TOR VERGATA
UNIVERSITÀ DEGLI STUDI DI ROMA

Quantitative Methods III - Practice 1
Simple Linear Regression - R Code

Prof. Lorenzo Cavallo: lorenzo.cavallo.480084@uniroma2.eu
Prof. Marianna Brunetti: marianna.brunetti@uniroma2.it

Exercise *Investments in Research and Development* (X, million EUR) and *Number of patents* (Y) are collected for 5 companies. The distribution of X and Y is as follows:

X	Y
5	7
6	8
7	8
9	10
9	9

1. Draw the scatterplot of the distribution: what kind of relationship do you expect?
2. Estimate the parameters of the regression line and explain their meaning.
3. Add the regression line to the graph produced in point (1).
4. Calculate the best goodness-of-fit indices of the estimated line to the observed data and comment on the results.
5. Calculate the correlation coefficient.

```
##### Upload the data for X and Y

X=c(5,6,7,9,9)

Y=c(7,8,8,10,9)

# Put the data in a table with cbind

cbind(X,Y)

length(X)
length(Y)

n=length(X)
n

####### Plot

plot(X,Y)

#beautify the graph

plot(X,Y,pch=1, col="blue",xlim=c(1,10),ylim=c(0,12))

###### statistics

#mean
X_bar=mean(X)
X_bar
Y_bar=mean(Y)
Y_bar

#variance and covariance
var_X=var(X)
var_X

var_Y=var(Y)
var_Y

cov_XY=var(X,Y)
cov_XY

#standard deviation
```

```r
sd(X)
sd(Y)

sqrt(var_X)
sqrt(var_Y)

#deviance
dev_X=var(X)*n
dev_X

dev_Y=var(Y)*n
dev_Y

codev_XY=var(X,Y)*n
codev_XY

####### regression with lm

reg_1 <- lm(Y ~ X)
summary(reg_1)

abline(reg_1,col="red")

####### regression "with calculation"

## beta_1

beta_1=cov_XY/var_X
beta_1
round(beta_1,4)

## beta_0
beta_0=Y_bar-beta_1*X_bar
beta_0
round(beta_0,4)

####### R2

summary(reg_1)

head(fitted(reg_1))
head(residuals(reg_1))
```

```
# R^2=ESS/TSS=1-RSS/TSS

TSS=sum(((Y - Y_bar))^2 )
TSS
ESS=sum((fitted(reg_1) - Y_bar)^2)
ESS
RSS=sum(residuals(reg_1)^2)
RSS

TSS
ESS+RSS

#R squared
ESS/TSS

1-RSS/TSS

####### SER

sqrt(sum(residuals(reg_1)^2)/(n-2))
```