Baruch College, STA-CIS 3920, Exercise#8 Anil Poonai 37

6 August 2020

**Exercise**

Code is in the appendix

This is the SAheart data set with famhist being the y variable. Both the cross-validating and BICq models agree.

A screenshot of a social media post

Description automatically generated

This is the Cardiac dataset with death being the y variable. They models don’t agree and only one has an actual variable.

A screenshot of a cell phone

Description automatically generated

This is the classification space graph. Now this makes sense only due to the bestglm function giving me one variable only so there most likely would not be a huge overlap between the variables I used.A screenshot of a survey

Description automatically generated

**Appendix**

R version 3.6.0 (2019-04-26) -- "Planting of a Tree"

Copyright (C) 2019 The R Foundation for Statistical Computing

Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Previously saved workspace restored]

> install.packages("tidyverse

+ ")

Installing package into ‘C:/Users/poona/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

--- Please select a CRAN mirror for use in this session ---

Warning message:

package ‘tidyverse

’ is not available (for R version 3.6.0)

> install.packages("bestglm")

Installing package into ‘C:/Users/poona/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

also installing the dependencies ‘iterators’, ‘foreach’, ‘shape’, ‘leaps’, ‘glmnet’, ‘grpreg’, ‘pls’

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/iterators\_1.0.12.zip'

Content type 'application/zip' length 343881 bytes (335 KB)

downloaded 335 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/foreach\_1.5.0.zip'

Content type 'application/zip' length 145971 bytes (142 KB)

downloaded 142 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/shape\_1.4.4.zip'

Content type 'application/zip' length 793405 bytes (774 KB)

downloaded 774 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/leaps\_3.1.zip'

Content type 'application/zip' length 111495 bytes (108 KB)

downloaded 108 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/glmnet\_4.0-2.zip'

Content type 'application/zip' length 2105877 bytes (2.0 MB)

downloaded 2.0 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/grpreg\_3.3.0.zip'

Content type 'application/zip' length 375494 bytes (366 KB)

downloaded 366 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/pls\_2.7-2.zip'

Content type 'application/zip' length 1230783 bytes (1.2 MB)

downloaded 1.2 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/bestglm\_0.37.3.zip'

Content type 'application/zip' length 999309 bytes (975 KB)

downloaded 975 KB

package ‘iterators’ successfully unpacked and MD5 sums checked

package ‘foreach’ successfully unpacked and MD5 sums checked

package ‘shape’ successfully unpacked and MD5 sums checked

package ‘leaps’ successfully unpacked and MD5 sums checked

package ‘glmnet’ successfully unpacked and MD5 sums checked

package ‘grpreg’ successfully unpacked and MD5 sums checked

package ‘pls’ successfully unpacked and MD5 sums checked

package ‘bestglm’ successfully unpacked and MD5 sums checked

The downloaded binary packages are in

C:\Users\poona\AppData\Local\Temp\RtmpAri0xV\downloaded\_packages

> library(tidyverse)

Error in library(tidyverse) : there is no package called ‘tidyverse’

> library(bestglm)

Loading required package: leaps

Warning messages:

1: package ‘bestglm’ was built under R version 3.6.3

2: package ‘leaps’ was built under R version 3.6.3

> data(SAheart)

> SAheart.Scruba=SAheart[,1:4]

> SAheart.Scrubb=SAheart[,6:10]

> SAheart.Scrub=cbind(SAheart.Scruba,SAheart.Scrubb)

> famhist=SAheart$famhist

> famhist = ifelse(famhist=='Present',1,0)

> SAheart.Scrub=cbind(SAheart.Scrub,famhist)

> CV.out=bestglm(SAheart.Scrub,IC="CV",family=binomial)

Morgan-Tatar search since family is non-gaussian.

> CV.out

CVd(d = 373, REP = 1000)

BICq equivalent for q in (0.0703191746407851, 0.895940044901826)

Best Model:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -1.75407801 0.330944614 -5.300216 1.156656e-07

age 0.02485654 0.007500281 3.314082 9.194469e-04

chd 0.91518643 0.216342727 4.230262 2.334192e-05

> best.out=bestglm(SAheart.Scrub,IC="BICq",family=binomial,TopModels=1)

Morgan-Tatar search since family is non-gaussian.

