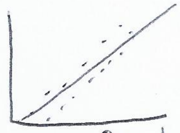
Lec 16 4/6/16 1)

OLS Assumptions

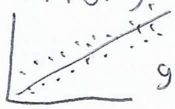
1. Linearity



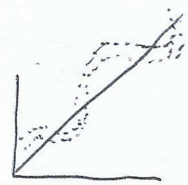
good

bad

2. Independence of errors



good



errors follow a pattern

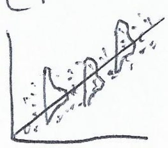
bad

3. Homoscedasticity

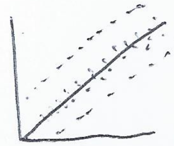
good

bad

4. Normality of Error Distribution

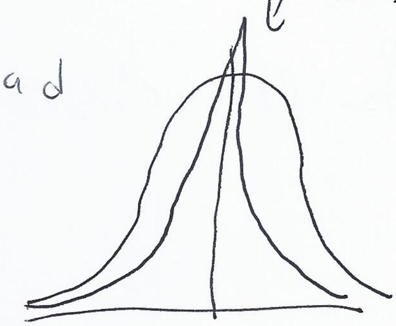


good



bad

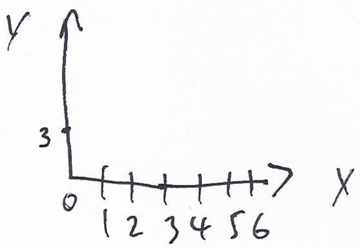
laplace distribution



OLS model

$$Y_i | X_i = x_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$$

all indep $\quad i \in \{1, \dots, n\}$



$$\underbrace{\substack{1 + 2x \\ \beta_0 + \beta_1 X}}_{f(x)} \qquad \sigma^2 = 1 \qquad n = 6$$

$$\theta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \sigma^2 \end{bmatrix}$$

$$\theta, \sigma^2 | X \propto p(\theta | \sigma^2, X)\, p(\sigma^2 | X) \quad P(\theta | \sigma^2) P(\sigma^2)$$

est target $\quad$ nuisance parameter

$$P(\theta | X) = \int_{\text{supp}[\sigma^2]} p(\theta, \sigma^2 | X)\, d\sigma^2 = T(\quad , \quad)$$

$$Y = \alpha + B_1 X + \gamma_1 z_1 + \gamma_2 z_2 + \dots + \gamma_p z_p$$ (no epsilon)

$z_1, \dots, z_p$ are noise

$$E[Y|X] = B_1 X + \underbrace{\left(\alpha + \sum_{i=1}^{p} \gamma_i E[z_i]\right)}_{B_0}$$

CLT

$$Var[Y|X] = \underbrace{\sum_{i=1}^{p} \gamma_i^2 \, var[z_i]}_{\sigma^2}$$

normality

MLE for B_0, B_1?

$(x-\mu)$

$$\mathcal{L}(B; B_0, \sigma^2; x, y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - B_0 - B_1 x_i)^2}$$

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - B_0 - B_1 x_i)^2}$$

$$\ell(\downarrow) = n\ln(\downarrow) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - B_0 - B_1 x_i)^2$$

$$\frac{\partial}{\partial B_1}[\downarrow] = -\frac{1}{\cancel{2}\sigma^2}\left(\sum_{i=1}^{n} -\cancel{2}x_i(y_i - B_0 - B_1 x_i)\right) = 0$$

$$\sum x_i y_i - B_0 \sum x_i - B_1 \sum x_i^2 = 0$$

$$B_1(S_x^2 + n\bar{x}^2) - 2nxy - 2n\bar{y}\bar{x} + 2B_0\bar{x}n = 0$$

$$\frac{\partial}{\partial B_0}[\ell()] = -\frac{1}{\cancel{2}\sigma^2}\left(-\cancel{2}\sum y_i - B_0 - B_1 x_i\right) = 0$$

$$\bar{y}n - nB_0 - nB_1\bar{x} = 0$$

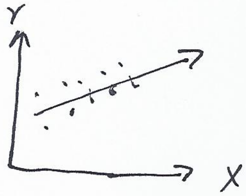
Same equation as algrithum $\sum \varepsilon_i^2$

MLE's = Least Square estimate

$B_1 = R\frac{S_x}{S_y} \sim ?$

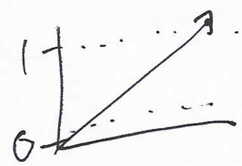
$B_0 = \bar{Y} - R\frac{S_x}{S_y}\bar{X} \sim ?$

$Y = f(x) + \epsilon = B_0 + B_1 X + \epsilon$

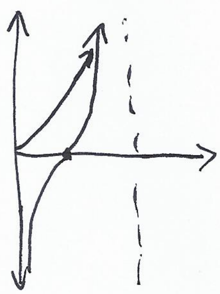


$Y \in \{0, 1\}$

$Y \in \mathbb{R}$



$\underbrace{P(Y=1|X)}_{\text{Target}}$ $\quad Y=1|X \sim \text{Bern}(\ )$



p $\quad \rightarrow P(Y=1|X)$

$p \in (0,1)$

$\frac{p}{1-p} \in (0, \infty)$

$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) \in (-\infty, \infty)$

link function

$\text{logit}(P(Y=1|X)) = B_0 + B_1 X$ (no epsilon)

$P(Y=1|X) = \text{logit}^{-1}(\ \downarrow\ )$

$= \frac{e^{B_0 + B_1 x}}{1 - e^{B_0 + B_1 x}}$

$\ln\left(\frac{p}{1-p}\right) = B_0 + B_1 X$

$\frac{p}{1-p} = e^{B_0 + B_1} \Rightarrow p = e^{B_0 + B_1 X} + p e^{B_0 + B_1 X}$

$p(1 - e') = e'$

$p = \frac{e'}{(1-e')}$

$X \sim Bern(p) := p^{X}(1-p)^{1-X}$

$$\mathcal{L}(\beta_0, \beta_1; Y, X) = \prod_{i=1}^{n} \left(\frac{e^{\beta_0 + \beta_1 X_i}}{1 - e^{\beta_0 + \beta_1 X_i}} \right)^{Y_i} \left(\frac{1}{1 - e^{\beta_0 + \beta_1 X_i}} \right)^{1 - Y_i}$$

$$= \prod_{i=1}^{n} \left(1 - e^{\beta_0 + \beta_1 X_i} \right)^{-1} e^{\sum_{i=1}^{n} Y_i (\beta_0 + \beta_1 X_i)}$$

Computer....

Normality

b_0, b_1

$\downarrow$

β_0, β_1