Lab 9

Wayne Stewart

March 16, 2018

Goal:
- Look at dataset that already tells you if someone is senile or not
- Use other variables in the data to predict if they are senile
- If not, then get a probability of if they are senile or not!
- Compare results to see if you predicted it correctly

- Analysis of the Wais data set
  - Discussion of the data
  - Read in the data
  - Task 1: Complete the Jags script to analyze the data
- Algebraic investigation of the logistic model
  - Task 2: Solve for p
  - Task 3: If $odds = 1$ then $p = \frac{1}{2}$?
  - Task 4: If $odds < 1$ then $p <$?
- Odds and odds ratios
  - If $a = 1$ then:
  - If $a > 1$ then:
- Logistic model and parameters as they relate to odds
- Task 5: Find the expression for odds(0)
- Task 6: Find the expression for $X = x(p = 0.5)$
- Task 7: Add the logical nodes necessary to track the $OR_{x+1,x}$, $odds(0)$ and $X = x(p = 0.5)$ and run the Jags script – give point and interval estimates. Interpret output.

Analysis of the Wais data set

Discussion of the data

1. We consider the data of Agresti (1990, pp. 122-123)
2. 54 elderly people completed a subtest of the Wechsler Adult Intelligence Scale (WAIS) resulting in a discrete score with range from 0 to 20.
3. Aim: identify people with senility symptoms (binary variable) using the WAIS score.
4. Interest also lies in calculating WAIS scores that correspond to increased probability of senility symptoms (i.e., with $p > 0.5$).

Read in the data

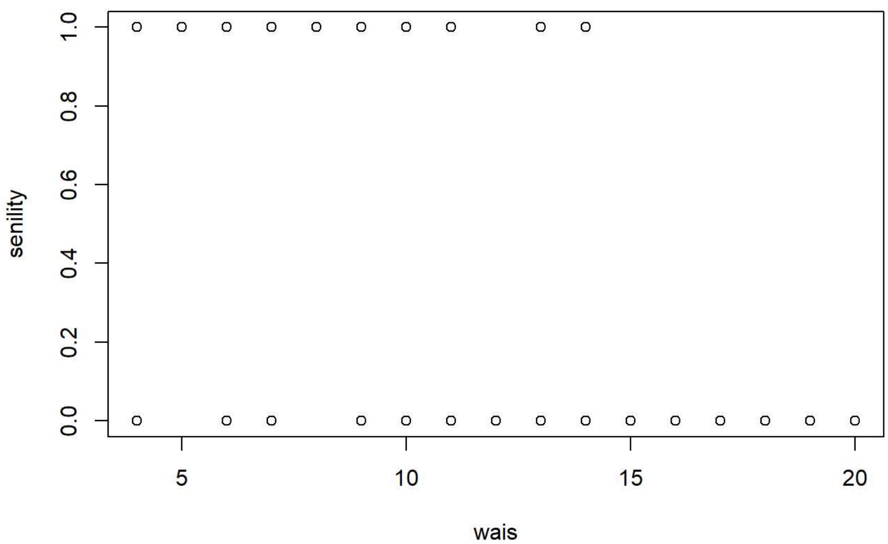
```
wais.df = read.csv("wais.csv")
head(wais.df)
```

```
##   wais senility
## 1    9        1
## 2   13        1
## 3    6        1
## 4    8        1
## 5   10        1
## 6    4        1
```

```
dim(wais.df)
```

```
## [1] 54  2
```

```
plot(senility ~ wais, data = wais.df)
```



Higher intelligence inversely relates to senility

Task 1: Complete the Jags script to analyze the data

```
library(rjags)
#Define the model:
modelString = "
model{
for(i in 1:Ntotal)
{

                                  Fill in this code


y[i] ~ dbin(p[i],1)


}




}
" # close quote for modelString                        Not much to change here
writeLines( modelString , con="TEMPmodel.txt" )

# Initialize the chains based on MLE of data.
# Option: Use single initial value for all chains:
#  thetaInit = sum(y)/length(y)
#  initsList = list( theta=thetaInit )

initsList = list(beta0 = -5, beta1 = 0.01)
# Run the chains:
jagsModel = jags.model( file="TEMPmodel.txt" , data=dataList , inits=initsList ,
                        n.chains=3 , n.adapt=500 )
list.samplers(jagsModel)

update( jagsModel , n.iter=500 )
codaSamples = coda.samples( jagsModel , variable.names=c("beta0", "beta1"),
                            n.iter=33340 )
save( codaSamples , file=paste0("lab9","Mcmc.Rdata") )

summary(codaSamples)

library(ggmcmc)
s = ggs(codaSamples)
ggs_density(s)

ggs_crosscorrelation(s)
```

