Rose-Hulman Undergraduate Mathematics Journal

Volume 24
Issue 2

Article 10

Divisibility Probabilities for Products of Randomly Chosen Integers

Noah Y. Fine
University of Maryland, noahyfine@gmail.com

Recommended Citation
Fine, Noah Y. (2023) "Divisibility Probabilities for Products of Randomly Chosen Integers," *Rose-Hulman Undergraduate Mathematics Journal*: Vol. 24: Iss. 2, Article 10.
Available at: https://scholar.rose-hulman.edu/rhumj/vol24/iss2/10

Divisibility Probabilities for Products of Randomly Chosen Integers

Cover Page Footnote

I am incredibly grateful to Professor Lawrence C. Washington from the University of Maryland Mathematics Department and to Ashvin Rajan for all of their advice and assistance in completing this paper. I would also like to thank the referee for carefully reading the paper and providing many helpful comments.

ℝose-Hulman Undergraduate Mathematics Journal

Volume 24, Issue 2, 2023

Divisibility Probabilities for Products of Randomly Chosen Integers

By *Noah Y. Fine*

Abstract. We find a formula for the probability that the product of n positive integers, chosen at random, is divisible by some integer d. We do this via an inductive application of the Chinese Remainder Theorem, generating functions, and several other combinatorial arguments. Additionally, we apply this formula to find a unique, but slow, probabilistic primality test.

1 Introduction

The problem that we solve in this paper is motivated by the following question from FiveThirtyEight.com [1]:

If we are given three random integers x, y, and z, what is the probability that their product xyz is divisible by 100?

In this paper, we derive a formula for the probability that a finite product of any n "randomly chosen" positive integers is divisible by a given positive integer d. This formula answers the question posed on FiveThirtyEight, and similar questions.

The ambiguity with this problem comes from how we define "randomly chosen". Based on their solution [3], FiveThirtyEight clearly intended that each positive integer be chosen using a uniform distribution. The remaining question is then: what is the set from which we are choosing uniformly?

Two answers seem plausible and in line with FiveThirtyEight's intent in their posing of the problem. The first is that each integer is chosen uniformly from $\{1,2,\dots,d\}$, or essentially the residue classes mod d, and the second is that each integer is chosen uniformly from $\{1,2,\dots,\mathrm{N}\}$, where N approaches infinity (assuming that this probability converges).

We will answer these two questions separately. As finding the formula in a slightly more generalized finite case can be used in solving the infinite case, we will find the probability described when choosing our random integers uniformly from $\{1,2,\dots,cd\}$, where c is an integer constant.

Beautifully, the answer in both the finite and infinite cases turns out to be the same:

Mathematics Subject Classification. 11N25, 11A51, 11B75

Keywords. probability, divisibility, primality test

Theorem 1.1. *The probability that $d|\mathrm{X}_n$, where $d = \prod_{i=1}^{\ell} p_i^{r_i}$ is the prime factorization of d and X_n is the product of n randomly chosen positive integers (by either definition of "randomly chosen" described above), is*

$$\prod_{i=1}^{\ell}\left(1-\sum_{k=0}^{r_i-1}\binom{n+k-1}{n-1}\left(1-\frac{1}{p_i}\right)^n\left(\frac{1}{p_i}\right)^k\right).$$

This formula gives the answer to FiveThirtyEight's original problem as 0.1243, which matches the solution given by FiveThirtyEight [3].

In order to complete this proof, we will review several well-known number theory definitions and theorems. First, we recall the definition of a multiplicative function: an arithmetic function f is called multiplicative if $f(mn) = f(m)f(n)$ whenever m and n are relatively prime positive integers [4, pp. 166–167]. Such functions are important because their values at all positive integers are determined by their values at prime powers. This fact inspires **Proposition 2.1**.

Second, we recall the Chinese Remainder Theorem: if $m_1, m_2, \ldots, m_r$ are pairwise relatively prime positive integers, then the system of congruence

$$x_1 \equiv a_1 \pmod{m_1}$$

$$x_2 \equiv a_2 \pmod{m_2}$$

$$\ldots$$

$$x_r \equiv a_r \pmod{m_r}$$

has a unique solution modulo $\mathrm{M} = m_1 m_2 \cdots m_r$ [4, pp. 107–108]. This theorem is used in the proof of **Lemma 2.4**.

