Math 262

Section 2.6.1 Day 14

1. Suppose that in a batch of 20 items, 3 are defective. If 5 of the items are sampled at random:

   (a) What is the probability that none of the sampled items are defective?

   (b) What is the probability that exactly 1 of the sampled items is defective?

   (c) What is the probability that exactly 4 of the sampled items are defective?

   (d) On average, how many defective items will be found in a random sample of 5 items?

   (e) What is the probability that the number of defective items sampled is within 2 standard deviations of the mean number?

2. Let X be a hypergeometric random variable with parameters n, M, N. Let Y be a Binomial random variable with parameters n and $p = \frac{M}{N}$. How does $E(X)$ compare to $E(Y)$? How does $\text{Var}(X)$ compare to $\text{Var}(Y)$?

3. Urn 1 contains 100 balls, 10 of which are red. Let X_1 be the number of red balls in a random sample of size 50 from Urn 1. Urn 2 contains 100 balls, 50 of which are red. Let X_2 be the number of red balls in a random sample of size 10 from Urn 2.

   (a) Use technology to compute the pmf of X_1. Display the values as a list or a table. Then do the same for the pmf of X_2. What do you notice?

(b) Change the numbers 100, 10, and 50 in this problem and recompute the pmfs of X_1 and X_2. What do you notice?

(c) Make a conjecture about when two hypergeometric random variables have the same pmf.
BONUS: Prove your conjecture.

4. An unnkown number, N, of animals inhabit a certain region. To estimate the size of the population, ecologists perform the following experiment: They first catch M of these animals, mark them in some way, and release them. After allowing the animals to disperse throughout the region, they catch n of the animals and count the number, X, of marked animals in this second catch.

   The ecologists want to make a *maximum likelihood estimate* of the population size N. This means that if the observed value of X is x, then they estimate the population size to be the integer N that maximizes the probability that $X = x$. Help them complete this estimate as follows.

   (a) What assumptions are necessary to say that X has a hypergeometric distribution?

   (b) Let $P_x(N)$ be the probability that $X = x$ given that $X \sim \text{Hypergeometric}(n, M, N)$. Write down a formula for $P_x(N)$.

   (c) Simplify the ratio $\dfrac{P_x(N)}{P_x(N-1)}$. *Hint*: use `FullSimplify` in Mathematica!

   (d) Show that $\dfrac{P_x(N)}{P_x(N-1)} \geq 1$ if and only if $N \leq \dfrac{Mn}{x}$.

   (e) Explain why $P_x(N)$ attains its maximum value when N is the largest integer less than or equal to $\dfrac{Mn}{x}$. What is the most likely population size N?

   (f) If $M = 30$, $n = 20$, and $x = 7$, what is the maximum likelihood estimate for N?