Mathematics and Statistics Research Competition

Question 8

Jiamu Li & Frank Tang & Edward Wang

July 26, 2022

Contents

A particle generator is emitting two types of particles (called X and Y) into a long tube. The particles will line up in order after entering the tube. Initially, the tube is empty. At each shot, either an X- or Y-particle is randomly emitted into the tube with equal probability. Different shots are assumed to be independent from each other. Suppose that n shots have been emitted.

PROBLEM 1

- What is the probability that no two X-particles are next to each other?

THEOREM 1. *The probability that no two X-particles are next to each other after n shots is given by*

$$\frac{F_{n+2}}{2^n},$$

where F_n is the nth Fibonacci number.

Proof. The probability we require can be calculated by dividing the total number of ways to arrange the contents of the tube such that there are no consecutive X-particles, by the total number of arrangements of the particles. That is to say:

$$\Pr(\text{No consecutive X-particles}) = \frac{\#\text{Arrangements w/o consecutive X-particles}}{\#\text{Total arrangements}}$$

CLAIM 1.1. *The number of arrangements with no consecutive X-particles is*

$$\sum_{k=0}^{n} \binom{n-k+1}{k}.$$

Proof. Consider a tube with n particles in it. Let the number of X-particles be equal to k, so that the number of Y-particles is $n-k$. Consider the tube without the X-particles, consisting solely of Y-particles in a line:

$$\underbrace{\text{YY}\ldots\text{YY}}_{n-k} \tag{1}$$

Now consider the 'gaps' between these Y-particles, indicated by a bar (|):

$$|\text{Y}|\text{Y}|\ldots|\text{Y}|\text{Y}| \tag{2}$$

Notice that there are exactly $n-k+1$ 'gaps'. Clearly, if we were to only place X-particles in the gaps, then there would never be any consecutive X-particles. This can be done in a total of

$$\binom{n-k+1}{k}$$

ways. However, we must consider this for any number of X-particles k, so we arrive at the sum

$$\#\text{Arrangements with no consecutive X-particles} = \sum_{k=0}^{n} \binom{n-k+1}{k}. \qquad \square$$

CLAIM 1.2.

$$\sum_{k=0}^{n} \binom{n-k+1}{k} = F_{n+2},$$

where F_n is the nth Fibonnaci number.

Proof. Figure 1 showcases a way to obtain the identity from Pascal's triangle. We will present an algebraic proof below.

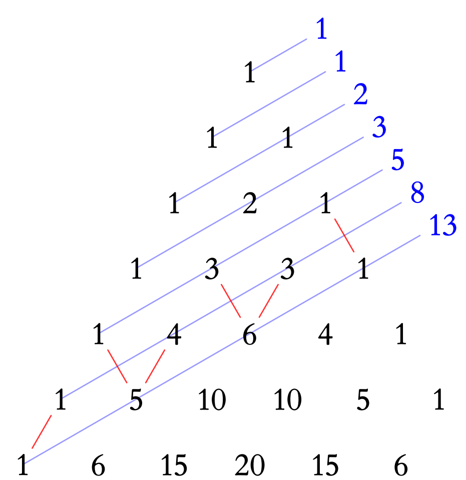

Figure 1: Each number in the row is the sum of the the numbers in the previous two rows.

Recall that the Fibonnaci numbers are defined as follows:

$$\begin{aligned} F_0 &= 0, \\ F_1 &= 1, \\ F_n &= F_{n-1} + F_{n-2}. && \text{(n > 1)} \end{aligned}$$

We will define a function $f(n) = \sum_{k=0}^{n} \binom{n-k+1}{n}$. It is sufficient to prove that $f(1) = F_3 = 2$, $f(2) = F_4 = 3$, and that $f(n) = f(n-1) + f(n-2)$, which would then imply the result by definition of the Fibonnaci numbers.

It is obvious that $f(1) = \binom{2}{0} + \binom{1}{1} = 2$, which is equal to F_3. Next, $f(2) = \binom{3}{0} + \binom{2}{1} + \binom{1}{2} = 3$. Notice that we define $\binom{n}{k} = 0$ when $n < k$, as it is impossible to choose k things from a set with elements less than k.