And finally, we recall the Squeeze Theorem, which states that if $\{x_n\}, \{y_n\}$, and $\{z_n\}$ are sequences of real numbers such that

$$x_n \le y_n \le z_n$$

for all n, and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n$, then $\lim_{n\to\infty} y_n$ exists and

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n [2, pp.66].$$

This theorem is used in the proof of **Theorem 3.1**.

2 On Finite Intervals

First, let random variables x_i be drawn uniformly from $\{1, 2, \ldots, \mathrm{N}\}$, and let $\mathrm{X}_n = \prod_{i=1}^{n} x_i$. We use the notation $\mathrm{P_N}(d|\mathrm{X}_n)$ to mean the probability that X_n is divisible by d. We will also use this notation for the probability of similar situations where the x_i are chosen uniformly from $\{1, 2, \ldots, \mathrm{N}\}$.

Note that, throughout the paper, $\mathrm{P}_\mathrm{N}(a|b)$ refers to the probability that a divides b, not any sort of conditional probability. Further, the notation $||$ is used in the standard manner: $p^a||q$ iff p^a is the largest power of p which divides q.

We want to find $\mathrm{P}_{cd}(d|\mathrm{X}_n)$. To do this, we first need a multiplicativity result:

Proposition 2.1. *Suppose $d = \prod_{i=1}^{\ell} p_i^{r_i}$. If c is a positive integer constant, then*

$$\mathrm{P}_{cd}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell} \mathrm{P}_{cd}(p_i^{r_i}|\mathrm{X}_n).$$

Remark 2.2. Our **Proposition 2.1** seems intuitive at first, as a natural extension of the Chinese Remainder Theorem. However, it is worth noting that we run into a counterexample very quickly if we try to show a more general version of this holds: $\mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell} \mathrm{P}_\mathrm{N}(p_i^{r_i}|\mathrm{X}_n)$.

We examine the case where $\mathrm{N} = 100$, $d = 5 \cdot 2^{30}$, $n = 5$. First, we evaluate $\mathrm{P}_\mathrm{N}(d|\mathrm{X}_n)$. There is only one possible tuple of values for $(x_1, \ldots, x_5)$ such that $2^{30}|\mathrm{X}_n$: namely, $(64, 64, 64, 64, 64)$. With any other tuple of values for $(x_1, \ldots, x_5)$, the product of the five variables does not contain enough powers of 2 in its prime factorization to be divisible by 2^{30}. However, if $(x_1, \ldots, x_5) = (64, \ldots, 64)$, then $5 \nmid \mathrm{X}_n$, and in turn $d \nmid \mathrm{X}_n$; since for no values of $x_1, \ldots, x_5$ can d divide X_n, we have $\mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) = 0$.

Now, we evaluate $\prod_{i=1}^{\ell} \mathrm{P}_\mathrm{N}(p_i^{r_i}|\mathrm{X}_n) = \mathrm{P}_\mathrm{N}(2^{30}|\mathrm{X}_n) \cdot \mathrm{P}_\mathrm{N}(5|\mathrm{X}_n)$. As only one tuple of possible values for $(x_1, \ldots, x_5)$ (out of 100^5 possible such tuples) satisfies $2^{30}|\mathrm{X}_n$, we have $\mathrm{P}_\mathrm{N}(2^{30}|\mathrm{X}_n) = \frac{1}{10^{10}}$. Further, the probability of none of the integers $x_1, \ldots, x_5$ being divisible by 5 is $\left(\frac{4}{5}\right)^5$, so the probability that at least one of them is divisible by 5 (and therefore $5|\mathrm{X}_n$) is $1 - \left(\frac{4}{5}\right)^5$. So, $\mathrm{P}_\mathrm{N}(2^{30}|\mathrm{X}_n) \cdot \mathrm{P}_\mathrm{N}(5|\mathrm{X}_n) = \frac{1}{10^{10}} \cdot \left(1 - \left(\frac{4}{5}\right)^5\right) \neq 0$. Therefore, in this case, $\mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) \neq \prod_{i=1}^{\ell} \mathrm{P}_\mathrm{N}(p_i^{r_i}|\mathrm{X}_n)$.

