MORE ON MIXED MODELS

I. In practice, situations arise where interest is in the range of variability of a factor, but it is not possible to choose a random number of levels. In such cases, treating the factor as random usually gives a better model than treating it as a fixed factor, but care must be taken i) to see that the choice of levels results in something that might be a random sample; ii) to apply the result only to a population from which the levels used could reasonably be a random sample; and iii) to use due caution in interpreting the result.

Example: Two protocols for physical therapy for a certain common physical condition are being compared at a large rehabilitation hospital where turnover of physical therapists is high. Interest is in variability of results over the range of physical therapists as well as in comparing the two protocols. The contracts with the physical therapists do not allow the researchers to randomly choose physical therapists to participate in the experiment, so they need to rely on volunteers. However, since the interest is in the variability over the range of physical therapists, treating physical therapist as a fixed effect would not be appropriate. Instead, the researchers use a mixed effects model with physical therapist as random effect. They check that the volunteers have a range of training and experience, so might reasonably result from random selection. In reporting their results, they caution that the results would only apply to physical therapists who might be hired by that hospital, and that since the sample of physical therapists was not random, there might be some qualities of volunteers that would make the results only apply to a smaller population of physical therapists.

II. Often blocking factors are random, so a mixed-effect analysis rather than a fixed-effect analysis may be appropriate for some block designs. Also, the considerations in I may apply.

Example: Contrast the following two examples.

A. Seedlings are being grown in a greenhouse. Since variables such as light, heat, humidity, and pests might vary according to location in the greenhouse, the greenhouse is divided into blocks by location, and treatments are randomly assigned within each block. Should blocks be treated as fixed or random?

B. The candle experiment (Problem 6, p. 326): Color of candle is the treatment factor, burning time of the candle is the response, and the blocks are the experimenters. Should blocks be considered as fixed or random?

More on the Candle experiment: Compare three ways of analyzing the experiment (but recall from previous homework that some caution is needed in interpreting results because of possible lack of model fit.)

1. Treating block as fixed:

Source	DF	SS	MS	F	P
BLOCK	3	151659	50553	29.58	0.000
COLOR	3	60345	20115	11.77	0.000
BLOCK*COLOR	9	15821	1758	1.03	0.431
Error	48	82025	1709		
Total	63	309850			

2. Treating block as random (unrestricted model):

Source	DF	SS	MS	F	P
BLOCK	3	151659	50553	28.76	0.000
COLOR	3	60345	20115	11.44	0.002
BLOCK*COLOR	9	15821	1758	1.03	0.431
Error	48	82025	1709		
Total	63	309850			

Source	Variance component	Error term	Expected Mean Square (using unrestricted model)
1 BLOCK		3	(4) + 4(3) + Q[1]
2 COLOR	1147.32	3	(4) + 4(3) + 16(2)
3 BLOCK*COLOR	12.25	4	(4) + 4(3)
4 Error	1708.85		(4)

3. Treating block as random (restricted model):

Source	DF	SS	MS	F	P
BLOCK	3	151659	50553	29.58	0.000
COLOR	3	60345	20115	11.44	0.002
BLOCK*COLOR	9	15821	1758	1.03	0.431
Error	48	82025	1709		
Total	63	309850			

Source	Variance component	Error term	Expected Mean Square (using restricted model)
1 BLOCK		3	(4) + 4(3) + 16Q[1]
2 COLOR	1150.38	4	(4) + 16(2)
3 BLOCK*COLOR	12.25	4	(4) + 4(3)
4 Error	1708.85		(4)

Questions:

1. How would the CI's for differences in burning time for the colors of candles compare treating blocks as a fixed vs a random effect? Why? Does this make sense intuitively?

2. Does it make sense to use the restricted model with this example? (Is it reasonable to expect that the interaction effects of color with experimenter are negatively correlated for the same experimenter?)

III. **Mixed Models with More than Two Factors**

Example: In the turbine example considered earlier as an example of three random factors, suppose that in fact temperature was a fixed factor. So we now have:

> The reading of the pressure drop across an expansion valve of a turbine is expected to be influenced by gas temperature on the inlet side, operator, and the pressure gauge used by the operator. A three-way design is used to study the effects of these three factors. Three temperatures are *fixed*. Four operators and three gauges are randomly selected. Two observations are taken at each treatment level.

A three-way complete model was used. Letting T, O, and G stand for temperature, operator, and gauge, respectively, the model equation is:

$Y_{ijkt} =$

Minitab output, with Temp fixed, other factors random:

Source	DF	SS	MS	F	P
Temp	2	1023.36	511.68	*	
Operator	3	423.82	141.27	*	
Gauge	2	7.19	3.60	*	
Temp*Operator	6	1211.97	202.00	14.59	0.000
Temp*Gauge	4	137.89	34.47	2.49	0.099
Operator*Gauge	6	209.47	34.91	2.52	0.081
Temp*Operator*Gauge	12	166.11	13.84	0.65	0.788
Error	36	770.50	21.40		
Total	71	3950.32			

Compare and contrast with treating all factors random:

Source	DF	SS	MS	F	P
Temp	2	1023.36	511.68	*	
Operator	3	423.82	141.27	*	
Gauge	2	7.19	3.60	*	
Temp*Operator	6	1211.97	202.00	14.59	0.000
Temp*Gauge	4	137.89	34.47	2.49	0.099
Operator*Gauge	6	209.47	34.91	2.52	0.081
Temp*Operator*Gauge	12	166.11	13.84	0.65	0.788
Error	36	770.50	21.40		
Total	71	3950.32			

So far nothing is different. From the output, it would be reasonable to test H_0^G: $\sigma_G^2 = 0$.

