Stat 1 Pal Worksheet 22: Sampling Distribution of the Sample Mean (Continued)

Name_____

In the previous worksheet you were introduced to the random variable *X* = the number of pets of the entire population of *N* students at a large university.

X	p(*x*) (or P(*X* = *x*))
0	0.6
1	0.3
2	0.1

a. Students at this university own μ_X = 0.5 pets, and that the standard deviation of the number of pets they own is σ_X = 0.6708. The distribution of the number of pets own by these students is skewed right.
b. The population $\bar{X}$ of ${}_NC_2$ possible means resulting from averaging the number of pets own by *n* = 2 students at a time has mean $\mu_{\bar{X}}$ = 0.5 and standard deviation $\sigma_{\bar{X}} = \sigma_X/\sqrt{2}$ = 0.4739.
c. The distribution of the population of sample means $\bar{X}$ is closer to a bell-shape in comparison to the distribution of *X*.

1. Repeat the work you did in the previous worksheet by using now samples of n = 3 students at a time: the number of all possible samples you could get with *n* = 3 students from the population of *N* students would be ${}_NC_3$. Suppose the number of students in this University is *N* = 20,000 and obtain ${}_{20,000}C_3$. That is, the *population of all possible samples* of *n* = 3 students that can be made with the *population of N = 20,000 students* is ${}_{20,000}C_3$.

2. Even though the population of possible samples of *n* = 3 students from this university is very large (${}_NC_3$), there are only 27 possible pairs of answers you would get when asking any two students the number of pets they have (each of which would repeat itself several times over the entire population of ${}_NC_3$ samples of 3 students). Compute the averages of the first 8 of these possible different answers would yield in the following table.

Answer of 1st Student X_1	Answer of 2nd Student X_2	Answer of 3rd Student X_3	Mean Number of Pets $\bar{X} = \frac{X_1 + X_2 + X_3}{3}$
0	0	0	
0	0	1	
0	1	0	
1	0	0	
0	1	1	
1	0	1	
1	1	0	
1	1	1	

2	2	2	
Etc.	Etc.	Etc.	Etc.

3. In the above table that if you tried to list the population of ${}_NC_3$ sample means $\bar{x}$ resulting when asking each possible group of 3 students the number of pets they own, there are only a few possible different values (27) for these sample means. List all 4 of them:

4. Obtain the proportion of times that each of the 4 different $\bar{x}$'s you obtained above appear in the population of ${}_NC_3$ $\bar{x}$'s:

Mean Number of Pets $\bar{X} = \frac{X_1 + X_2 + X_3}{3}$	p($\bar{x}$) (or P($\bar{X} = \bar{x}$)
Etc.	

5. The formulas mentioned in the previous worksheet, $\mu_{\bar{X}} = \mu_X$ and $\sigma_{\bar{X}} = \sigma_X/\sqrt{n}$ hold for samples of any size *n*. If you summarize the population of samples means $\bar{x}$'s you could get when averaging *n* values at a time from a population then: a. the average of the population of ${}_NC_n$ sample means is always equal to the mean of the population from where you are taking your samples of size n, and b. the standard deviation of the population of ${}_NC_n$ sample means is always smaller than the standard deviation of the values of the population from where you are taking your samples of size *n* by a factor of $1/\sqrt{n}$.

Therefore, note that you don't need to complete all the rows of the random variable $\bar{X} = \frac{X_2+X_2+X_3}{3}$ listed above to obtain its mean and standard deviation:

$\mu_{\bar{X}}$ =

$\sigma_{\bar{X}}$ =

6. In the previous worksheet that the shape of the distribution of $\bar{X} = \frac{X_2+X_2}{2}$ was closer to a bell-shape than the distribution of the number of pets own by the students at this university X. Based on this, how do you think the distribution of the random variable $\bar{X} = \frac{X_2+X_2+X_3}{3}$ looks like in comparison to the distribution of these other two random variables?

7. Suppose a very large number of students who did not know that on average all student at the university own μ_X = 0.5 pets, and each tried to estimate this value by averaging the number of pets own by a random sample of only three students. Will the proportion of students who get a sample mean $\bar{x}$ that is only 0.6 units (either above or below) from the population mean μ_X= 0.5 be smaller, larger or equal to the proportion of students in the previous worksheet who are also only 0.6 units off from the mean μ_X= 0.5 but who used samples of *n* = 2 students? Explain.