

Two-way Analysis of Variance

1. Two-way (fixed effects) ANOVA with equal group sizes

In the chapter on one-way ANOVA, we analyzed data from 2 fictitious experiments examining the effect of distraction on ability to solve mental arithmetic problems. Each of those experiments had 3 independent groups with $n=5$ observations per group. Now let us suppose that rather than 2 separate experiments, we have one experiment with 2 independent variables, as follows:

<u>Independent Variable</u>	<u>Level of Independent variable</u>
A. Difficulty of Problems	A_1 – easy problems A_2 – difficult problems
B. Number of Distractors	B_1 – 1 distractor B_2 – 2 distractors B_3 – 3 distractors

The factorial combination of these 2 independent variables (or “factors”) results in 6 independent groups of subjects, or 6 “cells”. The means of interest might be represented as follows:

	B_1	B_2	B_3	
A_1	$\overline{AB}_{11} = 3.8$	$\overline{AB}_{12} = 4.0$	$\overline{AB}_{13} = 4.0$	$\overline{A}_1 = 3.933$
A_2	$\overline{AB}_{21} = 4.8$	$\overline{AB}_{22} = 5.4$	$\overline{AB}_{23} = 10.4$	$\overline{A}_2 = 6.867$
	$\overline{B}_1 = 4.300$	$\overline{B}_2 = 4.700$	$\overline{B}_3 = 7.200$	$\overline{Y}_\bullet = 5.400$

Textbook authors often lay out the data in the 3 x 2 crosstabulation format shown above, and proceed to talk about row and column effects. I have always preferred to use A and B to refer to the two independent variables, however. The main reason for this is that if go on to three-way ANOVA, it is not clear to me what I should call the third independent variable, or how I should lay out the data if I’ve use Row and Column for the first two independent variables; but if I’ve used A and B (and the layout in Table 1), I simply add variable C.

Table 1 displays the data in a format that will be convenient, given the conceptual approach we will be taking (i.e., the first conceptual approach discussed in the one-way ANOVA chapter).

Table 1: Partitioning of SS_{Total} for a 2x3 factorial experiment.

Subject	A	B	Y	\bar{Y}_{\bullet}	\overline{AB}_{ij}	\bar{A}_i	\bar{B}_j	$(Y - \bar{Y}_{\bullet})^2$	$(\overline{AB}_{ij} - \bar{Y}_{\bullet})^2$	$(Y - \overline{AB}_{ij})^2$	$(\bar{A}_i - \bar{Y}_{\bullet})^2$	$(\bar{B}_j - \bar{Y}_{\bullet})^2$	$[(\overline{AB}_{ij} - \bar{Y}_{\bullet}) - (\bar{A}_i - \bar{Y}_{\bullet}) - (\bar{B}_j - \bar{Y}_{\bullet})]^2$
1	1	1	5	5.400	3.800	3.933	4.300	0.160	2.560	1.440	2.151	1.210	0.934
2	1	1	6	5.400	3.800	3.933	4.300	0.360	2.560	4.840	2.151	1.210	0.934
3	1	1	2	5.400	3.800	3.933	4.300	11.560	2.560	3.240	2.151	1.210	0.934
4	1	1	4	5.400	3.800	3.933	4.300	1.960	2.560	0.040	2.151	1.210	0.934
5	1	1	2	5.400	3.800	3.933	4.300	11.560	2.560	3.240	2.151	1.210	0.934
6	1	2	5	5.400	4.000	3.933	4.700	0.160	1.960	1.000	2.151	0.490	0.588
7	1	2	4	5.400	4.000	3.933	4.700	1.960	1.960	0.000	2.151	0.490	0.588
8	1	2	2	5.400	4.000	3.933	4.700	11.560	1.960	4.000	2.151	0.490	0.588
9	1	2	4	5.400	4.000	3.933	4.700	1.960	1.960	0.000	2.151	0.490	0.588
10	1	2	5	5.400	4.000	3.933	4.700	0.160	1.960	1.000	2.151	0.490	0.588
11	1	3	3	5.400	4.000	3.933	7.200	5.760	1.960	1.000	2.151	3.240	3.004
12	1	3	2	5.400	4.000	3.933	7.200	11.560	1.960	4.000	2.151	3.240	3.004
13	1	3	6	5.400	4.000	3.933	7.200	0.360	1.960	4.000	2.151	3.240	3.004
14	1	3	5	5.400	4.000	3.933	7.200	0.160	1.960	1.000	2.151	3.240	3.004
15	1	3	4	5.400	4.000	3.933	7.200	1.960	1.960	0.000	2.151	3.240	3.004
16	2	1	5	5.400	4.800	6.867	4.300	0.160	0.360	0.040	2.151	1.210	0.934
17	2	1	6	5.400	4.800	6.867	4.300	0.360	0.360	1.440	2.151	1.210	0.934
18	2	1	2	5.400	4.800	6.867	4.300	11.560	0.360	7.840	2.151	1.210	0.934
19	2	1	5	5.400	4.800	6.867	4.300	0.160	0.360	0.040	2.151	1.210	0.934
20	2	1	6	5.400	4.800	6.867	4.300	0.360	0.360	1.440	2.151	1.210	0.934
21	2	2	9	5.400	5.400	6.867	4.700	12.960	0.000	12.960	2.151	0.490	0.588
22	2	2	5	5.400	5.400	6.867	4.700	0.160	0.000	0.160	2.151	0.490	0.588
23	2	2	3	5.400	5.400	6.867	4.700	5.760	0.000	5.760	2.151	0.490	0.588
24	2	2	7	5.400	5.400	6.867	4.700	2.560	0.000	2.560	2.151	0.490	0.588
25	2	2	3	5.400	5.400	6.867	4.700	5.760	0.000	5.760	2.151	0.490	0.588
26	2	3	7	5.400	10.400	6.867	7.200	2.560	25.000	11.560	2.151	3.240	3.004
27	2	3	10	5.400	10.400	6.867	7.200	21.160	25.000	0.160	2.151	3.240	3.004
28	2	3	11	5.400	10.400	6.867	7.200	31.360	25.000	0.360	2.151	3.240	3.004
29	2	3	12	5.400	10.400	6.867	7.200	43.560	25.000	2.560	2.151	3.240	3.004
30	2	3	12	5.400	10.400	6.867	7.200	43.560	25.000	2.560	2.151	3.240	3.004
Sums of Squares:								243.200	159.200	84.000	64.533	49.400	45.267

The first step in performing a two-way ANOVA, oddly enough, is to ignore the factorial nature of the design, and treat it as a one-way ANOVA. In a one-way ANOVA, we saw that each score's deviation from the grand mean could be broken down into two further deviations: the deviation of the score's group mean from the grand mean; and the deviation of the score from its group mean. The only difference in two-way ANOVA is that the independent groups are often referred to as **cells**.

Table 2: Description of the squared deviations in Table 1.

