The ANOVA Approach to the Analysis of Linear Mixed-Effects Models

We begin with a relatively simple special case. Suppose

$$y_{ijk} = \mu + \tau_i + u_{ij} + e_{ijk}, \ (i = 1, \ldots, t; \ j = 1, \ldots, n; \ k = 1, \ldots, m)$$

$$\boldsymbol{\beta} = (\mu, \tau_1, \ldots, \tau_t)', \ \boldsymbol{u} = (u_{11}, u_{12}, \ldots, u_{tn})', \boldsymbol{e} = (e_{111}, e_{112}, \ldots, e_{tnm})',$$

$\boldsymbol{\beta} \in \mathbb{R}^{t+1}$, an unknown parameter vector,

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_u^2 \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \sigma_e^2 \boldsymbol{I} \end{bmatrix} \right),$$ where

$\sigma_u^2, \sigma_e^2 \in \mathbb{R}^+$ are unknown variance components.

- This is the standard model for a CRD with t treatments, n experimental units per treatment, and m observations per experimental unit.

- We can write the model as $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$, where

$$\boldsymbol{X} = [\mathbf{1}_{tnm\times 1},\ \boldsymbol{I}_{t\times t} \otimes \mathbf{1}_{nm\times 1}] \qquad \text{and} \qquad \boldsymbol{Z} = [\boldsymbol{I}_{tn\times tn} \otimes \mathbf{1}_{m\times 1}].$$

Special Case of $t = n = m = 2$

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix}, \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{u} = \begin{bmatrix} u_{11} \\ u_{12} \\ u_{21} \\ u_{22} \end{bmatrix}$$

Connection to ANOVA

Let $\boldsymbol{X}_1 = \mathbf{1}_{tnm \times 1}$, $\boldsymbol{X}_2 = [\mathbf{1}_{tnm \times 1},\ \boldsymbol{I}_{t \times t} \otimes \mathbf{1}_{nm \times 1}]$, $\boldsymbol{X}_3 = [\boldsymbol{I}_{tn \times tn} \otimes \mathbf{1}_{m \times 1}]$.

Note that $\mathcal{C}(\boldsymbol{X}_1) \subset \mathcal{C}(\boldsymbol{X}_2) \subset \mathcal{C}(\boldsymbol{X}_3)$, $\boldsymbol{X} = \boldsymbol{X}_2$, and $\boldsymbol{Z} = \boldsymbol{X}_3$.

As usual, let $\boldsymbol{P}_j = \boldsymbol{P}_{\boldsymbol{X}_j} = \boldsymbol{X}_j(\boldsymbol{X}_j'\boldsymbol{X}_j)^{-}\boldsymbol{X}_j'$ for $j = 1, 2, 3$.

An ANOVA Table

Sum of Squares	Degrees of Freedom
$\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}$	$\text{rank}(\boldsymbol{X}_2) - \text{rank}(\boldsymbol{X}_1) = t - 1$
$\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}$	$\text{rank}(\boldsymbol{X}_3) - \text{rank}(\boldsymbol{X}_2) = tn - t$
$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y}$	$\text{rank}(I) - \text{rank}(\boldsymbol{X}_3) = tnm - tn$
$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y}$	$\text{rank}(I) - \text{rank}(\boldsymbol{X}_1) = tnm - 1$

Shortcut for Obtaining DF from Source

Source	DF
$treatments$	$t-1$
$exp.units(treatments)$	$(n-1)t$
$obs.units(exp.units, treatments)$	$(m-1)nt$
$c.total$	$tnm-1$

Shortcut for Obtaining SS from DF

Source	DF	Sum of Squares
trt	$t-1$	$\sum_{i=1}^{t}\sum_{j=1}^{n}\sum_{k=1}^{m}(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot\cdot\cdot})^2$
$xu(trt)$	$t(n-1)$	$\sum_{i=1}^{t}\sum_{j=1}^{n}\sum_{k=1}^{m}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})^2$
$ou(xu,trt)$	$tn(m-1)$	$\sum_{i=1}^{t}\sum_{j=1}^{n}\sum_{k=1}^{m}(y_{ijk}-\bar{y}_{ij\cdot})^2$
$c.total$	$tnm-1$	$\sum_{i=1}^{t}\sum_{j=1}^{n}\sum_{k=1}^{m}(y_{ijk}-\bar{y}_{\cdot\cdot\cdot})^2$

Source	DF	Sum of Squares	Mean Square
trt	$t-1$	$nm\sum_{i=1}^{t}(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot\cdot\cdot})^2$	
$xu(trt)$	$tn-t$	$m\sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})^2$	SS/DF
$ou(xu,trt)$	$tnm-tn$	$\sum_{i=1}^{t}\sum_{j=1}^{n}\sum_{k=1}^{m}(y_{ijk}-\bar{y}_{ij\cdot})^2$	
$c.total$	$tnm-1$	$\sum_{i=1}^{t}\sum_{j=1}^{n}\sum_{k=1}^{m}(y_{ijk}-\bar{y}_{\cdot\cdot\cdot})^2$	

Expected Mean Squares

- Based on our linear mixed-effects model ($\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$ and associated assumptions), we can find the expected value of each mean square in the ANOVA table.

