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 random effects factor 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 fixed effect if we care only about those labs.
Lab is a random effect 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 fixed effect if we care only about those 4 forests.
Forest is a random effect 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 fixed effect if we are interested in those specific therapists
Therapist is a random effect 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.

<u>See picture drawn on white board</u>: Normal curve with $\mu_.$ as the mean and σ_μ^2 as the variance, each μ_i drawn from it. Then normal curves with each μ_i as the mean and σ^2 as the variance, 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.

<u>To summarize</u> fixed versus random effects model for one factor:

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.

Continuing random effects model: For two units from the *same* level (same *i*), covariance between Y_{ij} and $Y_{ij'}$ is σ_μ^2 and correlation is $\frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma^2} = \frac{\text{Variance of } \mu_i}{\text{Variance of } Y_{ij}}$ called the *intraclass correlation*. It represents the correlation between 2 population units in the *same* class or level of the factor. It's *large* if the variability is large *between* levels compared to *within* levels (large σ_μ^2 compared to σ^2). 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. It also represents the percent of variability in the Y_{ij} that can be attributed to the variability in the μ_i's. (See picture on white board.)

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 $\mu_.$, Use $\bar{Y}_{..}$ = the overall mean. To get a confidence interval, use MSR instead of MSE as estimate of variance, because now the natural variability in the Y's includes both σ^2 and σ_μ^2.

3. Estimate each component of the variances.

4. Estimate the intraclass correlation.

Details are messy – use computer!

EXAMPLE: Source http://lib.stat.cmu.edu/DASL/Datafiles/ReadingTestScores.html
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)* = **Factor**
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* = **Response variable**
6. POST2: Posttest score on second reading comprehension measure
7. POST3: Posttest score on third reading comprehension measure

Stata: COMPARING FIXED, RANDOM AND REGRESS ANOVA OUTPUT

```
. oneway post1 group
```

This is the usual one factor anova. It tests H_0: all $\mu_i = 0$ for fixed effects, H_0: $\sigma^2_\mu = 0$ for random effects

```
                        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
```

This is called "large oneway anova" and gives intraclass correlation.

```
                 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)
```

Estimate of σ^2_μ (Estimated SD of group effect)

Estimate of σ^2 (Estimated SD within group)

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

Remember this recodes string variables as numeric

```
      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)
```

Note that teaching method 3 is the "control"; est. mean = 7.772727

Mixed Models - Fixed And Random Effects; Simplest - Randomized Block Design

Blocks are similar to pairing when there are only two factor levels (groups). They are used to reduce extraneous variability in responses. Often, block = a person or "subject."

Example: Does skin response change for different *imagined* responses?
Y_{ij} = Skin potential for person (block) i and emotion j. Note that $i = 1$ to 8 and $j = 1$ to 4 *for each person*.

Model: $Y_{ij} = \mu_{..} + \rho_i + \tau_j + \varepsilon_{ij}$ where ε_{ij} are $N(0, \sigma^2)$

Blocks = Factor A = person, fixed or random. If fixed, $\sum_{i=1}^{a} \rho_i = 0$; if random, ρ_i are $N(0, \sigma^2_\rho)$

Treatment = Factor B = emotion. This is a fixed effect, and thus $\sum_{j=1}^{b} \tau_j = 0$.

Note: We cannot include *interaction* in the model because there is no way to estimate both interaction and error. This is because there is only one observation in each *ij* combination.

ANOVA Table for Randomized Block Design (blocks as fixed effect*):

Source	SS	df	MS	E{MS}	F*
Blocks (Subjects)	SSBL	a – 1	MSB	$\sigma^2 + \frac{b\sum \rho_i^2}{a-1}$	Not tested
Treatments	SSTR	b – 1	MSTR	$\sigma^2 + \frac{a\sum \tau_i^2}{b-1}$	$\frac{MSTR}{MSE}$
Error = Blks x Trts	SSE	(a – 1)(b – 1)	MSE	σ^2	
TOTAL	SSTO	ab – 1			

*When blocks considered to be random, the only difference is $E\{MSBL\} = \sigma^2 + b\,\sigma^2_\rho$

Example: H_0: all $\tau_j = 0$ (*population* mean skin response same for all 4 imagined emotions)

F* = 3.47, *p*-value = .034, reject the null hypothesis, conclude population means differ.

