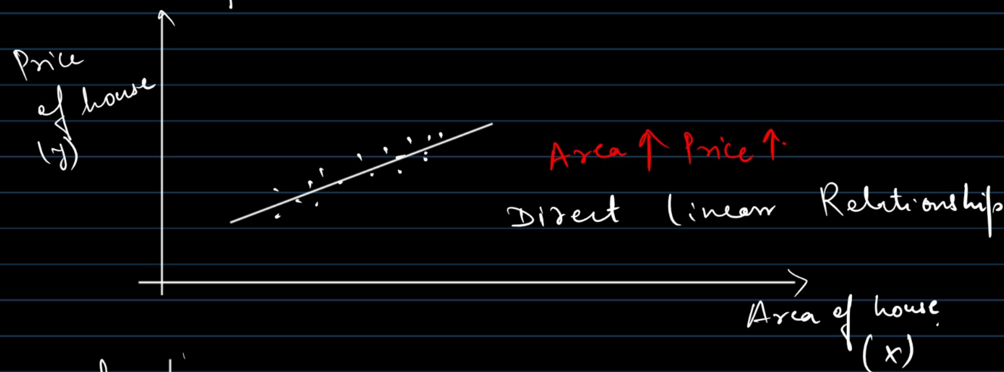
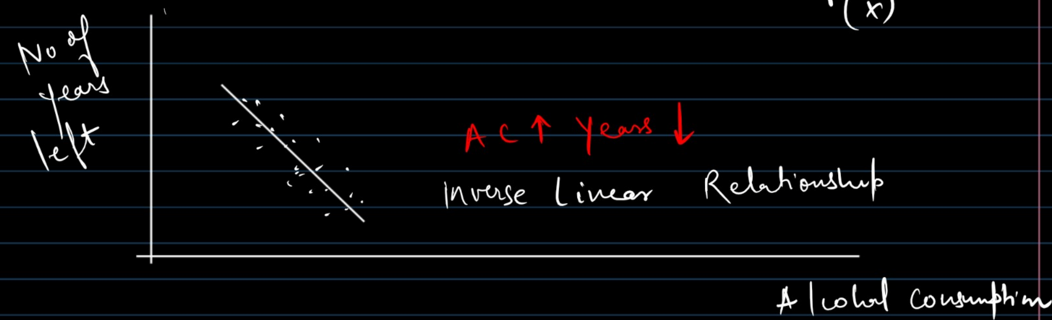
Assumption of linear Regression

① **Linearity** — X & Y should have linear relationship.





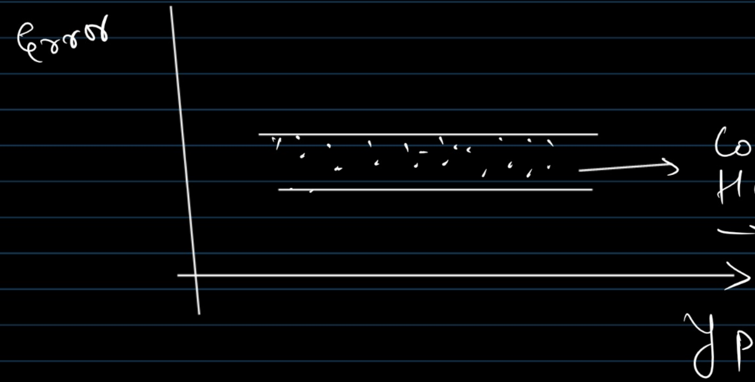
② **Independence** :- The observations (rows) are independent of each other.

Ht	Wt
5.2	80
6.1	90
-	-
-	-
-	-

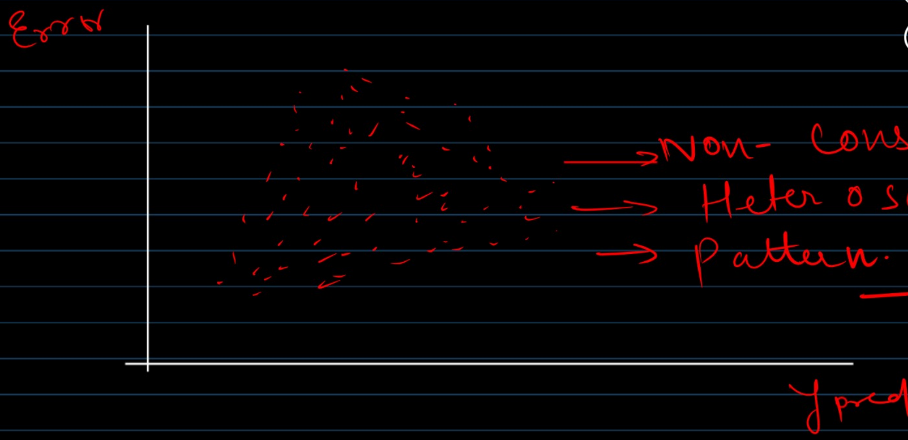
# of hours studied	y_{act} Marks obtained	y_{pred}
A 5	50	45
B 4	80	75
-	-	-
-	-	-

↳ Since the observations are independent, the errors should be independent.

③ **Homoscedasticity** → Also known as constant variance, this assumption means the variance of the errors are constant.



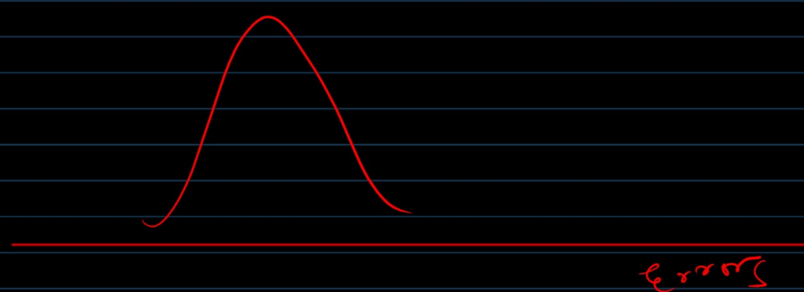
→ Constant variance
Homoscedasticity
→ No pattern



→ Non- constant Variance
→ Heteroscedasticity
→ pattern.

④ <u>Normality of Errors</u>

- Errors should be normally distributed.



⑤ The features should not be related or should have least relation. — <u>multicollinearity</u>

f_1 f_2 f_3 | y