# UGC/MATH 2018 (Dec math set-a), Q.116

T.Rohan - CS20BTECH11064

9th April, 2021

# Question

Suppose n units are drawn from a population of N units sequentially as follows. A random sample $U_1, U_2, ...U_N$ of size N, drawn from $U(0,1)$.The kth population unit is selected if $U_k < \frac{n-n_k}{N-k+1}$, k = 1, 2, ..N, where, $n_1 = 0, n_k =$ number of units selected out of first k-1 units for each k = 2, 3, ..N. Then,

1. The probability of inclusion of the second unit in the sample is $\frac{n}{N}$.
2. The probability of including the first and the second unit in the sample is $\frac{n(n-1)}{N(N-1)}$
3. The probability of not including the first and including the second unit in the sample is $\frac{n(N-n)}{N(N-1)}$
4. The probability of including the first and not including the second unit in the sample is $\frac{n(n-1)}{N(N-1)}$

## Explanation

- There are N units of population, out of which n are to be picked.
- From a uniform distribution in (0,1), N units are picked sequentially.
- kth unit corresponds to the kth person in the population.
- The kth person is picked if and only if the condition $U_k < \frac{n-n_k}{N-k+1}$ is satisfied. Where, $n_k$ are the number of units picked out of the first k-1 units.
- Let $X$ be a random variable, with $X \in \{1, 2, \ldots, N\}$, where X=i, when ith unit is included
- For the inclusion of the first unit in the sample, the condition is $U_1 < \frac{n-n_1}{N-(1)+1}$. Here $n_1 = 0$ is given in the question. Hence $X = 1$ if $U_1 < \frac{n-(0)}{N}$.

$$\therefore \Pr(X = 1) = \frac{n}{N} \tag{1}$$

# Option 1

The probability of inclusion of the second unit in the sample is $\frac{n}{N}$.

# Solution(Option 1)

For k=2, there are two cases. The first unit being included and excluded. When first unit is included, $n_2 = 1$,

$$U_2 < \frac{n - n_2}{N - 2 + 1} = \frac{n-1}{N-1} \tag{2}$$

$$\therefore \Pr(X = 2 \mid X = 1) = \frac{n-1}{N-1} \tag{3}$$

$$\Pr(X = 1, X = 2) = \Pr(X = 1) \times \Pr(X = 2 \mid X = 1) \tag{4}$$

From (1) and (3)

$$\therefore \Pr(X = 1, X = 2) = \frac{n(n-1)}{N(N-1)} \tag{5}$$

## Solution(Option 1) Contd.

When first unit is not included, $n_2 = 0$,

$$U_2 < \frac{n - n_2}{N - 2 + 1} = \frac{n}{N - 1} \tag{6}$$

$$\therefore \Pr(X = 2 \mid X \neq 1) = \frac{n}{N - 1} \tag{7}$$

$$\Pr(X \neq 1, X = 2) = \Pr(X \neq 1) \times \Pr(X = 2 \mid X \neq 1) \tag{8}$$

$$\tag{9}$$

From (1) and (7)

$$\therefore \Pr(X \neq 1, X = 2) = 1 - \frac{n}{N} \times \frac{n}{N - 1} = \frac{n(N - n)}{N(N - 1)} \tag{10}$$

From (5) and (10)

$$\Pr(X = 2) = \frac{n(n - 1)}{N(N - 1)} + \frac{n(N - n)}{N(N - 1)} = \frac{n}{N} \tag{11}$$

Hence, option 1 is correct.

# Option 2

The probability of including the first and the second unit in the sample is $\frac{n(n-1)}{N(N-1)}$

# Solution(Option 2)

From (5)

$$\Pr(X = 1, X = 2) = \frac{n(n-1)}{N(N-1)} \tag{12}$$

Hence, option 2 is correct.

# Option 3

The probability of not including the first and including the second unit in the sample is $\frac{n(N-n)}{N(N-1)}$

# Solution(Option 3)

From (10)

$$\Pr\left(X \neq 1, X = 2\right) = \frac{n(N-n)}{N(N-1)} \tag{13}$$

Hence, option 3 is correct.

# Option 4

The probability of including the first and not including the second unit in the sample is $\frac{n(n-1)}{N(N-1)}$

# Solution(Option 4)

$$\Pr(X = 1, X \neq 2) = \frac{n}{N} \times 1 - \frac{n}{N} = \frac{n(N-n)}{N^2} \tag{14}$$

Hence, option 4 is incorrect.
Therefore, Options 1, 2, 3 are correct