Learning econometrics through simulated data

Sandeep Gangarapu

2021-12-20

Contents

1 Prerequisites **5**

2 Introduction **7**

3 OLS **9**

3.1 single variable regression 9

4 Methods **13**

4.1 math example 13

5 Final Words **15**

Chapter 1

Prerequisites

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports, e.g., a math equation $a^2 + b^2 = c^2$.

The **bookdown** package can be installed from CRAN or Github:

```
install.packages("bookdown")
# or the development version
# devtools::install_github("rstudio/bookdown")
```

Remember each Rmd file contains one and only one chapter, and a chapter is defined by the first-level heading `#`.

To compile this example to PDF, you need XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): https://yihui.name/tinytex/.

Chapter 2

Introduction

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter 4.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2021) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

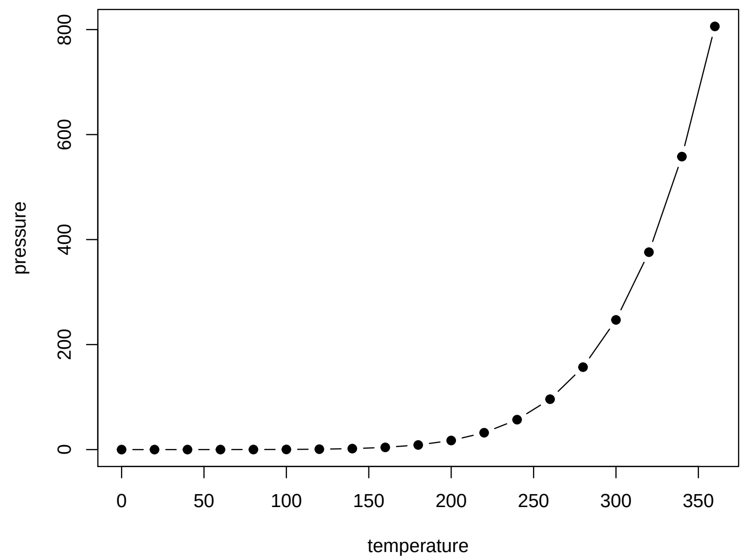

Figure 2.1: Here is a nice figure!

Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 3

OLS

3.1 single variable regression

y = 3x Y is caused by x. A unit change in x, changes y by 3 units.

3.1.1 Perfect fit

```
x = rnorm(100, 10, 2)

y =  3*x

reg <- lm(y~x)

summary(reg)
```

```
## Warning in summary.lm(reg): essentially perfect fit: summary may be unreliable
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.035e-13 -4.720e-16  9.610e-16  2.467e-15  5.921e-15
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
```

```
## (Intercept) 0.000e+00  5.446e-15 0.00e+00        1
## x           3.000e+00  5.208e-16 5.76e+15   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.072e-14 on 98 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:      1
## F-statistic: 3.318e+31 on 1 and 98 DF,  p-value: < 2.2e-16
```

This is essentially a perfect fit. There is barely any standard error. The pval is infinitesimal. However, in reality, that is not the case. There is always some error in measurement which is why we usually add in an error term.

3.1.2 Small error in measurement

```r
x = rnorm(100, 10, 2)
eps = rnorm(100,0,1)
y =  3*x + eps

reg <- lm(y~x)

summary(reg)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.38234 -0.56933 -0.06375  0.61628  2.62249
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.83778    0.47764   1.754   0.0826 .
## x            2.92894    0.04639  63.134   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9189 on 98 degrees of freedom
## Multiple R-squared:  0.976,  Adjusted R-squared:  0.9758
## F-statistic:  3986 on 1 and 98 DF,  p-value: < 2.2e-16
```

Now we are able to still recover estimate of x. The pvalues of says that the estimate is significatnlty different from zero. 95% confidence interval includes the true value.

3.1.3 Large error in measurement

```
x = rnorm(100, 10, 2)
eps = rnorm(100,0,100)
y =  3*x + eps

reg <- lm(y~x)

summary(reg)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -261.782  -55.651    3.665   59.063  229.111
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -164.321     50.008  -3.286 0.001411 **
## x             17.776      4.855   3.661 0.000407 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 91.64 on 98 degrees of freedom
## Multiple R-squared:  0.1203, Adjusted R-squared:  0.1113
## F-statistic:  13.4 on 1 and 98 DF,  p-value: 0.0004071
```

Chapter 4

Methods

We describe our methods in this chapter.

Math can be added in body using usual syntax like this

4.1 math example

p is unknown but expected to be around 1/3. Standard error will be approximated

$$SE = \sqrt(\frac{p(1-p)}{n}) \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

You can also use math in footnotes like this[1].

We will approximate standard error to 0.027[2]

---

[1] where we mention $p = \frac{a}{b}$

[2] p is unknown but expected to be around 1/3. Standard error will be approximated

$$SE = \sqrt(\frac{p(1-p)}{n}) \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

Chapter 5

Final Words

We have finished a nice book.

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2021). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.24.