# Topic 10. ANOVA models for random and mixed effects

**References:** ST&DT: Topic 7.5 p.152-153, Topic 9.9 p. 225-227, Topic 15.5 379-384. There is a good discussion in SAS System for Linear Models, 3rd ed. pages 191-198.

**Note:** do not use the rules for expected MS on ST&D page 381. We provide in the notes below updated rules from Chapter 8 from Montgomery D.C. (1991) "Design and analysis of experiments".

## 10. 1. Introduction

The experiments discussed in previous chapters have dealt primarily with situations in which the experimenter is concerned with making comparisons among ***specific factor levels***. There are other types of experiments, however, in which one wish to determine the ***sources of variation*** in a system rather than make particular comparisons. One may also wish to ***extend the conclusions beyond*** the set of specific factor levels included in the experiment. The purpose of this chapter is to introduce analysis of variance models that are appropriate to these different kinds of experimental objectives.

## 10. 2. Fixed and Random Models in One-way Classification Experiments

### 10. 2. 1. Fixed-effects model

A typical fixed-effect analysis of an experiment involves comparing treatment means to one another in an attempt to detect differences. For example, four different forms of nitrogen fertilizer are compared in a CRD with five replications. The analysis of variance takes the following familiar form:

| Source | df |
|---|---|
| Total | 19 |
| Treatment (i.e. among groups) | 3 |
| Error (i.e. within groups) | 16 |

And the linear model for this experiment is: $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$.

This experiment involves four **specific treatment levels** (e.g. urea, ammonia, ammonium nitrate, and 30% urea:ammonium nitrate solution). The experimenter decided these were the levels of interest, selected them specifically for investigation, and has no thought for any other N-fertilizers in the analysis. In other words, the experimenter's attention is fixed upon these four fertilizer treatments and no other. If the experiment were to be repeated, these same exact four forms of nitrogen fertilizer would be used again.

Treatment (or factor) effects are called ***fixed effects*** when the $\tau_i$'s remain the same in each replication of the experiment. Even if the experiment is not replicated, one can think of a probabilistic model in which there are very many possible outcomes and the actual experiment selects one of them. In this context, the term *fixed effects* means that in each possible outcome, the $\tau_i$'s have the same value. Also, the $\tau_i$'s must add to 0 since the mean

over all the treatments is μ. This ANOVA model is called the **fixed-effects model** or **Model I ANOVA**; and it is the one we have considered up to this point in the class. In such experiments, the $\varepsilon_{ij}$ (i.e. the residuals) are a random sample from a normally distributed population of errors with mean 0 and variance $\sigma^2$.

### 10.2.2 Random-effects model

In a random-effects model, ***the treatment levels themselves are a random sample*** drawn from a larger population of treatments. In this case, the linear model looks very similar:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

But now the treatment effect ($\tau_i$) is **a random variable**. While the *population* of these treatment effects has mean 0 and variance $\sigma_\tau^2$, for any given random sample of treatment levels, $\sum \tau_i \neq 0$.

The ***objective*** of the researcher is to ***extend the conclusions*** based on a sample of treatment levels to ***all treatment levels in the population***. Because the treatment levels are a random sample from a larger population of treatment levels (i.e. the treatment effect $\tau_i$ is a random variable), *conclusions about any one specific level are relatively useless*. Said another way, if the experiment were to be repeated, an entirely new random sample of treatment levels would be investigated; so clearly the objective is not to compare specific treatment levels to one another.

In fact, the null hypothesis of a random-effects ANOVA is quite different from its fixed-effects counterpart:

$H_0$ for a Fixed-effects ANOVA: $\tau_1 = \tau_2 = \ldots = \tau_t = 0$
$H_0$ for a Random-effects ANOVA: $\sigma_\tau^2 = 0$

When the null hypothesis is false, there will be an additional component of variance in the experiment equal to $r\sigma_\tau^2$. The aim of the researcher is to test for the presence of this additional component of variance and to estimate its magnitude.

As the linear models for the one-way ANOVA suggests, the computation is the same for the both the **fixed** and the **random** models. However, the **objectives** and the **conclusions** are quite different. Subsequent analysis following the initial test for significance is also different.

**Example**: In an animal breeding experiment conducted to estimate the breeding value (or combinability) of sires from a certain breed, several sires were randomly selected from a population and each sire was mated with several dams. The weights of all the newborn animals were recorded.

The experimental model is $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, where $\tau_i$ is the effect of the $i^{th}$ sire (i.e. the difference between the $i^{th}$ sire and the overall mean). The other terms are the same as the fixed-effects model.

Let us assume that four sires were chosen for the study and that each was mated with 6 dams. In this case, the four sires are of *no specific interest*; they are merely a random sample from a population of sires of the breed. They are interesting only to the extent that they represent their population. If the experiment were to be repeated, another set of sires very likely would be used.

The $\tau_i$'s are a random sample from a population of $\tau_i$'s with mean 0. In general, however, the $\tau_i$'s will be different in each replication of the experiment and will not sum to zero for any particular experiment. This lies at the crux of a **random-effects model** (sometimes called **Model II**).

