

## ejemplo glm

david

### Ejemplo GLM

Este archivo es un ejemplo de análisis de datos con un modelo lineal generalizado.

- Primero, copiamos los datos a la misma carpeta donde guardemos el archivo .Rmd
- Ahora, leemos los datos

```
tit <- read.csv("titanic_long.csv")
head(tit)
```

```
##   class    age   sex survived
## 1 first adult male      1
## 2 first adult male      1
## 3 first adult male      1
## 4 first adult male      1
## 5 first adult male      1
## 6 first adult male      1
```

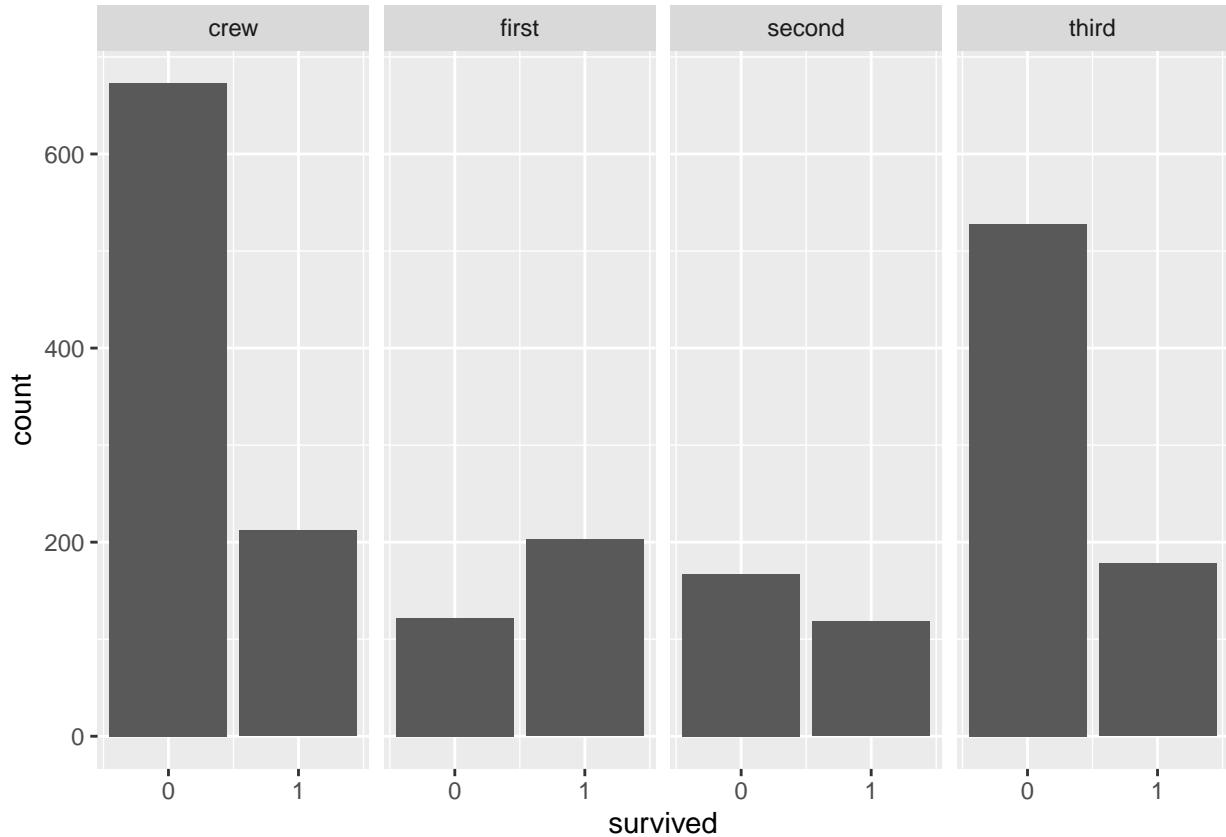
- Visualizamos los datos

```
table(tit$class)
```

```
##
##    crew   first  second  third
##    885     325     285    706
```

```
table(tit$survived)
```

```
##
##      0      1
## 1490  711
```



