History 595: Standard Error Calculations

Now that you have learned about the concept of the **standard error**, we want to use that new concept to reexamine some of the results of the SPSS runs that you have already done.

Remember that the **standard error** of a statistic from a **probability sample** provides a way to evaluate whether the **point estimate**, for example a mean for one group of cases, is really sufficiently different from another mean for a different group of cases that you can be sure that the difference isn't simply a function of the sample chosen.

That is, since we know that if we took another sample, we might get different means for both our groups. This variability between samples is called **sampling error**. It does NOT mean that the variability is a mistake. It simply means that the sampling process itself introduces some uncertainty or variability in the statistics. So we want to see if we can find a way to be sure that the any difference we have found is representative of patterns in the underlying **population**, not a function of our sampling procedure.

We now want to introduce the concept of **"statistical significance."** By that we mean that the results we have found has not occurred by **chance**, that is, there is a **low probability** that we would get a different result with repeated sampling.

Here's a more formal definition:

In statistics, a number that expresses the **probability** that the result of a given experiment or study could have occurred purely by **chance.** This number can be a **margin of error** ("The results of this public opinion poll are accurate to five percent"), or it can indicate a **confidence level** ("If this experiment were repeated, there is a probability of ninety-five percent that our conclusions would be substantiated").

We use the **"standard error"** to estimate the **interval** around our **point estimate** at a particular **probability level**, usually 95% in the social sciences. Since we know that the **sampling distribution** of a statistic calculated from random samples is **normal**, and the **standard error** represents the **standard deviation** of the **sampling distribution**, we can use the properties of the normal curve, the 68, 95, 99 percent rule to calculate the interval for the point estimate with 1, 2, or 3 standard errors.

This sounds complicated, but let's work on it in practice….