BIOS6611-Homework5-Randy-20191001

Randy

10/1/2019

##### Question1A: calculate the sensitivity and specificity

DN <- c(0,0,3,9,16,18)  
DY <- c(2,1,4,5,3,1)  
a <- sum(DY[1:4])  
b <- sum(DN[1:4])  
c <- sum(DY[5:6])  
d <- sum(DN[5:6])  
diagnosis <- as.table(matrix(c(a,b,c,d), ncol=2, byrow=TRUE))  
dimnames(diagnosis) <- list(CMMS=c("Less", "More"), ClinicDiagnosis=c("DementedN","DementedY"))  
diagnosis

## ClinicDiagnosis  
## CMMS DementedN DementedY  
## Less 12 12  
## More 4 34

sensitivity <- function(M){  
 sen <- M[1,1]/sum(M[, 1])  
 cat("The sensitivity is", sen, "\n")  
}   
sensitivity(diagnosis)

## The sensitivity is 0.75

specificity <- function(M){  
 spe <- M[2,2]/sum(M[, 2])  
 cat("The specificity is", spe, "\n")  
}   
specificity(diagnosis)

## The specificity is 0.7391304

##### Question1B: Make a table showing sensitivity and specificity

for cut-off values of 5, 10, 15, 20, 25, or 30.

DN <- c(0,0,3,9,16,18,0)  
DY <- c(2,1,4,5,3,1,0)  
TableSS <- function(n){  
 a <- sum(DY[1:n])  
 b <- sum(DN[1:n])  
 c <- sum(DY[(n+1):7])  
 d <- sum(DN[(n+1):7])  
 M <- as.table(matrix(c(a,b,c,d), ncol=2, byrow=TRUE))  
 sen <- M[1,1]/sum(M[, 1])  
 spe <- M[2,2]/sum(M[, 2])  
 return(list(sen, spe))  
}  
library(purrr)  
result1 <- unlist(map(1:6, TableSS))  
  
css1 <- as.table(matrix(result1, nrow=2))  
dimnames(css1) <- list( Attributes=c( "sensitivity","specificity"), The\_cut\_off\_value\_for\_CMMS\_Score= c("5","10","15","20","25","30"))  
css1

## The\_cut\_off\_value\_for\_CMMS\_Score  
## Attributes 5 10 15 20 25 30  
## sensitivity 0.1250000 0.1875000 0.4375000 0.7500000 0.9375000 1.0000000  
## specificity 1.0000000 1.0000000 0.9347826 0.7391304 0.3913043 0.0000000

##### Question1C: a possible consequence of a false positive, then a possible consequence of a false negative.

If the diagnosis is a flase positive, a healthy person might get mental stress and treatment unnecessary. Especially, there are certain side effects with Cholinesterase inhibitors, such as donepezil, rivastigmine, and galantamine. The side effects include nausea, vomiting and diarrhea. Other possible side effects include cardiovascular irregulation, fainting and sleep disturbances.

If the diagnosis is false negative, there will be a delay for the treatment and severe the demetia symptoms. Especially for the depression, parkinsonism oragitation, it costs the life quality for the patients and the family.

##### Question1D: select a cut-off for the CMMS, assuming CMMS false

positives and false negatives are equally undesirable

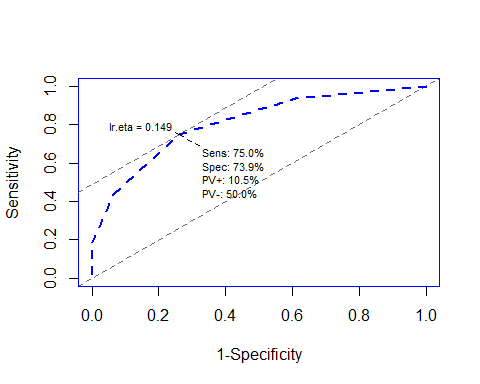
DN <- c(0,0,3,9,16,18,0)  
DY <- c(2,1,4,5,3,1,0)  
TableNP <- function(n){  
 a <- sum(DY[1:n])  
 b <- sum(DN[1:n])  
 c <- sum(DY[(n+1):7])  
 d <- sum(DN[(n+1):7])  
 M <- as.table(matrix(c(a,b,c,d), ncol=2, byrow=TRUE))  
 fnr <- M[2,1]/sum(M[, 1])  
 fpr <- M[1,2]/sum(M[, 2])  
 return(list(fnr, fpr))  
}  
result2 <- unlist(map(1:6, TableNP))  
css2 <- as.table(matrix(result2, nrow=2))  
dimnames(css2) <- list( Attributes=c( "FNR","FPR"), The\_cut\_off\_value\_for\_CMMS\_Score= c("5","10","15","20","25","30"))  
css2

