Exercises 10

Your name

2024-11-05

Exercises 10

```r
library(tidyverse)
```

1. A note on log transformations.

Notation:

In R, log computes logarithms, by default natural logarithms, log10 computes common (i.e., base 10) logarithms, and log2 computes binary (i.e., base 2) logarithms. The general form log(x, base) computes logarithms with base.("R Documentation")

Suppose we have a normal linear regression:

$Y = a + bX + e$

Level-Level Regression: Don't transform y or x.

Y is our dependent variable, a is the constant (intercept) term, b is the coefficient of interest and X is the independent variable. The error term is represented by e. This is known as a level-level regression.

So a unit increase in X (years of experience) is associated with a b*unit increase in Y (wage).

Level-Log Regression: Transform x.

Now, let us suppose that we had the (natural) log of X:

$Y = a + b\ln(X) + e$

Now we interpret the coefficient as a % increase in X, results in a (b/100)*unit increase in Y. This is known as a semi-elasticity or a level-log model.

Say that Y is wage, in thousands of pounds (£), and X is years of experience in the labour market. So a unit increase in X (years of experience) is associated with a b*unit increase in Y (wage).

In our example, this would mean that a 1% increase in years of experience results in a £(b/100) increase in wage.

Log-Level Regression: Transform y.

Next, let us turn to a model where the dependent variable is logged but the independent variable is not:

$\ln(Y) = a + bX + e$

This is known as a log-level model and the interpretation is that a unit increase in X results in a 100*b% increase in Y (we multiply by 100 because b is a percentage).

This is a rough approximation, assuming that b is small (approximately less than 0.15 in absolute value). More formally, we should exponeniate the coefficient, subtract one and multiply by 100: (exp(b)-1)*100.

This would mean that a year increase in experience is associated with a roughly 100*b% increase in wage.

Log-Log Regression: Transform y and x.

Our final model is a log-log model, with both dependent and independent variable appearing as (natural) logs:

$\ln(Y) = a + b\ln(X) + e$

This is interpreted as a 1% increase in X results in a b% increase in Y.

Therefore, for a 1% increase in experience we would expect wages to rise by b%.

Source: https://learneconomicsonline.com/blog/archives/1095

REGRESSION TYPE	REGRESSION	INTERPRETATION
Level-Level	Y = a + bX + e	A unit increase in X results in a b*unit increase in Y
Level-Log	Y = a + bln(X) + e	A percentage increase in X results in a (b/100)*unit increase in Y
Log-Level	ln(Y) = a + bX + e	A unit increase in X results in a 100*b% increase in Y
Log-Log	ln(Y) = a + bln(X) + e	A percentage increase in X results in a b% increase in Y

Transform x: level-log

Transform y: log-level

2. Data analysis.