TOPIC 51 ONE-WAY ANALYSIS OF VARIANCE

In the previous topic, you learned that the *t* test may be used to test the null hypothesis for the observed difference between two sample means. An alternative test for this problem is **analysis of variance** (often called **ANOVA**).[1] Instead of *t*, it yields a statistic called *F*, as well as degrees of freedom (*df*), sum of squares, mean square, and a *p* value, which indicates the probability that the null hypothesis is correct. As with the *t* test, the only value of interest to the typical consumer of research is the value of *p*. By convention, when *p* equals .05 or less (such as .01 or .001), we reject the null hypothesis and declare the result to be *statistically significant.*

Because the *t* test and ANOVA are based on the same theory and assumptions, when we compare two means, both tests yield exactly the same value of *p* and, hence, lead to the same conclusion regarding significance. So for two means, both tests are equivalent. Note, however, that a single *t* test can compare only two means, but a single ANOVA can compare a number of means, which is a great advantage.

Suppose, for example, we try three drugs designed to treat depression in an experiment and obtain the means in Table 1.

Table 1
Posttest Means: Depression Scores for Three Drugs

Drug A	Drug B	Drug C
$M = 6.00$	$M = 5.50$	$M = 2.33$

Note that the higher the score, the greater the depression. Inspection of the means shows that there are three observed differences:

1. Drug C is superior to Drug A.
2. Drug C is superior to Drug B.
3. Drug B is superior to Drug A.

The null hypothesis says that this *entire set* of three differences was created by sampling error. Through a series of computations that are beyond the scope of this book, an ANOVA for these data was conducted and this is the result: $F = 10.837$, $df = 2, 15$, $p < .05$. This result might be stated in a sentence or presented in a table such as Table 2, which is called an ANOVA table.

Table 2
ANOVA for Data in Table 1

Source of Variation	*df*	Sum of Squares	Mean Square	*F*
Between Groups	2	47.445	23.722	10.837*
Within Groups	15	32.833	2.189	
Total	17	80.278		

$^* p < .05$

While Table 2 contains many values, which were used to arrive at the probability in the footnote to the table, the typical consumer of research is only interested in the end result, which is the value of *p*. As you know, when the probability is .05 or less (as it is here), we reject the null hypothesis. This means that the *entire set* of differences is statistically significant at the .05 level. Note that the ANOVA does *not* tell us which of the three differences we listed are significant. In fact, it could be that only one, *or* only two, *or* all three are significant. This needs to be explored with additional tests known as *multiple comparisons tests*. There are a number of such tests based on different assumptions, which usually yield the same result. (There is still some controversy over which multiple comparisons test is most appropriate.) For the data we are considering, application of Scheffé's test (a popular multiple comparisons test) yields these probabilities:

1. for Drug C vs. A, $p < .05$
2. for Drug C vs. B, $p < .05$
3. for Drug B vs. A, $p > .05$

Thus, we have found that Drug C is significantly better than Drugs A and B because the probabilities are less than .05, but that Drugs B and A are not significantly different from each other because the probability is greater than .05.

In review, an ANOVA tells us whether a set of differences, *overall*, is significant. If so, we can use a multiple comparisons test to determine which individual pairs of means are significantly different from each other.

In this topic, we have been considering a **one-way ANOVA** (also known as a **single-factor ANOVA**). It is called this when the participants have been classified in only one way. In this case, they were classified only in terms of which drug they took (A, B, or C). In the next topic, we will consider the use of ANOVA when participants are classified in two ways.

[1] Because it yields a value of *F*, it is sometimes called an *F* test.