Dummy variables — qualitative variables
indicator variables
categorical variables
binary variables
dichotomous variables

So far, both independent & dependent variables — quantitative in nature

Sometimes the variable of interest may be qualitative in nature.

Quantitative	Qualitative
Income	Gender
Expenditure	Region
Height	Category
Weight	Membership status
Marks	
→ continuous variation	limited variation ←

⇓

captured numerically by using DUMMY VARIABLES

A dummy variable can take one of two possible numeric values (usually 1 and 0) — one for the presence of an attribute & the other for its absence.

Suppose I wish to capture gender using dummy variables. Then one possible formulation

$$D_i = 1 \text{ if female}$$
$$\quad = 0 \text{ if male}$$

Once the dummy variable D has been defined it can be used in regression analysis just like the quantitative variables

ANOVA MODEL

— a model in which there are only dummy variables as explanatory variables.

Consider the model,

$$Y_i = \beta_1 + \beta_2 D_i + u_i.$$

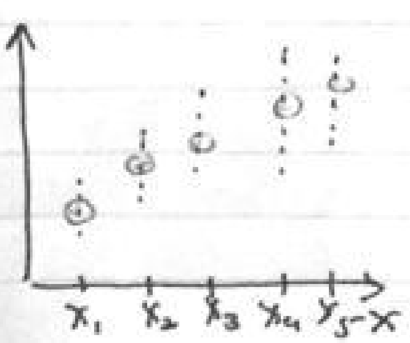
Y_i: Annual food expenditure (\$)

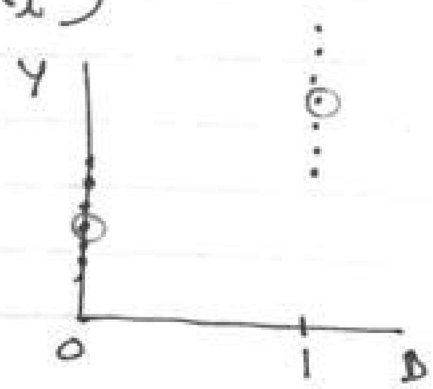
$D_i = 1$ if female
$\quad = 0$ if male.

The above model is saying that the variation in annual food expenditure can be explained by variation in gender.

So here there are 2 subpopulations — female + male (unlike the simple regression model where there were many subpopulations each corresponding to ~~ea~~ a different value of $\hat{X}_i$)

Simple Regression Model





ANOVA model

So like earlier, we can write an expression for the expected value of Y for each sub-population on the assumption that the error term satisfies the classical linear regression model (CLRM)

Simple regression	ANOVA
$Y_i = \beta_1 + \beta_2 X_i + u_i$	$Y_i = \beta_1 + \beta_2 D_i + u_i$
$E(Y_i/X_i) = \beta_1 + \beta_2 X_i$	$E(Y_i/D_i) = \beta_1 + \beta_2 D_i$
If $X_i = 10$ then $E(Y/X_i) = \beta_1 + 10\beta_2$	$E(Y_i/D_i=0) = \beta_1$
If $X_i = 12$ then $E(Y/X_i) = \beta_1 + 12\beta_2$	$E(Y_i/D_i=1) = \beta_1 + \beta_2$
+ so on for each possible value of X_i	Only 2 possible values of D_i

Coming back to the ANOVA model,

$E(Y_i/D_i=0)$ = Mean expenditure on food conditional on the observations belonging to the male sub population ($D_i = 0$)

$= \beta_1$

⇓

intercept of the regression.

$E(Y_i/D_i=1)$ = Mean expenditure on food conditional on observations belonging to the female sub population ($D_i = 1$)

$= \beta_1 + \beta_2$ = intercept + slope

$$E(Y_i / D_i = 0) = \beta_1 \qquad (1)$$
$$E(Y_i / D_i = 1) = \beta_1 + \beta_2 \qquad (2)$$

(2) − (1)

$\Downarrow$

$$E(Y_i / D_i = 1) - E(Y_i / D_i = 0) = (\beta_1 + \beta_2) - \beta_1$$
$$= \beta_2 .$$

$\Rightarrow$ slope gives the differential between the two mean food expenditures and is therefore called the differential intercept. coefficient.

In other words β_1 is the mean exp on food for ~~femes~~ males in the population while $\beta_1 + \beta_2$ is the mean exp on food for females.

A point to note is that of the two categories or sub populations, the one that is given the value 0 is called the base or reference category & in this case the male sub population is the reference category

Since β_2 is the difference between the two means in the population, it can be the basis of a testable hypothesis.

So two claims can be made about its value and can from our H_0 & H_a

$H_0: \beta_2 = 0$ $\quad$ $H_a: \beta_2 \neq 0$

H_0: No difference b/w the mean food exp of males + females

H_a: Difference b/w the mean food exp of males + females

Test Statistic: b_2 [estimatable using the OLS regression method]

$b_2 \sim N\left(\beta_2 \quad \frac{\sigma^2}{\sum d_i^2}\right)$ $\quad$ $b_2 = \frac{\sum y_i d_i}{\sum d_i^2}$

$$t = \frac{b_2 - \beta_2}{\sigma / \sqrt{\sum d_i^2}} \sim t_{n-2}$$

Choose α, compute $t_{\alpha/2, n-2}$, $-t_{\alpha/2, n-2}$

Find computed t

If t lies in the acceptance region — do not reject H_0

If t " " " rejection region — reject H_0.

