Analysis of covariance

Covariate (in this context): essentially a quantitative blocking variable, usually not under control of the investigator.

Examples: temperature, humidity, pre-test score, height, etc.

"Adjusting for" covariate might reveal significant effects of a treatment which otherwise would not be detected.

Ex. treatment A, two levels:

Sometimes (especially in observational studies) a quantitative covariate can reveal itself as a "lurking variable":

Interaction: lines are not parallel (slope depends on factor level).

Models

Single factor AOV

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

Covariate-only (regression)

$$Y_{ij} = \mu + \beta x_{ij} + \epsilon_{ij}$$

(or $Y_{ij} = \mu + \beta(x_{ij} - \overline{x}_{..}) + \epsilon_{ij}$ which uses the "centered" covariate values)

Both variables, without interaction

$$Y_{ij} = \mu + \alpha_i + \beta x_{ij} + \epsilon_{ij}$$

Both variables, with interaction

$$Y_{ij} = \mu + \alpha_i + \beta x_{ij} + \gamma_i x_{ij} + \epsilon_{ij}$$

$$(\beta + \gamma_i) x_{ij}$$

In SAS, one enters covariates directly into the MODEL statement, without declaring them as categorical in the CLASS statement.