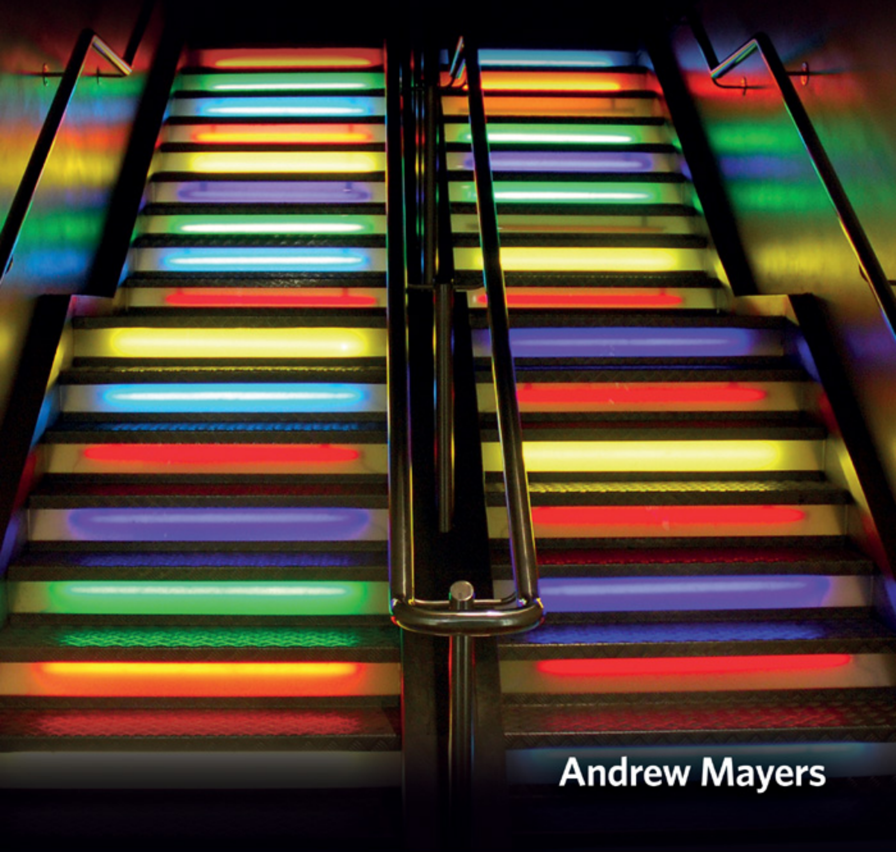
Introduction to Statistics and SPSS in Psychology



Andrew Mayers

ALWAYS LEARNING

PEARSON

Introduction to Statistics and SPSS in Psychology

Theory and rationale

Identifying differences

Essentially, all of the ANOVA methods have the same method in common: they assess **explained (systematic) variance** in relation to **unexplained (unsystematic) variance** (or 'error'). With an independent one-way ANOVA, the explained variance is calculated from the **group means** in comparison to the grand mean (the overall average dependent variable score from the entire sample, regardless of group). The unexplained (error) variance is similar to the standard error that we encountered in Chapter 4. In our example, if there is a significant difference in the number of hours spent in lectures between the university groups, the explained variance must be sufficiently larger than the unexplained variance. We examine this by partitioning the variance into the model sum of squares and residual sum of squares. We can see how this is calculated in Box 9.3. It is strongly recommended that you try to work through this manual example as it will help you understand how variance is partitioned and how this relates that to the statistical outcome in SPSS.

9.3 Calculating outcomes manually

Independent one-way ANOVA calculation

Table 9.1 No. of lecture hours attended per week

	Law (L)	Psychology (P)	Media (M)
	15	14	13
	10	13	12
	14	15	11
	15	14	11
	17	16	14
	13	15	11
	13	15	10
	19	18	9
	16	19	8
	16	13	10
Group mean	14.80	15.20	10.90
Standard deviation	2.486	1.990	1.792
Group variance	6.18	3.96	3.21
Grand mean	13.63	**Grand variance**	8.03

To illustrate how we can calculate the outcome of an independent one-way ANOVA, Table 9.1 presents some (fictitious) data based on the FUSS research question. The overall variance is measured by the total sum of squares. We need to 'partition' this into the model sum of squares and the residual sum of squares. The model sum of squares is found by squaring the difference between the group mean and the grand mean (the average score, regardless of group), multiplying that by the group size and summing the answer for each group (that is why it is called the sum of squares). The residual sum of squares is found from the variance of each group (which is the sum of the squared differences between each score and the group mean). These are expressed in terms of 'degrees of freedom' (the number of values that are 'free to vary' in the calculation, while everything else is held constant – see Chapter 6). This produces the model mean square and residual mean square. From these outcomes we find the F ratio, which can be compared with cut-off points to determine whether the between-group differences are significant.