Our concern is that we could somehow run into similar issues as in the example above even with the conditions laid out in **Proposition 2.1**. So, we will prove **Proposition 2.1** rigorously. Rather than attempting to apply the Chinese Remainder Theorem to X_n, we will inductively apply it to each x_j. To avoid such issues as in the example above, we will set conditions like $a < r$ in some of the lemmas below. These variables will be explained shortly.

To prove the proposition, we need the two lemmas below:

Lemma 2.3. *Let p be a prime and suppose $p^r||d$, and $a < r$. Then,*

$$\mathrm{P}_{cd}(p^a||\mathrm{X}_{k+1}) = \sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j}||\mathrm{X}_k) \cdot \left(1 - \frac{1}{p}\right) \cdot \frac{1}{p^j}.$$

Proof. First, observe that

$$\begin{aligned} \mathrm{P}_{cd}(p^a || \mathrm{X}_{k+1}) &= \mathrm{P}_{cd}\left(p^a || \prod_{i=1}^{k+1} x_{k+1}\right) \\ &= \mathrm{P}_{cd}(p^a || \mathrm{X}_k \cdot x_{k+1}) \\ &= \sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j} || \mathrm{X}_k \text{ and } p^j || x_{k+1}). \end{aligned}$$

Since all of the x_i's are independent and identically distributed (i.i.d.), the two events $p^{a-j} || \mathrm{X}_k$ and $p^j || x_{k+1}$ are independent. So, we have

$$\sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j} || \mathrm{X}_k \text{ and } p^j || x_{k+1}) = \sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j} || \mathrm{X}_k) \mathrm{P}_{cd}(p^j || x_{k+1}).$$

Now, we want to find $\mathrm{P}_{cd}(p^j || x_{k+1})$ when $0 \leq j \leq a$. We do this via counting:

x_{k+1} is chosen randomly and uniformly from the set $\{1, 2, \ldots, cd\}$. Of these integers, $p^j, 2p^j, \ldots,$ and $\left\lfloor \frac{cd}{p^j} \right\rfloor p^j$ are divisible by p^j. Further, of these, $p^{j+1}, 2p^{j+1}, \ldots,$ and $\left\lfloor \frac{cd}{p^{j+1}} \right\rfloor p^{j+1}$ are also divisible by p^{j+1}. So, $\left\lfloor \frac{cd}{p^j} \right\rfloor - \left\lfloor \frac{cd}{p^{j+1}} \right\rfloor$ of the possible values of x_{k+1} are divisible by p^j but not p^{j+1}.

Note, though, that both $\frac{cd}{p^j}$ and $\frac{cd}{p^{j+1}}$ are integers: since $j \leq a < r$, we know $j + 1 \leq r$. Then, $p^{j+1} | d$, so $\frac{cd}{p^j}$ and $\frac{cd}{p^{j+1}}$ are integers.

Then, since x_{k+1} takes on one of cd values uniformly, and $\frac{cd}{p^j} - \frac{cd}{p^{j+1}}$ of these values are such that $p^j || x_{k+1}$, we have

$$\begin{aligned} \mathrm{P}_{cd}(p^j || x_{k+1}) &= \frac{\frac{cd}{p^j} - \frac{cd}{p^{j+1}}}{cd} \\ &= \frac{1}{p^j} - \frac{1}{p^{j+1}}. \end{aligned}$$

Therefore,

$$\mathrm{P}_{cd}(p^j || x_{k+1}) = \left(1 - \frac{1}{p}\right) \cdot \frac{1}{p^j}. \tag{1}$$

Now, we can use this result in our earlier equation to find

$$\sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j} || \mathrm{X}_k) \mathrm{P}_{cd}(p^j || x_{k+1}) = \sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j} || \mathrm{X}_k) \cdot \left(1 - \frac{1}{p}\right) \cdot \frac{1}{p^j}$$

and we have shown

$$\mathrm{P}_{cd}(p^a || \mathrm{X}_{k+1}) = \sum_{j=0}^{a} \mathrm{P}_{cd}(p^{a-j} || \mathrm{X}_k) \cdot \left(1 - \frac{1}{p}\right) \cdot \frac{1}{p^j}.$$