If we let $\bar{Y}_i$ represent the mean breeding value of the $i^{th}$ sire and $\mu$ represent the average breeding value of sires from that population, then for any given sample of sires:

$$\sum \tau_i = \sum (\bar{Y}_i - \mu) \neq 0$$

This sum will not equal zero, just as indicated, *unless* the summation covers the entire population (i = 1,…,N).

Sometimes the determination of whether an effect is fixed or random is not obvious. Examples are laboratories or technicians in a comparative study, years in a multiple-year trial, locations in a multiple-location study, etc. These factors can be fixed or random depending on the objective of the study, the intended inferences to be made, and the process by which the levels of the factors were selected.

### 10. 2. 3. Differences between fixed and random-effects model

Although the linear models for the above two types of single-classification experiments (CRD) are similar, there are some fundamental differences:

1) **The objectives are different**. In the fertilizer experiment, each fertilizer is of specific interest. The four fertilizers are not a random sample from some larger population of fertilizers. The purpose of the study is to compare these particular treatment levels to one another. In the breeding study, however, the objective of the study is to estimate the combinability of a breed. The sires used in the study are merely **a sample** from which inferences are to be made concerning the **population**.

   The purpose of a fixed model is to test the hypothesis that the treatment effects are the same; the purpose of a random model is to estimate the component of variance $\sigma^2_\tau$.

2) **The sampling procedures are different**. In the fixed-effects experiment, the treatment levels are selected purposefully (i.e. not randomly) by the investigator. In the random-effects experiment, the treatment levels are selected randomly, and the unknown variance of the population of treatment effects contributes to the total sum of squares.

   If the first experiment is repeated, the four fertilizers will be used again and only the random errors change from experiment to experiment (i.e. the $\tau_i$'s are assumed to be constants; only the $\varepsilon_{ij}$'s change). In the second experiment, the four sires most likely

will differ each time the study is conducted. Thus not only the errors vary but the sire effects ($\tau_i$'s) vary as well.

3) **The expected sums of the effects are different**. Because the effects are constants, in the fixed-effects model:

$$\sum \tau_i = \sum (\bar{Y}_i - \mu) = 0$$

But for any given sample in the random-effects model:

$$\sum \tau_i = \sum (\bar{Y}_i - \mu) \neq 0$$

4) **And, therefore, the expected variances are different**

For Model I:

**Var of $Y_{ij}$** = variance of $\mu$ + variance of $\tau_i$ + variance of $\varepsilon_{ij}$
= variance of $\varepsilon_{ij}$
= $\sigma_\varepsilon^2$

For Model II:

**Var of $Y_{ij}$** = variance of $\mu$ + variance of $\tau_i$ + variance of $\varepsilon_{ij}$
= variance of $\tau_i$ + variance of $\varepsilon_{ij}$
= $r\sigma_\tau^2 + \sigma_\varepsilon^2$ ($\sigma_\tau^2$ and $\sigma_\varepsilon^2$ are called **variance components**)

The expected mean squares for these models:

| Source | df | MS | EMS | |
|---|---|---|---|---|
| | | | Fixed Model | Random Model |
| **Treatment** | t-1 | MST | $r\sum \frac{\tau_i^2}{t-1} + \sigma_\varepsilon^2$ | $r\sigma_\tau^2 + \sigma_\varepsilon^2$ |
| **Error** | t(r-1) | MSE | $\sigma_\varepsilon^2$ | $\sigma_\varepsilon^2$ |

**Example:** Suppose five cuts of meat are taken from each of three pigs, all from the same breed, and the fat content is measured in each cut. If the goal were to compare the fat content among these five cuts of meat, this would be an example of a fixed effects model with five treatment levels (cuts) and three replications (pigs).

If, however, the goal is to assess animal-to-animal and within-animal variation, the particular three pigs selected are not of interest. This would be an example of a *random effects model* with three treatment levels (pigs) and five replications (cuts).

Suppose the goal is indeed to assess components of variation. The analysis reveals a treatment mean square of 80 and an error mean square of 20. The resultant F ratio = 80/20 = 4, with 2 and 12 degrees of freedom. This number is larger than the critical value and hence is declared significant at the 5% level. We therefore conclude that $\sigma_\tau^2 > 0$; but how big is it?

The table above implies that:

$$80 = r\sigma_\tau^2 + \sigma_\varepsilon^2$$

Since r = 5 and $\sigma_\varepsilon^2 = 20$, we can solve:

$$80 = 5\sigma_\tau^2 + 20$$

$$\sigma_\tau^2 = 12$$

This number (12) is an estimate of $\sigma_\tau^2$, the variance component due to animal-to-animal differences. This study shows that fat content is nearly twice as variable among cuts within this breed of pig as it is among pigs of this breed.

## 10. 3. Two-way classification experiments

In the single-factor case, the model must specify whether the factor is characterized by either fixed or random effects. In the multifactor case these two possibilities are joined by a third: the ***mixed effects model*** in which some factors are fixed and some are random. An examination of this model is useful because it provides a good opportunity to contrast fixed and random effects.

