Se les preguntó si estaban o no de acuerdo con la frase:

"women should take care of running their homes and leave running the country up to men"

Datos

> frase

yearseducation genero conteos totales

1 0 males 4 6

2 1 males 2 3

3 2 males 4 5

4 3 males 6 9

5 4 males 5 10

6 5 males 13 20

7 6 males 25 34

8 7 males 27 42

9 8 males 75 124

10 9 males 29 58

11 10 males 32 77

12 11 males 36 95

13 12 males 115 360

14 13 males 31 101

15 14 males 28 107

16 15 males 9 32

17 16 males 15 125

18 17 males 3 32

19 18 males 1 29

20 19 males 2 15

21 20 males 3 23

22 0 female 4 6

23 1 female 1 2

24 2 female 1 2

25 3 female 6 7

26 4 female 10 11

27 5 female 14 21

28 6 female 17 22

29 7 female 26 42

30 8 female 21 45

31 9 female 30 65

32 10 female 55 122

33 11 female 50 112

34 12 female 190 593

35 13 female 17 109

36 14 female 18 99

37 15 female 7 41

38 16 female 13 128

39 17 female 3 31

40 18 female 1 21

41 19 female 1 3

42 20 female 2 6

MODELO CON 1 VARIABLE EXPLICATIVA

> mod1<-glm(cbind(conteos,totales-conteos)~yearseducation,family=binomial,data= frase)

> summary(mod1)

Call:

glm(formula = cbind(conteos, totales - conteos) ~ yearseducation,

family = binomial, data = frase)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.7542 -0.9588 -0.2149 0.8569 1.9917

Coefficients:

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

(Intercept) 2.13233 0.17658 12.08 <2e-16 \*\*\*

yearseducation -0.24199 0.01511 -16.02 <2e-16 \*\*\*

---

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

(Dispersion parameter for binomial family taken to be 1)

Baja bastante el valor de la Deviance

Null deviance: 361.456 on 41 degrees of freedom

Residual deviance: 49.084 on 40 degrees of freedom

AIC: 200.35

Number of Fisher Scoring iterations: 4

MODELO CON 1 VARIABLE EXPLICATIVA

> mod11<-glm(cbind(conteos,totales-conteos)~genero,family=binomial,data= frase)

> summary(mod11)

Call:

glm(formula = cbind(conteos, totales - conteos) ~ genero, family = binomial,

data = frase)

Deviance Residuals:

Min 1Q Median 3Q Max

No es significativa

-6.0102 -2.0486 0.7181 2.5276 5.6304

Coefficients:

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

(Intercept) -0.72049 0.05525 -13.041 <2e-16 \*\*\*

generomales 0.12675 0.07994 1.586 0.113

---

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

(Dispersion parameter for binomial family taken to be 1)

No baja mucho la Deviance

Null deviance: 361.46 on 41 degrees of freedom

Residual deviance: 358.94 on 40 degrees of freedom

AIC: 510.21

Number of Fisher Scoring iterations: 4

MODELO CON 2 VARIABLES EXPLICATIVAS

> mod2<-glm(cbind(conteos,totales-conteos)~genero+yearseducation,family=binomial,data= frase)

> summary(mod2)

Call:

glm(formula = cbind(conteos, totales - conteos) ~ genero + yearseducation,

family = binomial, data = frase)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.5751 -0.9477 -0.2613 0.7238 2.0209

Coefficients:

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

(Intercept) 2.08139 0.18243 11.409 <2e-16 \*\*\*

generomales 0.09141 0.08485 1.077 0.281

yearseducation -0.24124 0.01510 -15.976 <2e-16 \*\*\*

---

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

(Dispersion parameter for binomial family taken to be 1)

La Deviance es poco diferente a 49.084

Null deviance: 361.456 on 41 degrees of freedom

Residual deviance: 47.924 on 39 degrees of freedom

AIC: 201.2

Number of Fisher Scoring iterations: 4

MODELO CON INTERACCIÓN

> mod3<-glm(cbind(conteos,totales-conteos)~yearseducation\*genero,family=binomial,data= frase)

>

> summary(mod3)

Call:

glm(formula = cbind(conteos, totales - conteos) ~ yearseducation \*

genero, family = binomial, data = frase)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.4552 -0.8600 -0.1731 0.5884 2.1430

Coefficients:

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

(Intercept) 2.28618 0.27107 8.434 <2e-16 \*\*\*

yearseducation -0.25913 0.02313 -11.203 <2e-16 \*\*\*

generomales -0.26914 0.35693 -0.754 0.451

yearseducation:generomales 0.03174 0.03051 1.040 0.298

---

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

(Dispersion parameter for binomial family taken to be 1)

Bajo otro poquito la Deviance

Null deviance: 361.456 on 41 degrees of freedom

Residual deviance: 46.839 on 38 degrees of freedom

AIC: 202.11

Number of Fisher Scoring iterations: 4

SELECCIÓN DE VARIABLES

>

> anova(mod1 ,mod2,mod3)

Analysis of Deviance Table

Model 1: cbind(conteos, totales - conteos) ~ yearseducation

Model 2: cbind(conteos, totales - conteos) ~ genero + yearseducation

Model 3: cbind(conteos, totales - conteos) ~ yearseducation \* genero

Resid. Df Resid. Dev Df Deviance

Diferencias de las Deviance, modelos anidados.

1 40 49.084

2 39 47.924 1 1.1598

3 38 46.839 1 1.0859

> anova(mod1,mod3)

Analysis of Deviance Table

Model 1: cbind(conteos, totales - conteos) ~ yearseducation

Model 2: cbind(conteos, totales - conteos) ~ yearseducation \* genero

Resid. Df Resid. Dev Df Deviance

Diferencias modelo chico vs el más grande

1 40 49.084

2 38 46.839 2 2.2458

|  |
| --- |
| > 1-pchisq(2.24,2)  [1] 0.3262798  > 1-pchisq(1.15,1)  Estos son los valores de los p-values para esas diferencias de Deviance, ninguna sale significativa, es decir explican lo mismo el modelo grande y el chico, entonces se prefiere el modelo chico  [1] 0.2835491  > 1-pchisq(1.08,1)  [1] 0.2986976 |
|  |
| |  | | --- | |  | |