Column	Description	Column Sum
$(Y - \bar{Y}_{\bullet})^2$	Squared deviation of each score from the grand mean	SS_{Total}
$(\bar{AB}_{ij} - \bar{Y}_{\bullet})^2$	Squared deviation of each score's cell mean from the grand mean	$SS_{\text{Between-cells}}$
$(Y - \bar{AB}_{ij})^2$	Squared deviation of each score from its cell mean	$SS_{\text{Within-cells, or Error}}$
$(\bar{A}_i - \bar{Y}_{\bullet})^2$	Squared deviation of each score's A-mean from the grand mean	SS_A
$(\bar{B}_j - \bar{Y}_{\bullet})^2$	Squared deviation of each score's B-mean from the grand mean	SS_B
$[(\bar{AB}_{ij} - \bar{Y}_{\bullet}) - (\bar{A}_i - \bar{Y}_{\bullet}) - (\bar{B}_j - \bar{Y}_{\bullet})]^2$	Square of: [(cell mean – grand mean) – (A-mean – grand mean) – (B-mean – grand mean)]	SS_{AB}

For the moment, the only columns of interest in Table 1 are those headed $(Y - \bar{Y}_{\bullet})^2$, $(\bar{AB}_{ij} - \bar{Y}_{\bullet})^2$, and $(Y - \bar{AB}_{ij})^2$. As shown in Table 2, the sums at the bottom of these columns provide the sums of squares we would need to perform a one-way ANOVA on these data. Also note that $SS_{\text{Between-cells}} + SS_{\text{Within-cells}} = SS_{\text{Total}}$ ($159.2 + 84.0 = 243.2$). There are $2 \times 3 = 6$ cells, so $df_{\text{Between-cells}} = 5$. There are 30 subjects in total, so $df_{\text{Total}} = 29$. By subtraction, $df_{\text{Within-cells}} = 29 - 5 = 24$. Alternatively, there are 6 cells with 5 subjects in each: We lose one degree of freedom for each cell, so $df_{\text{Within-cells}} = 30 - 6 = 24$. This leads to the mean squares and F-ratio reported in Table 3.

Table 3: Partial ANOVA summary for data in Table 1.

Source of Variation	SS	df	MS	F	p
Between Cells	159.200	5	31.840	9.097	< .001
Within Cells	84.000	24	3.500		
Total	243.200	29			

The next step is to recognize that $SS_{\text{Between-cells}}$ can be partitioned further into 3 components. The first of these is based on the deviation of the A-means from the grand mean. Look at the column headed \bar{A}_i in Table 1. This column shows the A-mean for each score—i.e., the mean of all scores (including that score itself) that share the score's level of variable A. In Table 1, A=1 for

the first 15 scores, and $A=2$ for the next 15. Therefore, \bar{A}_1 = the mean of the 1st 15 scores, and \bar{A}_2 = the mean of the last 15 scores.

The column headed $(\bar{A}_i - \bar{Y}_{..})^2$ shows the squared deviation of the A-mean from the grand mean. The sum at the bottom of this latter column, therefore, shows SS_A , the portion of $SS_{\text{Between-cells}}$ that is due to differences in variable A: $SS_A = 64.533$.

Similarly, the column headed \bar{B}_j shows the B-mean for each score—the mean of all the scores that share that score's level of variable B. The column headed $(\bar{B}_j - \bar{Y}_{..})^2$ shows the squared deviation of each score's B-mean from the grand mean, and the sum at the bottom of this column gives SS_B , the portion of $SS_{\text{Between-cells}}$ that is due to differences in variable B. $SS_B = 49.400$.

The final column in Table 1 is headed $[(\bar{AB}_{ij} - \bar{Y}_{..}) - (\bar{A}_i - \bar{Y}_{..}) - (\bar{B}_j - \bar{Y}_{..})]^2$. The first part of this expression, $(\bar{AB}_{ij} - \bar{Y}_{..})$, gives the deviation of each cell mean from the grand mean, which is the basis of $SS_{\text{Between-cells}}$, as we saw earlier. Subtracted from this are the deviations of the A-mean from the grand mean and of the B-mean from the grand mean. So inside the square brackets, we have that part of the cell mean's deviation from the grand mean that is left over after we've taken into account the so-called **main effects** of A and B. Squaring and summing yields $SS_{A \times B}$, or $SS_{\text{Interaction}}$. Note that $SS_{A \times B} = SS_{\text{Between-cells}} - SS_A - SS_B$: $SS_{A \times B} = 159.200 - 64.553 - 49.400 = 45.267$.

This information can be summarized as shown in Table 4.

Table 4: Summary of two-way ANOVA for the data in Table 1.

Source of Variation	SS	df	MS	F	p
Between Cells	159.200	5			
A	64.553	1	64.553	18.444	< 0.001
B	49.400	2	24.700	7.057	0.004
AxB	45.267	2	22.634	6.467	0.006
Within Cells	84.000	24	3.500		
Total	243.200	29			

You may have noticed that in Table 4, no F-test is reported on the line labeled “Between Cells”. There would be nothing wrong with reporting an F-test here, but it is customary not to do so.

The degrees of freedom for a two-way ANOVA are computed as follows:

- df_A = the number of levels of variable A minus 1
- df_B = the number of levels of variable B minus 1
- $df_{A \times B} = df_A \times df_B$
- Alternatively, $df_{A \times B} = df_{\text{Between-cells}} - df_A - df_B$

The partitioning of SS_{Total} and df_{Total} is summarized in Figure 1.

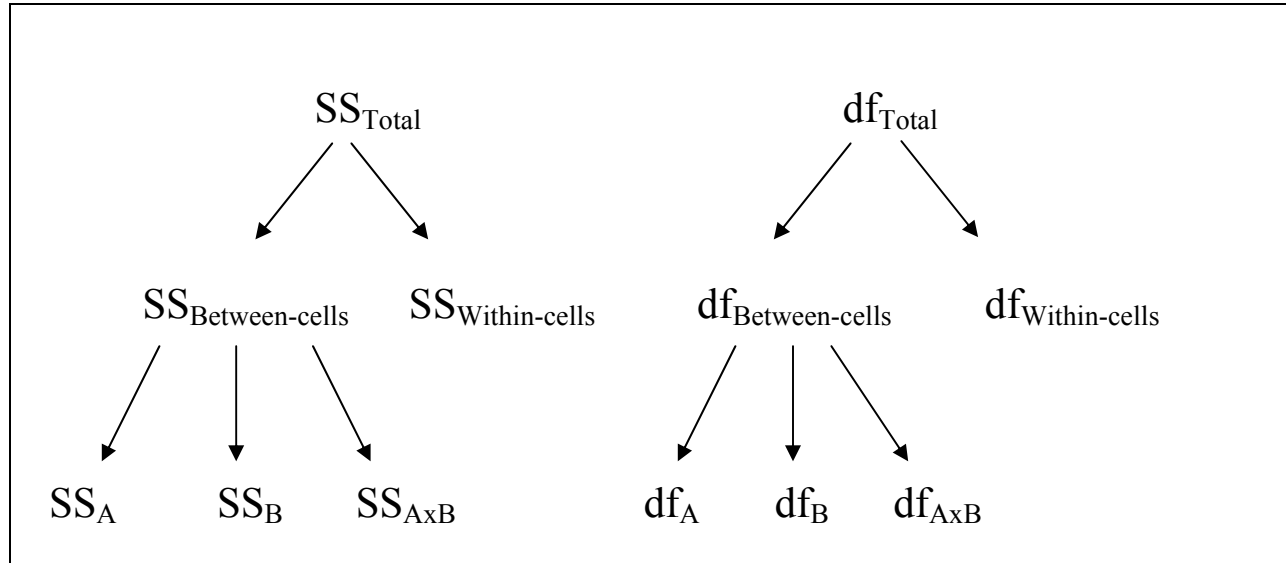


Figure 1: Partitioning diagram for two-way ANOVA.