- Examining these expected values helps us see ways to

  1) test hypotheses of interest by computing ratios of mean squares, and

  2) estimate variance components by computing linear combinations of mean squares.

Expected Mean Squares

- For balanced designs, there are shortcuts (not presented here) for writing down expected mean squares.
- Rather than memorizing shortcuts, I think it is better to know how to derive expected mean squares.
- Before going through one example derivation, we will prove a useful result that you may already be familiar with.

Expectation of Sample Variance Numerator

Suppose $w_1, \ldots, w_k \overset{ind}{\sim} (\mu_w, \sigma_w^2)$. Then

$$
\begin{aligned}
k\sigma_w^2 &= \sum_{i=1}^{k} \sigma_w^2 = \sum_{i=1}^{k} E(w_i - \mu_w)^2 = \sum_{i=1}^{k} E(w_i - \bar{w}_\cdot + \bar{w}_\cdot - \mu_w)^2 \\
&= E\left\{ \sum_{i=1}^{k} (w_i - \bar{w}_\cdot + \bar{w}_\cdot - \mu_w)^2 \right\} \\
&= E\left\{ \sum_{i=1}^{k} \left[(w_i - \bar{w}_\cdot)^2 + (\bar{w}_\cdot - \mu_w)^2 + 2(\bar{w}_\cdot - \mu_w)(w_i - \bar{w}_\cdot) \right] \right\} \\
&= E\left\{ \sum_{i=1}^{k} (w_i - \bar{w}_\cdot)^2 + \sum_{i=1}^{k} (\bar{w}_\cdot - \mu_w)^2 + 2(\bar{w}_\cdot - \mu_w) \sum_{i=1}^{k} (w_i - \bar{w}_\cdot) \right\}
\end{aligned}
$$

Expectation of Sample Variance Numerator (ctd.)

$$
\begin{aligned}
k\sigma_w^2 &= E\left\{\sum_{i=1}^{k}(w_i - \bar{w}.)^2 + \sum_{i=1}^{k}(\bar{w}. - \mu_w)^2\right\} \\
&= E\left\{\sum_{i=1}^{k}(w_i - \bar{w}.)^2 + k(\bar{w}. - \mu_w)^2\right\} \\
&= E\left\{\sum_{i=1}^{k}(w_i - \bar{w}.)^2\right\} + kE(\bar{w}. - \mu_w)^2 \\
&= E\left\{\sum_{i=1}^{k}(w_i - \bar{w}.)^2\right\} + k\mathrm{Var}(\bar{w}.) \\
&= E\left\{\sum_{i=1}^{k}(w_i - \bar{w}.)^2\right\} + k\sigma_w^2/k = E\left\{\sum_{i=1}^{k}(w_i - \bar{w}.)^2\right\} + \sigma_w^2
\end{aligned}
$$

Expectation of Sample Variance Numerator (ctd.)

We have shown

$$k\sigma_w^2 = E\left\{\sum_{i=1}^{k}(w_i - \bar{w}_{\cdot})^2\right\} + \sigma_w^2.$$

Therefore,

$$E\left\{\sum_{i=1}^{k}(w_i - \bar{w}_{\cdot})^2\right\} = (k-1)\sigma_w^2.$$

This is just a special case of the Gauss-Markov model result $E(\hat{\sigma}^2) = \sigma^2$. $\quad$ ($\boldsymbol{y} = [w_1, \ldots, w_k]'$, $\boldsymbol{X} = \mathbf{1}$, $\boldsymbol{\beta} = [\mu_w]$, $\sigma^2 = \sigma_w^2$)