- ajustar el modelo

```
m1 <- glm(survived ~ class,
           data = tit,
           family = "binomial")

summary(m1)

##
## Call:
## glm(formula = survived ~ class, family = "binomial", data = tit)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q       Max
## -1.3999   -0.7623   -0.7401    0.9702    1.6906
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.15516   0.07876 -14.667 < 2e-16 ***
## classfirst   1.66434   0.13902  11.972 < 2e-16 ***
## classsecond  0.80785   0.14375   5.620 1.91e-08 ***
## classthird   0.06785   0.11711   0.579    0.562
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```

## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 2769.5 on 2200 degrees of freedom
## Residual deviance: 2588.6 on 2197 degrees of freedom
## AIC: 2596.6
##
## Number of Fisher Scoring iterations: 4

```

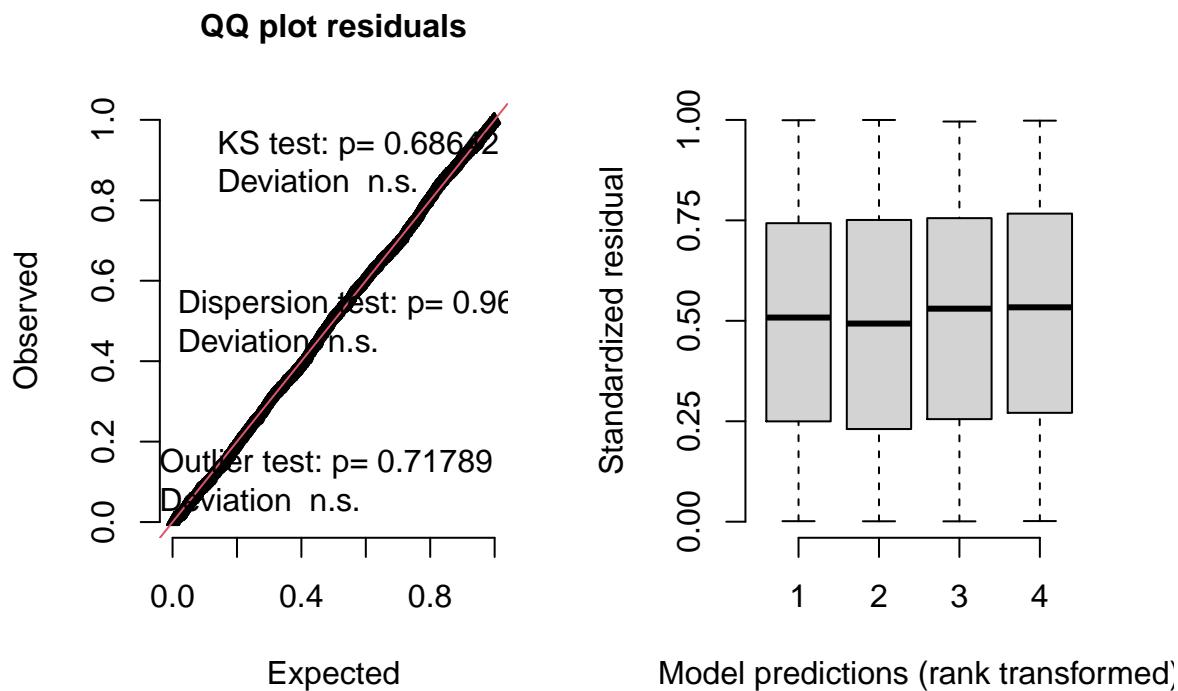
- comprobar los residuos

```

library(DHARMa)
simulateResiduals(m1, plot = TRUE)

```

## DHARMa residual diagnostics



```

## Object of Class DHARMa with simulated residuals based on 250 simulations with refit = FALSE . See ?DHARMa
##
## Scaled residual values: 0.8221129 0.8930727 0.5467574 0.4289614 0.358382 0.5792006 0.6313679 0.37602

```

- el modelo es razonable, así que interpretamos los coeficientes

```

library(effects)
allEffects(m1)

