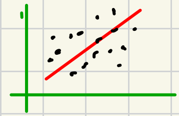
→ video in week 3 on Non-Parametric Density estimation

Linear Regression.

we assume $\varepsilon \sim N(0, \sigma^2)$

average distance of the pts to the line are 0



$$E(Y) = E[\beta_0 + \beta_1 x + \varepsilon]$$
$$= E[\beta_0 + \beta_1 x] + E[\varepsilon] \rightarrow = 0$$
$$= \beta_0 + \beta_1 x$$

→ x is not a R.V here

$$E[Y|X] = \beta_0 + \beta_1 x$$

assume ε & $\beta_0 + \beta_1 x$ to be independent of each other

$$\text{var}(Y) = \text{Var}(\beta_0 + \beta_1 x + \varepsilon)$$
$$= \text{var}(\beta_0 + \beta_1 x) + \text{var}(\varepsilon)$$
$$\text{var}[Y|X] = \text{var}(\varepsilon)$$

3) From a Sample $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ of N observations

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Can find β_0 & β_1 using least squares or MLE

Least Squares:

$$(\hat{\beta}_0, \hat{\beta}_1) = \arg\min_{\beta_0, \beta_1} \sum_{i=1}^{N} \left(y_i - (\beta_0 + \beta_1 x_i)\right)^2$$

↳ gives the value of β_0 & β_1 that minimise the error

Ex (cars) Assume $\beta_0 = 0$

$$y_i = \beta_1 x_i + \varepsilon_i$$
$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{N} (y_i - \beta x_i)^2 \quad \text{is s.t:}$$

take derivative.

$$-\sum_{i=1}^{N} 2x_i(y_i - \hat{\beta} x_i) = 0$$

(i.e) $\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$ Least Squares solution

R.V

A variable to take on a numerical value to represent a random event

A variable whose possible values are numerical outcomes to represent a random event.

So you have least squares & MLE to find the best parameters for a distribution

A MLE maximises the likelihood of the parameters that fit the data.

↳ So the parameters are the best parameters for the dist^n