Chapter 14.
One-Way Analysis of Variance for Independent Samples
Part 1

As most of the logic and procedure for this simplest version of the analysis of variance was developed in Chapter 13, the main portion of the present chapter can be fairly brief and to the point. Part 1 covers the essentials of one-way ANOVA for independent samples. Part 2 touches upon two matters that will take you a bit beyond the bare essentials.

This version of ANOVA applies to the case where you have one independent variable and three or more independent samples of subjects, each sample measured at a different level of the variable. To avoid having to repeat the cumbersome phrase "three or more," we will henceforth refer to the number of independent samples (which is the same as the number of levels of the independent variable) as k. Thus, with three groups of subjects and three levels of the independent variable, $k=3$; with four groups and four levels, $k=4$; and so on. We will illustrate the procedure with an example involving $k=4$.

One of the many complications in the lives of persons who have Alzheimer's disease is that they tend to suffer from frequent periods of intense agitation. Some investigators have theorized that this agitation stems from old aversive conditionings— the conditioned fears and anxieties that a person accumulates over the course of his or her life—with which the Alzheimer's patient is no longer able to cope, on account of severely diminished cognitive capacities.
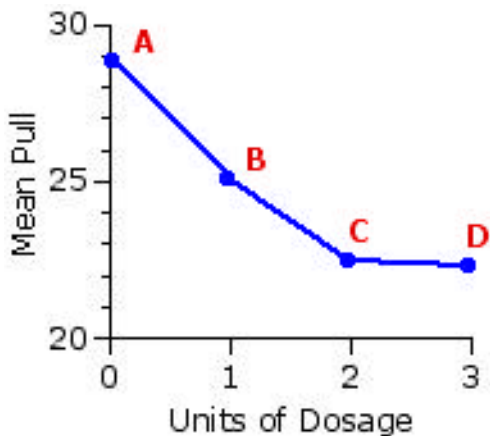
Against this background, a team of investigators has developed an experimental medication that they believe will substantially decrease the effects of old aversive conditionings. As a preliminary test of the medication, they trained 20 laboratory rats, by standard aversive conditioning procedures, to flee from a certain visual stimulus. These 20 subjects were then randomly and independently sorted into $k=4$ groups—A, B, C, and D— of 5 subjects each. In the subsequent experimental procedure, the members of each group received via sub-cutaneous injection one or another of four dosage levels of the medication. The members of group A, serving as a control group, received only an inert placebo containing zero units of the medication, while the members of groups B, C, and D received 1 unit, 2 units, and 3 units of the medication, respectively. Five minutes after receiving its injection, each subject was then presented with the aversively conditioned stimulus, and a measure was taken of how hard the subject pulled against a restraining harness in trying to move away from the stimulus. The smaller the pull, the smaller the degree of agitation presumed to be occasioned by the aversive stimulus.

The following table shows the measure of "pull" for each of the 5 subjects in each of the four groups. Also shown are the means of the groups along with M_T, the mean of the total array. This table is only our first pass at the data by way of an overview. A fuller listing of summary statistics will be given in a moment.

A 0 units	B 1 unit	C 2 units	D 3 units	Total Array
27.0 26.2 28.8 33.5 28.8	22.8 23.1 27.7 27.6 24.0	21.9 23.4 20.1 27.8 19.3	23.5 19.6 23.7 20.8 23.9	All groups combined.
$M_a = 28.86$	$M_b = 25.04$	$M_c = 22.50$	$M_d = 22.30$	$M_T = 24.68$

Figure 14.1 provides the same overview in graphical form. Whichever way you look at it, there clearly are differences among the means of the four groups, and these are consistent with what the investigators would have expected if the medication has the effect they suppose it to have. The greatest mean pull, hence presumably the greatest level of agitation, was found with the group that received only the placebo. For the group that received 1 unit of the medication, the mean pull was smaller; and for the groups that received 2 units and 3 units, it was smaller still.

Figure 14.1. Dosage Level and Mean Pull

But of course, here as elsewhere, there is always the possibility that the observed "effect" results from nothing more than mere random variability. And until that possibility is rationally assessed, no conclusions can be drawn, one way or the other. As indicated in Chapter 13, the one-way analysis of variance for independent samples performs that assessment by taking the ratio of two quantities

$$ F = \frac{MS_{bg}}{MS_{wg}} = \frac{\text{a measure of the aggregate differences among the means of the } k \text{ groups}}{\text{a measure of the amount of random variability that exists inside the } k \text{ groups}} $$

which is then referred to the appropriate sampling distribution of **F**, as defined by df_{bg} and df_{wg}. For purposes of practical computation, the first step is to calculate the values of $\sum X_i$ and $\sum X^2{}_i$ for each of the **k** groups and for all **k** groups combined. These, in conjunction with the relevant values of N (N_a, N_b, etc.), will then permit the calculation of all other quantities required for the analysis. The following table shows the full array of the preliminary summary statistics.

