22.

MIXED AND NESTED MODELS

22.1 INTRODUCTION

In the planning stages of an investigation of a psychological system the psychologist decides whether the levels of the factors considered will be set at *fixed* values or chosen at *random* from many levels. We have assumed in previous chapters that the levels were fixed.

A factor is *random* if the pscyhologist wishes to generalize the conclusion of the study to a larger population of levels than used in the experiment. The levels are determined using randomization and in a replication the levels to be used would be determined once again by randomization.

A factor is *fixed* if the psychologist wishes to limit the conclusions to the levels used in the experiment. Generalization can be made, however, these are subjective and not qualified by the statistical analysis. The levels are determined by the experimenter based on knowledge with the system and in a replication of the study the levels are the same as in the original experiment.

When all of the levels are *fixed* the model for the system is called a *fixed model*. When all the levels are chosen at *random* the model is called a *random model*. If some factors are fixed and others random, the model is called a *mixed model*.

22.2 ONE FACTOR MODELS

Consider the one factor model

$$\boxed{E(y \mid x) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k}$$

Whether fixed or random it is assumed that the residuals are normally and independently distributed

$e_i \sim N(0, \sigma_e^2)$

For the fixed model, the β_i are fixed constants, whereas for the random model the β_i are random variables which have a normal distribution with some unknown variance, say σ_β^2.

The analysis is the same in both cases as can be seen in the ANOVA tables.

ANOVA

Source	DF	SS	MS	EMS (Fixed)	EMS (Random)
Treatments	k-1	SSM	MSM	$\sigma^2 + n\theta_\beta$	$\sigma_e^2 + n\sigma_\beta^2$
Error	k(n-1)	SSE	MSE	σ_e^2	σ_e^2
Total	n-1				

The θ_β represents a fixed type of variance so that if there are no differences in the treatment means ($\theta_\beta = 0$) both mean squares estimate error variance (σ_e^2).

The hypothesis tested under the fixed model is

$$H_0\text{: } \beta_1 = \beta_2 = \cdots = \beta_k \text{ or } \theta_\beta = 0$$

whereas the hypothesis tested in the random model is

$$H_0\text{: } \sigma_\beta^2 = 0$$

From the table above it is clear that there is no difference in the mechanics of the F test, only in the conclusions. If H_0 is rejected the psychologist would say there is a difference in the means (β_i) for the fixed model and for the random model there are idfferences in the means (β_i) of which the k are a random sample.

22.3 TWO FACTOR MODELS

For two factors the model is completely described in Chapter 16.

The ANOVA table for two Factors A and B is given below. Three possibilities are considered

1. both A and B fixed (fixed model)
2. both A and B random (random model)
3. A fixed and B random (mixed model)

ANOVA

Source	DF	SS	MS
A	a-1	SSA	MSA
B	b-1	SSB	MSB
A x B	(a-1)(b-1)	SSAB	SMAB
Error	ab(n-1)	SSE	MSE

Source	EMS (Fixed)	EMS (Random)	EMS (Mixed)
A	$\sigma_e^2 + nb\theta_A$	$\sigma_e^2 + n\sigma_{AB}^2 + nb\sigma_A^2$	$\sigma_e^2 + n\sigma_{AB}^2 + nb\theta_A$
B	$\sigma_e^2 + na\theta_B$	$\sigma_e^2 + n\sigma_{AB}^2 + na\sigma_B^2$	$\sigma_e^2 + na\sigma_B^2$
A x B	$\sigma_e^2 + n\theta_{AB}$	$\sigma_e^2 + n\sigma_{AB}^2$	$\sigma_e^2 + n\sigma_{AB}^2$
Error	σ_e^2	σ_e^2	σ_e^2

Examining the EMS column for the fixed model, it is clear that the A, B and AB components are compared to the error. For the random model, the AB interaction is compared to error however the A and B components are compared to the AB interaction. The interaction hypothesis is tested by the error in the mixed model. Furthermore, the random B effect is also compared to error whereas the fixed A effect is compared to the interaction. The importance of the EMS column should be clearly evident to the reader. Although many designs in psychology utilize a fixed model a significant number do not. A set of simple rules will be given to assist in the determination of the EMS column.

22.4 EMS CALCULATION USING RULES

These rules are given as an easy method to the calculation of expected mean squares for a variety of models. With the EMS values, the psychologist in a great many situations can construct the appropriate F test. Consider a 2 Factor design where A is random, B is fixed and there are n observations per treatment.

1. Write the ANOVA model for the study omitting the mean.

   For example, the A x B factorial gives us

A_i	where	i = 1, 2, ..., a
B_j		j = 1, 2, ..., b
AB_{ij}		
$e_{k[ij]}$		k = 1, 2, ..., n

   representing the effects of A, B, A x B and error. The brackets in the e term represent the n observations at each of the i, j treatments.

