Hypoxemia Analysis

Christopher Thornton

2023-06-04

```r
library(dplyr)
data <- readr::read_csv("Q2 Data.csv")

data <- data %>% select(c(RecordNo,Age,`Gender(Male,Female)`,BMI,`Desats LT 90`,`Desats LT 80`,`Desats 

# data <- data %>% select(!c(Selected,RecordNo,`Recording Date`,Height,HeightUnit,Weight,WeightUnit,`Re

set.seed(06042023)
i <- sample(1:nrow(data),.8*nrow(data),replace=F)
train <- data[i,]
test <- data[-i,]

# glm0 <- glm(Arrhythmia~.,data=train,family = binomial)
# summary(glm0)
glm1 <- glm(Arrhythmia~`Desats LT 70`,data=train,family = binomial)
summary(glm1)
```

```
##
## Call:
## glm(formula = Arrhythmia ~ `Desats LT 70`, family = binomial,
##     data = train)
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.0489529  0.1296198  -8.093 5.84e-16 ***
## `Desats LT 70`  0.0006312  0.0326627   0.019    0.985
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 361.84  on 315  degrees of freedom
## Residual deviance: 361.84  on 314  degrees of freedom
## AIC: 365.84
##
## Number of Fisher Scoring iterations: 4
```

```r
glm2 <- glm(Arrhythmia~`Desats LT 80`,data=train,family = binomial)
summary(glm2)
```

```
##
## Call:
## glm(formula = Arrhythmia ~ `Desats LT 80`, family = binomial,
##     data = train)
```

```
## 
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.043732   0.130294  -8.011 1.14e-15 ***
## `Desats LT 80` -0.001427   0.006867  -0.208    0.835
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 361.84  on 315  degrees of freedom
## Residual deviance: 361.79  on 314  degrees of freedom
## AIC: 365.79
## 
## Number of Fisher Scoring iterations: 4
glm3 <- glm(Arrhythmia~`Desats LT 90`,data=train,family = binomial)
summary(glm3)
```

```
## 
## Call:
## glm(formula = Arrhythmia ~ `Desats LT 90`, family = binomial,
##     data = train)
## 
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.073599   0.136624  -7.858  3.9e-15 ***
## `Desats LT 90`  0.001020   0.001838   0.555    0.579
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 361.84  on 315  degrees of freedom
## Residual deviance: 361.54  on 314  degrees of freedom
## AIC: 365.54
## 
## Number of Fisher Scoring iterations: 4
pred_p <- predict(glm1,type="response",newdata = test)
roc_data <- pROC::roc(response=test$Arrhythmia,predictor=pred_p)
```

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

```
## Setting direction: controls < cases
```

```
plot(roc_data,main="Desats LT 70")
```

