Two Way Independent ANOVA: Theory

The Basic Idea

The name of this test can be broken down to tell us the type of design with which it is used. The 'two-way' part of the name simply means that two independent variables have been manipulated in the experiment. The 'independent' part of the name tells us that different participants were used in all conditions, and the ANOVA part tells us that we're comparing variances (i.e. we're looking at differences between means). Therefore, this design is used when you have two between-groups independent variables: each subject participates in only one experimental condition. When more than one independent variable is manipulated it is known as a *factorial* design. These designs are advantageous because they allow us to look at the interactions between variables.

Two-way ANOVA is conceptually very similar to a one-way ANOVA. Basically, we still find the total sum of squared errors (SS_T) and break this variance down into variance that can be explained by the experiment (SS_M) and variance that cannot be explained (SS_R). However, in two-way ANOVA, the variance explained by the model is made up of not one experimental manipulation but two. Therefore, we break the model sum of squares down into variance explained by the first independent variable (SS_A), variance explained by the second independent variable (SS_B) and variance explained by the interaction of these two variables ($SS_{A \times B}$).

How do you calculate two-way ANOVA?

Well, basically we start off in the same way as we did for a one-way ANOVA. That is, we calculate how much variability there is between scores when we ignore the experimental condition to which they belong. If we look at the example in Field (2000, section 8.2.1., p. 310) I describe an experiment to test the 'beer-goggles effect'. To recap, a psychologist was interested in the effects of alcohol on mate selection at night-clubs. She picked 48 students: 24 male and 24 female. She then took groups of 8 subjects to a night-club and gave them either no alcohol (subjects received placebo drinks of alcohol-free lager), 2 pints of strong lager, or 4 pints of strong lager. At the end of the evening she took a photograph of the person that the subject was chatting up. She then got a pool of independent raters to assess the attractiveness of the person in each photograph (out of 100). The data are presented in the table below.

Alcohol	**None**		**2 Pints**		**4 Pints**	
Gender	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
	65	50	70	45	55	30
	70	55	65	60	65	30
	60	80	60	85	70	30
	60	65	70	65	55	55
	60	70	65	70	55	35
	55	75	60	70	60	20
	60	75	60	80	50	45
	55	65	50	60	50	40
Total	**485**	**535**	**500**	**535**	**460**	**285**
Mean	**60.625**	**66.875**	**62.50**	**66.875**	**57.50**	**35.625**
Variance	**24.55**	**106.70**	**42.86**	**156.70**	**50.00**	**117.41**

© Dr. Andy Field, 2000

Step 1: Calculate SS_T

Remember from one-way ANOVA that SS_T is calculated using the following equation:

$$SS_T = s^2_{\text{grand}}(N-1)$$

The grand variance is simply the variance of all scores when we ignore the group to which they belong. So if we treated the data as one big group it would look as follows:

65	50	70	45	55	30
70	55	65	60	65	30
60	80	60	85	70	30
60	65	70	65	55	55
60	70	65	70	55	35
55	75	60	70	60	20
60	75	60	80	50	45
55	65	50	60	50	40

Grand Mean = 58.33

The variance of these scores is **190.78** (try this on your calculators). We used 48 scores to generate this value, and so *N* is 48. As such the equation becomes:

$$\begin{aligned} SS_T &= s^2_{\text{grand}}(N-1) \\ &= 190.78\,(48-1) \\ &= 8966.66 \end{aligned}$$

The degrees of freedom for this SS will be *N*–1, or 47.

Step 2: Calculate SS_M

The next step is to work out the model sum of squares. As I suggested earlier, this sum of squares is then further broken into three components: variance explained by the first independent variable (SS_A), variance explained by the second independent variable (SS_B) and variance explained by the interaction of these two variables ($SS_{A\times B}$).

Before we break down the model sum of squares into its component parts, we must first calculate it's value. We know we have 8966.66 units of variance to be explained and our first step is to calculate how much of that variance is explained by our experimental manipulations overall (ignoring which of the two independent variables is responsible). When we did one-way ANOVA we worked out the model sum of squares by looking at the difference between each group mean and the overall mean. We can do the same here. We effectively have 6 experimental groups if we combine all levels of the two independent variables (3 doses for the male participants and 3 doses for the females). So, given that we have 6 groups of different people we can then apply the equation for the model sum of squares that we used for one-way ANOVA (equation 7.4 in Field, 2000, p. 255):

$$SS_M = \sum n_i\left(\bar{x}_i - \bar{x}_{\text{grand}}\right)^2$$

The grand mean is the mean of all scores (above, 58.33), and *n* is the number of scores in each group (i.e. the number of participants in each of the 6 experimental groups; 8 in this case). Therefore, the equation becomes:

$$\begin{aligned} SS_M &= 8(60.625-58.33)^2 + 8(66.875-58.33)^2 + 8(62.5-58.33)^2 + 8(66.875-58.33)^2 + 8(57.5-58.33)^2 + 8(35.625-58.33)^2 \\ &= 8(2.295)^2 + 8\,(8.545)^2 + 8(4.17)^2 + 8\,(8.545)^2 + 8(-0.83)^2 + 8\,(-22.705)^2 \\ &= 42.1362 + 584.1362 + 139.1112 + 584.1362 + 5.5112 + 4124.1362 \\ &= 5479.167 \end{aligned}$$

