Linear Regression for Business Statistics

Errors and residuals in a regression model

Linear Regression for Business Statistics

Errors and residuals in a regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + error$$

Linear Regression for Business Statistics

Errors and residuals in a regression model

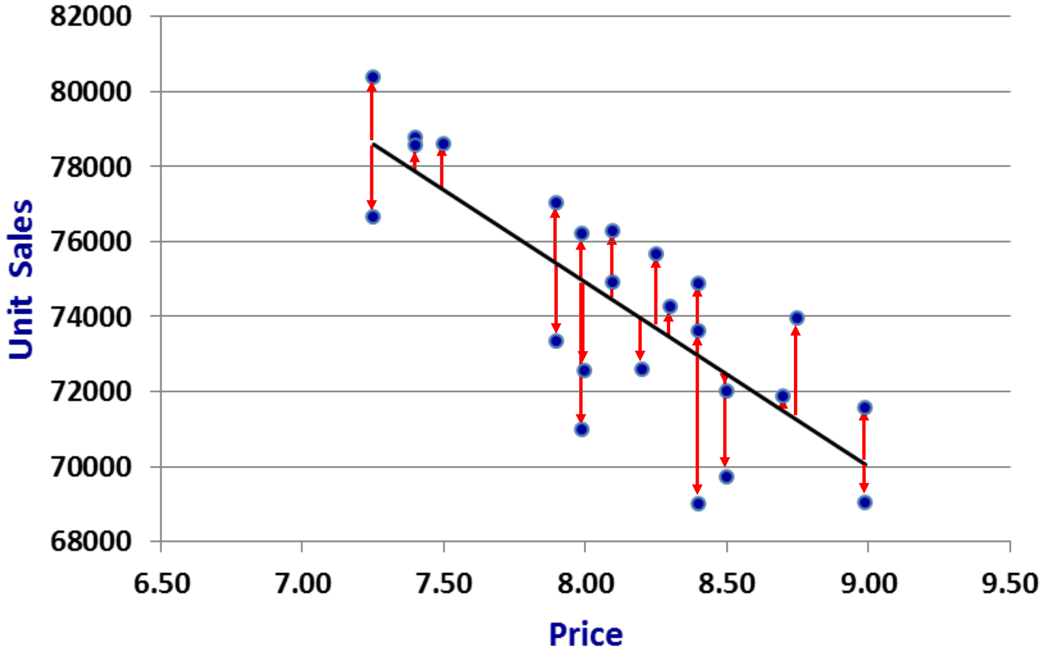
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + error$$

$$error \sim Normal(0, \quad some\ std)$$

RICE
JONES GRADUATE
SCHOOL OF BUSINESS

Linear Regression for Business Statistics

$$error \sim Normal(0,\ \ some\ std)$$



Linear Regression for Business Statistics

Importance of the Normality assumption about errors

Linear Regression for Business Statistics

Importance of the Normality assumption about errors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k$$

Linear Regression for Business Statistics

Importance of the Normality assumption about errors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k$$

Estimated Model :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

Linear Regression for Business Statistics

Importance of the Normality assumption about errors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

The *b*'s can be considered as random variables…,

Linear Regression for Business Statistics

Importance of the Normality assumption about errors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

The *b*'s can be considered as random variables…,

$$b_0 \sim Normal(\beta_0, \quad some\ std)$$

Linear Regression for Business Statistics

Importance of the Normality assumption about errors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

The *b*'s can be considered as random variables…,

$$b_0 \sim Normal(\beta_0, \quad some\ std)$$

$$b_1 \sim Normal(\beta_1, \quad some\ std)$$

Linear Regression for Business Statistics

Importance of the Normality assumption about errors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

The *b*'s can be considered as random variables…,

$$b_0 \sim Normal(\beta_0, \quad some\ std)$$

$$b_1 \sim Normal(\beta_1, \quad some\ std)$$

…

…

Linear Regression for Business Statistics

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Population Mean : μ

Sample Mean : $\bar{x}$

Linear Regression for Business Statistics

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Population Mean : μ ← Fixed but unknown

Sample Mean : $\bar{x}$

Linear Regression for Business Statistics

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Population Mean : μ ← Fixed but unknown

Sample Mean : $\bar{x}$ ← $\bar{x} \sim Normal(\mu,\ some\ std)$

Linear Regression for Business Statistics

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Population Mean : $\mu \leftarrow$ Fixed but unknown

Sample Mean : $\bar{x} \leftarrow$ $\bar{x} \sim Normal(\mu,\ some\ std)$

Linear Regression for Business Statistics

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

Estimated Model : $$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Population Mean : $\mu \leftarrow$ Fixed but unknown

Sample Mean : $\bar{x} \leftarrow$ $\bar{x} \sim Normal(\mu,\ some\ std)$

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of “X” variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of “X” variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of "X" variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of “X” variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of "X" variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of "X" variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

$$\frac{b_1 - \beta_1}{s_{b_1}} \sim t_{n-k-1}$$

s_{b_1} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of "X" variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

$$\frac{b_1 - \beta_1}{s_{b_1}} \sim t_{n-k-1}$$

s_{b_1} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of "X" variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

$$\frac{b_1 - \beta_1}{s_{b_1}} \sim t_{n-k-1}$$

s_{b_1} = *the standard error of* b_0

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of “X” variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

$$\frac{b_1 - \beta_1}{s_{b_1}} \sim t_{n-k-1}$$

s_{b_1} = *the standard error of* b_0

…

…

Linear Regression for Business Statistics

Some important results…

$$\frac{b_0 - \beta_0}{s_{b_0}} \sim t_{n-k-1}$$

n = number of observations

k = number of “X” variables

n-k-1 = residual degrees of freedom

s_{b_0} = *the standard error of* b_0

$$\frac{b_1 - \beta_1}{s_{b_1}} \sim t_{n-k-1}$$

s_{b_1} = *the standard error of* b_0

…

…

These results enable us to…,

- ***test stability and precision of coefficients***
- ***conduct hypothesis testing in a regression***

Linear Regression for Business Statistics

A note on notations…

Rice Jones Graduate School of Business

Linear Regression for Business Statistics

A note on notations…

- ***We have differentiated between the β's and the b's.***
- ***In common usage people tend to mix these notations.***
- ***However, as long as you understand the conceptual difference, you should be ok.***

Linear Regression for Business Statistics

A note on notations…

- ***We have differentiated between the β's and the b's.***
- ***In common usage people tend to mix these notations.***
- ***However, as long as you understand the conceptual difference, you should be ok.***

Linear Regression for Business Statistics

A note on notations…

- ***We have differentiated between the β's and the b's.***
- ***In common usage people tend to mix these notations.***
- ***However, as long as you understand the conceptual difference, you should be ok.***