HW9

IST 772 Homework 9

Due December 7th, 2021 at 8:00AM EDT

Homework 9 by Nora Lin: I produced the material below with no assistance.

```
#install.packages("~/Desktop/BaylorEdPsych_0.5.tar.gz", repos = NULL, type = "source")
library(BaylorEdPsych)
```

Excercise 1 p.234:

```
glmOut <- glm(vs~gear+hp,family=binomial(),data=mtcars)
summary(glmOut)
```

```
##
## Call:
## glm(formula = vs ~ gear + hp, family = binomial(), data = mtcars)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.76095  -0.20263  -0.00889   0.38030   1.37305
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 13.43752    7.18161   1.871   0.0613 .
## gear        -0.96825    1.12809  -0.858   0.3907
## hp          -0.08005    0.03261  -2.455   0.0141 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 43.860  on 31  degrees of freedom
## Residual deviance: 16.013  on 29  degrees of freedom
## AIC: 22.013
##
## Number of Fisher Scoring iterations: 7
```

```
#Analysis: The median residual is negative which suggests that that the distribution of residuals is po
```

Excercise 5 p.234:

```
PseudoR2(glmOut)
```

```
##         McFadden     Adj.McFadden        Cox.Snell       Nagelkerke
##        0.6349042        0.4525061        0.5811397        0.7789526
## McKelvey.Zavoina           Effron            Count        Adj.Count
##        0.8972195        0.6445327        0.8125000        0.5714286
##              AIC    Corrected.AIC
##       22.0131402       22.8702830
```

```
#Analysis:

#The Nagelkerke pseudo R-squared value is  0.7789526. For any given measure, you can loosely interpret
```

Excercise 6 p.234:

```
library(car)
```

```
## Loading required package: carData
```

```
data(Chile)
```

```
ChileN = Chile[Chile$vote=="N",]
ChileY = Chile[Chile$vote=="Y",]
ChileYN = rbind(ChileN, ChileY)
ChileYN = ChileYN[complete.cases(ChileYN),]
ChileYN$vote = factor(ChileYN$vote, levels=c("N","Y"))
str(ChileYN)
```

```
## 'data.frame':    1703 obs. of  8 variables:
##  $ region    : Factor w/ 5 levels "C","M","N","S",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ population: int  175000 175000 175000 175000 175000 175000 175000 175000 175000 175000 ...
##  $ sex       : Factor w/ 2 levels "F","M": 2 1 1 1 2 1 2 1 2 2 ...
##  $ age       : int  29 49 23 28 26 24 41 20 20 44 ...
##  $ education : Factor w/ 3 levels "P","PS","S": 2 1 3 1 2 3 1 2 2 2 ...
##  $ income    : int  7500 35000 35000 7500 35000 15000 15000 15000 35000 35000 ...
##  $ statusquo : num  -1.296 -1.032 -1.105 -1.047 -0.786 ...
##  $ vote      : Factor w/ 2 levels "N","Y": 1 1 1 1 1 1 1 1 1 1 ...
```

```
chileGLM = glm(vote~age+statusquo,family=binomial(),data=ChileYN)
summary(chileGLM)
```

```
##
## Call:
## glm(formula = vote ~ age + statusquo, family = binomial(), data = ChileYN)
##
## Deviance Residuals:
```

```
##     Min       1Q   Median       3Q      Max
## -3.2095  -0.2830  -0.1840   0.1889   2.8789
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.193759   0.270708  -0.716   0.4741
## age          0.011322   0.006826   1.659   0.0972 .
## statusquo    3.174487   0.143921  22.057   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2360.29  on 1702  degrees of freedom
## Residual deviance:  734.52  on 1700  degrees of freedom
## AIC: 740.52
##
## Number of Fisher Scoring iterations: 6
```

```
#Analysis:

#The coefficient on status quo predictor is statistically significant based on the 22.057 and p-vlaue o
```

```
exp(coef(chileGLM))
```

```
## (Intercept)         age   statusquo
##   0.8238564   1.0113863  23.9145451
```

```
#Analysis:

#The intercept represents odds of 0.82:1 for a Yes vote. For age the odds are close to 1:1, this means
```

```
#install.packages('MCMCpack')
library(MCMCpack)
```

```
## Loading required package: coda
```

```
## Loading required package: MASS
```

```
## ##
## ## Markov Chain Monte Carlo Package (MCMCpack)
```

```
## ## Copyright (C) 2003-2021 Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park
```