```

```

## model: survived ~ class
##

```

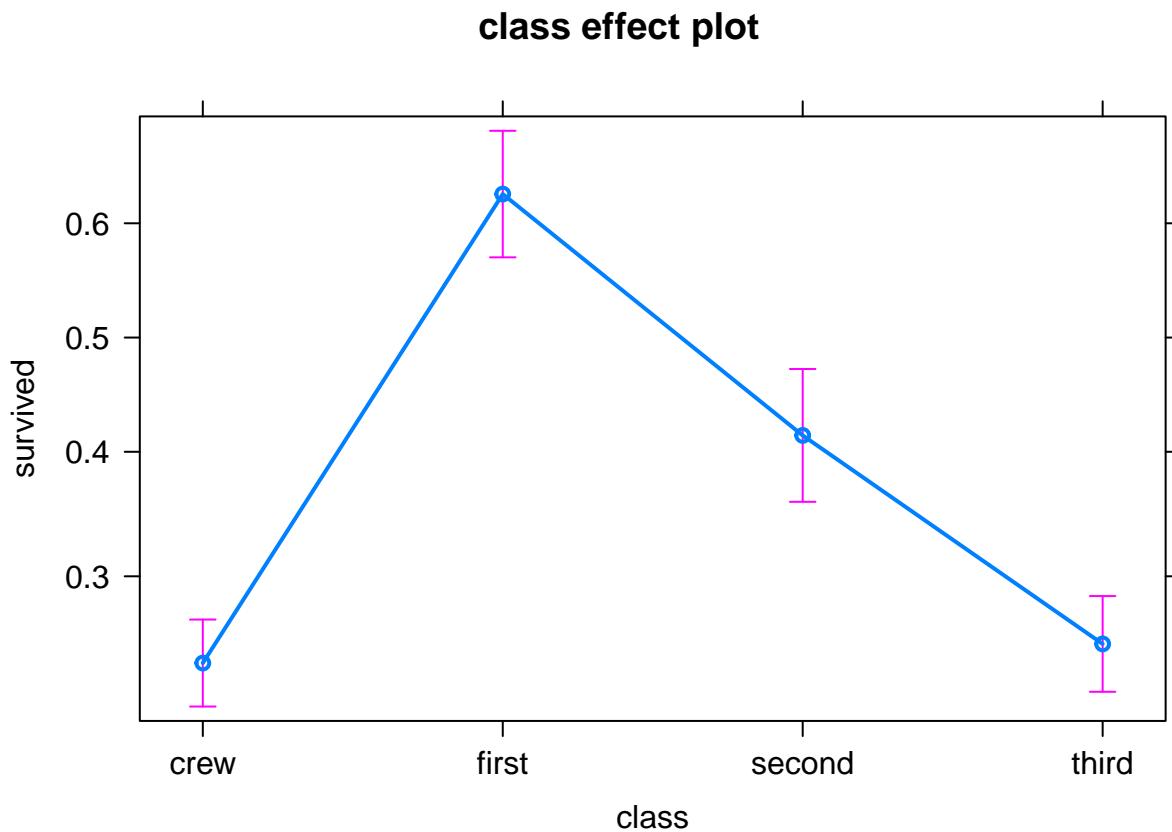
```

##  class effect
##  class
##    crew      first     second     third
## 0.2395480 0.6246154 0.4140351 0.2521246

```

- visualizamos los efectos

```
plot(allEffects(m1))
```



- mostramos una tabla con todos los coeficientes del modelo

```

library(broom)
knitr::kable(tidy(m1))