2. Null hypotheses for two-way ANOVA

I have not yet said anything about the null hypotheses for the 3 F-tests shown in Table 4. For the main effects, the null hypotheses are similar to what we saw for the one-way ANOVA:

Main effect of A: $H_0 : \mu_{A_1} = \mu_{A_2} \dots = \mu_{A_j}$ { where $j = \#$ of levels of variable A }

Main effect of B: $H_0 : \mu_{B_1} = \mu_{B_2} \dots = \mu_{B_k}$ { where $k = \#$ of levels of variable B }

The null hypothesis for the interaction effect can be stated in various ways. For example:

- The simple main effects of variable A are all the same—i.e., the effect of A is the same at all levels of variable B.
- The simple main effects of variable B are all the same—i.e., the effect of B is the same at all levels of variable A.

Imagine plotting the AxB cell means using a line graph. You could plot A on the X-axis and draw a line for each level of B. Or you could plot B on the X-axis and draw a line for each level of A. These two plots would correspond respectively to the 1st and 2nd statements of the null hypothesis shown above. And if the null hypothesis were true, the lines would be (more or less) parallel in both plots.

Another to think about interaction that I have always found useful is to consider how you would answer a simple question about one of the main effects. For example: *What is the effect of*

variable B? If you can give a simple answer, there is no interaction. But if you answer with, “It depends on which level of A you’re looking at,” then you have an interaction.

To make this more concrete, here’s a question about the fictitious data we have been using: *How does increasing the number of distractors affect problem solving performance?* In this case, I hope you would answer that **it depends on the difficulty of the problems**. For the easier problems, increasing the number of distractors had no statistically significant effect on problem solving performance; but for the more difficult problems, performance became poorer as the number of distractors was increased.

Finally, note that the null hypothesis for the interaction term does **not** state that there are no differences among the population means for the AxB cells.¹ There can be differences among those cells (e.g., one or both main effects may be significant), but no interaction. In other words, there may be a main effect of variable A, but that effect is the same at all levels of B; or vice versa.

3. Fixed and random effects

We have not yet addressed the distinction between **fixed** and **random** effects. This distinction also applies to one-way ANOVA designs, but I delayed discussion until now, because it becomes very important in the context of two-way ANOVA². Fixed and random effects can be distinguished as follows:

- | | |
|-----------------------|---|
| Fixed effect: | An independent variable is a fixed effect if the levels are deliberately chosen, and are the only levels of interest. In a replication of the study, the same levels would be used. |
| Random effect: | An independent variable is a random effect if the levels are randomly selected (the subjects in an experiment, for example). In a replication, a new set of randomly selected levels would be used. |

Actually, this is a bit of an oversimplification. Howell (1997) provides a useful discussion of this issue in terms of the sampling fraction, which he defines as follows:

The sampling fraction for a variable is the ratio of the number of levels of [that] variable that are *actually* used to the potential number of levels that *could have been* used.

Following Howell, I will use lowercase letters to indicate the number of levels actually used (e.g., a = the number of levels of independent variable A in a study), and uppercase letters to represent the total number of levels available for use—or the total number of levels of interest (e.g., A = the total number of levels of independent A available for study, or of interest).

If $a = A$, and $a/A = 1$, then independent variable A is fixed. That is, the levels of variable A under study are the only ones available, or the only ones of interest.

¹ That would be the null hypothesis for the “Between-cells” F-ratio in Table 4, had we chosen to report it.

² As we will see shortly, how one calculates the F-ratios depends on whether the independent variables are fixed or random effects.

For a random variable, on the other hand, the number of available levels (A) may be very large, even infinite in the case of continuous variables. The number of levels randomly selected for study (a) is typically very small in comparison. And so for all practical purposes, the sampling fraction $a/A = 0$.

Howell summarized the sampling fractions possible for a two-way ANOVA in his Table 13.6, which is reproduced here as Table 5. Note that Subjects is shown only as a random effect, because subjects are virtually always treated as a (random) sample from the population of subjects available for study.³

Table 5: Sampling fractions for the two-way ANOVA

Variable	Sampling fraction
A (fixed)	$a/A = 1$
A (random)	$a/A = 0$
B (fixed)	$b/B = 1$
B (random)	$b/B = 0$
Subjects (random)	$n/N = 0$

Using the sampling fraction, the expected mean squares for a two-way ANOVA can be summarized as shown in Table 6 (Table 13.7 in Howell).

Table 6: Expected mean squares for all two-way ANOVA designs.

Source	Expected MS
A	$(1 - n / N)\sigma_e^2 + n(1 - b / B)\sigma_{\alpha\beta}^2 + nb\sigma_\alpha^2$
B	$(1 - n / N)\sigma_e^2 + n(1 - a / A)\sigma_{\alpha\beta}^2 + na\sigma_\beta^2$
AB	$(1 - n / N)\sigma_e^2 + n\sigma_{\alpha\beta}^2$
Error	$(1 - n / N)\sigma_e^2$

Each expected MS in Table 6 contains the term $(1 - n / N)\sigma_e^2$. Because it is almost always assumed that subjects are sampled at random from a large population of potential subjects, n/N is essentially equal to 0. Therefore, $(1 - n / N)\sigma_e^2 = (1 - 0)\sigma_e^2 = \sigma_e^2$. Making similar substitutions for a/A and b/B (based on the information in Table 5), we can obtain the expected mean squares for models where:

- both A and B are fixed (Table 7)
- both A and B are random (Table 7)

³ Note too that it is possible to have a sampling fraction between 0 and 1. This may happen, for example, if the variable is something like school class, where the number of available levels available is relatively small. Howell suggests that in cases like this, the variable is usually treated as fixed.

- A is fixed and B is random (Table 8)
- A is random and B is fixed (Table 8)

Table 7: Expected mean squares for **fixed** and **random** effects two-way ANOVA models.

Source	Fixed effects model: A and B both Fixed	Random effects model: A and B both Random
<i>A</i>	$\sigma_e^2 + nb\sigma_\alpha^2$	$\sigma_e^2 + n\sigma_{\alpha\beta}^2 + nb\sigma_\alpha^2$
<i>B</i>	$\sigma_e^2 + na\sigma_\beta^2$	$\sigma_e^2 + n\sigma_{\alpha\beta}^2 + na\sigma_\beta^2$
<i>AB</i>	$\sigma_e^2 + n\sigma_{\alpha\beta}^2$	$\sigma_e^2 + n\sigma_{\alpha\beta}^2$
<i>Error</i>	σ_e^2	σ_e^2

Table 8: Expected mean squares for **mixed effects** two-way ANOVA models.

Source	Mixed effects model 1: A fixed, B random	Mixed effects model 2: A random, B fixed
<i>A</i>	$\sigma_e^2 + n\sigma_{\alpha\beta}^2 + nb\sigma_\alpha^2$	$\sigma_e^2 + nb\sigma_\alpha^2$
<i>B</i>	$\sigma_e^2 + na\sigma_\beta^2$	$\sigma_e^2 + n\sigma_{\alpha\beta}^2 + na\sigma_\beta^2$
<i>AB</i>	$\sigma_e^2 + n\sigma_{\alpha\beta}^2$	$\sigma_e^2 + n\sigma_{\alpha\beta}^2$
<i>Error</i>	σ_e^2	σ_e^2

In the two-way fixed effects model (i.e., both A and B fixed), which we saw in section 1 of this chapter, the 3 F-ratios are all computed with MS_{error} in the denominator. The reason for this is clear when you look at the expected mean squares in Table 7.

$$\text{Variable A: } E(F) = \frac{E(MS_A)}{E(MS_{error})} = \frac{\sigma_e^2 + nb\sigma_\alpha^2}{\sigma_e^2} \quad (1)$$

$$\text{Variable B: } E(F) = \frac{E(MS_B)}{E(MS_{error})} = \frac{\sigma_e^2 + na\sigma_\beta^2}{\sigma_e^2} \quad (2)$$

$$\text{AxB interaction: } E(F) = \frac{E(MS_{AB})}{E(MS_{error})} = \frac{\sigma_e^2 + n\sigma_{\alpha\beta}^2}{\sigma_e^2} \quad (3)$$

In each of equations 1-3, σ_e^2 appears in both the numerator and the denominator. And each numerator contains only **one additional term**. Furthermore, when the null hypothesis for a particular effect is true, the expected value of that additional term in the numerator is equal to 0.

