
## (Intercept)  VRDeck_log      Age  
##    1.7492904    0.7579412    0.9970847
```

Based on the logistic regression above, and using either the coefficients or the exponentiated coefficients, predict the probability of being transported for someone who is 20 years old and spent 250 at the VR Deck.