We proceed to prove that $f(n) = f(n-1) + f(n-2)$, where $n > 2$. Using the fact that $\binom{n}{0} = 1$, we rewrite $f(n)$ using Pascal's identity and linearity as

$$f(n) = 1 + \sum_{k=1}^{n} \binom{n-k+1}{k} = 1 + \sum_{k=1}^{n} \binom{n-k}{k} + \sum_{k=1}^{n} \binom{n-k}{k-1}.$$

Next, we simplify, getting

$$\begin{aligned}
f(n) &= \sum_{k=0}^{n} \binom{n-k}{k} + \sum_{k=1}^{n} \binom{n-k}{k-1} \\
&= \sum_{k=0}^{n-1} \binom{n-k}{k} + \binom{n-n}{k} + \sum_{k=1}^{n} \binom{n-k}{k-1} \\
&= \sum_{k=0}^{n-1} \binom{n-k}{k} + 0 + \sum_{k=0}^{n-1} \binom{n-k-1}{k} \\
&= \sum_{k=0}^{n-1} \binom{n-k}{k} + \sum_{k=0}^{n-2} \binom{n-k-1}{k} + \binom{n-(n-1)-1}{n-1} \\
&= f(n-1) + f(n-2) + 0.
\end{aligned}$$

Hence $f(n)$ must be equivalent to F_{n+2}. □

CLAIM 1.3. *The number of total arrangements of a tube with n particles is*

$$2^n.$$

Proof. Each particle in the tube can be either an X-particle or a Y-particle, meaning there are 2 choices for each of the n particles. Hence, there are a total of 2^n arrangements. □

Hence by dividing the number of arrangements where there are no consecutive X-particles by the total number of arrangements, we arrive at the formula

$$\frac{F_{n+2}}{2^n}$$

which gives the desired probability. □

PROBLEM 2

- Compute the average number of particles for $n = 2, 3, 4$.
- Can you find the pattern and establish an explicit formula for general n?

We found it most straightforward to proceed directly to finding a general formula. However, it should be noted that it is relatively simple to compute the averages for small n by considering every possible tube after n shots.

THEOREM 2. *The average number of particles after n shots is*

$$\frac{3}{4}n + \frac{1}{4}.$$

Proof. Let the average number of particles after n shots be T_n. Obviously, $T_1 = 1$. Next, consider the chance that after firing a shot, the number of particles *doesn't* decrease. This occurs only in the event that the last particle in the tube is an X-particle, and when the particle emitted is also an X-particle. Since both events have a probability 1/2, the probability that both occur is simply 1/4.

$$\begin{aligned} X &\to X : X \\ Y &\to X : YX \\ X &\to Y : XY \\ Y &\to Y : YY \end{aligned}$$

Hence, the probability that the number of particles *does* increase after firing a shot is $1 - 1/4 = 3/4$. Thus the expected number of particles increases by 3/4 after each shot, giving us the recursion

$$T_{n+1} = T_n + \frac{3}{4}.$$

Since $T_1 = 1$, we arrive at the formula

$$T_n = \frac{3}{4}n + \frac{1}{4}.$$ □

Using this, we can easily compute the average number of particles when $n = 2, 3, 4$:

n	T_n
1	1
2	1.75
3	2.5
4	3.25

PROBLEM 3

Suppose that at each shot, an X-particle is emitted with probability $p \in (0, 1)$.

- Under the same assumption as above, when n is very large, do you think the proportion of X-particles in the tube will eventually stabilise at a certain number? Why/why not?
- If so, can you compute this number explicitly?

THEOREM 3. *Suppose the probability of emitting an X-partical is $p \in (0, 1)$. The limiting ratio of X-particles in the tube, as the number of shots approaches infinity, is*

$$\frac{p}{1+p}.$$

Proof. We use a similar argument to that which is featured in Theorem 2 to arrive at a formula for the expected number of particles after n shots when the probability is p.

CLAIM 3.1. *The average number of particles after n shots when the probability of emitting an X-particle is $p \in (0, 1)$ is*

$$(1 - p^2)n + p^2.$$

Proof. Again, let the average number of particles after n shots be T_n. The number of particles *doesn't* increase when the last particle is an X-particle and the particle emitted is also an X-particle, which has a probabilty of p^2 of occuring. Hence, the probability that the number of particles *does* increase is $1 - p^2$, meaning that the expected number of particles increases by $1 - p^2$ after each shot, and thus we obtain

$$T_{n+1} = T_n + (1 - p^2).$$

Since $T_1 = 1$, the formula for T_n is

$$T_n = (1 - p^2)n + p^2. \qquad \square$$

We can now find the expected number of X-particles in the tube. Since the expected number of Y-particles in the tube is simply $(1 - p)n$, as the probability of emitting a Y-particle is $(1 - p)$ at each shot, the expected number of X-particles is

$$\begin{aligned}\#\text{X-particles} &= \#\text{Total particles} - \#\text{Y-particles} \\ &= T_n - (1-p)n \\ &= (1-p^2)n + p^2 - (1-p)n.\end{aligned}$$

We then divide this by the total number of particles to obtain the desired proportion and then take the limit as $n \to \infty$, to obtain

$$\lim_{n\to\infty} \frac{(1-p^2)n + p^2 - (1-p)n}{(1-p^2)n + p^2} = \lim_{n\to\infty} \frac{(1-p^2) + \frac{p^2}{n} - (1-p)}{(1-p^2) + \frac{p^2}{n}}.$$