If the null hypothesis for the main effect of variable A is true, for example, $nb\sigma_\alpha^2$ will equal 0; and so:

$$\text{Variable A when } H_0 \text{ is true: } E(F) = \frac{\sigma_e^2 + nb\sigma_\alpha^2}{\sigma_e^2} = \frac{\sigma_e^2 + 0}{\sigma_e^2} = 1 \quad (4)$$

In general, the numerator of the F -ratio for a particular effect must contain the same term(s) as the denominator, plus **one additional term**. The expected value of that additional term is 0 if the null hypothesis for the effect is true, and greater than 0 if the null hypothesis is false. Thus, the expected value of F is (about) 1 if H_0 is true, and greater than 1 if H_0 is false.

If we take this general principle and apply it to the information contained in Tables 7 and 8, we can see that the F -ratios for fixed, random and mixed designs ought to be computed as shown in Table 9.

Table 9: Computation of F -ratios in fixed, random, and mixed effects designs.

Effect	<u>Fixed Effects</u>	<u>Random Effects</u>	<u>Mixed Effects Models</u>	
	A fixed B fixed	A random B random	A fixed B random	A random B fixed
A	$F = \frac{MS_A}{MS_{error}}$	$F = \frac{MS_A}{MS_{AB}}$	$F = \frac{MS_A}{MS_{AB}}$	$F = \frac{MS_A}{MS_{error}}$
B	$F = \frac{MS_B}{MS_{error}}$	$F = \frac{MS_B}{MS_{AB}}$	$F = \frac{MS_B}{MS_{error}}$	$F = \frac{MS_B}{MS_{AB}}$
$A \times B$	$F = \frac{MS_{AB}}{MS_{error}}$	$F = \frac{MS_{AB}}{MS_{error}}$	$F = \frac{MS_{AB}}{MS_{error}}$	$F = \frac{MS_{AB}}{MS_{error}}$

In the case where both variables are fixed, all 3 F -ratios are computed with MS_{error} in the denominator, as we saw above. When both variables are random, the F -ratio denominator is MS_{AB} for the two main effects, and MS_{error} for the AB interaction.

For mixed effects models, things are slightly more interesting. The F -ratio denominator is MS_{error} for the AB interaction term and for the main effect of the **random factor**. But for the main effect of the **fixed factor**, the F -ratio denominator is MS_{AB} . This arrangement may seem counterintuitive. But if you review what we have discussed here, paying attention to the roll of the sampling fraction, I think you will find that it does indeed follow.

3.1 Pooling error terms

As we have just seen, in both random and mixed effects models, one or more F -ratios is computed with MS_{AB} as the denominator. Because df_{AB} is typically a lot lower than df_{error} , the

critical F -values for these tests will be substantially higher than for F -ratios computed using MS_{error} . Putting it another way, when F is computed with MS_{AB} as the denominator, the test will have relatively low power.

Howell (1997, p. 425) suggests the following solution to this problem:

One way out of this difficulty lies in first showing that there is no evidence to cause us to doubt that $\sigma_{\alpha\beta}^2 = 0$, and then dropping the interaction term from the model. If this is possible, we may now pool MS_{AB} and MS_{error} , forming a new error term, and use this to test the main effects.

In practice, this solution entails the following:

- Start with the full model that includes terms for the A and B main effects, and the AB interaction
- Test the null hypothesis for the interaction term using a relatively high alpha (e.g., Howell recommends 0.25)⁴
- If H_0 for the interaction term is not rejected, drop the interaction term from the model

Dropping the interaction term from the model essentially results in pooling it with the error term. The new (reduced) model will have terms for each main effect, and the error term, or residual term, will consist of everything left over.

$$MS_{residual} = \frac{SS_{AB} + SS_{within-cells}}{df_{AB} + df_{within-cells}} = \frac{SS_{residual}}{df_{residual}} \quad (5)$$

4. Two-way ANOVA with unequal sample sizes

We saw earlier that having unequal sample sizes is no great problem if you have a one-way ANOVA design, provided that the population variances are not too heterogeneous.⁵ But when you have a two-way design, unequal sample sizes can be quite problematic.

In a **balanced** two-way ANOVA design (where *balanced* means equal sample sizes for all cells), there is **no confounding** of A and B. Perhaps the easiest way to see what this means is to perform a χ^2 test of association on the cell counts for a balanced design, using equation (6).

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (6)$$

In equation (6), O = **observed frequency** (or the actual sample size for the cell), and E = **expected frequency**. As you no doubt recall, expected frequency for a cell is calculated as follows:

⁴ The reason for using a higher than normal alpha is that in this situation, we do not want to reject H_0 . And if we do fail to reject H_0 , we want to be confident that it is not simply for lack of statistical power. Raising alpha is one way to raise the power of the test.

⁵ As Howell (1997) put it, "heterogeneity of variance and unequal sample sizes do not mix."

$$\text{expected frequency} = \frac{(\text{row total})(\text{column total})}{(\text{grand total})} \quad (7)$$

The observed cell counts for the balanced two-way ANOVA we performed earlier in this chapter are shown in Table 10. Note that in every cell, *expected count* = $(15 \cdot 10)/30 = 5 = \text{observed count}$. Therefore, $\chi^2 = 0$. Note that this will be the case for any balanced design, because there is **no association between A and B**, or **no confounding of A and B**. Furthermore, it is precisely because of this **orthogonality**, or absence of confounding that:

- i. $SS_A + SS_B + SS_{AB} + SS_{\text{error}} = SS_{\text{Total}}$; and
- ii. the F -tests for A, B, and the AB interaction are all mutually independent.

Table 10: Observed cell counts for the balanced two-way design analyzed in section 1 of this chapter.

	B₁	B₂	B₃	Totals
A₁	5	5	5	15
A₂	5	5	5	15
Totals	10	10	10	30

Another way to think of this is illustrated using a Venn diagram, as in Figure 2. The rectangle represents SS_{Total} , and the 3 circles represent SS_A , SS_B , and SS_{AB} . The area inside the rectangle but outside all of the circles represents SS_{residual} (or SS_{error}), that portion of SS_{Total} that is not explained by A, B, or the AB interaction. The important thing to note about Figure 2 is that none of the 3 circles overlap, because A, B, and the AB interaction each explains a unique, non-overlapping portion of the total variability. This is what is meant by orthogonality.

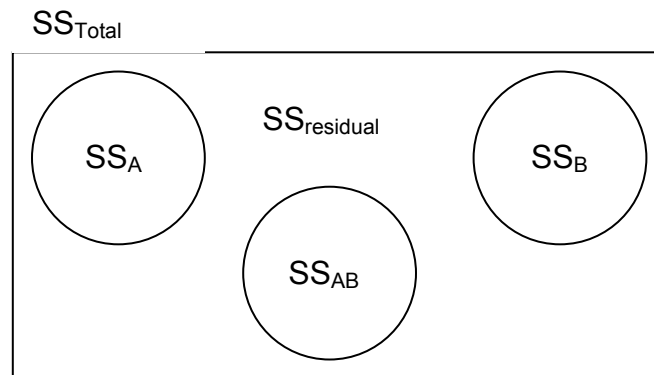


Figure 2: Venn diagram illustrating orthogonality in a two-way ANOVA design.

