Prime Conspiracies in the Classroom

Sri Ram and Shantha Bhushan

Oliver and Soundararajan have recently discovered an unexpected bias in the distribution of prime numbers. Interestingly, the authors embarked on this research after hearing about a counter-intuitive result obtained upon comparison of two elementary coin tossing experiments. This article describes these experiments and presents both numerical and analytical methods for exploring the source of inspiration for their research. Elaboration of such key motivations or themes provides a way of bringing the excitement of recent mathematical discoveries to the undergraduate classroom.

(left) Sri Ram teaches biology at Azim Premji University, Bengaluru. He is interested in microbial ecology and mathematical modelling.

(right) Shantha Bhushan teaches mathematics at Azim Premji University. She is interested in Knot theory and topology in general. She has also worked with Kalpavriksh on environmental issues, particular community-based conservation and biodiversity conservation.

1. Introduction

Oliver and Soundararajan have made an astounding discovery regarding behaviour of primes and this has created excitement in the mathematical community and media. As soon as their paper was uploaded to ARxiv (16 march 2016), it was picked up by media outlets, and many popular articles were written. One such article appeared in *Quanta Magazine* [2] with a catchy title, 'Mathematicians discover prime conspiracy' that began with the statement, "a previously unnoticed property of prime numbers seems to violate a longstanding assumption about how they behave."

Oliver and Soundararajan cite a coin tossing experiment by Alice and Bob as one of their key motivations. Bob is interested in the outcome HH (head followed by head) whereas Alice is interested in the outcome TH (tail followed by head). If a coin is tossed only twice, Alice and Bob have equal chances of winning. However, they con-

Keywords
Recursion, expectation, eigenvalues, experiment, simulation, distribution of primes.

tinue the experiment and toss coins until they get the desired outcome (HH for Bob, TH for Alice). If Bob gets outcomes such as THH or HTHH, TTHH etc., he stops the experiment. Similarly, if Alice gets results like HTH, HHTH, etc. HTTH, she stops the experiment. Naively, one would think that the mean or average number of tosses for success should be the same for Alice and Bob. But it turns out that the mean is 4 for Alice and 6 for Bob. This result seems counter-intuitive as one would expect the mean expectation to be the same.

Oliver and Soundararajan's paper itself is titled 'Unexpected biases in the distribution of consecutive primes'. They provide numerical evidence that the distribution of prime numbers shows bias. A prime number ending with the digit 1 is more likely to be followed by a prime number not ending with one. A prime number ending with 3 is more likely to be followed by a prime number whose last digit is not 3. This result goes against the intuition that a prime number ending with 1 should be followed by a prime number with equal chances of the last digit being $1, 3, 7, 9$.

How can we bring the excitement of such a new path-breaking discovery to our undergraduates? How can such discussions be made relevant in the context of the courses that they are studying? Many students come for such 'popular' talks but get discouraged because the discussions involve tools and results that seem beyond their reach. One way might be to pick out a key motivation or theme that leads to the final result and weave it into concepts they are familiar with.

How can we bring the excitement of a new path-breaking discovery to our undergraduates? How can such discussions be made relevant in the context of the courses that they are studying?

In this article, we describe several ways (both analytical and numerical) of bringing Alice and Bob's experiment into the classroom setting. Each approach can be tied to concepts that they are learning. More importantly, the approaches lend themselves to asking interesting questions that would motivate students. Analytically, we set up a recurrence relation and either use linear algebra or generating functions to solve the relation and find the mean. Numerically, the mean can be determined either using spreadsheets (Excel) or writing a simulation in a programming language such as Python.

2. Simulating the Experiments

To verify that the average number of tosses in the two experiments is different, one may simply ask students to perform them. With a class of 30 students and with 5 trials per student, around 150 data points can be easily obtained. To go much beyond this number of data points, it would be necessary to write a computer simulation. An example simulation written in the Python language is included in *Box* 1.

Box 1.

```
# Simulate Alice and Bob's coin tossing experiments
# Coin Tosses are simulated by generating a
# random 0 or 1 with equal probability
# 0 stands for heads, 1 for tails
import numpy

# Alice's coin toss experimen
# Returns the number of coin tosses in one experiment
def alice():
    # Generate first two coin tosses
    toss1 = numpy.random.random_integers(0,1)
    toss2 = numpy.random.random_integers(0,1)
    ntosses = 2 # tracks number of coin tosses

    # Keep tossing coins till one gets HT
    while True:
        if toss1==0 and toss2==1:
            break
        else:
            toss1 = toss2
            toss2 = numpy.random.random_integers(0,1)
            ntosses += 1
    return ntosses
```

Box 1. continued....