□

We need one more lemma before we can show multiplicativity:

Lemma 2.4. *Suppose $d = \prod_{i=1}^{\ell} p_i^{r_i}$, and let $\mathrm{S} \subseteq \{1,2,\dots,\ell\}$.*

Let $a_i < r_i$ for all $i \in \mathrm{S}$. Then,

$$\mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_n, \forall i \in \mathrm{S}) = \prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_n).$$

Proof. We show $\mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_n, \forall i \in \mathrm{S}) = \prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_n)$, when $a_i < r_i$ for all i, via induction on n:

First, we want to show **Lemma 2.4** holds in the base case, $n = 1$. Here, $\mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_n, \forall i \in \mathrm{S}) = \mathrm{P}_{cd}(p_i^{a_i} || x_1, \forall i \in \mathrm{S})$, where x_1 is an integer chosen uniformly from $\{1,2,\dots,cd\}$. By the Chinese Remainder Theorem, these events are independent, so we have

$$\mathrm{P}_{cd}(p_i^{a_i} || x_1, \forall i \in \mathrm{S}) = \prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i} || x_1) = \prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_n).$$

So, **Lemma 2.4** holds when $n = 1$.

Now, suppose **Lemma 2.4** holds when $n = k$. We want to show **Lemma 2.4** holds when $n = k+1$. Let's start with the right-hand side.

First, we focus on the $\mathrm{P}_{cd}(p_1^{a_1} || \mathrm{X}_{k+1})$ term on the right-hand side of **Lemma 2.3** with $n = k+1$. We apply **Lemma 2.3** for

$$\mathrm{P}_{cd}(p_1^{a_1} || \mathrm{X}_{k+1}) = \sum_{j=0}^{a_1} \mathrm{P}_{cd}(p_1^{a_1 - j} || \mathrm{X}_k) \cdot \left(1 - \frac{1}{p_1}\right) \cdot \frac{1}{p_1^j}$$

and apply this without loss of generality over all i in S to find

$$\prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i} || \mathrm{X}_{k+1}) = \prod_{i \in \mathrm{S}} \sum_{j=0}^{a_i} \mathrm{P}_{cd}(p_i^{a_i - j} || \mathrm{X}_k) \cdot \left(1 - \frac{1}{p_i}\right) \cdot \frac{1}{p_i^j}$$

$$= \sum_{m \in \mathrm{S}} \sum_{0 \le j_m \le a_m} \prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i - j_i} || \mathrm{X}_k) \cdot \left(1 - \frac{1}{p_i}\right) \cdot \frac{1}{p_i^{j_i}}.$$

Now, we turn our attention to the term on the inside of the sums,

$$\prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i - j_i} || \mathrm{X}_k) \cdot \left(1 - \frac{1}{p_i}\right) \cdot \frac{1}{p_i^{j_i}}.$$

First, since $a_i - j_i \le a_i < r_i$ for all i, we can apply our inductive hypothesis to say

$$\prod_{i \in \mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i - j_i} || \mathrm{X}_k) = \mathrm{P}_{cd}(p_i^{a_i - j_i} || \mathrm{X}_k, \forall i \in \mathrm{S}).$$

Additionally, by Equation (1) from the proof of **Lemma 2.3**, we have $\left(1-\frac{1}{p_i}\right)\cdot\frac{1}{p_i^{j_i}} = \mathrm{P}_{cd}(p_i^{j_i}||x_{k+1})$, for all i. So,

$$\prod_{i\in\mathrm{S}} \mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k)\cdot\left(1-\frac{1}{p_i}\right)\cdot\frac{1}{p_i^{j_i}} = \mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k, \forall i\in\mathrm{S})\prod_{i\in\mathrm{S}}\mathrm{P}_{cd}(p_i^{j_i}||x_{k+1}).$$

Further, by the Chinese Remainder Theorem, $\prod_{i\in\mathrm{S}}\mathrm{P}_{cd}(p_i^{j_i}||x_{k+1}) = \mathrm{P}_{cd}(p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S})$. So,

$$\mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k, \forall i\in\mathrm{S})\prod_{i\in\mathrm{S}}\mathrm{P}_{cd}(p_i^{j_i}||x_{k+1}) = \mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k, \forall i\in\mathrm{S})\cdot\mathrm{P}_{cd}(p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S})$$

$$= \mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k \text{ and } p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S})$$

as these events are independent.

