GATE BM 2023 Quetion 9

Snehil Singh EE22BTECH11050*

Question :
Out of 1000 individuals in a town, 100 unidentified individuals are covid positive. Due to lack of adequate covid-testing kits, the health authorities of the town devised a strategy to identify these covid-positive individuals. The strategy is to:

1) Collect saliva samples from all 1000 individuals and randomly group them into sets of 5.
2) Mix the samples within each set and test the mixed sample for covid.
3) If the test done in (ii) gives a negative result, then declare all the 5 individuals to be covid negative.
4) If the test done in (ii) gives a positive result, then all the 5 individuals are separately tested for covid.

Given this strategy, no more than testing kits will be required to identify all the 100 covid positive individuals irrespective of how they are grouped

1) 700
2) 600
3) 800
4) 1000

(GATE BM 2023)

Solution :
Given :

Parameter	Value
Number of individuals	1000
Strength of each group	5
Number of groups	200
Number of Covid positive individuals	100

At first we ll have to examine all the groups and then test all 5 individuals of each group who tested positive.

Using binomial expansion to find the probability of a group testing positive

Parameter	Value	Expression
N	1000	Total number of individuals
n	5	Number of individuals in group
p	0.1	Probability of individual testing positive
q	0.9	Probability of individual testing negative

Probability of a group of 5 with i covid negative individuals is given as
the PMF:

$$Pr(X = i) =^n C_i(q)^i(p)^{n-i} \tag{1}$$

$$=^5 C_i(0.9)^i(0.1)^{5-i} \tag{2}$$

Let us assume X is a random variable which denotes the number of individuals tested negative in a group.
Probability of group having atmost k negative tested individuals
the CDF :

$$F_X(k) = Pr(X \leq k) \tag{3}$$

$$= \Sigma_{i=0}^k (^n C_i) (q)^i(p)^{n-i} \tag{4}$$

$$= \Sigma_{i=0}^k (^5 C_i) (0.9)^i(0.1)^{5-i} \tag{5}$$

$$\tag{6}$$

If a group is tested positive, atleast one of the 5 individuals should be covid positive which means atmost
4 individuals are covid negative.
Finding probability of a group testing positive

$$Pr = F_X(4) \tag{7}$$

$$= \Sigma_{i=0}^4 (^5 C_i) (0.9)^i(0.1)^{5-i} \tag{8}$$

$$= 0.4096 \tag{9}$$

Parameter	Value
Number of tests in first round of testing	A
Number of groups tested positive in first round	b
Number of tests in second round of testing	B
Total number of tests in both rounds	T

In the first round we will test all the 200 groups.
Number of tests in first round(A) :

$$A = 200$$

Out of these 200 tests,assuming the estimated number of groups tested positive is b
If a group is tested positive, atleast one of the 5 is covid positive, which impies the group can only have
atmost 4 covid negative individual.
Calculating b :

$$\text{Pr(of group testing positive)} = \frac{b}{A} \tag{10}$$

$$F_X(4) = \frac{b}{A} \tag{11}$$

$$\implies b = 81.92 \tag{12}$$

In all the groups tested positive, atleast one individual is covid positive, to find all the covid positive
individuals, we will test all the 5 individuals of the positive groups seperately in round 2.

So estimated number tests in round 2 (B) :

$$B = 5b \tag{13}$$
$$= (5)(81.92) \tag{14}$$
$$= 409.6 \tag{15}$$

Total number of tests estimated (T) :

