20160915_Thompson_HW-01

Nick Thompson

September 15, 2016

1 Method of MLE

Find the Maximum Likelihood Estimate for π.

- The probability model of the binomial distribution is:

$$f(y|n,\pi) = \binom{n}{y}\pi^y(1-\pi)^{n-y}$$

- Step 1: Write down the likelihood function (I did this step for you):

$$\mathcal{L}(\pi|y,n) = p(y|n,\pi) = \binom{n}{y}\pi^y(1-\pi)^{n-y}$$

- Step 2: Take the log and reduce

$$log\mathcal{L}(\pi|y,n) = \sum_{i=1}^{n} log\pi^y * \sum_{i=1}^{n} log(1-\pi)^{n-y}$$

$$log\mathcal{L}(\pi|y,n) = \sum_{i=1}^{n} y\ log\pi * \sum_{i=1}^{n}(n-y)\ log(1-\pi)$$

- Step 3: Maximize and find the most likely value for π (Start with the log-likelihood and take the derivative with respect to the parameter).

$$\frac{dlog\mathcal{L}}{d\pi} = n\bar{y}\log\pi + (n-\bar{y})n\log(1-\pi) \rightarrow$$

$$\Rightarrow \frac{n\bar{y}}{\pi} - \frac{(n-y)n}{(1-\pi)} = 0 \rightarrow \frac{\bar{y}}{\pi} - \frac{(n-y)}{(1-\pi)} = 0 \rightarrow$$

$$\Rightarrow \frac{\bar{y}}{\pi} = \frac{(n-y)}{(1-\pi)} \rightarrow \bar{y}(1-\pi) = \pi(n-y) \rightarrow$$

$$\Rightarrow \bar{y} - \pi\bar{y} = \pi n - \pi\bar{y} \rightarrow \bar{y} = \pi n$$

2 R

Include your code. Report results only where you see *Report your results* printed.

2.1

The maximizing function we used in class used the "seq" function to generate a sequence from .01 to .99 in increments of .01.

```r
#this command
p.seq <- seq(0.01, 0.99, 0.01)
```

- How would you change the code to create a sequence from .01 to .99 in increments of .1?

```r
p.seq2 <- seq(0.01, 0.99, 0.1)
```

- How would you get R to print the first ten items in the sequence you just created? (look at subsetting)

```r
p.seq2_sub <- p.seq2[1:10]
```

2.2

The **csv** file located at this link [https://vincentarelbundock.github.io/Rdatasets/csv/datasets/presidents.csv]
contains self-reported votes in the 1992 U.S. presidential election. Additional documentation is available at
this link.

- Read the data in to R using **read.csv** command.
- Create a cross tab that shows vote choice (rows) by party identification (columns). *Report your results.*

```r
rm(list=ls())
cat("\014")

#presidents <- read.csv("http://vincentarelbundock.github.io/Rdatasets/csv/pscl/vote92.csv")
#save(presidents, file = ("presidents.RData"))
# if making the script from scratch, uncomment the presidents object immediately above then run load()
load("presidents.RData")

presidents$party[presidents$dem==1]<-"Democrat"
presidents$party[presidents$rep==1]<-"Republican"
presidents$party[presidents$rep==0 & presidents$dem==0]<-"Other"
my.table <- table(presidents[,"vote"],presidents[,"party"]); my.table
```

```
##
##            Democrat Other Republican
##    Bush          23    13        274
##    Clinton      356    30         30
##    Perot         65    31         87
```

- Use the **prop.table** command to get the column percentages. *Report your results*

```r
prop.table(my.table,2)*100
```

```
##
##             Democrat      Other Republican
##    Bush      5.180180 17.567568  70.076726
##    Clinton  80.180180 40.540541   7.672634
##    Perot    14.639640 41.891892  22.250639
```