

#### OUTPUT FOR QUESTION 4

```
donner=read.csv("C:/Users/Thuntida/OneDrive - University of
Calgary/dataset603/donner.txt", sep="")
head(donner,4)

##   age sex survivor
## 1  23   1         0
## 2  40   0         1
## 3  40   1         1
## 4  30   1         0

mylogit = glm(survivor~age+sex, data = donner, family = "binomial")
summary(mylogit)

##
## Call:
## glm(formula = survivor ~ age + sex, family = "binomial", data = donner)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.23041    1.38686   2.329   0.0198 *
## age         -0.07820    0.03728  -2.097   0.0359 *
## sex         -1.59729    0.75547  -2.114   0.0345 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 61.827  on 44  degrees of freedom
## Residual deviance: 51.256  on 42  degrees of freedom
## AIC: 57.256
##
## Number of Fisher Scoring iterations: 4
```

```

myinterlogit = glm(survivor~age+sex+age*sex, data = donner, family =
"binomial")
summary(myinterlogit)

##
## Call:
## glm(formula = survivor ~ age + sex + age * sex, family = "binomial",
##      data = donner)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  7.24638    3.20517   2.261   0.0238 *
## age         -0.19407    0.08742  -2.220   0.0264 *
## sex         -6.92805    3.39887  -2.038   0.0415 *
## age:sex      0.16160    0.09426   1.714   0.0312 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 61.827  on 44  degrees of freedom
## Residual deviance: 47.346  on 41  degrees of freedom
## AIC: 55.346
##
## Number of Fisher Scoring iterations: 5

```

```
anova(mylogit,myinterlogit,test="Chisq")
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: survivor ~ age + sex
```

```
## Model 2: survivor ~ age + sex + age * sex
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
## 1         42      51.256
```

```
## 2         41      47.346  1   3.9099    0.048 *
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lrtest(mylogit,myinterlogit)
```

```
## Likelihood ratio test
```

```
##
```

```
## Model 1: survivor ~ age + sex
```

```
## Model 2: survivor ~ age + sex + age * sex
```

```
##   #Df  LogLik Df  Chisq Pr(>Chisq)
```

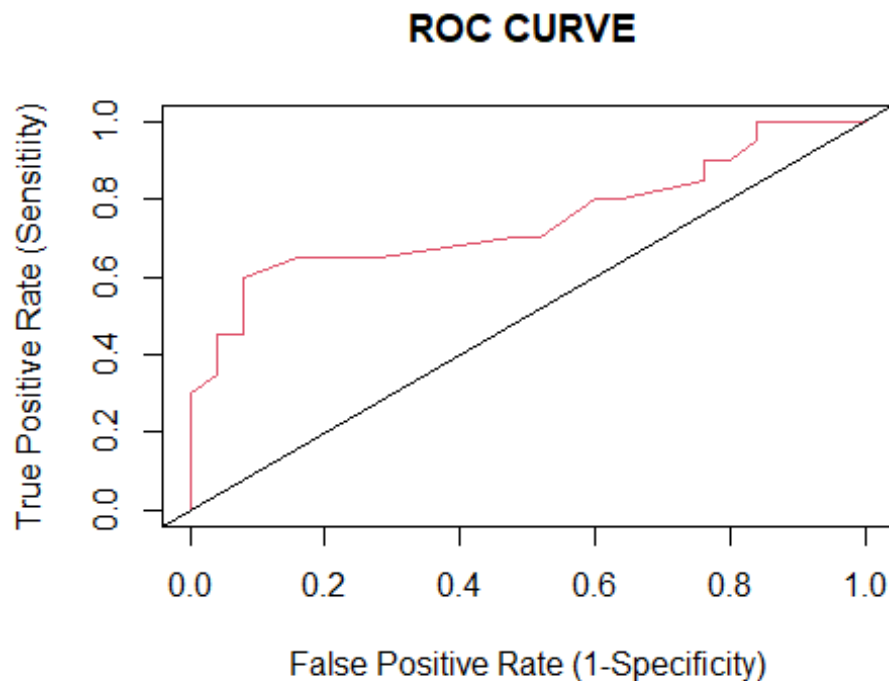
```
## 1    3 -25.628
```

```
## 2    4 -23.673  1 3.9099    0.048 *
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#-----ROC Curve for mylogit model-----
library(ROCR)
prob=predict(mylogit,type=c("response"))
pred<-prediction(prob,donner$survivor)
perf<-performance(pred,measure = "tpr",x.measure="fpr")
plot(perf,col=2,main="ROC CURVE ", xlab="False Positive Rate (1-
Specificity)",ylab="True Positive Rate (Sensitiity)")
abline(0,1)
```



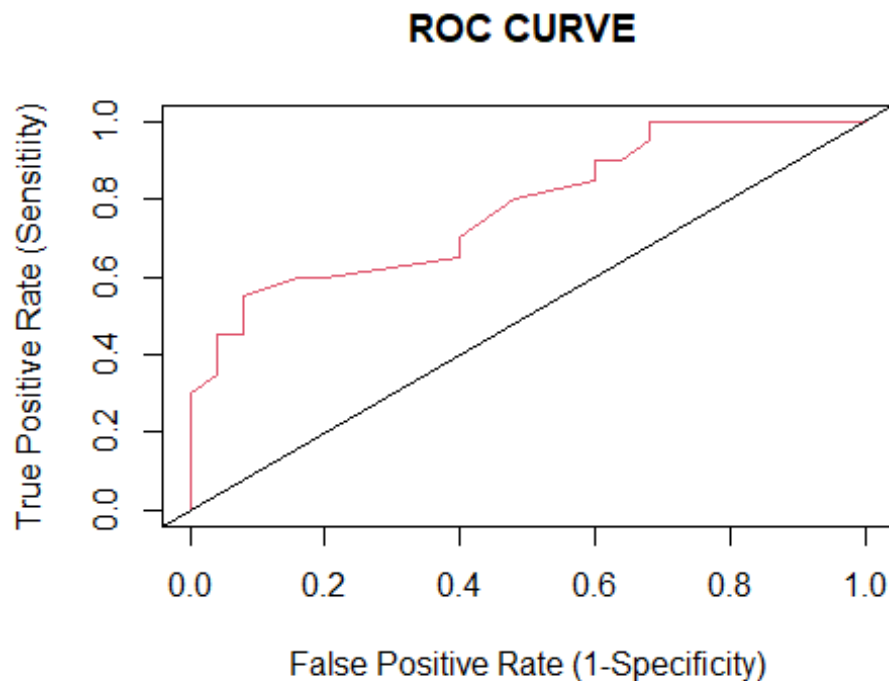
```
#-----AUC-----
library(pROC)
roc<-roc(donner$survivor,prob)

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

auc(roc)

## Area under the curve: 0.746
```

```
#-----ROC Curve for myinterLogit model-----
library(ROCR)
prob=predict(myinterlogit,type=c("response"))
pred<-prediction(prob,donner$survivor)
perf<-performance(pred,measure = "tpr",x.measure="fpr")
plot(perf,col=2,main="ROC CURVE ", xlab="False Positive Rate (1-
Specificity)",ylab="True Positive Rate (Sensitiity)")
abline(0,1)
```



```
#-----AUC-----
library(pROC)
roc<-roc(donner$survivor,prob)

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

auc(roc)

## Area under the curve: 0.778
```