Desats LT 70


```
pROC::auc(roc_data)
```
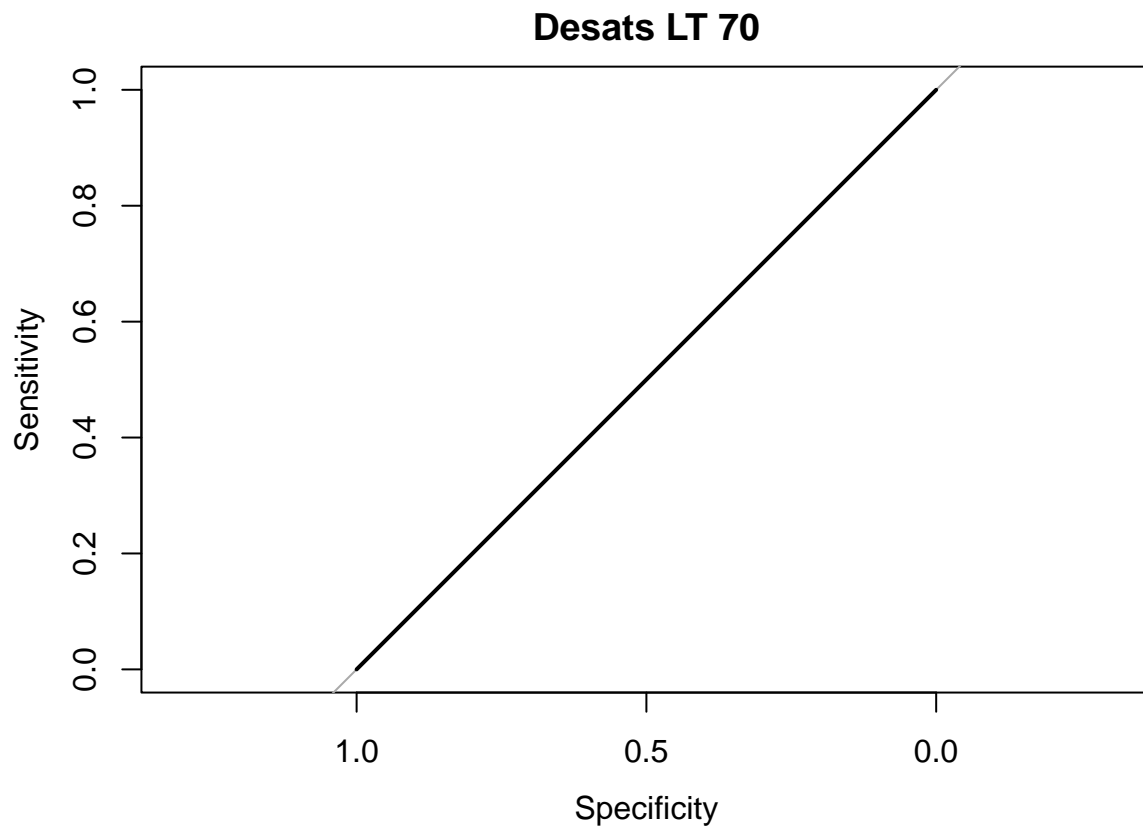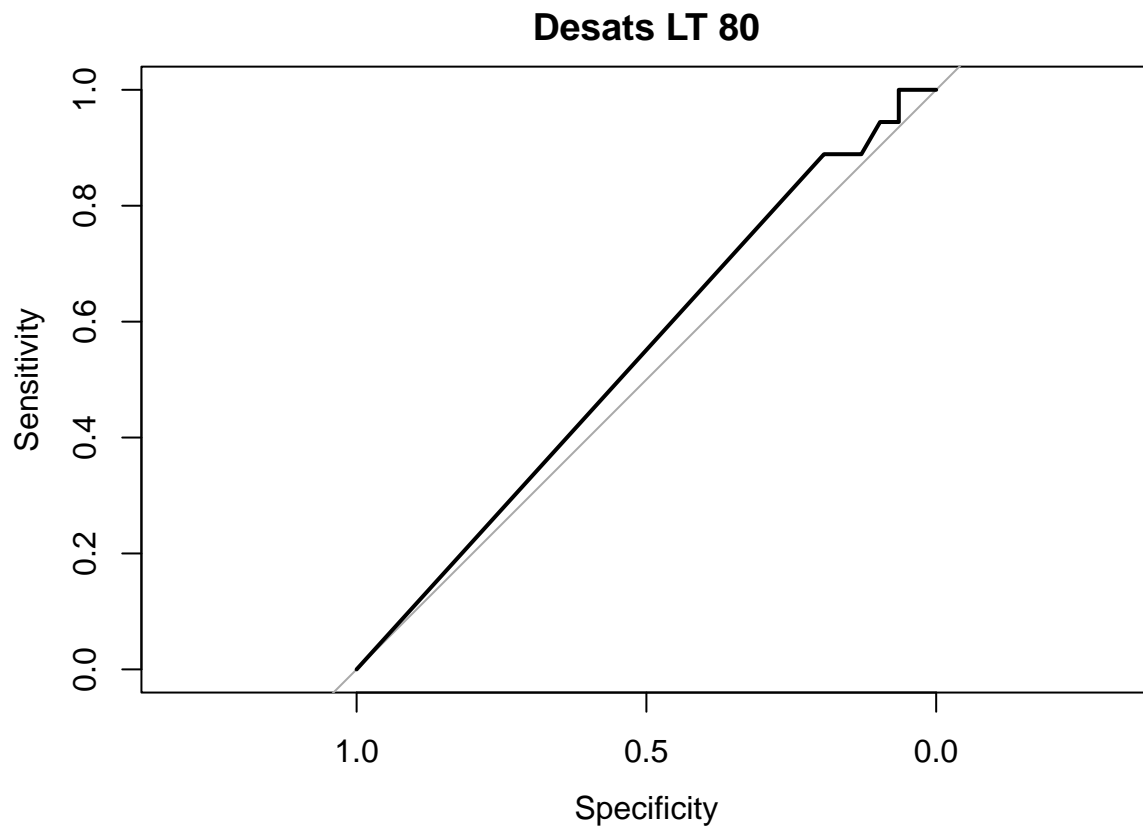
```
## Area under the curve: 0.5
```

```
pred_p <- predict(glm2,type="response",newdata = test)
roc_data <- pROC::roc(response=test$Arrhythmia,predictor=pred_p)
```

```
## Setting levels: control = 0, case = 1
## Setting direction: controls < cases
```

```
plot(roc_data,main="Desats LT 80")
```

