

6.3: Random Effects in Factorial and Nested Designs

Random effects can appear in both factorial and nested designs. By inspecting the EMS quantities, we can determine the appropriate F -statistic denominator for a given source. Let us look at two-factor studies.

Factorial Design

Recall the *Greenhouse* example in [section 5.1.1](#). In this example, there were two crossed factors (*fert* and *species*). We treated both factors as fixed and the SAS `proc mixed` ANOVA table was as follows:

Type 3 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
fert	3	745.437500	248.479167	Var(Residual) + Q(fert, fert*species)	MS(Residual)	40	73.10	<.0001
species	1	236.740833	236.740833	Var(Residual) + Q(species, fert*species)	MS(Residual)	40	69.65	<.0001
fert*species	3	50.584167	16.861389	Var(Residual) + Q(fert*species)	MS(Residual)	40	4.96	0.0051
Residual	40	135.970000	3.399250	Var(Residual)

If we inspect the EMS quantities in the output, we see that the correct denominator for all F -tests when both factors are fixed in the 2-factor crossed study is Error Mean Squares.

Now let us consider a case in which both factors A and B are random effects in the factorial design (i.e. factors A and B are crossed, and both are random effects). The expected mean squares for each of the source of variations in the ANOVA model would be as follows:

Source	EMS
A	$\sigma^2 + nb\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2$
B	$\sigma^2 + na\sigma_\beta^2 + n\sigma_{\alpha\beta}^2$
A \times B	$\sigma^2 + n\sigma_{\alpha\beta}^2$
Error	σ^2
Total	

The F -tests following from the EMS above would be:

Source	EMS	F
A	$\sigma^2 + nb\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2$	MSA / MSAB
B	$\sigma^2 + na\sigma_\beta^2 + n\sigma_{\alpha\beta}^2$	MSB / MSAB
A \times B	$\sigma^2 + n\sigma_{\alpha\beta}^2$	MSAB / MSE
Error	σ^2	
Total		

Here we can see the ramifications of having random effects. In fixed-effects models, the denominator for the F -statistics in significance testing was the mean square error (MSE). In random-effects models, however, we may have to choose different denominators depending on the term we are testing.

The F -statistic for testing the significance of a given effect, in general, is the ratio of the two MS values with MS of the effect as the numerator, and the denominator MS is chosen such that the F -statistic equals 1 if H_0 is true and greater than 1 if H_a is true.

Following this logic, we can see that when testing for the interaction effect of 2 random factors, the correct denominator is the error mean squares. Therefore the test statistic for testing $A \times B$ is $\frac{MSAB}{MSE}$. However, when we are testing for the main effect of factor A, the correct denominator would be $MSAB$.

Recall that the EMS quantities are the population counterparts for the MS values which actually are sample statistics. Examination of EMS expressions can therefore be used to choose the correct denominator for an F -statistic utilized for testing significance and will be discussed in detail in [Section 6.7](#).

Nested Design

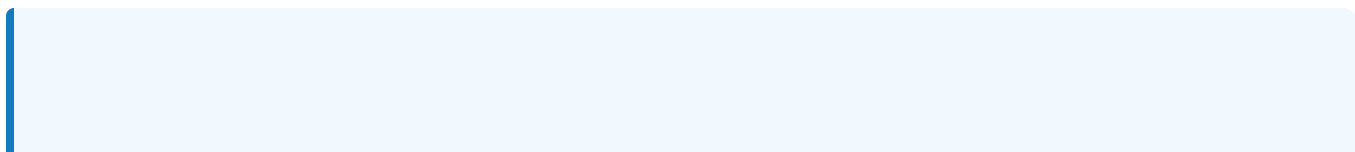
In the case of a nested design, where factor B is nested within the levels of factor A and both are random effects, the expected mean squares for each of the source of variations in the ANOVA model would be as follows:

Source	EMS
A	$\sigma^2 + bn\sigma_\alpha^2 + n\sigma_\beta^2$
B(A)	$\sigma^2 + n\sigma_\beta^2$
Error	σ^2
Total	

The F -tests follow from the EMS above:

Source	EMS	F
A	$\sigma^2 + bn\sigma_\alpha^2 + n\sigma_\beta^2$	MSA / MSB(A)
B(A)	$\sigma^2 + n\sigma_\beta^2$	MSB(A) / MSE
Error	σ^2	
Total		

Using R



? Greenhouse Data - Two Random Effects with Interaction

- Load the greenhouse data.
- Obtain the ANOVA for two random effects with interaction.

Show Detailed Steps

1. Load the greenhouse data by using the following commands:

```
setwd("~/path-to-folder/")
greenhouse_2way_data <- read.table("greenhouse_2way_data.txt", header=T)
attach(greenhouse_2way_data)
```

2. Obtain the ANOVA for two random effects with interaction by using the following commands:

```
library(lmerTest)
library(lme4)
greenhouse_anova<-lmer(height ~ (1 | fertilizer) + (1 | species) + (1 | fertiliz
summary(greenhouse_anova)
```

Linear mixed model fit by REML. t-tests use Satterthwaites method [*'lmerModLmerT*
Formula: height ~ (1 | fertilizer) + (1 | species) + (1 | fertilizer:species)
Data: greenhouse_2way_data

REML criterion at convergence: 216.7

#Scaled residuals:

#	Min	1Q	Median	3Q	Max
#	-2.46787	-0.38510	0.03012	0.38780	2.63056

#Random effects:

#	Groups	Name	Variance	Std.Dev.
#	fertilizer:species	(Intercept)	2.244	1.498
#	fertilizer	(Intercept)	19.301	4.393
#	species	(Intercept)	9.162	3.027
#	Residual		3.399	1.844

Number of obs: 48, groups: fertilizer:species, 8; fertilizer, 4; species, 2

#Fixed effects:

#	Estimate	Std. Error	df	t value	Pr(> t)	
#	(Intercept)	28.387	3.124	2.859	9.088	0.0034 **

*#Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

```
confint(greenhouse_anova)
#
```

	2.5 %	97.5 %
#.sig01	0.4327681	5.482701
#.sig02	0.0000000	10.319191
#.sig03	0.0000000	11.585745
#.sigma	1.5031328	2.335330
#(Intercept)	21.1262902	35.648887

Note that the command `lmer()` gives the ANOVA table only for the fixed effects. Therefore, in this example, since there are no fixed effects, we won't get the ANOVA table. In the "Random effects" section of the output, under the column variance we get the estimates for $\sigma_{\alpha\beta}^2$, σ_{α}^2 , σ_{β}^2 , and σ^2 which are equal to 2.244, 19.301, 9.162, and 3.399 respectively. In the "Fixed effects" section under the column estimate we get the estimate of μ , or the overall mean, which is equal to 28.387.

With the command `confint()` we will get confidence intervals for the standard deviations and the overall mean. If you take the square of the lower and upper bounds, you will get a confidence interval for the model variances.

Alternatively, we can use the command `aov()` which gives a partial ANOVA table.

```
greenhouse_anova1<-aov(height~Error(fertilizer+species+fertilizer:species),green
summary(greenhouse_anova1)
#Error: fertilizer
#           Df Sum Sq Mean Sq F value Pr(>F)
#Residuals  3   745.4    248.5

#Error: species
#           Df Sum Sq Mean Sq F value Pr(>F)
#Residuals  1   236.7    236.7

#Error: fertilizer:species
#           Df Sum Sq Mean Sq F value Pr(>F)
#Residuals  3    50.58    16.86

#Error: Within
#           Df Sum Sq Mean Sq F value Pr(>F)
#Residuals 40     136     3.399
detach(greenhouse_2way_data)
```

Note that both commands in R don't give the F -values and the p -values for the tests. Therefore, these must be done manually.