```
> rm(list=ls())
>
setwd("/home/alessandra/Documents/Lectures/ScienzePolitiche/MSRS/A
.A. 2017 2018/Lab")
>
> load("Dati/Grants.RData")
> ls()
[1] "grants"
> dim(grants)
[1] 2311    4
> head(grants)
   dropout hsgrade        S W
13       0       74 13812.38 1
17       0       99 14062.11 1
24       0       88 13568.72 1
36       0       91 13392.35 1
40       0       88 13473.67 1
41       0       98 14431.30 1
>
> attach(grants)
>
> ##MODELLO LINEARE NELLE PROBABILITA'
> summary(lm(dropout ~ hsgrade, data=grants))

Call:
lm(formula = dropout ~ hsgrade, data = grants)

Residuals:
    Min      1Q  Median      3Q     Max
-0.4120 -0.3576 -0.3001  0.6182  0.7089

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.5934366  0.0618823   9.590  < 2e-16 ***
hsgrade     -0.0030237  0.0007445  -4.062 5.04e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.474 on 2309 degrees of freedom
Multiple R-squared:  0.007094,  Adjusted R-squared:  0.006664
F-statistic:  16.5 on 1 and 2309 DF,  p-value: 5.036e-05

> ##MODELLO LOGIT
> m <- glm(dropout ~ hsgrade, family=binomial(link="logit"),
data=grants)
> #Dev = -2 log(l)
>
> summary(m)

Call:
```

```
glm(formula = dropout ~ hsgrade, family = binomial(link =
"logit"),
    data = grants)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0333  -0.9392  -0.8454   1.3881   1.5688

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.45588    0.27360   1.666   0.0957 .
hsgrade     -0.01341    0.00332  -4.039 5.36e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2978.8  on 2310  degrees of freedom
Residual deviance: 2962.4  on 2309  degrees of freedom
AIC: 2966.4

Number of Fisher Scoring iterations: 4

> confint.default(m)
                  2.5 %       97.5 %
(Intercept) -0.08035952  0.992118610
hsgrade     -0.01991740 -0.006903324
> exp(confint.default(m, "hsgrade"))
            2.5 %    97.5 %
hsgrade 0.9802796 0.9931204
>
> predict(m, data.frame(hsgrade=c(60, 70, 80, 90, 100)),
type="response")
        1         2         3         4         5
0.4136875 0.3815804 0.3504756 0.3205919 0.2921104
> predict(m, data.frame(hsgrade=c(60, 70, 80, 90, 100)),
type="response", se.fit=T)
$fit
        1         2         3         4         5
0.4136875 0.3815804 0.3504756 0.3205919 0.2921104

$se.fit
         1          2          3          4          5
0.02021728 0.01364881 0.01004617 0.01143087 0.01570076

$residual.scale
[1] 1

>
> m0 = glm(dropout ~ 1, family=binomial, data=grants)
```

```
> anova(m0,m, test="Chisq")
Analysis of Deviance Table

Model 1: dropout ~ 1
Model 2: dropout ~ hsgrade
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1      2310     2978.8
2      2309     2962.4  1    16.41 5.103e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> ##########################################
> #Modello di regressione logistica multipla
> ##########################################
>
> grants$S.1000<- S/1000
> attach(grants)
> m = glm(dropout ~ S.1000 + hsgrade + W, family=binomial, data =
grants)
> summary(m)

Call:
glm(formula = dropout ~ S.1000 + hsgrade + W, family = binomial,
    data = grants)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1519  -0.9395  -0.8480   1.3777   1.6780

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.413280   1.112091  -1.271   0.2038
S.1000       0.134046   0.076920   1.743   0.0814 .
hsgrade     -0.014633   0.003369  -4.343  1.4e-05 ***
W            0.220999   0.090362   2.446   0.0145 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2978.8  on 2310  degrees of freedom
Residual deviance: 2953.7  on 2307  degrees of freedom
AIC: 2961.7

Number of Fisher Scoring iterations: 4
>
> exp(m$coef)
(Intercept)      S.1000     hsgrade           W
  0.2433438   1.1434453   0.9854740   1.2473220
> confint.default(m)
```

```
                   2.5 %       97.5 %
(Intercept) -3.59293749  0.766377501
S.1000      -0.01671480  0.284806652
hsgrade     -0.02123546 -0.008029687
W            0.04389264  0.398105036
>
> #Per avere un'idea piu' precisa della probabilita' di abbandono
e' opportuno fornire una tabella di probabilita'
> X<- expand.grid(S.1000=c(13, 14, 15), hsgrade=c(60,80,100),
W=c(0,1))
> p = predict(m, X, type="response")
> cbind(X, p)
   S.1000 hsgrade W         p
1      13      60 0 0.3661806
2      14      60 0 0.3978117
3      15      60 0 0.4303201
4      13      80 0 0.3012634
5      14      80 0 0.3302084
6      15      80 0 0.3604995
7      13     100 0 0.2434347
8      14     100 0 0.2689622
9      15     100 0 0.2961189
10     13      60 1 0.4188153
11     14      60 1 0.4517525
12     15      60 1 0.4851178
13     13      80 1 0.3497155
14     14      80 1 0.3807788
15     15      80 1 0.4128495
16     13     100 1 0.2863983
17     14     100 1 0.3145580
18     15     100 1 0.3441511
>
> #Testare l'ipotesi nulla simultanea β1 = 0
> #(il parametro di S.1000) e β2 = 0 (il parametro di hsgrade)
> #Test del rapporto di verosimiglianza calcolato:
>
> mr = glm(dropout ~ W, family=binomial, data = grants)
> anova(mr,m, test="Chisq")
Analysis of Deviance Table

Model 1: dropout ~ W
Model 2: dropout ~ S.1000 + hsgrade + W
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1      2309     2975.8
2      2307     2953.7  2   22.016 1.657e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```