Displaying expected mean squares:

1. With Temp fixed:

```
Source                   Variance Error Expected Mean Square
                        component  term (using unrestricted model)
 1 Temp                            *    (8) + 2(7) + 8(5) + 6(4) + Q[1]
 2 Operator                -4.544  *    (8) + 2(7) + 6(6) + 6(4) + 18(2)
 3 Gauge                   -2.164  *    (8) + 2(7) + 6(6) + 8(5) + 24(3)
 4 Temp*Operator           31.359  7    (8) + 2(7) + 6(4)
 5 Temp*Gauge               2.579  7    (8) + 2(7) + 8(5)
 6 Operator*Gauge           3.512  7    (8) + 2(7) + 6(6)
 7 Temp*Operator*Gauge     -3.780  8    (8) + 2(7)
 8 Error                   21.403       (8)

* No exact F-test can be calculated.
```

2. With Temp random:

```
Source                   Variance Error Expected Mean Square
                        component  term (using unrestricted model)
 1 Temp                    12.044  *    (8) + 2(7) + 8(5) + 6(4) + 24(1)
 2 Operator                -4.544  *    (8) + 2(7) + 6(6) + 6(4) + 18(2)
 3 Gauge                   -2.164  *    (8) + 2(7) + 6(6) + 8(5) + 24(3)
 4 Temp*Operator           31.359  7    (8) + 2(7) + 6(4)
 5 Temp*Gauge               2.579  7    (8) + 2(7) + 8(5)
 6 Operator*Gauge           3.512  7    (8) + 2(7) + 6(6)
 7 Temp*Operator*Gauge     -3.780  8    (8) + 2(7)
 8 Error                   21.403       (8)
```

Note the differences in the row for the fixed factor Temp.

In this example, we see strong evidence of temperature by operator interaction, so it is not reasonable to test for the main effect of temperature. However, for purposes of illustration, if we had obtained no evidence for interaction of temperature with any of the other factors, and wanted to test the main effect of interaction, we would need to figure out and use the appropriate denominator.

From the expected mean squares table (1 above), we have

$$E(MST) = (8) + 2(7) + 8(5) + 6(4) + Q[1]$$
$$E(MSTO) = (8) + 2(7) + 6(4)$$
$$E(MSTG) = (8) + 2(7) + 8(5)$$
$$E(MSTOG) = (8) + 2(7)$$

If it were appropriate to test for main effect of T, then if the null hypothesis of no main effect of temperature were true, we would have Q[1] = 0, and so E(MST) would equal

E(MSTO + MSTG – MSTOG). Thus MSTO + MSTG – MSTOG would be the appropriate denominator. We can get Minitab to do the test as follows:

```
Approximate F-test with denominator: Temp*Operator + Temp*Gauge -
Temp*Operator*Gauge
Denominator MS = 222.63 with   7 degrees of freedom

      Numerator              DF       MS       F       P
      Temp                    2    511.7    2.30   0.171
```

Output from a later version of Minitab: (Temp fixed)

```
      Analysis of Variance for Drop

      Source                  DF        SS          MS        F       P
      Temp                     2    1023.36     511.68     2.30   0.171 x
      Operator                 3     423.82     141.27     0.63   0.616 x
      Gauge                    2       7.19       3.60     0.06   0.938 x
      Temp*Operator            6    1211.97     202.00    14.59   0.000
      Temp*Gauge               4     137.89      34.47     2.49   0.099
      Operator*Gauge           6     209.47      34.91     2.52   0.081
      Temp*Operator*Gauge     12     166.11      13.84     0.65   0.788
      Error                   36     770.50      21.40
      Total                   71    3950.32

Source                 Variance Error Expected Mean Square for Each
Term
                      component term (using unrestricted model)
 1 Temp                           *    (8) + 2(7) + 8(5) + 6(4) + Q[1]
 2 Operator              -4.544   *    (8) + 2(7) + 6(6) + 6(4) + 18(2)
 3 Gauge                 -2.164   *    (8) + 2(7) + 6(6) + 8(5) + 24(3)
 4 Temp*Operator         31.359   7    (8) + 2(7) + 6(4)
 5 Temp*Gauge             2.579   7    (8) + 2(7) + 8(5)
 6 Operator*Gauge         3.512   7    (8) + 2(7) + 6(6)
 7 Temp*Operator*Gauge   -3.780   8    (8) + 2(7)
 8 Error                 21.403        (8)

* Synthesized Test.

Error Terms for Synthesized Tests

Source                   Error DF   Error MS   Synthesis of Error MS
 1 Temp                      6.97     222.63    (4) + (5) - (7)
 2 Operator                  7.09     223.06    (4) + (6) - (7)
 3 Gauge                     5.98      55.54    (5) + (6) - (7)
```