You will find a Microsoft Excel spreadsheet associated with these calculations on the web page for this book.

Total sum of squares (SS_T):

$SS_T = S^2\text{grand}\,(N - 1) =$ grand variance $\times$ sample size (30) *minus* 1 $= 8.03 \times 29 = \mathbf{232.97}$

You will need to allow for slight 'rounding errors' due to decimal places throughout these calculations.

Grand variance: Deduct grand mean from each score, square it, repeat for all scores, add these up, divide by number of scores minus 1: $([15 - 13.63]^2 + [14 - 13.63]^2 + \ldots [10 - 13.63]^2) \div (30 - 1) = 8.03$

Model sum of squares (SS_M): The formula for model sum of squares: $SS_M = \sum n_k(\bar{x}_k - \bar{x}_{grand})^2$

Deduct grand mean from group mean, square it, multiply by no. of scores in group (10)

So $\mathbf{SS_M} = 10 \times (14.80 - 13.63)^2 + 10 \times (15.20 - 13.63)^2 + 10 \times (10.90 - 13.63)^2 = \mathbf{112.87}$

We have three groups, so degrees of freedom (*df*) for $SS_M(df_M) = 3 - 1 = 2$ (this is the numerator *df*)

Residual sum of squares (SS_R): Formula for the residual sum of squares: $(SS_R) = \sum s_k{}^2(n_k - 1)$

Multiply group variances by group size *minus* 1 ($10 - 1 = 9$):

$\mathbf{SS_R} = (6.18 \times 9) + (3.96 \times 9) + (3.21 \times 9) = \mathbf{120.15}$ (allow for decimal place rounding)

Group variance: Deduct group mean from each group score, square it, repeat for all group scores, add these up, divide by group size minus 1: e.g. for Gender: $([15 - 14.80]^2 + \ldots [16 - 14.80]^2) \div (10 - 1) = 6.18$

df for SS_R = sample size minus 1 $(30 - 1) - df_M(2)$: so $df_R = 29 - 2 = 27$ (this is the denominator *df*)

Mean squares: This is found by dividing model sum of squares and residual sum of squares by the relevant *df*:

Model mean square ($\mathbf{MS_M}$): $SS_M \div df_M = 112.87 \div 2 = \mathbf{56.43}$

Residual mean square ($\mathbf{MS_R}$): $SS_R \div df_R = 120.15 \div 27 = \mathbf{4.45}$

F ratio: This is calculated from model mean square divided by residual mean square:

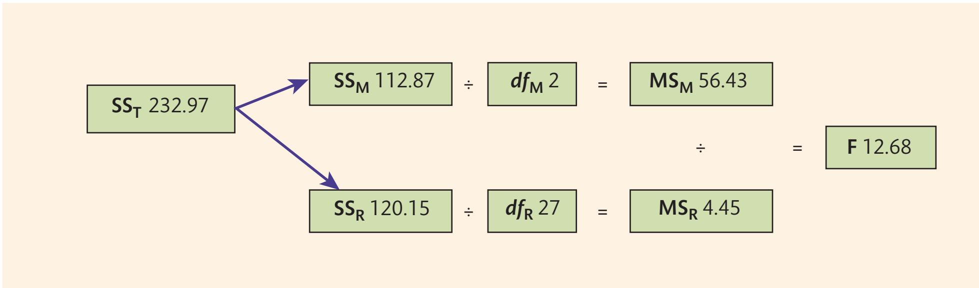
$$F = \frac{MS_M}{MS_R} = 56.43 \div 4.45 = \mathbf{12.68}$$

We compare that F ratio to **F-distribution tables** (Appendix 4), according to **numerator** and **denominator degrees of freedom**, at the agreed level of significance (usually $p = .05$). In our example, we have a numerator (2) and denominator (27) degrees of freedom ($df = 2, 27$) and we set significance as $p = .05$, so the F cut-off point (**critical value**) is 3.35. Our F ratio (12.68) is greater than that, so we have a significant difference in lecture hours across student groups.

We can also use Microsoft Excel to calculate the critical value of F and to provide the actual p value. You can see how to do that on the web page for this book. In our example, $p < .001$, You can also see how to perform the entire test in Excel.