Expected Value of MS_{trt}

$$
\begin{aligned}
E(MS_{trt}) &= \tfrac{nm}{t-1} \textstyle\sum_{i=1}^{t} E(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 \\
&= \tfrac{nm}{t-1} \textstyle\sum_{i=1}^{t} E(\mu + \tau_i + \bar{u}_{i\cdot} + \bar{e}_{i\cdot\cdot} - \mu - \bar{\tau}_{\cdot} - \bar{u}_{\cdot\cdot} - \bar{e}_{\cdot\cdot\cdot})^2 \\
&= \tfrac{nm}{t-1} \textstyle\sum_{i=1}^{t} E(\tau_i - \bar{\tau}_{\cdot} + \bar{u}_{i\cdot} - \bar{u}_{\cdot\cdot} + \bar{e}_{i\cdot\cdot} - \bar{e}_{\cdot\cdot\cdot})^2 \\
&= \tfrac{nm}{t-1} \textstyle\sum_{i=1}^{t} [(\tau_i - \bar{\tau}_{\cdot})^2 + E(\bar{u}_{i\cdot} - \bar{u}_{\cdot\cdot})^2 + E(\bar{e}_{i\cdot\cdot} - \bar{e}_{\cdot\cdot\cdot})^2] \\
&= \tfrac{nm}{t-1} [\textstyle\sum_{i=1}^{t} (\tau_i - \bar{\tau}_{\cdot})^2 + E\{\textstyle\sum_{i=1}^{t} (\bar{u}_{i\cdot} - \bar{u}_{\cdot\cdot})^2\} \\
&\qquad\qquad + E\{\textstyle\sum_{i=1}^{t} (\bar{e}_{i\cdot\cdot} - \bar{e}_{\cdot\cdot\cdot})^2\}]
\end{aligned}
$$

So, to simplify $E(MS_{trt})$ further, note that

$$\overline{u}_{1\cdot}, \ldots, \overline{u}_{t\cdot} \overset{i.i.d.}{\sim} N\left(0, \frac{\sigma_u^2}{n}\right).$$

Thus,

$$E\left\{\sum_{i=1}^{t}(\overline{u}_{i\cdot} - \overline{u}_{\cdot\cdot})^2\right\} = (t-1)\frac{\sigma_u^2}{n}.$$

Similarly,

$$\overline{e}_{1\cdot\cdot}, \ldots, \overline{e}_{t\cdot\cdot} \overset{i.i.d.}{\sim} N\left(0, \frac{\sigma_e^2}{nm}\right)$$

so that

$$E\left\{\sum_{i=1}^{t}(\overline{e}_{i\cdot\cdot} - \overline{e}_{\cdot\cdot\cdot})^2\right\} = (t-1)\frac{\sigma_e^2}{nm}.$$

It follows that

$$
\begin{aligned}
E(MS_{trt}) &= \frac{nm}{t-1}\left[\sum_{i=1}^{t}(\tau_i - \overline{\tau}_{\cdot})^2 + E\left\{\sum_{i=1}^{t}(\overline{u}_{i\cdot} - \overline{u}_{\cdot\cdot})^2\right\}\right. \\
&\qquad\qquad \left. + E\left\{\sum_{i=1}^{t}(\overline{e}_{i\cdot\cdot} - \overline{e}_{\cdot\cdot\cdot})^2\right\}\right] \\
&= \frac{nm}{t-1}\left[\sum_{i=1}^{t}(\tau_i - \overline{\tau}_{\cdot})^2 + (t-1)\frac{\sigma_u^2}{n}\right. \\
&\qquad\qquad \left. + (t-1)\frac{\sigma_e^2}{nm}\right] \\
&= \frac{nm}{t-1}\sum_{i=1}^{t}(\tau_i - \overline{\tau}_{\cdot})^2 + m\sigma_u^2 + \sigma_e^2.
\end{aligned}
$$

Similar calculations allow us to add an Expected Mean Squares (EMS) column to our ANOVA table.

Source	EMS
trt	$\sigma_e^2 + m\sigma_u^2 + \frac{nm}{t-1} \sum_{i=1}^{t}(\tau_i - \bar{\tau}_.)^2$
$xu(trt)$	$\sigma_e^2 + m\sigma_u^2$
$ou(xu, trt)$	σ_e^2

Expected Mean Squares (EMS) could be computed using

$$E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + E(\boldsymbol{y})'\boldsymbol{A}E(\boldsymbol{y}),$$

where

$$\boldsymbol{\Sigma} = \mathrm{Var}(\boldsymbol{y}) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R} = \sigma_u^2 \boldsymbol{I}_{tn\times\, tn} \otimes \mathbf{1}\mathbf{1}'_{m\times\, m} + \sigma_e^2 \boldsymbol{I}_{tnm\times\, tnm}$$

and

$$E(\boldsymbol{y}) = \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu + \tau_t \end{bmatrix} \otimes \mathbf{1}_{nm\times\, 1}.$$

Furthermore, with some nontrivial work, it can be shown that

$$\frac{\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2} \sim \chi^2_{t-1}\left(\frac{nm}{2(\sigma_e^2 + m\sigma_u^2)}\sum_{i=1}^{t}(\tau_i - \bar{\tau}.)^2\right),$$

$$\frac{\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2} \sim \chi^2_{tn-t},$$

$$\frac{\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y}}{\sigma_e^2} \sim \chi^2_{tnm-tn},$$

and that these three χ^2 random variables are independent.

