HW8

Sam Olson

Outline

- Q3: More explicit calculations
- Q4: Reconcile R code vs. manual calculations in R

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

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

and consequently

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

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

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

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 of $\hat{\beta}_\Sigma$ uses weights on $\bar{y}_{11.}$ and y_{121} proportionally to their inverse variances.

From the slides, we have:

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

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

(Since the two observations share the same random effect u_1)

For the single observation y_{121}, with Variance:

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

(As no other observations share the one random effect u_2)

As the weights are proportional to inverse variances, we have:

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

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

And the weight for y_{121}:

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

Thus, the weights match the given expressions:

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

And:

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

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

Let:

- Y_{ij}: observed rating given by person j on machine i
- $i = 1, 2, 3$: machine levels
- $j = 1, 2, \ldots, 6$: person levels
- $k = 1, 2, 3$: replications of experimental units; not necessarily balanced across all machine-person levels.

The Model formula is given by:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

Where:

- Y_{ijk}: Rating for the k-th observation of the i-th machine and j-th person.
- μ: Overall mean, no subscripts.
- α_i: Fixed effect of machine i.
- β_j: Fixed effect of person j.
- γ_{ij}: Fixed interaction effect between machine i and person j.
- ε_{ijk}: Residual error, with assumption $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

c)

Report the MSE.

Answer

Reported MSE is 0.872564.

The GLM Procedure

Dependent Variable: rating rating

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			

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	5	1011.053834	202.210767	231.74	<.0001
machine*person	10	404.315028	40.431503	46.34	<.0001

Figure 6: Type III SS

Answer All three fixed effects (machine, person, and machine-person interactions) are statistically significant at the 0.0001 level, meaning we have overwhelming evidence against the null hypothesis for each of the sources in the table. Within the context of the study, this is evidence that (1) There are meaningful differences in average ratings across machines, even after accounting for the effects of people and machine-people interactions (2) There are meaningful differences in average ratings across people, even after accounting for the effects of machine and machine-people interactions, and (3) There are meaningful interaction effects, i.e. that machine differences (effects) are not consistent across people.

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

Effect of Machine: There is a general upward trend across all lines from machine 1 to machine 3, though not always from machine 1 to machine 2. This is evidence that machine 3 on average received the highest ratings, followed by machine 2, and then machine 1. This is consistent with the ANOVA table where machine had a statistically significant F-statistic.

Effect of Person: The lines representing each person are distinctly separated, despite some intersections when moving from machine 1 to machine 2. Overall this shows that that different individuals tended to receive different ratings across machines. In particular, person 3 tended to receive the highest average ratings while person 6 tended to receive the lowest average ratings. This is consistent with the ANOVA table where person had a statistically significant F-statistic.

Interaction Effect: The lines are not parallel, changing slope or even intersecting (changing signage). The lines not being parallel indicate an interaction effect, i.e., that the effect of machine on the rating depends