Tukey multiple comparisons: Calm differs from Fearful, all other C.I.s cover 0.

See computer output for what happens if blocks are ignored. F* = 0.13, *p*-value = 0.941!

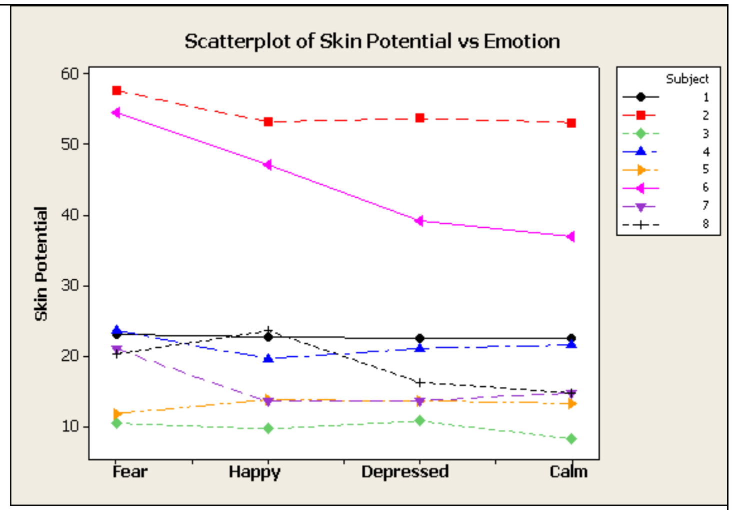
Clearly a good idea to use blocks. Natural variability *across* people is huge. Blocks control for that. See plot. Two people have very high skin potential measurements in general, others are much lower.

Statistics 110/201

Randomized Block Design (Subjects = Blocks, Emotions = Treatments)

Eight subjects were hypnotized and asked to imagine four emotions. Session order was randomized for each subject. Skin potential was measured as the response variable. Ref: "Physiological effects during hypnotically requested emotions" *Psychosomatic Medicine*, 1963, pp. 334-43. (Also Cobb, *Des. & Analysis of Exp.*)

Subj.	Calm	Depressed	Fearful	Happy	Mean
1	22.600	22.500	23.100	22.700	22.7
2	53.100	53.700	57.600	53.200	54.4
3	8.300	10.800	10.500	9.700	9.8
4	21.600	21.100	23.600	19.600	21.5
5	13.300	13.700	11.900	13.800	13.2
6	37.000	39.200	54.600	47.100	44.5
7	14.800	13.700	21.000	13.600	15.8
8	14.800	16.300	20.300	23.600	18.7
Mean:	23.2	23.9	27.8	25.4	



Analysis of Variance

Source	DF	SS	MS	F	P
Subject	7	7021.87	1003.12	102.86	0.000
Emotions	3	101.43	33.81	3.47	0.034
Subj.xEmotion	21	204.80	9.75		
Total	31	7328.10			

One-way Analysis of Variance (*Ignoring Blocks*)

Source	DF	SS	MS	F	P
Emotions	3	101	34	0.13	0.941
Error	28	7227	258		
Total	31	7328			

Tukey 95.0% Simultaneous Confidence Intervals

Response Variable Skin Potential

All Pairwise Comparisons among Levels of Emotions

```
Emotions = Calm subtracted from:
Emotions     Lower    Center     Upper   ----+---------+---------+---------+--
Depresse    -3.663    0.6875     5.038          (-------*--------)
Fear         0.287    4.6375     8.988                  (--------*--------)
Happy       -2.125    2.2250     6.575             (--------*--------)
                                         ----+---------+---------+---------+--
                                           -5.0       0.0       5.0      10.0

Emotions = Depresse subtracted from:
Emotions     Lower    Center     Upper   ----+---------+---------+---------+--
Fear        -0.400     3.950     8.300                (---------*--------)
Happy       -2.813     1.538     5.888           (---------*--------)
                                         ----+---------+---------+---------+--
                                           -5.0       0.0       5.0      10.0

Emotions = Fear subtracted from:
Emotions     Lower    Center     Upper   ----+---------+---------+---------+--
Happy       -6.763    -2.412     1.938   (---------*--------)
                                         ----+---------+---------+---------+--
```