

Structural Equation Models

An extension of multiple regression.
Can incorporate measurement error,
and more

Full generality later

First, introduction to multiple regression with measurement error

Linear Regression

$$Y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_{p-1} x_{i,p-1} + \epsilon_i$$

where $\epsilon_1, \dots, \epsilon_n$ are independent random variables with expected value zero and common variance σ^2 , and $x_{i,1}, \dots, x_{i,p-1}$ are fixed constants.

Matrix Form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where \mathbf{X} is an $n \times p$ matrix of known constants, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown constants, and $\boldsymbol{\epsilon}$ is multivariate normal with mean zero and covariance matrix $\sigma^2 \mathbf{I}_n$

Are X values really constants?

Double Expectation

$$E\{Y\} = E\{E\{Y|X\}\}$$

$E\{Y\}$ is a constant. $E\{Y|X\}$ is a random variable, a function of X .

$$E\{E\{Y|X\}\} = \int E\{Y|X = x\} f(x) dx$$

Beta-hat is (conditionally) unbiased

$$E\{\hat{\beta} | \mathbf{X} = \mathbf{x}\} = \beta$$

Unbiased unconditionally, too

$$E\{\hat{\beta}\} = E\{E\{\hat{\beta} | \mathbf{X} = \mathbf{x}\}\} = E\{\beta\} = \beta$$

Perhaps Clearer

$$\begin{aligned} E\{\hat{\boldsymbol{\beta}}\} &= E\{E\{\hat{\boldsymbol{\beta}}|\mathbf{X} = \mathbf{x}\}\} \\ &= \int \cdots \int E\{\hat{\boldsymbol{\beta}}|\mathbf{X} = \mathbf{x}\} f(\mathbf{x}) d\mathbf{x} \\ &= \int \cdots \int \boldsymbol{\beta} f(\mathbf{x}) d\mathbf{x} \\ &= \boldsymbol{\beta} \int \cdots \int f(\mathbf{x}) d\mathbf{x} \\ &= \boldsymbol{\beta} \cdot 1 = \boldsymbol{\beta}. \end{aligned}$$

Conditional size alpha test, Critical region A

$$Pr\{F \in A | \mathbf{X} = \mathbf{x}\} = \alpha$$

$$\begin{aligned} Pr\{F \in A\} &= \int \cdots \int \alpha f(\mathbf{x}) d\mathbf{x} \\ &= \alpha \int \cdots \int f(\mathbf{x}) d\mathbf{x} \\ &= \alpha. \end{aligned}$$