Units of Dosage				
0	1	2	3	Total Array
27.0 26.2 28.8 33.5 28.8	22.8 23.1 27.7 27.6 24.0	21.9 23.4 20.1 27.8 19.3	23.5 19.6 23.7 20.8 23.9	All groups combined.
$N_a = 5$	$N_b = 5$	$N_c = 5$	$N_d = 5$	$N_T = 20$
$\sum X_{ai} = 144.30$	$\sum X_{bi} = 125.20$	$\sum X_{ci} = 112.50$	$\sum X_{di} = 111.50$	$\sum X_{Ti} = 493.50$
$\sum X^2{}_{ai} =$ 4196.57	$\sum X^2{}_{bi} =$ 3158.50	$\sum X^2{}_{ci} =$ 2576.51	$\sum X^2{}_{di} =$ 2501.95	$\sum X^2{}_{Ti} =$ 12433.53
$M_a = 28.86$	$M_b = 25.04$	$M_c = 22.50$	$M_d = 22.30$	$M_T = 24.68$
$SS_a = 32.07$	$SS_b = 23.49$	$SS_c = 45.26$	$SS_d = 15.50$	$SS_T = 256.42$
(If it is not clear where the five values of **SS** are coming from, click <u>here</u> for an account of the computational details.)				

As indicated in the table, the raw measure of variability within the entire array of data, with all **k** groups combined, is $SS_T = 256.42$. This, in turn, is composed of two complementary components: SS_{bg}, which is the raw measure of the aggregate differences among the means of the **k** groups; and SS_{wg}, which is the raw measure of the variability that exists inside the **k** groups. With **k**= 4, the latter measure comes out as

$$SS_{wg} = SS_a + SS_b + SS_c + SS_d$$

$$= 32.07 + 23.49 + 45.26 + 15.50$$

$$= 116.32$$

Assuming that all number-crunching up to this point has been performed without error, you could then calculate the remaining component simply as

$$SS_{bg} = SS_T - SS_{wg}$$

$$= 256.42 - 116.32$$

$$= 140.10$$

However, it is always good practice to check the accuracy of one's calculations by calculating SS_{bg} also from scratch. As you saw in Chapter 13, the weighted squared deviate for each of the k group means can be calculated as

$$N_g(M_g - M_T)^2$$

where M_g is the mean of a particular group and N_g is the number of values of X_i on which that mean is based. Thus, for each of the four groups

A: $5(28.86 - 24.68)^2 = 87.36$

B: $5(25.04 - 24.68)^2 = 0.65$

C: $5(22.50 - 24.68)^2 = 23.76$ 140.09 versus 140.10.

D: $5(22.30 - 24.68)^2 = 28.32$ Close enough. The slight difference between the two derives from rounding errors in the present calculation. (Each mean value starts out rounded to two decimal places.)

$$SS_{bg} = 140.09$$

An alternative method for calculating SS_{bg} from scratch would be by way of the following computational formula. Although this rather cumbersome-looking device will not provide as clear an idea of the structure of SS_{bg}, it requires fewer computational steps and is also less susceptible to rounding errors.

$$SS_{bg} = \frac{(\sum X_{ai})^2}{N_a} + \frac{(\sum X_{bi})^2}{N_b} + \frac{(\sum X_{ci})^2}{N_c} + \frac{(\sum X_{di})^2}{N_d} - \frac{(\sum X_T)^2}{N_T}$$

$$= \frac{(144.3)^2}{5} + \frac{(125.2)^2}{5} + \frac{(112.5)^2}{5} + \frac{(111.5)^2}{5} - \frac{(493.5)^2}{20}$$

$$= 140.10$$

Here then, in summary, are our two component values of raw variability:

$$SS_{bg} = 140.10 \quad \text{and} \quad SS_{wg} = 116.32$$

The next step is to refine them into measures of **MS** through dividing each by its corresponding number of degrees of freedom. For the between- groups measure, degrees of freedom is the same as outlined in Chapter 13, except now we replace the phrase "number of groups" with the simple designation **k**.

$$df_{bg} = k - 1$$
$$= 4 - 1 = 3$$

Similarly for the within-groups measure, except now the groups are not A|B|C but A|B|C|D.

$$df_{wg} = (N_a - 1) + (N_b - 1) + (N_c - 1) + (N_d - 1)$$
$$= (5 - 1) + (5 - 1) + (5 - 1) + (5 - 1) = 16$$

However, once you clearly understand the structure of df_{wg}, its numerical value can be reached more simply through the algebraically equivalent formula

$$df_{wg} = N_T - k$$
$$= 20 - 4 = 16$$

Note that df_T, the number of degrees of freedom for the entire array of data, is
$$df_T = N_T - 1$$
$$= 20 - 1 = 19$$
and that $df_{bg} + df_{wg} = df_T$.

