**Sampling Distributions and the Central Limit Theorem**

Instructions: Refer to the accompanying spreadsheet as you read through each step of the worksheet. Complete the problems when indicated.

1. Suppose we know the underlying probability distribution for a population that only contains three possible observations. The probability distribution is given in the spreadsheet.

**PROBLEM**: Calculate the expected value of the population by completing the table and summing up the last columns

2. Suppose we are trying to estimate the mean of this population by making a sample of two observations. Because the observations are random, we can make any combination of observations, which all lead to different values of the sample mean. We will first enumerate every possible sample that can be drawn from this population by making a table of the possible observations.

3. Because each observation is independent of the subsequent observation (by assumption), we can calculate the probability of getting a specific sample as the product of probabilities of getting those particular observations:

Let **x**1 and **x**2 the first and second observations made from this population, then p(**x**1) \* p(**x**2) represents the probability of observing the sample {**x**1, **x**2 } in that specific order. p(**x**2) \* p(**x**1) would represent the probability of observing the sample {**x**2, **x**1 } .

**PROBLEM:** Complete the table on the spreadsheet by finding the probability of observing each particular sample. Examine the formulas in the first row and try to apply the logic to the next row in the table. You will not be able to drag any formulas. Each entry will have to be hand-typed.

4. Notice some of the samples produce equal sample means. We need to find the probability distribution of the sample mean by collecting all the values of the sample mean and their corresponding probabilities. The probability of observing a sample mean of a certain value is simply the probability in the fourth column of the previous table.

**PROBLEM:** Createa new table by listing all possible values of the sample mean and then their corresponding probabilities. Add together the probabilities of samples whose sample means yield the same value.

5. Now that we have the probability distribution of the sample mean, let's find some characteristics of the sample mean's distribution, namely it's expectation and its variance.

**PROBLEM:** Create a histogram of the sample mean's probability distribution. Comment on its shape. Is it symmetric? Skewed?

Next, create another histogram, this time of the population itself.

Next, calculate the expectation of the sample mean and the variance of the sample mean. Compare and contract the results against the actual expectation and variance of the population.

6. Switch the second spreadsheet and answer the questions found there concerning the Central Limit Theorem and sampling distributions.