4.1 Unequal but **proportional** cell counts

Table 11, which is an adaptation of Figure 18-1 in Kleinbaum et al (1988), shows another interesting case. Here, the cell counts are not equal, but they are **proportional**. In every row, the cell counts are in the ratio 4:2:3; and in every column, the counts are in the ratio 1:2:3:2.

The interesting thing about this table, or any other table with such proportionality, is that each expected count will be equal to the observed cell count. For the shaded cell with $O=6$, for example, $E = (27)(16) / 72 = 6$.

Table 11: Example of proportional cell counts in a 4x3 design.

	Factor 2			Totals
	4	2	3	
Factor 1	8	4	6	18
	12	6	9	27
	8	4	6	18
Totals	32	16	24	72

Because of this property of designs with proportional cell counts, the A and B main effects are independent of each other (orthogonal)—so in a Venn diagram, the A and B circles would not overlap. *But the interaction effect is not independent of the main effects.* In a Venn diagram, the AB interaction circle would overlap with one or both of the A and B circles. Therefore, you can calculate SS_A , and SS_B in the same way you would for a balanced design; but you must obtain SS_{AB} by subtraction, as follows:

$$SS_{AB|A,B} = SS_{\text{Between-cells}} - SS_A - SS_B \quad (8)$$

I used the notation $SS_{AB|A,B}$ for the interaction sum of squares to indicate that it is really SS_{AB} **given** that A and B are already in the model. In other words, this is the portion of SS_{AB} that does not overlap with SS_A or SS_B . (This is sometimes called the portion of variance that is **uniquely** explained by the AB interaction.)

For the sake of completeness, I should point out that I could have used this same kind of conditional notation for the balanced design we looked at earlier. That is:

$$SS_{\text{Between-cells}} = SS_A + SS_{B|A} + SS_{AB|A,B} \quad (9)$$

The reason we did not use that notation earlier is that for balanced designs, the conditional sums of squares in equation (9) are equal to the unconditional sums of squares. For **balanced** designs:

$$SS_{B|A} = SS_B \quad (10)$$

$$SS_{AB|A,B} = SS_{AB} \quad (11)$$

4.2 Unequal and non-proportional cell counts

The data in Table 12 are taken from Howell's (1997) Table 13.10. The dependent variable is the number of errors each driver made in a 30-minute session. Drivers from two different States participated in the study, and there was no reason to expect differences between the States.

Table 12: Illustration of loss of orthogonality in an unbalanced two-way design.

	Non-drinking	Drinking	Row Means
Michigan	13 15 14 16 12 $\bar{X}_{11} = 14$ $n_{11} = 5$	18 20 22 19 21 23 17 18 22 20 $\bar{X}_{12} = 20$ $n_{12} = 10$	$\bar{X}_{1\bullet} = 18$ $n_{1\bullet} = 15$
Arizona	13 15 18 14 10 12 16 17 15 10 14 $\bar{X}_{21} = 14$ $n_{21} = 11$	24 25 17 16 18 $\bar{X}_{22} = 20$ $n_{22} = 5$	$\bar{X}_{2\bullet} = 15.9$ $n_{2\bullet} = 16$
Column Means	$\bar{X}_{\bullet 1} = 14$ $n_{\bullet 1} = 16$	$\bar{X}_{\bullet 2} = 20$ $n_{\bullet 2} = 15$	

In this example, the cell means show pretty much what one would expect:

- In both States, drinking subjects make more errors on average than their non-drinking counterparts
- In both Conditions (drinking vs non-drinking), there is no difference between States

The column means are consistent with what we see in the cell means. But the row means are not. The row means show a difference between the States, with subjects from Michigan making more errors.

But note the confounding of Condition and State. One way to do this is to compute the expected cell counts using the marginal totals. For row 1 column 1, for example, $E = 15(16)/31 = 7.74$, which is not equal to the observed count. Note as well that Michigan, the State with the higher mean number of errors, has more Drinking than Non-drinking subjects (10 vs 5); and Arizona, the State with the lower mean, has fewer Drinking than Non-drinking subjects (5 vs 11). Thus, the apparent effect of State is really a Condition effect.

4.2.1 Unweighted means analysis

One solution to the problem of unequal sample sizes (and loss of orthogonality) is the so-called method of **unweighted means**. It is called “unweighted means” because the marginal means compared in the analysis are not weighted by sample size. Returning to Howell’s data, the marginal means shown in Table 12 (18 and 15.9) **are** weighted by sample size. The row means, for example, were computed as follows:

$$\bar{X}_{1\bullet} = \frac{n_{11}(\bar{X}_{11}) + n_{12}(\bar{X}_{12})}{n_{11} + n_{12}} = \frac{5(14) + 10(20)}{5 + 10} = \frac{270}{15} = 18.0 \quad (12)$$

$$\bar{X}_{2\bullet} = \frac{n_{21}(\bar{X}_{21}) + n_{22}(\bar{X}_{22})}{n_{21} + n_{22}} = \frac{11(14) + 5(20)}{11 + 5} = \frac{254}{16} = 15.9 \quad (13)$$

The **unweighted mean** for each row, on the other hand, is simply the **mean of the cell means** in that row, without any weighting by sample size—each cell mean in the row is weighted equally, regardless of sample size. For Howell’s data:

$$\text{Unweighted } \bar{X}_{1\bullet} = \frac{\bar{X}_{11} + \bar{X}_{12}}{2} = \frac{14 + 20}{2} = 17 \quad (14)$$

$$\text{Unweighted } \bar{X}_{2\bullet} = \frac{\bar{X}_{21} + \bar{X}_{22}}{2} = \frac{14 + 20}{2} = 17 \quad (15)$$

The unweighted row means allow a comparison between States that is not confounded with Condition (Non-drinking vs Drinking), and so are more appropriate.

In a nutshell, the method of unweighted means uses marginal means computed in exactly this way. It also uses an average sample size—actually the **harmonic mean** of the sample sizes—and proceeds as if each cell had that number of observations.⁶

Howell’s data set (Table 12) is made up, and is useful for demonstrating the difference between weighted and unweighted marginal means. But it is not ideal for demonstrating the method of unweighted means (because the 2 unweighted row means are identical, as are the 2 unweighted column means). Therefore, I will use the data in Kleinbaum et al’s (1988, 2nd ed.) Table 20-2 (Table 13 here) to demonstrate the method.

⁶ This applies to the calculation of the between-cells terms (SS_{cells} , SS_A , SS_B , and SS_{AB}) only. $SS_{\text{within-cells}}$ is calculated in the usual fashion.

Table 13: Satisfaction with medical care (Y) by patient worry and affective communication between patient and physician (from Kleinbaum et al, 1988, Table 20-2).