If H_0 is rejected ⇒ significant diff b/w the mean food exp of males + females

If H_0 is not rejected ⇒ no significant diff b/w the mean food exp of males + females

Given data on annual food expenditure for males + females for 2000 — 2001 for 12 individuals, the estimate regression is

$$E(Y_i) = \beta_1 + \beta_2 D_i$$

$$\hat{Y}_i = \underset{b_1}{3176.833} - \underset{b_2}{503.1667} D_i$$

se $\quad (233.0446) \qquad (329.5749)$

t $\quad (13.6318) \qquad (-1.5267)$

$r^2 = 0.1890$

Y – food expenditure ($)
$D_i = 1$ if female
$\quad = 0$ if male

- From the above regression
  $b_1 = 3176.833$ is the OLS estimate of the mean food expenditure for males

  $b_2 = -503.1667$ is the difference between the male + female mean food expenditure

  $\Rightarrow$ female mean food expenditure estimate is $3176.833 - 503.1667 = 2673.6663$

- Thus Male mean food exp $= 3176.833$
  Female " " " $= 2673.666$

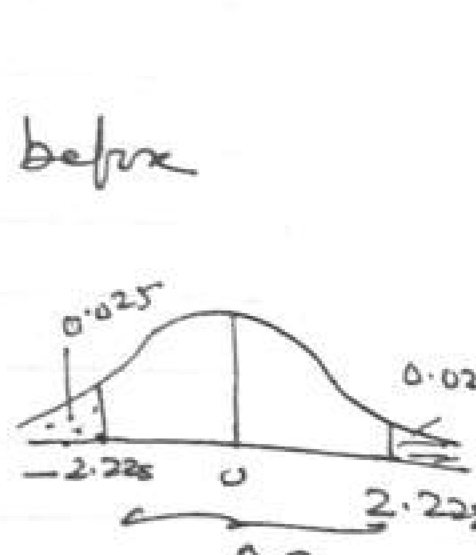
These are estimates + hence are random variables. They can be used to test the hypo:

$H_0: \beta_2 = 0 \qquad H_a: \beta_2 \neq 0$

using a t test as explained before

Computed $t = -1.526$

If $\alpha = 5\%$ then $t_{.025, 10} = 2.228$



Since computed t lies in the acceptance region, we do not reject H_0

Another way to carry out this test is by computing the p value of the computed t value

$p(-1.5267) = 0.15$

i.e 15%.

Thus can reject the null hypo only at significance levels higher than 15%.

Thus our t test indicates no significant difference between the male & female annual food expenditures.

It may be noted that the dummy variable regression is just a device to see if the two means are different. Looking at the data reveals that indeed the mean food exp for males is 3176.833 & that for females is 2673.663.

Another point to consider is what if the dummy variable was defined as

$D_i = 1$ if male
$= 0$ if female [reference category.

Since the data is the same, we must get the same values for estimated mean male food exp (3176.833) & estd mean female food exp (2673.663)

The estimated regression will now be

$$\hat{Y}_i = 2673.667 + 503.1667 D_i.$$

se $\quad (233.0446) \quad (329.5749)$

t $\quad (11.4227) \quad (1.5267) \qquad r^2 = 0.1890$

where $D_i = 1$ if male
$\quad = 0$ if female

As before the interpretation of the intercept is the mean food exp of the reference category which is now female

Thus now $b_1 = 2673.667$

The mean food exp of the other category is obtained as $b_1 + b_2$

$$\Rightarrow 2673.667 + 503.1667$$

$$= 3176.833$$

= mean food exp for males.

<u>Note:</u>

1. ~~E~~ r^2 remains unchanged
2. se's " "
3 t value changes.

1. Using data for 526 individuals the following model of wage determination was estimated

$$\text{Log } W_i = B_0 + B_1 D_i + B_2 EDU_i + B_3 (D_i * EDU_i) + u_i.$$

where W : daily wages in rupees
D : =1 for females
=0 for males
EDU : Years of education

The estimated =ⁿ is

$$\widehat{\text{Log } W_i} = \underset{(0.1190)}{0.3890} - \underset{(0.1680)}{0.2270} D_i + \underset{(0.0080)}{0.0820} EDU - \underset{(0.0131)}{0.0056} D_i^* EDU_i$$

std errors (0.1190) (0.1680) (0.0080) (0.0131)

a) Write the regression equations relating LOG W_i to EDU for males + females separately

b) The returns to education are measured by the % ↑ in wages due to an extra year of education. Using the results from (a) find the returns to education for males + females.

c) Is the difference between returns to education for males + females statistically significant at 5% level of significance?

2. To analyze the impact of takeovers on CEO compensation the following model was estimated

$$Comp_i = B_1 + B_2 SMP_i + B_3 D_i + B_4(D_i * SMP_i) + u_i$$

where

$Comp_i$ = Compensation in Rs hundreds

SMP : Index of stock market performance

D = 1 if a firm acquires another firm
= 0 otherwise

The following is the estimated equation using data on 34 firms

	Coefficient	Std error
Intercept	964.5202	69.1662
SMP	1868.567	288.0425
D	996.8745	111.9876
D*SMP	5157.474	545.9090

a) Using the regression results interpret the coefficients of D_i + $D_i * SMP_i$.

b) Test the hypothesis that the relation of compensation to the stock market performance remains the same irrespective of take-ov[er]s made by firms

3. Regression results for Korean savings–income data are presented for the period 1970–1995

$$\hat{Y}_i = 1.0161 + 152.4786 D_i + 0.0803 X_i + 0.0655 D_i X_i$$
$$t \quad (0.0504) \quad (4.6090) \quad (5.5413) \quad (-4.0963)$$

$$R^2 = 0.8819$$

where Y = savings X = Income

$D = 1$ for observations in 1982–95

$= 0$ " " 1970–81

a) Interpret the above regression results & obtain the regressions for the two periods 1970–81 & 1982–95.

b) What do you infer by the statistical significance of the differential intercept & the differential slope coefficients?

