Stat 541 Homework #8

Kenny Flagg

April 25, 2016

1. *Problem 14.4 (pages 637) but suppose, however, that five operators were randomly selected for each of the six jobs. Be sure to state the model and hypotheses to be tested, check the model assumptions, and provide estimates of all variance components. You can assume that Operator is also a random effect.*

   The appropriate model is

$$
\begin{aligned}
y_{ijk} &= \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)};\\
\tau_i &\overset{iid}{\sim} \mathrm{N}(0, \sigma_\tau^2),\\
\beta_j &\overset{iid}{\sim} \mathrm{N}(0, \sigma_\beta^2),\\
\epsilon_k &\overset{iid}{\sim} \mathrm{N}(0, \sigma^2)
\end{aligned}
$$

   where τ_i is the job effec and β_j is the operator effect within job.

   The hypotheses of interest are H_0: $\sigma_\tau^2 = 0$ and H_1: $\sigma_\tau^2 > 0$. With a test statistic of $F = 31.47$ and p-value < 0.0001, we have suffiecient evidence to reject H_0 at a 0.05 significance level. There is strong evidence that the mean time varies between jobs, so a common standard should not be used for all jobs.

Problem 14.4 ANOVA

The GLM Procedure
Tests of Hypotheses for Random Model Analysis of Variance

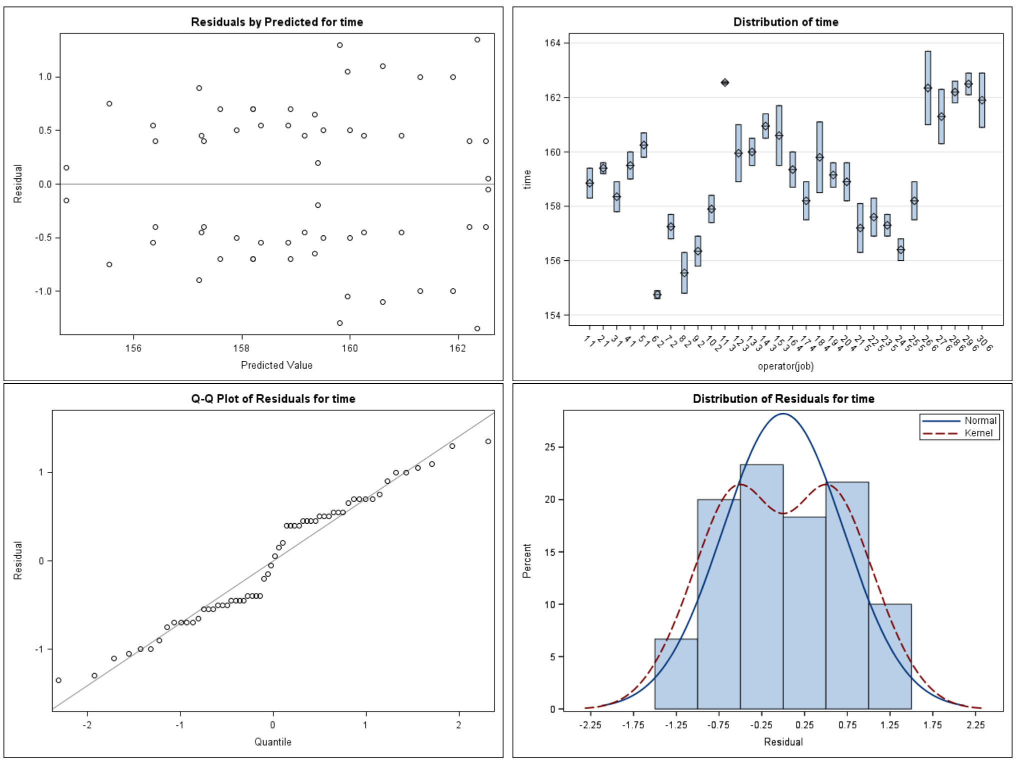
Dependent Variable: time

Source	DF	Type III SS	Mean Square	F Value	Pr > F
job	5	222.450833	44.490167	31.47	<.0001
Error	24	33.934000	1.413917		
Error: MS(operator(job))					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator(job)	24	33.934000	1.413917	1.44	0.1719
Error: MS(Error)	30	29.505000	0.983500		

REML Estimates	
Variance Component	**Estimate**
Var(job)	4.30762
Var(operator(job))	0.21521
Var(Error)	0.98350

The variance estimates are $\sigma_\tau^2 = 4.31$, $\sigma_\beta^2 = 0.215$ and $\sigma^2 = 0.954$. The job is by far the largest component of the variability in completion time.