Affective Communication	Worry		Row Marginals
	Negative	Positive	
High	2 5 8 6 2 4 3 10 ($n_{11} = 8$) ($\bar{Y}_{11} = 5$) ($SS_{11} = 58.00$)	7 5 8 6 3 5 6 4 5 6 8 9 ($n_{12} = 12$) ($\bar{Y}_{12} = 6$) ($SS_{12} = 34.00$)	($n_{1\bullet} = 20$) ($\bar{Y}_{1\bullet} = 5.6$)
Medium	4 6 3 3 ($n_{21} = 4$) ($\bar{Y}_{21} = 4$) ($SS_{21} = 6.00$)	7 7 8 6 4 9 8 7 ($n_{22} = 8$) ($\bar{Y}_{22} = 7$) ($SS_{22} = 16.00$)	($n_{2\bullet} = 12$) ($\bar{Y}_{2\bullet} = 6$)
Low	8 7 5 9 9 10 8 6 8 10 ($n_{31} = 10$) ($\bar{Y}_{31} = 8$) ($SS_{31} = 24.00$)	5 8 6 6 9 7 7 8 ($n_{32} = 8$) ($\bar{Y}_{32} = 7$) ($SS_{32} = 12.00$)	($n_{3\bullet} = 18$) ($\bar{Y}_{3\bullet} = 7.56$)
Column Marginals	($n_{\bullet 1} = 22$) ($\bar{Y}_{\bullet 1} = 6.18$)	($n_{\bullet 2} = 28$) ($\bar{Y}_{\bullet 2} = 6.57$)	($n_{\bullet\bullet} = 50$) ($\bar{Y}_{\bullet\bullet} = 6.40$)

I attempted to perform an ANOVA on these data using the same method of partitioning SS_{Total} we used earlier for a balanced design (i.e., the method illustrated in Table 1). The ANOVA summary is shown in Table 14.

Table 14: Summary of two-way ANOVA for the data in Table 13.

Source of Variation	SS	df	MS	F	p
Between Cells	72.000	5			
A (Aff. Comm.)	38.756	2	19.378	5.684	0.006
B (Worry)	1.87	1	1.87	0.549	0.463
AxB	28.604	2	14.302	4.195	0.021
Within Cells (Error)	150.000	44	3.409		
Total	222.000	49			

I hope you noticed that there is a problem in Table 14: SS_A , SS_B , and SS_{AB} do not sum to $SS_{\text{Between-cells}}$.

As mentioned previously, the unweighted means solution to the problem of unequal sample sizes is to analyze the data as if all sample sizes were equal to the **harmonic mean** of the sample sizes. For a variable X, the harmonic mean is defined as follows:

$$\bar{X}_h = \frac{k}{\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + \dots + \frac{1}{X_k}} \quad (16)$$

The harmonic mean of the sample sizes in Table 13 is:

$$\bar{n}_h = \frac{6}{\frac{1}{8} + \frac{1}{12} + \frac{1}{4} + \frac{1}{8} + \frac{1}{10} + \frac{1}{8}} = 7.423 \quad (17)$$

So when calculating SS_{Cells} , SS_A , SS_B , and SS_{AB} , we will proceed as if each cell has a sample size of 7.423; and we will use the cell means and unweighted marginal means shown in Table 15. SS_{error} , on the other hand, is calculated in the usual fashion (i.e., by summing the SS within each cell), so we can take SS_{error} from Table 14: $SS_{\text{error}} = 150$, and $df_{\text{error}} = 44$.

Table 15: Cell means and unweighted marginal means for the Table 13 data.

Affective Communication	Worry		Unweighted Row Marginals
	Negative	Positive	
High	$\bar{n}_h = 7.423$ $\bar{Y}_{11} = 5$	$\bar{n}_h = 7.423$ $\bar{Y}_{12} = 6$	$n_{1\bullet(\text{Unwght})} = 14.846$ $\bar{Y}_{1\bullet(\text{Unwght})} = 5.5$
Medium	$\bar{n}_h = 7.423$ $\bar{Y}_{21} = 4$	$\bar{n}_h = 7.423$ $\bar{Y}_{22} = 7$	$n_{2\bullet(\text{Unwght})} = 14.846$ $\bar{Y}_{2\bullet(\text{Unwght})} = 5.5$
Low	$\bar{n}_h = 7.423$ $\bar{Y}_{31} = 8$	$\bar{n}_h = 7.423$ $\bar{Y}_{32} = 7$	$n_{3\bullet(\text{Unwght})} = 14.846$ $\bar{Y}_{3\bullet(\text{Unwght})} = 7.5$
Unweighted Column Marginals	$n_{\bullet 1(\text{Unwght})} = 22.269$ $\bar{Y}_{\bullet 1(\text{Unwght})} = 5.667$	$n_{\bullet 2(\text{Unwght})} = 22.269$ $\bar{Y}_{\bullet 2(\text{Unwght})} = 6.667$	$n_{\bullet\bullet(\text{Unwght})} = 44.538$ $\bar{Y}_{\bullet\bullet(\text{Unwght})} = 6.167$

The calculation of SS_{cells} , SS_A , SS_B , and SS_{AB} is shown in Table 16. The first two columns show the 3 levels of A (Affective Communication) and 2 levels of B (Worry) respectively. The column headed \overline{AB}_{ij} lists the 6 cell means, and the column headed \overline{n}_h shows the harmonic mean of the 6 sample sizes. The next 3 columns show the unweighted marginal means, i.e., the unweighted marginal A-means, B-means, and the unweighted grand mean (i.e., the mean of the 6 cell means).

Table 16: Calculation of the between-cells sums of squares for unweighted means analysis.

A	B	\overline{AB}_{ij}	\overline{n}_h	Unweighed Means			Squared deviations x \overline{n}_h			
				$\overline{A}_{i(\text{unwght})}$	$\overline{B}_{j(\text{unwght})}$	$\overline{Y}_{\bullet(\text{unwght})}$	Cells	A	B	AB
1	1	5	7.423	5.5	5.667	6.167	10.109	3.302	1.856	0.000
1	2	6	7.423	5.5	6.667	6.167	0.207	3.302	1.856	0.000
2	1	4	7.423	5.5	5.667	6.167	34.858	3.302	1.856	7.423
2	2	7	7.423	5.5	6.667	6.167	5.151	3.302	1.856	7.423
3	1	8	7.423	7.5	5.667	6.167	24.940	13.190	1.856	7.423
3	2	7	7.423	7.5	6.667	6.167	5.151	13.190	1.856	7.423
							80.416 SS_{cells}	39.589 SS_A	11.135 SS_B	29.692 SS_{AB}

The last 4 columns in Table 16 show the same kinds of squared deviations we saw before, but this time, each one is multiplied by the harmonic mean of the sample size: Recall that we are proceeding as if there are 7.423 observations per cell. For the shaded cells, these were calculated as follows. In the column headed “Cells”:

$$\text{Cell entry} = \overline{n}_h (\overline{AB}_{12} - \overline{Y}_{\bullet(\text{unwght})})^2 = 7.423(8 - 6.167)^2 = 24.940 \quad (18)$$

In the column headed “A”:

$$\text{Cell entry} = \overline{n}_h (\overline{A}_{3(\text{unwght})} - \overline{Y}_{\bullet(\text{unwght})})^2 = 7.423(7.5 - 6.167)^2 = 13.19 \quad (19)$$

In the column headed “B”:

$$\text{Cell entry} = \overline{n}_h (\overline{B}_{1(\text{unwght})} - \overline{Y}_{\bullet(\text{unwght})})^2 = 7.423(5.667 - 6.167)^2 = 1.856 \quad (20)$$

And in the column headed “AB”:

$$\begin{aligned} \text{Cell entry} &= \overline{n}_h [(\overline{AB}_{31} - \overline{Y}_{\bullet(\text{unwght})}) - (\overline{A}_{3(\text{unwght})} - \overline{Y}_{\bullet(\text{unwght})}) - (\overline{B}_{1(\text{unwght})} - \overline{Y}_{\bullet(\text{unwght})})]^2 \\ &= 7.423[(8 - 6.167) - (7.5 - 6.167) - (5.667 - 6.167)]^2 = 7.423 \end{aligned} \quad (21)$$

The results of this unweighted means analysis are summarized in Table 17. Note that I have not included lines for Between-Cells and Total this time. These lines are customarily omitted from the summary table for an unweighted means analysis, because SS_{error} plus SS_{cells} computed using unweighted means do not sum to SS_{Total} .