Functional Forms

Linear Regression $\Rightarrow$ linear in parameters.

Linear in variables

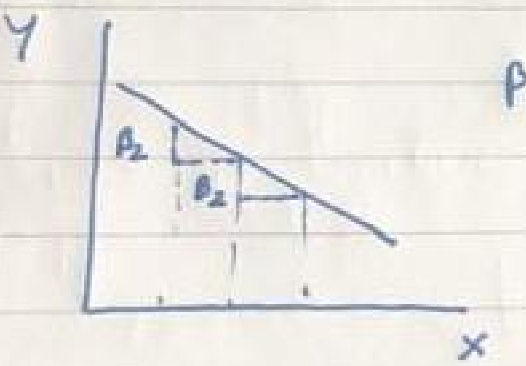
e.g $Y_i = \beta_1 + \beta_2 X_i + u_i$

$$\frac{dY_i}{dX_i} = \beta_2$$

$\downarrow$

constant

i.e. independent of X



$\beta_2 < 0$

Non-linear in variables

e.g $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$

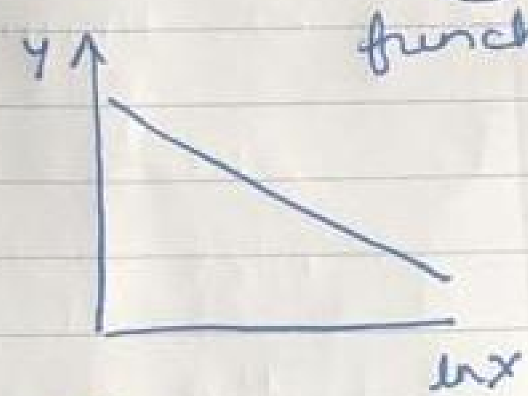
$$\frac{dY_i}{d\ln X_i} = \beta_2$$

$$\Rightarrow \frac{dY}{dX/X} = \beta_2$$

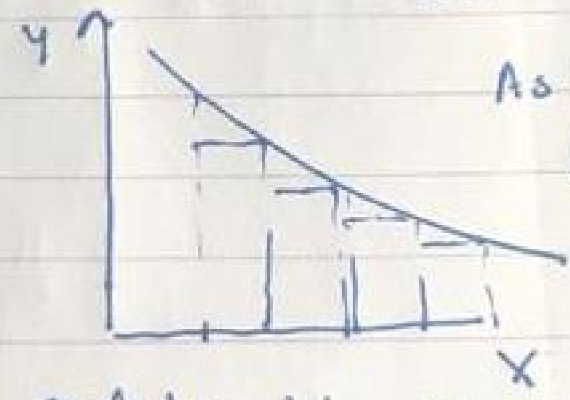
$$\Rightarrow \frac{dY}{dX} = \frac{\beta_2}{X}$$

$\Downarrow$

function of X

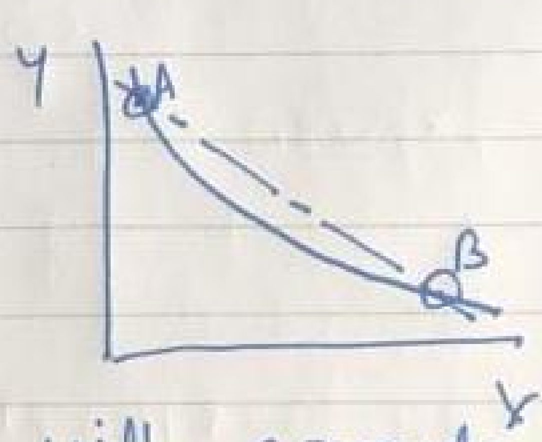


$\beta_2 < 0$.



As X $\uparrow$ $\frac{\beta_2}{X} \downarrow$.

So if we assume a linear relation b/w Y & X when actually



the fitted line will coincide with the true relation only twice (at A & B) as shown above.

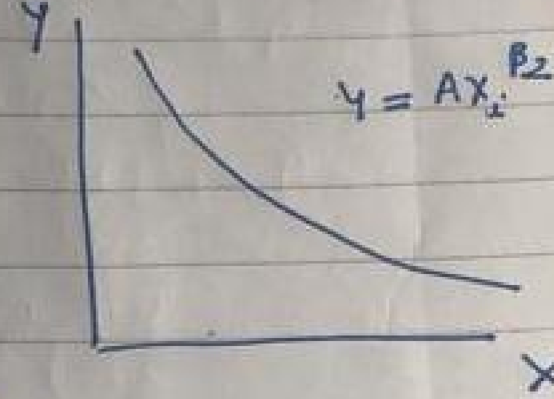
There are many relationships that are better postulated as non-linear relations b/w Y & X

$\Rightarrow$ we study some of them.