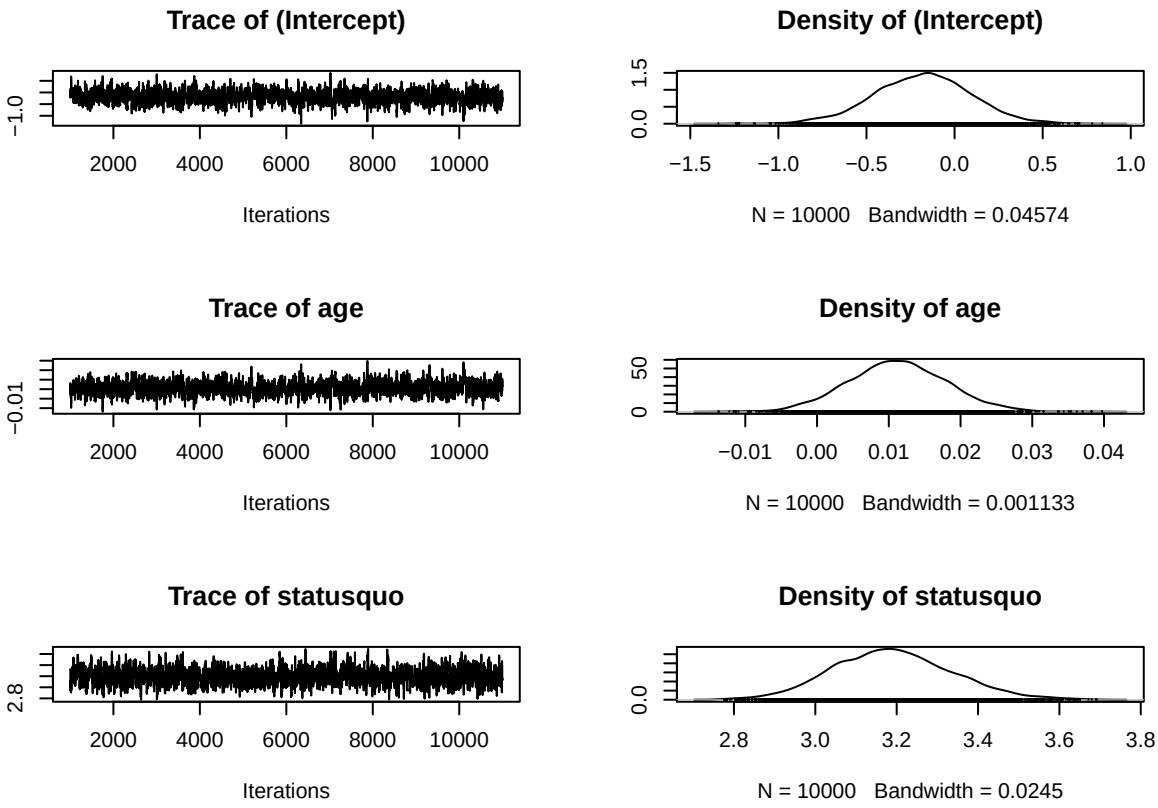
```
## ##
## ## Support provided by the U.S. National Science Foundation
```

```
## ## (Grants SES-0350646 and SES-0350613)
## ##
```

```r
ChileYN$vote = as.numeric(ChileYN$vote)-1
ChilesBayes=MCMClogit(vote~age+statusquo, data=ChileYN)
summary(ChilesBayes)
```

```
##
## Iterations = 1001:11000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                 Mean       SD  Naive SE Time-series SE
## (Intercept) -0.18272 0.272640 2.726e-03       0.008938
## age          0.01123 0.006817 6.817e-05       0.000223
## statusquo    3.19061 0.145853 1.459e-03       0.004993
##
## 2. Quantiles for each variable:
##
##                 2.5%       25%      50%        75%   97.5%
## (Intercept) -0.742761 -0.365241 -0.17552 -0.0003872 0.34439
## age         -0.002005  0.006733  0.01121  0.0157683 0.02499
## statusquo    2.914442  3.087259  3.18546  3.2847388 3.48698
```

