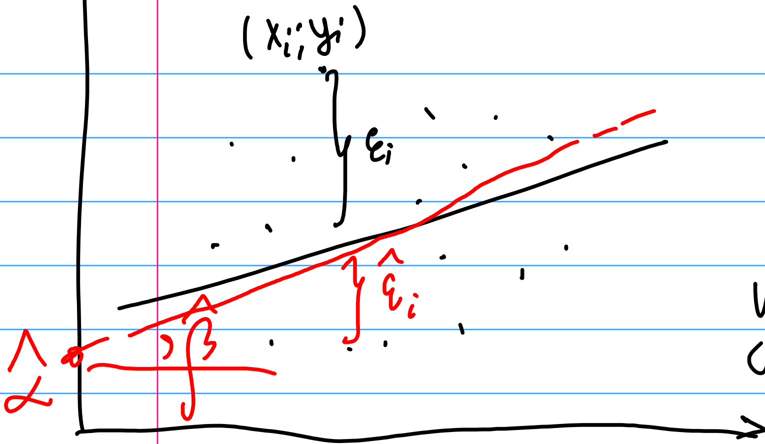
$X \uparrow 1 \quad y \uparrow \hat{\beta}$



$y_i = \alpha + \beta x_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

βx_i: fixed (deterministic)

ε_i: error

$y_i = \hat{\alpha} + \hat{\beta} x_i + \hat{\varepsilon}_i$

$\hat{\varepsilon}_i$: residual

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

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

estimators (function)

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

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

* estimate (number)

Properties of estimators:

1) Unbiasedness $\quad \underbrace{E(\hat{\theta} - \theta)}_{\text{Bias}} = 0 \qquad E(\hat{\theta}) = \theta$

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

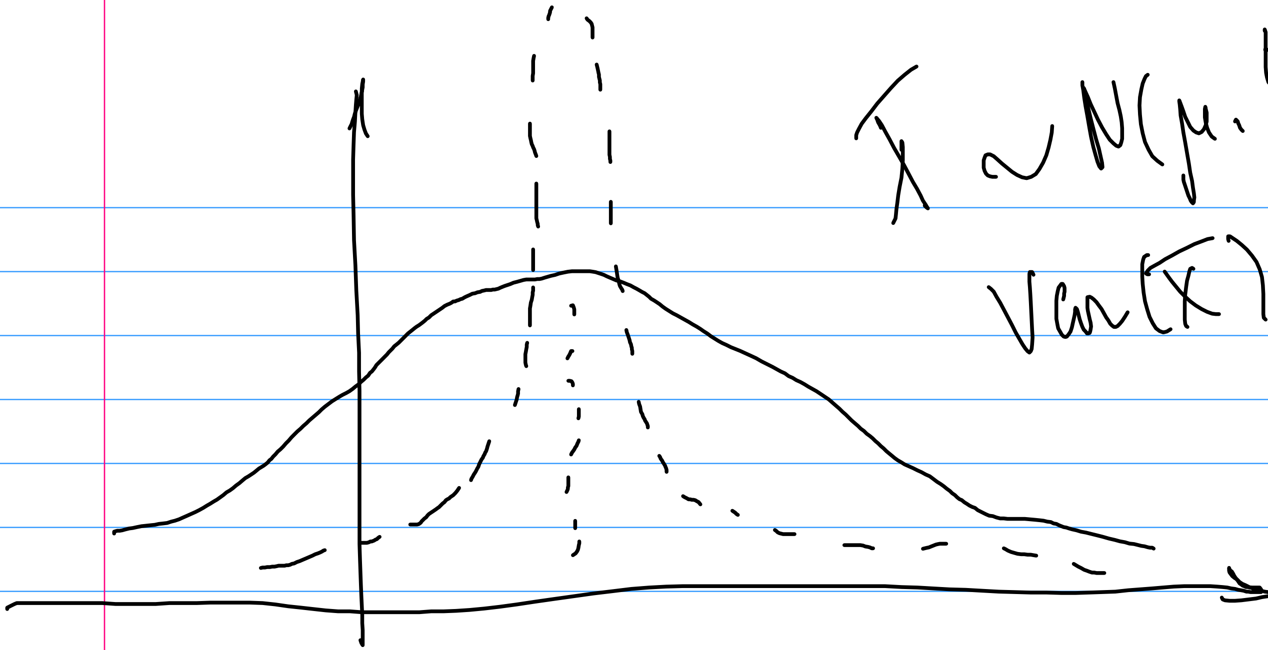
$$\lim_{n \to \infty} Pr(|\hat{\theta} - \theta| < \varepsilon) = 1$$