It follows that

$$
\begin{aligned}
F_1 &= \frac{MS_{trt}}{MS_{xu(trt)}} = \frac{\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}/(t-1)}{\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}/(tn-t)} \\
&= \frac{\left[\frac{\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2}\right]/(t-1)}{\left[\frac{\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2}\right]/(tn-t)} \\
&\sim F_{t-1,\, tn-t}\left(\frac{nm}{2(\sigma_e^2 + m\sigma_u^2)}\sum_{i=1}^{t}(\tau_i - \overline{\tau}_{\cdot})^2\right).
\end{aligned}
$$

Thus, we can use F_1 to test $H_0 : \tau_1 = \cdots = \tau_t$.

Also,

$$
\begin{aligned}
F_2 &= \frac{MS_{xu(trt)}}{MS_{ou(xu,trt)}} = \frac{\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}/(tn-t)}{\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y}/(tnm-tn)} \\
&= \left(\frac{\sigma_e^2 + m\sigma_u^2}{\sigma_e^2}\right) \frac{\left[\frac{\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2}\right]/(tn-t)}{\left[\frac{\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\sigma_e^2}\right]/(tnm-tn)} \\
&\sim \left(\frac{\sigma_e^2 + m\sigma_u^2}{\sigma_e^2}\right) F_{tn-t,\, tnm-tn}.
\end{aligned}
$$

Thus, we can use F_2 to test $H_0 : \sigma_u^2 = 0$.

Estimating σ_u^2

Note that

$$E\left(\frac{MS_{xu(trt)} - MS_{ou(xu,trt)}}{m}\right) = \frac{(\sigma_e^2 + m\sigma_u^2) - \sigma_e^2}{m} = \sigma_u^2.$$

Thus,

$$\frac{MS_{xu(trt)} - MS_{ou(xu,trt)}}{m}$$

is an unbiased estimator of σ_u^2.

- Although
$$\frac{MS_{xu(trt)} - MS_{ou(xu,trt)}}{m}$$
is an unbiased estimator of σ_u^2, this estimator can take negative values.

- This is undesirable because σ_u^2, the variance of the u random effects, cannot be negative.

- Later in the course, we will discuss likelihood based methods for estimating variance components that honor the parameter space.

Estimation of Estimable $\boldsymbol{C\beta}$

As we have seen previously,

$$\boldsymbol{\Sigma} \equiv \text{Var}(\boldsymbol{y}) = \sigma_u^2 \boldsymbol{I}_{tn\times\, tn} \; \otimes \; \mathbf{1}\mathbf{1}'_{m\times\, m} + \sigma_e^2 \boldsymbol{I}_{tnm\times\, tnm}.$$

It turns out that

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} = \hat{\boldsymbol{\beta}}.$$

Thus, the GLS estimator of any estimable $\boldsymbol{C\beta}$ is equal to the OLS estimator in this special case.

An Analysis Based on the Average for Each Experimental Unit

Recall that our model is

$$y_{ijk} = \mu + \tau_i + u_{ij} + e_{ijk}, \ (i = 1, \ldots, t; j = 1, \ldots, n; k = 1, \ldots, m)$$

The average of observations for experimental unit ij is

$$\bar{y}_{ij\cdot} = \mu + \tau_i + u_{ij} + \bar{e}_{ij\cdot}$$

If we define

$$\epsilon_{ij} = u_{ij} + \bar{e}_{ij\cdot} \ \forall \ i, j$$

and

$$\sigma^2 = \sigma_u^2 + \frac{\sigma_e^2}{m},$$

we have

$$\bar{y}_{ij\cdot} = \mu + \tau_i + \epsilon_{ij},$$

where the ϵ_{ij} terms are iid $N(0, \sigma^2)$. Thus, averaging the same number (m) of multiple observations per experimental unit results in a Gauss-Markov linear model with normal errors for the averages

$$\left\{\bar{y}_{ij\cdot} : i = 1, \ldots, t; j = 1, \ldots, n\right\}.$$

- Inferences about estimable functions of $\boldsymbol{\beta}$ obtained by analyzing these averages are identical to the results obtained using the ANOVA approach as long as the number of multiple observations per experimental unit is the same for all experimental units.

- When using the averages as data, our estimate of σ^2 is an estimate of $\sigma_u^2 + \frac{\sigma_e^2}{m}$.

- We can't separately estimate σ_u^2 and σ_e^2, but this doesn't matter if our focus is on inference for estimable functions of $\boldsymbol{\beta}$.