Suppose an experimenter conducts a field study including several varieties tested across a number of locations. In each location, a completely randomized design is used, with each variety replicated over ***r*** plots. Let $Y_{ijk}$ represent the plot yield of the $k^{th}$ plot of the $i^{th}$ variety at the $j^{th}$ location. Then the linear model is:

$$Y_{ijk} = \mu + \alpha_i + ß_j + (\alpha ß)_{ij} + \varepsilon_{ijk}$$

where $\mu$ is the overall mean yield, $\alpha_i$ is the effect on yield of the $i^{th}$ variety, $ß_j$ is the effect on yield of the $j^{th}$ location, $(\alpha ß)_{ij}$ is the effect on yield of the interaction between variety and location, and $\varepsilon_{ijk}$ is the usual random error term with mean 0 and variance $\sigma^2_\varepsilon$. The different possible models for this study are described below.

### 10. 3. 1. Fixed-effects model

If the experimenter is only interested in these particular varieties grown at these particular locations, and conclusions are not to be generalized to other varieties or locations, then the model is a ***fixed-effects model***. In this case:

1) μ is the overall mean yield in the study.
2) $\alpha_i = \overline{Y}_{i.} - \mu$ is the true effect of the $i^{th}$ variety and $\Sigma\alpha_i = 0$.
3) $\beta_j = \overline{Y}_{.j} - \mu$ is the true effect of the $j^{th}$ location and $\Sigma ß_j = 0$.
4) $(\alpha ß)_{ij}$ = is the specific interaction effect due to the $i^{th}$ variety and the $j^{th}$ location

The variance of $Y_{ijk}$ is $\sigma^2_\varepsilon$. In this model, the experimenter is interested in estimating and testing hypotheses about $\alpha_i$, $ß_j$, and $(\alpha ß)_{ij}$.

### 10. 3. 2. Random-effects model

Now suppose the varieties are randomly chosen from a population of varieties and the investigator is interested in characterizing the yield variation within a population of varieties (e.g. pre-Green Revolution wheat varieties). Similarly, the locations are randomly selected from numerous possible testing sites representing some larger region; therefore, the *specific* locations in the trial are of no particular interest. In this situation:

1. μ is the overall mean yield of all varieties at all possible locations (only a *sample* of those varieties and locations were included in the study).
2. $\alpha_i = \overline{Y}_{i.} - \mu$ is a random effect from a population with mean 0 and variance $\sigma^2_\alpha$.
3. $\beta_j = \overline{Y}_{.j} - \mu$ is a random effect from a population with mean 0 and variance $\sigma^2_\beta$.
4. $(\alpha ß)_{ij}$ = is the random interaction effect from a population of all possible interaction effects between varieties and locations, with mean 0 and variance $\sigma^2_{\alpha\beta}$.

The variance of $Y_{ijk}$ in this model is $\sigma^2_\alpha + \sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2_\varepsilon$. In this model, the experimenter is interested in estimating and testing hypotheses about $\sigma^2_\alpha$, $\sigma^2_\beta$, and $\sigma^2_{\alpha\beta}$.

### 10. 3. 3. Mixed-effects model

Now suppose the *varieties* are specifically chosen for comparisons but the *locations* are randomly selected from many possible locations in order to examine the consistency of the varietal yields under various environmental conditions. In this case:

1. μ is the overall mean yield of these particular varieties at all possible locations.
2. $\alpha_i = \overline{Y}_{i.} - \mu$ is the true effect of the $i^{th}$ variety; and $\sum \alpha_i = 0$.
3. $\beta_j = \overline{Y}_{.j} - \mu$ is the random effect of the $j^{th}$ location drawn from a population with mean 0 and variance $\sigma^2_\beta$.

4. $(\alpha ß)_{ij}$ = is the interaction effect of the ith variety in the jth location. Since locations are random, this interaction is usually considered to be a random effect with mean 0 and variance $\sigma^2_{\alpha\beta}$.

Here the variance of $Y_{ijk}$ is $\sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2_\varepsilon$. In this model, the experimenter is interested in estimating and testing hypotheses about $\alpha_i$, $\sigma^2_\beta$, and $\sigma^2_{\alpha\beta}$.

### 10. 3. 4. Expected mean squares and *F* tests

In all of the examples of fixed models discussed in previous topics, the error term has always been on the last line, whatever unexplained variation is left over after partitioning the total sum of squares. The proper tool for determining the appropriate error variances for more complex situations is the set of **expected mean squares**. Expected mean squares are algebraic expressions which specify the underlying model parameters that are estimated by the calculated mean squares (i.e. the mean squares which result from partitioning the sum of squares for a particular dataset). Generally, these expected mean squares are linear functions of elements that represent:

1. The error variances
2. Functions of variances of random effects
3. Functions of sums of squares and products (quadratic forms) of fixed effects