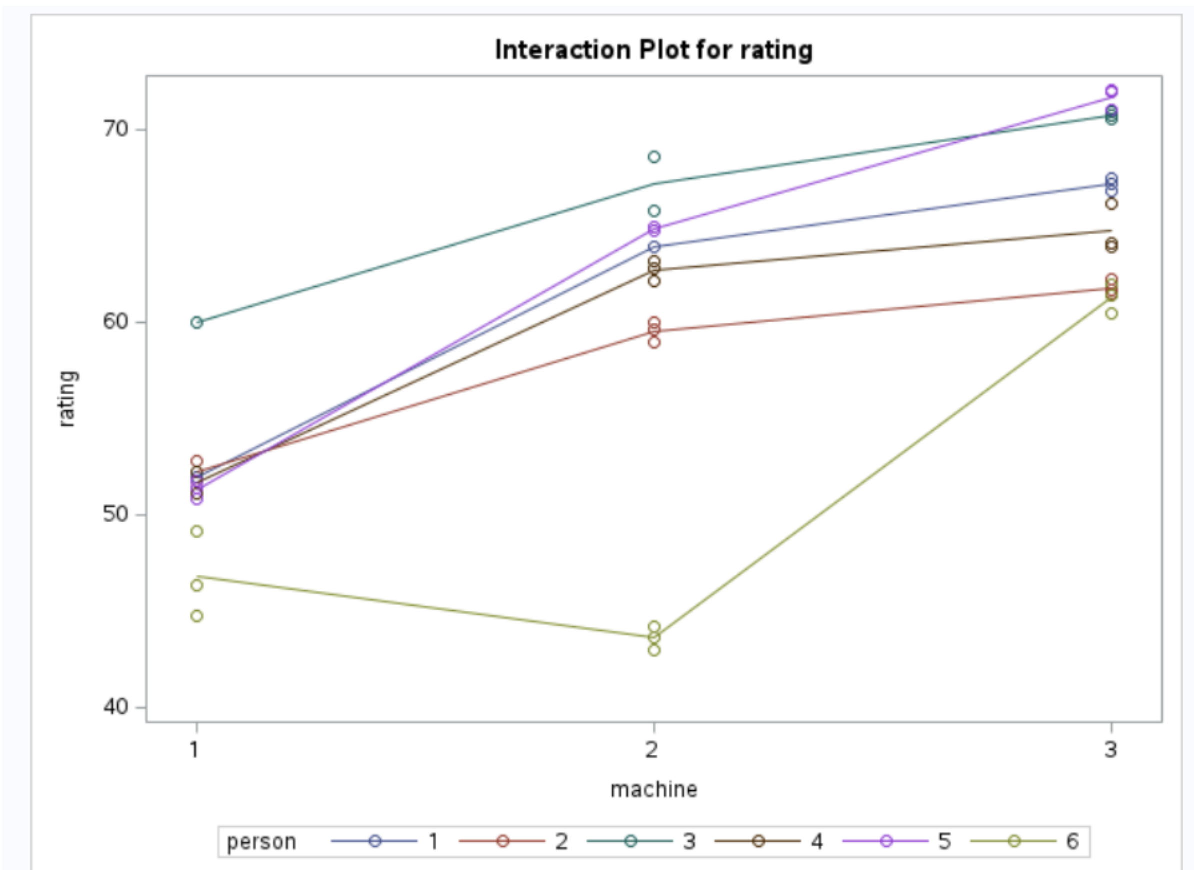


Figure 7: InteractionPlot

on the person. This is also consistent with the ANOVA table where person-machine interaction effect had a statistically significant F-statistic.

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

Let:

- Y_{ij}: observed rating given by person j on machine i
- $i = 1, 2, 3$: machine levels
- $j = 1, 2, \ldots, 6$: person levels
- $k = 1, 2, 3$: replications of experimental units; not necessarily balanced across all machine-person levels.

The Model formula is given by:

$$Y_{ij} = \mu + \alpha_i + u_j + \varepsilon_{ij}$$

Where:

- μ is the overall mean rating
- α_i is the fixed effect of machine i, with sum-to-zero constraint $\sum_i \alpha_i = 0$
- $u_j \sim \mathcal{N}(0, \sigma_u^2)$ is the random effect of person j
- $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ is the residual error, assumed independent of u_j

f)

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

Covariance Parameter Estimates	
Cov Parm	Estimate
person	24.3840
Residual	11.8517

Figure 8: Model 2 Residual

Answer

The MSE for Model 2 is 11.8517, as reported in the table. This is larger than the MSE for Model 1. Though expected due to the inclusion of person as a random effect and the exclusion of the interaction term in the model, is does appear Model 2 has a "worse" fit compared to Model 1 in having a larger MSE.

g)

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

Answer

Model 1: The F-statistic for the fixed effect of machine is 709.52, with $p < .0001$.

Model 2: The F-statistic for machine drops to 58.41, though the p-value remains $< .0001$.