Statistical significance: putting it into context

The outcome shown in Box 9.3 suggests that there is a significant difference in the number of hours spent in lectures across the university groups. When we perform the test in SPSS later, we can refer to the F ratio and significance outcome to confirm these findings. However, although it would be easy for you to simply look at that outcome, it would help your understanding of what the result means to view this in terms of explained vs. unexplained variance and relate that to the sum of squares and mean squares outcome from Box 9.3. The explained variance is that which relates to differences between the groups; it is illustrated by the model sum of squares (SS_M). The unexplained variance is the error, or that which is not related to the differences between the groups; it is shown by the residual sum of squares (SS_R). To take account of the number of groups and sample size, we need to express the model and residual sum of squares in terms of the degrees of freedom (*df*) – see Box 9.3. This produces the model mean square (MS_M) and residual mean square (MS_R). We divide MS_M by MS_R to get the F ratio. In our case, MS_M is larger than MS_R, which means that we have 'explained' most of the overall variance in lecture hours (the between-group differences will never be significant if there is more unexplained than explained variance). To assess whether the outcome is statistically significant, we compare the F ratio to F-distribution tables – if we exceed the relevant cut-off point we know that the observed between-group differences are significant.



9.4 Exercise

Partitioning sum of squares mini-exercise

Using the data set shown below, calculate the sum of squares, mean squares and F ratio. State whether there is a significant difference in quality of life scores between the town residents.

Table 9.2 Quality of life scores

Town A	Town B	Town C
36	94	26
18	56	67
81	20	83
30	53	94
36	56	69
76	56	67
31	76	77
29	24	66
55	69	38
16	59	42

Exercise outcome

To assess whether there is a significant difference in quality of life scores between the towns, you should have started by calculating the group means, group variance (you may have needed some help from Chapter 4 there) and grand mean. Those outcomes are shown in Table 9.3.

Table 9.3 Mean and variance outcomes for quality of life scores

	Quality of life scores		
	Town A	Town B	Town C
Group mean	40.80	56.30	62.90
Group variance	509.96	483.34	452.10
Grand mean 53.33			

Using the information in Table 9.3 and the guidelines in Box 9.3, you should have found the following:

Model sum of squares (SS_M): 2574.07
Residual sum of squares (SS_R): 13008.60
Model degrees of freedom (*df*): 3 groups *minus* 1 = 2
Residual *df*: sample size *minus* 1 *minus* model *df* = 27
Model mean square (MS_M): $SS_M \div 2 = 1287.03$
Residual mean square (MS_R): $SS_R \div 27 = 481.80$
F ratio: $MS_M \div MS_R = 2.67$

Using F-distribution tables (or the Excel function shown in the associated spreadsheet) we see that the cut-off point for F when $df = 2, 27$ is 3.35. The calculated F ratio (2.67) is *less* than that, so there is *not* a significant difference in quality of life scores between the towns (exact p value = .087).

Finding the source of difference

The F ratio will indicate whether we have a significant difference in mean scores between the groups – we call this the 'omnibus' ANOVA. However, *if* we have three or more groups, this tells us only that most of the overall variance is explained by differences across those groups; it will not indicate *where* those differences are. If we have a significant difference across two groups, we can simply compare the mean scores to tell where the differences are. If we have significant difference across three or more groups, we need more information.

When we explored the FUSS data just now, we saw that there was a significant difference in the number of lecture hours attended between three university course groups. The mean data suggest that media students attended for fewer hours than psychology and law students (who appear to be similar in the number of hours attending lectures). However, we need to do more than make visual comparisons: we must explore the differences statistically. To do this, we need to perform additional analyses, using either **planned contrasts** or ***post hoc* tests**. We use planned contrasts if we have predicted a *specific* outcome about differences between the groups; otherwise we must use *post hoc* tests.

Planned contrasts, *post hoc* tests and multiple comparisons

The key difference between planned contrasts and *post hoc* tests rests on the way in which they account for multiple comparisons. Usually, we base significance testing on the probability that there is less than a 5% likelihood that the observed differences occurred by chance (see Chapter 4). The more tests we run, the greater the possibility we have of finding a significant outcome simply by chance factors alone (known as the, '**familywise error**'). To account for that possibility, we may need to adjust the level at which we declare statistical significance. For every additional test that we run, we may need to divide the significance cut-off according to how many of those tests we undertake. As we will see throughout the next few sections, the way in which we handle **multiple comparisons** will depend on what type of test we employ to examine those differences.

But remember, neither planned contrasts nor *post hoc* tests are needed if, a) there are only two groups, or b) the overall outcome is not significant.

