ST553 HW8

Nick Sun

May 29, 2019

Problem 1

We have 2 machines, 4 operators (which is our random effect), and 2 boxes per machine-operator combination. In notation, this is $a = 2, b = 4, n = 2$. If we define γ to be the vector of random effect parameters, we have

$$\gamma = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ (\alpha\beta)_{11} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{13} \\ (\alpha\beta)_{14} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{22} \\ (\alpha\beta)_{23} \\ (\alpha\beta)_{24} \end{pmatrix}$$

and our design matrix Z would be

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

where there would be 2 copies of each unique row (one for each of the two replicates for the operator-machine combos).

Interesting note; `knitr` *only supports matrices that are 10 columns or less by default - you can change this in the preamble by including* `\setcounter{MaxMatrixCols}{20}`

Problem 2

a. From Problem 12.2, define the model with assumptions and constraints

We can define the model as follows:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \epsilon_{ijk}$$

where y_{ijk} is the *lth* dental filling from the *ith* condensation method, *jth* gold alloy, *kth* dentist, α_i is the fixed effect of condensation method i, β_j is the fixed effect of gold alloy j, γ_k is the random effect of dentist k, $(\alpha\beta)_{ij}$ is the fixed interaction effect of the gold alloy j and the condensation method i, $(\alpha\gamma)_{ik}$ is the random

interaction effect of the condensation method and the dentist k, $(\beta\gamma)_{jk}$ is the random interaction effect of the gold alloy and dentist k, and ϵ_{ijk} is the random error.

The assumptions are that $\gamma_k \sim N(0, \sigma^2_\gamma)$, $(\alpha\gamma)_{ik} \sim N(0, \sigma^2_{\alpha\gamma})$, $(\beta\gamma)_{ik} \sim N(0, \sigma^2_{\beta\gamma})$, $\epsilon_{ijk} \sim N(0, \sigma^2)$ and all of these terms are independent from one another.

Lastly, the constraints on the fixed effects are that

$$\sum_{i=1}^{3} \alpha_i = \sum_{j=1}^{8} \beta_j = \sum_{i=1}^{3} (\alpha\beta)_{ij} = \sum_{j=1}^{8} (\alpha\beta)_{ij} = 0$$

b. Calculate the covariance in terms of the model parameters for observations from the same dentist

For observations that had the same dentist and gold alloy, but different condensation method

$$cov(y_{ijk}, y_{i'jk}) = cov(\gamma_k, \gamma_k) + cov((\beta\gamma)_{ik}, (\beta\gamma)_{ik}) = \sigma^2_\gamma + \sigma^2_{\beta\gamma}$$

For obserations that have the same dentist and condensation method, but different gold alloy

$$cov(y_{ijk}, y_{ij'k}) = cov(\gamma_k, \gamma_k) + cov((\alpha\beta)_{ik}, (\alpha\beta)_{ik}) = \sigma^2_\gamma + \sigma^2_{\alpha\gamma}$$

For obserations that had the same dentist but different gold alloys and condensation methods,

$$cov(y_{ijk}, y_{i'j'k}) = cov(\gamma_k, \gamma_k) = \sigma^2_\gamma$$

c. Derive the variance for the difference in sample means between condensation methods 1 and 2

$$.4\sigma^2_{\alpha\gamma} + .05\sigma^2$$

Work shown in appendix.

Problem 3

For Problem 12.6, write down the model

Here our experimental units are the gallons of milk. Each of our 8 gallons of milk are subjected to one of four abuse treatments. After the treatments are applied, each gallon is split into 5 groups and sent to one of five randomly selected laboratories.

The treatments are random effects since a repeat of the experiment would probably select different abuse durations. The laboratories are a random effect since a repeat of this experiment would probably select different laboratories. The samples which are sent to the labs are nested inside of our experimental units. Since each lab gets a different set of 5 samples, the samples are nested inside of the laboratories as well.

$$y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{j(i)k} + \epsilon_{ijkl}$$

where y_{ijkl} is the counts of bacteria in the *lth* replicate from the *kth* lab, *jth* sample nested in the *ith* abuse treatment, μ is the overall mean, α_i sit eh random effect of the *ith* abuse treatment, $\beta_{j(i)}$ is the *jth* sample nested in the *ith* abuse treatment, γ_k is the random effect of the *kth* laboratory, $(\alpha\gamma)_{ik}$ si the random

interaction of the *kth* laboratory and the *ith* treatment, $(\beta\gamma)_{j(i)k}$ is the random interaction between the *jth* sample and the *kth* laboratory, and ϵ_{ijkl} is the random error.

The assumptions are that

$$\begin{aligned}
\alpha_i &\sim N(0, \sigma^2_\alpha) \\
\beta_{j(i)} &\sim N(0, \sigma^2_(\beta(\alpha))) \\
(\alpha\gamma)_{ik} &\sim N(0, \sigma^2_{\alpha\gamma}) \\
(\beta\gamma)_{j(i)k} &\sim N(0, \sigma^2_{\beta(\alpha)\gamma}) \\
\epsilon_{ijkl} &\sim N(0, \sigma^2)
\end{aligned}$$

There are no coefficient constraints since there are no fixed effects in this model.

The output of `PROC MIXED` is displayed below:

Type 3 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
abuse	3	11518698815	3839566272	Var(Residual) + 2 Var(sample*lab(abuse)) + 4 Var(abuse*lab) + 10 Var(sample(abuse)) + 20 Var(abuse)	MS(sample(abuse)) + MS(abuse*lab) - MS(sample*lab(abuse))	11.718	54.86	<.0001
sample(abuse)	4	761700	190425	Var(Residual) + 2 Var(sample*lab(abuse)) + 10 Var(sample(abuse))	MS(sample*lab(abuse))	16	0.19	0.9420
lab	4	608473263	152118316	Var(Residual) + 2 Var(sample*lab(abuse)) + 4 Var(abuse*lab) + 16 Var(lab)	MS(abuse*lab)	12	2.15	0.1372
abuse*lab	12	849857647	70821471	Var(Residual) + 2 Var(sample*lab(abuse)) + 4 Var(abuse*lab)	MS(sample*lab(abuse))	16	69.34	<.0001
sample*lab(abuse)	16	16342850	1021428	Var(Residual) + 2 Var(sample*lab(abuse))	MS(Residual)	40	0.90	0.5746
Residual	40	45409000	1135225	Var(Residual)	.	.	.	.

We see that we have significant evidence to say that the abuse treatments does have an influence on the count of bacteria in milk (pvalue $< .0001$). However, we do not have significant evidence to say that the lab has an effect on the count of bacteria in milk (pvalue $= .1372$).