O Modelo

$$Y_i \sim N(\mu, \sigma^2)$$

ou

$$Y_i = \mu + \epsilon_i$$

onde $\epsilon_i \sim N(0, \sigma^2)$

Propriedades de Y_i:

$E(Y_i) = \mu$

$Var(Y_i) = Var(\epsilon_i) = \sigma^2$

μ e σ^2 são parâmetros desconhecidos e serão estimados a partir dos dados.

Estimadores:

(a) $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}$

(b) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$

(c) $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$

Propriedades de $\bar{Y}$:

$E(\bar{y}) = \mu \rightarrow$ é um estimador não-viesado.

$Var(\bar{y}) = \frac{\sigma^2}{n}$ $\boxed{\bar{y} \sim N(\mu; \frac{\sigma^2}{n})}$

De onde vem a equação?

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

(a) $\hat{\mu} = \arg\min_{\mu} \sum_{i=1}^{n}(y_i - \mu)^2$

(b) $\hat{\mu} = \arg\max_{\mu} \prod_{i=1}^{n} P(Y_i = y_i)$

(b.2) $\hat{\mu} = \arg\max_{\mu} \sum_{i=1}^{n} \log P(Y_i = y_i)$

Vício do estimador variância amostral

✓ $E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$

✓ $E(S^2) = \sigma^2$

Faixa de Referência p/ Y_i:

$FR(Y_i, \alpha): \mu \pm Z_{\alpha/2} \cdot \sigma$

Faixa de Referência p/ $\bar{Y}$:

$FR(\bar{y}, \alpha): \mu \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

Intervalo de confiança p/ μ:

$IC(\mu, \alpha): \bar{y} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

Intervalo de confiança p/ μ considerando σ desconhecido:

$IC(\mu, \alpha): \bar{y} \pm t_{\alpha/2, n-1} \times \frac{S}{\sqrt{n}}$

Teorema Central do Limite
Teorema do Limite Central

$Y_i \sim ?\ (E(Y_i) = \mu;\ Var(Y_i) = \sigma^2)$

$$\bar{Y} \underset{n \to +\infty}{\sim} Normal\left(\mu, \frac{\sigma^2}{n}\right)$$

Capítulo 2 - Regressão Linear Simples

O Modelo (gerador dos dados)

$$Y_i | x_i \sim \text{Normal}(\mu = \beta_0 + \beta_1 x_i \,;\, \sigma^2)$$

ou

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

onde $\epsilon_i \sim \text{Normal}(0, \sigma^2)$