Now, consider the event $p_i^{a_i-j_i}||\mathrm{X}_k \cap p_i^{j_i}||x_{k+1}$. Since $p_i^{a_i-j_i}||\mathrm{X}_k$ and $p_i^{j_i}||x_{k+1}$ if and only if $p_i^{a_i}||\mathrm{X}_{k+1}$ and $p_i^{j_i}||x_{k+1}$, we find

$$\mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k \text{ and } p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S}) = \mathrm{P}_{cd}(p_i^{a_i}||\mathrm{X}_{k+1} \text{ and } p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S}).$$

So, returning to our original sum, we can now find

$$\sum_{m\in\mathrm{S}}\sum_{0\le j_m\le a_m}\prod_{i\in\mathrm{S}}\mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k)\cdot\left(1-\frac{1}{p_i}\right)\cdot\frac{1}{p_i^{j_i}}$$

$$= \sum_{m\in\mathrm{S}}\sum_{0\le j_m\le a_m}\mathrm{P}_{cd}(p_i^{a_i}||\mathrm{X}_{k+1} \text{ and } p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S}).$$

Each term represents the probability that $p_i^{a_i}||\mathrm{X}_{k+1}$ and $p_i^{j_i}||x_{k+1}$ for $j_i = 0,1,\ldots,a_i$, for all $i\in\mathrm{S}$. Note that this encompasses all cases such that $p_i^{a_i}||\mathrm{X}_{k+1}$ is true, as in this event it must be that $p_i^{j_i}||x_{k+1}$ for one of $j_i = 0,1,\ldots,a_i$. So,

$$\sum_{m\in\mathrm{S}}\sum_{0\le j_m\le a_m}\mathrm{P}_{cd}(p_i^{a_i}||\mathrm{X}_{k+1} \text{ and } p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S})$$

$$= \sum_{m\in\mathrm{S}}\sum_{0\le j_m\le a_m}\mathrm{P}_{cd}(p_i^{a_i-j_i}||\mathrm{X}_k \text{ and } p_i^{j_i}||x_{k+1}, \forall i\in\mathrm{S}) = \mathrm{P}_{cd}(p_i^{a_i}||\mathrm{X}_{k+1}, \forall i\in\mathrm{S})$$

and we have shown that **Lemma 2.4** holds for $n = k$ implies that **Lemma 2.4** holds for $n = k+1$.

By induction, therefore, we have shown $\mathrm{P}_{cd}(p_i^{a_i}||\mathrm{X}_n, \forall i \in \mathrm{S}) = \prod_{i\in\mathrm{S}}\mathrm{P}_{cd}(p_i^{a_i}||\mathrm{X}_n)$, when $a_i < r_i$ for all $i\in\mathrm{S}$. □

With the two lemmas above, we can now show multiplicativity: -

Proof of Proposition 2.1. Let's start with the right-hand side.

Note first that $p_i^{r_i}|\mathrm{X}_n$ if and only if $p_i^0||\mathrm{X}_n$, $p_i^1||\mathrm{X}_n$, ..., and $p_i^{r_i-1}||\mathrm{X}_n$ are all not true. More formally, the event $p_i^{r_i}|\mathrm{X}_n$ is equivalent to the complement of the union of the events $p_i^a||\mathrm{X}_n$ for all $a \in \{0,1,\dots,r_i-1\}$. So,

$$\prod_{i=1}^{\ell} \mathrm{P}_{cd}(p_i^{r_i}|\mathrm{X}_n) = \prod_{i=1}^{\ell}\left(1-\sum_{j=0}^{r_i-1}\mathrm{P}_{cd}(p_i^j||\mathrm{X}_n)\right)$$

$$= \sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\prod_{i\in\mathrm{I}}\sum_{j=0}^{r_i-1}\mathrm{P}_{cd}(p_i^j||\mathrm{X}_n)$$

$$= \sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\sum_{m\in\mathrm{I}}\sum_{0\le j_m\le r_m-1}\prod_{i\in\mathrm{I}}\mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n).$$

