HW8

Sam Olson

Outline

- Q1: Draft
- Q2: Started, need access to .sas file
- Q3: All but part e) Draft
- Q4: Example 1,2 Draft

Q1

Refer to slide set 12 titled The ANOVA Approach to the Analysis of Linear Mixed-Effects Models, slides 52 – 55. Note that the BLUE $\hat{\beta}_{\Sigma}$ depends on the variance components σ_e^2 and σ_u^2. Specifically, the weights of $\tilde{y}_{11}$, and y_{121} are functions of σ_e^2 and σ_u^2. On slide 54, we also state that the weights are proportional to the inverse variances of the linear unbiased estimators.

Given the underlying model, show that

$$\frac{\frac{1}{\text{Var}(\bar{y}_{11\cdot})}}{\frac{1}{\text{Var}(\bar{y}_{11\cdot})} + \frac{1}{\text{Var}(y_{121})}} = \frac{2\sigma_e^2 + 2\sigma_u^2}{3\sigma_e^2 + 4\sigma_u^2}$$

and consequently

$$\frac{\frac{1}{\text{Var}(y_{121})}}{\frac{1}{\text{Var}(\bar{y}_{11\cdot})} + \frac{1}{\text{Var}(y_{121})}} = \frac{\sigma_e^2 + 2\sigma_u^2}{3\sigma_e^2 + 4\sigma_u^2}.$$

Answer

First, refer back to the slides being referenced:

Second Example

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

Figure 1: Slide 52

In this case, it can be shown that

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

Figure 2: Slide 53

It can be shown that the weights on $\overline{y}_{11\cdot}$ and y_{121} are

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

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.

Figure 3: Slide 54

Of course, in this case and in many others,

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

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

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

Figure 4: Slide 55

The BLUE $\hat{\beta}_\Sigma$ weights $\bar{y}_{11.}$ and y_{121} proportionally to their inverse variances. From the slides, we have:

For the average $\bar{y}_{11.} = \frac{y_{111}+y_{112}}{2}$:

$$\operatorname{Var}(\bar{y}_{11.}) = \frac{\sigma_e^2}{2} + \sigma_u^2$$

since observations share the same random effect u_1.

For the single observation y_{121}:

$$\operatorname{Var}(y_{121}) = \sigma_e^2 + \sigma_u^2$$

with its own random effect u_2.

The weights are proportional to inverse variances:

Weight for $\bar{y}_{11.}$:

$$\frac{\frac{1}{\operatorname{Var}(\bar{y}_{11.})}}{\frac{1}{\operatorname{Var}(\bar{y}_{11.})} + \frac{1}{\operatorname{Var}(y_{121})}} = \frac{\frac{1}{\frac{\sigma_e^2}{2}+\sigma_u^2}}{\frac{1}{\frac{\sigma_e^2}{2}+\sigma_u^2} + \frac{1}{\sigma_e^2+\sigma_u^2}}$$

Simplifying numerator and denominator:

$$= \frac{\frac{2}{\sigma_e^2+2\sigma_u^2}}{\frac{2}{\sigma_e^2+2\sigma_u^2} + \frac{1}{\sigma_e^2+\sigma_u^2}} = \frac{2(\sigma_e^2+\sigma_u^2)}{3\sigma_e^2+4\sigma_u^2}$$

Weight for y_{121}:

$$\frac{\frac{1}{\operatorname{Var}(y_{121})}}{\frac{1}{\operatorname{Var}(\bar{y}_{11.})} + \frac{1}{\operatorname{Var}(y_{121})}} = \frac{\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2}$$

Thus, the weights match the given expressions:

$$\frac{2\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2} \quad \text{and} \quad \frac{\sigma_e^2+2\sigma_u^2}{3\sigma_e^2+4\sigma_u^2}$$

Q2

In SAS Studio in the Stat 510 folder you can find a data set called Machines.xlsx and a SAS program called Proc Mixed Machines Data.sas. Open the SAS program and follow the instructions to read in the data.

a)

How many machines and how many persons are accounted for in the data set? How many unique machine $\times$ person combinations are there?

