Assignment 12

Gautam Singh (CS21BTECH11018)

May 26, 2022

Outline

Problem Statement

(Papoulis/Pillai Exercise 5-52) A box contains n white and m black marbles. Let X represent the number of draws needed for the rth white marble.

1. If sampling is done with replacement, show that X has a negative binomial distribution with parameters r and $p = \frac{n}{m+n}$.
2. If sampling is done without replacement, then show that for $r \leq k \leq m+n$,

$$\Pr(X = k) = \binom{k-1}{r-1} \frac{\binom{m+n-k}{n-r}}{\binom{m+n}{n}} \tag{1}$$

3. For a given k and r, show that the probability distribution in (1) tends to a negative binomial distribution as $n + m \to \infty$. Thus, for large population size, sampling with or without replacement is the same.

Solution

1 Suppose $X = k$, $r \leq k \leq m + n$. Then, the k^{th} marble drawn must be white. The other $r - 1$ white marbles can be drawn in any of the previous draws, and we must also have $k - r$ black marbles drawn. Since marbles are drawn with replacement, the probability of drawing a white marble at any instant is given by

$$p = \frac{n}{m + n} \tag{2}$$

We can now get the PMF of X (here, $q = 1 - p$)

PMF of X, With Replacement

$$\Pr(X = k) = \begin{cases} \binom{k-1}{r-1} p^r q^{k-r}, & r \leq k \leq m + n \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

2 If sampling is done without replacement, the value of p in (2) keeps changing with each draw. We then write out the PMF as follows

PMF of X, Without Replacement

For $r \leq k \leq m+n$,

$$\Pr(X = k) = \binom{k-1}{r-1} \frac{[n \dots (n-r+1)][m \dots (m-k+r-1)]}{[(m+n) \dots (m+n-k+1)]} \tag{4}$$

$$= \binom{k-1}{r-1} \frac{n!}{(n-r)!} \frac{m!}{(m+r-k)!} \frac{(m+n-k)!}{(m+n)!} \tag{5}$$

$$= \binom{k-1}{r-1} \frac{\binom{m+n-k}{n-r}}{\binom{m+n}{n}} \tag{6}$$

$$\tag{7}$$

Therefore, we can write the PMF of X completely as follows

$$\Pr\left(X = k\right) = \begin{cases} \binom{k-1}{r-1}\frac{\binom{m+n-k}{n-r}}{\binom{m+n}{n}}, & r \le k \le m+n \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

3 Letting $m + n \to \infty$, and using (2) and (4),

$$\Pr\left(X = k\right) \approx \binom{k-1}{r-1}\frac{n^r m^{k-r}}{(m+n)^k} \tag{9}$$

$$= \binom{k-1}{r-1}p^r q^{k-r} \sim NB(r, p) \tag{10}$$

as desired. The limit is verified in `codes/12_1.py`.