> best.out

BICq(q = 0.25)

BICq equivalent for q in (0.0703191746407851, 0.895940044901826)

Best Model:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -1.75407801 0.330944614 -5.300216 1.156656e-07

age 0.02485654 0.007500281 3.314082 9.194469e-04

chd 0.91518643 0.216342727 4.230262 2.334192e-05

> #Models agree

> Cardiac <- read\_csv("C:/Users/poona/Downloads/Cardiac.csv")

Error in read\_csv("C:/Users/poona/Downloads/Cardiac.csv") :

could not find function "read\_csv"

> library(Readr)

Error in library(Readr) : there is no package called ‘Readr’

> install.packages("Readr")

Installing package into ‘C:/Users/poona/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning messages:

1: package ‘Readr’ is not available (for R version 3.6.0)

2: Perhaps you meant ‘readr’ ?

> 2

[1] 2

> install.packages("readr")

Installing package into ‘C:/Users/poona/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

also installing the dependencies ‘assertthat’, ‘glue’, ‘utf8’, ‘digest’, ‘cli’, ‘ellipsis’, ‘fansi’, ‘lifecycle’, ‘magrittr’, ‘pillar’, ‘pkgconfig’, ‘rlang’, ‘vctrs’, ‘Rcpp’, ‘tibble’, ‘hms’, ‘R6’, ‘clipr’, ‘crayon’, ‘BH’

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/assertthat\_0.2.1.zip'

Content type 'application/zip' length 55094 bytes (53 KB)

downloaded 53 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/glue\_1.4.1.zip'

Content type 'application/zip' length 153714 bytes (150 KB)

downloaded 150 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/utf8\_1.1.4.zip'

Content type 'application/zip' length 215250 bytes (210 KB)

downloaded 210 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/digest\_0.6.25.zip'

Content type 'application/zip' length 249457 bytes (243 KB)

downloaded 243 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/cli\_2.0.2.zip'

Content type 'application/zip' length 397909 bytes (388 KB)

downloaded 388 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/ellipsis\_0.3.1.zip'

Content type 'application/zip' length 46066 bytes (44 KB)

downloaded 44 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/fansi\_0.4.1.zip'

Content type 'application/zip' length 223777 bytes (218 KB)

downloaded 218 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/lifecycle\_0.2.0.zip'

Content type 'application/zip' length 101652 bytes (99 KB)

downloaded 99 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/magrittr\_1.5.zip'

Content type 'application/zip' length 157499 bytes (153 KB)

downloaded 153 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/pillar\_1.4.6.zip'

Content type 'application/zip' length 182418 bytes (178 KB)

downloaded 178 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/pkgconfig\_2.0.3.zip'

Content type 'application/zip' length 22334 bytes (21 KB)

downloaded 21 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/rlang\_0.4.7.zip'

Content type 'application/zip' length 1149552 bytes (1.1 MB)

downloaded 1.1 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/vctrs\_0.3.2.zip'

Content type 'application/zip' length 1190773 bytes (1.1 MB)

downloaded 1.1 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/Rcpp\_1.0.5.zip'

Content type 'application/zip' length 3125634 bytes (3.0 MB)

downloaded 3.0 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/tibble\_3.0.3.zip'

Content type 'application/zip' length 415753 bytes (406 KB)

downloaded 406 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/hms\_0.5.3.zip'

Content type 'application/zip' length 109695 bytes (107 KB)

downloaded 107 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/R6\_2.4.1.zip'

Content type 'application/zip' length 59304 bytes (57 KB)

downloaded 57 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/clipr\_0.7.0.zip'

Content type 'application/zip' length 52042 bytes (50 KB)

downloaded 50 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/crayon\_1.3.4.zip'

Content type 'application/zip' length 750158 bytes (732 KB)

downloaded 732 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/BH\_1.72.0-3.zip'

Content type 'application/zip' length 18270741 bytes (17.4 MB)

downloaded 17.4 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.6/readr\_1.3.1.zip'