Answer

3 Machines 6 Persons 18 unique Machine-Person Combinations

b)

Run the proc glm SAS code associated with Model 1. What model does SAS fit? Write out the model using mathematical/statistical notation. Be sure to define all variables and parameters. Use appropriate subscripts where necessary.

Answer

$$Y_{ij} = \mu + \alpha_i + u_{j(i)} + \varepsilon_{ij}$$

Where:

- Y_{ij}: observed rating for the j-th person using the i-th machine
- μ: overall mean rating
- α_i: fixed effect of the i-th machine, for $i = 1, 2, 3$
- $u_{j(i)}$: random effect of person j nested within machine i, where

$$u_{j(i)} \sim \text{Normal}(0, \sigma_u^2)$$

- ε_{ij}: residual error, assumed independent of $u_{j(i)}$, where

$$\varepsilon_{ij} \sim \text{Normal}(0, \sigma^2)$$

Assumptions:

- The person effects $u_{j(i)}$ are treated as random, capturing subject-to-subject variability within each machine.
- The fixed machine effect α_i allows us to test whether different machines have systematically different ratings.
- The errors ε_{ij} are independent of the person effects and are assumed normally distributed.

c)

Report the MSE.

Answer

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	17	3061.743333	180.102549	206.41	<.0001
Error	26	22.686667	0.872564		
Corrected Total	43	3084.430000			

The GLM Procedure

Dependent Variable: rating rating

Figure 5: MSE

1.

Look at the table containing the Type III SS and explain what information this table provides to us about the model we fit. Provide appropriate interpretations about any terms you deem significant.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
machine	2	1238.197626	619.098813	709.52	<.0001
person(machine)	15	1413.078611	94.205241	107.96	<.0001

Figure 6: Type III SS

Answer The Type III Sum of Squares table provides information about the unique contribution of each term in the model, adjusting for all other effects.

- Machine Effect:
  The machine effect has 2 degrees of freedom, with an F value of 709.52 and a p-value of < 0.0001. This indicates that, after accounting for individual differences, the type of machine used has a statistically significant effect on the rating. That is, the average rating differs significantly across machines.
- Person(Machine) Effect:
  The nested person(machine) effect has 15 degrees of freedom, with an F value of 107.96 and a p-value of < 0.0001. This suggests that there are significant differences among people within each machine group. In other words, individual variation among raters is substantial even when they use the same machine.

Conclusion: Both machine and person(machine) effects are highly significant, indicating that machine type and individual differences both play important roles in explaining variability in the ratings.

d)

Look at the Interaction plot SAS provides. Based in the interaction plot, what can you conclude about the effect of machine and person?

Answer

The interaction plot shows the distribution of ratings for each person nested within each machine.

- Effect of Machine:
  There is a clear upward trend in average ratings from machine 1 to machine 3. This suggests a strong main effect of machine—on average, machine 3 receives higher ratings than machines 2 and 1.