Because

$$E(\boldsymbol{y}) = \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu + \tau_t \end{bmatrix} \otimes \mathbf{1}_{nm \times 1},$$

the only estimable quantities are linear combinations of the treatment means $\mu + \tau_1, \mu + \tau_2, \ldots, \mu + \tau_t$, whose Best Linear Unbiased Estimators are $\bar{y}_{1\cdot\cdot}, \bar{y}_{2\cdot\cdot}, \ldots, \bar{y}_{t\cdot\cdot}$, respectively.

Thus, any estimable $\boldsymbol{C\beta}$ can always be written as

$$\boldsymbol{A} \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_2 \\ \vdots \\ \mu + \tau_t \end{bmatrix} \text{ for some matrix } \boldsymbol{A}.$$

It follows that the BLUE of $\boldsymbol{C\beta}$ can be written as

$$\boldsymbol{A} \begin{bmatrix} \bar{y}_{1\cdot\cdot} \\ \bar{y}_{2\cdot\cdot} \\ \vdots \\ \bar{y}_{t\cdot\cdot} \end{bmatrix}.$$

Now note that

$$\begin{aligned}
\mathrm{Var}(\bar{y}_{i\cdot\cdot}) &= \mathrm{Var}(\mu + \tau_i + \bar{u}_{i\cdot} + \bar{e}_{i\cdot\cdot}) \\
&= \mathrm{Var}(\bar{u}_{i\cdot} + \bar{e}_{i\cdot\cdot}) \\
&= \mathrm{Var}(\bar{u}_{i\cdot}) + \mathrm{Var}(\bar{e}_{i\cdot\cdot}) \\
&= \frac{\sigma_u^2}{n} + \frac{\sigma_e^2}{nm} \\
&= \frac{1}{n}\left(\sigma_u^2 + \frac{\sigma_e^2}{m}\right) \\
&= \frac{\sigma^2}{n}.
\end{aligned}$$

Thus

$$\mathrm{Var}\left(\begin{bmatrix} \bar{y}_{1\cdot\cdot} \\ \bar{y}_{2\cdot\cdot} \\ \cdot \\ \cdot \\ \cdot \\ \bar{y}_{t\cdot\cdot} \end{bmatrix}\right) = \frac{\sigma^2}{n}\boldsymbol{I}_{t\times t}$$

which implies that the variance of the BLUE of $\boldsymbol{C\beta}$ is

$$\mathrm{Var}\left(\boldsymbol{A}\begin{bmatrix} \bar{y}_{1\cdot\cdot} \\ \cdot \\ \cdot \\ \cdot \\ \bar{y}_{t\cdot\cdot} \end{bmatrix}\right) = \boldsymbol{A}\left(\frac{\sigma^2}{n}\boldsymbol{I}_{t\times t}\right)\boldsymbol{A}' = \frac{\sigma^2}{n}\boldsymbol{A}\boldsymbol{A}'.$$

- Thus, we don't need separate estimates of σ_u^2 and σ_e^2 to carry out inference for estimable $\boldsymbol{C\beta}$.
- We do need to estimate $\sigma^2 = \sigma_u^2 + \frac{\sigma_e^2}{m}$.
- This can equivalently be estimated by

$$\frac{MS_{xu(trt)}}{m}$$

or by the MSE in an analysis of the experimental unit means

$$\left\{\bar{y}_{ij\cdot} : i = 1, \ldots, t; j = 1, \ldots, n.\right\}.$$

For example, suppose we want to estimate $\tau_1 - \tau_2$. The BLUE is $\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}$ whose variance is

$$\begin{aligned}
\text{Var}(\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}) &= \text{Var}(\bar{y}_{1\cdot\cdot}) + \text{Var}(\bar{y}_{2\cdot\cdot}) \\
&= 2\frac{\sigma^2}{n} = 2\left(\frac{\sigma_u^2}{n} + \frac{\sigma_e^2}{nm}\right) \\
&= \frac{2}{nm}(\sigma_e^2 + m\sigma_u^2) \\
&= \frac{2}{nm}E(MS_{xu(trt)})
\end{aligned}$$

Thus,

$$\widehat{\text{Var}}(\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}) = \frac{2MS_{xu(trt)}}{nm}.$$

A $100(1-\alpha)\%$ confidence interval for $\tau_1 - \tau_2$ is

$$\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot} \pm t_{t(n-1),1-\alpha/2}\sqrt{\frac{2MS_{xu(trt)}}{nm}}.$$

A test of $H_0 : \tau_1 = \tau_2$ can be based on

$$t = \frac{\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}}{\sqrt{\frac{2MS_{xu(trt)}}{nm}}} \sim t_{t(n-1)}\left(\frac{\tau_1 - \tau_2}{\sqrt{\frac{2(\sigma_e^2 + m\sigma_u^2)}{nm}}}\right).$$

- What if the number of observations per experimental unit is not the same for all experimental units?

