



ECONOMETRICS-I Lecture Notes

Chapter
6

EXTENSIONS OF THE TWO-VARIABLE LINEAR REGRESSION MODEL

Regression Through the Origin

Population Model: $Y_i = \beta_2 X_i + u_i$; Sample Model: $Y_i = \hat{\beta}_2 X_i + \hat{u}_i$

$$RSS(\hat{\beta}_2) = \sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_2 X_i)^2$$

$$\frac{d RSS(\hat{\beta}_2)}{d \hat{\beta}_2} = 2 \cdot \sum (Y_i - \hat{\beta}_2 X_i) \cdot (-X_i) = 0$$

$$\sum X_i Y_i - \hat{\beta}_2 \cdot \sum X_i^2 = 0$$

$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

The other formulas are as follows:

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_i^2} \quad \text{and} \quad \hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-1}$$

→ df is $n-1$ because there's only one β in the model.

Notes:

* $\sum \hat{u}_i^2$ is NOT necessarily 0.

* r^2 may be negative. So, we compute

$$\text{raw } r^2 = \frac{\sum (X_i Y_i)^2}{\sum X_i^2 \sum Y_i^2}$$

* Raw r^2 and r^2 are NOT comparable

* Unless there's very strong a priori expectation, intercept-present model is advised.

6.2. The following regression results were based on monthly data over the period January 1978 to December 1987:

$$\begin{aligned} \hat{Y}_t &= 0.00681 + 0.75815X_t \\ \text{se} &= (0.02596) \quad (0.27009) \\ t &= (0.26229) \quad (2.80700) \\ p \text{ value} &= (0.7984) \quad (0.0186) \quad r^2 = 0.4406 \\ \hat{Y}_t &= 0.76214X_t \\ \text{se} &= (0.265799) \\ t &= (2.95408) \\ p \text{ value} &= (0.0131) \quad r^2 = 0.43684 \end{aligned}$$

where Y = monthly rate of return on Texaco common stock, %, and X = monthly market rate of return, %.

- What is the difference between the two regression models?
- Given the preceding results, would you retain the intercept term in the first model? Why or why not?
- How would you interpret the slope coefficients in the two models?
- What is the theory underlying the two models?
- Can you compare the r^2 terms of the two models? Why or why not?
- The Jarque-Bera normality statistic for the first model in this problem is 1.1167 and for the second model it is 1.1170. What conclusions can you draw from these statistics?
- The t value of the slope coefficient in the zero intercept model is about 2.95, whereas that with the intercept present is about 2.81. Can you rationalize this result?

6.2) a) First model is usual two-variable regression model whereas second one is regression through the origin.

b) For the first Model, we have;

(i) $H_0: \beta_1 = 0$

$H_A: \beta_1 \neq 0$

$\alpha = 0.05$

(ii) $t = \frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)}$

(iii) Reject H_0 if $p\text{-value} < \alpha$

(iv) $p\text{-value} = 0.7984 > 0.05$

(v) Do NOT Reject H_0 at any reasonable α

Intercept term may be omitted from the model, data supports second model.



c) Second Model (first one is also similar): A one percent increase in the monthly market rate of return leads on the average about 0,762% increase in the rate of return on Texaco common stock.

d) Common stock rate of return is (positively) related to market rate of return.

e) They're NOT comparable because second is raw r^2 , whose formula is different from the usual r^2 .

f) Jarque-Bera (JB) Test of Normality:

$$JB = n \cdot \left[\frac{S^2}{6} + \frac{(K-3)^2}{24} \right] \sim \chi^2_2 \text{ where } S: \text{Skewness and } K: \text{Kurtosis}$$

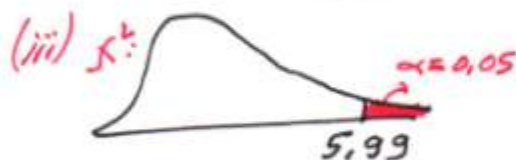
is used to make a normality test of residuals.

(i) H_0 : Normality Assumption is Valid
 H_1 : Residuals do NOT follow Normal distribution

$$\alpha = 0,05$$

(iv) $JB = 1,12$

(ii) $JB \sim \chi^2_2$ pdf



Reject H_0 if $JB > 5,99$

Do NOT Reject H_0 . Normality assumption is valid at $\alpha = 0,05$
 (Note that $p\text{-value} > 0,50$)

g) Greater t model is better since evidence is stronger B_2 to be Nonnegative (Validity of the model). This result also supports the second model.

