Problem 4 for Wednesday, March 22: This problem is intended to give you some idea of why the population correction factor (discussed in class Monday, March 8 – notes are on the web) is needed when we sample with replacement from a finite population that is not sufficiently larger than the sample. You will need to draw on your knowledge from probability.

To keep things relatively simple, we will just consider counts, and samples of size 2. So suppose we have the following situation:

We have collection of N individuals, with m "successes" and N-m "non-successes."

We consider samples of size 2.

X = the number of successes in a sample of size 2.

a. If we sample *with* replacement, X is a binomial random variable. For this part, use the usual formulas for mean and variance of a binomial random variable.

i. What is the mean of X?

ii. What is the variance of X?

b. Now suppose we sample *without* replacement. Then

$$P(X = k) = \frac{\binom{m}{k}\binom{N-m}{2-k}}{\binom{N}{2}}$$

i. Explain how this formula was derived. That is, explain what each part stands for and why the parts are arranged as they are.

ii. Use the formula to help you find E(X).

iii. How does your answer compare with your answer in part (i) of (a), when you sampled *with* replacement?

iv. Use the formula to help you find the Variance of X. (Probably the formula $Var(X) = E(X^2) - [E(X)]^2$ is the easiest way to do this.

v. Compare your answer with your answer in part (ii) of (a), when you sampled *with* replacement.

vi. Based on your answers to (ii) of part (a) and (iv) of part (b), what is the Finite Population Correction Factor for variance in this situation?