

```
#####  
# Jiri Beran Multivariate Statistics Exercise 9: Discriminant analysis  
#####
```

```
library(rrcov)  
library(classifly)  
data(hemophilia)
```

```
#####  
# 1.a)  
#####
```

```
#Discriminant functions  
LinCla<-LdaClassic(gr~,data=hemophilia)  
LinRob<-Linda(gr~,data=hemophilia)  
QdaCla<-QdaClassic(gr~,data=hemophilia)  
QdaRob<-QdaCov(gr~,data=hemophilia)
```

```
#Misclassification  
misclassLinCla<-1-  
sum(diag(table(predict(LinCla)@classification, hemophilia$gr)))/nrow(hemophilia)  
misclassLinRob<-1-sum(diag(table(predict(LinRob)@classification, hemophilia$gr))  
/nrow(hemophilia)  
misclassQdaCla<-1-sum(diag(table(predict(QdaCla)@classification, hemophilia$gr))  
/nrow(hemophilia)  
misclassQdaRob<-1-sum(diag(table(predict(QdaRob)@classification, hemophilia$gr))  
/nrow(hemophilia)
```

```
misclassLinCla  
misclassLinRob  
misclassQdaCla  
misclassQdaRob
```

```
#####  
# 1.b)  
#####
```

```
#vectors for boxplots  
vectorLinCla<-c();vectorLinRob<-c();vectorQdaCla<-c();vectorQdaRob<-c();
```

```
## four loop
```

```
for(i in 1:100){  
  train<-sample(1:nrow(hemophilia),round(1/2*nrow(hemophilia)))
```

```
  traindata<-hemophilia[train,]  
  testdata<-hemophilia[-train,]
```

```
  LinCla2<-LdaClassic(gr~,data=traindata)  
  LinRob2<-Linda(gr~,data=traindata)  
  QdaCla2<-QdaClassic(gr~,data=traindata)  
  QdaRob2<-QdaCov(gr~,data=traindata)
```

```
  missclassLinCla2<-1-  
sum(diag(table(predict(LinCla2,newdata=testdata[,1:2])@classification, hemophilia[-train,3]
```

```
))/nrow(hemophilia)-length(train))
  missclassLinRob2<-1-
sum(diag(table(predict(LinRob2,newdata=testdata[,1:2])@classification,hemophilia[-train,3]
))/nrow(hemophilia)-length(train))
  missclassQdaCla2<-1-
sum(diag(table(predict(QdaCla2,newdata=testdata[,1:2])@classification,hemophilia[-train,3]
))/nrow(hemophilia)-length(train))
  missclassQdaRob2<-1-
sum(diag(table(predict(QdaRob2,newdata=testdata[,1:2])@classification,hemophilia[-train,3]
))/nrow(hemophilia)-length(train))
```

```
vectorLinCla[i]<-missclassLinCla2
vectorLinRob[i]<-missclassLinRob2
vectorQdaCla[i]<-missclassQdaCla2
vectorQdaRob[i]<-missclassQdaRob2}
```

```
#Summary
```

```
summary(vectorLinCla)
summary(vectorLinRob)
summary(vectorQdaCla)
summary(vectorQdaRob)
```

```
#Boxplots
```

```
lmts <- range(vectorLinCla,vectorLinRob,vectorQdaCla,vectorQdaRob)
par(mfrow=c(1,4))
boxplot(vectorLinCla,main="LinCla",ylim=lmts)
boxplot(vectorLinRob,main="LinRob",ylim=lmts)
boxplot(vectorQdaCla,main="QdaCla",ylim=lmts)
boxplot(vectorQdaRob,main="QdaRob",ylim=lmts)
```

```
#####
```

```
# 2.a)
```

```
attach(olives)
```

```
mydata<-olives[,1:9]
mydata<-mydata[,-2]
LinCla2<-LdaClassic(Region~.,data=mydata)
LinRob2<-Linda(Region~.,data=mydata)
QdaCla2<-QdaClassic(Region~.,data=mydata)
QdaRob2<-QdaCov(Region~.,data=mydata)
```

```
misclassLinCla2<-1-
sum(diag(table(predict(LinCla2)@classification,olives$Region)))/nrow(olives)
misclassLinRob2<-1-sum(diag(table(predict(LinRob2)@classification,olives$Region))
)/nrow(olives)
misclassQdaCla2<-1-sum(diag(table(predict(QdaCla2)@classification,olives$Region))
)/nrow(olives)
misclassQdaRob2<-1-sum(diag(table(predict(QdaRob2)@classification,olives$Region))
)/nrow(olives)
```

```
misclassLinCla2
misclassLinRob2
misclassQdaCla2
misclassQdaRob2
```

```
#####  
# 2.b)  
#####  
  
## foor loop  
vectorLinCla2<-c();vectorLinRob2<-c();vectorQdaCla2<-c();vectorQdaRob2<-c();  
for(i in 1:100){  
  train<-sample(1:nrow(olives),round(1/2*nrow(olives)))  
  traindata<-mydata[train,]  
  testdata<-mydata[-train,]  
  
  LinCla2<-LdaClassic(Region~.,data=traindata)  
  LinRob2<-Linda(Region~.,data=traindata)  
  QdaCla2<-QdaClassic(Region~.,data=traindata)  
  QdaRob2<-QdaCov(Region~.,data=traindata)  
  
  misclassLinCla2<-1-  
  sum(diag(table(predict(LinCla2,newdata=testdata[,2:8])@classification,testdata$Region)))/(  
  nrow(olives)-length(train))  
  misclassLinRob2<-1-  
  sum(diag(table(predict(LinRob2,newdata=testdata[,2:8])@classification,testdata$Region)))/(  
  nrow(olives)-length(train))  
  misclassQdaCla2<-1-  
  sum(diag(table(predict(QdaCla2,newdata=testdata[,2:8])@classification,testdata$Region)))/(  
  nrow(olives)-length(train))  
  misclassQdaRob2<-1-  
  sum(diag(table(predict(QdaRob2,newdata=testdata[,2:8])@classification,testdata$Region)))/(  
  nrow(olives)-length(train))  
  
  vectorLinCla2[i]<-misclassLinCla2  
  vectorLinRob2[i]<-misclassLinRob2  
  vectorQdaCla2[i]<-misclassQdaCla2  
  vectorQdaRob2[i]<-misclassQdaRob2  
}  
  
lmts1 <- range(vectorLinCla2,vectorLinRob2,vectorQdaCla2,vectorQdaRob2)  
  
par(mfrow=c(1,4))  
boxplot(vectorLinCla2,main="LinCla",ylim=lmts1)  
boxplot(vectorLinRob2,main="LinRob",ylim=lmts1)  
boxplot(vectorQdaCla2,main="QdaCla",ylim=lmts1)  
boxplot(vectorQdaRob2,main="QDARob",ylim=lmts1)  
  
#Comparison boxplots of misclassification without and with test data  
par(mfrow=c(2,4))  
  
boxplot(vectorLinCla,main="LinCla",ylim=lmts)  
boxplot(vectorLinRob,main="LinRob",ylim=lmts)  
boxplot(vectorQdaCla,main="QdaCla",ylim=lmts)  
boxplot(vectorQdaRob,main="QDARob",ylim=lmts)  
  
boxplot(vectorLinCla2,main="LinCla",ylim=lmts1)  
boxplot(vectorLinRob2,main="LinRob",ylim=lmts1)  
boxplot(vectorQdaCla2,main="QdaCla",ylim=lmts1)  
boxplot(vectorQdaRob2,main="QDARob",ylim=lmts1)
```