

Chapter 2 - The Simple Regression Model

Thalles Quinaglia Liduares

11/02/2022

Exercise 2.7

Upload packages

```
library(wooldridge)
```

```
library(dplyr)
```

```
library(lmreg)
```

Upload database

```
data7<-wooldridge::charity
```

```
attach(data7)
```

```
detach(data7)
```

```
View(data7)
```

Use the data in CHARITY.RAW [obtained from Franes and Paap (2001)] to answer the following questions:

(i) What is the average gift in the sample of 4,268 people (in Dutch guilders)? What percentage of people gave no gift?

```
ave_gift<-round(mean(data7$gift),2)
```

```
ave_gift
```

```
## [1] 7.44
```

```
num_with_gift<- data7 %>%
  filter(respond == 1) %>%
  summarise(n=n())

num_tot<- data7 %>%
  filter(respond == 1 | respond == 0) %>%
  summarise(n=n())

share_with_gift<-(num_with_gift/num_tot)*100

share_with_gift
```

```
##           n
## 1 39.99531
```

The average number of gifts is 7.4 and the percentual of person without gift is aproximately 60%.

(ii) What is the average mailings per year? What are the minimum and maximum values?

```
ave_mail<-round(mean(data7$mailsyear),2)

ave_mail
```

```
## [1] 2.05
```

```
summary(data7$mailsyear)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.25    1.75    2.00    2.05    2.50    3.50
```

The minimum value is 0.25, the maximum is 3.50 e mean value is 2.05.

(iii) Estimate the model

$$gift = \beta_0 + \beta_1 mailsyear + \varepsilon$$

by OLS and report the results in the usual way, including the sample size and R-squared.

```
lm1<-lm(data7$gift~data7$mailsyear)

summary(lm1)
```

```
##
## Call:
## lm(formula = data7$gift ~ data7$mailsyear)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.287  -7.976  -5.976   2.687  245.999
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.0141     0.7395   2.724  0.00648 **
## data7$mailsyear  2.6495     0.3431   7.723  1.4e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.96 on 4266 degrees of freedom
## Multiple R-squared:  0.01379,    Adjusted R-squared:  0.01356
## F-statistic: 59.65 on 1 and 4266 DF,  p-value: 1.404e-14
```

The estimated equation is given by

$$gift = 2.01 + 2.64mailsyear + \varepsilon$$

The sample size is 4268 and $R^2 = 0.013$. Hence, only 1.3% of variability in number of gifts is explained by the number of mails per year.