Computing Standard Errors
Overview

- In practice, we typically rely on statistical software to compute the needed standard errors (SEs)

- Why the need to understand their functional form?
  1. Understanding SEs can help in determining the best approach to increase power
  2. WARNING: A fixed effects procedure that uses least squares for estimation may not produce the correct SEs

- This topic runs through some examples of calculating SEs for reference purposes as well as demonstrating a method for calculation

Computing Standard Errors
One-way Random Effects

\[ y_{ij} = \mu + \tau_i + \epsilon_{ij} \]

\[ \begin{align*}
\tau_i &\sim N(0, \sigma^2_\tau) \text{ and } \epsilon_{ij} \sim N(0, \sigma^2) \\
\{\tau_i\} &\text{ and } \{\epsilon_{ij}\} \text{ independent}
\end{align*} \]

\[ \text{Var}(\overline{y}_{..}) = \text{Var}(\mu + \overline{\tau} + \overline{\epsilon}_{..}) \]
\[ = \text{Var}(\mu) + \text{Var}(\overline{\tau}) + \text{Var}(\overline{\epsilon}_{..}) \]
\[ = 0 + \sigma^2_\tau / a + \sigma^2 / an \]
\[ = (n\sigma^2_\tau + \sigma^2) / an \]

Since \( E(\text{MS}_{\text{Trit}}) = n\sigma^2_\tau + \sigma^2 \), we use this mean square and its degrees of freedom when constructing a confidence interval or performing a hypothesis test.
One-way Random Effects

- What if the $n_i$ vary?

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \left\{ \begin{array}{l} i = 1, 2 \ldots a \\ j = 1, 2, \ldots n_i \end{array} \right.$$  

$$\tau_i \sim N(0, \sigma^2 \tau)$$

$\{\tau_i\}$ and $\{\epsilon_{ij}\}$ independent

- Based on model, $\bar{y}_i \sim N(\mu, \sigma^2 \tau + \sigma^2 / n_i)$

- Estimate of $\mu$ is then a weighted average of the trt means

- Standard error is $\sqrt{1/\sum w_i}$ where $w_i = 1/(\hat{\sigma}^2 \tau + \hat{\sigma}^2 / n_i)$

- Degrees of freedom based on Satterthwaite’s approximation
Two Factor Random Effects Model

\[ y_{ijk} = \mu + \tau_i + \beta_j + (\tau \beta)_{ij} + \epsilon_{ijk} \]

\( i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, n \)

\( \tau_i \sim N(0, \sigma^2_\tau) \) and \( \beta_j \sim N(0, \sigma^2_\beta) \)

\( (\tau \beta)_{ij} \sim N(0, \sigma^2_{\tau \beta}) \) and \( \epsilon_{ijk} \sim N(0, \sigma^2) \)

\{\tau_i\}, \{\beta_j\}, \{(\tau \beta)_{ij}\} \) and \{\epsilon_{ijk}\} \) independent

\[
\text{Var}(\bar{y}_{..}) = \text{Var}(\mu + \bar{\tau} + \bar{\beta} + (\tau \beta)_{..} + \bar{\epsilon}...)
= 0 + \sigma^2_\tau / a + \sigma^2_\beta / b + \sigma^2_{\tau \beta} / ab + \sigma^2 / abn
= (bn\sigma^2_\tau + an\sigma^2_\beta + n\sigma^2_{\tau \beta} + \sigma^2) / abn
\]

In this case, there is no expected mean square equal the numerator. As a result, the combination \( \text{MS}_A + \text{MS}_B - \text{MS}_{AB} \) is used to estimate the variance and Satterthwaite’s degrees of freedom formula is used to approximate the degrees of freedom.
Two Factor Mixed Effects Model

\[ y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \]

\[ \sum \tau_i = 0 \quad \text{and} \quad \sum_i (\tau\beta)_{ij} = 0 \text{ for } \beta \text{ level } j \]

\[ \beta \sim N(0, \sigma_{\beta}^2) \quad \text{and} \quad \epsilon_{ijk} \sim N(0, \sigma^2) \]

\[
\text{Var}(\bar{y}_{i..}) = \text{Var}(\mu + \tau_i + \bar{\beta} + (\tau\beta)_{i.} + \bar{\epsilon}_{i..}) \\
= 0 + \sigma_{\beta}^2/b + (a-1)\sigma_{\tau\beta}^2/ab + \sigma^2/bn
\]

The unrestricted model would not have this \( \frac{a-1}{a} \) coefficient in front of the \( \sigma_{\tau\beta}^2 \). In either case, the estimate of the variance can only be written in the form \( p_1 MS_1 + p_2 MS_2 + ... + p_k MS_k \) where some of the \( p_i \) are different than \( \pm 1 \). The Satterthwaite formula