Box 1. continued....

```
# Bob's coin toss experiment
# Returns the number of coin tosses in one experiment
def bob():
    # Generate first two coin tosses
    toss1 = numpy.random.random_integers(0,1)
    toss2 = numpy.random.random_integers(0,1)
    ntosses = 2 # tracks number of coin tosses

    # Keep tossing coins till one gets HH
    while True:
        if toss1==0 and toss2==0:
            break
        else:
            toss1 = toss2
            toss2 = numpy.random.random_integers(0,1)
            ntosses += 1
    return ntosses

# Perform Alice and Bob's experiments multiple times
ntrials = 100000 # Number of repetitions of experiment
alice_trials = numpy.zeros(ntrials)
bob_trials = numpy.zeros(ntrials)

for i in range(ntrials):
    alice_trials[i] = alice()
    bob_trials[i] = bob()

print("Avg. tosses (Alice): "+ str(numpy.mean(alice_trials)))
print("Avg. tosses (Bob): "+ str(numpy.mean(bob_trials)))
```

Using simulations, students can indeed verify that the average number of tosses in Bob's experiment is higher. They can also plot the distribution of number of coin tosses and find other ways of exploring the simulation data. What is the most likely number of coin tosses? Is it the same as the average? If Alice and Bob start their experiments simultaneously and take the same time to toss a coin, what is the probability that Bob finishes before Alice? Modern computers are powerful enough that it is possible to perform a million independent trials of the two coin tossing experiments within a few seconds.

Given this computing power and the availability of excellent programming tools, it is not surprising that computers have become an indispensable tool in mathematics research. For instance, the starting point for Oliver and Soundararajan's work was a numerical investigation of prime numbers.

Computers have become an indispensable tool in mathematics research.

3. Identifying Patterns

After verifying that the average in the two experiments is different either by actual coin tossing or via computer simulations, the next insight can come from enumerating the toss sequences that Alice or Bob can observe. This enumeration can be used to deduce a recurrence relation and that can be used to calculate the average number of coin tosses.

Coin toss sequences can be thought of as strings consisting of letters H and T. We call a sequence of length n favourable if the experiment ends upon observing that sequence. A favourable sequence (FS) for Alice's experiment must end in HT, but the first $n-1$ tosses should not have an H followed by a T (otherwise, the experiment would have ended earlier). In the rest of the article, we use the following notation:

- n: The number of coin tosses in an experiment.
- F_n: The number of favourable coin toss sequences of length n.
- H_n: The number of favourable coin toss sequences of length n beginning with an H.
- T_n: The number of favourable coin toss sequences of length n beginning with a T.
- p_n: The probability of the experiment ending after n tosses.

Tosses (n)	Favourable Sequences of Length n	Number (F_n)	Probability (p_n)
2	HT	1	$\frac{1}{4}$
3	THT, HHT	2	$\frac{2}{8}$
4	TTHT, THHT, HHHT	3	$\frac{3}{16}$
5	TTTHT, TTHHT, THHHT, HHHHT	4	$\frac{4}{32}$
6	TTTTHT, TTTHHT, TTHHHT, THHHHT, HHHHHT	5	$\frac{5}{64}$
7	TTTTTHT, TTTTHHT, TTTHHHT, TTHHHHT, THHHHHT, HHHHHHT	6	$\frac{6}{128}$

Table 1. Favourable sequences in Alice's experiment.

From the definitions above, it is easy to see that $F_n = T_n + H_n$ and $p_n = \frac{F_n}{2^n}$.

3.1 *Alice's Experiment*

Alice tosses a coin till she gets an H followed by a T. *Table* 1 lists favourable sequences of lengths 2–7. Clearly, an FS has to be of length ≥ 2. There is only one such sequence of length 2 (HT), so that $F_2 = 1$, $H_2 = 1$ and $T_2 = 0$. To find a recurrence relation, we observe the following:

- The last $n - 1$ tosses of any FS of length $n > 2$ form another FS.
- To form an FS of length $n > 2$ beginning with T, we can prefix T to any FS of length $n - 1$. This will not generate the pattern HT in the beginning.
- To form an FS of length $n > 2$ beginning with H, we can prefix H to any FS of length $n-1$ beginning with an H (but not T).

