no office hour today: 3/28/25

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

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

end lecture 24
Wednesday
03-26-25

Expected Mean Squares (EMS) could be computed using

$$\mathrm{E}(\boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{y}) = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + \mathrm{E}(\boldsymbol{y})^\top \boldsymbol{A}\mathrm{E}(\boldsymbol{y}),$$

where

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

and

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

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

accounts for treatment

1. $$\frac{\boldsymbol{y}^\top(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2} \sim \chi^2_{t-1}\left(\underbrace{\frac{nm}{2(\sigma_e^2 + m\sigma_u^2)} \sum_{i=1}^{t} (\tau_i - \bar{\tau}_\cdot)^2}\right),$$

2. now we account for random effect u

$$\frac{\boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2} \sim \chi^2_{tn-t},$$

under Ho: no treatment effect

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

$\tau_1 = \tau_2 = \ldots \tau_t$

$\Rightarrow ncp = 0$

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

It follows that

$$
\begin{aligned}
F_1 &= \frac{MS_{trt}}{MS_{xu(trt)}} = \frac{\boldsymbol{y}^\top(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}/(t-1)}{\boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}/(tn-t)} \\
&= \frac{\left[\frac{\boldsymbol{y}^\top(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}}{\cancel{\sigma_e^2 + m\sigma_u^2}}\right]/(t-1)}{\left[\frac{\boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\cancel{\sigma_e^2 + m\sigma_u^2}}\right]/(tn-t)} \\
&\sim F_{t-1,\, tn-t}\left(\underbrace{\frac{nm}{2(\sigma_e^2 + m\sigma_u^2)}\sum_{i=1}^{t}(\tau_i - \bar{\tau}_{\cdot})^2}_{\text{under } H_0 \ \ ncp = 0}\right).
\end{aligned}
$$

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

Also,

observe

$$
\begin{aligned}
F_2 &= \frac{MS_{xu(trt)}}{MS_{ou(xu,trt)}} = \boxed{\frac{\boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}/(tn-t)}{\boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y}/(tnm-tn)}} \\
&= \boxed{\left(\frac{\sigma_e^2 + m\sigma_u^2}{\sigma_e^2}\right)} \frac{\left[\frac{\boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}}{\sigma_e^2 + m\sigma_u^2}\right]/(tn-t)}{\left[\frac{\boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y}}{\sigma_e^2}\right]/(tnm-tn)} \\
&\sim \left(\frac{\sigma_e^2 + m\sigma_u^2}{\sigma_e^2}\right) F_{tn-t,\, tnm-tn}.
\end{aligned}
$$

Scaled F-distribut.

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

Estimating σ_u^2 : linear Combinations of Mean Squares

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,

$$\frac{MS_{xu(trt)} - MS_{ou(xu,trt)}}{m} = \hat{\sigma}_u^2$$

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. REML

Estimation of Estimable $C\beta$

As we have seen previously,

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

It turns out that

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

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

An Analysis Based on the Average for Each Experimental Unit

is possible when we have balanced data & focus is on $C\beta$

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

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

If we define

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

and

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

we have

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

implication: we no longer are able to seperate σ_u^2 & σ_e^2/m

where the ϵ_{ij} terms are $iid\ \mathcal{N}(0, \sigma^2)$. 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

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

- 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 $\boldsymbol{\beta}$.

Estimation of $C\beta$ (treatment effect)

Because

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

estimable: $\mu + \tau_i$ + any linear combination: $(\mu+\tau_1)-(\mu+\tau_2)$ $\tau_1 - \tau_2$ for example

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 $\overline{y}_{1\cdot\cdot}, \overline{y}_{2\cdot\cdot}, \ldots, \overline{y}_{t\cdot\cdot}$, 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

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

Now note that

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

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

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

- 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

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

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

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

Thus,

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

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

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

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

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

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

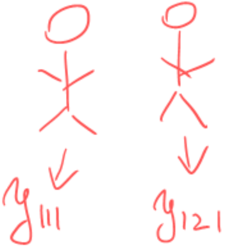
- Let us look at two miniature examples to understand how this type of unbalancedness affects estimation and inference.

First Example

2 treatments, 3 exp. units, 4 obs.

$$\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{121} \\ \hline y_{211} \\ y_{212} \end{bmatrix} \begin{array}{l} \text{trt 1} \\ \\ \text{trt 2} \end{array}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

trt 1:



$\Rightarrow$ 2 ind. exp. units but no replication: $m = 1$

trt 2:

1 exp. unit with $m = 2$ replications

trt1 trt 2

1st exp. unit

2nd exp. unit

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

3rd exp. unit

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

intercept only model

$C(X_1)$

trt into account

$C(X_2)$

accounting for random effect associated with each exp. unit

$C(X_3)$

First Example

End lecture 25
3-28-25

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

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

Recall that

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