- Let us look at two miniature examples to understand how this type of unbalancedness affects estimation and inference.

First Example

$$\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ y_{211} \\ y_{212} \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

First Example

$$\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ y_{211} \\ y_{212} \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{X}_1 = \mathbf{1}, \qquad \boldsymbol{X}_2 = \boldsymbol{X}, \qquad \boldsymbol{X}_3 = \boldsymbol{Z}$$

First Example

$$\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ y_{211} \\ y_{212} \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{X}_1 = \boldsymbol{1}, \qquad \boldsymbol{X}_2 = \boldsymbol{X}, \qquad \boldsymbol{X}_3 = \boldsymbol{Z}$$

Recall that

$$SS(j+1|j) = \boldsymbol{y}'(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{y} = ||\boldsymbol{P}_{j+1}\boldsymbol{y} - \boldsymbol{P}_j\boldsymbol{y}||^2.$$

First Example

$$\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ y_{211} \\ y_{212} \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{X}_1 = \boldsymbol{1}, \qquad \boldsymbol{X}_2 = \boldsymbol{X}, \qquad \boldsymbol{X}_3 = \boldsymbol{Z}$$

$$\boldsymbol{P}_1 \boldsymbol{y} = \begin{bmatrix} \bar{y}_{\cdots} \\ \bar{y}_{\cdots} \\ \bar{y}_{\cdots} \\ \bar{y}_{\cdots} \end{bmatrix}, \quad \boldsymbol{P}_2 \boldsymbol{y} = \begin{bmatrix} \bar{y}_{1\cdot 1} \\ \bar{y}_{1\cdot 1} \\ \bar{y}_{21\cdot} \\ \bar{y}_{21\cdot} \end{bmatrix}, \quad \boldsymbol{P}_3 \boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ \bar{y}_{21\cdot} \\ \bar{y}_{21\cdot} \end{bmatrix}$$

$$\boldsymbol{P}_1\,\boldsymbol{y} = \begin{bmatrix} \bar{y}_{\cdots} \\ \bar{y}_{\cdots} \\ \bar{y}_{\cdots} \\ \bar{y}_{\cdots} \end{bmatrix}, \quad \boldsymbol{P}_2\,\boldsymbol{y} = \begin{bmatrix} \bar{y}_{1\cdot 1} \\ \bar{y}_{1\cdot 1} \\ \bar{y}_{21\cdot} \\ \bar{y}_{21\cdot} \end{bmatrix}, \quad \boldsymbol{P}_3\,\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ \bar{y}_{21\cdot} \\ \bar{y}_{21\cdot} \end{bmatrix}$$

Thus,

$$\begin{aligned} SS_{trt} &= \boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y} = ||\boldsymbol{P}_2\,\boldsymbol{y} - \boldsymbol{P}_1\,\boldsymbol{y}||^2 \\ &= (\bar{y}_{1\cdot 1} - \bar{y}_{\cdots})^2 + (\bar{y}_{1\cdot 1} - \bar{y}_{\cdots})^2 + (\bar{y}_{21\cdot} - \bar{y}_{\cdots})^2 + (\bar{y}_{21\cdot} - \bar{y}_{\cdots})^2 \\ &= 2(\bar{y}_{1\cdot 1} - \bar{y}_{\cdots})^2 + 2(\bar{y}_{21\cdot} - \bar{y}_{\cdots})^2 = (\bar{y}_{1\cdot 1} - \bar{y}_{21\cdot})^2, \end{aligned}$$

where the last line follows from

$$\bar{y}_{1\cdot 1} - \bar{y}_{\cdots} = \bar{y}_{1\cdot 1} - (\bar{y}_{1\cdot 1} + \bar{y}_{21\cdot})/2 = (\bar{y}_{1\cdot 1} - \bar{y}_{21\cdot})/2$$

and

$$\bar{y}_{21\cdot} - \bar{y}_{\cdots} = \bar{y}_{21\cdot} - (\bar{y}_{1\cdot 1} + \bar{y}_{21\cdot})/2 = -(\bar{y}_{1\cdot 1} - \bar{y}_{21\cdot})/2.$$

