

Analysis of Variance

Mahdi Akbarzadeh

Department of Biostatistics and Epidemiology

Introduction

◆ The Purpose of Analysis of Variance

In general, the purpose of analysis of variance (ANOVA) is to test for significant differences between means.

If we are only comparing two means, ANOVA will produce the same results as the t test for independent.

Introduction

◆ Why the name analysis of variance?

It may seem odd that a procedure that compares means is called analysis of variance. However, this name is derived from the fact that in order to test for statistical significance between means, we are actually comparing (i.e., analyzing) variances.

The Partitioning of Sums of Squares

At the heart of ANOVA is the fact that variances can be divided, that is, partitioned. Remember that the variance is computed as the sum of squared deviations from the overall mean, divided by $n-1$ (sample size minus one). Thus, given a certain n , the variance is a function of the sums of (deviation) squares, or SS for short. Partitioning of variance works as follows. Consider this data set:

The Partitioning of Sums of Squares

	Group 1	Group 2
Observation 1	2	6
Observation 2	3	7
Observation 3	1	5
Mean	2	6
Sums of Squares (SS)	2	2
Overall Mean	4	
Total Sums of Squares	28	

The means for the two groups are quite different (2 and 6, respectively). The sums of squares within each group are equal to 2. Adding them together, we get 4. If we now repeat these computations ignoring group membership, that is, if we compute the total SS based on the overall mean, we get the number 28. In other words, computing the

The Partitioning of Sums of Squares

	MAIN EFFECT				
	SS	df	MS	F	p
Effect	24.0	1	24.0	24.0	.008
Error	4.0	4	1.0		

As can be seen in the above table, the total SS (28) was partitioned into the SS due to within-group variability ($2+2=4$) and variability due to differences between means ($28-(2+2)=24$).

The Partitioning of Sums of Squares

SS Error and SS Effect. The within-group variability (SS) is usually referred to as Error variance. This term denotes the fact that we cannot readily explain or account for it in the current design. However, the SS Effect we can explain. Namely, it is due to the differences in means between the groups. Put another way, group membership explains this variability because we know that it is due to the differences in means.

In SPSS

Thank You