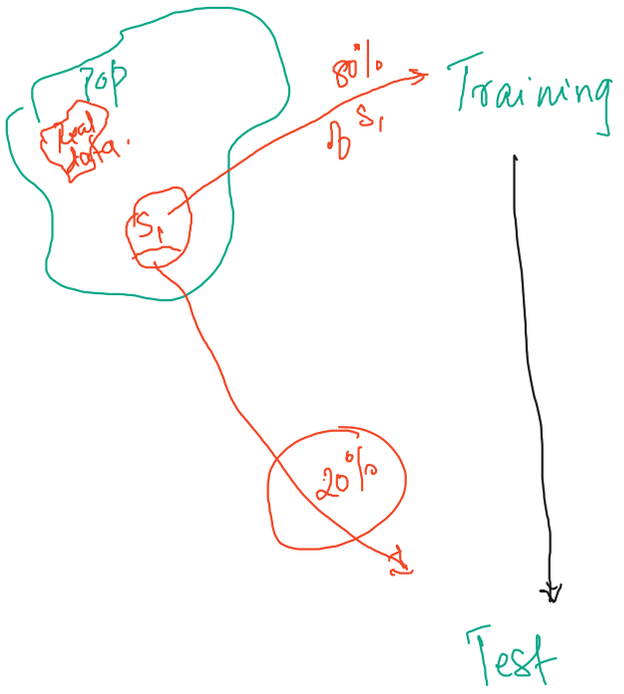


$\checkmark Y = \checkmark f(\checkmark X)$

?

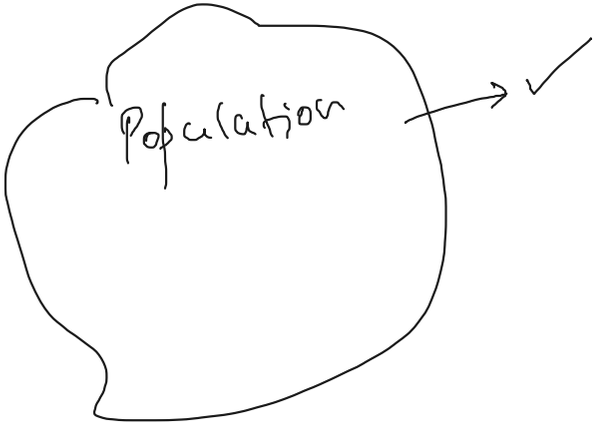
- Regression model ... NEURAL NET
- classification " ... Linear Regression
- clustering

$\checkmark Y = \checkmark \beta_0 + \checkmark \beta_1 \checkmark X$

?

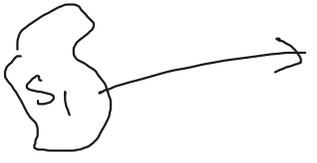
$\checkmark Y = \checkmark f(\checkmark X)$

EXPECTED VALUE:



- mean
- std dev
- Variance

} PARAMETERS

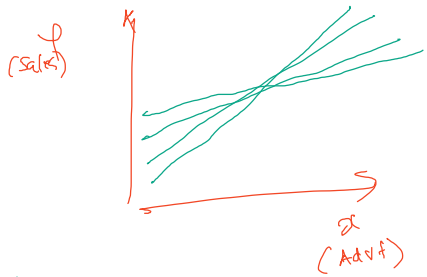


- mean
- std dev
- variance

} STATISTICS

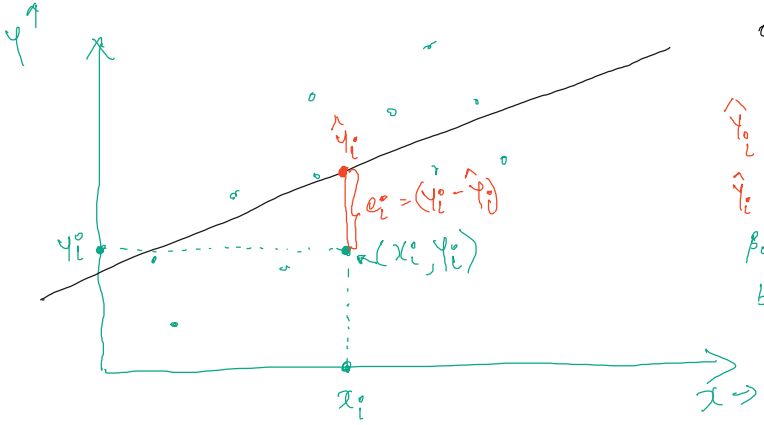
Using statistics we want to ESTIMATE Parameters.

There will always be ERRORS of Estimation & it depends on the Sample size 'n'



S_1, S_2, S_3, S_4

Different samples will result in different regression models...



obs → S_1

$\hat{y}_i = \beta_0 + \beta_1 x_i$ (or) $\hat{y}_i = b + a x_i$

β_0 & β_1 (or b & a) — Regression Coefficients

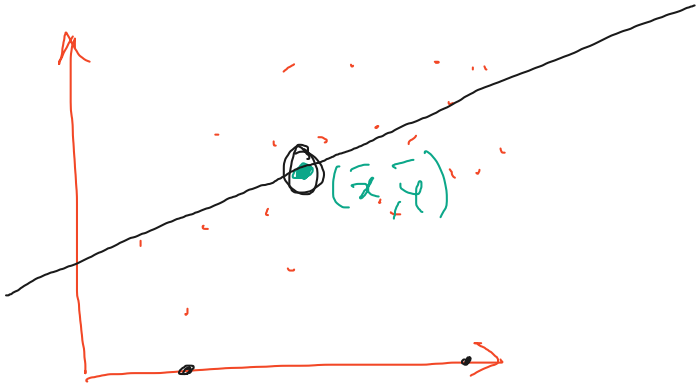
$\sum e_i = 0$ $\quad \sum e_i^2 = SSE$

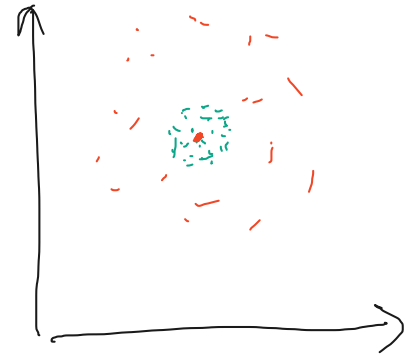
find b & a such that $\min(SSE)$ — objective.

$$a = \frac{\overline{xy} - \bar{x}\bar{y}}{(\overline{x^2} - \bar{x}^2)}$$

$$b = \frac{\bar{y}\overline{x^2} - \bar{x} \cdot \overline{xy}}{\overline{x^2} - \bar{x}^2}$$

Derivation uploaded to Moodle.





Simplest Model = Mean

IDEA !

Can I use the mean as a model.

- Given (x_i^o, y_i^o)
- Find out $\bar{y}$
- For all x_i^o the prediction is $\bar{y}$

$\Rightarrow$ I am using the 'mean' as the model.