

Analysis of Variance

By David R. Larsen

Analysis of variance, often called ANOVA is a technique that is used to test multiple samples or multiple levels of variables within a sample. It is very tempting to simply apply a **t test** multiple times to the data. In the case of multiple samples this is invalid.

In analysis of variance we are simply partitioning the variance into categories that influence the variation in the sample design. Let us consider the following data.

Table 1. Data used in the ANOVA example

<u>Sample 1</u>	<u>Sample2</u>
6	7
5	8
7	6
8	5
5	7
6	8
6	9
5	7
4	6
5	7
6	8
7	8
6	8
5	7
3	6
6	9
5.625	7.25
1.204159	1.125463

These data are from two samples and may be from two populations. The following graph is a histogram of the data set.



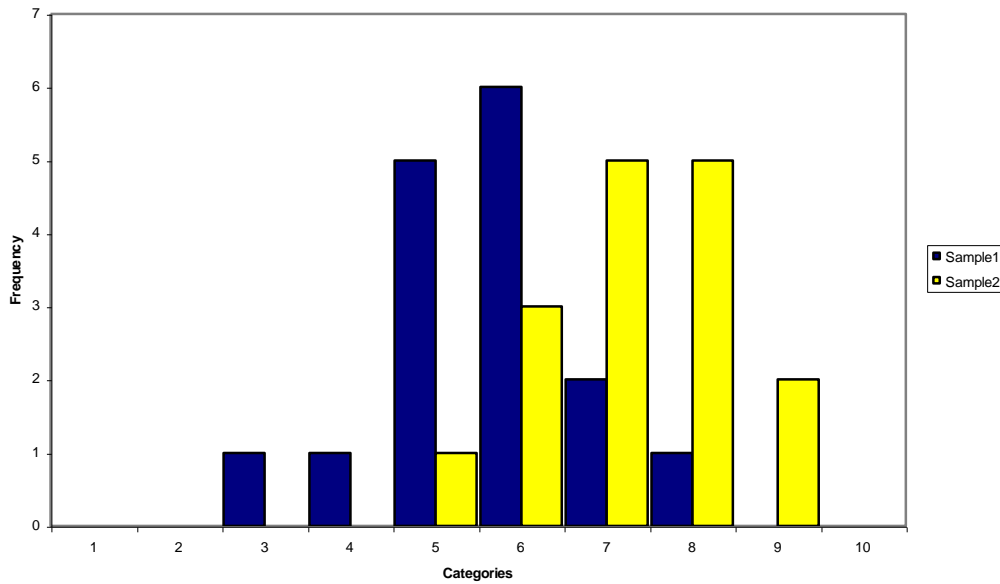


Figure 1. Histogram of the Example data set with the means plotted as lines

Using a Spreadsheet like EXCEL® we can generate a standard ANOVA table as below. The summary section just lists the sample statistics of n, mean, and variance. The second table lists the **SS** or Sums of Squares for between the groups (sample 1 or 2), within each group, and the total Sums of Squares. Next is the column for degrees of freedom **df**, the means Square **MS**, and the **F test**, which is the ration of the between group Sum of Squares and the within group Sums of Squares. The next column is the **p-value** which is the probability that the **F test** is really less than the **F crit** value due to sampling error.

The larger the **p-value** the less sure you can be in the result.

This table indicates that the two samples are significantly different and we are very sure of this result (p = 0.000446).

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
6	15	84	5.6	1.542857143
7	15	109	7.266666667	1.352380952

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	20.83333333	1	20.83333333	14.39144737	0.000728755	4.195982228
Within Groups	40.53333333	28	1.447619048			
Total	61.36666667	29				

Also See:

Chapter 10 - More on the Testing of Hypotheses pages 125-149 in:



Phillips, J. L. 2000. How to think about statistics. W. H. Freeman and Co. New York. 202 pp. ISBN 0-7167-3654-3

Chapter 10 - Single factor Analysis of Variance pages 180-191 in:

Zar, J. H. 1984. Biostatistical Analysis. Prentice-Hall, Inc. Englewood Cliffs, New Jersey. 718 pp.

