1. a) Suppose that a sample of 3 people are to be chosen at random from a group of 3 women and 2 men. Determine the probability than women outnumber men in the sample.

There are $\binom{5}{3} = 10$ ways to choose the sample. Women outnumber men if all 3 women are chosen, which can happen in 1 way, or if 2 women and 1 man are chosen, which can happen in $\binom{3}{2} \binom{2}{1} = 6$ ways. So, the probability that women outnumber men is $7/10 = .7$.

b) Repeat for choosing a sample of 3 people from a group of 6 women and 4 men.

Now there are $\binom{10}{3} = 120$ ways to choose the sample. Women outnumber men if all 3 women are chosen, which can happen in $\binom{6}{3} = 20$ ways, or if 2 women and 1 man are chosen, which can happen in $\binom{6}{2} \binom{4}{1} = 60$ ways. So, the probability that women outnumber men is $80/120 \approx .667$.

Now suppose that a sample of 3 people are to be chosen at random from a population with $3k$ women and $2k$ men, where $k$ is a positive integer.

c) Express the probability than women outnumber men in the sample, as a function of $k$.

Now there are $\binom{5k}{3}$ ways to choose the sample. Women outnumber men if all 3 women are chosen, which can happen in $\binom{3k}{3}$ ways, or if 2 women and 1 man are chosen, which can happen in $\binom{3k}{2} \binom{2k}{1}$ ways. So, the probability that women outnumber men is $\frac{\binom{3k}{3} + \binom{3k}{2} \binom{2k}{1}}{\binom{5k}{3}}$.

d) Produce a graph of this probability as a function of $k$, for values of $k$ from 1 to 100.
e) Report the values of this probability, to five decimal places, for $k = 1, 2, 10, 100, 1000, 1$ million.

$k = 1$: prob = .70000  
$k = 2$: prob = .66667  
$k = 10$: prob = .65102  
$k = 100$: prob = .64829  
$k = 1000$: prob = .64803  
$k = 1,000,000$: prob = .64800

f) Summarize what these calculations reveal about the role of the population size on this probability.

Notice that women always comprise 60% of the population. We see that population size has very little impact on this probability once the population size exceeds a few hundred, as the probability quickly approaches and converges to .648.

2. A variation of the capture/recapture technique has been applied to the problem of correcting for undercount when conducting a census. Suppose that when an initial census count is done in a certain community, it counts 851 residents. Then an intensive follow-up survey of 100 residents is conducted, and it turns out that 89 of them were counted in the initial census. Let $N$ be the actual total number of residents in the city.

a) Express the probability that the follow-up survey of 100 would have found 89 who had been counted the first time, as a function of $N$. [Hint: Be clear about specifying the input values of $N$ for which your function applies.]

$$
\Pr(89 \text{ of } 100 \text{ in } 2^{nd} \text{ sample would have been included in } 1^{st} \text{ sample}) \\
= \frac{\binom{851}{89} \binom{N-851}{11}}{\binom{N}{100}} \text{ for } N = 862, 863, \ldots
$$

b) Graph this probability as a function of $N$. 

![Graph of probability distribution](image.png)
c) Identify the value(s) of $N$ that maximize this probability, and also report the value of that probability.

This probability function is maximized at $N = 956$, for which this probability (to 5 decimal places) is .13371.

_Not to hand in:_ Consider working on exercises 1-4, 6-13, 16-18 at the end of section 1.8 in the text.