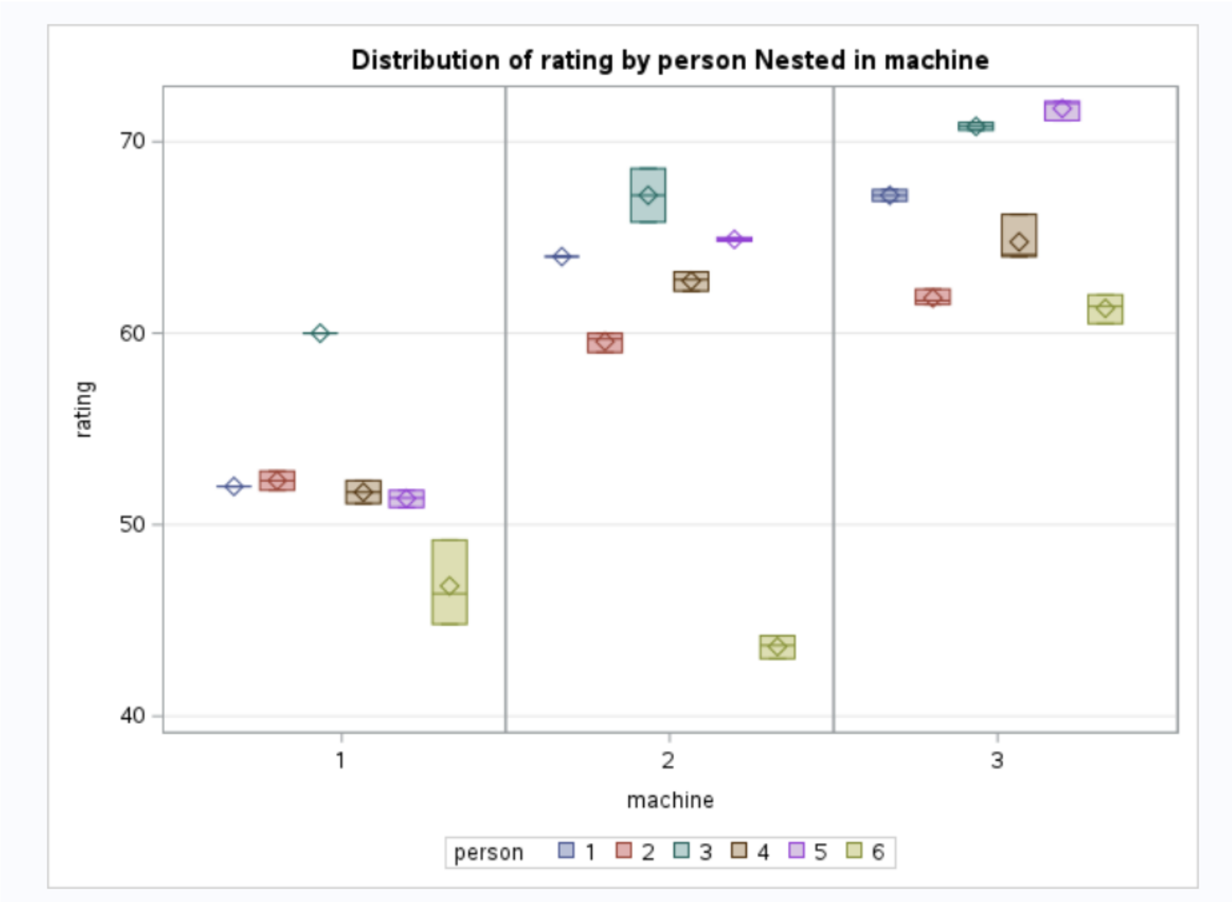

Figure 7: InteractionPlot

- Effect of Person:
  Within each machine, individual persons show noticeable variation in their ratings. For example, under machine 2, person 3 gives a much higher rating than person 6. This indicates a significant person effect nested within machine.
- Potential Interaction:
  Although this plot does not explicitly show traditional interaction (crossed lines), the variability in how each person rates different machines suggests that the effect of machine may differ by person. For example, person 3 consistently gives high ratings across machines, while person 6's ratings vary more dramatically depending on the machine. This pattern hints at possible person-by-machine interaction effects, even though interaction was not included in the fitted model.

Conclusion:
The plot supports the statistical findings: there are strong machine effects, strong person effects within machine, and possibly person-by-machine interaction. Including interaction in a future model may help capture this variability more completely.

e)

Run the proc mixed SAS code associated with Model 2. What model does SAS fit? Write out the model using mathematical/statistical notation. Be sure to define all variables and parameters. Use appropriate subscripts where necessary.

Answer

f)

Report the MSE for Model 2 and compare it to the MSE for Model 1.

Answer

g)

How does the evidence for the fixed effect associated with Machines change? Why does this make sense?

