

Research Methods II: Autumn Term 2001

Using SPSS: One-way Independent-Measures ANOVA:

In this handout, I will show you how to do the following:

- (a) perform a one-way independent-measures ANOVA;
- (b) obtain descriptive statistics for the data on which the statistics are performed;
- (c) perform post-hoc tests to determine exactly which differences between groups are giving rise to the overall statistical significance shown by the ANOVA.

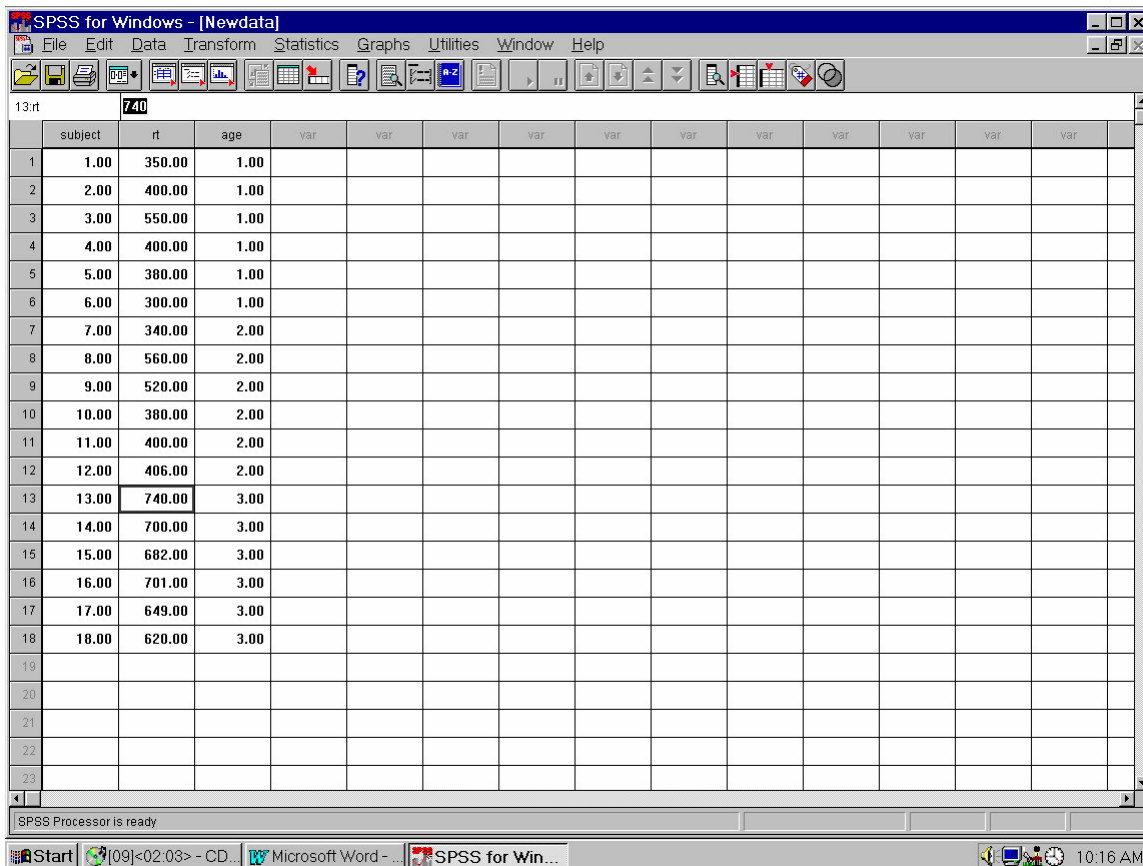
Step-by-step one-way independent-measures ANOVA:

1. Enter the data:

You need to arrange the data in the following way. One column contains the dependent variable, and another column contains code-numbers which tell SPSS which group the subject belongs to (i.e. the independent variable). Note that you can give your columns meaningful names by double clicking on the top of the column (double-click where the columns have already been given default names like "var00001").

A concrete example might make this clear. Imagine we are interested in age-differences in reaction-time. We have three groups: "young" people, "middle-aged" people and "old" people. Let's call the "young" people group "1"; the middle-aged people group "2"; and the "old" people group "3". Suppose we have six subjects in each group¹.

Here's what our SPSS data-window might look like.



