

Ignoring Measurement Error in Regression¹

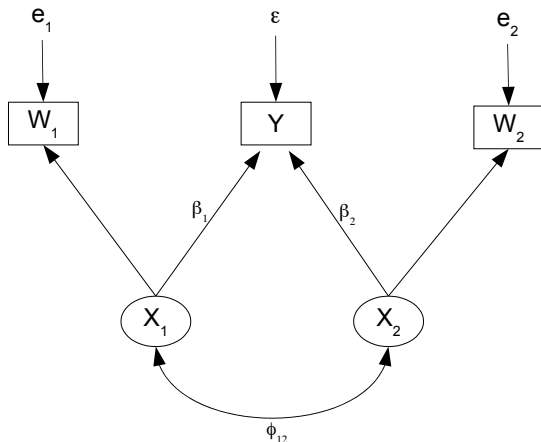
STA2053 Fall 2022

¹See last slide for copyright information.

Measurement Error

- Snack food consumption
- Exercise
- Income
- Cause of death
- Even amount of drug that reaches animal's blood stream in an experimental study
- Is there anything that is *not* measured with error?

Measurement error in two explanatory variables



Want to assess the relationship of X_2 to Y controlling for X_1 by testing $H_0 : \beta_2 = 0$.

Statement of the model

Independently for $i = 1, \dots, n$

$$\begin{aligned}Y_i &= \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i \\W_{i,1} &= X_{i,1} + e_{i,1} \\W_{i,2} &= X_{i,2} + e_{i,2},\end{aligned}$$

where

$$E(X_{i,1}) = \mu_1, E(X_{i,2}) = \mu_2, E(\epsilon_i) = E(e_{i,1}) = E(e_{i,2}) = 0,$$

$$\text{Var}(\epsilon_i) = \psi, \text{Var}(e_{i,1}) = \omega_1, \text{Var}(e_{i,2}) = \omega_2,$$

The errors $\epsilon_i, e_{i,1}$ and $e_{i,2}$ are all independent,

$X_{i,1}$ and $X_{i,2}$ are independent of $\epsilon_i, e_{i,1}$ and $e_{i,2}$, and

$$\text{cov} \begin{pmatrix} X_{i,1} \\ X_{i,1} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{pmatrix}.$$

True Model versus Naive Model

Independently for $i = 1, \dots, n$

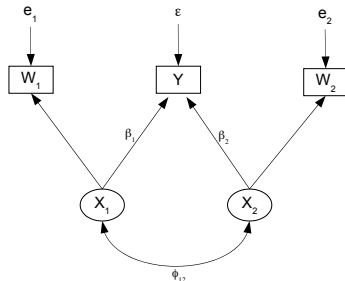
True model:

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \epsilon_i$$

$$W_{i,1} = X_{i,1} + e_{i,1}$$

$$W_{i,2} = X_{i,2} + e_{i,2},$$

Naive model: $Y_i = \beta_0 + \beta_1 W_{i,1} + \beta_2 W_{i,2} + \epsilon_i$



Copyright Information

This slide show was prepared by **Jerry Brunner**, Department of Statistics, University of Toronto. It is licensed under a **Creative Commons Attribution - ShareAlike 3.0 Unported License**. Use any part of it as you like and share the result freely. The \LaTeX source code is available from the course website:
<http://www.utstat.toronto.edu/brunner/oldclass/2053f22>