From last time:

3. An unknown 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.

$X \sim \text{Hypergeometric}(n, M, N)$

Let $P_x(N)$ be the probability that $X = x$ given that N is the true population size.

So: $P_x(N) = \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ and $\dfrac{P_x(N)}{P_x(N-1)} = \dfrac{(N-M)(N-n)}{N(N+x-M-n)}$

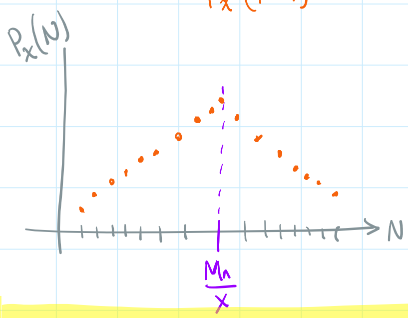
$$\frac{P_x(N)}{P_x(N-1)} \geq 1 \quad \text{iff} \quad \frac{(N-M)(N-n)}{N(N+x-M-n)} \geq 1$$

$$\text{iff} \quad (N-M)(N-n) \geq N(N+x-M-n)$$

$$\cancel{N^2} - \cancel{MN} - \cancel{Nn} + \underline{Mn} \geq \cancel{N^2} + \underline{Nx} - \cancel{NM} - \cancel{Nn}$$

$$\text{iff} \quad Mn \geq Nx$$

$$\frac{P_x(N)}{P_x(N-1)} \geq 1 \quad \text{iff} \quad \boxed{\frac{Mn}{x} \geq N}$$



If $N \leq \frac{Mn}{x}$, then $P_x(N) \geq P_x(N-1)$, so population size N is more likely than population size $N-1$.

If $N > \frac{Mn}{x}$, then population size N is less likely than population size $N-1$

THUS: The most likely population size is the largest integer N that is less than or equal to $\frac{Mn}{x}$.

(f) $M = 30, \quad n = 20, \quad x = 7$

the max. likelihood estimate for N is:

$$\frac{Mn}{x} = \frac{30(20)}{7} \approx 85.7 \qquad \text{so} \quad \underline{N = 85}$$

NEGATIVE BINOMIAL DISTRIBUTION

An experiment consists of a sequence of independent trials. Each trial results in either "success" or "failure." The probability of success is p for each trial. The experiment stops when a certain number, r, of successes have occurred. Let X be the number of trials necessary to achieve r successes.

Then $X \sim \textbf{NegativeBinomial}(r, p)$

pmf: $$P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

If $r = 1$, then $X \sim \textbf{Geometric}(p)$

pmf: $$P(X = x) = \underbrace{(1-p)^{x-1}}_{\text{prob. of } x-1 \text{ failures}} \underbrace{p}_{\text{prob. of success on trial } x}$$