Lecture Notes

# Chapter 11: Analysis Of Variance

## Learning Objectives

11.1 Explain the application of a one-way analysis of variance (ANOVA) model.

11.2 Define the concepts of between and within total variance.

11.3 Calculate and interpret a test for two or more sample cases with means.

11.4 Determine the significance of an F-ratio test statistic.

11.5 Interpret output for ANOVA.

## Chapter Outline

1. Understanding Analysis of Variance
   1. **Analysis of variance** (ANOVA) is an inferential statistics technique designed to test for a significant relationship between two variables in two or more groups or samples.
   2. “*t*”test examines the difference between two means, , while the null hypothesis assumes that there is no difference between them: μ1 = μ2.
   3. The logic of ANOVA is same but extending to two or more groups.
   4. When ANOVA procedures are applied to data with one dependent and one independent variable, it is called a **one-way ANOVA**.
   5. To determine whether the differences are significant, ANOVA examines:
      1. The differences betweenthe samples.
      2. The differences withina single sample.
      3. The differences can also be referred to as variance or variation, which is why ANOVA is the analysis of variance.
      4. ANOVA allows us to determine whether the variance between samples is larger than the variance within the samples.
2. The Structure of Hypothesis Testing With ANOVA
   1. The Assumptions
      1. Independent random samples are used.
      2. The dependent variable is an interval-ratio level of measurement. Some researchers also apply ANOVA to ordinal-level measurements.
      3. Although we cannot confirm whether the populations are normal, given that our *N* is so small, we must assume that the population is normally distributed to proceed with our analysis.
      4. The population variances are equal. Based on calculations, the sample variances, although not identical, are relatively homogeneous.
   2. Stating the Research and the Null Hypotheses and Setting Alpha
      1. Research hypothesis (H1) proposes, at least one of the means is different.
      2. Alpha: Probability of rejecting null hypothesis.
   3. The Concepts of Between and Within Total Variance
      1. Primary set of calculations has to do with the two types of variance:
         1. Between-group variance.
         2. Within-group variance.
      2. The **between-group sum of squares**or *SSB* measures the difference in average years of education between our four groups.
      3. Formula: **Within-group sum of squares** or *SSW* measures the variation of scores within a single sample.
3. Formula: 
   * 1. The **total sum of squares** or SSTcan be represented by 
     2. Estimating the between-group variance by calculating **mean square between** which is the sum of squares between divided by its corresponding degrees of freedom. SSB/*df*b
     3. Estimating the within-group variance by calculating **mean square within** which is the sum of squares between divided by its corresponding degrees of freedom. SSW/*df*w
   1. The *F* Statistic
      1. Together the mean square between and mean square within compose **the F ratio obtained** or ***F* statistic**.
      2. 
   2. Making a decision
      1. Calculating *F* statistic and its distribution: To determine the probability of calculating an *F* statistic, we rely on the distribution of the *F* statistic.
      2. Degrees of freedom:
         1. There are two degrees of freedom, *df*1 = *df*b and *df*2 = *df*w.
         2. Because of the two degrees of freedom, we’ll have to determine the probability of our *F* obtained differently than we did with *t* test or chi-square.
      3. *F*-critical and *F*-obtained:
4. ***F* critical:** The *F*-test statistic that corresponds to the alpha level, *df*w, and *df*b.
5. **F obtained:** The *F*-test statistic that is calculated.
6. If the *F* obtained is greater than the *F* critical, we know that its probability is <.05, extending into the shaded area.
7. If the *F* obtained was lesser than the *F* critical, we could determine that its probability was greater than our alpha.
8. Statistics in Practice: The Ethical Consumer
   1. By purchasing fair labor, organic, or environmentally safe products, consumers express their politics through their purchases, this has been referred to as ethical consumerism.
   2. The GSS2014 included a series of questions on what it takes to be a good citizen.
   3. The ANOVA output includes two tables: (1) Descriptives and (2) ANOVA.
      1. In the Descriptives table, the *N*, mean, and standard deviation are reported for each group and the entire sample, along with the 95% confidence interval for each mean.
      2. The *F* obtained is reported in the ANOVA table, along with its level of significance (or probability).
      3. Since the level of significance is greater than our alpha, we fail to reject the null hypothesis.
      4. Although the means are different, the differences are not significant.
9. Reading the Research Literature: College Satisfaction Among Latino Students
   1. Like bivariate, *t* test or chi-square analyses, ANOVA can help us understand how the categories of experience, race, age, class, and/or gender shape our social lives. ANOVA allows us to investigate a variety of social categories by comparing the differences between them.
   2. Research: How ANOVA is interpreted in the social science literature:
      1. Young Kim, Liz Rennick, and Marla Franco investigated the levels of college engagement and satisfaction among Latino undergraduate students attending highly selective colleges and universities in comparison with other racial/ethnic groups.
      2. Kim and their colleagues limited their analyses to the outcomes of 33,415 junior and senior students.
      3. All factors were measured on a 6-point Likert-type scale: 1 = *very poor* to 6 = *excellent*.
      4. The findings:
         1. In terms of satisfaction with the educational experience, quality of instruction and courses in the major, advising and out-of-class contact, and library support, Latino students and all other racial/ethnic groups rated equally.
         2. Sense of belonging and overall college satisfaction: White students reported higher levels of satisfaction.
      5. Conclusion: The researchers may argue that the college experiences of Latino students at highly selective research institutions are favorable and meaningful, which may contribute to development in desired educational outcomes at these types of institutions.