The degrees of freedom for this SS will be the number of groups used minus 1 (see handout on one-way ANOVA). We used 6 groups and so the *df* are 5.

At this stage we know that the model (our experimental manipulations) can explain 5479.167 units of variance out of the total of 8966.66 units. The next stage is to further break down this model sum of squares to see how much variance is explained by our independent variables separately.

Step 2a: Calculate SS_A

To work out the variance accounted for by the first independent variable (in this case Gender) we need to group the scores in the data set according to which gender they belong. So, basically we ignore the amount of drink that has been drunk, and we just place all of the male scores into one group and all of the female scores into another. So, the data will look like this (note that the first box contains the 3 female columns from our original table and the second box contains the male columns):

A_1: Female		
65	70	55
70	65	65
60	60	70
60	70	55
60	65	55
55	60	60
60	60	50
55	50	50

Mean Female = 60.21

A_2: Male		
50	45	30
55	60	30
80	85	30
65	65	55
70	70	35
75	70	20
75	80	45
65	60	40

Mean Male = 56.46

We can then apply the equation for the model sum of squares that we used for one-way ANOVA (equation 7.4 in Field, 2000, p. 255):

$$SS_A = \sum n_i(\bar{x}_i - \bar{x}_{\text{grand}})^2$$

The grand mean is the mean of all scores (above), and *n* is the number of scores in each group (i.e. the number of males and females; 24 in this case). Therefore, the equation becomes:

$$\begin{aligned} SS_{\text{Gender}} &= 24(60.21 - 58.33)^2 + 24(56.46 - 58.33)^2 \\ &= 24(1.88)^2 + 24(-1.87)^2 \\ &= 84.8256 + 83.9256 \\ &= 168.75 \end{aligned}$$

The degrees of freedom for this SS will be the number of groups used minus 1 (see handout on one-way ANOVA). We used 2 groups and so the *df* are 1.

Step 2b: Calculate SS_B

To work out the variance accounted for by the second independent variable (in this case Alcohol) we need to group the scores in the data set according to how much alcohol was consumed. So, basically we ignore the gender of the participant, and we just place all of the scores after no drinks in one group, the scores after 2 pints in another group, and the scores after 4 pints in a third group. So, the data will look like this:

B_1: None	
65	50
70	55
60	80
60	65
60	70
55	75
60	75
55	65

Mean None = 63.75

B_2: 2 Pints	
70	45
65	60
60	85
70	65
65	70
60	70
60	80
50	60

Mean 2 Pints = 64.6875

B_3: 4 Pints	
55	30
65	30
70	30
55	55
55	35
60	20
50	45
50	40

Mean 4 Pints = 46.5625

We can then apply the equation for the model sum of squares that we used for one-way ANOVA (equation 7.4 in Field, 2000, p. 255):

$$SS_B = \sum n_i\left(\bar{x}_i - \bar{x}_{\text{grand}}\right)^2$$

The grand mean is the mean of all scores (above), and *n* is the number of scores in each group (i.e. the number of scores in each of the boxes above, in this case 16). Therefore, the equation becomes:

$$\begin{aligned} SS_{\text{alcohol}} &= 16(63.75 - 58.33)^2 + 16(64.6875 - 58.33)^2 + 16(46.5625 - 58.33)^2 \\ &= 16(5.42)^2 + 16(6.3575)^2 + 16(-11.7675)^2 \\ &= 470.0224 + 646.6849 + 2215.5849 \\ &= 3332.292 \end{aligned}$$

The degrees of freedom for this SS will be the number of groups used minus 1 (see handout on one-way ANOVA). We used 3 groups and so the *df* are 2.

