Behavior Of Regression Coefficients

$$y(i) = \beta 0 \; + \beta 1 * x(i,1) \;\; + \ldots + \beta k \; * x(i,p) \;\; + \varepsilon(i)$$

Features

Coefficients

Samples

$$y_{n \times 1} = X_{n \times p} \; \beta_{p \times 1}$$

- βhat =(X’X)^-1 X’(Xβ+ε) = β+(X’X)^−1 X’ ε
  - Betas and epsilons are vectors
  - X are matricies
  - ‘ means transpose
  - ^-1 means inverse
- Thus, the distribution of βhat depends on the distribution of the error terms.

Normally Distributed Errors:

If the error terms ε are independently and identically distributed (i.i.d.) as normal with mean zero and variance σ^2, i.e., $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$, then:

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X^T X)^{-1})$$

This result indicates that $\hat{\beta}$ is normally distributed with mean β and covariance matrix $\sigma^2 (X^T X)^{-1}$.

Normally Distributed Errors:

If the error terms ε are independently and identically distributed (i.i.d.) as normal with mean zero and variance σ^2, i.e., $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$, then:

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X^T X)^{-1})$$

This result indicates that $\hat{\beta}$ is normally distributed with mean β and covariance matrix $\sigma^2 (X^T X)^{-1}$.

Non-Normally Distributed Errors:

If the error terms are not normally distributed but satisfy the conditions of the Central Limit Theorem (e.g., they are i.i.d. with finite variance), then for large sample sizes (n large), the sampling distribution of $\hat{\beta}$ will approximate a normal distribution due to the Central Limit Theorem.

Implication

- The normality of the regression coefficient underlies the hypothesis testing
- When the normality assumption about the error terms is met, exact inference is possible.
- In large samples, even if the error terms are not normal, inference based on the asymptotic normality of βhat is frequently justified.

Assumptions About Independent, Predictor Variables

- Linearity – Between predictor and dependent (proportion change)
- Measurement without error in predictor
- Fixed or non-random (esp. in experimental designs)
- No perfect multicollinearity
- Non constant value for predictor