Below is a table of the expected mean squares of three two-way classification experiments, featuring **a** varieties, **b** locations, and **r** replications:

| Source | df | MS | Fixed Model | Random Model | Mixed Model<br>Fixed Var, Random Loc |
|---|---|---|---|---|---|
| Var | a-1 | MSV | $\sigma^2_\varepsilon + br\sum\frac{\alpha^2}{a-1}$ | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta} + br\sigma^2_\alpha$ | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta} + br\sum\frac{\alpha^2}{a-1}$ |
| Loc | b-1 | MSL | $\sigma^2_\varepsilon + ar\sum\frac{\beta^2}{b-1}$ | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta} + ar\sigma^2_\beta$ | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta} + ar\sigma^2_\beta$ |
| V*L | (b-1)(a-1) | MSVL | $\sigma^2_\varepsilon + r\sum\frac{(\alpha\beta)^2}{(a-1)(b-1)}$ | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta}$ | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta}$ |
| Error | ba(r-1) | MSE | $\sigma^2_\varepsilon$ | $\sigma^2_\varepsilon$ | $\sigma^2_\varepsilon$ |

There is controversy among statisticians regarding the inclusion of interactions between fixed and random factors in the EMS of the random factors. Steel, Torrie, and Dickey, for example, exclude the $r\sigma^2_{\alpha\beta}$ term from the EMS of the random variable. However, Hocking (*Journal of the American Statistical Association* 1975 **70**: 706-712) pointed out that exclusion of $r\sigma^2_{\alpha\beta}$ is inconsistent with results commonly reported for the unbalanced case. Hocking's approach is recommended by SAS; and in this syllabus we will include the interactions between fixed and random factors in the EMS of random factors in a

mixed model. Note also that the EMS for interactions in the mixed model in ST&D (p 226, 380) include an additional correction factor of the form x/(x-1). This factor is not used in most statistic books and is not recommended by SAS. ST&D also point this out (p 382).

Based on the expected mean squares, correct denominators can be identified to perform the appropriate F tests on locations, varieties, and their interaction. The appropriate F tests for the three previous models are shown in the table below:

| Source | Fixed | Random | Mixed |
|---|---|---|---|
| Variety | MSV/**MSE** | MSV/MSVL | MSV/MSVL |
| Location | MSL/**MSE** | MSL/MSVL | MSL/MSVL |
| V*L | MSVL/**MSE** | MSVL/**MSE** | MSVL/**MSE** |

**Note that the appropriate F tests change, depending on the model.**

The underlying principle of an F test on a set of fixed-effect parameters is that the expected mean square for the *denominator* contains a linear function of variances of random effects, whereas the expected mean square for the *numerator* contains the same function of these variances **plus** a quadratic form of the parameter being tested.

In **fixed models**, each mean square other than the MSE estimates the residual error variance plus a quadratic form of the parameter in question. Hence the proper denominator for all tests is the MSE. That is why determination of the expected mean squares is usually not required for fixed models.

In **random and mixed models**, however, it is essential to determine the expected mean squares in order to select the correct denominators for all F tests. In the **random model**, one common objective is to extend one's conclusion to the complete population of treatment levels (e.g. locations, as in the examples above). Therefore, it makes sense to divide by the MS of the *interaction*. Why? Because, for the conclusions about varieties to be valid across all the locations from the region we want to extrapolate our conclusions, the variety effects must be larger than any genotype x environment interaction.

In Topic 9.7.2, it was pointed out that a significant interaction in a **fixed-effects model** leads us to forfeit tests of hypotheses for main effects in favor of tests of hypotheses for simple effects. However, for a **mixed-effects model**, we may not be at all interested in simple effects for a fixed effect, since they are measured at *randomly selected levels of another factor*. Instead, we remain interested in the main effects, even when there is an interaction between the random and the fixed factors.

### 10. 3. 4. 1. Rules for determining Expected Mean Squares (EMS)

To perform an analysis of variance, one must first determine the sum of squares for each component in the model as well as the number of degrees of freedom associated with

each sum of squares. Then, to construct appropriate test statistics (see box below), the expected mean squares must be determined. In complex design situations, particularly those involving random or mixed models, it is frequently helpful to have a formal procedure for this process.

**The appropriate test statistic (F) is a ratio of mean squares that is chosen such that the expected value of the *numerator* differs from the expected value of the *denominator* only by the specific variance component or fixed factor being tested.**

The following set of rules is appropriate for the manual calculation of the expected mean squares for any balanced factorial, nested, or nested factorial experiment. (Note that partially balanced arrangements, such as Latin squares and incomplete block designs, are specifically excluded.) To illustrate the application of these rules, the two-factor factorial model is used.

RULE 1. The error term in the model, $\varepsilon_{ij...m}$, is written as $\varepsilon_{(ij...)\,m}$, where the subscript m denotes the replication subscript. For the two-factor model, this rule implies that $\varepsilon_{ijk}$ becomes $\varepsilon_{(ij)\,k}$.

RULE 2. In addition to the error term, the model contains all the main effects and any interactions that the experimenter assumes exist. In other words, the model is the linear model; it contains all the effects that would appear in the Model statement of the Proc GLM.

