

Laboratorium 7

1. Wielowymiarowa regresja logistyczna

W pewnym szpitalu przeprowadzono badanie skuteczności pewnego testu medycznego na obecność pewnej choroby. Zgromadzono 47 pacjentów chorych na tę chorobę oraz 79 pacjentów, u których wykluczono tę chorobę. Następnie każdy z pacjentów został poddany testowi medycznemu którego duże wartości świadczą o chorobie. Zbadano również każdego pacjenta pod kontem obecności pewnego czynnika ryzyka dla badanej choroby. Wyniki przedstawione są w pliku Dane 7.1. Korzystając z modelu regresji logistycznej wyznacz funkcję szacującą prawdopodobieństwo wystąpienia choroby na podstawie zebranych danych. Dla każdego predyktora oblicz iloraz szans wraz z przedziałami ufności. Zobrazuj wartość prognostyczną dwuwymiarowego predyktora za pomocą krzywej ROC.

```
library(readxl)
```

```
library(caret)
```

```
library(pROC)
```

```
Dane <- read_excel("Dane 7.1.xlsx")
```

```
D<-Dane$"D"
```

```
T<-Dane$"T"
```

```
G<-Dane$"G"
```

```
n<-length(D)
```

```
lik<-function(x)
```

```
{
```

```
  x1 <- x[1]
```

```
  x2 <- x[2]
```

```
  x3 <- x[3]
```

```
  L<-D*log(1/(1+exp(-(x1+x2*T+x3*G))))+(1-D)*log(1-1/(1+exp(-(x1+x2*T+x3*G))))
```

```
  return(-sum(L))
```

```
}
```

```
Opt<-optim(c(0,0,0), lik)
```

```
C<-Opt$par
```

```
P<-1/(1+exp(-(C[1]+C[2]*T+C[3]*G)))
```

```
V<-cbind(T,G,P,D)
```

```
V=V[order(V[,1],decreasing=FALSE),]
```

```
V
```

```
J<-c()
```

```
for(i in 1:n)
```

```
  J[i]=1
```

```
X<-cbind(J,T,G)
```

```
I<-matrix(nrow = 3, ncol = 3)
```

```
for(a in 1:3)
```

```
{
```

```
  for(b in 1:3)
```

```
  {
```

```
    w<-C[1]+C[2]*X[,2]+C[3]*X[,3]
```

```
    I[a,b]=sum(X[,a]*X[,b]*exp(2*w)/(1+exp(w))^3)+
```

```
      sum(X[,a]*X[,b]*exp(-2*w)/(1+exp(-w))^3);
```

```
  }
```

```
}
```

```
II<- solve(I)
```

```
ORT<-exp(C[2])
```

```
LORT<-exp(C[2]-II[2,2]^0.5*qnorm(0.975,0,1))
```

```
RORT<-exp(C[2]+II[2,2]^0.5*qnorm(0.975,0,1))
```

```
ORG<- exp(C[3])
```

```
LORG<-exp(C[3]-II[3,3]^0.5*qnorm(0.975,0,1))
```

```
RORG<-exp(C[3]+II[3,3]^0.5*qnorm(0.975,0,1))
```

```
ORT
```

```
LORT
```

```
RORT
```

```
ORG
```

```
LORG
```

```
RORG
```

```
ROC_DT <- roc(D,P,ci=TRUE,plot=TRUE,
```

```
  auc.polygon=TRUE, max.auc.polygon=TRUE,
```

```
  grid=TRUE, print.auc=TRUE)
```

```
coords(ROC_DT, "best", ret="threshold",
```

```
transpose = FALSE, best.method="youden")
```

Gotowiec dla logitu:

```
model <- glm(D ~ T+G,
```

```
data = Dane, family = binomial)
```

```
S<-summary(model)$coef
```

S

```
exp(S[2,1])
```

```
exp(S[2,1]-S[2,2]*qnorm(0.975,0,1))
```

```
exp(S[2,1]+S[2,2]*qnorm(0.975,0,1))
```

```
exp(S[3,1])
```

```
exp(S[3,1]-S[3,2]*qnorm(0.975,0,1))
```

```
exp(S[3,1]+S[3,2]*qnorm(0.975,0,1))
```