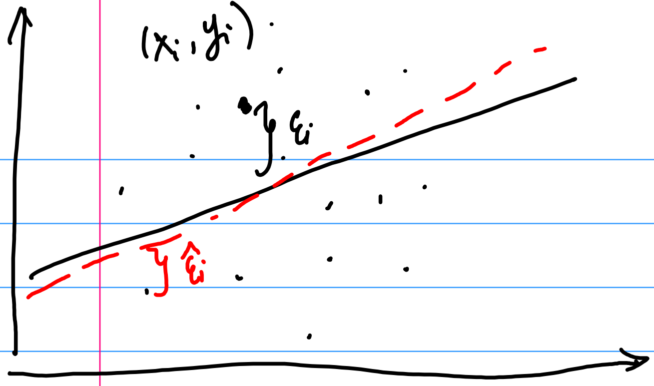


$y_i = \alpha + \beta x_i + \varepsilon_i$

βx_i: fixed (deterministic); ε_i: error

$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

$y_i = \hat{\alpha} + \hat{\beta} x_i + \hat{\varepsilon}_i$ ($\hat{\varepsilon}_i$: residual)

$y = \alpha + \beta X$ – true regression line

$y = \hat{\alpha} + \hat{\beta} x$ – estimated regression line

estimators (functions)

$$\hat{\beta} = \frac{\widehat{Cov}(x,y)}{\widehat{Var}(x)} = \frac{\cancel{\frac{1}{n-1}} \sum (x_i - \bar{x})(y_i - \bar{y})}{\cancel{\frac{1}{n-1}} \sum (x_i - \bar{x})^2}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x}$$

(*) estimates (numbers)

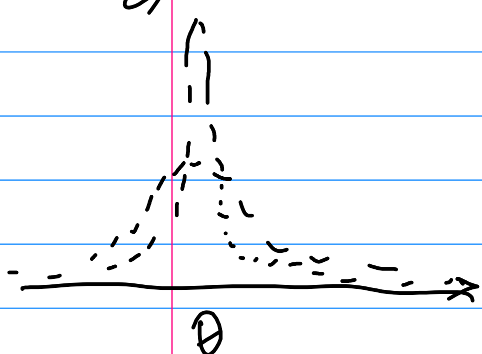
Properties of estimators:

1) Unbiased $E(\hat{\theta}) = \theta$

$E(\text{Bias}) = E(\hat{\theta} - \theta) = 0$

2) Consistency $\operatorname*{plim}_{n \to \infty} \hat{\theta} = \theta$

$$\lim_{n \to \infty} \Pr(|\hat{\theta} - \theta| < \varepsilon) = 1$$



1. Unbiased (at least asympt.)
2. $\operatorname*{Var}_{n \to \infty}(\hat{\theta}) \to 0$

3) Efficiency $var(\hat{\theta}) \leq var(\tilde{\theta})$ $\forall \tilde{\theta} \in \mathcal{U}_{LUE}$

Classic Linear Regression Assumptions:

1) Specification is correct: Endogeneity: omitted var.; simultaneity; measur. error

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni} + \varepsilon_i \quad (\cancel{\ast})$$

2) $X_1, \dots, X_n$ – fixed, not linearly dependent (no perfect multicollinearity)

3) $E\,\varepsilon_i = 0$

4) $\text{Var}\,\varepsilon_i = \sigma^2_\varepsilon$ Heteroscedasticity $(\cancel{\ast})$

5) $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ Autocorrelation $(\cancel{\ast})$

GMT: if 1-5 are true $\Rightarrow$

$$\hat{\beta}_{OLS} = \begin{bmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_n \end{bmatrix} \text{ – BLUE (BUE Hansen 2022)}$$

