Linear and Logistic Regression Questions

Jared Cross

2024-05-01

Spaceship Titanic

Linear Regression

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

summary(m1)
```

```
##
## Call:
## lm(formula = Transported ~ CryoSleep, data = train)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8176 -0.3289  0.1824  0.1824  0.6711
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.32892    0.00599   54.91   <2e-16 ***
## CryoSleepTRUE  0.48866    0.01001   48.84   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4417 on 8474 degrees of freedom
##   (217 observations deleted due to missingness)
## Multiple R-squared:  0.2196, Adjusted R-squared:  0.2195
## F-statistic:  2385 on 1 and 8474 DF,  p-value: < 2.2e-16
```

1. Please write an equation to predict the probability of being Transported from CryoSleep.

2. Can you reject the null hypothesis that there is no association between CryoSleep and Transported?

3

. Please interpret the equation that you wrote in #1.

Room Service?

```
# Note:
# RoomService_log = log(RoomService + 1)

m2 = lm(Transported ~ RoomService_log,
        data=train)

coef(m2)
```

```
##     (Intercept) RoomService_log
##      0.62207810     -0.06629989
```

4. Please write an equation describing the relationship between room service and the chance of being transported.

5. Predict the chance of being transported for someone who spent 0 on room service.

6. Predict the chance of being transported for someone who spent 1000 on room service.

```
m3 = lm(Transported ~
          RoomService_log + CryoSleep,
        data=train)

coef(m3)
```

```
##     (Intercept) RoomService_log   CryoSleepTRUE
##      0.42014576     -0.03258604      0.39864845
```

7. Predict the chance of being transported for someone who spent 0 on room service and who was *not* in CryoSleep.

8. Predict the chance of being transported for someone who spent 1000 on room service and who was in CryoSleep.

Logistic Regression

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

coef(m4)
```

```
##   (Intercept) CryoSleepTRUE
##    -0.7130704     2.2131285
```

9. Write an equation to predict the *log odds* of being transported based on Cryosleep.

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

```
##   (Intercept) CryoSleepTRUE
##      0.490137      9.144279
```

10. Based on this logistic regression, write an equation to predict the odds of being transported based on Cryosleep.

11. Based on this logistic regression, predict the odds of being transported for someone who was in Cryosleep.

12. Based on this logistic regression, predict the probability of being transported for someone who was in Cryosleep.

Bonus

```
m5 = glm(Transported ~ CryoSleep +
           RoomService_log,
         data=train,
         family="binomial")

coef(m5)
```

```
##     (Intercept)   CryoSleepTRUE RoomService_log
##      -0.3133030       1.8215024      -0.1570825
```

11. Based on the logistic regression above, predict the probabilty of being transported for someone who was not in cryosleep and who spent 1000 on room serive.