## The\_cut\_off\_value\_for\_CMMS\_Score  
## Attributes 5 10 15 20 25  
## FNR 0.87500000 0.81250000 0.56250000 0.25000000 0.06250000  
## FPR 0.00000000 0.00000000 0.06521739 0.26086957 0.60869565  
## The\_cut\_off\_value\_for\_CMMS\_Score  
## Attributes 30  
## FNR 0.00000000  
## FPR 1.00000000

The cut-off score of 20 is equally undesirable for the for false positive and false negative values. It has really both higher FPR 0.2608696 and higher FNR css2[1,4]

##### Question1E: Plot the ROC curve and obtain the AUC using R

library(Epi) #load library  
DN <- c(0,0,3,9,16,18)  
DY <- c(2,1,4,5,3,1)  
cutoff <- seq(5,30, by=5)  
DementY <- rep(cutoff, DY)  
y <- DementY  
y[] <- 1  
DementX <- rep(cutoff, DN)  
x <- DementX  
x[] <- 0  
DemInf <- c(x,y)  
CMMS <- c(DementX, DementY)  
DemCMMS <- as.data.frame(cbind(CMMS,DemInf))  
par(col = "blue", lty = 2)  
ROC(form = DemInf~CMMS, data = DemCMMS, plot = "ROC", PV = T, MI = F, grid = F, AUC = F)



library(pROC)

## Type 'citation("pROC")' for a citation.

##   
## Attaching package: 'pROC'

## The following objects are masked from 'package:stats':  
##   
## cov, smooth, var

auc(DemInf, CMMS)

## Setting levels: control = 0, case = 1

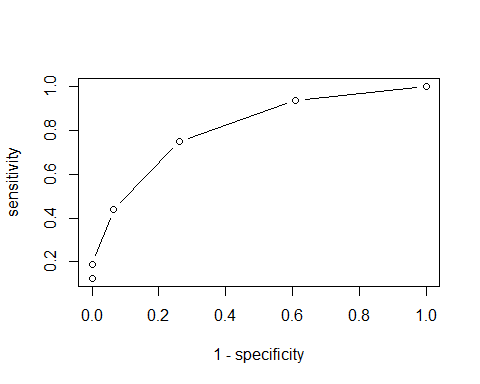
## Setting direction: controls > cases

## Area under the curve: 0.8091

Because the plots and the ROC show the same shape, I would assume that the AUC are the same for these two tests. The AUC is 0.8091; it is larger than 0.5 but not quite near to 1. CMMS can discriminates between people with and without dementia, but not very accurate.

##### Question1F: Extra Credit: Plot the ROC curve using your answers in B. How similar is it to the AUC from part E?

sensitivity <- css1[1,]  
specificity <- css1[2,]  
plot(1-specificity, sensitivity, type = "b")



part.auc <- function(a){  
 s <- 1/2\*(sensitivity[a]+sensitivity[a+1])\*(specificity[a]-specificity[a+1])   
}  
n <- 1:5  
unlist(map(n, part.auc))

## 5 10 15 20 25   
## 0.00000000 0.02038043 0.11616848 0.29347826 0.37907609

trapezoid <- sum(unlist(map(n, part.auc)))  
trapezoid

## [1] 0.8091033

Based on the trapezoid rule, the result of the AUC is the same as the result from pROC package, auc() function. They are exactly the same.