Planned contrasts

Planned contrasts can be used only if *specific* (one-tailed) predictions have been made about outcomes between the groups. For example, using the FUSS research example, a specific prediction might state that psychology students will spend more time in lectures than the other two groups, while there will be no difference in attendance between law students and media students. In that scenario we can use planned contrasts. If FUSS predicts only that there will be difference between the groups, planned contrasts cannot be employed; *post hoc* tests must be undertaken instead. The type of planned contrast used depends on whether one of the groups being analysed represents a **control group** (this is the group to which outcome across **experimental groups** are compared). If we do have a control group, we explore between-group differences using 'orthogonal' planned contrasts; if there is no control group, a non-orthogonal test must be employed. Outcomes from orthogonal planned contrasts can be reported without adjusting for multiple comparisons, otherwise (with non-orthogonal analyses) we must adjust the significance cut-off point by the number of additional tests that we run.

Orthogonal planned contrasts

When we have a control group, we need to know how the other experimental groups compare with that control group (we call that Contrast 1). We also need to know how the experimental groups compare with each other. If we have two experimental groups we explore that in one further test (Contrast 2); if we have more than two additional groups we would need several extra contrasts. To enter the data into SPSS we need to assign values to the contrasts. For every positive value entered we need corresponding negative values, so that the overall values sum to zero. The weight of the values will depend on how many groups there are. Box 9.5 shows how we would allocate values to some example planned contrasts.

9.5 Nuts and bolts
Value allocation in orthogonal planned contrast

Say we examine one control group (Group 1) against two experimental groups (Groups 2 and 3). In this scenario, we allocate −2 for the control group (because there are two experimental groups to compare with). To balance back to 0, we allocate +1 to each of the experimental groups. Then we compare the experimental groups with each other. The control group is now redundant, so is given the value 0. The remaining experimental groups receive values of −1 and +1 respectively (so balance to 0 once again).

Value allocation: one control group and two experimental groups

	Compared groups	Redundant group	Contrast values		
Contrast 1	Control vs. Groups 1 and 2	None	−2	+1	+1
Contrast 2	2 vs. 3	Control	0	−1	+1

If we have three experimental groups (Groups 2–4) to compare to the single control group (Group 1), we have additional comparisons to undertake. We allocate −3 to the control group and +1 to each of the experimental groups. We then perform three additional contrasts allocating −1 and +1 to each pair of groups that we need to compare, leaving the control group and remaining experimental group redundant.

Value allocation: one control group and three experimental groups

	Compared groups	Redundant groups	Contrast values			
Contrast 1	Control vs. Groups 1, 2, and 3	None	−3	+1	+1	+1
Contrast 2	2 vs. 3	Control, 4	0	−1	+1	0
Contrast 3	2 vs. 4	Control, 3	0	−1	0	+1
Contrast 4	3 vs. 4	Control, 2	0	0	−1	+1

Non-orthogonal planned contrasts

If we have made a specific prediction about the outcomes between the groups, but none of the groups is being used as a control group, we must use a non-orthogonal planned contrast. This means that the contrasts are no longer independent. The method for calculating values in a non-orthogonal planned contrast is shown in Box 9.6. Effectively, it is the same as the three-group example in Box 9.5, but without the control condition. Once we have created the contrast values, and calculated outcomes across the contrasts, we must adjust the significance level to account for multiple comparisons. If we have three groups, we have three pairwise comparisons, so we divide the significance cut-off point by three. Assuming overall significance is set at $p < .05$, contrast outcomes will be significant only where $p < .016$ ($0.05 \div 3 = 0.016$). If we have four groups, there will be six contrasts; significance occurs only where $p < .008$, and so on. It could be argued that the additional work undertaken to run non-orthogonal contrasts is not worth the effort. We gain nothing in terms of improving chances of finding significant outcomes. As we will see soon, *post hoc* tests (in SPSS) are performed with little fuss and (most) automatically adjust for multiple comparisons.

9.6 Nuts and bolts

Value allocation in non-orthogonal planned contrast

In this scenario we examine pairs of groups, such as Group 1 vs. Group 2 (Contrast 1), 1 vs. 3 (Contrast 2), and 2 vs. 3 (Contrast 3). We must assign values for each of the groups in each contrast, using −1 and +1 for the comparison pair and 0 for the redundant group. We will see how to enter this into SPSS later.

	Compared groups	Redundant group	Contrast values		
Contrast 1	1 vs. 2	3	−1	+1	0
Contrast 2	1 vs. 3	2	−1	0	+1
Contrast 3	2 vs. 3	1	0	−1	+1