While both models show overwhelming evidence for a machine effect (both having $p < 0.0001$), the F-statistic is much larger in Model 1 because it treats person and interaction effects as fixed, artificially reducing residual variance. Model 2 accounts for person-level variability (main effect of person) as random, which is a more conservative estimate of the model variability. In a word: The smaller F-statistic in Model 2 reflects better partitioning of the variance components, which should not (probably) be conflated as "weaker evidence".

h)

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

Covariance Parameter Estimates	
Cov Parm	**Estimate**
person	24.3840
Residual	11.8517

Figure 9: Model 2 Residual, Again

Answer

The two estimated variance components for this model are:

Person (random effect): 24.3840

Residual (error term): 11.8517

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

Let:

- Y_{ij}: observed rating given by person j on machine i
- $i = 1, 2, 3$: machine levels
- $j = 1, 2, \ldots, 6$: person levels
- $k = 1, 2, 3$: replications of experimental units; not necessarily balanced across all machine-person levels.

The Model formula is given by:

$$Y_{ijk} = \mu + \alpha_i + u_{ij} + \varepsilon_{ijk}$$

Where:

- μ: overall mean rating
- α_i: fixed effect of machine i, with $\sum_i \alpha_i = 0$
- $u_{ij} \sim \mathcal{N}(0, \sigma_u^2)$: random effect for person j nested within machine i
- $\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)$: residual error, independent of the random effects for person

j)

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

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
Intercept	person(machine)	36.6803
Residual		0.8721

Figure 10: Model 3 Residual

Answer

Model 3 MSE is 0.8721.

This is nearly identical to the MSE from Model 1 (0.8726) and substantially lower than the MSE from Model 2 (11.8517). This similarity makes sense because both Models 1 and 3 account interaction effects between person and machine, albeit somewhat differently–Model 1 through a fixed interaction, and Model 3 through a random effect nested within machine. In both cases, our model has smaller MSE through the inclusion of an interaction effect of some kind.

k)

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

Dimensions	
Covariance Parameters	2
Columns in X	4
Columns in Z	6
Subjects	1
Max Obs per Subject	44

Figure 11: Model 2 Dimensions

Dimensions	
Covariance Parameters	2
Columns in X	4
Columns in Z per Subject	1
Subjects	18
Max Obs per Subject	3

Figure 12: Model 3 Dimensions

Answer

Dimensions tables for Models 2 and 3 are provided. There is a noticeable differences between the two models considered in this question:

Model 2 treats `person` as a single random effect, meaning each person has one random intercept across all machines, hence why the Dimensions table reports 18 subjects, each with up to 3 observations per person-machine combination.

By contrast, Model 3 uses `person(machine)` as the subject for the random effect, treating each person–machine combination as a unique random level, which in turn is why the respective Dimensions table reports only 1 subject, but with 44 observations, and 6 columns in Z, indicating 6 random effects to account for nesting `person` within `machine`.

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

In Model 3, the F-statistic for the fixed effect of `machine` remains highly significant, which is consistent with Models 1 and 2. The strength of evidence does not meaningfully decrease compared to Model 1, and it is stronger than in Model 2, based on comparisons of F-statistics.

Similarity with Model 1 makes sense because Model 3 accounts for possible interaction effects between person and machine. There is also some sense to be made by them being somewhat different too, as Model 1 has fixed interaction effects whereas Model 3 has random interaction effects, in part due to the nesting of `person` within `machine`.

In comparison with Model 2, Model 3 is better able to partition the sources of variation through the interaction effect. By contrast, Model 2 treats `person` as a single random effect without allowing variation across machines.