The screenshot shows the SPSS data editor window with the following data:

	subject	rt	age	var	var	var	var	var	var	var	var	var	var	var
1	1.00	350.00	1.00											
2	2.00	400.00	1.00											
3	3.00	550.00	1.00											
4	4.00	400.00	1.00											
5	5.00	380.00	1.00											
6	6.00	300.00	1.00											
7	7.00	340.00	2.00											
8	8.00	560.00	2.00											
9	9.00	520.00	2.00											
10	10.00	380.00	2.00											
11	11.00	400.00	2.00											
12	12.00	406.00	2.00											
13	13.00	740.00	3.00											
14	14.00	700.00	3.00											
15	15.00	682.00	3.00											
16	16.00	701.00	3.00											
17	17.00	649.00	3.00											
18	18.00	620.00	3.00											
19														
20														
21														
22														
23														

I have three columns, labelled "subject", "rt" and "age". (The "subject" column isn't necessary for the ANOVA, I've included it here in order to emphasise how each row represents one person's data.).

In the column of the SPSS data-window labelled "rt", I have entered the reaction-times for the subjects. In the column labelled "age", I've entered the number 1, 2 or 3, to designate which group the subject belongs to.

(Merely for the sake of clarity, it's best to enter all of the scores for one group, and then all of the scores for the next group, and so on, as shown here. However, there's nothing to stop you entering the scores all mixed up in the column, since SPSS will use the codes in the "age" column to sort out which scores belong to which group).

It is always a good idea to save data you have entered. To do this, put a floppy disk in a drive. Click on "File" and then "Save".

2. Perform the ANOVA:

There's more than one way to perform a one-way independent-measures ANOVA with SPSS. The method described here has the advantage of enabling you to get descriptive statistics for the groups (means, etc.) without having to ask for them separately. It also allows you to perform post-hoc tests on the data, to pinpoint the source of the significant differences shown by the ANOVA (assuming there are any!).

(a): Click on "Analyze", then "Compare Means", and then "One-way ANOVA":

subject	rt	age	var	var	var	var	var	var	var	var	var	var	var
1	1.00												
2	2.00												
3	3.00												
4	4.00												
5	5.00												
6	6.00												
7	7.00												
8	8.00												
9	9.00												
10	10.00												
11	11.00	400.00	2.00										
12	12.00	406.00	2.00										
13	13.00	740.00	3.00										
14	14.00	700.00	3.00										
15	15.00	682.00	3.00										
16	16.00	701.00	3.00										
17	17.00	649.00	3.00										
18	18.00	620.00	3.00										
19													
20													
21													
22													
23													

(b) Highlight the dependent variable on which you want to perform the ANOVA. In this case, we've only got one to choose from: "rt", the column containing the reaction-time scores. Click on the upper arrow-button to move "rt" into the box entitled "dependent list". Basically, we have now told SPSS where our data, our scores, are located.

(c) We now need to tell SPSS about the independent variable in our experiment. Our independent variable is "age". Highlight "age" and use the lower arrow-button to move it into the box labelled "factor".

(d) We could now click on "OK", and we would get our ANOVA. However, we want descriptive statistics for each group, and we want post-hoc tests to see which of the three groups are significantly different from each other.

To get the descriptive statistics, click on "Options...". A dialog box will appear. Click on "Descriptive", to put a little tick in the box next to it; and then click on "Continue". You will then be returned to the previous dialog box.

To get the post-hoc tests, click on "Post Hoc...". A dialog box will appear, giving a whole range of different post-hoc tests - different ways of making comparisons between pairs of groups in your study. You can

click on any of these to have SPSS work them out for you: at the moment, just click on "S-N-K" (short for Student Newman-Keuls). A tick should appear next to it, to show that it has been selected. Click on "continue" to close the dialog box and return to the previous one.

(f) Now click on "OK", and SPSS will perform the calculations. The window will swap from the data-window to the output window, and you will see the results of the analyses

3. Testing the assumptions of ANOVA

ANOVA assumes that the population distribution for each of your groups is normal. To check this you should plot a histogram for the data in each group. On the SPSS controls at the top of the screen, click on "Analyze"; this will produce a menu. On this menu, click on "Descriptive Statistics". Another menu will appear. On this one, click on "Explore". A dialog box will now appear. Enter your dependent variable in the dependent list and your independent variable in the factor list. Click on "Plots..." and click against "Histogram". Click on "Continue" and then "OK".