Table 17: Summary of *unweighted means* two-way ANOVA for the data in Table 13.

Source of Variation	SS	df	MS	F	p
A (Aff. Comm.)	39.589	2	19.378	5.806	0.006
B (Worry)	11.135	1	1.87	3.266	0.078
AxB	29.692	2	14.486	4.355	0.019
Within Cells (Error)	150.000	44	3.409		

4.3 Regression analyses and GLM's for unbalanced designs (continued)

As we saw in the chapter on ANCOVA, through the use of dummy variables, any ANOVA problem can be treated as a multiple regression problem. Many authors, including Kleinbaum and Howell, recommend the use of regression analysis for unbalanced two-way (or higher order) designs. Rather than embarking on a long discussion of this, I will carry on with the same dataset, and present some annotated output from SPSS that illustrates how this works. The syntax used to generate this output is found in the file *anova2.SPS*, which can be downloaded from my SPSS webpage:

<http://www.angelfire.com/wv/bwhomedir/spss.html/>

* ===== .

* Kleinbaum text suggests using regression to analyze these data;
we need to use a series of regression models to get the
required sums of squares.

* First, compute dummy variables .

```
compute a2 = (a=2).
compute a3 = (a=3).
compute b2 = (b=2).
compute a2b2 = a2*b2.
compute a3b2 = a3*b2.
exe.
formats a2 to a3b2 (f2.0).
```

```

REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS CI R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT y
  /METHOD= test (a2 a3)
           /test (a2 a3) (b2)
           /test (a2 a3) (b2) (a2b2 a3b2)
  /RESIDUALS HIST(ZRESID) .

```

Regression

Variables Entered/Removed^d

Model	Variables Entered	Variables Removed	Method
1	A3, A2	.	Test
2	B2	.	Test
3	A3B2, A2B2	.	Test

a. Dependent Variable: DV

Model Summary^d

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.418 ^a	.175	.139	1.97	.175	4.970	2	47	.011
2	.448 ^b	.201	.149	1.96	.026	1.520	1	46	.224
3	.569 ^c	.324	.248	1.85	.123	4.016	2	44	.025

a. Predictors: (Constant), A3, A2

b. Predictors: (Constant), A3, A2, B2

c. Predictors: (Constant), A3, A2, B2, A3B2, A2B2

d. Dependent Variable: DV

ANOVA^e

Model			Sum of Squares	df	Mean Square	F	Sig.	R Square Change
1	Subset Tests	A2, A3	38.756	2	19.378	4.970	.011 ^a	.175
	Regression		38.756	2	19.378	4.970	.011 ^b	
	Residual		183.244	47	3.899			
	Total		222.000	49				
2	Subset Tests	A2, A3	42.747	2	21.373	5.543	.007 ^a	.193
		B2	5.861	1	5.861	1.520	.224 ^a	.026
	Regression		44.617	3	14.872	3.857	.015 ^c	
	Residual		177.383	46	3.856			
	Total		222.000	49				
3	Subset Tests	A2, A3	63.273	2	31.636	9.280	.000 ^a	.285
		B2	4.800	1	4.800	1.408	.242 ^a	.022
		A2B2, A3B2	27.383	2	13.692	4.016	.025 ^a	.123
	Regression		72.000	5	14.400	4.224	.003 ^d	
	Residual		150.000	44	3.409			
	Total		222.000	49				

a. Tested against the full model.

b. Predictors in the Full Model: (Constant), A3, A2.

c. Predictors in the Full Model: (Constant), A3, A2, B2.

d. Predictors in the Full Model: (Constant), A3, A2, B2, A3B2, A2B2.

e. Dependent Variable: DV

- * $SS(A) = 38.756 = SS \text{ for subset } (A2, A3) \text{ in Model 1 .}$
- * $SS(B) = 5.861 = SS \text{ for subset } (B2) \text{ in Model 2 .}$
- * $SS(AB) = 27.383 = SS \text{ for subset } (A2B2, A3B2) \text{ in Model 3 .}$

* Note that $SS(A) + SS(B) + SS(AB) = 72 = SS(\text{cells})$.

* Kleinbaum et al use the F-tests from Models 1, 2, and 3 for the effects of A, B, and AB respectively; in other words, they do not use the same error term for all of these tests.

* Note that these F-tests are identical to the F Change tests presented in the Model Summary output table.

- * In SPSS, there is an easier way to get essentially the same results: i.e., using the GLM UNIANOVA procedure with **Type I SS**; I say essentially the same, because the SS for A, B, and AB will be exactly as we saw above; but **the F-ratios will be calculated using a common error term (the error term from Model 3 above)** .
- * NOTE: Using GLM UNIANOVA also negates the need to compute dummy variables--they are generated internally by UNIANOVA .

UNIANOVA

```

y BY a b /* Y is the DV; A and B the IV's */
/METHOD = SSTYPE(1) /* Type III sum of squares is the default */
/INTERCEPT = INCLUDE
/CRITERIA = ALPHA(.05)
/print etasq
/DESIGN = a b a*b.

```

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Affective Communication	1	A1- High	20
	2	A2- Medium	12
	3	A3- Low	18
Patient Worry	1	B1- Negative	22
	2	B2- Positive	28

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type I Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	72.000 ^a	5	14.400	4.224	.003	.324
Intercept	2048.000	1	2048.000	600.747	.000	.932
A	38.756	2	19.378	5.684	.006	.205
B	5.861	1	5.861	1.719	.197	.038
A * B	27.383	2	13.692	4.016	.025	.154
Error	150.000	44	3.409			
Total	2270.000	50				
Corrected Total	222.000	49				

a. R Squared = .324 (Adjusted R Squared = .248)

**** One problem with this approach is that the results depend on the order in which you enter the variables; note how things change when we enter B, then A, then AB .

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI R ANOVA CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT y
/METHOD=      test (b2)
               /test (b2) (a2 a3)
               /test (b2) (a2 a3) (a2b2 a3b2)
/RESIDUALS HIST(ZRESID) .
```

Regression

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	B2	.	Test
2	A2, A3	.	Test
3	A3B2, A2B2	.	Test

a. Dependent Variable: DV

Model Summary^d

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.092 ^a	.008	-.012	2.14	.008	.408	1	48	.526
2	.448 ^b	.201	.149	1.96	.193	5.543	2	46	.007
3	.569 ^c	.324	.248	1.85	.123	4.016	2	44	.025

a. Predictors: (Constant), B2

b. Predictors: (Constant), B2, A2, A3

c. Predictors: (Constant), B2, A2, A3, A3B2, A2B2

d. Dependent Variable: DV

ANOVA^e

Model			Sum of Squares	df	Mean Square	F	Sig.	R Square Change
1	Subset Tests	B2	1.870	1	1.870	.408	.526 ^a	.008
	Regression		1.870	1	1.870	.408	.526 ^b	
	Residual		220.130	48	4.586			
	Total		222.000	49				
2	Subset Tests	B2	5.861	1	5.861	1.520	.224 ^a	.026
		A2, A3	42.747	2	21.373	5.543	.007 ^a	.193
	Regression		44.617	3	14.872	3.857	.015 ^c	
	Residual		177.383	46	3.856			
	Total		222.000	49				
3	Subset Tests	B2	4.800	1	4.800	1.408	.242 ^a	.022
		A2, A3	63.273	2	31.636	9.280	.000 ^a	.285
		A2B2, A3B2	27.383	2	13.692	4.016	.025 ^a	.123
	Regression		72.000	5	14.400	4.224	.003 ^d	
	Residual		150.000	44	3.409			
	Total		222.000	49				

a. Tested against the full model.

b. Predictors in the Full Model: (Constant), B2.

c. Predictors in the Full Model: (Constant), B2, A2, A3.

d. Predictors in the Full Model: (Constant), B2, A2, A3, A3B2, A2B2.

e. Dependent Variable: DV

* SS(A) = 42.747 = SS for subset (A2,A3) in Model 2 .
 * SS(B) = 1.870 = SS for subset (B2) in Model 1 .
 * SS(AB) = 27.383 = SS for subset (A2B2,A3B2) in Model 3 .

* Now the same analysis using UNIANOVA and Type I SS .

UNIANOVA