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 13: 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,

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

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

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

Figure 14: 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

- $$\begin{aligned} MS_{trt} &= \boldsymbol{y}^\top(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y} = 2(\overline{y}_{1\cdot 1} - \overline{y}_{\dots})^2 + 2(\overline{y}_{21.} - \overline{y}_{\dots})^2 \\ &= (\overline{y}_{1\cdot 1} - \overline{y}_{21.})^2 \end{aligned}$$
- $$\begin{aligned} MS_{xu(trt)} &= \boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y} = (y_{111} - \overline{y}_{1\cdot 1})^2 + (y_{121} - \overline{y}_{1\cdot 1})^2 \\ &= \frac{1}{2}(y_{111} - y_{121})^2 \end{aligned}$$
- $$\begin{aligned} MS_{ou(xu,trt)} &= \boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y} = (y_{211} - \overline{y}_{21.})^2 + (y_{212} - \overline{y}_{21.})^2 \\ &= \frac{1}{2}(y_{211} - y_{212})^2. \end{aligned}$$

Figure 15: Slide 42

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

Figure 16: 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 17: 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 18: 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 $P_1 y, P_2 y$ and $P_3 y$. Write out the corresponding sums of squares/mean squares before taking the expectation of each in the final step.

Answer

As given in the slides:

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

Projection Matrices

$P_1 = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$

$P_2 = [\mathbf{X}\ \mathbf{Z}]\left([\mathbf{X}\ \mathbf{Z}]^T[\mathbf{X}\ \mathbf{Z}]\right)^{-1}[\mathbf{X}\ \mathbf{Z}]^T$

$P_3 = \mathbf{I} - P_2$

Probably need to actually calculate those matrices, don't I?

Sums of Squares

Treatment sum of squares:

$$SS_{\text{trt}} = \mathbf{y}^T P_1 \mathbf{y}$$

Subject within treatment, random effect:

$$SS_{\text{xu(trt)}} = \mathbf{y}^T (P_2 - P_1)\mathbf{y}$$

Residual:

$$SS_{\text{Residual}} = \mathbf{y}^T (\mathbf{I} - P_2)\mathbf{y}$$

Mean Squares are the same as Sum of Squares, because…

Degrees of freedom: $\text{df}_{\text{trt}} = 1$, $\text{df}_{\text{xu(trt)}} = 1$, $\text{df}_{\text{Residual}} = 1$

Expected Mean Squares

$$\begin{aligned} E[\text{MS}_{\text{trt}}] &= (\tau_1 - \tau_2)^2 + 1.1667\sigma_u^2 + \sigma_e^2 \\ E[\text{MS}_{\text{xu(trt)}}] &= 1.3333\sigma_u^2 + \sigma_e^2 \\ E[\text{MS}_{\text{Residual}}] &= \sigma_e^2 \end{aligned}$$

Primarily based on SAS output, need to actually calculate these

b)

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

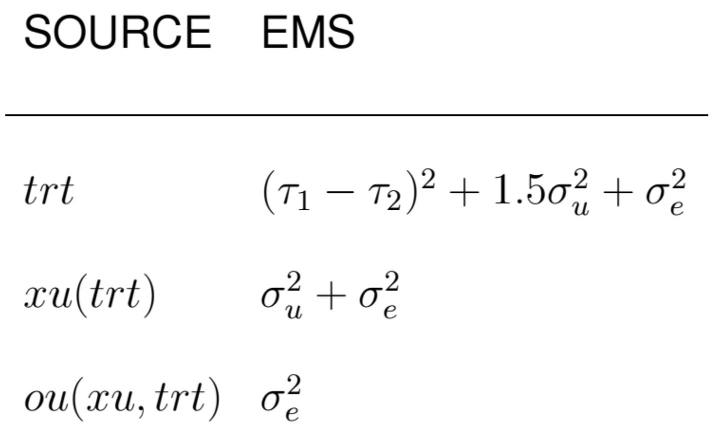


Figure 19: Slide 45

Answer

Source	df	Expected Mean Square
Treatment	1	$(\tau_1 - \tau_2)^2 + \frac{7}{6}\sigma_u^2 + \sigma_e^2$
Subject(Treatment)	1	$\frac{4}{3}\sigma_u^2 + \sigma_e^2$
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:

$$F = \frac{MS_{\text{trt}}}{\widehat{V}}$$

Where:

$\widehat{V}$ is an unbiased estimator of the variance component portion of $E[MS_{\text{trt}}]$.