Some observations with large predicted values also have large residuals, but there is not a strong increasing trend. The boxplots for individual operators show no signs of a serious heterogeneity of variance problem. The Q-Q plot shows some deviation from the diagonal line near the center, but it is not surprising to see this sort of pattern with only two replicates for each operator. The histogram of residuals looks reasonably bell-shaped, so it is reasonable to assume Normality.

2. *Two-stage nested design with 12 machines, 4 spindles nested within each machine, and 5 replicates within each spindle*

   (a) *Fill in the d.f. column assuming all 240 values were collected.*

Source	d.f.
Machine	11
Spindle(Machine)	36
Error	192
Total	239

   (b) *Suppose the three* **X** *values were missing for Spindle 3. Fill in the partial ANOVA table again.*

Source	d.f.
Machine	11
Spindle(Machine)	36
Error	189
Total	236

   (c) *Suppose the five* **Y** *values for Spindle 3 from Machine 12 are missing. Fill in the partial ANOVA table again.*

Source	d.f.
Machine	11
Spindle(Machine)	35
Error	188
Total	234

   (d) *Suppose the the three* **X** *values and the five* Y *values were missing. Fill in the partial ANOVA table again.*

Source	d.f.
Machine	11
Spindle(Machine)	35
Error	185
Total	231

3. *Using the data in Problem 14.3 (page 637), what are the model matrix* $\mathbf{X}$ *and the corresponding vector* y *assuming* $\sum_{i=1}^{3} \alpha_i = 0$ *and* $\sum_{j(i)=1}^{2} \beta_j = 0$ *for* $i = 1, 2, 3$*? Write* $\mathbf{X}$ *in so that there is a column for each model parameter.*

$$\mathbf{X} = \begin{bmatrix}
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix} \qquad y = \begin{bmatrix}
12 \\ 9 \\ 11 \\ 12 \\ 8 \\ 9 \\ 10 \\ 8 \\ 14 \\ 15 \\ 13 \\ 14 \\ 12 \\ 10 \\ 11 \\ 13 \\ 14 \\ 10 \\ 12 \\ 11 \\ 16 \\ 15 \\ 15 \\ 14 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix}$$

4. *Redo the previous problem with the reduced column form for* **X**.

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & -1 \\ 1 & -1 & -1 & 0 & 0 & -1 \\ 1 & -1 & -1 & 0 & 0 & -1 \\ 1 & -1 & -1 & 0 & 0 & -1 \end{bmatrix} \qquad y = \begin{bmatrix} 12 \\ 9 \\ 11 \\ 12 \\ 8 \\ 9 \\ 10 \\ 8 \\ 14 \\ 15 \\ 13 \\ 14 \\ 12 \\ 10 \\ 11 \\ 13 \\ 14 \\ 10 \\ 12 \\ 11 \\ 16 \\ 15 \\ 15 \\ 14 \end{bmatrix}$$

Appendix: SAS Code

Problems 1

```
DM 'LOG; CLEAR; OUT; CLEAR;';
OPTIONS CENTER NONUMBER NODATE;
ODS PDF FILE='hw8prob1.pdf';
ODS GRAPHICS ON / IMAGENAME='hw8p1' RESET=INDEX;

*******************************************************;
*** PROBLEM 14-4, MONTGOMERY but with five operators ***;
*** This will assign unique operator numbers 1 to 30 ***;
*** to the Operators: 1-5 for Job 1, 6-10 for job 2, ***;
*** 11-15 for Job 3, and so on.                      ***;
*******************************************************;

DATA in;
  RETAIN operator 0;
  DO job=1 TO 6;
    DO oper = 1 TO 5;
      operator=operator+1;
      DO rep=1 TO 2;
        INPUT time @@;
        OUTPUT;
      END;
    END;
  END;
DATALINES;
158.3 159.4  159.2 159.6  158.9 157.8  160.0 159.0  159.8 160.7
154.6 154.9  157.7 156.8  154.8 156.3  155.8 156.9  158.4 157.4
162.5 162.6  161.0 158.9  160.5 159.5  160.5 161.4  161.7 159.5
160.0 158.7  157.5 158.9  161.1 158.5  159.6 158.7  158.2 159.6
156.3 158.1  158.3 156.9  157.7 156.9  156.8 156.0  158.9 157.5
163.7 161.0  162.3 160.3  162.6 161.8  162.9 162.1  162.9 160.9
;

PROC GLM DATA=in PLOTS(UNPACK)=ALL;
  CLASS job operator;
  MODEL time = job operator(job) / SS3;
  RANDOM job operator(job) / TEST;
  MEANS operator(job);
TITLE 'Problem 14.4 ANOVA';

RUN;

PROC VARCOMP DATA=in METHOD=REML;
  CLASS job operator;
  MODEL time = job operator(job);
TITLE 'Problem 14.4 Variance Components Estimtes';

RUN;

ODS GRAPHICS OFF;
ODS PDF CLOSE;
QUIT;
```