

regression

Akira MATSUI

Using Regression to a report

When you want to use R as a data analysis tool, you must be writing a paper for some propose (for example, a term paper, report ..etc). In this exercise, you firstly conduct a simple regression analysis by using the built in iris data set . Next, you will wrap up an analyzed data and write a paper on LaTeX format by using stargaizer.

Questions

- Regress Sepal.Length with Sepal.Width, Petal.Length and Petal.Width
- Use stargaizer your result to a latex format
- convert latex to pdf

Sample Answer

Iris Data

R already has Iris data set. You can check it by

```
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1         5.1         3.5         1.4         0.2  setosa
## 2         4.9         3.0         1.4         0.2  setosa
## 3         4.7         3.2         1.3         0.2  setosa
## 4         4.6         3.1         1.5         0.2  setosa
## 5         5.0         3.6         1.4         0.2  setosa
## 6         5.4         3.9         1.7         0.4  setosa
```

.

If you want to see Sepal.Length of iris data, you can access it by

```
head(iris["Sepal.Length"])
```

```
##   Sepal.Length
## 1         5.1
## 2         4.9
## 3         4.7
## 4         4.6
## 5         5.0
## 6         5.4
```

Regression

R provide regression function `lm()`

```
lm(dependent variable ~ independent variable0 + independent variable1 + ... , data="Data Frame" )
```

In this case, Data Frame is iris. The regression analysis we are supposed to conduct is

```
result <- lm (Sepal.Length ~ Sepal.Width + Petal.Length + Petal.Width , data=iris)
summary(result)
```

```
##
## Call:
## lm(formula = Sepal.Length ~ Sepal.Width + Petal.Length + Petal.Width,
##     data = iris)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.82816 -0.21989  0.01875  0.19709  0.84570
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.85600    0.25078   7.401 9.85e-12 ***
## Sepal.Width    0.65084    0.06665   9.765 < 2e-16 ***
## Petal.Length   0.70913    0.05672  12.502 < 2e-16 ***
## Petal.Width   -0.55648    0.12755  -4.363 2.41e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3145 on 146 degrees of freedom
## Multiple R-squared:  0.8586, Adjusted R-squared:  0.8557
## F-statistic: 295.5 on 3 and 146 DF,  p-value: < 2.2e-16
```

Stargayzer

Stargayzer makes your result attractive. If you do not have stargayzer library, you can get and load it via

```
install.packages('stargayzer')
library(stargazer)
```

It convert result to latex format.

```
stargazer(result)
```

It also provides a text format

```
stargazer(result, type="text")
```

```
=====
                        Dependent variable:
-----
                        Sepal.Length
-----
Sepal.Width              0.651***
                        (0.067)

Petal.Length             0.709***
```

	(0.057)
Petal.Width	-0.556*** (0.128)
Constant	1.856*** (0.251)

Observations	150
R2	0.859
Adjusted R2	0.856
Residual Std. Error	0.315 (df = 146)
F Statistic	295.539*** (df = 3; 146)

=====

Note: *p<0.1; **p<0.05; ***p<0.01