2. Construct a table with each term in the model equal to a row.

Source	
A_i	
B_j	
AB_{ij}	
$e_{k[ij]}$	

3. Form a column for each subscript in the model.

Source	i	j	k
A_i			
B_j			
AB_{ij}			
$e_{k[ij]}$			

4. Over each subscript write F if the factor is fixed and R if random. Write the number of levels of each subscript.

	a	b	n
	R	F	R
Source	i	j	k
A_i			
B_j			
AB_{ij}			
$e_{k[ij]}$			

5. For each row enter the number of levels under each subscript providing the subscript does not appear in the row heading.

	a	b	n
	R	F	R
Source	i	j	k
A_i		b	n
B_j	a		n
AB_{ij}			n
$e_{k[ij]}$			

6. For any terms in the model with bracketed subscripts, place a 1 under those subscripts inside the brackets.

	a	b	n
	R	F	R
Source	i	j	k
A_i		b	n
B_j	a		n
AB_{ij}			n
$e_{k[ij]}$	1	1	

7. Fill in the remaining cells of a column with a 0 or a 1 depending upon whether the subscript represents a Fixed (F) of Random (R) factor.

	a	b	n
	R	F	R
Source	i	j	k
A_i	1	b	n
B_j	a	0	n
AB_{ij}	1	0	n
$e_{k[ij]}$	1	1	1

This is called the *mean square table* for this design.

22.5 HOW TO FIND EMS FOR ANY TERM

1. Examine the subscript(s).
2. Eliminate any row that does not have the subscript(s) (ie. for A_i omit row B_j)
3. Cover each column of the table headed by a nonbracketed subscript of the term. (ie. for A_i cover column i, $e_{k[ij]}$ cover column k)
4. Multiply the remaining uncovered entries in each row to obtain the coefficients of terms in the expected mean square.

For example, to find the EMS for the A term

Apply rules 2 and 3 to the mean square table.
rule 2: omit row B_j
rule 3: cover column i

	a	b	n	
	R	F	R	
Source	i	j	k	
A_i	1	b	n	$bn\sigma_A^2$
B_j	a	0	n	omit
AB_{ij}	1	0	n	0
$e_{k[ij]}$	1	1	1	σ_e^2

rule 4: $bn\sigma_A^2 + omit + 0 + \sigma_e^2$

the EMS for term A is therefore $bn\sigma_A^2 + \sigma_e^2$

These rules give the following EMS for the design:

	a	b	n	
	R	F	R	
Source	i	j	k	EMS
A_i	1	b	n	$\sigma_e^2 + bn\sigma_A^2$
B_j	a	0	n	$\sigma_e^2 + n\sigma_{AB}^2 + an\theta_B$
AB_{ij}	1	0	n	$\sigma_e^2 + n\sigma_{AB}^2$
$e_{k[ij]}$	1	1	1	σ_e^2

where θ_B is a fixed type of variance. What is the appropriate F statistic to test the effect of B? $F = \dfrac{MSB}{MSAB}$.

22.6 Example 2

Consider a 3 factor design with factors A and B fixed and factor C random. The mean square table and EMS are given below.

	a	b	c	n	
	F	F	R	R	
Source	i	j	k	l	EMS
A_i	0	b	c	n	$\sigma_e^2 + bn\sigma_{AC}^2 + bcn\theta_A$
B_j	a	0	c	n	$\sigma_e^2 + an\sigma_{BC}^2 + acn\theta_B$
AB_{ij}	a	b	1	n	$\sigma_e^2 + n\sigma_{ABC}^2 + cn\theta_{AB}$
C_k	0	0	c	n	$\sigma_e^2 + abn\sigma_c^2$
AC_{jk}	0	b	1	n	$\sigma_e^2 + bn\sigma_{AC}^2$
BC_{jk}	a	0	1	n	$\sigma_e^2 + an\sigma_{BC}^2$
ABC_{ijk}	0	0	1	n	$\sigma_e^2 + n\sigma_{ABC}^2$
$e_{l[ijk]}$	1	1	1	1	σ_e^2

The F statistic to test the effect of B is

$$F = \frac{MSB}{MSBC}$$

Another use of EMS common in psychology is the estimation of the *component of variance* associated with terms in the model.

Exercises

Write out the appropriate F statistic to test the terms in the model above.

22.7 PSUEDO-F TEST

Sometimes the EMS for a psychological study has no exact F-test for one or more terms in the model.

On method of testing hypotheses in these situations is to construct a mean square that is a linear combination of the mean squares in the study. This procedure was developed by Satterthwaite (1946). Gaylor and Hopper (1969) discuss some properties of this procedure. The test in this case is called a *psuedo-F or* F' test.

22.8 NESTED MODELS

In a *nested* experiment levels of one factor are subsamples or are nested within levels of another factor. For example an educational psychologist examined reading ability of Grade 4 students. Three schools were sampled in the district. She randomly examined 2 classes; 5 students were randomly selected from each class.

Schools	1		2	
Classes	1	2	1	2
Students	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

The two factors considered here are school (A) and class (B), however, the factors are not crossed as in the factorial design. There is a *hierarchical* nature to the sampling since the two classes in school 1 are different from the two classes in school 2. In other words classes are nested within schools. We write this as B(A). Note that the errors are nested within B. There is no interaction term in the model since factors are not crossed. If we treat both factors A, B as random, we obtain the following ANOVA for B nested within A.

ANOVA

Source	DF	MS	EMS (random)
A_i	a-1	MSA	$\sigma_e^2 + n\sigma_B^2 + nb\sigma_A^2$
$B_j(A_i)$	a(b-1)	MSB(A)	$\sigma_e^2 + n\sigma_B^2$
Error [$e_{k(ij)}$]	ab(n-1)	MSE	σ_e^2
Total	abn-1		

22.9 Exercises