RULE 3. For each term in the model, divide the subscripts into three classes:

a) **Live**: those subscripts that are present in the term and are not in parentheses.
b) **Dead**: those subscripts that are present in the term and are in parentheses (usually subscripts of nested factors).
c) **Absent**: those subscripts that are present in the model but not in that particular term.

For example, in $(\alpha\beta)_{ij}$, i and j are live and k is absent; and in $\varepsilon_{(ij)\,k}$ k is live and i and j are dead.

RULE 4. The number of degrees of freedom for any term in the model is the product of the number of levels associated with each dead subscript and the number of levels minus 1 associated with each live subscript.

For example: df $(\alpha\beta)_{ij}$= (a-1)(b-1); df $\varepsilon_{(ij)\,k}$=ab(r-1)

RULE 5. Each effect has either a variance component (random effect) or a fixed factor (fixed effect) associated with it. If an interaction contains at least one random effect, the entire interaction is considered to be a random effect. A variance component has Greek letters as subscripts to identify the particular random effect. Thus, in a two-factor mixed model with factor A fixed and factor B random, the variance component for B is $\sigma^2_{\beta}$ and the variance component for AB is $\sigma^2_{\alpha\beta}$. A fixed effect is always represented by the sum of squares of the model components associated with that factor, divided by its degrees of freedom. In our example, the fixed effect of A is:

$$\frac{\sum_{i=1}^{a}\alpha_i^2}{a-1}$$

RULE 6. To obtain the **expected mean squares**, prepare a table with a row for each model component (mean square) and a column for each subscript. Over each subscript, write the number of levels of the factor associated with that subscript and whether the factor is fixed (F) or random (R). Replicates and subsamples are always considered to be random.

(a) In each row, write 1 if one of the dead subscript in the row component matches the subscript in the column:

(b)

| Fixed or Random<br>Number of levels<br>Factor | F<br>a<br>i | R<br>b<br>j | R<br>r<br>k |
|---|---|---|---|
| $\alpha_i$ | | | |
| $\beta_j$ | | | |
| $(\alpha\beta)_{ij}$ | | | |
| $\varepsilon_{(ij)k}$ | 1 | 1 | |

(c) In each row, if any of the subscripts of the row component match the subscript in the column, write 0 if the column is headed by a fixed factor and 1 if the column is headed by a random factor. For interactions involving at least one random factor, write 1 in the columns that match the row subscript, independent of the nature (fixed or random) of that column:

| Fixed or Random<br>Number of levels<br>Factor | F<br>a<br>i | R<br>b<br>j | R<br>r<br>k |
|---|---|---|---|
| $\alpha_i$ | 0 | | |
| $\beta_j$ | | 1 | |
| $(\alpha\beta)_{ij}$ | **1** | 1 | |
| $\varepsilon_{(ij)k}$ | 1 | 1 | 1 |

Note the inclusion of a **1** in the a column for the interaction term $(\alpha\beta)_{ij}$ . This will determine the presence of $r\sigma^2_{\alpha\beta}$ in the EMS of the random Factor $\beta$.

(d) In the remaining empty row positions, write the number of levels shown above the column heading:

| Fixed or Random<br>Number of levels<br>Factor | F<br>a<br>i | R<br>b<br>j | R<br>r<br>k |
|---|---|---|---|

| | | | |
|---|---|---|---|
| $\alpha_i$ | 0 | b | r |
| $\beta_j$ | a | 1 | r |
| $(\alpha\beta)_{ij}$ | **1** | 1 | r |
| $\varepsilon_{(ij)\,k}$ | 1 | 1 | 1 |

(e) To obtain the expected mean square for any model component, first cover all columns headed by live subscripts associated with that component. Then, in each row that contains at least the same subscripts as those of the component being considered, take the product of the visible numbers and multiply by the appropriate fixed or random factor from Rule 5. The sum of these quantities is the EMS of the model component being considered.

To find $EMS_A$, for example, cover column I (i is the index associated with factor A). The product of the visible numbers in the rows that contain at least subscript i are br (row 1), r (row 3), and 1 (row 4). Note that i is missing in row 2.

The expected mean squares derived using these rules for the two-way mixed model are:

| Fixed or Random<br>Number of levels<br>Factor | F<br>a<br>i | R<br>b<br>j | R<br>r<br>k | Expected Mean Squares |
|---|---|---|---|---|
| $\alpha_i$ | 0 | b | r | $\sigma^2_\varepsilon + \mathbf{r\sigma^2_{\alpha\beta}} + br\Sigma\alpha^2/(a-1)$ |
| $\beta_j$ | a | 1 | r | $\sigma^2_\varepsilon + \mathbf{r\sigma^2_{\alpha\beta}} + ar\sigma^2_\beta$ |
| $(\alpha\beta)_{ij}$ | **1** | 1 | r | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta}$ |
| $\varepsilon_{(ij)\,k}$ | 1 | 1 | 1 | $\sigma^2_\varepsilon$ |

