AGENDA

- "LINEAR" — LINEAR IN THE VARIABLES / LINEAR IN THE PARAMETERS
- BASIC IDEA ABOUT LINEAR REGRESSION ANALYSIS

LINEAR IN THE VARIABLES

- THE REGRESSION CURVE IS A STRAIGHT LINE.
- X'S ARE RAISED TO THE FIRST POWER ONLY.

$$E(Y|X_i) = \beta_1 + \beta_2 X_i .$$

EX: $Y = \beta_1 + \beta_2 X .$

LINEAR IN THE PARAMETERS

- THE CONDITIONAL MEAN OF Y IS A LINEAR FUNCTION OF THE PARAMETERS, β's.

EX:

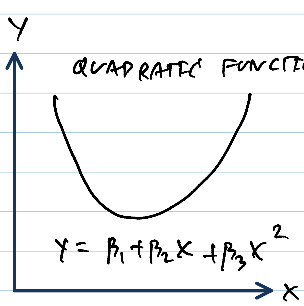
$$Y = \beta_1 + \beta_2 X + \beta_3 X^2$$

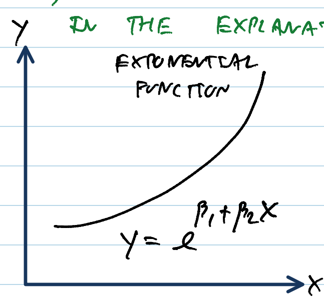
$$Y = \beta_1 + \beta_2 X + \beta_3 X^2 + \beta_4 X^3$$

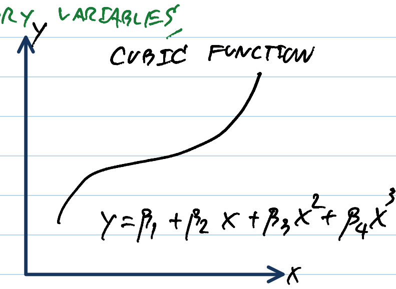
$$Y = e^{\beta_1 + \beta_2 X} .$$

> LINEAR REGRESSION MODEL REFERS TO A MODEL THAT IS LINEAR IN PARAMETERS. IT MAY NOT BE LINEAR IN VARIABLES.

i.e., MAY OR MAY NOT BE LINEAR IN THE EXPLANATORY VARIABLES







EX: TOTAL COST CURVE IN THE SHORT RUN PRODUCTION

ALL FUNCTIONS HERE ARE LINEAR IN THE PARAMETERS, β's. AS A RESULT, THEY ARE LINEAR REGRESSION FUNCTIONS (OR MODELS).

LINEAR IN THE VARIABLES

	YES	NO
[illegible]	[illegible]	[illegible]