With the values of **SS** and **df** for between-groups and within-groups, we can then calculate

$$MS_{bg} = \frac{SS_{bg}}{df_{bg}} = \frac{140.10}{3} = 46.70$$
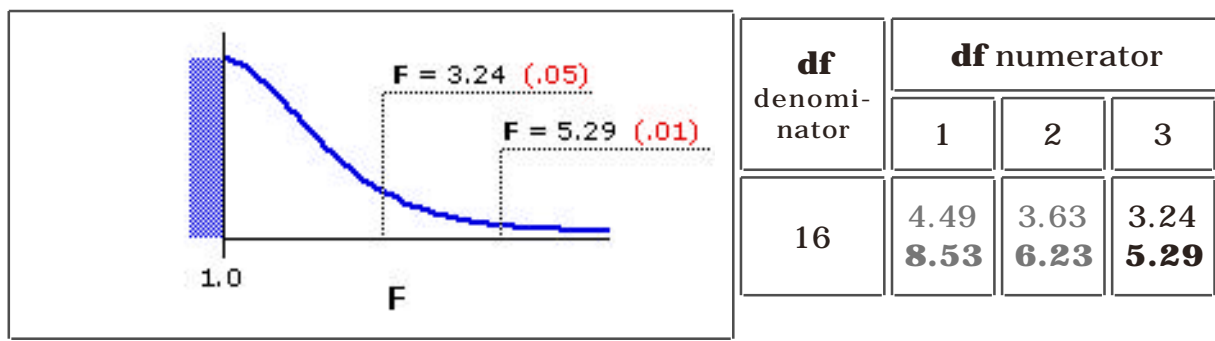
and

$$MS_{wg} = \frac{SS_{wg}}{df_{wg}} = \frac{116.32}{16} = 7.27$$

And this, in turn, permits the calculation of the **F**-ratio as

$$F = \frac{MS_{bg}}{MS_{wg}} = \frac{46.70}{7.27} = 6.42 \qquad \text{with } \mathbf{df} = 3,16$$

Figure 14.2 shows the sampling distribution of **F** for $\mathbf{df} = 3,16$. As indicated, $\mathbf{F} = 3.24$ and $\mathbf{F} = 5.29$ mark the points beyond which fall 5% and 1%, respectively, of all possible mere-chance outcomes, assuming the null hypothesis to be true. These are the same numbers that appear in the table of critical values of **F** (Appendix D), the relevant portion of which is shown adjacent to the graph.

Figure 14.2. Sampling Distribution of F for df=3,16

df denomi- nator	df numerator		
	1	2	3
16	4.49 **8.53**	3.63 **6.23**	3.24 **5.29**

As the observed value of $\mathbf{F} = 6.42$ falls to the right of $\mathbf{F} = 5.29$, our investigators can regard the aggregate differences among the means of their four samples as significant beyond the .01 level. Recall once again that this term "significant" always has an If/Then logical structure embedded within it, and that the center-point of the structure is always the null hypothesis. For the present example the structure is this: If the null hypothesis were true— if the differences among the means of the four samples were occasioned by nothing more than random variability— then the likelihood of ending up with an **F**-ratio this large or larger would be less than 1%

($P<.01$). The investigators can accordingly reject the null hypothesis with a level of confidence somewhat greater than 99%, provisionally concluding that their experimental medication does have the effect they supposed it to have.

It is possible to calculate the probability associated with an **F**-ratio in more fine-grained detail than is provided by the .05 and .01 critical values conventionally listed in tables of **F**. Although the details of this calculation are more complex than most users of statistical methods would care to go into, it can be easily performed with the built-in functions of certain computer spreadsheet applications such as Microsoft Excel. For the present example, with **F**$=6.42$ and **df**$=3,16$, the more fine-grained probability value is **P**$=.005$.

• ANOVA Summary Tables

When reporting the results of an analysis of variance, it is good practice to present a summary table such as the following. Clearly identifying each component of **SS**, **df**, and **MS**, it allows the reader to take in the main details of the analysis at a single glance.

Source	SS	df	MS	F	P
between groups ("effect")	140.10	3	46.70	6.42	<.01*
within groups ("error")	116.32	16	7.27		
TOTAL	256.42	19			

*As indicated above, the probability value for this analysis
  could also be reported as **P**$=.005$.