These observations lead to the following recurrence relations:

$$\begin{aligned} T_n &= F_{n-1} = T_{n-1} + H_{n-1} \,, \\ H_n &= H_{n-1} \,, \\ F_n &= T_n + H_n \,. \end{aligned} \tag{1}$$

3.2 *Bob's Experiment*

Bob is looking for two consecutive heads. The favourable sequences of lengths 2–7 are listed in *Table* 2. As with Alice's experiment, there is only one FS of length 2 (HH), so that $F_2 = 1$, $H_2 = 1$ and $T_2 = 0$. Once again, the last $n-1$ tosses of an FS of length $n > 2$ form another FS. An FS of length n can be generated by prefixing T to any FS of length $n-1$ or by prefixing H to an FS of length $n-1$ starting with T. Thus, we have:

$$\begin{aligned} T_n &= F_{n-1} = T_{n-1} + H_{n-1} \,, \\ H_n &= T_{n-1} \,, \\ F_n &= T_n + H_n \,. \end{aligned} \tag{2}$$

4. Computing Averages

The average number of coin tosses (the expected value)

Table 2. Favourable sequences in Bob's experiment

Tosses (n)	Favourable Sequences of Length n	Number (F_n)	Probability (p_n)
2	HH	1	$\frac{1}{4}$
3	THH	1	$\frac{1}{8}$
4	TTHH,HTHH	2	$\frac{2}{16}$
5	TTTHH, THTHH, HTTHH	3	$\frac{3}{32}$
6	TTTTHH, TTHTHH, THHTTH, HTTTHH, HTHTHH	5	$\frac{5}{64}$
7	TTTTTHH, TTTHTHH, TTHHTTH, THTTTHH, THTHTHH, HTTTTHH, HTTHTHH, HTHHTTH	8	$\frac{8}{128}$

is given by $\sum_{n=0}^{\infty} np_n$. Once the recurrence relation has been identified, the average number of coin tosses can be computed either numerically using a spreadsheet program like Excel or analytically using a variety of techniques. We illustrate these methods below.

4.1 *Numerical Computation*

The simplest way to compute average values numerically is to use a spreadsheet program. *Figure* 1 illustrates this approach. If one looks at recurrence relations (1) and (2), one can observe that computing H_n, T_n and p_n using a spreadsheet only requires values in the same row or the previous row. This property makes the computation quite straightforward. Spreadsheets also make it easy to plot the computed values. Sample plots are shown in *Figure* 2. One can see that $p_n \to 0$ more rapidly in Alice's experiment. This is in accordance with the fact that the average number of coin tosses for Alice's experiment is smaller.

Figure 1. Computing averages using a spreadsheet program. Column headings are as follows: Tn denotes T_n, Hn denotes H_n, pn denotes p_n, n × pn is np_n, and cumulative average respresents $\sum_{i=0}^{n} ip_i$.

Calculating Expected Number of Coin Tosses

Tosses	Alice's Experiment					Bob's Experiment				
(n)	Tn	Hn	pn	n X pn	Cumulative Average	Tn	Hn	pn	n X pn	Cumulative Average
2	0	1	0.25000	0.50000	0.500	0	1	0.25000	0.50000	0.500
3	1	1	0.25000	0.75000	1.250	1	0	0.12500	0.37500	0.875
4	2	1	0.18750	0.75000	2.000	1	1	0.12500	0.50000	1.375
5	3	1	0.12500	0.62500	2.625	2	1	0.09375	0.46875	1.844
6	4	1	0.07813	0.46875	3.094	3	2	0.07813	0.46875	2.313
7	5	1	0.04688	0.32813	3.422	5	3	0.06250	0.43750	2.750
8	6	1	0.02734	0.21875	3.641	8	5	0.05078	0.40625	3.156
9	7	1	0.01563	0.14063	3.781	13	8	0.04102	0.36914	3.525
10	8	1	0.00879	0.08789	3.869	21	13	0.03320	0.33203	3.857
11	9	1	0.00488	0.05371	3.923	34	21	0.02686	0.29541	4.153
12	10	1	0.00269	0.03223	3.955	55	34	0.02173	0.26074	4.414
13	11	1	0.00146	0.01904	3.974	89	55	0.01758	0.22852	4.642
14	12	1	7.93E-04	0.01111	3.985	144	89	0.01422	0.19910	4.841
15	13	1	4.27E-04	0.00641	3.992	233	144	0.01151	0.17258	5.014
16	14	1	2.29E-04	0.00366	3.995	377	233	0.00931	0.14893	5.163
17	15	1	1.22E-04	0.00208	3.997	610	377	0.00753	0.12801	5.291
18	16	1	6.48E-05	0.00117	3.999	987	610	0.00609	0.10966	5.400
19	17	1	3.43E-05	0.00065	3.999	1597	987	0.00493	0.09364	5.494
20	18	1	1.81E-05	0.00036	4.000	2584	1597	0.00399	0.07975	5.574
21	19	1	9.54E-06	2.00E-04	4.000	4181	2584	0.00323	0.06774	5.641
22	20	1	5.01E-06	1.10E-04	4.000	6765	4181	0.00261	0.05741	5.699
23	21	1	2.62E-06	6.03E-05	4.000	10946	6765	0.00211	0.04856	5.747
24	22	1	1.37E-06	3.29E-05	4.000	17711	10946	0.00171	0.04099	5.788
25	23	1	7.15E-07	1.79E-05	4.000	28657	17711	0.00138	0.03455	5.823