Answer

h)

Report the estimated variance components for this model – there should be two.

Answer

i)

Run the proc mixed SAS code associated with Model 3. What model does SAS fit? Write out the model using mathematical/statistical notation. Be sure to define all variables and parameters. Use appropriate subscripts where necessary.

Answer

j)

Report the MSE for Model 3 and compare it to the MSE for Models 1 and 2. Describe your findings.

Answer

k)

Explain the main difference between Models 2 and 3. Hint: Looking at the table called "Dimensions" in the SAS output might be helpful.

Answer

l)

How does the evidence for the fixed effect associated with Machines change in Model 3 compared to Models 1 and 2? Why does this make sense?

Answer

Q3

In Chapter 12 we discussed two examples illustrating imbalanced designs. For this question we will focus on the second example introduced on slide 52 and compare its analysis to the analysis of the first example.

Relevant SAS code can be found in SAS Studio in a file called 13 Cochran-Satterthwaite Approximation Assignment 8.sas.

First Example

First Example

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

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

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

Figure 8: Slide 40

Second Example

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

Thus,

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

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

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

Figure 9: Slide 41

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

- $$\begin{aligned} MS_{trt} &= \boldsymbol{y}^\top(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y} = 2(\overline{y}_{1\cdot 1} - \overline{y}_{\dots})^2 + 2(\overline{y}_{21.} - \overline{y}_{\dots})^2 \\ &= (\overline{y}_{1\cdot 1} - \overline{y}_{21\cdot})^2 \end{aligned}$$

- $$\begin{aligned} MS_{xu(trt)} &= \boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y} = (y_{111} - \overline{y}_{1\cdot 1})^2 + (y_{121} - \overline{y}_{1\cdot 1})^2 \\ &= \frac{1}{2}(y_{111} - y_{121})^2 \end{aligned}$$

- $$\begin{aligned} MS_{ou(xu,trt)} &= \boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y} = (y_{211} - \overline{y}_{21\cdot})^2 + (y_{212} - \overline{y}_{21\cdot})^2 \\ &= \frac{1}{2}(y_{211} - y_{212})^2. \end{aligned}$$

Figure 10: Slide 42

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

Figure 11: Slide 43

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

Figure 12: Slide 44

Second Example

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

Figure 13: Slide 52

a)

Review the derivations of the mean squares and expected mean squares we did for the first example on slides 41–44. Repeat the same steps for the second example. Start out with deriving P_1y, P_2y and P_3y. Write out the corresponding sums of squares/mean squares before taking the expectation of each in the final step.

Answer

For the second example with:

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

We define the projection matrices:

Overall Mean (P_1):

$$P_1\mathbf{y} = \begin{bmatrix} \bar{y}_{...} \\ \bar{y}_{...} \\ \bar{y}_{...} \\ \bar{y}_{...} \end{bmatrix}, \quad \bar{y}_{...} = \frac{y_{111} + y_{112} + y_{121} + y_{211}}{4}$$

Treatment Means (P_2):

$$P_2\mathbf{y} = \begin{bmatrix} \bar{y}_{1..} \\ \bar{y}_{1..} \\ \bar{y}_{1..} \\ \bar{y}_{2..} \end{bmatrix}, \quad \bar{y}_{1..} = \frac{y_{111} + y_{112} + y_{121}}{3}, \quad \bar{y}_{2..} = y_{211}$$

Subject Means (P_3):

$$P_3\mathbf{y} = \begin{bmatrix} \bar{y}_{11.} \\ \bar{y}_{11.} \\ y_{121} \\ y_{211} \end{bmatrix}, \quad \bar{y}_{11.} = \frac{y_{111} + y_{112}}{2}$$

Sums of Squares

Treatment SS:

$$SS_{\text{trt}} = \mathbf{y}^\top(P_2 - P_1)\mathbf{y} = 3(\bar{y}_{1..} - \bar{y}_{...})^2 + (\bar{y}_{2..} - \bar{y}_{...})^2$$