Scaling and Units of Measurement

let Y = Gross Private Domestic Investment (GPDI) and
 X = Gross Domestic Product (GDP)

$$y_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_i + \hat{u}_i$$

Defining w_1 and w_2 "scaling factors";

$$y_i^* = w_1 \cdot y_i \quad \text{and} \quad x_i^* = w_2 \cdot x_i$$

Consider the model:

$$y_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* \cdot x_i^* + \hat{u}_i^*$$

Note that: $\hat{u}_i^* = y_i^* - \hat{y}_i^* = w_1 y_i - w_1 \hat{y}_i = w_1 (y_i - \hat{y}_i) = w_1 \hat{u}_i$

$$\hat{u}_i^* = w_1 \cdot \hat{u}_i$$

Remember; $\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2} \quad \hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2}$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \cdot \bar{x} \quad \text{Var}(\hat{\beta}_1) = \frac{\sum x_i^2}{n \cdot \sum x_i^2} \cdot \sigma^2$$

So, we have;

$$\hat{\beta}_2^* = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}} = \frac{\sum w_2 x_i w_1 y_i}{\sum w_2^2 x_i^2} = \frac{w_1 \cdot w_2}{w_2^2} \frac{\sum x_i y_i}{\sum x_i^2} = \left(\frac{w_1}{w_2} \right) \hat{\beta}_2$$

$$\hat{\beta}_1^* = \bar{y}^* - \hat{\beta}_2^* \cdot \bar{x}^* = w_1 \bar{y} - \frac{w_1}{w_2} \cdot \hat{\beta}_2 \cdot w_2 \bar{x} = w_1 (\bar{y} - \hat{\beta}_2 \bar{x}) = w_1 \hat{\beta}_1$$

$$\hat{\sigma}^{*2} = \frac{\sum \hat{u}_i^{*2}}{n-2} = \frac{\sum (w_1 \hat{u}_i)^2}{n-2} = w_1^2 \frac{\sum \hat{u}_i^2}{n-2} = w_1^2 \cdot \hat{\sigma}^2$$

likewise; $\text{Var}(\hat{\beta}_2^*) = \left(\frac{w_1}{w_2} \right)^2 \cdot \text{Var}(\hat{\beta}_2)$; $\text{Var}(\hat{\beta}_1^*) = w_1^2 \text{Var}(\hat{\beta}_1)$

and $r_{xy}^2 = r_{x^*y^*}^2$ \rightarrow Explanatory Power of Regression is independent of scaling.



Example Consider the model where both GDPDI and GDP are measured in billions of dollars:

$$\widehat{GDPDI}_t = -1026,498 + 0,3016 GDP_t$$

$$s.e = (257,5874 \quad 0,0399) \quad r^2 = 0,8772$$

Obtain the models where the measures are

- Both GDPDI and GDP in millions of dollars
- GDPDI in billions, GDP in millions of dollars
- GDPDI in millions, GDP in billions of dollars

Answer a) $w_1 = w_2 = 1000$. New Model is;

$$\widehat{GDPDI}_t = -1026498 + 0,3016 GDP_t$$

$$s.e = (257587,4) \quad (0,0399) \quad r^2 = 0,8772$$

b) $w_1 = 1 \quad w_2 = 1000$. New Model is;

$$\widehat{GDPDI}_t = -1026,498 + 0,0003016 GDP_t$$

$$s.e \quad (257,5874) \quad (0,0000399) \quad r^2 = 0,8772$$

c) $w_1 = 1000 \quad w_2 = 1$. New Model is;

$$\widehat{GDPDI}_t = -1026498 + 301,6 GDP_t$$

$$s.e = (257587,4) \quad (39,9) \quad r^2 = 0,8772$$

Regression on Standardized Variables

A variable is said to be *standardized* if we subtract the mean value of the variable from its individual values and divide the difference by standard deviation. Namely;

$$y_i^* = \frac{y_i - \bar{y}}{s_y} = \frac{y_i}{s_y} \quad \text{and} \quad x_i^* = \frac{x_i - \bar{x}}{s_x} = \frac{x_i}{s_x}$$

6.5. Consider the following models:

Model I: $Y_i = \beta_1 + \beta_2 X_i + u_i$

Model II: $Y_i^* = \alpha_1 + \alpha_2 X_i^* + u_i$

where Y^* and X^* are standardized variables. Show that $\hat{\alpha}_2 = \hat{\beta}_2 (s_x/s_y)$ and hence establish that although the regression slope coefficients are independent of the change of origin they are not independent of the change of scale.

