Lecture 7: Models, Conditions, Analysis

Models
* Two independent samples
  \[ Y = \mu_i + \varepsilon \]
  * \( Y \) is the observed value.
  * \( \mu_i \) is the population mean for the treatment \( i \).
  * \( \varepsilon \) is the random error.

Models
* Two independent samples
  \[ Y = \mu + \tau_i + \varepsilon \]
  * \( Y \) is the observed value.
  * \( \mu \) is the grand population mean.
  * \( \tau_i \) is the effect of treatment \( i \).
  * \( \varepsilon \) is the random error.

Models
Observed value = “true” value + residual error

Conditions
* “true” value
  * Constant.
  * Pieces add.

Conditions
* Residual (random) error
  * Add to zero.
  * Same standard deviation (\( \sigma \)).
  * Independent.
  * Normally distributed.

Conditions
* The conditions on the random error are exactly the same as we saw in Stat 401.
Lecture 7: Models, Conditions, Analysis

Analysis

* Informal.
* Formal.
* Check Conditions (secondary).

Informal Analysis

* Estimate the “true” values (parameters) using the data from the experiment.

Estimates

* Use the overall sample mean \( \bar{Y}_{++} \) to estimate the overall population mean \( (\mu) \).

Estimates

* Use the treatment sample mean \( \bar{Y}_{i+} \) to estimate the treatment population mean \( (\mu_i) \).

Estimates

* Use the difference \( (\hat{\tau}_i = \bar{Y}_{i+} - \bar{Y}_{++}) \) to estimate the effect of treatment \( i \) \( (\tau_i = \mu_i - \mu) \).

Visualizing the Model

Overall Mean + Treatment Effect + Residual Error
Sodium in diet experiment

* Analysis
  * Informal.
  * Formal.
  * Check Conditions.

Informal Analysis

* Compute estimates of the “true” values (parameters) from the experimental data.

Informal Analysis

\[
\bar{Y}_{++} = \frac{\sum Y}{N} = \frac{3061}{20} = 153.05 \text{ mmHg}
\]

* The average blood pressure for all 20 men in the experiment.

Informal Analysis

* 50 mmol Na/day
  * \( \bar{Y}_{50+} = 143.10 \) mmHg
  * 200 mmol Na/day
  * \( \bar{Y}_{200+} = 163.00 \) mmHg

Informal Analysis

* Effect of 50 mmol Na/day
  * \( \bar{Y}_{50+} - \bar{Y}_{++} = -9.95 \) mmHg
  * Effect of 200 mmol Na/day
  * \( \bar{Y}_{200+} - \bar{Y}_{++} = +9.95 \) mmHg

Oneway Analysis of Systolic (mmHg) By Group
Lecture 7: Models, Conditions, Analysis

Formal Analysis

* The Analysis of Variance – ANOVA
* Quantify the total amount of variability and split it among the sources.

Partitioning Variability

Total Variability

\[ \text{Total Variability} = \text{Variability due to treatments effects} + \text{Variability due to chance (random) error}. \]

Planned Systematic

Measurement

Experimental Material

Total Variability

\[ \sum (Y - \bar{Y}_++)^2 = \frac{4140.95}{19} \]

\[ \frac{SS_{Total}}{df_{Total}} \]

Partitioning Variability

Sum of Squares

C Total

\[ \text{Sum of Squares Treatments} + \text{Sum of Squares Error} \]

Degrees of Freedom

C Total

\[ \text{Degrees of Freedom Treatments} + \text{Degrees of Freedom Error} \]