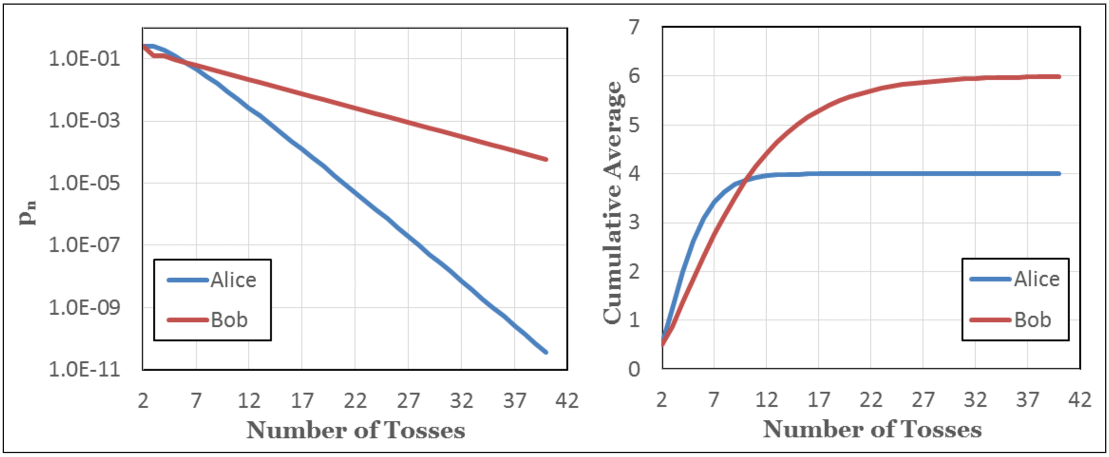

Figure 2. p_n (left) and $\sum_{i=1}^{n} ip_i$ (right) as a function of n for Alice and Bob's experiments. Values were obtained using the spreadsheet shown in Figure 1.

4.2 *Using Generating Functions*

Given a sequence $\{s_n\}_{n=0}^{\infty}$, the corresponding generating function is given by $s(x) = \sum_{n=0}^{\infty} s_n x^n$. In cases where a closed form expression for $s(x)$ can be obtained and the term $s_n x^n$ can be identified with a probability distribution for a suitable value for x, manipulations of the power series can be used to compute expected values (and even higher moments) as will be shown below.

4.2.1 ***Alice's Experiment***: From recurrence relations (1), we find that $H_n = H_{n-1} = ... = H_2 = 1$, so that $T_n = T_{n-1} + 1$. Since $T_2 = 0$, we can get by induction that $T_n = n - 2$. Hence, $F_n = T_n + H_n = n - 1$. Thus, we have:

$$\begin{aligned}\overline{p} = \sum_{n=2}^{\infty} np_n &= \sum_{n=2}^{\infty} n\frac{F_n}{2^n} \\ &= \sum_{n=2}^{\infty} n\frac{(n-1)}{2^n} \\ &= \sum_{n=0}^{\infty} n\frac{(n-1)}{2^n} \ .\end{aligned}$$

In going from the second to the third line above, we have used the fact that $n(n-1) = 0$ for $n = 0, 1$. The last line corresponds to setting $x = \frac{1}{2}$ in the generating

This calculation assumes that the generating function for the Fibonacci sequence converges for $x = 1/2$.

function $s(x)$ for the sequence $\{s_n = n(n-1)\}_{n=0}^{\infty}$. In (A5) of *Appendix* A, $s(x)$ is shown to be $\frac{2x^2}{(1-x)^3}$, so that the expected value is $\overline{p} = s\left(\frac{1}{2}\right) = 4$.