6.5) Note that; $\sum y_i^* = \sum \frac{y_i - \bar{y}}{s_y} = \frac{1}{s_y} (\sum y_i - n\bar{y}) = 0$

So; $\bar{y}_i^* = \bar{x}_i^* = 0$ and $y_i^* = \frac{y_i}{s_y}$; $x_i^* = \frac{x_i}{s_x}$

We have; $\hat{\alpha}_2 = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}} = \frac{\sum \frac{x_i}{s_x} \frac{y_i}{s_y}}{\sum \frac{x_i^2}{s_x^2}} = \frac{s_x^2}{s_y s_x} \cdot \frac{\sum x_i y_i}{\sum x_i^2} = \hat{\beta}_2 \frac{s_x}{s_y}$

So; $\hat{\alpha}_2 = \hat{\beta}_2 \cdot \frac{s_x}{s_y}$

$\hat{\alpha}_1 = \bar{y}^* - \hat{\alpha}_2 \cdot \bar{x}^* = 0$ $y_i^* = \alpha_2 x_i^* + u_i$

Note that, we interpret $\hat{\alpha}_2$ as follows:

"If X_i increases by one standard deviation, the expected change in Y_i is $\hat{\alpha}_2$ standard deviations."

Functional forms of Regression Models.

Remember, Slope is derivative of y w.r.t X . $m = \frac{dy}{dX}$

Elasticity of y w.r.t X is "the percentage change in y for a given (small) percentage change in X ."

Namely;
$$E = \frac{dy/y}{dX/X} = \frac{dy}{dX} \cdot \frac{X}{y}$$

Note that, $\Delta X = X_t - X_{t-1}$ refers to "change" and dX refers to "small change" (as limit goes to 0) small changes

We have; (i) Absolute change: $X_t - X_{t-1} = \Delta X = dX$

(ii) Relative change: $\frac{X_t - X_{t-1}}{X_{t-1}} = \frac{\Delta X}{X_{t-1}} = \frac{dX}{X}$

(iii) Percentage change: $\frac{X_t - X_{t-1}}{X_{t-1}} \cdot 100 = \frac{\Delta X}{X_{t-1}} \cdot 100$

We'll use these definitions in functional forms.

(I) The log-log Model:

$$y_i = \beta_1 \cdot X_i^{\beta_2} \cdot e^{u_i}$$

$$\ln y_i = \ln \beta_1 + \beta_2 \cdot \ln X_i + u_i$$

$\underbrace{\ln y_i}_{y_i^*} = \underbrace{\ln \beta_1}_{\alpha} + \beta_2 \cdot \underbrace{\ln X_i}_{X_i^*} + u_i$

$$y_i^* = \alpha + \beta_2 \cdot X_i^* + u_i$$

Taking the derivative of both sides;

$$\frac{dy/dx}{y} = \beta_2 \cdot \frac{1}{x} \Rightarrow \frac{dy}{dx} = \frac{y}{x} \cdot \beta_2$$

Slope: $m = \frac{y}{x} \cdot \beta_2$; Elasticity: $E = \frac{dy}{dx} \cdot \frac{x}{y} = \frac{y}{x} \cdot \beta_2 \cdot \frac{x}{y} = \beta_2$

Since Elasticity is β_2 , we call the log-log model as "Constant elasticity model."

Example; PCEXP: Total personal consumption expenditure

EXPDUR: Expenditure on durable goods

Consider the following SRF:

$$\widehat{\ln EXPDUR_t} = -9,6971 + 1,9056 \ln PCEXP_t$$

s.e. = (0,4341) (0,0514)

t = (-22,337) (37,096)

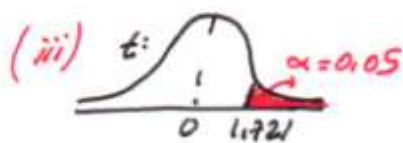
$r^2 = 0,9849$
 $n = 23$

- Interpret the ~~slope~~ coefficient: $\hat{\beta}_2$
- Is ~~PCEXP~~ EXPDUR elastic with respect to PCEXP?

