5.04 Analysis of Variance: Factorial ANOVA

In this video we'll discuss the basics of factorial ANOVA, with not one, but two or more factors and one dependent variable. Some sources refer to factorial ANOVA as a multivariate technique because it involves more than two variables. However, the term multivariate is more commonly reserved for techniques involving two or more *dependent* variables, like multivariate ANOVA or factor analysis.
Factorial ANOVA - with just one dependent variable - combines the levels of all factors, allowing us to test whether the means differ between the factor levels for each factor separately, but also if there is a combined effect of the factors, called and interaction effect.

The factors are categorical independent variables, each with two or more categories. In a full factorial design the factors are crossed, meaning that each level of a factor is combined with each of the levels of the other factors.

Suppose we want to investigate the effect of raw meat versus canned food and the effect of portion control versus free-feeding on cat health, rated on a scale from zero to ten. We cross these factors by creating four groups of cats that are randomly assigned to be fed 1) raw meat in controlled portions - appropriate for their age, sex and breed 2) raw meat fed freely - they're allowed as much food as they want 3) canned food in controlled portions, and 4) canned food fed freely.

With two factors - each with two levels - we have two times two equals four groups. We call this a two-factor analysis with a two-by-two design. The number of factors and levels can be larger, For example, if we included the factor hairiness - with the levels hairless, shorthair and longhair - we would have a three-way analysis with a two-by-two-by-three design and two times two times three equals twelve groups or cells.

Let's keep things simple and consider the two-by-two example. We can compare the groups in several ways. First of all we can consider the effect of diet by collapsing over feeding pattern and comparing the *marginals* for diet. So we ignore the distinction between controlled versus free feeding and compare all cats fed on raw meat with all cats fed on canned food. This is referred to as checking the main effect of diet.

We can test whether the mean health scores differ between the two diets using the same principle as in one-way ANOVA, by estimating the between- and within-group variance and seeing whether the between-group variance is larger, indicating a difference in the means. The only thing different from one-way ANOVA is that we have to keep careful track of what means we're comparing.

For the main effect of diet the corresponding F-test numerator (the between-group variance) is determined by calculating the sum of squares for the diet groups compared to the grand mean and dividing by the appropriate degrees of freedom, the number of groups minus one:
$MS_{Diet} = \frac{(\bar{y}_r - \bar{y})^2 + (\bar{y}_c - \bar{y})^2}{2-1}$.
The denominator is the within-group variance: The sum of squares calculated by comparing each observation with its group mean for each of the four individual groups - not the marginals! - and dividing by the appropriate degrees of freedom. These are the number of observations minus the total number of groups:

$$MS_{within} = \frac{\Sigma(\bar{y}_{ij} - \bar{\bar{y}}_j)^2}{n-4}, \qquad F = \frac{MS_{Diet}}{MS_{within}} = \frac{\frac{(\bar{y}_r - \bar{y})^2 + (\bar{y}_c - \bar{y})^2}{2-1}}{\frac{\Sigma(\bar{y}_{ij} - \bar{\bar{y}}_j)^2}{n-4}}.$$

We can test the main effect of feeding pattern the same way. We calculate the between-group variance by comparing the marginal means of the controlled and the free-fed cats and divide by the same within-group variance we just calculated for the main effect of diet. We can now test whether mean health score differs for cats fed in controlled portions and free-fed cats:

$$F = \frac{MS_{FP}}{MS_{within}} = \frac{\frac{(\bar{y}_p - \bar{y})^2 + (\bar{y}_f - \bar{y})^2}{2-1}}{\frac{\Sigma(\bar{y}_{ij} - \bar{\bar{y}}_j)^2}{n-4}}.$$

Finally we can check for an interaction effect. We can see if the effect of one factor influences the effect of the other factor, for example if portion control strengthens a positive effect of a raw meat diet. Another example could be that canned food negatively affects health, but only if cats are free-fed. Fully crossed, factorial ANOVA allows us to test interactions, which is not possible with two separate t-tests or separate one-way ANOVAs.

The within-group variance for the interaction is calculated in the same way as before. The between-group variance is calculated by taking the individual group means - not the marginals! - subtracting the corresponding marginal means for that groups level of diet and feeding pattern and adding the grand mean. The degrees of freedom for the interaction effect are the number of levels for the factor diet minus one, times the number of levels for the factor feeding pattern minus one:

$$F = \frac{MS_{Diet\,x\,FP}}{MS_{within}} = \frac{\frac{(\bar{y}_{pr}-\bar{y}_{p}-\bar{y}_{r}+\bar{y})^2+(\bar{y}_{pc}-\bar{y}_{p}-\bar{y}_{c}+\bar{y})^2+(\bar{y}_{fr}-\bar{y}_{f}-\bar{y}_{r}+\bar{y})^2+(\bar{y}_{fc}-\bar{y}_{f}-\bar{y}_{c}+\bar{y})^2}{(2-1)\cdot(2-1)}}{\frac{\Sigma(\bar{y}_{ij}-\bar{\bar{y}}_{j})^2}{n-4}}.$$

You won't be asked to calculate the sums of squares for main and interaction effects manually. However, you should be able to calculate the F-values based on sums of squares provided by software:

$$F = \frac{MS_{Factor}}{MS_{within}} = \frac{\frac{SS_{Factor}}{df_{Factor}}}{\frac{SS_{within}}{df_{within}}}.$$

It helps to know that the total sum of squares (all the variation in the response variable) is equal to the sum of squares for diet, plus the sum of squares for feeding pattern plus the sum of squares fro the interaction plus the within-group or error sum of squares: $SS_{tot} = SS_{D} + SS_{FP} + SS_{D\,x\,FP} + SS_{within}$.

If the interaction is significant it can help to interpret it by following up with pairwise comparisons of the individual groups to see what causes the interaction. This is referred to as testing the simple effects. We'll discuss follow-up comparisons later.

If we represent the group means visually, parallel lines represent a lack of interaction, and non-parallel lines represent an interaction. If an interaction effect is significant it might not make sense to interpret the main effects.
For example if a raw meat diet results in higher health compared to canned food, but only if portions are controlled, then even if a main effect of diet is significant it's meaningless, because we already know this effect occurs only for a certain type of feeding pattern.
Main effects also don't require interpretation if we find a crossing interaction, for example if cats free-fed on raw meat and on portion-fed canned food are in good health and cats portion-fed on raw meat and free-fed on canned food are in poor health. In these cases the main effects are cancelled out by the interaction.
In other cases a main effect can be meaningful, even if an interaction is present, for example if a raw meat diet results in better health for both feeding patterns, just more so when portions are controlled.