Step 2c: Calculate $SS_{A \times B}$

The final stage is to calculate how much variance is explained by the interaction of the two variables. The simplest way to do this, is to remember that the SS_M is made up of three components (SS_A, SS_B, and $SS_{A \times B}$), therefore, given that we know SS_A and SS_B we can calculate the interaction term using subtraction:

$$SS_{A\times B} = SS_M - SS_A - SS_B$$

Therefore, for these data, the value is:

$$\begin{aligned} SS_{A\times B} &= SS_M - SS_A - SS_B \\ &= 5479.167 - 168.75 - 3332.292 \\ &= 1978.125 \end{aligned}$$

The degrees of freedom can be calculated in the same way, but are also the product of the degrees of freedom for the main effects (either method works):

$$\begin{aligned} df_{A\times B} &= df_M - df_A - df_B \\ &= 5 - 1 - 2 \\ &= 2 \end{aligned}$$

$$\begin{aligned} df_{A\times B} &= df_A \times df_B \\ &= 1 \times 2 \\ &= 2 \end{aligned}$$

Step 3: Calculate SS_R

The residual sum of squares is calculated in the same way as for one-way ANOVA and again represents individual differences in performance or the variance that can't be explained by factors that were systematically manipulated. We saw in one-way ANOVA that the value is calculated by taking the squared error between each data point and its corresponding group mean. An alternative way to express this was as:

$$SS_R = s^2_{\text{group1}}(n_1 - 1) + s^2_{\text{group2}}(n_2 - 1) + s^2_{\text{group3}}(n_3 - 1) \ldots s^2_{\text{group n}}(n_n - 1)$$

So, we use the individual variances of each group and multiply them by one less than the number of people within the group (n). We have the individual group variances in our original table of data (first page of this handout) and there were 8 people in each group (therefore, n = 8) and so the equation becomes:

$$\begin{aligned} SS_R &= s^2_{\text{group1}}(n_1 - 1) + s^2_{\text{group2}}(n_2 - 1) + s^2_{\text{group3}}(n_3 - 1) + s^2_{\text{group4}}(n_4 - 1) + s^2_{\text{group5}}(n_5 - 1) + s^2_{\text{group6}}(n_6 - 1) \\ &= (24.55)(8-1) + (106.7)(8-1) + (42.86)(8-1) + (156.7)(8-1) + (50)(8-1) + (117.41)(8-1) \\ &= (24.55 \times 7) + (106.7 \times 7) + (42.86 \times 7) + (156.7 \times 7) + (50 \times 7) + (117.41 \times 7) \\ &= 171.85 + 746.9 + 300 + 1096.9 + 350 + 821.87 \\ &= 3487.52 \end{aligned}$$

The degrees of freedom for each group will be one less than the number of scores per group (i.e. 7). Therefore, if we add the sums of squares for each group, we get a total of 6 × 7 = 42.

Step 4: Calculate Mean Squares and F-ratios

Just as in one-way ANOVA, the mean squares for each SS is calculated by dividing by the degrees of freedom:

$$MS_A = \frac{SS_A}{df_A} = \frac{168.75}{1} = 168.75$$

$$MS_B = \frac{SS_B}{df_B} = \frac{3332.292}{2} = 1666.146$$

$$MS_{A\times B} = \frac{SS_{A\times B}}{df_{A\times B}} = \frac{1978.125}{2} = 989.062$$

$$MS_R = \frac{SS_R}{df_R} = \frac{3487.52}{42} = 83.036$$

The F-ratios for the two independent variables and their interactions are then calculated by dividing their mean squares by the residual mean squares.

$$F_A = \frac{MS_A}{MS_R} = \frac{168.75}{83.036} = 2.032$$

$$F_B = \frac{MS_B}{MS_R} = \frac{1666.146}{83.036} = 20.065$$

$$F_{A\times B} = \frac{MS_{A\times B}}{MS_R} = \frac{989.062}{83.036} = 11.911$$

Step 5: Construct a Summary Table and Calculate the Significance of the Fs

All of these values can be summarised in a table as we did for the one-way ANOVA. AS before, we can look up the critical value for each of the three *F*-ratios by using their degrees of freedom (that is the degrees of freedom for the effect, and the degrees of freedom for the error). The summary table for this example will be as follows (using the values we've calculated):

Source	SS	df	MS	F	Sig.
Model	5479.17	5	1095.833		
A (gender)	168.75	1	168.75	2.03	Ns
B (alcohol)	3332.29	2	1666.146	20.06*	< 0.001
A × B (alcohol × gender)	1978.13	2	989.062	11.91*	<0.001
Residual (Error)	3487.52	42	83.036		
Total	8966.66	47			

Compare this table with SPSS output 8.12 in Field (2000, p. 318). This handout explains where these values come from and should help you to understand the process by which the model variance is split into its component parts.

Now, look at the Handout on conducting two-way Independent ANOVA on SPSS.

- You will **not** be expected to conduct two-way ANOVA by hand in your exam. This handout is simply additional reading material for the lecture.

This handout is based around material from:

Field, A. P. (2004). *Discovering statistics using SPSS: advanced techniques for the beginner (second edition)*. London: Sage.

If you want a copy of the full chapter, just email me, or see Field (2000, chapter 8) for details on how to use SPSS to analyse factorial ANOVAs.