Chapter 10 – Logistic Regression

Instructor: Zach Zhizhong ZHOU,
Shanghai Jiao Tong University
主讲教师：周志中，上海交通大学

Data Mining for Business Intelligence

Shmueli, Patel & Bruce

DeathPenalty - 1

```
setwd("C:/BA/Logit")

dpen <- read.csv("DeathPenalty.csv",header = TRUE)

dpen[1:4,]

#Agg表示罪行严重程度，越严重数值越大。VRace表示种族，1表示受害人是白人，0表示受害人是黑人。Death表示罪犯是否被判死刑，1代表被判死刑，0代表未被判死刑。

m1=glm(Death~VRace+Agg,family=binomial,data=dpen)

#使用Logit回归，GLM表示广义线性回归，binomial表示二项分布。

m1

summary(m1)
```

DeathPenalty - 2

```
## calculating logits

exp(m1$coef[2])

exp(m1$coef[3])

## plotting probability of getting death penalty as a
function of

## aggravation separately for black (in black) and white (in
red)

## victim

fitBlack=dim(501)

fitWhite=dim(501)

ag=dim(501)
```

DeathPenalty - 3

- for (i in 1:501) { ag[i]=(99+i)/100
fitBlack[i]=exp(m1$coef[1]+ag[i]*m1$coef[3])/(1+exp(m1$coef[1]+ag[i]*m1$coef[3]))

- fitWhite[i]=exp(m1$coef[1]+m1$coef[2]+ag[i]*m1$coef[3])/(1+exp(m1$coef[1]+m1$coef[2]+ag[i]*m1$coef[3]))}

- plot(fitBlack~ag,type="l",col="black",ylab="Prob[Death]",

- xlab="Aggravation",ylim=c(0,1),

- main="red line for white victim; black line for black victim")

- points(fitWhite~ag,type="l",col="red")