4.2.2 *Bob's Experiment*: For Bob's experiment, the recurrence relations (2) yield $T_n = F_{n-1}$ and $H_n = T_{n-1} = F_{n-2}$ so that $F_n = T_n + H_n = F_{n-1} + F_{n-2}$. Since $F_1 = 0$ and $F_2 = 1$, we can identify F_n with the famous Fibonacci sequence with an offset of 1, that is, $F_n = f_{n-1}$. Here, $\{f_n\}_{n=0}^{\infty}$ denotes the Fibonacci sequence (defined in *Appendix* B). In *Appendix* B, we derive the generating function $f(x) = \sum_{n=0}^{\infty} f_n x^n$ for the Fibonacci sequence to be $\frac{x}{1-x-x^2}$. Formal differentiation of this power series yields

$$f'(x) = \frac{1+x^2}{(1-x-x^2)^2} = \sum_{n=1}^{\infty} n f_n x^{n-1}\,.$$

The expected number of coin tosses in Bob's experiment is:

$$\begin{aligned}
\overline{p} = \sum_{n=2}^{\infty} n p_n &= \sum_{n=2}^{\infty} n \frac{F_n}{2^n} \\
&= \sum_{n=2}^{\infty} n \frac{f_{n-1}}{2^n} \\
&= \sum_{n=1}^{\infty} (n+1) \frac{f_n}{2^{n+1}} \\
&= \frac{1}{4} \sum_{n=1}^{\infty} n f_n \left(\frac{1}{2}\right)^{n-1} + \frac{1}{2} \sum_{n=0}^{\infty} f_n \left(\frac{1}{2}\right)^{n} \\
&\quad \text{using } f_0 = 0 \\
&= \frac{1}{4} f'\left(\frac{1}{2}\right) + \frac{1}{2} f\left(\frac{1}{2}\right) \\
&= \frac{1}{4} \cdot 20 + \frac{1}{2} \cdot 2 = 6\,.
\end{aligned}$$

4.3 *Using Linear Algebra*

Instead of adopting the generating functions approach, recurrence relations (1) and (2) can also be solved using the machinery of linear algebra. This alternative way is illustrated below.

The higher eigenvalue of the matrix A is the so-called golden ratio, which has an illustrious mathematical history.

4.3.1 ***Bob's Experiment***: Recurrence relations (2) can be represented in matrix notation as follows:

$$\begin{pmatrix} T_n \\ H_n \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} T_{n-1} \\ H_{n-1} \end{pmatrix} . \tag{3}$$

Denoting the matrix $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ by A, we get by induction:

$$\begin{pmatrix} T_n \\ H_n \end{pmatrix} = A^{n-2} \begin{pmatrix} T_2 \\ H_2 \end{pmatrix} = A^{n-2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{4}$$

Using a similarity transform, the matrix A can be diagonalized as $A = M^{-1} \left[\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right] M$ for some suitable invertible matrix M. Here λ_1, λ_2 are the two eigenvalues of A. $\lambda_1 = \frac{1+\sqrt{5}}{2}$, $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Thus, (4) can be rewritten as:

$$\begin{pmatrix} T_n \\ H_n \end{pmatrix} = M^{-1} \begin{bmatrix} \lambda_1^{n-2} & 0 \\ 0 & \lambda_2^{n-2} \end{bmatrix} M \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{5}$$

From here, one can see that both T_n and H_n are linear combinations of λ_1^{n-2} and λ_2^{n-2}. Since $F_n = T_n + H_n$, F_n is also a linear combination of λ_1^{n-2} and λ_2^{n-2}. If we write $F_n = \alpha\lambda_1^{n-2} + \beta\lambda_2^{n-2}$, we find using $F_2 = 1$ and $F_3 = 1$ that $\alpha = \frac{1}{2} + \frac{\sqrt{5}}{10}$, $\beta = \frac{1}{2} - \frac{\sqrt{5}}{10}$. The average number of coin tosses is then

$$\begin{aligned} \overline{p} &= \sum_{n=2}^{\infty} n p_n = \sum_{n=2}^{\infty} n \frac{F_n}{2^n} \\ &= \sum_{n=2}^{\infty} n \left(\frac{\alpha\lambda_1^{n-2} + \beta\lambda_2^{n-2}}{2^n} \right) \\ &= \frac{\alpha}{2\lambda_1} \sum_{n=2}^{\infty} n \left(\frac{\lambda_1}{2} \right)^{n-1} + \frac{\beta}{2\lambda_2} \sum_{n=2}^{\infty} n \left(\frac{\lambda_2}{2} \right)^{n-1} \end{aligned}$$

Unlike the other analytical methods presented in this article, the linear algebra method does require dealing with irrational numbers explicitly.

$$= \frac{\alpha}{2\lambda_1}\left(-1+\sum_{n=1}^{\infty} n\left(\frac{\lambda_1}{2}\right)^{n-1}\right) + \frac{\beta}{2\lambda_2}\left(-1+\sum_{n=1}^{\infty} n\left(\frac{\lambda_2}{2}\right)^{n-1}\right).$$

Noting that $|\frac{\lambda_1}{2}| < 1$ and $|\frac{\lambda_2}{2}| < 1$, we can put $x = \frac{\lambda_1}{2},\frac{\lambda_2}{2}$ in equation (A2) of *Appendix* A to get:

$$\overline{p} = \frac{\alpha}{2\lambda_1}\left(-1+\frac{1}{\left(1-\frac{\lambda_1}{2}\right)^2}\right) + \frac{\beta}{2\lambda_2}\left(-1+\frac{1}{\left(1-\frac{\lambda_2}{2}\right)^2}\right). \tag{6}$$

Plugging the values of α,β,λ_1,λ_2 into (6) yields $\overline{p} = 6$.