By **Lemma 2.4**, $\prod_{i\in\mathrm{I}}\mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n) = \mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n, \forall i \in \mathrm{I})$, and we can apply this to the line above to obtain

$$\sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\sum_{m\in\mathrm{I}}\sum_{0\le j_m\le r_m-1}\prod_{i\in\mathrm{I}}\mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n)$$

$$= \sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\sum_{m\in\mathrm{I}}\sum_{0\le j_m\le r_m-1}\mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n, \forall i \in \mathrm{I}).$$

Now, observe that $p_i^{r_i}\nmid\mathrm{X}_n$ if and only if $p_i^{j_i}||\mathrm{X}_n$ for some $0 \le j_i \le r_i - 1$. So, we have $\sum_{m\in\mathrm{I}}\sum_{0\le j_m\le r_m-1}\mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n, \forall i\in\mathrm{I}) = \mathrm{P}_{cd}(p_i^{r_i}\nmid\mathrm{X}_n, \forall i \in \mathrm{I})$, and therefore

$$\sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\sum_{m\in\mathrm{I}}\sum_{0\le j_m\le r_m-1}\mathrm{P}_{cd}(p_i^{j_i}||\mathrm{X}_n, \forall i \in \mathrm{I}) = \sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\mathrm{P}_{cd}(p_i^{r_i}\nmid\mathrm{X}_n, \forall i \in \mathrm{I}).$$

Then, by the inclusion-exclusion principle,

$$\sum_{\mathrm{I}\subseteq\{1,2,\dots,\ell\}}(-1)^{|\mathrm{I}|}\mathrm{P}_{cd}(p_i^{r_i}\nmid\mathrm{X}_n, \forall i \in \mathrm{I}) = \mathrm{P}_{cd}(p_i^{r_i}|\mathrm{X}_n, \forall i \in \{1,\dots,\ell\})$$

$$= \mathrm{P}_{cd}(d|\mathrm{X}_n).$$

We have shown that $\mathrm{P}_{cd}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell}\mathrm{P}_{cd}(p_i^{r_i}|\mathrm{X}_n)$. This completes the proof of **Proposition 2.1**. □

Now that we have shown that the formula for $\mathrm{P}_{cd}(d|\mathrm{X}_n)$ is multiplicative, we need to find its value when d is a prime power. Then, for any d, we can simply use **Proposition 2.1** to obtain the value of $\mathrm{P}_{cd}(d|\mathrm{X}_n)$.

Lemma 2.5. *Suppose that n is a fixed positive integer, p is a prime number, and $r \in \mathbb{N}$ is such that $p^r || d$. Then,*

$$\mathrm{P}_{cd}(p^r|\mathrm{X}_n) = 1 - \sum_{k=0}^{r-1} \binom{n+k-1}{n-1} \left(1-\frac{1}{p}\right)^n \left(\frac{1}{p}\right)^k.$$

Proof. Let m_j be such that $p^{m_j}||x_j$, and $m = \sum_{j=1}^n m_j$.

Observe that

$$\mathrm{P}_{cd}(p^r|\mathrm{X}_n) = \mathrm{P}_{cd}(m \geq r) = 1 - \mathrm{P}_{cd}(m < r) = 1 - \sum_{k=0}^{r-1} \mathrm{P}_{cd}(m = k).$$

So, we can obtain $\mathrm{P}_{cd}(p^r|\mathrm{X}_n)$ by obtaining the values of $\mathrm{P}_{cd}(m = k)$ for all k integers less than r. This avoids the problems described in Remark 2.2.