LOG-linear model

We are familiar with the Cobb Douglas function.

$$Y_i = A X_i^{\beta_2}$$



Since $\frac{dY}{dX} = \beta_2 A X_i^{\beta_2 - 1}$

$\downarrow$

function of X

$\therefore \downarrow$

Non-linear in X

By taking logs we can transform it into a linear in variables function

$$\log Y_i = \log A + \beta_2 \log X_i$$

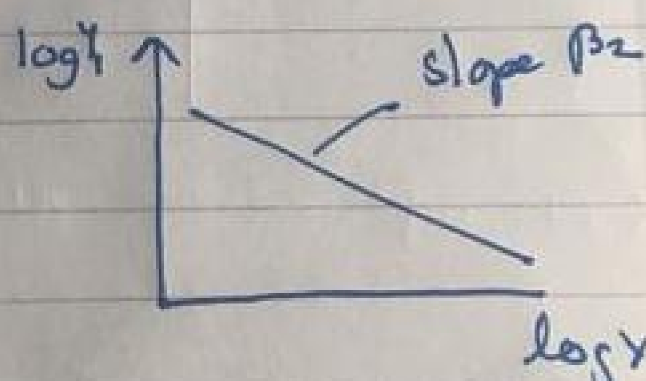
Adding an error term, the PRF becomes

$$\log Y_i = \log A + \beta_2 \log X_i + u_i.$$

$$\frac{d \log Y_i}{d \log X_i} = \beta_2$$

$\downarrow$

constant



This model is called the double-log model or the log-linear model as it is linear in the logs of the variables

Let $\log Y = Y^*$ & $\log X = X^*$ & $\log A = \beta_1$
Then:

$$Y_i^* = \beta_1 + \beta_2 X_i^* + u_i.$$

The above $=^n$ is estimated by OLS & if the assumptions of the CLRM are satisfied the estimates are BLUE

Now let us interpret the parameters

$$\beta_2 = \frac{dY_i^*}{dX_i^*}$$

$$= \frac{d(\log Y)}{d(\log X)} = \frac{dY/Y}{dX/X} = \frac{dY}{dX} \cdot \frac{X}{Y}$$

$\Downarrow$

elasticity of Y w.r.t X

Thus the slope coefficient estimate is an estimated elasticity.

— this model hypothesizes a constant elasticity. & so is also called a constant elasticity model

β_1 = value of Y_i^* when $\log X_i = 0$

SAT score eg:

$$\widehat{\ln Y_i} = 4.8877 + 0.1258 \ln X_i$$

log SAT score — log annual income

se	(0.1573)	(0.0148)
t	(31.074)	(8.5095)
p	(1.25×10^{-9})	(2.79×10^{-5})

$r^2 = 0.9005$

elasticity = 0.13

=> as annual income ↑s by 1%. Math SAT score on average ↑ by 0.13 %.

=> Math SAT score is inelastic

Comparing Linear & Log linear regⁿ models

— empirical question

guidelines

1) — plot the data
   — not possible in multiple reg

2) on the basis of r^2

prob'

need the dep var to be of the same form

— linear model Y

— log linear model $\ln Y$

The linear model measure Δ in absolute value of Y
whereas the log linear model measure the ~~Δ~~ proportional or % Δ

Even if dep var in the 2 models is the same so that r^2 are comparable
— not advisable to choose on the basis of a high r^2 criterion.
— need to carefully see the relevance of
LV, expected sign, stat significance
etc

In the LIV model elas^n varies from pt to pt
— avg elas^n = $\frac{\Delta Y}{\Delta x} \frac{\bar{X}}{\bar{Y}}$

$Y_i = A X_2^{\beta_2} X_3^{\beta_3}$

Multiple log-linear regression models

$\ln Y_i = \beta_1 + \beta_2 \ln X_{2i}^{\text{-lab}} + \beta_3 \ln X_{3i}^{\text{-cap}} + u_i$

partial elasticities

$\widehat{\ln Y_t} = 1.6524 + 0.3397 \ln X_{2t} + 0.8460 \ln X_{3t}$

se	(0.6062)	(0.1857)	(0.09343)
t	(−2.73)	(1.83)	(9.06)
p	(0.014)	(0.085)	(0.00)

$R^2 = 0.995$
$F = 1719.23$
(0.000^{**})

$b_2 + b_3 \rightarrow$ estimate of returns to scale

Q. **5.10**

$n = 11$

Model A :

$$\hat{Y}_t = 2.6911 - 0.4795 X_t$$
$$se \quad (0.1216) \quad (0.1140)$$
$$r^2 = 0.6628$$

Model B :

$$\ln \hat{Y}_t = 0.7774 - 0.2530 \ln X_t$$
$$se \quad (0.0152) \quad (0.0494)$$
$$r^2 = 0.7448$$

Y - cups of coffee consumed per day
X - price of coffee in dollars per pound

a) slope coeff
— interpretations

$\bar{Y} = 2.43 \quad \bar{X} = 1.11$

At these mean values what is the elasticity of price for model A

$$e = \frac{dY}{dX} \cdot \frac{\bar{X}}{\bar{Y}}$$

$$= -0.4795 \times \frac{1.11}{2.43} = -0.2190$$

1955 - 1974

S.11 (7)

$$\widehat{\ln Y_t} = -1.6524 + 0.3397 \ln \overset{lab}{X_{2t}} + 0.8460 \ln \overset{K}{X_{3t}}$$

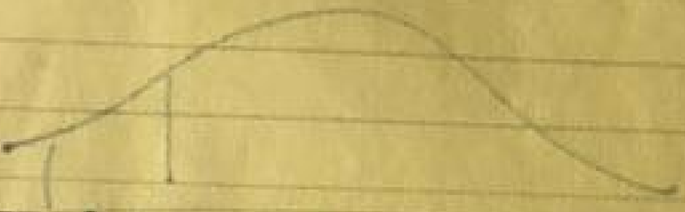
se (0.6062) (0.1857) (~~0.93~~ 0.09343)

t (−2.73) (1.83) (9.06)

p (0.014) (0.085) (0.000)

$R^2 = 0.995$

$F = 1719.93$



a) Interpret coeff.

