Module 4A: Test Statistics and Analysis

by Dr. Tim Brophy

Hello, my name is Tim Brophy and welcome back to Passport to Great Teaching. This is module 4 where we're going to discuss and introduce test statistics and analysis. So we have three goals for this module and this module consists of three short videos.

The first goal is to take some time to describe, explain, and calculate some descriptive statistics like mean, median, mode, variance, and standard deviation. Then we're going to review again validity and test reliability that we discussed in an earlier module in this series. And then we're going to take some time to really reinforce and analyze and interpret item difficulty and discrimination, and also test reliability. So let's get started.

In module 4a, I'll introduce to you descriptive test statistics. This may be a reminder or a refresher for some of you who already know some of these terms, but it doesn't hurt for us to take some time just to review them before we launch into the rest of the module. Let's talk first about populations versus samples.

Now a population usually refers to a score set that is a complete score set for an entire group of people that we may be testing their scores and analyzing those scores. Now, from that population we can draw samples. They can be larger or smaller depending on what the size of the sample is that we wish to take, and also its ability to represent the entire population. So it's very important that when we select samples, that they are actually representative of the larger population so that we can later make some inferences about the results we get statistically from the sample to the entire population.

Let's talk now about measures of central tendency. Some of these terms you've probably heard before, but the term "distribution" has to do with the entire set of scores that we may have for a population. So that's a distribution. It could also serve as the entire set of scores we have for a sample from a population.

Now, the measures of central tendency that we're going to define now are mean, median, and mode, terms you may already know. But remember the mean is the arithmetic mean of a distribution of numbers. So if we have a group of scores from a quiz or a test and we add them all up and divide by the number of scores that we have, then we get the mean or the average. Then there is the mode. That happens to be the most frequently occurring number within our distribution that we're studying. And the median is actually the midpoint of the distribution, which happens to be the point at which there are 50% above the median and 50% of the scores below the median.

Continuing with basic definitions, we know that scores on a test vary based on individual performance on the items on the test. So our students are going to produce different scores that vary widely or not so widely depending on the items that we're assessing them

on. So variance happens to be a numerical representation of the degree to which those scores vary. So the implication here is that the larger the variance number, the greater the differences are amongst the scores.

Now to calculate variance for a complete data set you'll use this formula. Now, we're not going to ask you to commit that to memory because many statistical programs will do this for you very easily, but you can see the key on the left-- excuse me-- on the right hand side of your screen that describes all of the different components of that particular formula.

Now to calculate for the variance for data sets that are subsets or samples of the larger population, then we use this formula. Now notice the denominator is different and the population sample-- or the population variance formula-- the large N represents the entire set of scores, whereas in the subset formula the variance is n minus 1, which means it's the sample minus 1 because there is 1 less degree of freedom for scores to vary in that sample. And that's all we need to talk about in that particular area at this time.

Continuing with our basic definitions, the standard deviation, all right, is the square root of the variance. So this could be calculated for the entire population or it can be calculated just for the sample. And many times we do this for a sample only and then use that standard deviation-- if our sample has been selected correctly-- to represent the population standard deviation, or at least to provide an inference of what that could be, all right. So what the statistic does is provide some information about the distribution of the scores around the mean and here's what it tells us.

For all score distributions, whether they be samples or populations, 68% of scores fall within one standard deviation of the mean and 95% of the scores will fall within two standard deviations from the mean. 99% will fall within three standard deviation of the mean, and so forth. Now, descriptive statistics can be calculated really very quickly and easily using these free online calculators. I've linked them here, you can bookmark those on your own website or your own computer and you'll have them to use whenever you need them.

So let's review quickly what these statistics tell us. First mean, median, and mode, these simply describe the score set that we're analyzing. These are sample dependent, they're only representative of the sample or the population that we are getting those statistics for. And they give you an idea also of the skewness of the score set, by skewness we mean the degree to which scores tend to be on the positive or the negative side of the mean, all right.

Then there's variation. This is the numerical representation of the degree to which the scores vary. So again, the implication here is that the larger of the variance number, the greater the differences among the scores. And then the standard deviation is the square root of the variance, and remember a 68% of the scores will fall within plus or minus one standard deviation of the mean.

So let's think and practice for just a minute. First of all, I'd like to ask you a question. What value does knowing the descriptive statistics for your quizzes and tests add to your teaching? All right. And then to practice this, use the score set-- using the score set in the online calculators I've provided you in the module resources for this, module 4a. You can calculate the statistics, all right, using the information that I provided. Or you can do this for a test that you've given yourself.

So the question here is, what do these tests statistics tell you about the scores on the assessment? So now that concludes module 4a and our review of these basic statistical terms, and we're going to move on now to module 4b.