Content type 'application/zip' length 1551085 bytes (1.5 MB)

downloaded 1.5 MB

package ‘assertthat’ successfully unpacked and MD5 sums checked

package ‘glue’ successfully unpacked and MD5 sums checked

package ‘utf8’ successfully unpacked and MD5 sums checked

package ‘digest’ successfully unpacked and MD5 sums checked

package ‘cli’ successfully unpacked and MD5 sums checked

package ‘ellipsis’ successfully unpacked and MD5 sums checked

package ‘fansi’ successfully unpacked and MD5 sums checked

package ‘lifecycle’ successfully unpacked and MD5 sums checked

package ‘magrittr’ successfully unpacked and MD5 sums checked

package ‘pillar’ successfully unpacked and MD5 sums checked

package ‘pkgconfig’ successfully unpacked and MD5 sums checked

package ‘rlang’ successfully unpacked and MD5 sums checked

package ‘vctrs’ successfully unpacked and MD5 sums checked

package ‘Rcpp’ successfully unpacked and MD5 sums checked

package ‘tibble’ successfully unpacked and MD5 sums checked

package ‘hms’ successfully unpacked and MD5 sums checked

package ‘R6’ successfully unpacked and MD5 sums checked

package ‘clipr’ successfully unpacked and MD5 sums checked

package ‘crayon’ successfully unpacked and MD5 sums checked

package ‘BH’ successfully unpacked and MD5 sums checked

package ‘readr’ successfully unpacked and MD5 sums checked

The downloaded binary packages are in

C:\Users\poona\AppData\Local\Temp\RtmpAri0xV\downloaded\_packages

> library(readr)

Warning message:

package ‘readr’ was built under R version 3.6.3

> Cardiac <- read\_csv("C:/Users/poona/Downloads/Cardiac.csv")

Parsed with column specification:

cols(

.default = col\_double()

)

See spec(...) for full column specifications.

> death=Cardiac$death

> death= ifelse(death==1,0,1)

> Cardiac.Scrub=Cardiac[,1:15]

> Cardiac.Scrub=cbind(Cardiac.Scrub,death)

> Cardiac.Scrub$`%mphr(b)`=NULL

> CV.out=bestglm(Cardiac.Scrub,IC="CV",family=binomial)

Morgan-Tatar search since family is non-gaussian.

There were 50 or more warnings (use warnings() to see the first 50)

> CV.out

CVd(d = 454, REP = 1000)

BICq equivalent for q in (0, 0.437766660377008)

Best Model:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -3.102342 0.2086608 -14.86787 5.327953e-50

> best.out=bestglm(Cardiac.Scrub,IC="BICq",family=binomial,TopModels=1)

Morgan-Tatar search since family is non-gaussian.

> best.out

BICq(q = 0.25)

BICq equivalent for q in (0, 0.437766660377008)

Best Model:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -0.8594004 0.8922678 -0.9631642 0.33546515

sbp -4.3733355 1.7955096 -2.4357071 0.01486271

> #Models don't agree

> #Alright so because they only gave one variable, I'm going to use that plus any of the other variables that were admitted from the model.

> glm.fit=glm(death~sbp+bhr,data=Cardiac.Scrub,family=binomial)

> newdata=ProbeX

> X=Cardiac.Scrub[,c(1,5)]

> StdX=apply(X,2,scale)

> dfX=as.data.frame(StdX)

> glm.probs=predict(glm.fit,newdata=dfX,type="response")

> length(glm.probs)

558

> StCard2=as.data.frame(cbind(dfX,death))

> glm.fit2=glm(death~sbp+bhr,data=StCard2,family=binomial)

> names(dfX)

“bhr” “sbp”

> names(StProbeX)[1]="bhr"

> names(StProbeX)[2]="sbp"

> glm.probe=predict(glm.fit2,newdata=StProbeX,type="response")

> glm.y=glm.probe

> glm.y[glm.probe>.5]=1

> glm.y[glm.probe<.5]=0

> ProbeGlm(ProbeX=StProbeX,ProbeYhat = c(glm.y),InX = dfX,InY = Cardiac.Scrub$death,xr=c(-3,3),yr=c(-3,3))

>