Business Statistics Mr. Nelson

INFERENCE CONDITIONS EXPLAINED FURTHER

RETURN *to Example Step Three*

*** Random Sample: A random sample was conducted to insure every member of the population was equally likely to be selected.**

Allows us to use probability theory to calculate the conditional probability that is known as the P-value.

*** Normal Sampling Distribution: The sampling distribution of all possible sample means has an approximately normal shape because the sample was of sufficient size, over 30 (per the Central Limit Theorem).**

A. Normal shape is required to allow us to calculate the P-value.

B. There are two ways to determine that the sampling distribution is normally shaped.

(1) The population distribution (from which the sample is drawn) has a normal shape. In this case, the size of the sample is not important. Even a small sample of only 2 or 3 drawn from a normal population distribution would have a normally shaped sampling distribution. This shape would be exactly normal in shape, not just approximately normal. The practical problem with this option is that most real world sampling is performed without knowing the shape of the population being studied.

OR

(2) The sample size is sufficiently large to insure that the sampling distribution will have an approximately normally shape even if the population distribution is highly skewed, a relationship established by the "Central Limit Theorem". Generally, a sample size of 30 or more is sufficient for us to employ the Central Limit Theorem and assume the sampling distribution has an approximately normal shape. Note that under this approach the sampling distribution is near normal but not exactly normal. With a sufficiently large sample size however, this very slight deviation form is not significant and does not alter the calculation of P-values described herein

*** Independence: The lack of replacement is not a problem in this case because the number of subjects in the population is more than ten times the sample size.**

Consider the lack of replacement and its impact on the probability calculations. The Indepedence condition has many sides, but most significant is the procedural fact that we do not replace a member of the population once they have been selected in the sample. Without replacement, each selection reduces the denominator of the probability for each successive selection by one. To illustrate, examine the example below where a sample of three is drawn from a population of ten. Here the change of probability from one selection to the next is too great for us to assume independence in our calculations of P-value. In particular, it violates the Multiplication Rule.

First Selection: Every Member of Population Has Probability of $\frac{1}{10}$

Second Selection: Every Member of Population Has Probability of $\frac{1}{9}$

Third Selection: Every Member of Population Has Probability of $\frac{1}{8}$

In the example below, the differences in probability from selection to selection is too significant for us to overlook and apply the Multiplication Rule. Now consider an example where the size of the population (assumed to be 30) is ten times the sample size (assumed to be 3).

First Selection: Every Member of Population Has Probability of $\frac{1}{30}$

Second Selection: Every Member of Population Has Probability of $\frac{1}{29}$

Third Selection: Every Member of Population Has Probability of $\frac{1}{28}$

In this second example, the lack of replacement causes the probabilities for the three selections to be different but by a much smaller degree than the first example. These relatively small differences in probability are not significant, and it allows us to use the Multiplication Rule in calculating P-values. In most cases, one can determine that the population size is at least ten times the sample size to meet the Independence condition for one sample tests.