Answer a) $E = 1,91$: If total personal expenditure goes up by 1%, on average, the expenditure on durable goods goes up by about 1,91%.

b) (i) $H_0: \beta_2 \leq 0$
 $H_A: \beta_2 > 0$
 $\alpha = 0,05$

(ii) $t = \frac{\hat{\beta}_2 - \beta_2}{SE(\hat{\beta}_2)}$; $df = n - 2 = 21$



Reject H_0 if $t > 1,721$

(iv) $t = \frac{1,9056 - 1}{0,0514} = 17,62$

(v) Reject H_0 . Yes, it is elastic at $\alpha = 0,05$

(II) Log-Lin Model (the Growth Rate)

Remember: Compound interest formula

$$Y_t = Y_0 (1+r)^t$$

$$\ln Y_t = \underbrace{\ln Y_0}_{=\beta_1} + \underbrace{t}_{=X_t} \cdot \underbrace{\ln(1+r)}_{=\beta_2}$$

$$\ln Y_t = \beta_1 + \beta_2 \cdot X_t + u_t$$

In this model, the slope coefficient β_2 measures the constant proportional or relative change in Y for a given absolute change in the value of the regressor

$$\beta_2 = \frac{\text{relative change in regressand}}{\text{absolute change in regressor}}$$

* In the Model $\ln Y_t = \beta_1 + \beta_2 t + u_t$; β_2 gives the instantaneous rate of growth. To find the compound (over a period of time) rate of growth, we use

$$\text{Compound Rate of Growth} = (e^{\hat{\beta}_2} - 1) \cdot 100\%$$

Example: EXS: Expenditure on services

$$\widehat{\ln EXS_t} = 7,7890 + 0,00743t$$

$$s.e. = (0,0023) \quad (0,00017)$$

$$t = (3387,619) \quad (46,2826)$$

$$n = 23$$

$$(t = 1, 2, \dots, 23)$$

$$r = 0,9894$$

- what is the elasticity at time 20?
- what is the (average) elasticity?
- what is Compound Rate of Growth?

Answer Note that: $\ln Y_t = \beta_1 + \beta_2 \cdot X_t$

$$\frac{dY/dX}{Y} = \beta_2 \Rightarrow \frac{dY}{dX} = Y \cdot \beta_2$$

$$E = \frac{dY}{dX} \cdot \frac{X}{Y} = Y \cdot \beta_2 \cdot \frac{X}{Y} = X \cdot \beta_2$$

a) $E(t) = t \cdot \hat{\beta}_2 \Rightarrow E(20) = 20 \cdot 0,00743 = 0,1486$

b) $E(\bar{t}) = \bar{E}(12) = 12 \cdot 0,00743 = 0,0892$

$$\left(\bar{t} = \frac{1+23}{2} = 12 \right)$$

c) Growth rate = 0,743%
Compound Rate of Growth = $(e^{0,00743} - 1) \cdot 100 = 0,746\%$

(III) Lin-log Model

$$Y_i = \beta_1 + \beta_2 \cdot \ln X_i + u_i$$

$$\beta_2 = \frac{\text{Change in } Y}{\text{Change in } \ln X} = \frac{\text{Change in } Y}{\text{Relative Change in } X} = \frac{\Delta Y}{\Delta X/X}$$

Relative

$$\beta_2 = \frac{\Delta Y}{\Delta X/X} \Rightarrow \boxed{\Delta Y = \beta_2 \cdot \frac{\Delta X}{X}}$$

absolute change = $\frac{\beta_2}{100} \cdot \frac{\Delta X}{X} \cdot 100$ *percentage change*

* Note that, if $\beta_2 = 500$, the absolute change in Y is $\frac{500}{100} = 5$ units. So, do NOT forget to divide β_2 by 100. This implies the interpretation:
"if X increases by 1%, the expected change in Y is $\frac{\hat{\beta}_2}{100}$ units."



Example 4 Consider the estimated regression function of Food Expenditure on Total Expenditure:

$$\widehat{\text{Food Exp}_i} = -1283,91 + 257,27 \ln \text{Total Exp}_i$$

$t = \quad (-4,38) \quad (5,66) \quad r^2 = 0,377$

- a) Interpret the slope coefficient.
- b) Estimate the Elasticity of food expenditure with respect to Total Expenditure when total Expenditure is 460.