b) Is coeff of X_2 stat diff fm 1

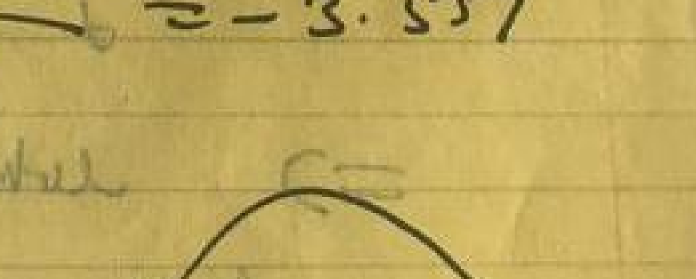
$H_0: B_2 = 1$ $H_1: B_2 \neq 1$

$$t = \frac{b_2 - 1}{se(b_2)}$$

$$= \frac{0.3397 - 1}{0.1857} = -3.557$$

$n = 20$

$df = n - 3 = 17$



$t_{c,0.05} = 2.110$

$t_{c,0.01}$ 2.898

Since $t_{computed}$ is greater than t_c

reject H_0: $B_2 = 1$

SAT score eg:

$$\widehat{\ln Y_i} = 4.8877 + 0.1258 \ln X_i$$

log SAT score; log annual income

se	(0.1573)	(0.0148)
t	(31.074)	(8.5095)
p	(1.25×10^{-9})	(2.79×10^{-5})

$r^2 = 0.9005$

elasticity = 0.13

=> as annual income ↑s by 1%. Math SAT score on average ↑ by 0.13 %.

=> Math SAT score is inelastic

Comparing Linear & Log linear regn models

— empirical question

guidelines

1) — plot the data
— not possible in multiple regn

2) on the basis of r^2

prob:

need the dep var to be of the same form

— linear model Y

— log linear model ln Y

Lin – Log Model

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

$$\beta_2 = \frac{dY_i}{d\ln X_i} = \frac{dY_i}{dX_i/X_i} = \frac{\text{absolute } \Delta \text{ in } Y}{\text{proportionate } \Delta \text{ in } Y}$$

Dividing both sides by 100

$$\frac{\beta_2}{100} = \frac{dY_i}{\frac{dX_i}{X_i} \times 100} = \frac{\text{absolute } \Delta \text{ in } Y}{\% \ \Delta \text{ in } Y}$$

So consider the following

$$Y_t = \beta_1 + \beta_2 \ln X_{2t} + u_t$$

Y_t ↓ expenditure on services

X_{2t} ↓ Total personal consumption expenditure

This was estimated for the period (1975–2006)

$$\hat{Y}_t = -12564.8 + 1844.22 \ln X_t$$

Thus $b_2 = 1844.22$

When $\ln X_t$ ↑ by 1 $\quad Y$ ↑ by 1844.22

When X ↑ by 1% $\quad Y$ ↑ by $\frac{b_2}{100} = 18.4422$

$$\frac{\frac{d\,(\text{US pop})}{\text{US pop}}}{dt} = 0.0107$$

Multiply both sides by 100

$$\frac{\frac{d\,(\text{US pop}^n)}{\text{US pop}} \times 100}{dt} = 0.0107 \times 100$$

$$= 1.07\% \text{ per year}$$

So the US pop^n grows at the rate of 1.07% per year.

called the instantaneous growth rate

Thus in general terms

$$\frac{d \log Y_t}{dt} = \beta_1$$

$$\frac{dY_t / Y_t}{dt} = \beta_1$$

$$\frac{\frac{dY_t}{Y_t} \times 100}{dt} = \beta_1 \times 100$$

rate of growth of Y per year.

Semi-log models are also known as growth models.

Going back to our example

$$b_0 = 5.3593$$

$$\log Y_0 = 5.3593$$

Taking antilogs on both sides

$$Y_0 = \text{antilog } 5.3593 = 212.5761$$

Linear Trend Models

$$Y_t = \beta_1 + \beta_2 t + u_t$$

If $\beta_2 > 0 \Rightarrow$ upward trend

$\beta_2 < 0 \Rightarrow$ downward trend

$$\widehat{US\ pop^n}_t = 209.6731 + 2.7570\, t$$

per year the US $pop^n \uparrow^s$ by an absolute amt of 2.757 mln.

Thus for $t=0$ the popⁿ was
212.5761 → initial population

Also $b_1 = \ln(1+r)$
Taking antilogs
antilog $b_1 = 1+r$
$r =$ antilog $b_1 - 1$
$= 1.0108 - 1 = 0.010757$.

i.e. the compound growth rate of growth of the US popⁿ is 1.0757% p.a

Another possible example for semi-log model
– relation between earnings or wages & age or years of service.

Because typically contracts specify that for each add^al year of service, a worker gets a certain % ↑ in his/her earnings & regⁿ analysis may be used for this purpose as shown above

$$\widehat{\ln \text{earnings}} = 2.453 + 0.0128 \text{ Experience}$$
↓
For every add^al yr of experience earnings ↑ by 1.28%.

Semi-log model (Log-lin model)

→ regression analysis used to measure the growth rate of economic variables

Consider the following

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

Y_t: pop^n at time t

Y_0: pop^n at time 0 i.e. initial population

$(1+r)^t$: compound rate of growth.

taking logs

$$\log Y_t = \log Y_0 + t \log(1+r)$$

$$\log Y_t = \beta_0 + \beta_1 t$$

where $\beta_1 = \log(1+r)$

Adding the error term

$$\log Y_t = \beta_0 + \beta_1 t + u_t$$

t takes values 1, 2, 3 etc.