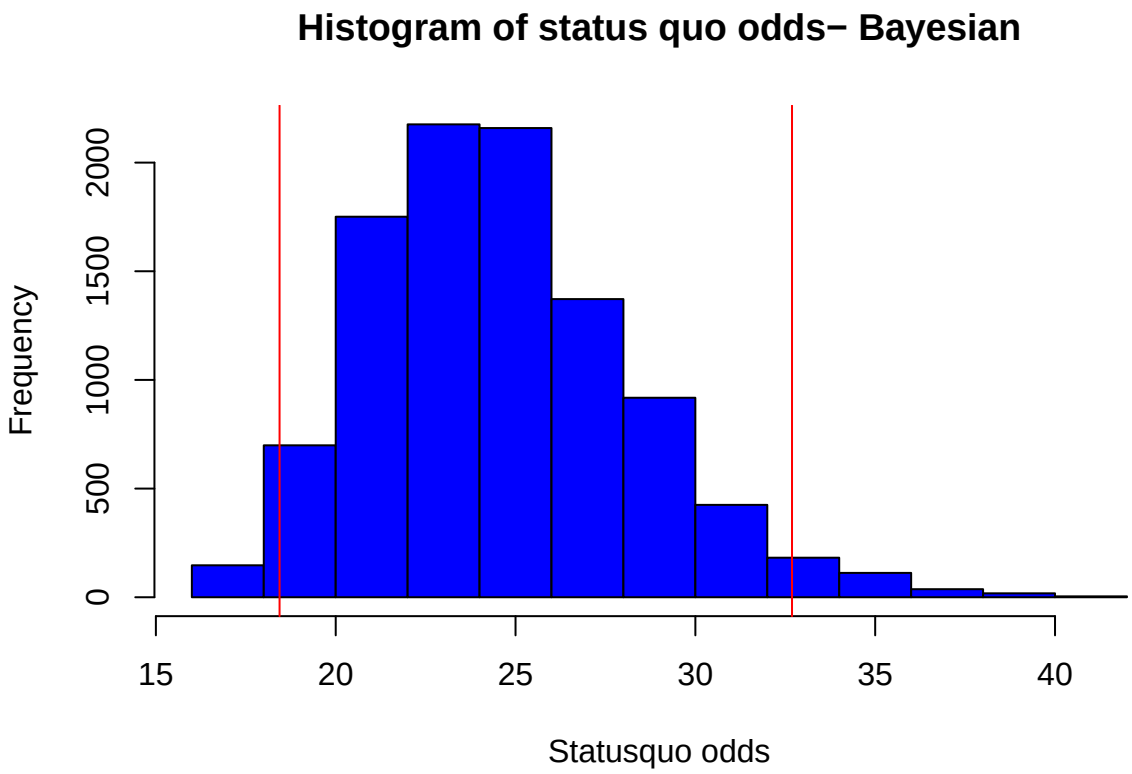
```r
plot(ChilesBayes)
```



Excercise 7 p.234:

```
OddsHistogram <-function(mcmc_out,seq){
  logOdds <- as.matrix(mcmc_out[,3])
  odds <- apply(logOdds,1,exp)
  hist(odds, col="blue",
       main= "Histogram of status quo odds- Bayesian",
       xlab= 'Statusquo odds')
abline(v=quantile(odds,c(0.025)),col="red")
abline(v=quantile(odds,c(0.975)),col="red")
}

OddsHistogram(ChilesBayes,3)
```




```
#Analysis:
#The histogram above shows the posterior distribution of odds for the status quo predictor variable. Th
```