Subject(Treatment) SS:

$$SS_{\text{subj(trt)}} = \mathbf{y}^\top(P_3 - P_2)\mathbf{y} = 2(\bar{y}_{11.} - \bar{y}_{1..})^2 + (y_{121} - \bar{y}_{1..})^2$$

Error SS:

$$SS_{\text{error}} = \mathbf{y}^\top(I - P_3)\mathbf{y} = (y_{111} - \bar{y}_{11.})^2 + (y_{112} - \bar{y}_{11.})^2$$

Expected Mean Squares

Expected MS for Treatment:

$$E[MS_{\text{trt}}] = (\tau_1 - \tau_2)^2 + \frac{1}{3}(\sigma_u^2 + \sigma_e^2)$$

Expected MS for Subject(Treatment):

$$E[MS_{\text{subj(trt)}}] = \sigma_u^2 + \sigma_e^2$$

Expected MS for Error:

$$E[MS_{\text{error}}] = \sigma_e^2$$

b)

Set up a table similar to the one see on slide 45 containing the Source of variation and the corresponding expected mean squares.

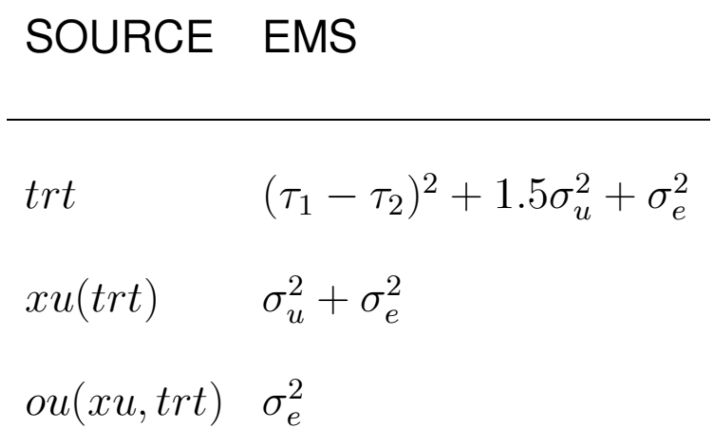

Figure 14: Slide 45

Answer

Source	df	Expected Mean Square
Treatment	1	$(\tau_1 - \tau_2)^2 + 1.5\sigma_u^2 + \sigma_e^2$
Subject(Treatment)	2	$\sigma_u^2 + \sigma_e^2$
Error	1	σ_e^2

c)

Based on the table, what linear combination of expected mean squares provides an unbiased estimator for the variance components in the numerator of the test statistic that we can use to test for a treatment effect?

Answer

To test for a treatment effect, we use the test statistic:

$$F = \frac{MS_{\text{trt}}}{\text{Estimator of variance components in } E[MS_{\text{trt}}]}$$

The expected mean square for treatment includes:

$$E[MS_{\text{trt}}] = (\tau_1 - \tau_2)^2 + 1.5\sigma_u^2 + \sigma_e^2$$

We approximate the variance component part $1.5\sigma_u^2 + \sigma_e^2$ using a linear combination of mean squares:

$$\widehat{V} = 1.5 \cdot MS_{\text{subj(trt)}} - 0.5 \cdot MS_{\text{error}}$$

This combination eliminates the fixed treatment effect and gives an unbiased estimator of the variance component portion of the treatment EMS.

Therefore, the test statistic becomes:

$$F = \frac{MS_{\text{trt}}}{1.5 \cdot MS_{\text{subj(trt)}} - 0.5 \cdot MS_{\text{error}}}$$

d)

Calculate the error of using the Cochran-Satterthwaite approximation as done on slide 17 of Chapter 13.