- unbiased (asymptotically)
- $\lim_{n \to \infty} Var(\hat{\theta}) = 0$

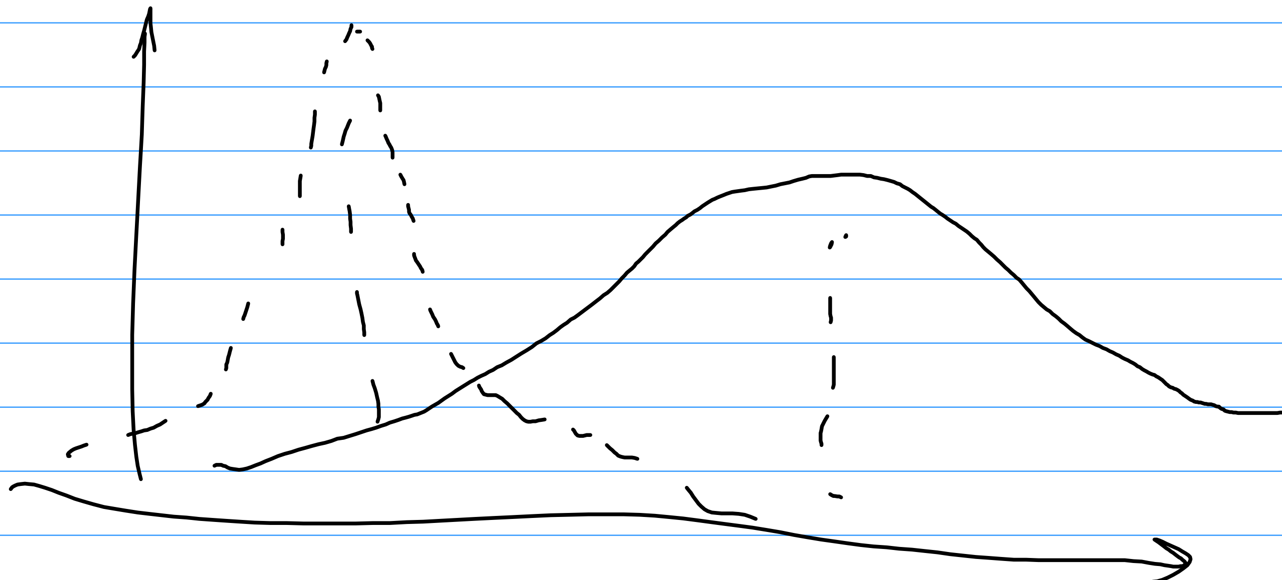
3) Efficiency $\quad Var(\hat{\theta}) \leq Var(\tilde{\theta}) \quad \forall \tilde{\theta} \in \mathcal{U}_{LUE}$

$\bar{X} \sim N(\mu, \sigma^2/n)$

$Var(\bar{X}) = \frac{\sigma^2}{n}$



θ



$\theta \leftarrow \hat{\theta}_n$

Classical Linear Regression Assumptions

(Biased Inconsistent)

1) Correct Specification: Endogeneity → Omitted variables → simultaneity → measurement error

$$y_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_k X_{ki} + \varepsilon_i$$

2) $X_1, \ldots, X_k$ – fixed (deterministic)

not linearly dependent (no perfect multicollinearity)

3) $E\,\varepsilon_i = 0$

4) $Var\,\varepsilon_i = \sigma^2_\varepsilon$ – const Heteroscedasticity (efficiency)

5) $Cov(\varepsilon_i, \varepsilon_j) = 0$ Autocorrelation (efficiency)

GMT: if 1-5 are true ⇒

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

$$\hat{\beta}_{OLS} = (X'X)^{-1} X' y$$

The only downside? BLUE is a much cooler acronym than BUE. I'll miss it.

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

$$P(X, Y) = P(X)\,P(Y)$$ independence

$$E(X|Y) = E(X) \qquad E(Y|X) = E(Y)$$ unpredictability

$$E(XY) = E(X)\,E(Y)$$ uncorrelatedness

$$\rho_{xy} = \frac{\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y} = 0 \qquad \mathrm{COV}(X,Y) = E(XY) - E(X)E(Y) = 0$$