Unbalanced two-factor ANOVA

The term "unbalanced" means that the sample sizes n_{ij} are not all equal. A balanced design is one in which all $n_{ij} = n$.

In the unbalanced case, there are two ways to define sums of squares for factors A and B.

Method 1 is the default in Stata, and goes by the name "partial SS" or "adjusted SS." In SAS it's called Type III sums of squares. For this method, here are the full and reduced models being tested for Factor A, Factor B, and the AB interaction:

(I've left the subscripts off in all of the following model statements. They should be obvious.)

Factor A:
Full model is $Y = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
Reduced model is $Y = \mu + \beta + \alpha\beta + \varepsilon$

Factor B:
Full model is $Y = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
Reduced model is $Y = \mu + \alpha + \alpha\beta + \varepsilon$

AB interaction:
Full model is $Y = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
Reduced model is $Y = \mu + \alpha + \beta + \varepsilon$

Method 2 is called sequential sums of squares, and is the default in R. To get it in Stata, you need to add the option ", sequential" to the end of the anova command. It's called Type I sums of squares in SAS.

Factor A:
Full model is $Y = \mu + \alpha + \varepsilon$
Reduced model is $Y = \mu + \varepsilon$

Factor B:
Full model is $Y = \mu + \alpha + \beta + \varepsilon$
Reduced model is $Y = \mu + \alpha + \varepsilon$

AB interaction: Same as above.

For Method 2, notice that it matters what order you use to name the factors, whereas it doesn't matter in Method 1.

Random Effects Models

A factor is called a <u>random effects factor</u> if the levels of the factor represent a larger set of interest.

Examples:

1. Medicine: How accurate are labs for testing for a certain disease? Do labs differ in their accuracy? Suppose we have people tested at 3 different labs.
Factor = Lab
Unit = a person having a medical test
Y_{ij} = accuracy rating of the test for person j and lab i

n_i = number of people tested at Lab i

Lab is a <u>fixed effect</u> if we care only about those labs.
Lab is a <u>random effect</u> if the 3 labs are a random sample of all such labs.

2. Ecology: How much species diversity is there in some forests of interest? (Species diversity is number of different species.) Suppose we have 4 forests to use.
Factor = Forest
Unit = a randomly selected one acre square within a forest
Y_{ij} = diversity for acre j in forest i.

n_i = number of different squares examined in Forest i.

Forest is a <u>fixed effect</u> if we care only about those 4 forests.
Forest is a <u>random effect</u> if those forest are randomly sampled from a larger set of interest.

3. Psychology: Compare therapists for effectiveness.
Factor = therapist
Unit = patient
Y_{ij} = change in score on depression test after one year of therapy for patient j with therapist i.

Therapist is a <u>fixed effect</u> if we are interested in those specific therapists
Therapist is a <u>random effect</u> if the therapists are randomly selected from all therapists of interest.

RANDOM EFFECTS MODEL (One factor only):

$Y_{ij} = \mu_i + \varepsilon_{ij} = \mu_. + \tau_i + \varepsilon_{ij}$ for $i = 1, \ldots, k; j = 1, \ldots, n_i$

Where:
1. ε_{ij} is $N(0, \sigma^2)$ as before
2. μ_i is $N(\mu_., \sigma_\mu^2)$, so the μ_i are *not* considered fixed as they were before.
3. $\mu_.$ is a fixed constant.
4. All μ_i and ε_{ij} are independent.

Picture: Normal curve with $\mu_.$ as the mean, each μ_i drawn from it. Then normal curves with each μ_i as the mean, each Y_{ij} drawn from that.

EXAMPLE: Therapists and depression scores

$\mu_.$ is the overall mean change in depression scores for the *population* of *all possible patients* (not just those treated) for *all possible therapists* of the type the sample was drawn from.

μ_i is the overall mean change in depression scores for the *population* of *all possible patients* if they were to have therapist i.

σ^2 is the variance of the changes in depression scores for the population for *any* particular therapist.

σ_μ^2 is the variance of the *means* μ_i for all possible therapists (not just the ones in the study), so it's the variability *across* therapists in the population.

The variance of Y_{ij} is made of *two* components. It's variance of $(\mu_i + \varepsilon_{ij}) = \sigma^2 + \sigma_\mu^2$ = variance across all possible patients and all possible therapists.

