From the Expert: Two-way ANOVA

From the Expert: Two-way Analysis of Variance (Equal sample sizes)

Two-way analysis of variance is also known as two-way ANOVA or two-factor ANOVA. This method is used to compare the means of two-factor experiments. The two factors are commonly referred to as Factor A (with a levels) and Factor B (with b levels).

An example of a two-factor experiment: the effect of diet (diet1, diet2, diet3) and gender (male, female) on weight loss. Diet (Factor A) has 3 levels, and gender (Factor B) has 2 levels. So, there are `a x b = 3 x 2 = 6` treatment combinations. If all these combinations have equal sample sizes, then this design will be called a ***balanced complete two-factor factorial*** design. In short, factorial design is a study where each independent variable is crossed to create all possible combinations. This notation can be extended to multi-factor designs such as 3 factors factorial design is `a x b x c`, when `a,b,c` are the number of levels of the first factor, second factor, and third factor, respectively. This chapter discusses a two-factor design.

Why a two-factor experiment?

Besides understanding the effects of individual factors on the response, two-factor design allows researchers to examine the joint effects of two factors. Joint effects of two factors are also known as interactions. This is part of the ANOVA power that can test and estimate interaction effects.

In summary, there are 2 kinds of effects:

- Main effects – refers to primary factors in the experiment. A change in the levels of a factor can affect the response (e.g. effects of Factor A on the response, effects of Factor B on the response). It is the effect of one factor, when the other factor is held constant. In other words, it is the effects of one factor averaged over all levels of the other factor.
- Interaction effects – the effect of one factor on the response depending on the level of the other factor. In other words, it is combined effects of both factors (A and B) on the response.

Data Summary Notation

http://www.math.montana.edu/~cherry/st412/pdf_files/CourseNotes14.pdf

- the number of levels of factor A
- the number of levels of factor B
- the number of replications at each cell (combination of factor levels)

- the observation from combinations of factors A and B levels
- total number of observations (=abn)
- (the sum of all observations)
- (the sum for treatment *i* of factor A)
- (the sum for treatment *j* of factor B)
- (the grand mean of all observations)
- (the treatment mean of factor A)
- (the treatment mean of factor B)
- (the cell mean of factor A and factor B)

ANOVA notation

- (the sum of squares for factor A, with degree of freedom)
- (the mean square for factor A)
- (the sum of squares for factor B, with degree of freedom)
- (the mean square for factor B)
- (the sum of squares interaction between A and B, with degree of freedom)
-
- (the sum of squared error)
- (the mean square error)
- (the sum of squared total)

ANOVA Table

Source of Variation	Sum of Squares	d.f.	Mean Square	F Ratio
Total				

Let be the overall mean

be the effet for *i*th level of factor A
be the effect for *j*h level of factor B
be the (*i,j*)th interaction effect of Factor A and B
be the random error
be the response of *k*th observation at *i*th level of factor A and *j*h level of factor B
Statistical Model can be formulated as:

Hypothesis Testing

The testing procedure for two-way ANOVA can be very complicated. Follow these steps to perform F-tests:

- Test for interactions (suggested):

all (no interaction between factor A and factor B)
all (interaction terms are not all zero)

Test statistic ,with d.f. =
If is rejected, then there is a significant interaction. Skip step 2 and 3 below and (usually) continue to compare treatment means using multiple comparison procedure such as as Bonferroni or Tukey. If the results show no significant interaction, it is justified to compare the factor levels (as given in step 2 and 3)

In case of no interaction, two factors can examine separately one at a time to find best level of factor A and best level of factor B.

- Test for Factor A (main effects):

no effect of factor A on response variable)
: not all

Test statistic ,with d.f. =
If is rejected, then average response differ for A (i.e. average weight loss differ for 3 diets)

- Test for Factor B (main effects):

no effect of factor B on response variable)
: not all

Test statistic ,with d.f. =
If is rejected, then average response differs for B (i.e. average weight loss differs for males and females)

2. Interaction Effects

To understand the interaction effects, we usually do interaction plots. Interaction plot is a plot of means against levels of one factor, with different lines for the other factor. For example, in the study on different advertising agencies, and different advertising mediums on sale, there are two possible plots:

- plot sale against agency, with different lines representing different medium
- plot sale against medium, with different lines representing different agency

Interaction plots that are parallel or near parallel imply there is no significant interaction (e.g. effect of one factor is the same at all levels of the other factor). Otherwise, it may indicate interaction because the effect of one factor is different at least on one level of the other factor)

Interaction plot between advertising agency (A) and the advertising medium (B)

Source: http://www.math.montana.edu/~cherry/st412/pdf_files/CourseNotes14.pdf

These plots indicate an interaction because the lines are not parallel.

Examples of different types of main effects and interaction effects can be found at:

EXPERIMENTAL RESEARCH DESIGNS

(page 11-12)

3. Two-way ANOVA Assumptions

Assumptions are the same as one-way ANOVA (i.e. normally distributed data and homogeneity of variance). Therefore, we can use the same procedures as one-way ANOVA. Fortunately, two-way ANOVA is quite robust if sample sizes are equal.

4. Two-way ANOVA procedure

- Check assumptions (normality, independence, homogeneity)
- Do F-test for interaction effect
- If there is no interaction, do main effects testing on each factor. Compare individual levels of a factor by applying t-test with Bonferroni correction for number of comparisons made.
- If there is an interaction, use interaction plot to visualize the graph. Create combination of factors. Then, apply Tukey's HSD to examine which groups are different.

5. Reminder:

If it appears that there is an interaction between two factors A and B, the main effects should also be included in the model. Because interaction indicates that both factors are important.

Marginal means are average over the levels of the other factor.

Examples of Two-way ANOVA using R include:

http://www.stat.columbia.edu/~martin/W2024/R8.pdf

References:

Cherry, S. Two Factor ANOVA – Equal Sample Sizes. Montana State University.Retrieved from:http://www.math.montana.edu/~cherry/st412/pdf_files/CourseNotes14.pdf

Lindquist, A.M. (2009) Introduction to Statistics, course Notes. Columbia University.
Retrieved from: http://www.stat.columbia.edu/~martin/W2024/R8.pdf

Stevens, J.J. Interaction Effects in ANOVA. University of Oregon. Retrieved from:
http://pages.uoregon.edu/stevensj/interaction.pdf

Activity Details

Task: View this topic