```

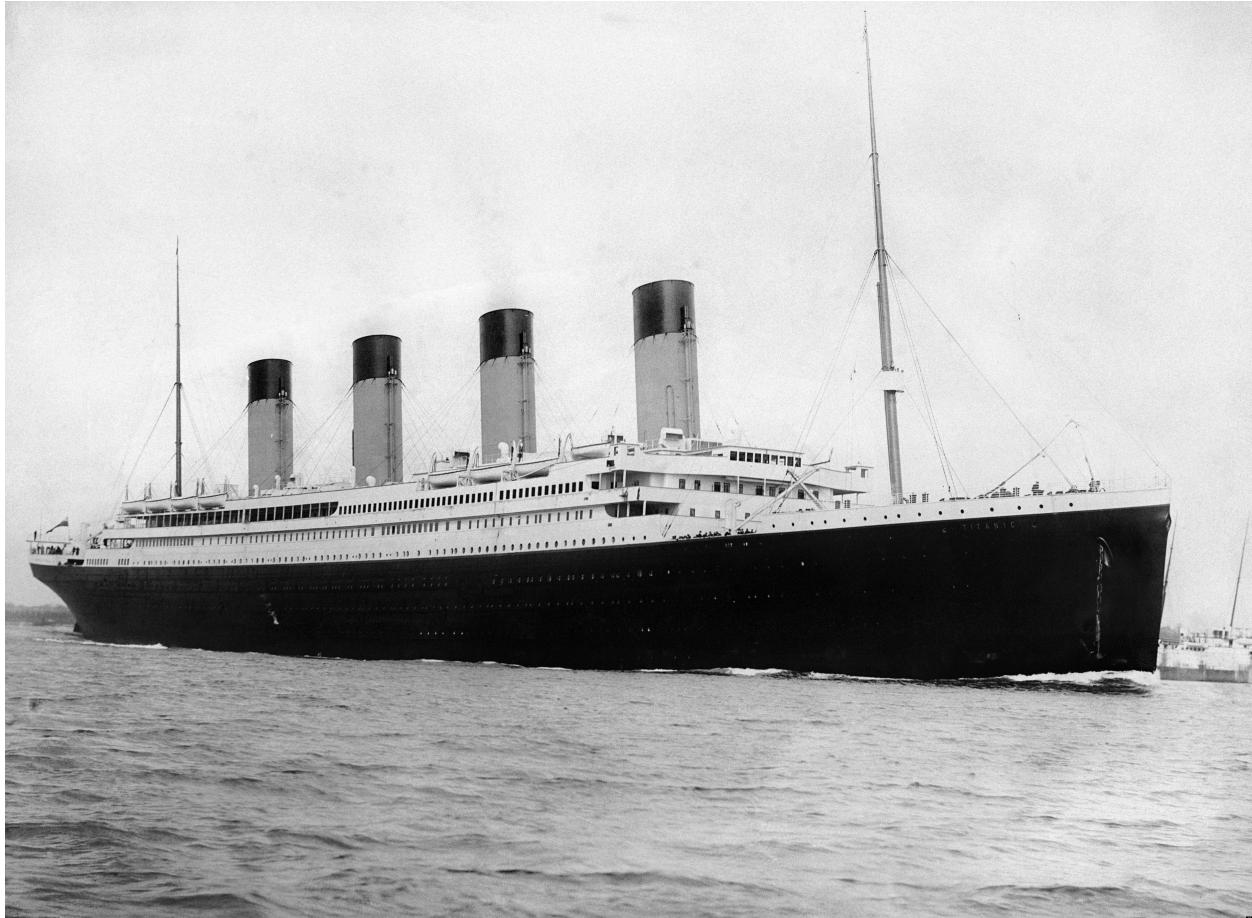
term	estimate	std.error	statistic	p.value
(Intercept)	-1.1551591	0.0787582	-14.6671580	0.0000000
classfirst	1.6643440	0.1390168	11.9722466	0.0000000
classsecond	0.8078499	0.1437548	5.6196358	0.0000000
classthird	0.0678463	0.1171100	0.5793383	0.5623609

Ejecutamos una regresión logística con función de enlace logit, sobre 2201 observaciones, en la que analizamos la probabilidad de supervivencia en función de la clase de los pasajeros.

Los pasajeros de primera clase tienen una probabilidad de supervivencia de 0.67.

Podemos incluir imágenes externas

```
knitr::include_graphics("titanic.jpg")
```



NOTA:

- Para abrir los archivos asociados al trabajo, usad

```
read.csv("nombre-apellido.txt", sep = " ")
```

- otras funciones

```
table(ej$rio)
rio1 <- subset(ej, rio == "Corumbel")
table(rio1$enfermo)
```