[Academic Script]

Others Functional Forms

Subject:	Business Economics
Course:	B. A. (Hons.), 3rd Semester, Undergraduate
Paper No. & Title:	Paper – 304 Basic Econometrics
Unit No. & Title:	Unit – 4 Others Functional Forms
Lecture No. & Title:	Lecture – 1 Others Functional Forms

Academic Script

1. Introduction

Regression analysis is a technique of predicting a dependent variable using one or more independent variables.

We have already studied **Simple regression.** The model has one dependent and one independent variable. It's equation is given as

$$Y = b_0 + b_1 x + e$$

where Y is the response or dependent variable and X is the independent or explanatory or predictor variable.

Here b_0 is the intercept of the line, b_1 is the regression coefficient and e is the residual or error term.

Eg.

- Independent: First year mileage for a certain car model;
  Dependent: Maintenance cost.

2. Logarithmic transformation of variables

Considering the simple bivariate linear model $Y_i = b_0 + b_1 X_i$, there are four possible combinations of transformations involving logarithms: the linear case with no transformations, the linear-log model, the log-linear model, and the log-log model.

Y	**X**	
	X	**logX**
Y	*linear* $Y_i = b_0 + b_1 X_i$	*linear – log* $Y_i = b_0 + b_1 logX_i$
logY	*log-linear* $logY_i = b_0 + b_1 X_i$	*log-log* $logY_i = b_0 + b_1 logX_i$

(Remember that we are using natural logarithms, where the base is e ≈ 2.71828.)

3. Why use logarithmic transformations of variables?

Logarithmically transforming variables in a regression model is a very common way to handle situations where a non-linear relationship exists between the independent and dependent variables. Using the logarithm of one or more variables instead of the un-logged form makes the effective non-linear relationship, while still preserving the linear model. Logarithmic transformations are also a convenient means of transforming a highly skewed variable into one that is more approximately normal.

4. Interpreting coefficients in models with logarithmic transformations

1 Linear model: $Y_i = b_0 + b_1 X_i$

The coefficient b_1 gives us directly the change in Y for a one-unit change in X.

2 Linear-log model: $Y_i = b_0 + b_1 log X_i$

In the linear-log model, the literal interpretation of the estimated coefficient b_1 is that a one-unit increase in logX will produce an expected increase of b_1 units in Y.

The expected change in Y associated with a p% increase in X can be calculated as b_1 · log ([100 + p]/100). So to work out the expected change associated with a 10% increase in X, multiply b_1

by log (110/100) = log (1.1) = .095. In other words, $0.095b_1$ is the expected change in Y when X is multiplied by 1.1, i.e. when X increases by 10%.

3 Log-linear model: $logY_i = b_0 + b_1X_i$

In the log-linear model, the literal interpretation of the estimated coefficient b_1 is that a one-unit increase in X will produce an expected increase of b_1 units in log Y.

The effect of a c-unit increase in X is to multiply the expected value of Y by e*c b_1. So the effect for a 5-unit increase in X would be e*5 b_1.

4 Log-log model: $logY_i = b_0 + b_1logX_i$

The interpretation is given as an expected percentage change in Y when X increases by some percentage. Such relationships, where both Y and X are log-transformed, are commonly referred to as elastic in econometrics, and the coefficient of log X is referred to as an elasticity. So in terms of effects of changes in X on Y (both unlogged):

- multiply X by e will multiply expected value of Y by e* b_1
- To get the proportional change in Y associated with a p percent increase in X, calculate a = log ([100 + p]/100) and take e* ab_1

5. Reciprocal Transformation

$$Y_i = b_0 + b_1\left(\frac{1}{X_i}\right)$$

If the relationship between Y and X is curvilinear, as in the case of the Phillips curve, this model generally gives a good fit.

6. Example

An experiment result of number of bacteria as per the number of dose is given below. Fit a curve $y = e^{a+bx}$ and estimate the number of bacteria for 16th dose.

Dose	**Bacteria**
1	35500
2	21100
3	19700
4	16600
5	14200
6	10600
7	10400
8	6000
9	5600
10	3800
11	3600
12	3200
13	2100
14	1900
15	1500

Solution:

Dose is the independent variable x and Number of bacteria is the dependent variable y .

The curve to be fitted is $y = e^{a+bx}$.

Taking natural logarithmic transformation, we get $Y = a + b\,x$ where Y = $\ln y$.

Dose x	**Bacteria y**	$Y = ln\,y$
1	35500	10.4773
2	21100	9.95703
3	19700	9.88837
4	16600	9.71716
5	14200	9.561

6	10600	9.26861
7	10400	9.24956
8	6000	8.69951
9	5600	8.63052
10	3800	8.24276
11	3600	8.18869
12	3200	8.07091
13	2100	7.64969
14	1900	7.54961
15	1500	7.31322

This data was analyzed using SPSS software and following results were obtained.

Table 1: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.994	.988	.987	.110017408

Table 1 gives the value of R square which is 0.988, indicating that the data exhibit strong regression relationship. The value of adjusted R Square is 0.987.

Table 2: ANOVA

Model		Sum of Squares	D f	Mean Square	F	Sig.
	Regression	13.359	1	13.359	1103.677	.000[b]
1	Residual	.157	13	.012		
	Total	13.516	14			

Table 2 gives the ANOVA analysis which provides the statistical tests for the overall model fit in terms of the F ratio. Here F ratio is 1103.677 and significance level is 0.000 which indicates that the model is a good fit.

Table 3: Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.578	.060		176.958	.000
	Dose	-.218	.007	-.994	-33.222	.000

Using Table 3 we get the fitted equation:

Ln bacteria = 10.579 -0.218 (**Dose**)

Here -0.218 is the regression coefficient of dose which indicates that with every increase of a dose, the ln- bacteria decreases by 0.218.

We want to estimate number of bacteria for 16 doses,

Ln bacteria = 10.579 -0.218 (**16**) = 7.091

Thus, number of bacteria = $e^{7.091}$ = 1201.10 ≈ 1201

7. Summary

Regression analysis is a technique of predicting a dependent variable using one or more independent variables.

Logarithmic transformation of variables

Considering the simple bivariate linear model $Y_i = b_0 + b_1 X_i$, there are four possible combinations of transformations involving logarithms: the linear case with no transformations, the linear-log model $Y_i = b_0 + b_1 log X_i$, the log-linear model $log Y_i = b_0 + b_1 X_i$ and the log-log model $log Y_i = b_0 + b_1 log X_i$.

Why use logarithmic transformations of variables?

Logarithmically transforming variables in a regression model is a very common way to handle situations where a non-linear relationship exists between the independent and dependent variables. Using the logarithm of one or more variables instead of the un-logged form makes the effective relationship non-linear, while still preserving the linear model. Logarithmic transformations are also a convenient means of transforming a highly skewed variable into one that is more approximately normal.

Reciprocal Transformation: $Y_i = b_0 + b_1\left(\frac{1}{X_i}\right)$

If the relationship between Y and X is curvilinear, as in the case of the Phillips curve, this model generally gives a good fit.