4.3.2 *Alice's Experiment*: We can represent recurrence relations (1) in matrix form as follows:

$$\begin{pmatrix} T_n \\ H_n \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} T_{n-1} \\ H_{n-1} \end{pmatrix}. \tag{7}$$

For Alice's experiment, the matrix $A = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ is defective (non-diagonalizable). For such defective matrices, the general solution for T_n, H_n and F_n would be of the form $(\alpha+\beta n)\lambda^n$, where λ is the repeated eigenvalue and $(\alpha+\beta n)$ is a linear function of n with α, β being determined by initial conditions. In this case, since $\lambda = 1$, F_n turns out to be a purely linear function ($F_n = n-1$). Once F_n is determined, we get:

$$\begin{aligned} \overline{p} &= \sum_{n=2}^{\infty} n\frac{F_n}{2^n} = \sum_{n=2}^{\infty} n\frac{(n-1)}{2^n} \\ &= \frac{1}{4}\sum_{n=2}^{\infty} n(n-1)\left(\frac{1}{2}\right)^{n-2} \\ &= \frac{1}{4}\cdot 16 = 4\,. \end{aligned}$$

Here, the sum has been evaluated by putting $x = \frac{1}{2}$ in equation (A3) of *Appendix* A.

Students may find the use of recurrence relations somewhat abstract. An alternative approach is to explicitly enumerate favourable sequences and calculate the expected value after that.

5. Explicit Enumeration of Favourable Sequences

Students may find the use of recurrence relations to evaluate F_n somewhat abstract. An alternative approach is to explicitly enumerate favourable sequences and calculate the expected value after that. This approach is illustrated below.

5.1 *Alice's Experiment*

Alice's experiment ends upon getting a heads followed by a tails. She will continue tossing till she first gets an H (i.e., gets a sequence of tails followed by head). Once an H is obtained, she will keep tossing as long as she keeps getting heads, but will stop the first time she gets a tails. Thus favourable sequences in her experiment must be of the form $\mathrm{T}^{n_1}\mathrm{H}\mathrm{H}^{n_2}\mathrm{T}$, where T^{n_1} denotes a string of n_1 tails, H^{n_2} denotes a string of n_2 heads, $0 \leq n_1 < \infty$ and $0 \leq n_2 < \infty$. With this enumeration, the average number of tosses becomes:

$$\begin{aligned}
\overline{p} &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}(n_1+n_2+2)\left(\frac{1}{2}\right)^{n_1+n_2+2} \\
&= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} n_1 \left(\frac{1}{2}\right)^{n_1+n_2+2} \\
&\qquad + \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} n_2 \left(\frac{1}{2}\right)^{n_1+n_2+2} \\
&\qquad + 2\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \left(\frac{1}{2}\right)^{n_1+n_2+2} \\
&= \frac{1}{4}\sum_{n_1=0}^{\infty} n_1 \left(\frac{1}{2}\right)^{n_1} \left(\sum_{n_2=0}^{\infty}\left(\frac{1}{2}\right)^{n_2}\right)
\end{aligned}$$

We enumerate the favourable sequences in Bob's experiment based on the number of heads in them.

$$+\frac{1}{4}\sum_{n_2=0}^{\infty} n_2 \left(\frac{1}{2}\right)^{n_2} \left(\sum_{n_1=0}^{\infty}\left(\frac{1}{2}\right)^{n_1}\right)$$
$$+\frac{1}{2}\left(\sum_{n_1=0}^{\infty}\left(\frac{1}{2}\right)^{n_1}\right)\left(\sum_{n_2=0}^{\infty}\left(\frac{1}{2}\right)^{n_2}\right).$$

The sums within parentheses can be shown to be equal to 2 by putting $x = \frac{1}{2}$ in (A1) of *Appendix* A. Thus, we get:

$$\overline{p} = \frac{1}{2}\sum_{n_1=0}^{\infty} n_1 \left(\frac{1}{2}\right)^{n_1} + \frac{1}{2}\sum_{n_2=0}^{\infty} n_2 \left(\frac{1}{2}\right)^{n_2} + \frac{1}{2}\cdot 4$$
$$= \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 2 + 2 = 4\,.$$

To evaluate the sums, we have put $x = \frac{1}{2}$ in (A4) of *Appendix* A.

