

CHAPTER FOUR

ANOVA

4.0 INTRODUCTION

Chapter 4 focuses on the nature of statistical data, ordering and manipulation of data, concepts of statistical inference, and hypothesis testing. The aim of the exercises is to ensure that students are able to use the manual calculation and SPSS programme to analyze the statistical theories and calculations. Throughout the discussion, students would be exposed to various data sets, and required to analyze types of data.

4.1 ANALYSIS OF VARIANCE: ANOVA

According to Caldwell (2007), many statisticians think of Analysis of Variance (commonly abbreviated as ANOVA) as an extension of the difference of means test because it is based, in part, on a comparison of sample means. At the same time, however, the procedure involves a comparison of different estimates of population variance—hence the name analysis of variance. Because ANOVA is appropriate for research involving three or more samples, it has wide applicability.

In the field of experimental psychology, for example, results from three or more samples, are often referred to as treatment groups. Imagine an educational psychologist wanting to know if students exposed to three different treatments conditions and learning environments (positive sanction, negative sanction and sanction neutral) exhibit different test scores. Assuming the test scores are based on an interval/ratio scale of measurement, ANOVA would be an appropriate approach to the problem. In short, ANOVA allows the comparison of multiple samples in a single application.

4.2 The Logic of ANOVA

Imagine for a moment if scores in an aptitude test actually vary for students in different types of schooling environment—home schooling, public schooling, and private schooling. The research question will involve a comparison of more than two groups. Assume that the test scores are measured at the interval/ratio level, the situation will be tailor-made for an application of ANOVA. The study could easily be one that asks whether or not aptitude test scores vary on the basis of school environment.

Another way to look at the question is whether or not the type of school environment is a legitimate classification scheme when it comes to the matter of aptitude test scores. If aptitude test scores really do vary on the basis of school environment—if there is a significant difference between the scores in the three environments—then it is probably legitimate to speak in terms of school environments when looking at test scores. If there is no significant difference between the scores, then question the legitimacy of a classification scheme.

The purpose of ANOVA is to measure whether there is more variation between groups than within groups. It examines the legitimacy of a classification scheme.

4.3 From Curves to Data Distributions

In essence, ANOVA allows calculating a ratio of the variation between groups to the variation within groups. This ratio is referred to as the F ratio (named after its developer Sir Ronald Fisher). Assume that in search of significant results in a hypothesis-testing situation, the more variation between the means of several groups, the more relative to the variation within the groups. In short, there will be more variation between than within. Because the F ratio is an expression of the between-to-within ratio, look for a large F value because all factors being equal, the larger the F ratio, the greater the probability that rejects the null hypothesis. Look for more variations between the samples than within the samples to achieve significant results.

4.4 The Different Means

The two types of mean which comes into play in ANOVA is the grand mean and the individual sample means. Each sample mean, or group mean, is a function in part, of the number of cases in the sample.

4.5 From Different Means to Different Types of Variation

The concept of variation typically involves the extent to which various scores in a distribution deviate from the mean of the distribution. The variation between-groups are an expression of the amount of deviation of sample means from the grand mean. The variation within-groups is an expression of the amount of deviation of sample scores from sample means.

4.6 The Null Hypothesis

The null hypothesis simply states that the means of the regions are equal. It can be stated as follows:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

In terms of the F ratio, recall that there has to be more variation between the regions than within the regions for the F ratio to be significant. It all goes back to the notion that the F ratio is an expression of the ratio of the variation between groups to the variation of within groups; the larger the F ratio, the more likely it is to be significant. If all the sample means are equal, there would not be any between-groups variation. Calculate the F ratio (test statistic) as a test statistic (the F ratio) meets or exceeds the critical value; reject the null hypothesis (with a known probability) of having committed a Type 1 error). When the null hypothesis would be rejected, the result showed that there are significance differences between groups. However, it does not indicate where the difference lies. In order to find out where the difference lies, a ‘post-hoc test’ should be conducted’.

4.7 Post Hoc Test

In order to determine which groups differ from each other, a post-hoc test can be conducted after the initial ANOVA test is completed. However please note that the post-hoc test should only be carried out if only the initial ANOVA problem is significant (rejected the null). If the hypothesis fails to reject the null, then there are no differences to find.

For the Tukey’s post-hoc test, the differences between the means of all of our groups will be determined. This difference score will be compared to a critical value to see if the difference is significant. The critical value in this case is the HSD (honestly significant difference) and it must be computed. It is the point when a mean difference becomes honestly significantly different.

$$HSD = q \sqrt{\frac{MS_{wg}}{n}}$$

Note that “ q ” is a table value, and n is the number of values we are dealing with in each group (not total n). MS_{wg} is the mean square error from the overall F-test that from the ANOVA you already computed. The Turkey test is usually recommended in the post hoc test because studies show it has greater power than the other tests under most circumstances and it is readily available in computer packages such as SPSS.

4.8 Procedure for one-way between-groups ANOVA with post-hoc tests using SPSS

- Select **Analyze** and click on **Compare Means**, then on **One-way ANOVA**.
- Click on **dependent variable** (continuous) and move into the box of **Dependent List**.
- Click on your **independent variable** (categorical) and move into the box of **Factor**.

- Click on **Options**, select **Descriptive**, **Homogeneity of Variance test**, **Brown-Forsythe**, **Welsh** and **Means Plot**.
- Click on **Post Hoc** and select **Tukey**
- Click on **Continue** and then **OK**