Using data for US pop^n 1975-2007 the following reg^n results were obtained

$$\widehat{\log (\text{US pop}^n)} = 5.3593 + 0.0107t$$

$$\frac{d\log(\text{US pop})}{dt} = 0.0107$$

Linear Trend Models

$$Y_t = \beta_1 + \beta_2 t + u_t$$

If $\beta_2 > 0 \Rightarrow$ upward trend

$\beta_2 < 0 \Rightarrow$ downward trend

$$\widehat{\text{US pop}^n}_t = 209.6731 + 2.7570\,t$$

per year the US pop^n ↑s by an absolute amt of 2.757 mln.

Logarithms in Regⁿ

Case	Regⁿ Spfⁿ	Interpretation of B_2
I.	$Y_i = B_1 + B_2 \ln X_i + u_i$.	A 1% Δ in X is associated with a Δ in Y of $0.01B_2$ i.e. $B_2/100$
II	$\ln(Y_i) = B_1 + B_2 X_i + u_i$	A Δ in X by 1 unit is associated with a $100B_2$ % Δ in Y
III	$\ln(Y_i) = B_1 + B_2 \ln(X_i) + u_i$.	A 1% Δ in X is ass with a B_2 % Δ in Y so B_2 = elasⁿ of Y wrt X.

$\beta_2 = \frac{dY}{dX} = \frac{dY}{\frac{dX}{X}}$

dY

$\frac{B_2}{100}$ β_2

β_2. $\frac{x}{100}$

So let us transform our data to create a new variable $Z_i = \frac{1}{X_i}$

+ then regress Y_i on Z_i

i.e. Y_i on $\frac{1}{X_i}$

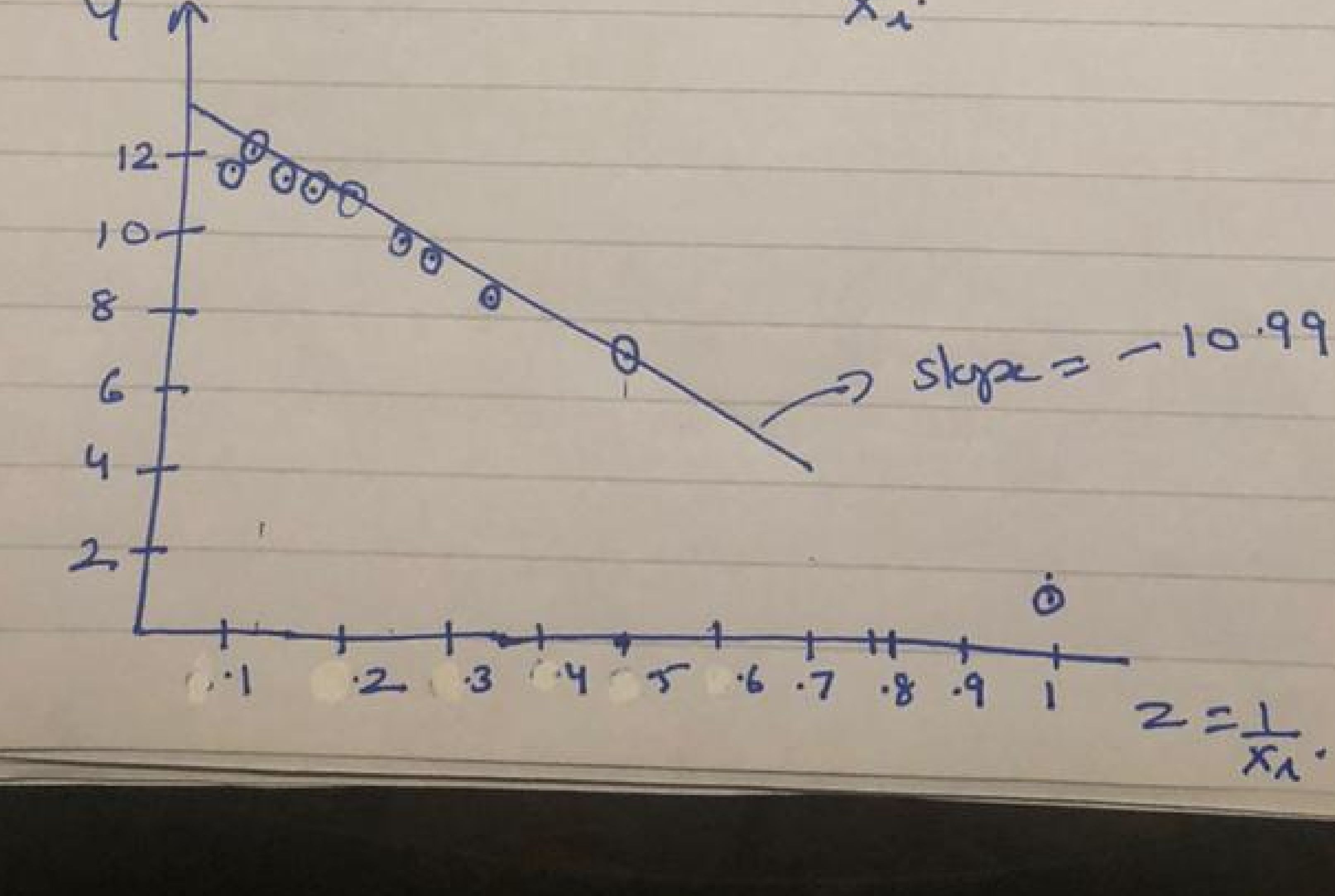
$\Downarrow$

Reciprocal model

Y	X	$Z_i = \frac{1}{X_i}$	$\hat{Y}$	e_i
1.71	1	1.000	1.49	0.22
6.88	2	0.500	6.985	−0.105
8.25	3	0.333	8.817	−0.567
9.52	4	0.250	9.733	−0.213
9.81	5	0.200	2.198	7.612
11.43	6	0.167		
11.09	7	0.143		
10.87	8	0.125		
12.15	9	0.111		
10.94	10	0.100		