To summarize:
Fixed effects model, Y_{ij} is $N(\mu_i, \sigma^2)$ and all independent.

Random effects model, Y_{ij} is $N(\mu_., \sigma^2 + \sigma_\mu^2)$ and *not* all independent. For two units from the *same* level, Covariance is σ_μ^2 and correlation is $\dfrac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma^2}$ called the *intraclass correlation* because it represents the correlation between 2 population units in the *same* class or level of the factor. Note it's *large* if the variability is large *between* levels compared to *within* levels. This makes sense, because then knowing the value for one unit within a level gives good info about other units in that same level. If there's lots of variability *within* a level, then knowing one value doesn't help so much with others.

QUESTIONS OF INTEREST FOR THE RANDOM EFFECTS MODEL:

1. Test H_0: $\sigma_\mu^2 = 0$ vs Ha: $\sigma_\mu^2 > 0$. If the null hypothesis is true, all *possible* μ_i are equal.

Example: Are all therapists equally effective, on average?

How to do this: EASY! Still use the same F test as in fixed effects one-way ANOVA, $F^* = MSR/MSE$. In this case, $E\{MSE\} = \sigma^2$ and $E\{MSR\} = \sigma^2 + n' \sigma_\mu^2$ where n' is a weighted sample size.

2. Estimate μ. Use $\bar{Y}_{..}$, the overall mean. To get a confidence interval, use MSR instead of MSE as estimate of variance.

3. Estimate each component of the variances.

4. Estimate the intraclass correlation.

Details are messy – use computer!

EXAMPLE:
Datafile Name: Reading Test Scores
Datafile Subjects: Education
Story Names: Reading Comprehension
Reference: Moore, David S., and George P. McCabe (1989). *Introduction to the Practice of Statistics*. Original source: study conducted by Jim Baumann and Leah Jones of the Purdue University Education Department.
Authorization:
Description: Data from a study of the effect of three different methods of instruction on reading comprehension in children. Participants were given a reading comprehension test before and after receiving the instruction.
Number of cases: 66
Variable Names:

1. Subject: Subject number
2. Group: Type of instruction that student received (Basal, DRTA, or Strat)
3. PRE1: Pretest score on first reading comprehension measure
4. PRE2: Pretest score on second reading comprehension measure
5. POST1: Posttest score on first reading comprehension measure
6. POST2: Posttest score on second reading comprehension measure
7. POST3: Posttest score on third reading comprehension measure

We will use the POST1 scores as the response variable. Three methods in Stata (next page)

COMPARING FIXED, RANDOM AND REGRESS ANOVA OUTPUT

```
. oneway post1 group

                        Analysis of Variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      108.121212      2   54.0606061      5.32     0.0073
 Within groups           640.5     63   10.1666667
------------------------------------------------------------------------
    Total           748.621212     65   11.5172494

Bartlett's test for equal variances:  chi2(2) =   3.7349  Prob>chi2 = 0.155
```

```
. loneway post1 group

                 One-way Analysis of Variance for post1: POST1

                                              Number of obs =         66
                                                  R-squared =     0.1444

    Source                SS         df      MS            F     Prob > F
-------------------------------------------------------------------------
Between group          108.12121      2    54.060606      5.32     0.0073
Within group               640.5     63    10.166667
-------------------------------------------------------------------------
Total                  748.62121     65    11.517249

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ------------------------------------------------
            0.16405     0.17156       0.00000     0.50031

         Estimated SD of group effect           1.412508
         Estimated SD within group              3.188521
         Est. reliability of a group mean        0.81194
              (evaluated at n=22.00)
```

```
. encode group, generate(method)
. anova post1 method, regress

      Source |       SS       df       MS              Number of obs =      66
-------------+------------------------------           F(  2,    63) =    5.32
       Model |  108.121212     2  54.0606061           Prob > F      =  0.0073
    Residual |       640.5    63  10.1666667           R-squared     =  0.1444
-------------+------------------------------           Adj R-squared =  0.1173
       Total |  748.621212    65  11.5172494           Root MSE      =  3.1885

------------------------------------------------------------------------------
       post1        Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------------------------------------------------------
_cons            7.772727    .679795    11.43   0.000     6.414265    9.131189
method
          1     -1.090909   .9613753    -1.13   0.261    -3.012064    .8302461
          2             2   .9613753     2.08   0.042     .0788448    3.921155
          3     (dropped)
```