Computing Standard Errors
can be generalized to approximate this situation. It is simply

\[ \text{df} = \frac{(\sum p_i \text{MS}_i)^2}{\sum p_i^2 \text{MS}_i^2 / df_i} \]

For a pairwise treatment comparison

\[
\text{Var}(\bar{y}_{i..} - \bar{y}_{i*..}) = \text{Var}(\tau_i - \tau_{i*} + (\tau\beta)_{i.} - (\tau\beta)_{i*} + \epsilon_{i..} - \epsilon_{i*..}) = 2\sigma^2_{\tau\beta}/b + 2\sigma^2/bn
\]

In both the restricted and unrestricted models, the variance of the difference between two treatment means is the same. Here we would use MS_{AB} and its degrees of freedom when performing hypothesis tests.
Split Plot Design

Will consider the unrestricted mixed model and look at both pooled and unpooled subplot error. Will also focus on RCBD in whole plot with no replication.

There are several comparisons that may be of interest.

1. Main effect in whole plot
2. Main effect in subplot
3. Interaction with $i$ fixed
4. Interaction with $k$ fixed

Pooled

$$y_{ijk} = \mu + B_j + A_i + AB_{ij} + C_k + AC_{ik} + \epsilon_{ijk}$$

$$\begin{align*}
\sum A_i &= 0 & \text{and} & & B &\sim N(0, \sigma_B^2) \\
AB_{ij} &\sim N(0, \sigma_{AB}^2) & \text{and} & & \epsilon_{ijk} &\sim N(0, \sigma^2) \\
\sum C_k &= 0 & \text{and} & & \sum \sum AC_{ik} &= 0
\end{align*}$$
\[ \bar{y}_{..} = \mu + \bar{B} + A_i + \bar{AB}_i + \bar{C} + \bar{AC}_i + \bar{\epsilon}_{..} \]
\[ \bar{y}_{..k} = \mu + \bar{B} + \bar{A} + \bar{AB}_{..} + C_k + \bar{AC}_{.k} + \bar{\epsilon}_{..k} \]
\[ \bar{y}_{i,k} = \mu + \bar{B} + A_i + \bar{AB}_i + C_k + AC_{ik} + \bar{\epsilon}_{i,k} \]

1 Use MS_{AB} in calculations
\[ \text{Var}(\bar{y}_{..} - \bar{y}_{..}) = \text{Var}(A_i - A_i* + \bar{AB}_i - \bar{AB}_{i*} + \bar{\epsilon}_{..} - \bar{\epsilon}_{..}) = 2(\sigma^2_{AB}/b + \sigma^2/bc) \]

2 Use MS_E in calculations
\[ \text{Var}(\bar{y}_{..k} - \bar{y}_{..k}) = \text{Var}(C_k - C_k* + \bar{\epsilon}_{..k} - \bar{\epsilon}_{..k}) = 2\sigma^2/ab \]

3 Use MS_E in calculations
\[ \text{Var}(\bar{y}_{i,k} - \bar{y}_{i,k}) = \text{Var}(C_k - C_k* + AC_{ik} - AC_{ik*} + \bar{\epsilon}_{i,k} - \bar{\epsilon}_{i,k}) = 2\sigma^2/b \]

4 Use linear combination \((c - 1)MS_E + MS_{AB}\)
\[ \text{Var}(\bar{y}_{i,k} - \bar{y}_{i,k*}) = \text{Var}(A_i - A_i* + \bar{AB}_i - \bar{AB}_{i*} + AC_{ik} - AC_{ik*} + \bar{\epsilon}_{i,k} - \bar{\epsilon}_{i,k*}) = 2(\sigma^2_{AB}/b + \sigma^2/b) \]
Unpooled

\[ y_{ijk} = \mu + B_j + A_i + AB_{ij} + C_k + AC_{ik} + BC_{jk} + \epsilon_{ijk} \]

1 Same: Use MS_{AB} in calculations

2 Use MS_{BC} in calculations

\[
\text{Var}(\bar{y}_{..k} - \bar{y}_{..k*}) = \text{Var}(C_k - C_{k*} + \overline{BC}.k - \overline{BC}.k* + \\
\quad \overline{\epsilon}_{..k} - \overline{\epsilon}_{..k*}) \\
= 2(\sigma^2_{BC}/b + \sigma^2/ab)
\]

3 Use linear combination \((a - 1)\text{MS}_E + \text{MS}_{BC}\)

\[
\text{Var}(\bar{y}_{i.k} - \bar{y}_{i.k*}) = \text{Var}(C_k - C_{k*} + AC_{ik} - AC_{ik*} + +\overline{BC}.k - \overline{BC}.k* + \\
\quad \overline{\epsilon}_{i.k} - \overline{\epsilon}_{i.k*}) \\
= 2(\sigma^2_{BC}/b + \sigma^2/b)
\]

4 Same as before