Question 10

Assuming : Number of people infected is distributed uniformly. Suppose no tests have been conducted yet and we do not know anything about covid. We have assumed that it is equal for the number of +ve cases to be 1 or 100 or 199.

1. Here we start by assuming some m in [k,n-p-k] where p is the population tested, n is the total population, k is the observed positives.  
   Now we calculate the probability that for these m +ves out of n, k appear in the selected p.  
   And we can say that this is proportional to the probability of m being chosen given k (using Bayes' Rule).  
   And so we have a distribution for P(POS=m|k).  
   The **estimate will be the most probable value**   
   The **expectation will be sum(m\*P(POS=m|k))** for all valid m.
2. And using the distribution, we can calculate P(POS>exp cases+1|k).

As we increase n from 200 to 400 and sample size from 10 to 20, (POS>exp cases+1|k) for k=2 decreases. This means that our confidence interval increases. This can be seen because the E(POS) increases as well.