From part c), to get an unbiased estimate of $1.1667\sigma_u^2 + \sigma_e^2$, we have a system of linear equations with real-valued constants a and b such that:

$$\widehat{V} = a \cdot MS_{\text{subj(trt)}} + b \cdot MS_{\text{error}}$$

And taking expectation:

$$E[\widehat{V}] = a(1.3333\sigma_u^2 + \sigma_e^2) + b(\sigma_e^2) = 1.1667\sigma_u^2 + \sigma_e^2$$

For σ_u^2: $1.3333a = 1.1667 \rightarrow a = \frac{1.1667}{1.3333} \approx 0.875$, and

For σ_e^2: $a + b = 1 \rightarrow b = 1 - 0.875 = 0.125$

So we estimate the variance component part using:

$$\widehat{V} = 0.875 \cdot MS_{\text{subj(trt)}} + 0.125 \cdot MS_{\text{error}}$$

So the test statistic becomes:

$$F = \frac{MS_{\text{trt}}}{0.875 \cdot MS_{\text{subj(trt)}} + 0.125 \cdot MS_{\text{error}}}$$

d)

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

The Cochran-Satterthwaite formula for the approximate degrees of freedom associated with the linear combination of mean squares defined by M is

$$d = \frac{M^2}{\sum_{i=1}^k a_i^2 M_i^2 / d_i} = \frac{\left(\sum_{i=1}^k a_i M_i\right)^2}{\sum_{i=1}^k a_i^2 M_i^2 / d_i}.$$

Figure 20: Slide 4

Answer

From part c), we have:

$$F = \frac{MS_{\text{trt}}}{\frac{7}{8} MS_{\text{subj(trt)}} + \frac{1}{8} MS_{\text{error}}}$$

Via the Cochran–Satterthwaite approximation, we use the approximate F-Test to determine the approximated degrees of freedom:

$$d = \frac{\left(\frac{7}{8} MS_{\text{subj(trt)}} + \frac{1}{8} MS_{\text{error}}\right)^2}{\left(\frac{7}{8} MS_{\text{subj(trt)}}\right)^2 + \left(\frac{1}{8} MS_{\text{error}}\right)^2}$$

Simplify this further?

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

Figure 21: Slide 17

Some Actual Data

From SAS Output:

$MS_1 = MS_{\text{subj(trt)}} = 9.626667$, $df_1 = 1$

And:

$MS_2 = MS_{\text{error}} = 2.42$, $df_2 = 1$

Numerator:

$$(0.875 \cdot 9.626667 + 0.125 \cdot 2.42)^2 = (8.422333 + 0.3025)^2 = (8.724833)^2 \approx 76.1208$$

Denominator:

$$(0.875)^2 \cdot \frac{9.626667^2}{1} + (0.125)^2 \cdot \frac{2.42^2}{1} = 0.765625 \cdot 92.6827 + 0.015625 \cdot 5.8564 \approx 70.9728 + 0.0915 = 71.0643$$

Calculating:

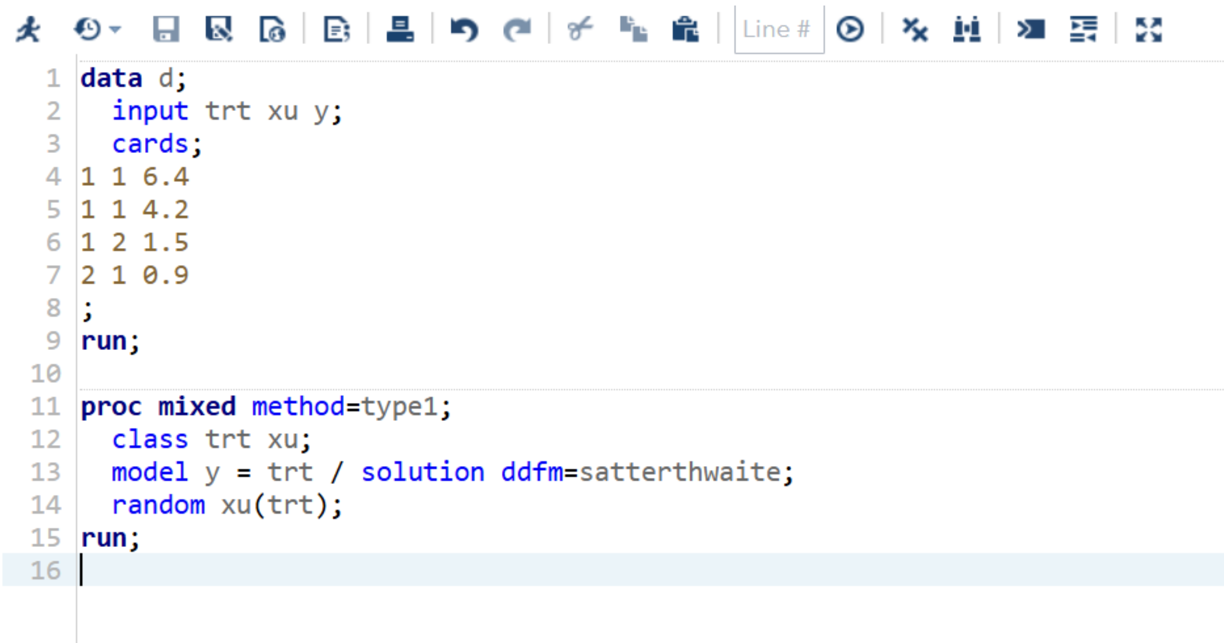
$$d = \frac{76.1208}{71.0643} \approx 1.0717$$

e)

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

Answer

Do I need to point at things?



```
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 22: 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 23: 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 24: SAS Output 2

```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
model <- lmer(y ~ trt + (1 | subject), data = d2, REML = FALSE)

anova_table <- anova(model, type = 1)
anova_table
# model2 <- lmer(y ~ trt + (1 | subject), data = d2)
# anova(model2, type = 1, ddf = "Satterthwaite")
# anova(model2)
# VarCorr(model2)
```

```r
y_mean <- mean(d2$y)
y_trt1 <- mean(d2$y[d2$trt == 1])
y_trt2 <- mean(d2$y[d2$trt == 2])
SS_trt <- 3 * (y_trt1 - y_mean)^2 + 1 * (y_trt2 - y_mean)^2
SS_trt
```

```
## [1] 7.363333
```

```r
model2 <- lmer(y ~ trt + (1 | subject), data = d2)
```

```r
anova_out <- anova(model2, type = 1)
ms_trt <- anova_out$`Mean Sq`[1]
ss_trt <- anova_out$`Sum Sq`[1]

vc <- as.data.frame(VarCorr(model2))
ms_error <- vc$vcov[vc$grp == "Residual"]
ss_error <- ms_error

ms_subject <- 1.3333 * vc$vcov[vc$grp == "subject"] + ms_error
ss_subject <- ms_subject

ss_trt
```

```
## [1] 1.499852
```

```r
ss_subject
```

```
## [1] 9.626487
```

```r
ss_error
```

```
## [1] 2.42
```

For Type 1 SS Specifically:

Source	df	Sums of Squares
Treatment	1	7.363333
Subject(Treatment)	1	9.626487
Error	1	2.42

Notably, the Cochran-Satterthwaite Df Approximation from R is 1.1019. This is somewhat different, but approximately equal, to SAS because of differences in calculation. From what I gathered, it is REML-related.

Bonus - Example 1

```r
library(nlme)
library(lmerTest)

d1 <- data.frame(
  trt = factor(c(1, 1, 2, 2)),
  xu = factor(c(1, 2, 1, 1)),
  y = c(6.4, 4.2, 1.5, 0.9)
)

model1 <- lmer(y ~ trt + (1 | subject), data = d1)
summary(model1)
anova(model1)
mod <- lmer(y ~ trt + (1 | trt:xu), data = d1)
summary(mod)
```