Pr. 1 X, Y independent $\Rightarrow$ $\rho_{X,Y} = 0$

$P(X,Y) = P(X)P(Y)$ independency

$E(X \mid Y) = E(X)$ $E(Y \mid X) = E(Y)$ unpredictability

$E(X \cdot Y) = E(X)E(Y)$ uncorrelatedness

$$\underbrace{COV(X,Y)}_{=0} = E(XY) - E(X)E(Y)$$

$$\underbrace{\rho_{X,Y}}_{=0} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$$

P2.2 $\{X_1, \ldots, X_n\}$ $\qquad \bar{\mu}_x = \frac{\sum X_i}{n}$

(a) Biasedness? (yes)

$E(X_i)$

$X_i \sim Law$ $\hookrightarrow \{X_1, X_3, \ldots, X_{2m-1}\}$ - odd indexes

$E(X) = \mu_x$ $\quad \hat{\mu}_x = \frac{X_1 + X_3 + \ldots + X_{2m-1}}{m}$

$$E\hat{\mu}_x = \frac{E(X_1) + \ldots + E(X_{2m-1})}{m} = \frac{\sum_{j=1}^{m} E(X_j)}{m} = \frac{m\mu_x}{m} = \mu_x$$

(b) Efficiency: (NO).

$$Var(\hat{\mu}) = Var\left(\frac{X_1 + \ldots + X_{2m-1}}{m}\right) =$$

$$= \frac{1}{m^2} \sum_{j=1}^{m} Var(X_j) = \frac{m \cdot \sigma_x^2}{m^2} = \frac{\sigma_x^2}{m}$$

$$\vee$$

$$Var(\bar{\mu}) = \frac{\sigma_x^2}{n}$$

P2.2 $\quad X \qquad E(X) = \mu_x \qquad Var(X) = \sigma_x^2$

$S = \frac{b_1}{1-q}$

$$\{X_1, \dots, X_n\}$$

$$\hat{\mu}_x = \frac{1}{2} X_1 + \frac{1}{4} X_2 + \dots + \frac{1}{2^n} X_n$$

is $\hat{\mu}_x$ - consistent?

a) $E(\hat{\mu}_x) = \frac{1}{2}\mu_x + \frac{1}{4}\mu_x + \dots + \frac{1}{2^n}\mu_x < \mu_x$

$\hat{\mu}_x$ - biased

$$\lim_{n\to\infty} E(\hat{\mu}_x) = \frac{1/2}{1-1/2} \cdot \mu_x = \mu_x$$

$\hat{\mu}_x$ - asymptotically unbiased

b) $Var(\hat{\mu}_x) = \frac{1}{4}\sigma_x^2 + \frac{1}{16}\sigma_x^2 + \dots + \frac{1}{2^{2n}}\sigma_x^2$

$$\lim_{n\to\infty} (\hat{\mu}_x) = \frac{1/4}{1-1/4}\sigma_x^2 = \frac{1}{3}\sigma_x^2$$

$\Rightarrow \hat{\mu}_x$ - not consistent

P. 3. X $\quad E(X) = \mu \quad Var(X) = \sigma^2$

$\{X_1, \dots, X_n\}$

$$\hat{\mu} = \frac{n+2}{n^2+3n+1} \sum X_i =$$

$\bar{X}$ – consistent estimator

$$= \frac{n^2+2n}{n^2+3n+1} \cdot \frac{\sum X_i}{n} = \frac{n^2+2n}{n^2+3n+1} \cdot \bar{X}$$

$$E\hat{\mu} = \frac{n^2+2n}{n^2+3n+1} \mu$$ – biased

$$\lim_{n\to\infty} E\hat{\mu} = \lim_{n\to\infty} \frac{1+\frac{2}{n}}{1+\frac{3}{n}+\frac{1}{n^2}} \mu = \mu$$ – asympt. unbiased

$$Var\,\hat{\mu} = \left(\frac{n+2}{n^2+3n+1}\right)^2 Var\left(\sum X_i\right) = \left(\frac{n+2}{n^2+3n+1}\right)^2 \cdot n \cdot \sigma_X^2 \longrightarrow 0$$

$\Rightarrow$ consistent