Algebraic investigation of the logistic model

$$odds = \frac{p}{1-p}$$

Probability of success
/
prob. of failure

Task 2: Solve for p

$$p = odds(1-p)$$

$$p(1+odds) = odds$$

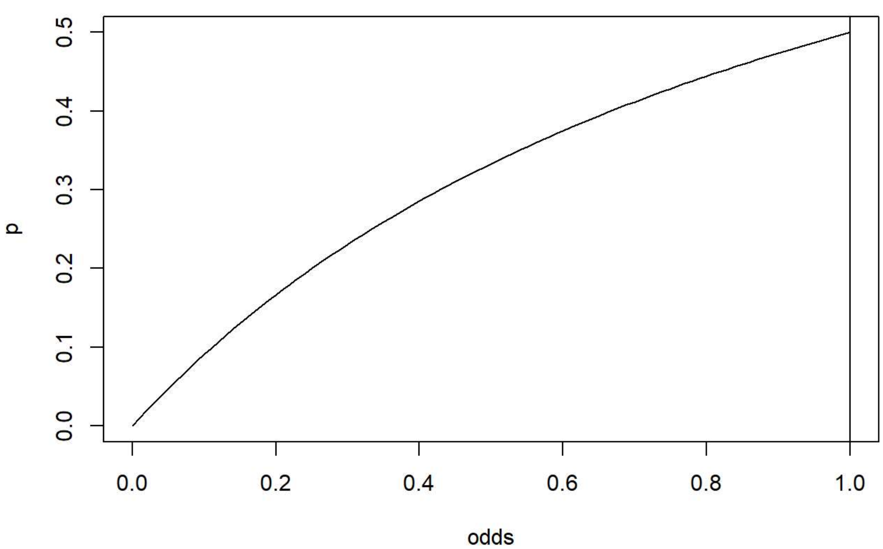
$$p = \frac{odds}{1+odds}$$

Task 3: If $odds = 1$ then $p = \frac{1}{2}$?

Task 4: If $odds < 1$ then $p <$?

$$p < 0.5$$

```
curve(x/(1+x),xlim=c(0,1),xlab="odds",ylab="p")
abline(v=1)
```



Plotted:

Probability of success
/
prob. of failure

Odds and odds ratios

The probability of a success is $p = P(Y=1)$, but if we use a linear predictor to estimate Y then p will be a function of covariates. Suppose we have just on x, then

How logit is laid out

$$logit(p_i) = \beta_0 + \beta_1 x_i$$

Interpretation:
Probability of success is dependent on...

Notice that p_i is function of x_i, you see that by noticing the index i.

$$Y_i \sim Bin(p_i, n_i)$$

Y is binomial

and using the definition of $logit$ we can write

$$log(odds_i) = \beta_0 + \beta_1 x_i$$

Actual model is a log of odds.

Remember log() is log base e!!!!

If we wish to compare the odds under two different X values:

We define:

$$odds(X=x) = \frac{P(Y=1|X=x)}{P(Y=0|X=x)}$$

Left X is random. little x is particular value of the random distribution

So that:

$$OR_{12} = \frac{odds(X=1)}{odds(X=2)} = a$$

Ratio of the odds.

Interpretation:
"Odds when X is 1 / odds when X is 2"

If $a = 1$ then:

$$odds(X=1) = odds(X=2)$$

If $a > 1$ then:

$$odds(X=1) > odds(x=2)$$

Logistic model and parameters as they relate to odds

$$logit(p) = log(odds(x)) = \beta_0 + \beta_1 x$$

Then:

$$exp(logit(p)) = exp(log(odds(x))) = exp(\beta_0 + \beta_1 x) = B_0 B_1^x$$

Get rid of front log, so Exponentiate

$$odds(x) = B_0 B_1^x$$

$$OR_{x+1,x} = \frac{odds(x+1)}{odds(x)} = \frac{B_0 B_1^{x+1}}{B_0 B_1^x} = B_1 = e^{\beta_1}$$

$$Percentage\ increase = \frac{final - initial}{initial} \times 100$$

$$percentage\ increase\ in\ odds\ for\ 1\ unit\ increase\ in\ x = \frac{B_0 B_1^x (B_1 - 1)}{B_0 B_1^x} \times 100 = (B_1 - 1) \times 100$$

Task 5: Find the expression for odds(0)

$$odds(0) = B_0 B_1^0$$

$$odds(0) = B_0 = exp(\beta_0)$$

Task 6: Find the expression for $X = x(p = 0.5)$

$$log(0.5/(1-0.5) = \beta_0 + \beta_1 x$$

Interpretation:
When p is 0.5 (probability of 50%), then what is the probability of being senile?

$$0 = \beta_0 + \beta_1 x$$

$$x = -\frac{\beta_0}{\beta_1}$$

What is the independent variable when the probability of success is 50%?

This is for the case when disease probability is 0.5.

E.g., what is the WAIS score when the probability of senility is 50%?

Task 7: Add the logical nodes necessary to track the $OR_{x+1,x}$, $odds(0)$ and $X = x(p = 0.5)$ and run the Jags script – give point and interval estimates. Interpret output.

To answer task 7 you will need to be able to interpret the meaning of these nodes.