Deriving the other sums of squares similarly and noting that $r_1 = 1$, $r_2 = 2$, and $r_3 = 3$ so that the degrees of freedom for each sum of squares is 1, we have

$$MS_{trt} = \boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y} = 2(\bar{y}_{1\cdot 1} - \bar{y}_{\cdots})^2 + 2(\bar{y}_{21\cdot} - \bar{y}_{\cdots})^2 = (\bar{y}_{1\cdot 1} - \bar{y}_{21\cdot})^2$$

$$MS_{xu(trt)} = \boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y} = (y_{111} - \bar{y}_{1\cdot 1})^2 + (y_{121} - \bar{y}_{1\cdot 1})^2 = \frac{1}{2}(y_{111} - y_{121})^2$$

$$MS_{ou(xu,trt)} = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y} = (y_{211} - \bar{y}_{21\cdot})^2 + (y_{212} - \bar{y}_{21\cdot})^2 = \frac{1}{2}(y_{211} - y_{212})^2.$$

$$
\begin{aligned}
E(MS_{trt}) &= E(\overline{y}_{1\cdot 1} - \overline{y}_{21\cdot})^2 \\
&= E(\tau_1 - \tau_2 + \overline{u}_{1\cdot} - u_{21} + \overline{e}_{1\cdot 1} - \overline{e}_{21\cdot})^2 \\
&= (\tau_1 - \tau_2)^2 + \text{Var}(\overline{u}_{1\cdot}) + \text{Var}(u_{21}) + \text{Var}(\overline{e}_{1\cdot 1}) + \text{Var}(\overline{e}_{21\cdot}) \\
&= (\tau_1 - \tau_2)^2 + \frac{\sigma_u^2}{2} + \sigma_u^2 + \frac{\sigma_e^2}{2} + \frac{\sigma_e^2}{2} \\
&= (\tau_1 - \tau_2)^2 + 1.5\sigma_u^2 + \sigma_e^2
\end{aligned}
$$

$$
\begin{aligned}
E(MS_{xu(trt)}) &= \tfrac{1}{2}E(y_{111} - y_{121})^2 \\
&= \tfrac{1}{2}E(u_{11} - u_{12} + e_{111} - e_{121})^2 \\
&= \tfrac{1}{2}(2\sigma_u^2 + 2\sigma_e^2) \\
&= \sigma_u^2 + \sigma_e^2 \\
E(MS_{ou(xu,trt)}) &= \tfrac{1}{2}E(y_{211} - y_{212})^2 \\
&= \tfrac{1}{2}E(e_{211} - e_{212})^2 \\
&= \sigma_e^2
\end{aligned}
$$

SOURCE	EMS
trt	$(\tau_1 - \tau_2)^2 + 1.5\sigma_u^2 + \sigma_e^2$
$xu(trt)$	$\sigma_u^2 + \sigma_e^2$
$ou(xu, trt)$	σ_e^2

With some nontrivial work, it can be shown that

$$F = \left(\frac{MS_{trt}}{1.5\sigma_u^2 + \sigma_e^2}\right) \Big/ \left(\frac{MS_{xu(trt)}}{\sigma_u^2 + \sigma_e^2}\right) \sim F_{1,1}\left(\frac{(\tau_1 - \tau_2)^2}{3\sigma_u^2 + 2\sigma_e^2}\right).$$

The test statistic that we used to test

$$H_0 : \tau_1 = \cdots = \tau_t$$

in the balanced case is not F distributed in this unbalanced case.

$$\frac{MS_{trt}}{MS_{xu(trt)}} \sim \frac{1.5\sigma_u^2 + \sigma_e^2}{\sigma_u^2 + \sigma_e^2} F_{1,1}\left(\frac{(\tau_1 - \tau_2)^2}{3\sigma_u^2 + 2\sigma_e^2}\right)$$

A Statistic with an Approximate F Distribution

- We'd like our denominator to be an unbiased estimator of $1.5\sigma_u^2 + \sigma_e^2$ in this case.

- Consider $1.5MS_{xu(trt)} - 0.5MS_{ou(xu,trt)}$
  The expectation is

$$1.5(\sigma_u^2 + \sigma_e^2) - 0.5\sigma_e^2 = 1.5\sigma_u^2 + \sigma_e^2.$$

- The ratio

$$\frac{MS_{trt}}{1.5MS_{xu(trt)} - 0.5MS_{ou(xu,trt)}}$$

  can be used as an approximate F statistic with 1 numerator DF and a denominator DF obtained using the Cochran-Satterthwaite method.

- The Cochran-Satterthwaite method will be explained in the next set of notes.
- We should not expect this approximate F-test to be reliable in this case because of our pitifully small dataset.

Best Linear Unbiased Estimates in this First Example

What do the BLUEs of the treatment means look like in this case? Recall

$$\boldsymbol{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \mathrm{Var}(\boldsymbol{y}) = \boldsymbol{ZGZ}' + \boldsymbol{R} = \sigma_u^2 \boldsymbol{ZZ}' + \sigma_e^2 \boldsymbol{I} \\
&= \sigma_u^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} + \sigma_e^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_u^2 & 0 & 0 & 0 \\ 0 & \sigma_u^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & \sigma_u^2 \\ 0 & 0 & \sigma_u^2 & \sigma_u^2 \end{bmatrix} + \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_u^2 + \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_u^2 + \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 + \sigma_e^2 & \sigma_u^2 \\ 0 & 0 & \sigma_u^2 & \sigma_u^2 + \sigma_e^2 \end{bmatrix}
\end{aligned}
$$

It follows that

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}} &= (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}\boldsymbol{y} = \begin{bmatrix} \bar{y}_{1\cdot 1} \\ \bar{y}_{21\cdot} \end{bmatrix}\end{aligned}$$

Fortunately, this is a linear estimator that does not depend on unknown variance components.

