

PartC_3

Diego Castro

11/2/2021

```
load("C:/Users/usuario/OneDrive - University of East Anglia/PhD/First
Semestre/Econometrics/Assigment/dt_wages.RData")

library(stargazer)

## Warning: package 'stargazer' was built under R version 4.1.1
##
## Please cite as:
## Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary
Statistics Tables.
## R package version 5.2.2. https://CRAN.R-project.org/package=stargazer

library(data.table)

## Warning: package 'data.table' was built under R version 4.1.1

library(GGally)

## Warning: package 'GGally' was built under R version 4.1.1
## Loading required package: ggplot2
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg      ggplot2

library(MASS)

## Warning: package 'MASS' was built under R version 4.1.1

library(ggplot2)
library(Metrics)

## Warning: package 'Metrics' was built under R version 4.1.1

#Look at the summaries
stargazer(dt.wages, type = "text")

##
## =====
## Statistic  N    Mean    St. Dev.  Min    Pctl(25) Pctl(75)  Max
## -----
```

```
## wage      526  5.896   3.693   0.530   3.330   6.880   24.980
## educ      526 12.563   2.769    0      12      14      18
## exper     526 17.017  13.572    1      5      26      51
## tenure    526  5.105   7.224    0      0      7      44
## nonwhite   526  0.103   0.304    0      0      0      1
## female     526  0.479   0.500    0      0      1      1
## married    526  0.608   0.489    0      0      1      1
## numdep     526  1.044   1.262    0      0      2      6
## smsa       526  0.722   0.448    0      0      1      1
## northcen   526  0.251   0.434    0      0      0.8      1
## south      526  0.356   0.479    0      0      1      1
## west       526  0.169   0.375    0      0      0      1
## construc  526  0.046   0.209    0      0      0      1
## ndurman    526  0.114   0.318    0      0      0      1
## trcommpu   526  0.044   0.205    0      0      0      1
## trade      526  0.287   0.453    0      0      1      1
## services   526  0.101   0.301    0      0      0      1
## profserv   526  0.259   0.438    0      0      1      1
## profocc    526  0.367   0.482    0      0      1      1
## clerocc    526  0.167   0.374    0      0      0      1
## servocc    526  0.141   0.348    0      0      0      1
## lwage      526  1.623   0.532  -0.635  1.203   1.929   3.218
## expersq    526 473.435 616.045    1      25     676   2,601
## tenursq    526 78.150 199.435    0      0      49   1,936
## -----
```

```
dt.wages[, list(avg_wage=mean(wage) , sd_wage=sd(wage) ) ]
```

```
##      avg_wage sd_wage
## 1: 5.896103 3.693086
```

#Generate a confidence interval

```
conf.int <- function(wage) {
  n <- length(wage)
  error <- qt (0.975, df=526-1) * sd(wage) / sqrt (526)
  mean.X <- mean(wage)
  return (list(lower = mean.X - error , upper = mean.X + error ))
}
```

#In the section on testing can you reproduce the findings on the equality of the samples?

```
dt.wages[ female == 1, t.test (wage, mu=5)]
```

```
##
## One Sample t-test
##
## data: wage
## t = -2.5879, df = 251, p-value = 0.01022
```

```
## alternative hypothesis: true mean is not equal to 5
## 95 percent confidence interval:
##  4.273855 4.901462
## sample estimates:
## mean of x
##  4.587659

t.test ( dt.wages[ female == 1, wage] , dt.wages[ female == 0, wage] )

##
##  Welch Two Sample t-test
##
## data:  dt.wages[female == 1, wage] and dt.wages[female == 0, wage]
## t = -8.44, df = 456.33, p-value = 4.243e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.096690 -1.926971
## sample estimates:
## mean of x mean of y
##  4.587659  7.099489
```