Margin of Error

When a sample is used to draw inferences about a larger population, there is always a possibility that the sample is *non-representative*, i.e. does not reflect the true attributes of the population. We call this sampling error. While selection bias is the most common cause of sampling error, it is also possible to get sampling error even when there is no discernible reason for selection bias. We are after all, dealing with a sample. It is possible to get sampling error due to *random fluctuation*.

Suppose we wanted to use a sample of 25 students to estimate the average GPA of all students. Now suppose we choose our sample by random sampling and by some wild fluke we end up with all 25 being members of Phi Kappa Phi (Honors Society). This is highly improbable, of course, but not impossible. Such a sample would give us an estimate for GPA that would be far too large. While we can never be certain that a sample is representative of the larger population, we can calculate probabilities for getting sampling error due to random fluctuation.

Confidence

A *confidence level* is essential a probability that we have chosen a sample whose sample mean is within some *margin of error* of the true population mean. This gives us a precise way of making estimates based on a sample.
Example: Based on a sample of 200 households, the average household income for Toledo is $32,315 with a margin of error of $756 at a 95% confidence level.
Interpretation: There is a 95% probability that the true mean household income for all Toledo families is within $756 of $32,315.

Confidence levels are typically 90% or higher. In practice, the confidence level is chosen and the the margin of error is calculated from the formula

$$E = \frac{\sigma z_{\alpha/2}}{\sqrt{n}}$$

where $z_{\alpha/2}$ is calculated from the confidence level and the normal distribution, σ is the population standard deviation and n is the sample size. Often the sample standard deviation s is used in lieu of an unknown population standard deviation. Note that the sample size n appears in the denominator. This tells us that small samples tend to produce large margins or error and larger samples produce smaller margins.

Often we see research which compares a sample mean to a known population mean. If the sample mean differs enough from the population mean, we conclude that there must be some systematic cause for that difference. That's often the point, to provide evidence for some cause/effect relationship. Generally, in order to be convincing, the difference between the population mean and the sample mean must be *more than the margin of error* at a reasonable confidence level.

Example: Composite scores on the ACT test have a mean of 21.1 and a standard deviation of 4.8 nationwide. A high school principal reports that based on the scores of 70 recent graduates, average scores for his school are 22.0 with a margin of error of 1.124 at a 95% confidence level. He claims therefore that his school is doing a better than average job of preparing students for the ACT. Should we believe him?

1. Yes
2. No

Notice that the population mean of 21.1 is well within the margin of error (1.124) of the sample mean of 22.0.

$$22.0 - 1.124 = 20.876 < 21.1$$

In a case like this where the population mean is within a reasonable margin of error of the sample mean, the difference between the two may be due to random fluctuation, and not to any systematic cause (like superior teaching).

Confidence Interval

A *confidence interval estimate* is another way of conveying information about the sample mean. A confidence interval is the mean plus or minus the margin of error, that is

$$(\bar{x} - E, \bar{x} + E)$$

For example, the mean of 22.0 and margin of error of 1.124 might be reported as

$$(22.0 - 1.124, 22.0 + 1.124) = (20.876, 23.124)$$

Note that the population mean of 21.1 falls within this interval.

Note that the margin of error and confidence interval estimates do not in any way account for selection bias. If we have reason to believe there was some form of systematic bias in selecting the sample, we simply cannot trust the results.