## 10. 4. Expected Mean squares for three-way ANOVA

Consider a three-factor factorial experiment with a levels of factor A, b levels of factor B, c levels of factor C, and r replicates. The analysis of this design assuming that A is fixed and B and C are random is given below. The appropriate statistical model is:

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

| Fixed or Random<br>Number of levels<br>Factor | F<br>a<br>i | R<br>b<br>j | R<br>c<br>k | R<br>r<br>l | **Expected Mean Squares** | **F** |
|---|---|---|---|---|---|---|
| $\alpha_i$ | 0 | b | c | r | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta\gamma} + br\sigma^2_{\alpha\gamma} + cr\sigma^2_{\alpha\beta} + bcr\Sigma\alpha^2/(a-1)$ | **?** |
| $\beta_j$ | a | 1 | c | r | $\sigma^2_\varepsilon + \mathbf{r\sigma^2_{\alpha\beta\gamma}} + ar\sigma^2_{\beta\gamma} + \mathbf{cr\sigma^2_{\alpha\beta}} + acr\sigma^2_\beta$ | **?** |
| $\gamma_k$ | a | b | 1 | r | $\sigma^2_\varepsilon + \mathbf{r\sigma^2_{\alpha\beta\gamma}} + ar\sigma^2_{\beta\gamma} + \mathbf{br\sigma^2_{\alpha\gamma}} + abr\sigma^2_\gamma$ | **?** |
| $(\alpha\beta)_{ij}$ | **1** | 1 | c | r | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta\gamma} + cr\sigma^2_{\alpha\beta}$ | $MS_{\alpha\beta}/MS_{\alpha\beta\gamma}$ |
| $(\alpha\gamma)_{ik}$ | **1** | b | 1 | r | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta\gamma} + br\sigma^2_{\alpha\gamma}$ | $MS_{\alpha\gamma}/MS_{\alpha\beta\gamma}$ |
| $(\beta\gamma)_{jk}$ | a | 1 | 1 | r | $\sigma^2_\varepsilon + \mathbf{r\sigma^2_{\alpha\beta\gamma}} + ar\sigma^2_{\beta\gamma}$ | $MS_{\beta\gamma}/MS_{\alpha\beta\gamma}$ |
| $(\alpha\beta\gamma)_{ijk}$ | **1** | 1 | 1 | r | $\sigma^2_\varepsilon + r\sigma^2_{\alpha\beta\gamma}$ | $MS_{\alpha\beta\gamma}/MS_\varepsilon$ |
| $\varepsilon_{(ijk)l}$ | 1 | 1 | 1 | 1 | $\sigma^2_\varepsilon$ | |

Identical results to those presented in the previous table can be obtained in SAS using the **RANDOM** statement within **PROC GLM**. When the **TEST** option is used within the RANDOM statement, SAS will automatically determine the appropriate forms for all tests and provide probabilities for them.

For this procedure to work in SAS, **all random effects and all interactions that involve at least one random effect must be explicitly designated as "random" in the RANDOM statement**.

**SAS Program (A=fix, B and C= Random)**

```
model Y= A B C A*B  A*C  B*C  A*B*C;
random B  C  A*B  A*C  B*C  A*B*C;.
```

**Caution**: Please be aware, however, that **contrasts** are not corrected by the **RANDOM** statement, and there is no way within **Proc GLM** for the user to specify a synthetic denominator for a contrast (SAS provides an alternative procedure for this purpose, called **Proc Mixed**, which is not covered in this class).

**Example SAS Program** (A is a fixed factor with 3 levels; B is a random factor with 5 levels, and C is a random factor with 2 levels, and there are 5 replications (r=5):

```
Proc GLM;
   Class A B C;
   Model Y = A B C A*B A*C B*C A*B*C;
   Random B C A*B A*C B*C A*B*C / test;
```

or to simplify

```
Proc GLM;
   Class A B C;
   Model Y = A|B|C;
   Random B C A*B A*C B*C A*B*C / test;
```

To two model statements above are the same. The **RANDOM** statement generates the following table of estimated mean squares, identical to the one produced manually using the method outlined in the previous section:

```
Class          Levels    Values
A                   3    1 2 3
B                   5    1 2 3 4 5
C                   2    1 2
With 5 replications= 3*5*2*5= 150 observations
```

| Source | Type III Expected Mean Square |
|---|---|
| A | Var(Error) + 5 Var(A*B*C) + 25 Var(A*C) + 10 Var(A*B) + Q(A)[1] |
| B | Var(Error) + 5 Var(A*B*C) + 15 Var(B*C) + 10 Var(A*B) + 30 Var(B) |
| C | Var(Error) + 5 Var(A*B*C) + 15 Var(B*C) + 25 Var(A*C) + 75 Var(C) |
| A*B | Var(Error) + 5 Var(A*B*C) + 10 Var(A*B) |
| A*C | Var(Error) + 5 Var(A*B*C) + 25 Var(A*C) |
| B*C | Var(Error) + 5 Var(A*B*C) + 15 Var(B*C) |
| A*B*C | Var(Error) + 5 Var(A*B*C) |

[1]SAS uses **Q** to summarize the fixed effects $Q(A)= bcr\Sigma\alpha^2/(a-1)$

Proper F tests for the two-factor interactions and the three-factor interaction can be performed using different MS from this table. For the three-factor interaction the appropriate error is "`Var(Error)`", and for the two-factor interactions is "`Var(Error) + 5 Var(A*B*C)`".

