Lab Assignment: Logistic Regression

Maggie Schweihs

Nov., 2017

Create a Word document from this R Markdown file for the following exercises. Submit the R markdown file and resulting Word document via D2L Dropbox.

## Exercise 1

A study was conducted whereby the type of anesthetic (A or B), nausea after the surgery (Yes or No), the amount of pain medication taken during the recovery period, and age for a random sample of 72 patients undergoing reconstructive knee surgery.

The data is in the file anesthesia.rda.

### Part 1a

Use R to create a two-way table with the type of anesthetic defining the rows and nausea after the surgery as the columns and also produce the output for a chi-square test for independence.

Is there an association between these two categorical variables at a 5% level of significance?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1a -|-|-|-|-|-|-|-|-|-|-|-

require(DS705data)  
data("anesthesia")  
nausea\_response <- table(anesthesia$anesthetic, anesthesia$nausea)  
nausea\_response

##   
## No Yes  
## A 13 26  
## B 23 10

chi\_result <- chisq.test(nausea\_response, correct = FALSE)  
chi\_result

##   
## Pearson's Chi-squared test  
##   
## data: nausea\_response  
## X-squared = 9.4545, df = 1, p-value = 0.002106

Nausea after surgery is not associated with the type of anethesia administered.

Nausea after surgery is associated with the type of anethesia administered.

Conclusion: Reject at . There is sufficient evidence in this sample to claim that nausea after surgery is associated with the type of anethesia administered ( 0.0021063).

### Part 1b

Obtain the output from R (including the Wald tests for coefficients - so use "summary" function) for the logistic regression model with nausea as the dependent variable and the type of anesthetic as the predictor variable.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1b -|-|-|-|-|-|-|-|-|-|-|-

attach(anesthesia)  
model <- glm(nausea~anesthetic, family = "binomial")  
summary(model)

##   
## Call:  
## glm(formula = nausea ~ anesthetic, family = "binomial")  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.4823 -0.8497 0.0254 0.9005 1.5453   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.6931 0.3397 2.041 0.04129 \*   
## anestheticB -1.5261 0.5088 -2.999 0.00271 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 99.813 on 71 degrees of freedom  
## Residual deviance: 90.133 on 70 degrees of freedom  
## AIC: 94.133  
##   
## Number of Fisher Scoring iterations: 4

### Part 1c

What is the outcome of the hypothesis test that the coefficient of **anesthetic** is "zero" vs "not zero" at a 5% level of significance? (use the Wald test from the R output from the logistic regression you performed)

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1c -|-|-|-|-|-|-|-|-|-|-|-

Outcome: Reject that **anesthetic** is "zero" at . There is sufficient evidence in this sample to claim that the coefficient for **anesthetic** is not zero. ( 0.002705).

### Part 1d

Convert the estimated coefficient of **anestheticB** to an odds ratio and interpret it in the context of the problem.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1d -|-|-|-|-|-|-|-|-|-|-|-

coeff\_anestheticB <- summary(model)$coefficients[2,1]  
odds\_anestheticB <- exp(coeff\_anestheticB)  
odds\_anestheticB

## [1] 0.2173913

The odds of having nausea after surgery where anesthetic B was administrered is only 21.7% as large as the odds of having nausea after surgery where anesthetic A was administered. That is, the odds of nausea after surgery is 78.3% less with anesthetic B.

### Part 1e

Install the package "mosaic" (if you don't have it installed already), then load it. Use the oddsRatio function to compute the odds ratio for having nausea for anesthetic A vs B. You may have to refer back to Week 8 for details on odds ratios and the oddsRatio function in R.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e -|-|-|-|-|-|-|-|-|-|-|-

require(mosaic)  
cp <- matrix(c(nausea\_response[1,], nausea\_response[2,]), nrow = 2)  
cp\_1 <- matrix(c(cp[2,], cp[1,]), nrow=2)  
rownames(cp\_1)<-c('A','B')  
colnames(cp\_1)<-c('Yes','No')  
cp\_1

## Yes No  
## A 26 13  
## B 10 23

oddsRatio(cp\_1, verbose = TRUE)

##   
## Odds Ratio  
##   
## Proportions  
## Prop. 1: 0.6667   
## Prop. 2: 0.303   
## Rel. Risk: 0.4545   
##   
## Odds  
## Odds 1: 2   
## Odds 2: 0.4348   
## Odds Ratio: 0.2174   
##   
## 95 percent confidence interval:  
## 0.2589 < RR < 0.7982   
## 0.0802 < OR < 0.5893   
## NULL

## [1] 0.2173913

The odds ratio of having nausea for those who were administered anesthetic A compared to those who were administered anesthetic B is .2174.