$$T = A + B \tag{16}$$
$$= 200 + 409.6 \tag{17}$$
$$= 609.6 \tag{18}$$

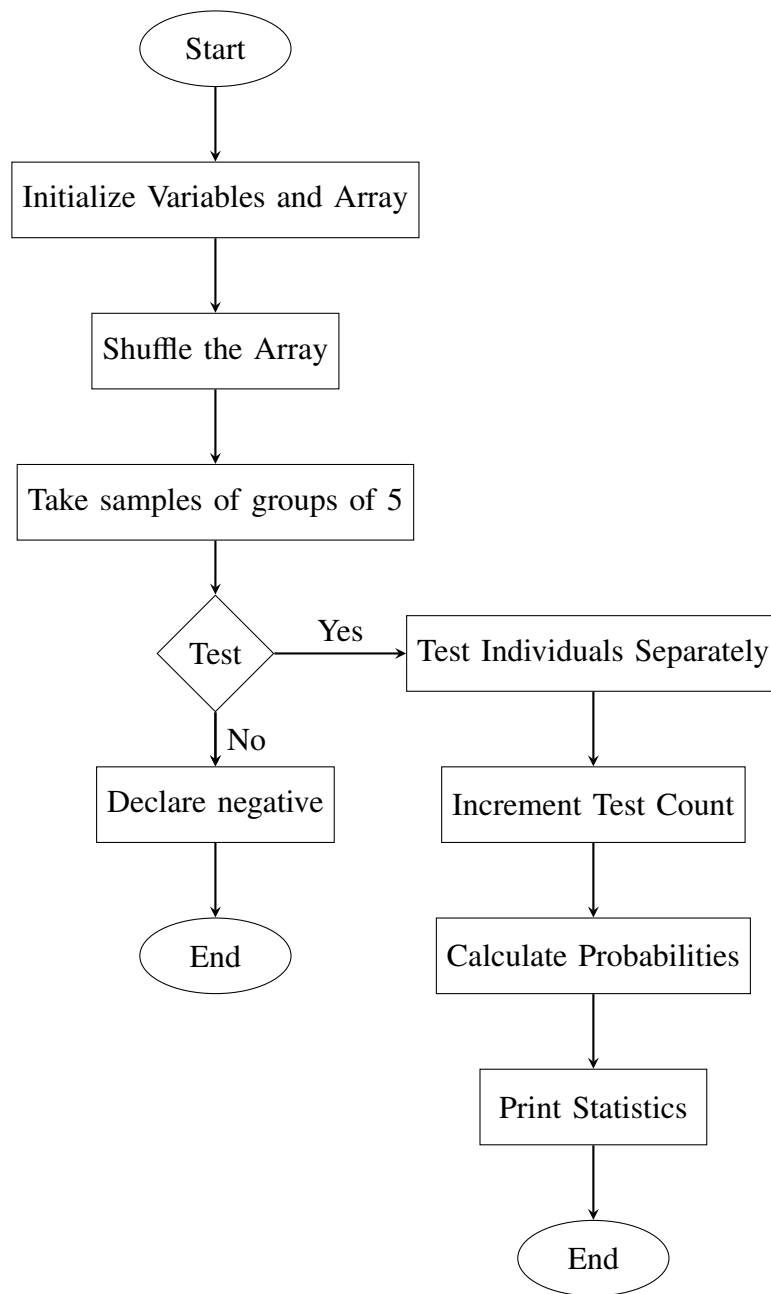$\therefore$ So 609.6 is the expected number of tests required to detect all the covid positive individuals. According to the options, no more than 700 tests will be required.
So, the correct answer is option A 700.

Simulation :

1) Including standard C libraries - `stdio.h` (for input/output), `stdlib.h` (for dynamic memory allocation), `time.h` (for time-related functions), and `math.h` (for mathematical functions).

2) Defining constants for the total number of individuals (`TOTAL_INDIVIDUALS`), the number of initially positive individuals (`POSITIVE_INDIVIDUALS`), and the group size used in testing (`GROUP_SIZE`).

3) Declared a function named `covidTestingStrategy` that simulates a COVID-19 testing strategy using group testing. The function takes an array of individuals, the total number of individuals, and the number of initially positive individuals as parameters.

4) Initialized variables `testCount` and `groupTestTrueCount` to keep track of the total number of tests and the number of group tests that turned out to be true and shuffled the array of individuals randomly using the Fisher-Yates shuffle algorithm to have randomness into the order of testing.