5.2 *Bob's Experiment*

We enumerate the favourable sequences in Bob's experiment based on the number of heads in them. Since Bob's experiment terminates upon getting two consecutive heads, there must be at least 2 heads in favourable sequences. Consider favourable sequences with $q + 2$ heads (the last 2 heads must be the ending HH), $0 \leq q < \infty$. The first of the $q + 2$ heads can be preceded by zero or more tails. In addition, the first q heads must be followed by one or more tails (if they are followed by zero tails, then there would be two consecutive heads before the ending HH). Thus, the favourable sequences with $q + 2$ heads must be of the form
$\mathrm{T}^{n_1}\mathrm{H}_1\mathrm{T}^{1+n_2}\mathrm{H}_2...\mathrm{T}^{1+n_q}\mathrm{H}_q T^{1+n_{q+1}}\mathrm{HH}$, where $\mathrm{H}_1, ..., \mathrm{H}_q$ denote the first q heads and $0 \leq n_i < \infty$ for $1 \leq i \leq q+1$. The contribution of the sequences with $q + 2$ heads to the average is:

$$\overline{p_q} = \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}...\sum_{n_{q+1}=0}^{\infty} \Delta \left(\frac{1}{2}\right)^{\Delta},$$

where $\Delta = n_1 + 1 + n_2 + ... + 1 + n_{q+1} + q + 2$. We therefore get:

$$\overline{p_q} = \left(\frac{1}{2}\right)^{2q+2} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} ... \sum_{n_{q+1}=0}^{\infty} n_1 \left(\frac{1}{2}\right)^{n_1+n_2+...+n_{q+1}}$$
$$+ ... + \left(\frac{1}{2}\right)^{2q+2} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} ... \sum_{n_{q+1}=0}^{\infty} n_{q+1} \left(\frac{1}{2}\right)^{n_1+n_2+...+n_{q+1}}$$
$$+ \left(\frac{1}{2}\right)^{2q+2} (2q+2) \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} ... \sum_{n_{q+1}=0}^{\infty} \left(\frac{1}{2}\right)^{n_1+n_2+...+n_{q+1}} \tag{8}$$

By putting $x = \frac{1}{2}$ in (A1) and (A4) of *Appendix* A, we get

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 2$$

and

$$\sum_{n=0}^{\infty} n \left(\frac{1}{2}\right)^n = 2.$$

The first $q + 1$ sums in (8) evaluate to

$$\left(\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n\right)^q \left(\sum_{n=0}^{\infty} n \left(\frac{1}{2}\right)^n\right) = 2^{q+1}$$

, whereas the last sum evaluates to

$$\left(\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n\right)^{q+1} = 2^{q+1} .$$

Thus, we get:

$$\overline{p_q} = \left(\frac{1}{2}\right)^{2q+2} (q + 1)2^{q+1} + \left(\frac{1}{2}\right)^{2q+2} (2q + 2)2^{q+1} .$$
$$= 3(q + 1) \left(\frac{1}{2}\right)^{q+1} .$$

The expected number of coin tosses can be calculated by summing over all possible values of q:

$$\overline{p} = \sum_{q=0}^{\infty} \overline{p_q} = \sum_{q=0}^{\infty} 3(q+1)\left(\frac{1}{2}\right)^{q+1}$$
$$= \frac{3}{2}\sum_{q=0}^{\infty} q\left(\frac{1}{2}\right)^{q} + \frac{3}{2}\sum_{q=0}^{\infty}\left(\frac{1}{2}\right)^{q}$$
$$= \frac{3}{2}\cdot 2 + \frac{3}{2}\cdot 2 = 6$$

Once again, the sums over index q have been evaluated by putting $x = \frac{1}{2}$ in (A1) and (A4) of *Appendix* A.

6. Further Explorations

We conclude this article by considering a third coin tossing experiment. Carol keeps tossing coins till the total number of heads exceeds the number of tails by 1. Some of the favourable sequences in her experiment include H, THH, THTHH, TTHHH. As with Alice and Bob's experiments, we can start with computer simulations. The first observation that one is likely to make is that the simulations take a longer time to complete[1]. Different students might get vastly different averages and the average would seem to increase with the number of trials instead of converging to a fixed number. What is happening? It turns out that for Carol's experiment, $\sum_{n=1}^{\infty} p_n = 1$, but $\sum_{n=1}^{\infty} np_n$ is not a finite number. In other words, while the simulation is guaranteed to end, the computed average would be a misleading number.

