

Homework 2:

A Literacy Study: Effect of Literacy and Scientific Articles on GDP per capita

Koray Yenil
March 10, 2020

Introduction

GDP per capita (GDPpc) is a measure of a country's development level. It is a snapshot of a country's economic health and it is affected by various factors. In this report, World Bank's publicly available data is used to extract and analyze three variables: GDP, literacy rate and Scientific and technical journal articles. The report uses literacy rate and Scientific and technical journal articles as explanatory variables and GDP as the response variable.

Null Hypothesis

There exists no relationship between a country's GDPpc and its corresponding literacy rate and scientific articles.

Methodology

Python is used to scrape data from the World Bank's website. A multiple linear regression model is used to estimate the beta values, and standard errors and 95% confidence interval.

First, Beta values were found by using the formula

$$\beta_i = (X'X)^{-1} * X'Y$$

By knowing the Beta values, we get regression equation in the form of

$$y = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2$$

From the formula, we obtain predicted y values, get the error values and standard error.

Using scipy library, Beta values and standard error, we calculate the 95% confidence interval using t values and degrees of freedom. The following equation was used:

$$\sigma^2 = (e'e) / (n - k - 1)$$

Summary Table

	Estimates	Standard Errors	Lower CI	Upper CI
Bo	43715.165686	251856.266301	-454985.396094	542415.727465
B1	-2168.104919	0.152571	-2168.407025	-2167.802813
B2	9909.688397	251005.449128	-487106.170403	506925.547197

Upper and Lower Bounds

```
lower=[l1,l2,l3]
upper=[u1,u2,u3]
```

```
Lower interval:
[[-454985.39609373]
 [ -2168.4070251 ]
 [-487106.17040346]]
Upper interval:
[[542415.72746511]
 [ -2167.80281288]
 [506925.54719707]]
```

Conclusion

According to the summary table, the estimates for β_1 , β_2 and β_3 are $4.37151657e+04$, $-2.16810492e+03$, $9.90968840e+03$, where their standard error values are 251,856 , 0.15 and 251,005. Therefore, we fail to reject the hypothesis, meaning that we do not have enough evidence to state that there is no relationship between a country's literacy rate, scientific articles and its GDPpc.

References

[1] "The World Bank Indicators API". [Online]. Available: <https://datahelpdesk.worldbank.org/knowledgebase/articles/889392-about-the-indicators-api-documentation> [Accessed: 7- Mar- 2020].