Second Example

$$
\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{211} \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

In this case, it can be shown that

$$
\begin{aligned}
\hat{\beta}_{\boldsymbol{\Sigma}} &= (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \\
&= \begin{bmatrix} \frac{\sigma_e^2+\sigma_u^2}{3\sigma_e^2+4\sigma_u^2} & \frac{\sigma_e^2+\sigma_u^2}{3\sigma_e^2+4\sigma_u^2} & \frac{\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{211} \end{bmatrix} \\
&= \begin{bmatrix} \frac{2\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2}\,\bar{y}_{11\cdot} + \frac{\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2}\,y_{121} \\ y_{211} \end{bmatrix}.
\end{aligned}
$$

It is straightforward to show that the weights on $\bar{y}_{11\cdot}$ and y_{121} are

$$\frac{\frac{1}{\mathrm{Var}(\bar{y}_{11\cdot})}}{\frac{1}{\mathrm{Var}(\bar{y}_{11\cdot})} + \frac{1}{\mathrm{Var}(y_{121})}} \text{ and } \frac{\frac{1}{\mathrm{Var}(y_{121})}}{\frac{1}{\mathrm{Var}(\bar{y}_{11\cdot})} + \frac{1}{\mathrm{Var}(y_{121})}}, \text{ respectively.}$$

This is a special case of a more general phenomenon: the BLUE is a weighted average of independent linear unbiased estimators with weights for the linear unbiased estimators proportional to the inverse variances of the linear unbiased estimators.

Of course, in this case and in many others,

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}} = \begin{bmatrix} \frac{2\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2}\,\bar{y}_{11\cdot} & + & \frac{\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2}\,y_{121} \\ & y_{211} & \end{bmatrix}$$

is not an estimator because it is a function of unknown parameters.

Thus, we use $\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}$ as our estimator (i.e., we replace σ_e^2 and σ_u^2 by estimates in the expression above).

- $\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}$ is an approximation to the BLUE.
- $\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}$ is not even a linear estimator in this case.
- Its exact distribution is unknown.
- When sample sizes are large, it is reasonable to assume that the distribution of $\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}$ is approximately the same as the distribution of $\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}$.

$$
\begin{aligned}
\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}) &= \mathrm{Var}[(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}] \\
&= (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\mathrm{Var}(\boldsymbol{y})[(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}]' \\
&= (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} \\
\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}) &= \mathrm{Var}[(\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y}] = ???? \approx (\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}
\end{aligned}
$$

Summary of Main Points

- Many of the concepts we have seen by examining special cases hold in greater generality.

- For many of the linear mixed models commonly used in practice, balanced data are nice because...

1. It is relatively easy to determine degrees of freedom, sums of squares, and expected mean squares in an ANOVA table.

2. Ratios of appropriate mean squares can be used to obtain exact F-tests.

3. For estimable $\boldsymbol{C\beta}$, $\boldsymbol{C}\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}} = \boldsymbol{C}\hat{\boldsymbol{\beta}}$. (OLS $=$ GLS).

4. When $\text{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) = \text{constant} \times E(MS)$, exact inferences about $\boldsymbol{c}'\boldsymbol{\beta}$ can be obtained by constructing t-tests or confidence intervals based on
$$t = \frac{\boldsymbol{c}'\hat{\boldsymbol{\beta}} - \boldsymbol{c}'\boldsymbol{\beta}}{\sqrt{\text{constant} \times (MS)}} \sim t_{DF(MS)}.$$

5. Simple analysis based on experimental unit averages gives the same results as those obtained by linear mixed model analysis of the full data set.

When data are unbalanced, the analysis of linear mixed may be considerably more complicated.

1. Approximate F-tests can be obtained by forming linear combinations of Mean Squares to obtain denominators for test statistics.

2. The estimator $\boldsymbol{C}\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}$ may be a nonlinear estimator of $\boldsymbol{C}\boldsymbol{\beta}$ whose exact distribution is unknown.

3. Approximate inference for $\boldsymbol{C}\boldsymbol{\beta}$ is often obtained by using the distribution of $\boldsymbol{C}\hat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}$, with unknowns in that distribution replaced by estimates.

Whether data are balanced or unbalanced, unbiased estimators of variance components can be obtained using linear combinations of mean squares from the ANOVA table.