Define the generating function

$$\mathrm{F}(\mathrm{T}) = \sum_{k=0}^{\infty} \mathrm{P}_{cd}(m = k)\mathrm{T}^k.$$

Then,

$$\sum_{k=0}^{\infty} \mathrm{P}_{cd}(m = k)\mathrm{T}^k = \sum_{k=0}^{\infty} \sum_{k_1+\ldots+k_n=k} \mathrm{P}_{cd}\big((m_1,\ldots,m_n) = (k_1,\ldots,k_n)\big)\mathrm{T}^k$$

$$= \sum_{k=0}^{\infty} \sum_{k_1+\ldots+k_n=k} \mathrm{P}_{cd}(m_1 = k_1)\mathrm{T}^{k_1} \cdots \mathrm{P}_{cd}(m_n = k_n)\mathrm{T}^{k_n} \qquad \text{(as the } x_i\text{'s are i.i.d.)}$$

$$= \sum_{0 \leq k_1+\ldots+k_n < \infty} \mathrm{P}_{cd}(m_1 = k_1)\mathrm{T}^{k_1} \cdots \mathrm{P}_{cd}(m_n = k_n)\mathrm{T}^{k_n}$$

$$= \prod_{i=1}^{n} \sum_{k_i=0}^{\infty} \mathrm{P}_{cd}(m_i = k_i)\mathrm{T}^{k_i}$$

$$= \prod_{i=1}^{n} \sum_{k=0}^{\infty} \mathrm{P}_{cd}(m_i = k)\mathrm{T}^{k}.$$

We use this generating function only to find coefficients of terms of degree less than r. As discussed above, we do not care about coefficients beyond T^{r-1}, nor do we want to work with them because of problems similar to those described in Remark 2.2. We can remove unneeded terms by saying

$$\prod_{i=1}^{n} \sum_{k=0}^{\infty} \mathrm{P}_{cd}(m_i = k)\mathrm{T}^k = \left(\prod_{i=1}^{n} \sum_{k=0}^{r-1} \mathrm{P}_{cd}(m_i = k)\mathrm{T}^k\right) + \mathcal{O}(\mathrm{T}^r)$$

where $\mathcal{O}(\mathrm{T}^r)$ denotes an unspecified sum of powers T^r and higher in a power series.

Now, we can apply Equation (1) from the proof of **Lemma 2.3** to say $\mathrm{P}_{cd}(m_i = k) = \frac{1}{p^k}\left(1-\frac{1}{p}\right)$, so

$$\left(\prod_{i=1}^{n}\sum_{k=0}^{r-1}\mathrm{P}_{cd}(m_i = k)\mathrm{T}^k\right) + \mathscr{O}(\mathrm{T}^r) = \left(\prod_{i=1}^{n}\sum_{k=0}^{r-1}\frac{1}{p^k}\left(1-\frac{1}{p}\right)\mathrm{T}^k\right) + \mathscr{O}(\mathrm{T}^r).$$

We can "pull terms out" of the $\mathscr{O}(\mathrm{T}^r)$ to change the summation bound from $r-1$ to ∞ to find

$$\begin{aligned}\left(\prod_{i=1}^{n}\sum_{k=0}^{r-1}\frac{1}{p^k}\left(1-\frac{1}{p}\right)\mathrm{T}^k\right) + \mathscr{O}(\mathrm{T}^r) &= \left(\prod_{i=1}^{n}\sum_{k=0}^{\infty}\frac{1}{p^k}\left(1-\frac{1}{p}\right)\mathrm{T}^k\right) + \mathscr{O}(\mathrm{T}^r)\\ &= \left(\sum_{k=0}^{\infty}\frac{1}{p^k}\left(1-\frac{1}{p}\right)\mathrm{T}^k\right)^n + \mathscr{O}(\mathrm{T}^r)\\ &= \left(1-\frac{1}{p}\right)^n\left(\sum_{k=0}^{\infty}\frac{1}{p^k}\mathrm{T}^k\right)^n + \mathscr{O}(\mathrm{T}^r).\end{aligned}$$

Now, let $\mathrm{S} = \frac{\mathrm{T}}{p}$. Then,

$$\begin{aligned}\left(1-\frac{1}{p}\right)^n\left(\sum_{k=0}^{\infty}\frac{1}{p^k}\mathrm{T}^k\right)^n + \mathscr{O}(\mathrm{T}^r) &= \left(1-\frac{1}{p}\right)^n\left(\sum_{k=0}^{\infty}\mathrm{S}^k\right)^n + \mathscr{O}(\mathrm{T}^r)\\ &= \left(1-\frac{1}{p}\right)^n\left(\frac{1}{1-\mathrm{S}}\right)^n + \mathscr{O}(\mathrm{T}^r)\end{aligned}$$

by sum of a geometric series.