		YES	NO
LINEAR IN THE PARAMETERS	YES	LRM	LRM
	NO	NLRM	NLRM

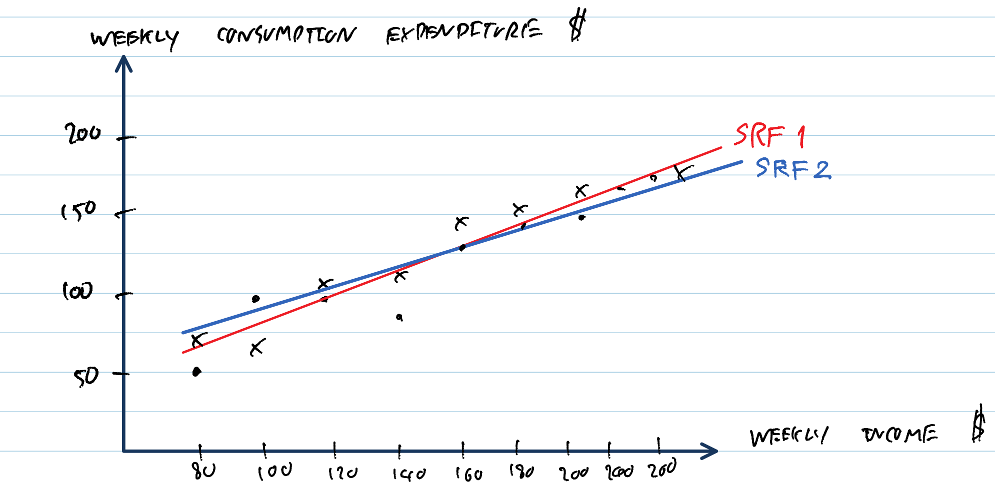
LRM = LINEAR REGRESSION MODEL

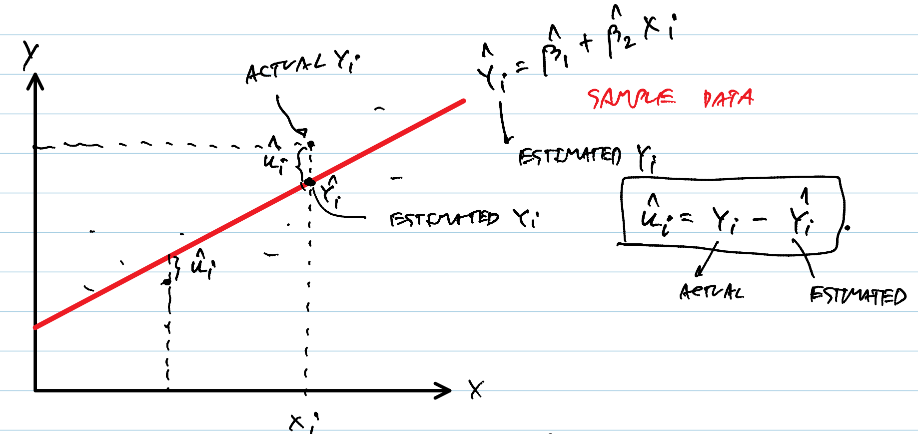
NLRM = NON LINEAR REGRESSION MODEL

(EX : $Y = \beta_1 + \beta_2^2 X$ → NON LINEAR IN PARAMETER & LINEAR IN VARIABLE

$Y = \beta_1 + \beta_2^2 X^3$ → NON LINEAR IN BOTH PARAMETER AND VARIABLE

• BASIC IDEA ABOUT LINEAR REGRESSION MODEL.





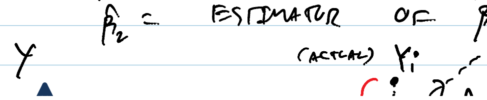
SAMPLE REGRESSION FUNCTION (SRF)

$$\boxed{\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i}$$

$\hat{Y}_i$ = ESTIMATOR OF $E(Y_i | X_i)$

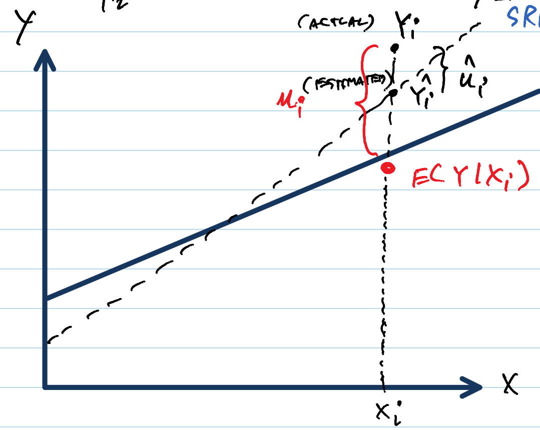
$\hat{\beta}_1$ = ESTIMATOR OF β_1

$\hat{\beta}_2$ = ESTIMATOR OF β_2



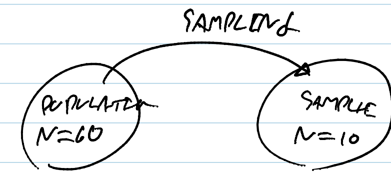
SRF: $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_i$

$\hat{\beta}_2$ = ESTIMATOR OF β_2



SRF: $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_i$

PRF: $E(Y|X_i) = \beta_1 + \beta_2 X_i$
(WE DON'T KNOW THIS LINE)