By the algebraic limit theorem, we obtain

$$\begin{aligned}\frac{(1-p^2)+0-(1-p)}{(1-p^2)+0} &= \frac{p-p^2}{1-p^2}\\ &= \frac{p(1-p)}{(1+p)(1-p)}\\ &= \frac{p}{1+p}\end{aligned}$$

□

Code Implementation

```
#include <future>
#include <iostream>
#include <random>
#include <string>
#include <thread>
#include <vector>

using namespace std;

random_device rd;
mt19937 rng(rd());

const char particles[] = "XY";

unsigned int num_threads = thread::hardware_concurrency();

string gen_tube(int length, double p) {
  discrete_distribution<int> pick{p, 1 - p};
  string tube;
  for (int i = 0; i < length; i++) {
    tube += particles[pick(rng)];
  }
  return tube;
}

string annihilate(string tube) {
  string output = "";
  for (int i = 0; i < tube.length() - 1; i++) {
    if (tube[i] != tube[i + 1] && tube[i] == 'X')
      output += tube[i];
    else if (tube[i] != 'X')
      output += tube[i];
  }
  output.push_back(tube.back());
  return output;
}

bool check_consec_x(string tube) {
  for (int i = 0; i < tube.length() - 1; i++) {
    if (tube[i] == tube[i + 1] && tube[i] == 'X')
      return 0;
  }
  return 1;
}

double prob_no_two_consec_after_n(int n, int runs) {
  int count = 0;
  for (int i = 0; i < runs; i++) {
    if (check_consec_x(gen_tube(n, 0.5)))
      count++;
  }
  return double(count) / runs;
}

double average_len_after_n(int n, int runs) {
```

```
  long long count = 0;
  for (int i = 0; i < runs; i++) {
    count += annihilate(gen_tube(n, 0.5)).length();
  }
  return count / double(runs);
}

double threaded_avg_len_after_n(int n, int runs) {
  double avg = 0;
  vector<future<double>> threads;
  for (int i = 0; i < num_threads; i++) {
    threads.push_back(
        async(launch::async, average_len_after_n, n, runs / num_threads));
  }
  for (auto &t : threads) {
    avg += t.get();
  }
  return avg / num_threads;
}

double proportion_x(int runs, double p) {
  string tube = "";
  vector<future<string>> threads;
  for (int i = 0; i < num_threads; i++) {
    string small_tube = gen_tube(runs / num_threads, p);
    threads.push_back(async(launch::async, annihilate, small_tube));
  }
  for (auto &t : threads) {
    tube += t.get();
  }
  int len = tube.length();
  return (len - (1 - p) * runs) / len;
}

int main() {

  unsigned long long runs = 1000000;

  cout << "q1\n";
  for (int i = 1; i <= 10; i++) {
    cout << "n = " << i << " prob = " << prob_no_two_consec_after_n(i, runs)
         << '\n';
  }
  cout << "q2\n";
  for (int i = 1; i <= 10; i++) {
    cout << "n = " << i << " avg length = " << threaded_avg_len_after_n(i, runs)
         << '\n';
  }
  cout << "q3\n";
  for (double p = 0; p <= 10; p++) {
    cout << "p = " << p / 10 << " proportion = " << proportion_x(runs, p / 10)
         << '\n';
  }

  return 0;
}
```

This code will output something similar (as it uses a random number generator) to:

```
q1
n = 1 prob = 1
n = 2 prob = 0.750097
n = 3 prob = 0.624354
n = 4 prob = 0.500569
n = 5 prob = 0.406112
n = 6 prob = 0.327949
n = 7 prob = 0.265757
n = 8 prob = 0.215079
n = 9 prob = 0.173859
n = 10 prob = 0.141454
q2
n = 1 avg length = 1
n = 2 avg length = 1.7502
n = 3 avg length = 2.49697
n = 4 avg length = 3.24526
n = 5 avg length = 3.99283
n = 6 avg length = 4.74027
n = 7 avg length = 5.49196
n = 8 avg length = 6.24677
n = 9 avg length = 6.99458
n = 10 avg length = 7.74433
q3
p = 0 proportion = 0
p = 0.1 proportion = 0.0908549
p = 0.2 proportion = 0.166695
p = 0.3 proportion = 0.231008
p = 0.4 proportion = 0.285435
p = 0.5 proportion = 0.334107
p = 0.6 proportion = 0.374551
p = 0.7 proportion = 0.412772
p = 0.8 proportion = 0.445532
p = 0.9 proportion = 0.472449
p = 1 proportion = 1
```

This output matches with the theoretical values obtained earlier.