$\therefore \quad \hat{Y} = 12.48 - 10.99Z$

i.e. $\hat{Y} = 12.48 - 10.99 \frac{1}{X_i}$



If we fit the model

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \text{(linear in variables)}$$

$$\hat{Y}_i = \underset{(1.26)}{4.62} + \underset{(0.20)}{0.84} X_i \qquad R^2 = 0.69$$

Residual analysis shows that the errors are non random
— looks like a curve would fit the data better
— a curve that shows the ΔY decreasing for successive ΔX's

Consider running the model

$$Y_i = \beta_1 + \frac{\beta_2}{X_i} + u_i$$

$$\frac{dY_i}{dX_i} = \frac{-\beta_2}{X_i^2}$$

If $\beta_2 < 0$

then $\dfrac{dY_i}{dX_i} = \dfrac{\text{positive constant}}{X_i^2}$

i.e rate of change of Y as X Δs is a function of X + a decreasing f^n of X.

As X $\uparrow$ $\dfrac{dY}{dX_i}$ $\downarrow$.

IL $\Rightarrow$ and this seems to be a good way to describe the dashed curve in Fig 1. $\rightarrow$ a concave curve

$\Rightarrow$ the $\uparrow$ in Y for successive equal increments in X is smaller + smaller

Much better fit.

Comparing the two models

$$\hat{Y}_i = 4.62 + 0.84X_i$$

& $$\hat{Y}_i = 12.48 - 10.99\frac{1}{X_i}$$

The sign of b_2 is negative in the second eqn while it is positive in the first.

But both are in agreement with the data where consn of bananas ↑s as income ↑s. ($\because$ as $X_i \uparrow$ $\frac{1}{X_i} \downarrow \Rightarrow Y \downarrow$)

A negative relation b/w Y & $\frac{1}{X_i}$

$$\Updownarrow$$

positive relation b/w Y & X_i

A key feature of Reciprocal Models

→ there is an asymptotic or limiting value of the dependent variable (Y_i) in them

Consider

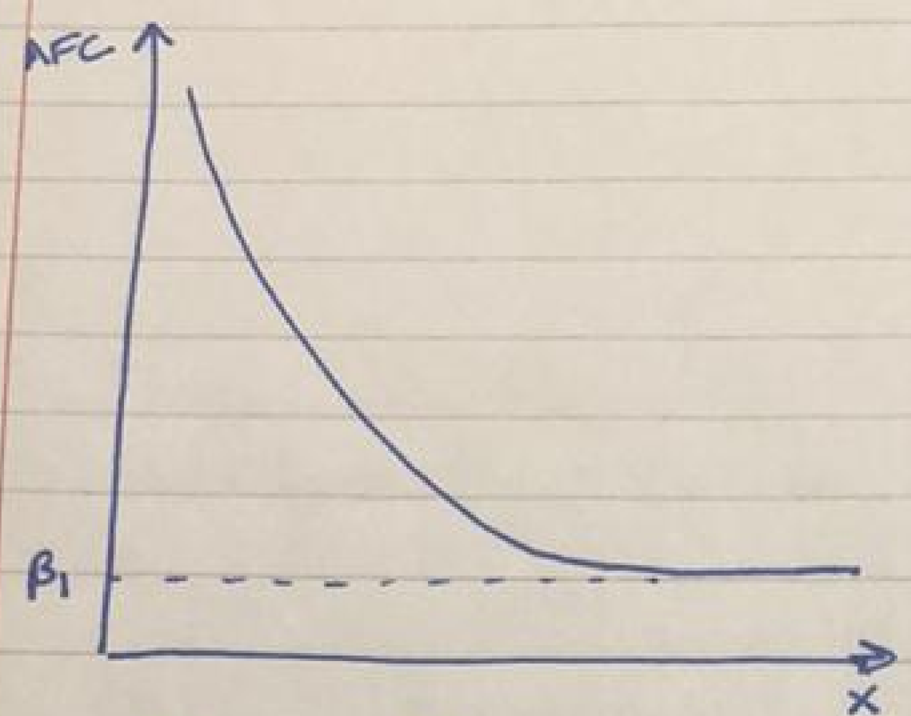
$$Y_i = \beta_1 + \beta_2\frac{1}{X_i} + u_i.$$

As $X_i \uparrow$ $\frac{1}{X_i} \rightarrow 0$ $\Rightarrow Y_i \rightarrow \beta_1$

(asymptotic value of Y)

Other examples of reciprocal models.