##### Question1G: Comment on the difference in the sets of predictive values for the two prevalence values.

DN <- c(0,0,3,9,16,18,0)  
DY <- c(2,1,4,5,3,1,0)  
TablePV <- function(n){  
 a <- sum(DY[1:n])  
 b <- sum(DN[1:n])  
 c <- sum(DY[(n+1):7])  
 d <- sum(DN[(n+1):7])  
 M <- as.table(matrix(c(a,b,c,d), ncol=2, byrow=TRUE))  
 ppv <- M[1,1]/sum(M[1, ])  
 npv <- M[2,2]/sum(M[2, ])  
 return(list(ppv, npv))  
}  
result3 <- unlist(map(1:4, TablePV))  
css3 <- as.table(matrix(result3, nrow=2))  
dimnames(css3) <- list( Attributes=c( "PPV","NPV"), The\_cut\_off\_value\_for\_CMMS\_Score= c("5","10","15","20"))  
css3

## The\_cut\_off\_value\_for\_CMMS\_Score  
## Attributes 5 10 15 20  
## PPV 1.0000000 1.0000000 0.7000000 0.5000000  
## NPV 0.7666667 0.7796610 0.8269231 0.8947368

sensitivity <- css1[1,]  
specificity <- css1[2,]  
PPV\_10 <- sensitivity\*0.1/((sensitivity\*0.1)+(1-specificity)\*0.9)  
PPV\_40 <- sensitivity\*0.4/((sensitivity\*0.4)+(1-specificity)\*0.6)  
NPV\_10 <- specificity\*0.9/((1-sensitivity)\*0.1+specificity\*0.9)  
NPV\_40 <- specificity\*0.6/((1-sensitivity)\*0.4+specificity\*0.6)  
rbind( PPV\_10, css3["PPV",], PPV\_40, NPV\_10, css3["NPV",], NPV\_40)

## 5 10 15 20 25 30  
## PPV\_10 1.0000000 1.0000000 0.4270557 0.2421053 0.1461245 0.100000  
## 1.0000000 1.0000000 0.7000000 0.5000000 1.0000000 1.000000  
## PPV\_40 1.0000000 1.0000000 0.8172589 0.6571429 0.5066079 0.400000  
## NPV\_10 0.9113924 0.9171975 0.9373297 0.9637795 0.9825625 NaN  
## 0.7666667 0.7796610 0.8269231 0.8947368 0.7666667 0.779661  
## NPV\_40 0.6315789 0.6486486 0.7136929 0.8160000 0.9037657 NaN

“PPV and NPV are very dependent on prevalence. Using study prevalence is not appropriate when calculating their values, unless it’s from a large prospective study from which prevalence can be well estimated.” from lecture 9 notes.

##### Question1Hi: Is there a single cutoff that makes the CMMS a good test for both ruling in and ruling out dementia?

LR\_plus <- sensitivity/(1-specificity)  
LR\_minus <- (1-sensitivity)/specificity  
LR <- as.table(matrix(c(LR\_plus, LR\_minus), nrow=2, byrow=TRUE))  
dimnames(LR) <- list(Attributes=c("LR+", "LR-"), The\_cut\_off\_value\_for\_CMMS\_Score= c("5","10","15","20","25","30"))  
LR

## The\_cut\_off\_value\_for\_CMMS\_Score  
## Attributes 5 10 15 20 25 30  
## LR+ Inf Inf 6.7083333 2.8750000 1.5401786 1.0000000  
## LR- 0.8750000 0.8125000 0.6017442 0.3382353 0.1597222

There is not a single cutoff that makes the CMMS a good test for both ruling in and ruling out dementia.

##### Question1Hii: Obtain the posterior odds of dementia and the posterior odds of no dementia for the various combinations of LR+ and LR-.

prior\_odd <- 0.3   
prior\_p <- prior\_odd/(1+prior\_odd)  
postior\_odd\_D <- prior\_odd \* LR\_plus  
postior\_odd\_N <- (1-prior\_p)/prior\_p \* 1/LR\_minus  
postior\_odd\_D; postior\_odd\_N

## 5 10 15 20 25 30   
## Inf Inf 2.0125000 0.8625000 0.4620536 0.3000000

## 5 10 15 20 25 30   
## 3.809524 4.102564 5.539452 9.855072 20.869565 NaN

Postior odds for dementia decreas with the cutoff scores, closely associated with LR+; in the contrast, the postior odds for non-dementia increases with cutoff socres, associated with LR-.