However, no exact tests exist for the main effects of A, B, or C. That is, if we wish to test the hypothesis $Q(A)= bcr\Sigma\alpha^2/(a-1) = 0$, we cannot form a ratio of two expected mean squares such that the only term in the numerator that is not in the denominator is Q(A). However, it is likely that tests on the main effects are of central importance to the experimenter. The solution of this problem is to create synthetic errors, which are considered in the next section. As a first reference, below is the SAS result from the TEST statement:

```
                                The GLM Procedure
                Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: Y
Source                      DF     Type III SS     Mean Square    F Value    Pr > F
A                            2        0.372114        0.186057      -2.80         .
Error                   0.0957       -0.006351       -0.066377
Error: MS(A*B) + MS(A*C) - MS(A*B*C) + 24E-16*MS(Error) The last term is ~0, ignore it


Source                      DF     Type III SS     Mean Square    F Value    Pr > F
B                            4        2.530022        0.632506       1.57    0.4693
Error                    1.484        0.599150        0.403747
Error: MS(A*B) + MS(B*C) - MS(A*B*C) - 44E-17*MS(Error)


Source                      DF     Type III SS     Mean Square    F Value    Pr > F
C                            1        0.531676        0.531676       7.47    0.8437
Error                   0.0542        0.003857        0.071153
Error: MS(A*C) + MS(B*C) - MS(A*B*C) - 44E-17*MS(Error)
```

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| A*B | 8 | 2.963840 | 0.370480 | 0.78 | 0.6329 |
| A*C | 2 | 0.075772 | 0.037886 | 0.08 | 0.9240 |
| B*C | 4 | 2.032042 | 0.508011 | 1.07 | 0.4312 |
| **Error: MS(A*B*C)** | 8 | 3.797947 | 0.474743 | | |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| A*B*C | 8 | 3.797947 | 0.474743 | 0.46 | 0.8827 |
| **Error: MS(Error)** | 120 | 124.154374 | 1.034620 | | |

## 10. 5. Approximate F tests: synthetic errors

In factorial experiments with three or more factors, at least one of which is a random factor, and certain other, more complex designs, there are frequently **no exact test statistics** for certain effects in the models.

One possible "solution" to this dilemma is to assume that certain interactions are negligible. Although this seems an attractive possibility, there must be something in the nature of the system (or some strong prior knowledge) to justify such an assumption. In general, this assumption is not easily made, nor should it be taken lightly. We should not eliminate certain interactions from the model without conclusive evidence that it is appropriate to do so.

A procedure advocate by some experimenters is to test the interactions first and then set to zero those interactions that are found to be not significant. Once this is done, one can then assume that those non-significant interactions are zero when testing other effects in the same experiment. Although sometimes done in practice, this procedure has its own risks because any decision regarding an interaction is subject to both Type I and Type II errors.

If we cannot assume that certain interactions are negligible but we still need to make inferences about those effects for which exact tests do not exist, a procedure attributed to **Satterthwaite** (1946) can be employed. Satterthwaite's method utilizes **linear combinations of extant mean squares**, for example:

$$MS' = MS_r + \cdots + MS_S$$

and

$$MS'' = MS_U + \cdots + MS_V$$

where:

a) The mean squares are chosen so that no MS appear simultaneously in MS' and MS''
b) E(MS′) - E(MS″) is equal to the effect (the model parameter or variance component) considered in the null hypothesis. Then the test statistic would be

$$F = MS'/MS''$$

which is distributed approximately as $F_{p,q}$, where $p$ and $q$ are the ***effective degrees of freedom*** for the numerator and denominator, respectively**:**

$$p = \frac{(MS_r + \cdots + MS_S)^2}{MS_r^2/df_r + \cdots + MS_S^2/df_S} \qquad q = \frac{(MS_u + \cdots + MS_V)^2}{MS_U^2/df_U + \cdots + MS_V^2/df_V}$$

In these expressions for $p$ and $q$, $df_i$ is the number of degrees of freedom associated with the mean square $MS_i$. There is no assurance that $p$ and $q$ will be integers, so it will be necessary to interpolate in the tables of the F distribution.

