Why Are There Five Million Types of Statistics?

Mike Aamodt

Reprinted with permission of the *Assessment Council News* (ACN), December 1998, published by the International Personnel Management Association Assessment Council (IPMAAC), 1617 Duke Street, Alexandria, VA 22314, http://www.ipmaac.org.

Over the past few years, we have received several calls from HR folks asking about how to interpret one type of statistic or another. This article provides the big picture about statistical analysis and future issues will occasionally contain easy-to-understand explanations of specific statistics. If this idea excites you, seek professional help immediately!!!

Why Are There Five Million Types of Statistics?

Actually, it isn t that bad. Although one does encounter many different statistical analyses when reading research articles, they are done for one of only four reasons: describe data, determine if two or more groups differ on some variable, determine if two or more variables are related, or reduce data. The paragraphs below are intended to briefly explain these reasons.

To Describe Data

The most simple type of statistical analysis is conducted to describe a data set. For example, if an employee survey was conducted, one might want to report the number of employees who responded to each question (sample size or n), how the typical employee responded to each question (mean, median, mode), and the extent to which the employees answered the questions in similar ways (variance, standard deviation, range).

To Determine If Two or More Groups Differ on Some Variable

Once descriptive statistics are obtained, a commonly asked question is whether certain groups differ. For example, did females perform better in training than males? Were older employees as likely to accept the new benefit plan as their younger counterparts? To answer questions such as these, if our descriptive statistic was a mean, we might use a *t-test* if there were only two groups (e.g., male, female) or an *analysis of variance*

(ANOVA) if there were more than two groups (e.g., south, north, east, west) or more than two variables (e.g., race and gender). If the descriptive statistic was a frequency count, we might use a *chi-square*. Regardless of what statistic we use, the question is the same: Do our groups differ?

To Determine If Two or More Variables Are Related

Often a question asked in research is the extent to which two or more variables are related, rather than different. For example, we might ask if a test score is related to job performance, if job satisfaction is related to employee absenteeism, or if the amount of money spent on recruitment is related to the number of qualified applicants that apply. To determine if variables are related, we might use *correlation*. If we wanted to be a bit more precise or are interested in how several different variables predict performance, we might use *regression* or *causal modeling*.

To Reduce Data

At times we have lots of data that we think can be simplified. For example, we might have a 100-item questionnaire. Rather than separately analyzing all 100 questions, we think that the 100 questions represent five major themes/categories/factors. To reduce data we might use a *factor analysis* or a *cluster analysis*.

What's The Deal With Significance Levels?

Significance levels are one of the nice things about statistical analysis. If you are reading an article about the effectiveness of a new training technique, and don't care a thing about statistics, you can move through the alphabet soup describing the type of analysis used e.g., (ANOVA, t-test, MANOVA, ANCOVA) and go right to the significance level which will be written something like $p < .03$. What this is telling you is that the difference in performance between two or more groups (e.g., trained versus not trained) is significantly different at some level of chance.

Let me explain. Suppose that you walked into a training room and asked the people on the right side of the room how old they were and then did the same to people sitting on the left side of the room. You found that people on the right side of the room averaged 37.6 years whereas the people on the left side of the room averaged 39.3 years of age. Does this difference make you want to submit a paper on the subject? Could it be that older people sit closer to the door so they don't have to walk as much? Could be, but probably not. Anytime we collect data from two or more groups, the numbers will never be identical. The question becomes then, if the numbers are never identical, how much of a difference does it take before we can say that something is going on? This is where significance levels come in. Based on a variety of factors such as sample size and variance, the end result of any statistical analysis is a significance level that indicates the probability that our differences occurred by chance alone. If our analysis indicates that the groups differ at $p < .03$, we would conclude that there are 3 chances in 100 that the differences we obtained were the result of fate, karma, or chance. In the social sciences, we have a very dumb rule that if the probability is less than 5 in 100 ($p < .05$) that our results could be due to chance, we say that our results are "statistically significant."

So, Significance Levels Tell Us The Importance of Our Study?

Unfortunately, no. Significance levels only tell us if we are allowed to "pay attention" to our results. If our results are statistically significant, we get to talk about them (not that anyone will listen). If they are not statistically significant, we start again.

If our results are statistically significant, we then ask about the "practical significance" of our findings. This is usually done by looking at *effect sizes,* which can include d scores, correlations (r), omega-squared, and a host of other awful sounding terms. Effect sizes are important because we can obtain statistical significance with large sample sizes but have results with no practical significance. For example, suppose that we conduct a study with 1,000,000 people and find that women score an average of 86 on a math test and men score an average of 87. With such a big sample size, we would probably find the difference to be statistically significant. However, what would we conclude about the practical significance of a one-point difference between genders on a 100-point exam? Are men "superior" to women in math? Will we have adverse impact? Should I discourage my daughter from a career in science? Probably not. The statistical significance allows me to confidently say that there is little difference between men and women on this variable. If I compute an effect size, I can say this in a more precise way.

This Stuff is Awful, I Thought You Said It Would Be Easy?

No, I never said it would be easy. I just said that I would do my best to make statistics easier. When I was in graduate school, I took five statistics classes and thought I knew a lot about statistics. In fact, I had a professor who told us that "if you understand correlation, t-tests, and analysis of variance, you will be able to read 85% of all the published research articles." That was then. These days, doctoral students probably took five statistics courses before they graduated elementary school! I can't think of the last time I saw a simple ANOVA or correlation in the *Journal of Applied Psychology or Personnel Psychology.* I can't think of the last time I saw an article containing the name Schmidt or Hunter that I actually understood or that had more English words than statistical symbols! No, statistics is not an easy topic to master, but it is easy enough to understand so you can read journal articles and dress as a nerd next Halloween.

Author

Mike Aamodt is a professor of I/O Psychology at Radford University in Radford, VA. He is the author of *Applied Industrial/Organizational Psychology* (3rd ed) published in 1999 by Brooks/Cole. He operates a part- time consulting business and is active in such professional associations as IPMAAC, SIOP, SHRM, and the Society for Police and Criminal Psychology.