¶Assumptions of the One-Way ANOVA for Independent Samples

This particular version of the analysis of variance makes the following assumptions about the data that are being fed into it:

1. that the scale on which the dependent variable is measured has the properties of an equal interval scale;

2. that the **k** samples are independently and randomly drawn from the source population(s);

3. that the source population(s) can be reasonably supposed to have a normal

distribution; and

4. that the **k** samples have approximately equal variances.

You have already encountered the first three of these assumptions in connection with the **t**-test for independent samples. The fourth assumption will require a bit of explaining. Listed below are sums of squared deviates within each of our 4 samples. Dividing each by N_g-1 (recall that the subscript "g" means "any particular group") will give an estimate of the variance of the population from which the sample comes. The analysis we have just performed in Part 1 of this chapter assumes that these four variance estimates are all approximately equal. As a practical rule of thumb, you can take the phrase "all approximately equal" to entail that the ratio of the largest sample variance to the smallest should not exceed 1.5.

Group	SS	N_g-1	Variance Estimate
A	32.07	4	8.02
B	23.49	4	5.87
C	45.26	4	11.32
D	15.50	4	3.88

As you can see, this equal-variance assumption is potentially a bit of bad news. The largest of our sample variances (11.32) is nearly three times as great as the smallest (3.88) and nearly twice as great as the next smallest (5.87). This bad news does not apply only to the present example. It very often happens in real-life applications of ANOVA that sample variances do not satisfy the assumption of being "approximately equal."

But now for the good news. The analysis of variance is a very **robust** test, in the sense that it is relatively unperturbed when the equal-variance assumption is not met. This is especially so when the **k** samples are all of the same size, as in the present example. Hence, for this or any other version of ANOVA, it is always a good idea to ensure that all samples are of the same size. (When the several samples are of different sizes, the rule of thumb mentioned above remains in force: the ratio of the largest sample variance to the smallest should not exceed 1.5.)

When the samples are of the same size, the analysis of variance is also robust with respect to the assumption that the source populations are normally distributed. So, in brief, the one-way ANOVA for independent samples can be applied to virtually any set of data that will fit into it, providing that all **k** of the samples are of equal size and that the first two of the assumptions are met:

1. that the scale on which the dependent variable is measured has the properties of an equal interval scale; and

2. that the **k** samples are independently and randomly drawn from the source population(s).

Indeed, with equal sample sizes even the first of these assumptions can be safely flexed, up to a point. In some precincts of research you will quite often find ANOVA

procedures usefully applied to data that derive from rating scales, which as indicated in Chapter 2 do not intrinsically have the properties of an equal interval scale. However, I would urge the beginning student to apply ANOVA procedures to data of this sort only under the guidance of an experienced researcher.

¶Step-by-Step Computational Procedure: One-Way Analysis of Variance for Independent Samples

I will show the procedures for the case where the number of groups is **k** = 4. The modifications required for different values of **K** will be fairly obvious. The steps listed below assume that you have already done the basic number-crunching to get $\sum X_i$ and $\sum X^2{}_i$ for each of the **k** groups and for all **k** groups combined.

Step 1. Combining all **k** groups together, calculate

$$SS_T = \sum X^2{}_i - \frac{(\sum X_i)^2}{N_T}$$

Step 2. For each of the **k** groups separately, calculate the sum of squared deviates within the group ("g") as

$$SS_g = \sum X^2{}_{gi} - \frac{(\sum X_{gi})^2}{N_g}$$

Step 3. Take the sum of the SS_g values across all **k** groups to get

$$SS_{wg} = SS_a + SS_b + SS_c + SS_d$$

Step 4. Calculate SS_{bg} as

$$SS_{bg} = SS_T - SS_{wg}$$

Step 4a. Check your calculations up to this point by calculating SS_{bg} separately as

$$SS_{bg} = \frac{(\sum X_{ai})^2}{N_a} + \frac{(\sum X_{bi})^2}{N_b} + \frac{(\sum X_{ci})^2}{N_c} + \frac{(\sum X_{di})^2}{N_d} - \frac{(\sum X_T)^2}{N_T}$$

Step 5. Calculate the relevant degrees of freedom as

$$df_T = N_T - 1$$
$$df_{bg} = k - 1$$
$$df_{wg} = N_T - k$$

Step 6. Calculate the relevant mean-square values as

$$MS_{bg} = \frac{SS_{bg}}{df_{bg}}$$

and

$$MS_{wg} = \frac{SS_{wg}}{df_{wg}}$$

Step 7. Calculate **F** as

$$F = \frac{MS_{bg}}{MS_{wg}}$$

Step 8. Refer the calculated value of **F** to the table of critical values of **F** (Appendix D), with the appropriate pair of numerator/denominator degrees of freedom, as described earlier in this chapter.

End of Chapter 14, Part 1.

Home	Click this link only if the present page does not appear in a frameset headed by the logo **Concepts and Applications of Inferential Statistics**