For example, in the three-factor mixed effects model discussed above, it is relatively easy to see that an appropriate test statistic for $H_0$: $\alpha_1=...=\alpha_t=0$ would be

$$F = \frac{MS_A + MS_{ABC}}{MS_{AB} + MS_{AC}}$$

Why? Let's re-express this equation in terms of expected mean squares (refer back to the EMS table on page 12):

$$F = \frac{\sigma_\varepsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + br\sigma_{\alpha\gamma}^2 + cr\sigma_{\alpha\beta}^2 + bcr\sum\frac{\alpha^2}{a-1} + \sigma_\varepsilon^2 + r\sigma_{\alpha\beta\gamma}^2}{\sigma_\varepsilon^2 + r\sigma_{\alpha\beta\gamma}^2 + br\sigma_{\alpha\gamma}^2 + cr\sigma_{\alpha\beta}^2 \quad + \quad \sigma_\varepsilon^2 + r\sigma_{\alpha\beta\gamma}^2}$$

Notice that F =1 if and only if the effect of factor A ( $bcr\sum\frac{\alpha^2}{a-1}$ ) is 0. This ratio of linear combinations of mean squares is, therefore, an appropriate F test for Factor A.

The degrees of freedom for this F test can be estimated using the above equations for ***p*** and ***q***. Since no mean square appears in both the numerator and denominator, the numerator and denominator are independent from one another.

In general, it is better to construct linear combinations of mean squares in the numerator and denominator through addition rather than through subtraction [e.g. $MS_A$ / ($MS_{AB}$ + $MS_{AC}$ - $MS_{ABC}$)], because negative signs in such linear functions can lead to difficulties (ST&D pg. 380) such as close to 0 estimates for some MSE. Nevertheless, SAS uses the subtraction method for ease of calculation, and we will accept this approach in this class. Just disregard MSE estimates close to 0, such as the one produced above in the estimation of the error term for A:

```
Error: MS(A*B) + MS(A*C) - MS(A*B*C) + 24E-16*MS(Error) The last term is ~0, ignore it
```

## 10.5.2. An additional example with nested effects

The rules outlined in section 10.3.4.1 also apply to nested factorial experiments (i.e. factorial experiments with subsamples). The reason for this is that subsamples force the experimental units into the linear model, and experimental units (i.e. replications) are random, by definition. So even when all factors are fixed effects, nesting transforms the model into a mixed model.

To illustrate the use of the rules with a nested factorial experiment, consider a two-way factorial (fixed A, random B) with C (random) replications in each A x B combination and D subsamples measured on each replication. For all factors other than the error term, subscripts in parentheses give specific meaning to the subscript which precedes the parentheses (i.e. Rep(A*B) indicates that the ID variable "Rep" only has meaning once a specific combination of A and B levels is stated).

| Fixed or Random<br>Number of levels<br>Factor | F<br>a<br>i | R<br>b<br>j | R<br>c<br>k | R<br>d<br>l | **Expected Mean Squares** | **F** |
|---|---|---|---|---|---|---|
| $\alpha_i$ | 0 | b | c | d | $\sigma^2_\varepsilon + d\sigma^2_{\gamma(\alpha\beta)} + cd\sigma^2_{\alpha\beta} + bcd\Sigma\alpha^2/(a-1)$ | $MS_\alpha/MS_{\alpha\beta}$ |
| $\beta_j$ | a | 1 | c | d | $\sigma^2_\varepsilon + d\sigma^2_{\gamma(\alpha\beta)} + cd\sigma^2_{\alpha\beta} + acd\sigma^2_\beta$ | $MS_\beta/MS_{\alpha\beta}$ |
| $(\alpha\beta)_{ij}$ | **1** | 1 | c | d | $\sigma^2_\varepsilon + d\sigma^2_{\gamma(\alpha\beta)} + cd\sigma^2_{\alpha\beta}$ | $MS_{\alpha\beta}/MS_{\gamma(\alpha\beta)}$ |
| $\gamma(\alpha\beta)_{k\ (ij)}$ | 1 | 1 | 1 | d | $\sigma^2_\varepsilon + d\sigma^2_{\gamma(\alpha\beta)}$ | $MS_{\gamma(\alpha\beta)}/MS_\varepsilon$ |
| $\varepsilon_{(ijk)l}$ | 1 | 1 | 1 | 1 | $\sigma^2_\varepsilon$ | |

Note that the results are logical:

1. To test the interaction, we must use the MS of the factor (the EU) nested within the interaction, as with any previous nested design: $MS_{\alpha\beta}/MS_{\gamma(\alpha\beta)}$ .

2. To test A or B, we must use the interaction ($MS_{\alpha\beta}$) because our intention is to extend conclusions across the full population of treatment levels from which we drew the B sample of treatment levels. For effects A or B to be significant means they must be *significantly larger than their interaction*, the differences in responses to A at the different levels of B.

**Note on VARCOMP**

If you want to specify a nested effect within the interaction of two fixed factors, and use

```
model number= Block D R D*R plot(D*R*Block);
random plot(D*R*Block);
```

SAS ASSUMES that the D*R interaction is also random
To avoid that in the PROC VARCOMP you can use the option /fixed=n; in the model

FIXED=n tells the VARCOMP procedure that **the first n effects in the MODEL statement are fixed** effects. The remaining effects are assumed to be random. By default, PROC VARCOMP assumes that all effects are random in the model.

Example

```
proc varcomp Method= Type1;
  class D R Block plot;
  model number= D R D*R Block plot(D*R*Block)  / fixed = 3;
```

The examples is from the factorial lecture "Subsamples in a factorial RCBD"