Q6 – Analysis of binary

(a) calculate y\_i/n\_i’s and compare their averages for seed and Extract

> data = read.csv('extract.csv',encoding = 'latin1')

> colnames(data)=c('seed', 'extract', 'n' ,'yes')

> data['no'] = data$n-data$yes

> data['p'] = data$yes/data$n

> ##SeedA

> mean(data[data$seed==0,]$p)

[1] 0.559039

> ##SeedB

> mean(data[data$seed==1,]$p)

[1] 0.3972767

> ##Extract1

> mean(data[data$extract==0,]$p)

[1] 0.3487189

> ##Extract2

> mean(data[data$extract==1,]$p)

[1] 0.6031824

> ##SeedA & Extract1

> mean(data[data$seed==0 & data$extract==0,]$p)

[1] 0.371406

> ##SeedA & Extract2

> mean(data[data$seed==0 & data$extract==1,]$p)

[1] 0.7153998

> ##SeedB & Extract1

> mean(data[data$seed==1 & data$extract==0,]$p)

[1] 0.3260317

> ##SeedB & Extract2

> mean(data[data$seed==1 & data$extract==1,]$p)

[1] 0.4685216

(b) 2-way ANOVA

1) 2-way ANOVA without interaction

> summary(aov(p ~ seed + extract,data=data))

Df Sum Sq Mean Sq F value Pr(>F)

seed 1 0.1371 0.1371 6.254 0.02227 \*

extract 1 0.3205 0.3205 14.625 0.00124 \*\*

Residuals 18 0.3945 0.0219

---

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

1) 2-way ANOVA with interaction

> summary(aov(p ~ seed + extract + seed\*extract,data=data))

Df Sum Sq Mean Sq F value Pr(>F)

seed 1 0.1371 0.1371 6.823 0.018221 \*

extract 1 0.3205 0.3205 15.954 0.000939 \*\*\*

seed:extract 1 0.0530 0.0530 2.636 0.122845

Residuals 17 0.3415 0.0201

---

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

(c) logistic regression

1) without interaction

> summary(glm(cbind(yes,no)~seed+extract,family = binomial,data =data))

Call:

glm(formula = cbind(yes, no) ~ seed + extract, family = binomial,

data = data)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.3919 -0.9949 -0.3744 0.9831 2.4766

Coefficients:

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

(Intercept) -0.4300 0.1137 -3.781 0.000156 \*\*\*

seed -0.2705 0.1547 -1.748 0.080435 .

extract 1.0647 0.1442 7.383 1.55e-13 \*\*\*

---

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 98.719 on 20 degrees of freedom

Residual deviance: 39.686 on 18 degrees of freedom

AIC: 122.28

Number of Fisher Scoring iterations: 4

2) without interaction

> summary(glm(cbind(yes,no)~seed+extract +seed\*extract,family = binomial,data =data))

Call:

glm(formula = cbind(yes, no) ~ seed + extract + seed \* extract,

family = binomial, data = data)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.01617 -1.24398 0.05995 0.84695 2.12123

Coefficients:

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

(Intercept) -0.5582 0.1260 -4.429 9.46e-06 \*\*\*

seed 0.1459 0.2232 0.654 0.5132

extract 1.3182 0.1775 7.428 1.10e-13 \*\*\*

seed:extract -0.7781 0.3064 -2.539 0.0111 \*

---

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 98.719 on 20 degrees of freedom

Residual deviance: 33.278 on 17 degrees of freedom

AIC: 117.87

Number of Fisher Scoring iterations: 4