Answer 4 a) $\hat{\beta}_2 = 257$ means that an increase in the total food expenditure of 1%, on average, leads to about $\frac{\hat{\beta}_2}{100} = \frac{257}{100} = 2,57\%$ increase in the expenditure on food.

b) $Y = \beta_1 + \beta_2 \ln X$

$$\frac{dY}{dX} = \beta_2 \cdot \frac{1}{X} \Rightarrow E = \frac{dY}{dX} \cdot \frac{X}{Y} = \beta_2 \cdot \frac{1}{X} \cdot \frac{X}{Y} = \beta_2 \cdot \frac{1}{Y}$$

$$\hat{E} = \hat{\beta}_2 \cdot \frac{1}{\hat{Y}}$$

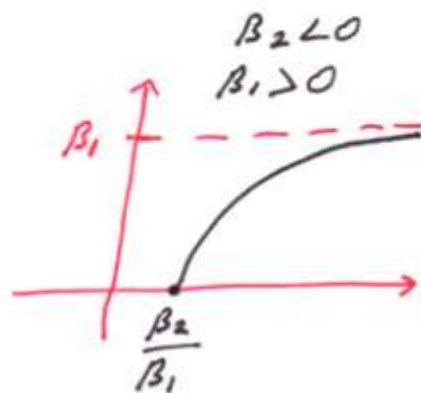
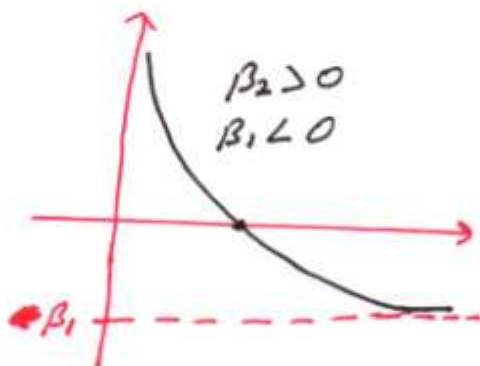
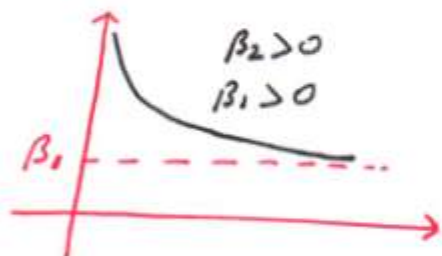
$$(\hat{Y} | X=460) = -1283,91 + 257,27 \cdot \ln(460) = 293,5$$

$$\hat{E} = 257,27 \cdot \frac{1}{293,5} = 0,877$$

* Note that, if NO X (or Y) value is given for elasticity, Estimate elasticity for \bar{X} (or \bar{Y}). Surely, we'll know $\sum X_i$ (or $\sum Y_i$)

(IV) Reciprocal Models.

$$Y_i = \beta_1 + \beta_2 \cdot \frac{1}{X_i} + U_i$$



Notes:

- * As X increases indefinitely, Y approaches β_1
- * Slope = $\frac{dY}{dX} = -\beta_2 \cdot \frac{1}{X^2}$: slope has converse sign with β_2
- * Elasticity is: $E = \frac{dY}{dX} \cdot \frac{X}{Y} = -\beta_2 \cdot \frac{1}{X^2} \cdot \frac{X}{Y} = -\beta_2 \cdot \frac{1}{XY}$

Example Consider the model that Regresses Child Mortality (CM) on Per Capita Gross National Product (PGNP)

$$\widehat{CM}_i = 81,79 + 27273,17 \cdot \left(\frac{1}{PGNP_i} \right)$$

$$s.e = (10.83) \quad (3760)$$

$$t = 7,55 \quad 7,25 \quad r^2 = 0,459$$

a) If Per Capita GNP increases indefinitely, what is the limiting expected Child Mortality?

b) By how much, 1 \$ increase in PGNP decreases the child mortality, When PGNP is 1000 \$?

Answer \leftarrow a) $\hat{\beta}_1 = 81,79 \approx 82$ deaths / thousand

b) Slope = $-\hat{\beta}_2 \cdot \frac{L}{X^2} = -27273,17 \cdot \frac{L}{1000^2} = -0,027$

Decrease is 0,027 deaths / thousand

Choice of functional form

(i) The underlying theory (i.e. the Phillips curve) may suggest a particular functional form.

(ii) If the y_i 's are the same (and the same function), one may compare r^2 values to compare models.

(iii) Note that, statistical significance (higher t values), the signs of the coefficients and theoretical underpinning of the model are much more important than r^2 values.