[1] It is advisable to write the code in a manner that there is a cutoff in the number of tosses per trial. If the cutoff is crossed, the trial should end even if the number of heads is not one more than the number of tails. A recommended cutoff is 10,00,000.

Acknowledgements

The authors would like to thank Prof. Rajaram Nityananda at the School of Liberal Studies in Azim Premji University for valuable feedback and comments on the article.

Suggested Reading

For Oliver and Soundararajan's Result:

[1] R J L Oliver and K Soundararajan, Unexpected biases in the distribution of consecutive primes. http://arxiv.org/abs/1603.03720

[2] E Klarreich, Mathematicians discover prime conspiracy. https://www.quantamagazine.org/20160313-mathematicians-discover-prime-conspiracy/

[3] T Tao, Biases between consecutive primes. https://terrytao.wordpress.com/2016/03/14/biases-between-consecutive-primes/

For Coin Tossing:

[4] E W Weisstein, Coin Tossing. http://mathworld.wolfram.com/CoinTossing.html

For Probability and Counting:

[5] Combinatorics. https://brilliant.org/math/combinatorics/

[6] H S Wilf, generatingfunctionology. https://www.math.upenn.edu/~wilf/DownldGF.html

Address for Correspondence
Sri Ram and
Shantha Bhushan
School of Liberal Studies
Azim Premji University
Bengaluru 560 100, India.
Email: sri.ram@apu.edu.in
shantha.bhushan@apu.edu.in

Appendix A. Generating Functions for $s_n = n$ and $s_n = n(n-1)$

We start with the following identity which holds for $|x| < 1$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots . \tag{A1}$$

Differentiating (A1) with respect to x, we get $\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + 5x^4 + \dots$. That is,

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2} . \tag{A2}$$

Differentiating (A2) with respect to x yields $\frac{2}{(1-x)^3} = 2+3\cdot 2x+4\cdot 3x^2+5\cdot 4x^3+6\cdot 5x^5+\dots$. Thus,

$$\sum_{n=2}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3} . \tag{A3}$$

The generating function for the sequence $\{s_n = n\}_{n=0}^{\infty}$ is given by:

$$\begin{aligned} s(x) &= \sum_{n=0}^{\infty} nx^n \\ &= \sum_{n=1}^{\infty} nx^n && \text{using } nx^n = 0 \text{ for } n = 0 \\ &= x\sum_{n=1}^{\infty} nx^{n-1} \\ &= \frac{x}{(1-x)^2} && \text{using equation (A2)} . \end{aligned} \tag{A4}$$

The generating function for the sequence $\{s_n = n(n-1)\}_{n=0}^{\infty}$ is given by:

$$\begin{aligned} s(x) &= \sum_{n=0}^{\infty} n(n-1)x^n \\ &= \sum_{n=2}^{\infty} n(n-1)x^n && \text{using } n(n-1) = 0 \text{ for } n = 0, 1 \\ &= x^2\sum_{n=2}^{\infty} n(n-1)x^{n-2} \\ &= \frac{2x^2}{(1-x)^3} && \text{using equation (A3)} . \end{aligned} \tag{A5}$$

Appendix B. Generating Function for Fibonacci Sequence

The Fibonacci sequence f_n is defined using the following recurrence relation:

$$\begin{aligned} f_0 &= 0 \\ f_1 &= 1 \\ f_n &= f_{n-1} + f_{n-2} \text{ for } n \geq 2\,. \end{aligned}$$

The generating function $f(x)$ for the Fibonacci series is given by $f(x) = \sum_{n=0}^{\infty} f_n x^n$. With this definition, we have:

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} f_n x^n \\ &= x + \sum_{n=2}^{\infty} f_n x^n && \text{(using } f_0 = 0,\ f_1 = 1\text{)} \\ &= x + \sum_{n=2}^{\infty} (f_{n-1} + f_{n-2}) x^n && \text{(using } f_n = f_{n-1} + f_{n-2} \text{ for } n \geq 2\text{)} \\ &= x + x \sum_{n=2}^{\infty} f_{n-1} x^{n-1} + x^2 \sum_{n=2}^{\infty} f_{n-2} x^{n-2} \\ &= x + x f(x) + x^2 f(x) && \text{(using } f(x) = \textstyle\sum_{n=0}^{\infty} f_n x^n\text{)}\,. \end{aligned}$$

Rearranging $f(x) = x + xf(x) + x^2 f(x)$, we get

$$f(x) = \frac{x}{1 - x - x^2}.$$