Given a sample of say only 12 people in a group, you can't expect the distribution to look very normal. In any case, ANOVA is reasonably robust to some violation of normality. What you are checking for is that the data are basically bunched together (not two separate groups, and also no outliers), reasonably symmetrically (not one tail much longer than the other). Particularly be on the lookout for outliers – an observation far removed from the main body of the data. These may be recording errors or they may be real, but they are disproportionately influential on the mean and standard deviation, and so should be removed from the analysis (just make sure you report that you have removed any data).

ANOVA also assumes that the variances are roughly equal in each group (the posh name for this assumption is "homogeneity of variance"). ANOVA is reasonably robust to violations of this assumption. If your largest variance is no more than five times your smallest then the assumption is satisfied adequately (assuming equal numbers of subjects in each group). The standard deviation for each group is given by the histogram; square the standard deviation to obtain the variance.

If either assumption is violated then you would normally need to do something about it (i.e. use Kruskal-Wallis). However, for the purposes of this module, you will use ANOVA anyway in order to learn the technique. But in your write-up you must indicate that you are aware of whether the assumptions were satisfied or not.

Note that Explore also automatically produces two other types of plots you may not be familiar with but are useful: stem-and-leaf plots and box-and-whisker plots (boxplots). Use "help" on SPSS to see an explanation of what a boxplot is - ask us in class about anything you don't understand. You do not need to include boxplots or stem-and-leaf plots in your write-up, but they can help you get a feel for your data.

A stem-and-leaf plot is an excellent way of looking at the distribution of your data without losing all information about the values of individual cases. In this example, the "stems" are multiples of 100 ms, and the "leaves" give the tens of milliseconds, with one leaf for each subject. For example for group one, in the plot:

```
3 . 0
3 . 58
4 . 00
(extremes >= 550)
```

the stems are the numbers before the dot, and the leaves are the numbers after the dot. In this case, all values between 300 and the 340's will be represented by the first stem, all values between the 350's and the 390's will be represented by the second stem, all values between the 400s and the 440s will be represented by the second stem, and so on. The single "0" leaf after the first stem indicates there was one subject with a value of 300 (strictly, in this case: with a value between 300 and 309), the "5" and the "8" in the second stem indicates that there was one subject with a value in the 350s and one with a value in the 380s. Finally, there were two subjects with a value between 400 and 409, and one extreme score (an extreme score is a type of outlier). The stem-and-leaf is simply an alternative to a histogram for looking at the shape of your data: You can use one or both plots to make your decision about normality and the presence of outliers.

4. Interpreting the output:

The bits in normal type below are the output you would see; the bits in bold italics, enclosed within square brackets, are my explanations of what the output means.

Oneway

[Below are the descriptive statistics for the three groups in our experiment. The meaning of "N", "mean" and "standard deviation" should all be self-evident to you, and these are the important bits to

concentrate on in the output below. "Standard error" and "95% confidence intervals for the mean" might puzzle you, although we have covered them in your first-year course.

Many procedures in SPSS provide estimates of "standard error". Imagine you had repeated your experiment an indefinite number of times, and calculated the mean each time; the standard error is an estimate of the standard deviation of that distribution of means. Basically, the larger it is, the more likely it is that taking different samples of subjects would produce different values for the mean. Remember that the sample mean is a kind of "spot-estimate" of the true mean of the population from which the sample was taken. In other words, if I did the experiment again, with different subjects, how likely would I be to get similar group means? If the standard error values are small, the answer is: quite likely to do so. If the standard error values are large, the answer would be: not very likely! So this gives you a rough-and-ready idea of how confident you can be about the values of the group means.

The 95% confidence intervals do a similar job in a way that may be easier to grasp intuitively. The 95% confidence interval is, informally speaking, the interval within which you can be 95% sure the true population mean lies. (More precisely, the interval is defined like this: If you repeated your experiment an indefinite number of times, the 95% confidence interval would contain the true population mean 95% of the time. But let's stick with the informal definition for now.) Take the mean for group 1: it's 397. This looks pretty precise! However, you know it is only a sample estimate, and the population mean may be a little bit bigger or smaller. But how much bigger or smaller? It's likely to be close to the sample mean; it's far less likely to be 20000, for example. In fact, we are 95% sure it is not bigger than 485, nor smaller than 308. But we can't with any confidence rule out any of the numbers in between.

