# HW20

# Wenhui Yang

# 5/5/2019

setwd("/Users/ninenins/Desktop/R/数据集")

college <- read.csv("college(1).csv")

college.1 <- college[,c(-1,-7,-8,-9,-10,-11,-12,-14,-15)]

college.1$dummy\_pc <- as.numeric(college.1$perc.alumni > mean(college.1$perc.alumni))

logistic <- glm((Private=="Yes") ~ S.F.Ratio + perc.alumni+college.1$dummy\_pc,

data=college.1,family="binomial")

summary(logistic)

coef(logistic)

p.hat <- fitted(logistic)

y.hat <- round(p.hat)

table(y.hat, y.true=college.1$Private)

outcome:

Call:

glm(formula = (Private == "Yes") ~ S.F.Ratio + perc.alumni +

college.1$dummy\_pc, family = "binomial", data = college.1)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.5751 -0.5807 0.3249 0.6514 3.5789

Coefficients:

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

(Intercept) 3.69475 0.57308 6.447 1.14e-10 \*\*\*

S.F.Ratio -0.27803 0.03117 -8.920 < 2e-16 \*\*\*

perc.alumni 0.06456 0.01868 3.457 0.000547 \*\*\*

college.1$dummy\_pc 0.57343 0.38409 1.493 0.135454

---

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 910.75 on 776 degrees of freedom

Residual deviance: 642.65 on 773 degrees of freedom

AIC: 650.65

Number of Fisher Scoring iterations: 5

> coef(logistic)

(Intercept) S.F.Ratio perc.alumni

3.69474646 -0.27803352 0.06456032

college.1$dummy\_pc

0.57342818

> p.hat <- fitted(logistic)

> y.hat <- round(p.hat)

> table(y.hat, y.true=college.1$Private)

y.true

y.hat No Yes

0 119 45

1 93 520