5) We mixed the samples within each group using Bitwise "OR" and tested them.

   a) If the mixed sample is negative (`groupResult == 0`), declares all 5 individuals in the group as COVID negative with one test.
   b) If the mixed sample is positive, tests each individual separately, and increments the test count accordingly.

6) Calculated the probability of the group test being true based on the initial probability of an individual being positive and the group size and Calculated the expected number of tests for one set, total expected number of sets, and total expected number of tests based on the calculated probabilities.

7) Printed various statistics related to the testing strategy.

8) Seed the random number generator with the current time. Initialized an array representing individuals, marking the first `POSITIVE_INDIVIDUALS` as positive. Calls the `covidTestingStrategy` function with the array and other parameters.

Flowchart to explain the function `covidTestingStrategy` is given below :

```
                    ┌─────────────┐
                    │    Start    │
                    └──────┬──────┘
                           │
            ┌──────────────▼──────────────┐
            │  Initialize Variables and Array │
            └──────────────┬──────────────┘
                           │
                  ┌────────▼────────┐
                  │ Shuffle the Array │
                  └────────┬────────┘
                           │
            ┌──────────────▼──────────────┐
            │  Take samples of groups of 5 │
            └──────────────┬──────────────┘
                           │
                       ◇ Test ◇ ── Yes ──► Test Individuals Separately
                           │                        │
                          No                        │
                           │                        ▼
                  Declare negative          Increment Test Count
                           │                        │
                           ▼                        ▼
                          End               Calculate Probabilities
                                                    │
                                                    ▼
                                             Print Statistics
                                                    │
                                                    ▼
                                                   End
```

Start

Initialize Variables and Array

Shuffle the Array

Take samples of groups of 5

Test

Yes → Test Individuals Separately

No

Declare negative

End

Increment Test Count

Calculate Probabilities

Print Statistics

End

```
10 // Custom power function for integer exponentiation
11 double customPow(double base, int exponent) {
12     double result = 1.0;
13
14     for (int i = 0; i < exponent; i++) {
15         result *= base;
16     }
17
18     return result;
19 }
20
21 // Function to simulate the testing strategy
22 void covidTestingStrategy(int individuals[], int total, int positive) {
23     int testCount = 0;
24     int groupTestTrueCount = 0;
25
26     // Shuffle the individuals randomly
27     for (int i = total - 1; i > 0; --i) {
28         int j = rand() % (i + 1);
29         // Swap individuals[i] and individuals[j]
30         int temp = individuals[i];
31         individuals[i] = individuals[j];
32         individuals[j] = temp;
33     }
34
35     // Perform testing in groups
36     for (int i = 0; i < total; i += GROUP_SIZE) {
37         int groupResult = 0;
38
39         // Mix the samples within each set
40         for (int j = i; j < i + GROUP_SIZE; ++j) {
41             groupResult |= individuals[j];
42         }
43
44         // Test the mixed sample for covid
45         if (groupResult == 0) {
46             // Declare all 5 individuals to be covid negative
47             testCount++;
48         } else {
49             // Test each individual separately if the mixed sample is positive
50             for (int j = i; j < i + GROUP_SIZE; ++j) {
51                 if (individuals[j] == 1) {
52                     // Individual is covid positive
53                     testCount++;
54                 }
55             }
56             groupTestTrueCount++;
57         }
58     }
59
60     // Probability of the group test being true
61     double p = (double)positive / total;
62     double groupTestTrueProbability = 1.0 - customPow(1.0 - p, GROUP_SIZE);
63
64     // Expected number of tests for one set
65     double expectedTestsOneSet = 1 + GROUP_SIZE * groupTestTrueProbability;
66
67     // Total expected number of sets
68     double totalExpectedSets = (double)total / GROUP_SIZE;
69
70     // Total expected number of tests
71     double totalExpectedTests = totalExpectedSets * expectedTestsOneSet;
```

Fig. 1: Input

Fig. 2: Output