Summary of Slopes and Elasticity

Model	Equation	Slope ($= \frac{dy}{dx}$)	Elasticity ($= \frac{dy}{dx} \cdot \frac{x}{y}$)
Linear	$y = \beta_1 + \beta_2 X$	β_2	$\beta_2 \cdot \frac{X}{y}$
Log-log	$\ln y = \beta_1 + \beta_2 \ln X$	$\beta_2 \cdot \frac{y}{X}$	β_2
Log-lin	$\ln y = \beta_1 + \beta_2 X$	$\beta_2 \cdot y$	$\beta_2 \cdot X$
Lin-log	$y = \beta_1 + \beta_2 \ln X$	$\beta_2 \cdot \frac{1}{X}$	$\beta_2 \cdot \frac{1}{y}$
Reciprocal	$y = \beta_1 + \beta_2 \cdot \frac{1}{X}$	$-\beta_2 \cdot \frac{1}{X^2}$	$-\beta_2 \cdot \frac{1}{Xy}$

6.13. You are given the data in Table 6.7. Fit the following model to these data and obtain the usual regression statistics and interpret the results:

$$\frac{100}{100 - Y_i} = \beta_1 + \beta_2 \left(\frac{1}{X_i} \right)$$

TABLE 6.7

Y_i	86	79	76	69	65	62	52	51	51	46
X_i	3	7	12	17	25	35	45	55	70	120

6.13)

Y_i	X_i	$Y_i^* = \frac{100}{100 - Y_i}$	$X_i^* = \frac{1}{X_i}$	Y_i^{*2}	X_i^{*2}	$Y_i^* \cdot X_i^*$
86	3	7,143	0,333	51,020	0,111	2,381
79	7	4,762	0,143	22,676	0,020	0,680
76	12	4,167	0,083	17,361	0,007	0,347
69	17	3,226	0,059	10,406	0,003	0,190
65	25	2,857	0,040	8,163	0,002	0,114
62	35	2,632	0,029	6,925	0,001	0,075
52	45	2,083	0,022	4,340	0,000	0,046
51	55	2,041	0,018	4,165	0,000	0,037
51	70	2,041	0,014	4,165	0,000	0,029
48	120	1,923	0,008	3,698	0,000	0,016
TOTAL =		32,874	0,750	132,920	0,145	3,916

$\hat{y}_i^* = \hat{\beta}_1 + \hat{\beta}_2 \cdot X_i^*$

$\sum y_i^*$ $\sum x_i^*$ $\sum y_i^{*2}$ $\sum x_i^{*2}$ $\sum x_i^* y_i^*$ $n=10$

$$\sum x_i^{*2} = 0,145 - \frac{0,750^2}{10} = 0,0888$$

$$\sum x_i^* y_i^* = 3,916 - \frac{32,874 \cdot 0,750}{10} = 1,45$$

$$\sum y_i^{*2} = 132,92 - \frac{32,874^2}{10} = 24,85$$

$$\hat{\beta}_2 = \frac{1,45}{0,0888} = 16,33 \quad ; \quad \hat{\beta}_1 = \frac{32,874}{10} - 16,33 \cdot \frac{0,750}{10} = 2,06$$

$$\frac{100}{100 - \hat{y}_i} = 2,06 + 16,33 \cdot \frac{1}{X_i}$$



How do we interpret $\hat{\beta}_2$?

A unit increase in X_i^* , on the average, is expected to increase y^* by 16,33 units.

What about the relationship between X and y ?

We need to calculate the slope.

$$\frac{100}{100-y} = \beta_1 + \beta_2 \cdot \frac{1}{X}$$

$$\frac{dy}{dX} \cdot \frac{100}{100-y} = -\beta_2 \cdot \frac{1}{X^2} \Rightarrow \frac{dy}{dX} = -\beta_2 \cdot \frac{100-y}{100X^2}$$

Then, for example, what is the expected change in y when X increases from 40 to 41?

$$X_0 = 40 \Rightarrow \frac{100}{100 - \hat{y}_0} = 2,06 + 16,33 \cdot \frac{1}{40} = 2,468$$

$$100 - \hat{y}_0 = \frac{100}{2,468} = 40,5$$

$$\hat{y}_0 = 59,5$$

$$\text{Then; } \left. \frac{dy}{dX} \right|_{X_0=40} = -16,33 \cdot \frac{100 - 59,5}{100 \cdot 40^2} = -0,004$$

y is expected to decrease by 0,004 units.

By the same manner, we can estimate Elasticity of y with respect to X when $X=40$.

$$E = \frac{dy}{dX} \cdot \frac{X}{y} = -\beta_2 \cdot \frac{100-y}{100X^2} \cdot \frac{X}{y} = -\beta_2 \cdot \frac{100-y}{100Xy}$$

$$\hat{E}(X=40) = -16,33 \cdot \frac{100 - 59,5}{100 \cdot 40 \cdot 59,5} = -0,00278$$