### Part 1f

When logistic regression coefficients are negative, the interpretation sometimes has more impact when we switch the perspective and use the reciprocal of the exponentiated coefficient. Find the odds ratio for having nausea for anesthetic B compared to anesthetic A (i.e. take the reciprocal of the odds ratio you computed in part **1d**).

Interpret this odds ratio in the context of the problem.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1f -|-|-|-|-|-|-|-|-|-|-|-

1/(exp(coeff\_anestheticB))

## [1] 4.6

Patients who are administered anesthetic A are 4.6 times more likely to experience nausea post-surgery.

### Part 1g

Compute the predicted probability of a reconstructive knee surgery patient having nausea after surgery when anesthetic A was used.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1g -|-|-|-|-|-|-|-|-|-|-|-

newdata <- data.frame(anesthetic='A')  
predict(model, newdata, type = "response")

## 1   
## 0.6666667

The patient has a .667 chance of experiencing nausea post surgery.

### Part 1h

Compute a 95% confidence interval for the predicted probability of a reconstructive knee surgery patient having nausea after surgery when anesthetic A was used.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1h -|-|-|-|-|-|-|-|-|-|-|-

out <- predict(model, newdata, se.fit = TRUE)  
C = .95 #confidence level  
crit = qnorm(1-(1-C)/2) #critical values  
lower = exp(out$fit - crit\*out$se.fit)/(1+exp(out$fit - crit\*out$se.fit))  
upper = exp(out$fit + crit\*out$se.fit)/(1+exp(out$fit + crit\*out$se.fit))  
c(lower, upper)

## 1 1   
## 0.5068447 0.7955831

With 95% confidence, the probability of the patient experiencing post-surgery nausea after receivin anesthesia A is between 0.507 and 0.796.

## Exercise 2

Continue using the anesthesia.rda data set to do the following.

### Part 2a

Obtain the output from R (including the Wald tests for coefficients - so use "summary" function) for the logistic regression model with nausea as the dependent variable and the amount of pain medication taken as the predictor variable.

At , is there a statistically significant relationship between nausea and the amount of pain medication taken?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 2a -|-|-|-|-|-|-|-|-|-|-|-

model2 <- glm(nausea~painmed, family = "binomial")  
model2\_sum <- summary(model2)  
model2\_sum

##   
## Call:  
## glm(formula = nausea ~ painmed, family = "binomial")  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.8555 -0.6167 -0.1072 0.8206 1.7894   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.062742 0.764501 -4.006 6.17e-05 \*\*\*  
## painmed 0.037487 0.008833 4.244 2.20e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 99.813 on 71 degrees of freedom  
## Residual deviance: 68.049 on 70 degrees of freedom  
## AIC: 72.049  
##   
## Number of Fisher Scoring iterations: 5

At , there is a statistically significant relationship between post-surgery nausea and the amount of pain medication taken ( 2.195535210^{-5}).

### Part 2b

Convert the estimated coefficient of **painmed** to an odds ratio and interpret it in the context of the problem.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 2b -|-|-|-|-|-|-|-|-|-|-|-

exp(model2\_sum$coefficients[2,1])

## [1] 1.038199

The odds of experiencing post-surgery nausea increases by 1.04 for each unit of pain medication taken when no other variables are taken into account.

### Part 2c

Compute the predicted probabilities of a reconstructive knee surgery patient having nausea in the recovery time after surgery for when 50 units of pain medication are used and also for when 100 units of pain medication are used.

Comment on these two probabilities.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 2c -|-|-|-|-|-|-|-|-|-|-|-

newdata1 <- data.frame(painmed=50)  
predict1 <- predict(model2, newdata1, type = "response")  
newdata2 <- data.frame(painmed=100)  
predict2 <- predict(model2, newdata2, type = "response")  
predict1;predict2

## 1   
## 0.2335485

## 1   
## 0.6650716

The patient has a .234 chance of experiencing post-surgery nausea when 50 units of pain medication are used. The odds increase to .665 when 100 units of pain medication are used.