```
y BY a b /* Y is the DV; A and B the IV's */
/METHOD = SSTYPE(1) /* Type III sum of squares is the default */
/INTERCEPT = INCLUDE
/CRITERIA = ALPHA(.05)
/print etasq
/DESIGN = b a a*b.
```

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Affective Communication	1	A1- High	20
	2	A2- Medium	12
	3	A3- Low	18
Patient Worry	1	B1- Negative	22
	2	B2- Positive	28

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type I Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	72.000 ^a	5	14.400	4.224	.003	.324
Intercept	2048.000	1	2048.000	600.747	.000	.932
B	1.870	1	1.870	.549	.463	.012
A	42.747	2	21.373	6.270	.004	.222
A * B	27.383	2	13.692	4.016	.025	.154
Error	150.000	44	3.409			
Total	2270.000	50				
Corrected Total	222.000	49				

a. R Squared = .324 (Adjusted R Squared = .248)

**** Type I SS is sometimes called the "**hierarchical**" sum of squares, or the "**sequential**" sum of squares; as you can see above, $SS(A) + SS(B) + SS(AB) = SS(\text{Corrected Model}) = SS(\text{cells})$; but the values of $SS(A)$ and $SS(B)$ depend on the order in which you enter A and B in the model; this is because the portion of variance that is explained by BOTH A and B will be included as part of $SS(A)$ when A is entered first, and as part of $SS(B)$ when B is entered first.

**** An alternative approach is to use a method that includes only that portion of the variance that is **UNIQUE** to each of A, B, and AB; this is what you get if you select **Type III** sums of squares (which is the default option in SPSS); because it is including only the unique portions of variance, the order in which variables are entered no longer matters .

Sums of Squares in GLM (from SPSS Help files)

For the model, you can choose a type of sums of squares. Type III is the most commonly used and is the default.

Type I. This method is also known as the hierarchical decomposition of the sum-of-squares method. Each term is adjusted for only the term that precedes it in the model. Type I sums of squares are commonly used for:

- A balanced ANOVA model in which any main effects are specified before any first-order interaction effects, any first-order interaction effects are specified before any second-order interaction effects, and so on.
- A polynomial regression model in which any lower-order terms are specified before any higher-order terms.
- A purely nested model in which the first-specified effect is nested within the second-specified effect, the second-specified effect is nested within the third, and so on. (This form of nesting can be specified only by using syntax.)

Type II. This method calculates the sums of squares of an effect in the model adjusted for all other "appropriate" effects. An appropriate effect is one that corresponds to all effects that do not contain the effect being examined. The Type II sum-of-squares method is commonly used for:

- A balanced ANOVA model.
- Any model that has main factor effects only.
- Any regression model.
- A purely nested design. (This form of nesting can be specified by using syntax.)

Type III. The default. This method calculates the sums of squares of an effect in the design as the sums of squares adjusted for any other effects that do not contain it and orthogonal to any effects (if any) that contain it. The Type III sums of squares have one major advantage in that they are invariant with respect to the cell frequencies as long as the general form of estimability

remains constant. Hence, this type of sums of squares is often considered useful for an unbalanced model with no missing cells. In a factorial design with no missing cells, this method is equivalent to the Yates' weighted-squares-of-means technique. The Type III sum-of-squares method is commonly used for:

- Any models listed in Type I and Type II.
- Any balanced or unbalanced model with no empty cells.

Type IV. This method is designed for a situation in which there are missing cells. For any effect F in the design, if F is not contained in any other effect, the Type IV = Type III = Type II. When F is contained in other effects, Type IV distributes the contrasts being made among the parameters in F to all higher-level effects equitably. The Type IV sum-of-squares method is commonly used for:

- Any models listed in Type I and Type II.
- Any balanced model or unbalanced model with empty cells.

* GLM using Type III SS .

```
UNIANOVA
  y BY a b                                /* Y is the DV; A and B the IV's */
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /CRITERIA = ALPHA(.05)
  /print etasq
  /DESIGN = a b a*b.
```

Univariate Analysis of Variance

Between-Subjects Factors			
		Value Label	N
Affective Communication	1	A1- High	20
	2	A2- Medium	12
	3	A3- Low	18
Patient Worry	1	B1- Negative	22
	2	B2- Positive	28

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	72.000 ^a	5	14.400	4.224	.003	.324
Intercept	1693.608	1	1693.608	496.792	.000	.919
A	44.577	2	22.289	6.538	.003	.229
B	11.134	1	11.134	3.266	.078	.069
A * B	27.383	2	13.692	4.016	.025	.154
Error	150.000	44	3.409			
Total	2270.000	50				
Corrected Total	222.000	49				

a. R Squared = .324 (Adjusted R Squared = .248)

*** Now change order of entry (on the /DESIGN line) to B, A, AB, and show that results do not change .**

UNIANOVA

```

y BY a b /* Y is the DV; A and B the IV's */
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/CRITERIA = ALPHA(.05)
/print etasq
/DESIGN = b a a*b.

```

Univariate Analysis of Variance**Between-Subjects Factors**

		Value Label	N
Affective	1	A1- High	20
Communication	2	A2- Medium	12
	3	A3- Low	18
Patient Worry	1	B1- Negative	22
	2	B2- Positive	28

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	72.000 ^a	5	14.400	4.224	.003	.324
Intercept	1693.608	1	1693.608	496.792	.000	.919
B	11.134	1	11.134	3.266	.078	.069
A	44.577	2	22.289	6.538	.003	.229
A * B	27.383	2	13.692	4.016	.025	.154
Error	150.000	44	3.409			
Total	2270.000	50				
Corrected Total	222.000	49				

a. R Squared = .324 (Adjusted R Squared = .248)

**** SS(A) = 44.577, SS(B) = 11.134, and SS(AB) = 27.383
for both of the foregoing analyses: The order in
which terms are entered does not matter when you
use Type III (or 'unique') SS .

**** Finally, note that **GLM UNIANOVA with Type III SS produces results that are approximated by the method of "unweighted means"** .

* ~~~~~ .

References

Howell, D.C. (1997). *Statistical methods for Psychology* (4th Ed.). Belmont, CA: Duxbury.

Kleinbaum, D.G., Kupper, L.L., Muler, K.E., & Nizam, A. (1998). *Applied regression analysis and other multivariable methods* (3rd Ed.). Pacific Grove, CA: Duxbury.