Desats LT 80


```
pROC::auc(roc_data)
```

```
## Area under the curve: 0.5403
```
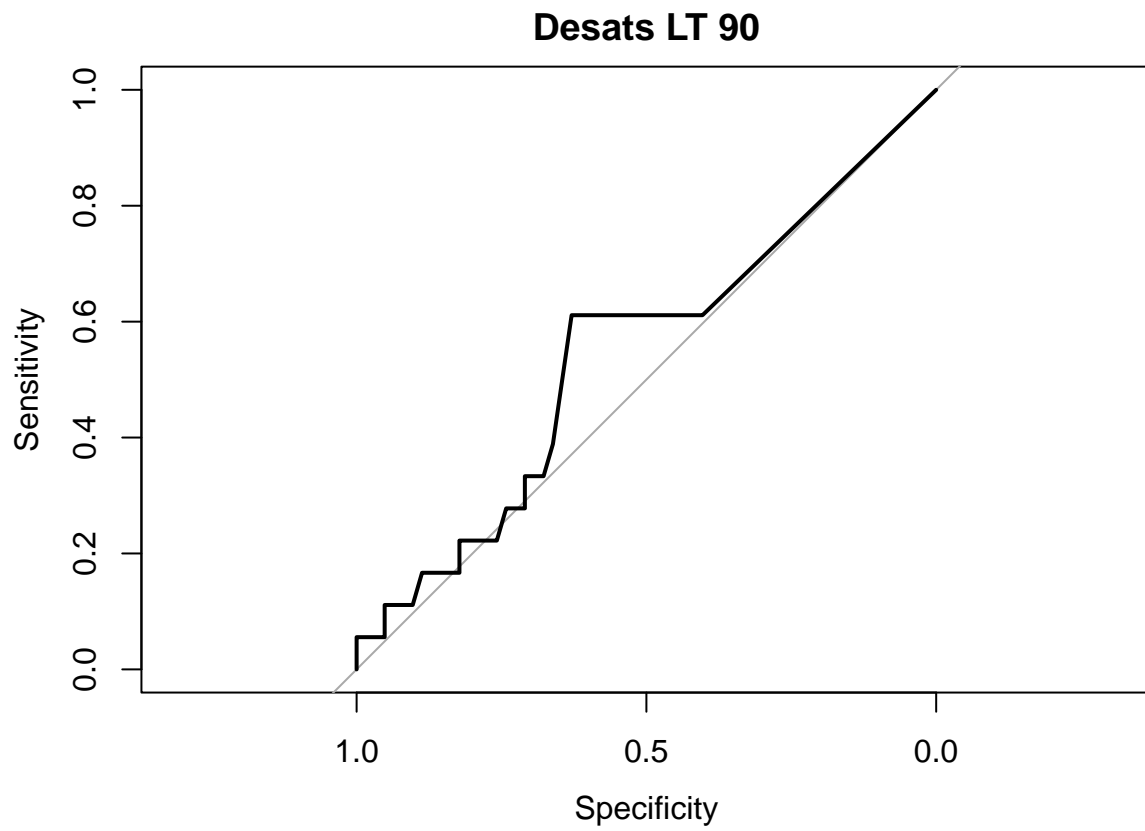
```
pred_p <- predict(glm3,type="response",newdata = test)
roc_data <- pROC::roc(response=test$Arrhythmia,predictor=pred_p)
```

```
## Setting levels: control = 0, case = 1
## Setting direction: controls < cases
```

```
plot(roc_data,main="Desats LT 90")
```

Desats LT 90


```
pROC::auc(roc_data)
```

```
## Area under the curve: 0.5439
```