Simply looking at the means, we might think group 1 and group 2 are different from each other; however, looking at the confidence interval values makes us more aware of how much potential overlap there might be between these groups. For example, in this case it is quite likely that, with different groups of subjects, the mean for group 2 might be the same or even lower than the mean for group 1! In contrast, it is highly unlikely that, with different groups of subjects to those used here, there would be any overlap between group 3 and either of the other two groups.]

Descriptives

RT

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00	6	396.6667	84.0635	34.3188	308.4475	484.8859	300.00	550.00
2.00	6	434.3333	85.9806	35.1014	344.1022	524.5645	340.00	560.00
3.00	6	682.0000	42.3840	17.3032	637.5208	726.4792	620.00	740.00
Total	18	504.3333	147.4537	34.7552	431.0063	577.6603	300.00	740.00

[This is called an ANOVA table. The Mean Square for Between Groups is often called "Mean Square Treatment", or *MSt*. The Mean Square for Within Groups is often called "Mean Square Error", or *Mse*. Note that each mean square is the relevant Sum of Squares divided by its degrees of freedom (*D.F.*). The *F*-ratio is *MSt/Mse*. *F*-prob. is the significance level for the *F*-ratio. SPSS only shows this to three digits - so the fact that it says .000 here means that the *F*-ratio is so highly significant that the *p* value is less than .0005 and hence too small to show, NOT that it is actually zero!]

ANOVA

RT

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	288345.333	2	144172.667	26.607	.000
Within Groups	81278.667	15	5418.578		
Total	369624.000	17			

[Here are the results of the post-hoc tests that we asked for - in this case, we asked for Newman-Keuls tests. SPSS presents the groups in sets: groups shown within a set are not significantly different from each other. In this example, groups 1 and 2 are not significantly different, and so they are lumped together. Group 3 is different from the others, and thus is placed in a subset by itself.]

----Post Hoc Tests

Homogeneous Subsets

RT

Student-Newman-Keuls

	N	Subset for alpha = .05	
AGE		1	2
1.00	6	396.6667	
2.00	6	434.3333	
3.00	6		682.0000
Sig.		.389	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 6.000

[To save the Output, click on "File" and then "Save".]

The conclusions that would be drawn from this output:

What would we conclude from these results? The overall (or "omnibus") ANOVA is significant, which tells us that there is *some* difference, somewhere, between our three groups, in terms of reaction time. Inspection of the means suggests that the "old" group are slower than the two younger groups, and that there is little difference between the two younger groups. The Newman-Keuls test confirms this impression statistically: the "old" subjects are indeed significantly slower than the "young" and "middle-aged" subjects. The "middle-aged" subjects are not significantly slower than the "young" subjects.

An Example Results Section

Table 1 displays the means and standard deviations of the RT scores (in milliseconds) for the different groups.

Table 1.

Group:	Mean:	Standard Deviation:
Young	397	84
Middle-aged	434	86
Old	682	42

Note: the use of meaningful group names and not numbers to label the groups.

Inspection of the histograms for each group revealed that the data were roughly normally distributed. The variances were also roughly equal in each group.

A one-way between-subjects (age(young vs middle-aged vs old)) ANOVA on RT scores was significant, $F(2,15) = 26.61$, $p < .0005$.

Note: the name of the independent variable ('age') is given in brackets, followed by the names of its levels in further brackets. The first number after the F is the degrees of freedom for treatment (i.e. between group variability) and the second number is the degrees of freedom for error (i.e. within group variability). Both of these are obtained from the analysis of variance summary table. F is given to two decimal places, and p would be given to two significant figures, e.g. $p = .0016$.

A Student Newman-Keuls indicated the following pattern:

Young Middle-aged Old

Note: the convention used here is to write out the names of the groups in order of their means, not necessarily in order of your code numbers 1,2,3. You then underline those groups that were non-significantly different from each other. For example:

Young Middle-aged Old

would mean that the young differed from the old, but there were no other significant differences. That is, you would have to run more subjects to know whether the middle-aged are more like the young or the old, or neither.

You need not use this visual convention if you think it is clearer to state the pattern in plain English; e.g. "A Student Newman-Keuls indicated that each group was significantly different from all others". Note also that the Student Newman-Keuls is only reported if the one-way ANOVA was significant.

This is all that appears in the results section. Raw data, histograms, and the ANOVA summary table should appear in the Appendix.