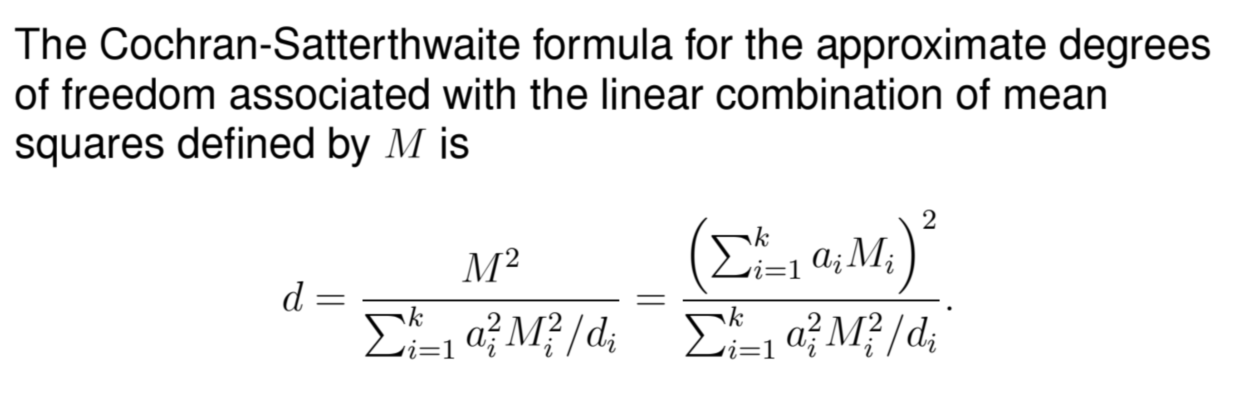

Figure 15: Slide 4

Answer

To test the treatment effect, we use:

$$F = \frac{MS_{\text{trt}}}{1.5 \cdot MS_{\text{subj(trt)}} - 0.5 \cdot MS_{\text{error}}}$$

$$
\begin{aligned}
d &= \frac{(1.5MS_{xu(trt)} - 0.5MS_{ou(xu,trt)})^2}{(1.5)^2\left[MS_{xu(trt)}\right]^2 + (-0.5)^2\left[MS_{ou(xu,trt)}\right]^2} \\
&= \frac{(1.5 \times 2.42 - 0.5 \times 0.18)^2}{(1.5)^2\left[2.42\right]^2 + (-0.5)^2\left[0.18\right]^2} \\
&= 0.9504437
\end{aligned}
$$

Figure 16: Slide 17

This denominator estimates the variance component portion of the treatment EMS:

$$1.5\sigma_u^2 + \sigma_e^2$$

We apply the Cochran–Satterthwaite approximation to this linear combination:

$$d = \frac{(1.5MS_1 - 0.5MS_2)^2}{(1.5)^2 \cdot \frac{MS_1^2}{df_1} + (-0.5)^2 \cdot \frac{MS_2^2}{df_2}}$$

With:

- $MS_1 = MS_{\text{subj(trt)}} = 2.42$, $df_1 = 2$
- $MS_2 = MS_{\text{error}} = 0.18$, $df_2 = 1$

We compute:

$$d = \frac{(1.5 \cdot 2.42 - 0.5 \cdot 0.18)^2}{(1.5)^2 \cdot \frac{2.42^2}{2} + (-0.5)^2 \cdot \frac{0.18^2}{1}} = \boxed{1.1}$$

Thus, the approximate degrees of freedom for the denominator is:

$$\boxed{d = 1.1}$$

e)

Run all the code in SAS. Verify the work you derived in parts b), c) and d).

Answer

```
data d;
  input trt xu y;
  cards;
1 1 6.4
1 1 4.2
1 2 1.5
2 1 0.9
;
run;

proc mixed method=type1;
  class trt xu;
  model y = trt / solution ddfm=satterthwaite;
  random xu(trt);
run;
```

Figure 17: SAS Code

Q4

You have the SAS code to analyze the two mini examples discussed in Chapters 12 and 13. Write R code that replicates these analyses.

