**CSC 423/324 -- Project 4**

1. **Input and print the [HeartAttack](http://facweb.cdm.depaul.edu/sjost/csc423/projects/heart-attack.txt) dataset.**

 > data <- read.table("c:/DataSets/heartAttack.txt", header = T)

> print(data)

ha2 ang sco

1 1 1 70

2 1 1 80

3 1 1 50

4 1 0 60

5 1 0 40

6 1 0 65

7 1 0 75

8 1 0 80

9 1 0 70

10 1 0 60

11 0 1 65

12 0 1 50

13 0 1 45

14 0 1 35

15 0 1 40

16 0 1 50

17 0 0 55

18 0 0 45

19 0 0 50

20 0 0 60

> ha2 <- data$ha2

> ang <- data$ang

> sco <- data$sco

1. **Type the logit model equation for predicting the probability that a patient has a second attack (ha2) from ang and sco. (See the dataset for the meanings of these variables.)**

> #Logistic Regression Equation

> logitEq <- glm(ha2 ~ ang + sco, data = data, family = "binomial")

> summary(logitEq)

Call:

glm(formula = ha2 ~ ang + sco, family = "binomial", data = data)

Deviance Residuals:

Min 1Q Median 3Q Max

-1.52106 -0.68746 0.00424 0.70625 1.88960

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -6.36347 3.21362 -1.980 0.0477 \*

ang -1.02411 1.17101 -0.875 0.3818

sco 0.11904 0.05497 2.165 0.0304 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 27.726 on 19 degrees of freedom

Residual deviance: 18.820 on 17 degrees of freedom

AIC: 24.82

Number of Fisher Scoring iterations: 4

1. **Find the predicted probability of having a second heart attack for each observation.**

> #New Data with just independent variables

> newdata <- data.frame(ang, sco)

> print(newdata)

ang sco

1 1 70

2 1 80

3 1 50

4 0 60

5 0 40

6 0 65

7 0 75

8 0 80

9 0 70

10 0 60

11 1 65

12 1 50

13 1 45

14 1 35

15 1 40

16 1 50

17 0 55

18 0 45

19 0 50

20 0 60

>

> #Predicted Probabilities

> predictedProbs <- predict(logitEq, newdata, type="response")

> print(predictedProbs)

1 2 3 4 5 6

0.72022303 0.89435490 0.19226955 0.68551307 0.16774789 0.79809811

7 8 9 10 11 12

0.92856807 0.95930576 0.87757611 0.68551307 0.58670090 0.19226955

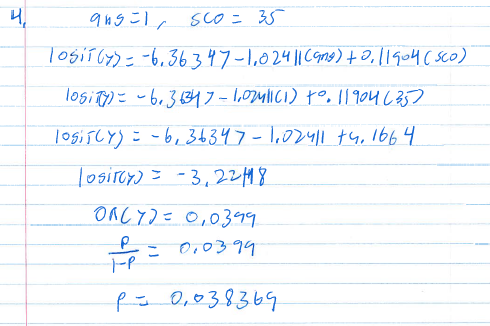
13 14 15 16 17 18

0.11603202 0.03838292 0.06749763 0.19226955 0.54587077 0.26767543

19 20

0.39861864 0.68551307

1. **If a new patient has ang=1 and sco=35, calculate the probability of a second heart attack, both by hand and using SAS or R.**



> #New observation and predicted probability of new observation

> newdata2 <- data.frame(ang = 1, sco = 35)

> predictedProbs2 <- predict(logitEq, newdata2, type="response")

> print(predictedProbs2)

1

0.03838292