Regression Models: Assignment 2

Daniel Alonso

January 11th, 2020

Installing libraries used

```r
packages = c("dplyr","MuMIn","MASS","leaps","glmnet","car")
for (package in packages) {
    install.packages(package)
}
```

Importing libraries

```r
library(dplyr)
library(MuMIn)
library(MASS)
library(leaps)
library(glmnet)
library(car)
```

Exercise 1

1 - Model and parameter interpretation

Y = Binary variable representing whether the customer will buy a car or not *income* = annual family income

Therefore we model the response as:

$Y = \beta_0 + \beta_1 X + \epsilon$

And our values will be:

$Y = -1.98079 + 0.04342X + \epsilon$

The higher the family income, the higher the likeliness that a family will purchase a new car.

```
y = function(x) (-1.98079 + 0.04342 * x)
p = function(y) (exp(y)/(exp(y)+1))
```

2 - 95%-CI for the probability that a family with annual income of 60 thousand dollars will purchase anew car next year.

We calculate the asymptotic $(1-\alpha)\%$ confidence interval:

$\hat{\beta}_j \pm z_{\frac{\alpha}{2}} S.E.(\hat{\beta}_j)$

With our values we get:

```
# CI calculation
l_bound <- 0.04342 - qnorm(0.975) * 0.02011
u_bound <- 0.04342 + qnorm(0.975) * 0.02011
```

$0.004005124 \leq \hat{\beta}_j \leq 0.08283488$

```
p(l_bound)
#> [1] 0.5010013
p(u_bound)
#> [1] 0.5206969
```

$0.5010013 \leq p \leq 0.5206969$

Therefore for 60 thousand dollar annual income households:

```
L <- -1.98079 + 0.004005124 * 60
U <- -1.98079 + 0.08283488 * 60
p(L)
#> [1] 0.1492517
p(U)
#> [1] 0.9520885
```

$0.1492517 \leq p_{60k} \leq 0.9520885$