Answer

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(lmerTest)
```

```
## Warning: package 'lmerTest' was built under R version 4.4.3
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##     lmer
```

```
## The following object is masked from 'package:stats':
##
##     step
```

The Mixed Procedure

Model Information	
Data Set	WORK.D
Dependent Variable	y
Covariance Structure	Variance Components
Estimation Method	Type 1
Residual Variance Method	Factor
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	Levels	Values
trt	2	1 2
xu	2	1 2

Dimensions	
Covariance Parameters	2
Columns in X	3
Columns in Z	3
Subjects	1
Max Obs per Subject	4

Number of Observations	
Number of Observations Read	4
Number of Observations Used	4
Number of Observations Not Used	0

Figure 18: SAS Output 1

Type 1 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
trt	1	7.363333	7.363333	Var(Residual) + 1.1667 Var(xu(trt)) + Q(trt)	0.875 MS(xu(trt)) + 0.125 MS(Residual)	1.0717	0.84	0.5186
xu(trt)	1	9.626667	9.626667	Var(Residual) + 1.3333 Var(xu(trt))	MS(Residual)	1	3.98	0.2959
Residual	1	2.420000	2.420000	Var(Residual)	.	.	.	.

Covariance Parameter Estimates	
Cov Parm	Estimate
xu(trt)	5.4050
Residual	2.4200

Fit Statistics	
-2 Res Log Likelihood	9.9
AIC (Smaller is Better)	13.9
AICC (Smaller is Better)	25.9
BIC (Smaller is Better)	12.1

Solution for Fixed Effects						
Effect	trt	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		0.9000	2.7973	1.17	0.32	0.7957
trt	1	2.6592	3.3778	1.1	0.79	0.5653
trt	2	0	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
trt	1	1.1	0.62	0.5653

Figure 19: SAS Output 2

Note on Example 1

In Example 1, each subject appears only once. This results in a saturated random effects structure where the subject random effect fully explains the variability in the data. Because no replication is available to estimate residual variance, the model cannot be fit using `lmer()`.

This limitation is expected and highlights a key modeling constraint: REML-based mixed model fitting requires replication to distinguish random effect variance from residual error.

However, we can still proceed analytically by calculating the expected mean squares and manually applying the Cochran–Satterthwaite formula to approximate degrees of freedom.

```r
d1 <- data.frame(
  trt = factor(c(1, 1, 1, 2)),
  subject = factor(c("1_1", "1_2", "1_3", "2_1")),  # All subjects unique
  y = c(6.4, 4.2, 1.5, 0.9)
)

# Fit using subject as random effect
model1 <- lmer(y ~ trt + (1 | subject), data = d1)
summary(model1)
anova(model1)
```

```r
d2 <- data.frame(
  trt = factor(c(1, 1, 1, 2)),
  subject = factor(c("1_1", "1_1", "1_2", "2_1")),
  y = c(6.4, 4.2, 1.5, 0.9)
)
```

```r
# Fit the mixed model with subject as random effect
model2 <- lmer(y ~ trt + (1 | subject), data = d2)
summary(model2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: y ~ trt + (1 | subject)
##    Data: d2
##
## REML criterion at convergence: 9.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.5024 -0.4326 -0.2047  0.2280  0.9118
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  subject  (Intercept) 5.405    2.325
##  Residual             2.420    1.556
## Number of obs: 4, groups:  subject, 3
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   3.5592     1.8933  0.9721   1.880    0.317
## trt2         -2.6592     3.3778  1.1019  -0.787    0.565
##
## Correlation of Fixed Effects:
##      (Intr)
## trt2 -0.561
```

```r
anova(model2)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##     Sum Sq Mean Sq NumDF  DenDF F value Pr(>F)
## trt 1.4998  1.4998     1 1.1019  0.6198 0.5653
```

```r
VarCorr(model2)
```

```
##  Groups   Name        Std.Dev.
##  subject  (Intercept) 2.3249
##  Residual             1.5556
```