PRF: $Y_i = \beta_1 + \beta_2 X_i + u_i$

SRF: $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$

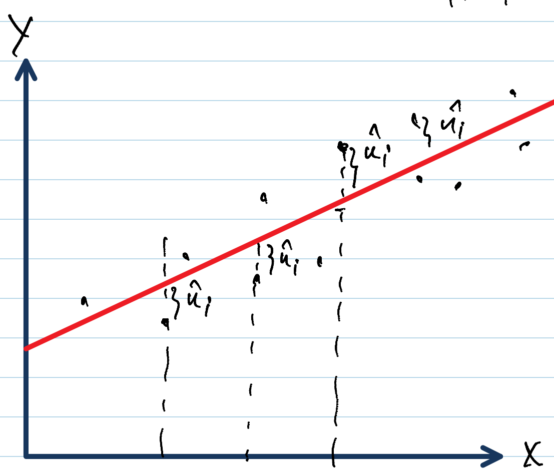
TASK: ESTIMATE PRF FROM SRF!

BASED ON PRF: $Y_i = E(Y|X_i) + u_i$

OR $Y_i = \beta_1 + \beta_2 X_i + u_i$

BASED ON SRF: $Y_i = \hat{Y}_i + \hat{u}_i$

OR $Y_i = (\hat{\beta}_1 + \hat{\beta}_2 X_i) + \hat{u}_i$



SAMPLE

BEST FITTED LINE

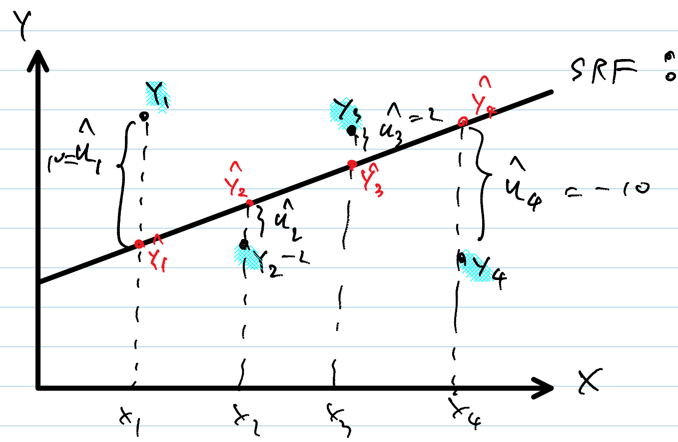
WE WANT TO ESTIMATE β_1 AND β_2 FROM SRF:

$$Y_i = \underbrace{\hat{\beta}_1 + \hat{\beta}_2 X_i}_{\hat{Y}_i} + \hat{u}_i$$

$$Y_i = \hat{Y}_i + \hat{u}_i$$

$$\hat{u}_i = Y_i - \hat{Y}_i$$

$$\hat{u}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$$



SRF : $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

OBJ : WE WANT THE LINE TO BE CLOSE TO Y_i AS POSSIBLE.

WHAT WOULD BE THE RULE OR CLEVER METHOD FOR THIS OBJ?

PROPOSAL #1 : ~~CHOOSE THE~~ LINE SUCH THAT

$$\cancel{\sum_{i=1}^{n} \hat{u}_i = 0}.$$

PROPOSAL #2 :

$$\sum_{i=1}^{n} \hat{u}_i^2 = 0$$

(SUM OF SQUARED ERRORS)

PRINCIPLE OF ORDINARY LEAST SQUARES (OLS)

FIND ESTIMATOR $\hat{\beta}_1$ AND $\hat{\beta}_2$ SUCH THAT

THE SUM OF SQUARED ERRORS IS THE LEAST :

$$\min \sum_{i=1}^{n} \hat{u}_i^2 \quad \text{OR} \quad \min \hat{u}_i^2 (\hat{\beta}_1, \hat{\beta}_2)$$

$$\min \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$\min_{(\hat{\beta}_1, \hat{\beta}_2)} \sum_{i=1}^{n} (Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_i))^2$$