1) Y — AFC $\quad X$ — output



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

$$\frac{dY}{dX} = \frac{-\beta_2}{X_i^2}$$

Asymptotic value β_1

Since $\frac{dY}{dX} < 0$

$\Rightarrow -\frac{\beta_2}{X_i^2} < 0$

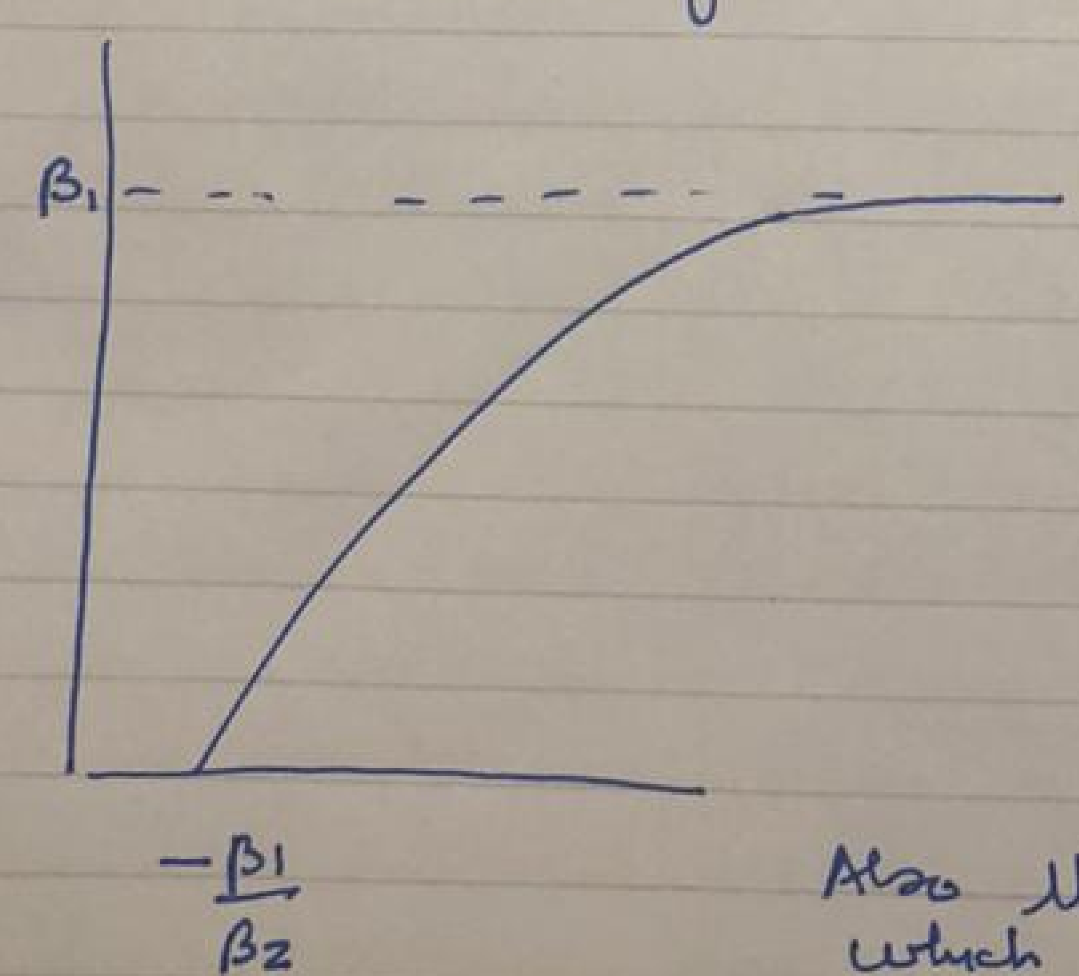
$\Rightarrow \beta_2 > 0$

and $\beta_1 > 0$

slope depends on X + ↓s as X ↑s

2) Y — expenditure on a commodity $\quad X$ — income

Like the banana-income example



$\frac{dY}{dX} > 0$

$\Rightarrow -\frac{\beta_2}{X_i^2} > 0$

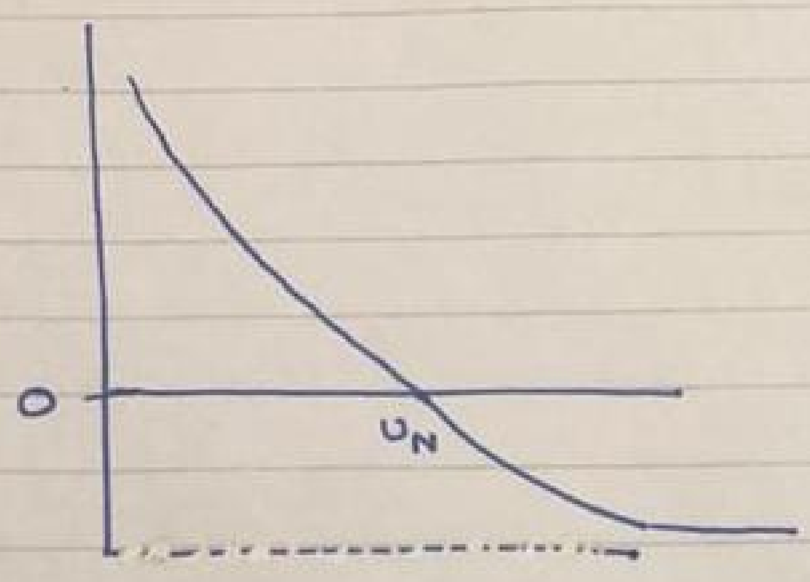
$\Rightarrow \beta_2 < 0$.

and $\beta_1 > 0$

Also the income level below which exp on the commodity is zero

$0 = \beta_1 + \beta_2 \frac{1}{X_i} \Rightarrow X_i = -\frac{\beta_1}{\beta_2}$

3) Y – % Δ in index of hourly earnings
X – civilian unemployment rate



$\frac{dY}{dX} < 0$

$\frac{-\beta_2}{X_i^2} < 0$

$\Rightarrow \beta_2 > 0$

and $\beta_1 < 0$

$\Downarrow$

as unemployment rate $\uparrow$s the % Δ in index of hourly earnings $\downarrow$s but at a decreasing rate till it reaches its asymptotic value of β_1. Further $\uparrow$s in the unemployment rate do not $\downarrow$ the index of hourly earnings

$\Rightarrow \beta_1$ – is the wage floor.