# Transformations to Linearize the Model

Dr. Kiah Wah Ong

# Transformations to Linearize the Model

Recall the major assumptions we have made in linear regression models

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i, \quad i = 1, 2, \cdots, n$$

are

- The relationship between the response and regressors is linear.
- The error terms $\epsilon_i$ have mean zero.
- The error terms $\epsilon_i$ have constant variance $\sigma^2$ (homoscedasticity)
- The error terms $\epsilon_i$ are normally distributed.
- The error terms $\epsilon_i$ and $\epsilon_j$ are uncorrelated for $i \neq j$.
- The regressors $x_1, \cdots, x_k$ are nonrandom.
- The regressors $x_1, \cdots, x_k$ are measured without error.
- The regressors are linearly independent.

# Transformations to Linearize the Model

Even though the usual starting point in regression analysis is to assume $y$ and the regressors are linearly related.

Occasionally, this assumption is inappropriate either because

- nonlinearity is detected during model diagnostic.
- prior experience or theoretical consideration suggest the relationship is nonlinear.

# Transformations to Linearize the Model

In this lesson, we will be looking at some of the **linearizable** models.

These are models with nonlinear relationship between the response and regressors, but can be changed into linear relationship through the use of transformation.

# Transformations to Linearize the Model

A model often used to describe a growth process is

$$y = \alpha x^{\beta}$$

This functional form is linearizable since

$$\begin{aligned} \log y &= \log \alpha + \beta \log x \\ &= \beta_0 + \beta_1 \log x \end{aligned}$$

Another model often used to describe a growth process is

$$y = \alpha \beta^{x}$$

This functional form is also linearizable since

$$\begin{aligned} \log y &= \log \alpha + (\log \beta) x \\ &= \beta_0 + \beta_1 x \end{aligned}$$

# Transformations to Linearize the Model

There is another possibility for transformation involving logarithm where the response variable $y$ is not logged, but the regressor $x$ is.

This situation arises when

$$\exp(y) = \alpha x^{\beta}$$

Logging both sides gives the relationship

$$\begin{aligned} y &= \log \alpha + \beta \log x \\ &= \beta_0 + \beta_1 \log x \end{aligned}$$

ILLINOIS TECH

# Transformations to Linearize the Model

Common linearizable models and the required transformations are shown below:

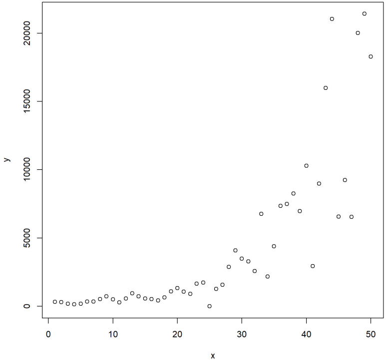
| True relationship | Transformation | Linearized model |
|---|---|---|
| $y = \beta_0 x^{\beta_1}$ | $y' = \log y, x' = \log x$ | $y' = \log \beta_0 + \beta_1 x'$ |
| $y = \beta_0 e^{\beta_1 x}$ | $y' = \ln y$ | $y' = \ln \beta_0 + \beta_1 x$ |
| $y = \beta_0 + \beta_1 \log x$ | $x' = \log x$ | $y = \beta_0 + \beta_1 x'$ |
| $y = \frac{x}{\beta_0 x - \beta_1}$ | $y' = \frac{1}{y}, x' = \frac{1}{x}$ | $y' = \beta_0 - \beta_1 x'$ |

# Transformations to Linearize the Model

## Example

Let us look at the data in the file $\mathrm{VT1.CSV}$.

```
VarTrans<-read.csv("VT1.CSV", header=TRUE, sep=",")
x<-VarTrans$x
y<-VarTrans$y
plot(x,y)
```




We can see clearly that the relationship between $x$ and $y$ are nonlinear. Suppose, a priori that we know the relationship between the $x$ and $y$ is of the form:

$$y = \alpha\beta^x$$

# Transformations to Linearize the Model

Just like what is shown in this table below. We will try to refit the model using the change of variable

$$y \to \ln y.$$

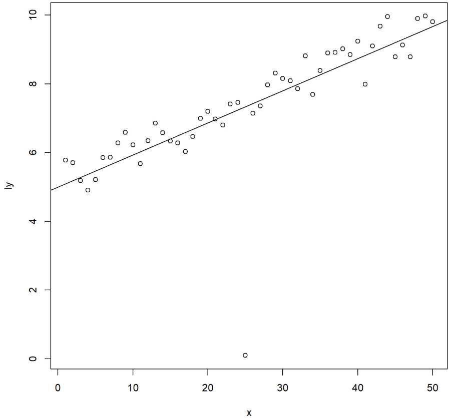
| True relationship | Transformation | Linearized model |
|---|---|---|
| $y = \beta_0 x^{\beta_1}$ | $y' = \log y, x' = \log x$ | $y' = \log \beta_0 + \beta_1 x'$ |
| $y = \beta_0 e^{\beta_1 x}$ | $y' = \ln y$ | $y' = \ln \beta_0 + \beta_1 x$ |
| $y = \beta_0 + \beta_1 \log x$ | $x' = \log x$ | $y = \beta_0 + \beta_1 x'$ |
| $y = \frac{x}{\beta_0 x - \beta_1}$ | $y' = \frac{1}{y}, x' = \frac{1}{x}$ | $y' = \beta_0 - \beta_1 x'$ |

# Transformations to Linearize the Model

```
ly=log(y)
plot(x, ly)
model1=lm(ly~x)
abline(model1)
summary(model1)
par(mfrow=c(2,2))
plot(model1)
```

Using $R$ to set up the regression model

$$\ln y = \beta_0 + \beta_1 x + \epsilon$$



# Transformations to Linearize the Model

The diagnostic plots for the regression model

$$\ln y = \beta_0 + \beta_1 x + \epsilon$$

are pretty satisfactory.