From the binomial expansion of $(1-\mathrm{S})^{-n} = \sum_{k=0}^{\infty}\binom{-n}{k}(-\mathrm{S})^k$, and the identity $\binom{-n}{k} = (-1)^k\binom{n+k-1}{k}$, we have found the generating function for $\mathrm{P}_{cd}(m = k)$ as a function of the formal variable T to be

$$\mathrm{F}(\mathrm{T}) = \sum_{k=0}^{\infty}\binom{n+k-1}{n-1}\left(1-\frac{1}{p}\right)^n\left(\frac{\mathrm{T}}{p}\right)^k + \mathscr{O}(\mathrm{T}^r).$$

We can now use the generating function to find $\mathrm{P}_{cd}(m \geq r)$:

$$\mathrm{P}_{cd}(m \geq r) = 1 - \mathrm{P}_{cd}(m < r) = 1 - \sum_{k=0}^{r-1}\mathrm{P}_{cd}(m = k) = 1 - \sum_{k=0}^{r-1}\binom{n+k-1}{n-1}\left(1-\frac{1}{p}\right)^n\left(\frac{1}{p}\right)^k.$$

Equivalently,

$$\mathrm{P}_{cd}(p^r | \mathrm{X}_n) = 1 - \sum_{k=0}^{r-1}\binom{n+k-1}{n-1}\left(1-\frac{1}{p}\right)^n\left(\frac{1}{p}\right)^k.$$

□

With this lemma, we can now obtain a formula as desired in the finite case:

Theorem 2.6. *Suppose $d = \prod_{i=1}^{\ell} p_i^{r_i}$. If c is a positive integer constant, then*

$$\mathrm{P}_{cd}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell} \left(1 - \sum_{k=0}^{r_i - 1} \binom{n+k-1}{n-1} \left(1-\frac{1}{p}\right)^n \left(\frac{1}{p}\right)^k\right).$$

Proof. By **Proposition 2.1**,

$$\mathrm{P}_{cd}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell} \mathrm{P}_{cd}(p_i^{r_i}|\mathrm{X}_n)$$

and we can apply **Lemma 2.5** to obtain

$$\prod_{i=1}^{\ell} \mathrm{P}_{cd}(p_i^{r_i}|\mathrm{X}_n) = \prod_{i=1}^{\ell} \left(1 - \sum_{k=0}^{r_i - 1} \binom{n+k-1}{n-1} \left(1-\frac{1}{p}\right)^n \left(\frac{1}{p}\right)^k\right).$$

□

We have now found a formula for the probability that d divides the product of n integers chosen randomly from $\{1,\ldots,d\}$ - simply apply **Theorem 2.6** with $c = 1$.

We can restate the result from **Theorem 2.6** to obtain the following, which will be useful in **Corollary 4.1**:

Corollary 2.7. *Suppose $d = \prod_{i=1}^{\ell} p_i^{r_i}$. If c is a positive integer constant, then*

$$\mathrm{P}_{cd}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell} \sum_{k=0}^{n-1} \binom{r_i+n-1}{k} \left(1-\frac{1}{p_i}\right)^k \left(\frac{1}{p_i}\right)^{r_i+n-1-k}.$$

This corollary follows from setting $x = \frac{1}{p}$ in the following lemma.

Lemma 2.8. *Let n be a positive integer, r a non-negative integer, and x a real number with $0 < x < 1$. Then,*

$$1 - \sum_{k=0}^{r-1} \binom{n+k-1}{n-1} (1-x)^n x^k = \sum_{k=0}^{n-1} \binom{r+n-1}{k} (1-x)^k x^{r+n-1-k}.$$

Proof. We show this via induction on r.

First, we want to show **Lemma 2.8** holds in the base case, $r = 0$. Here, the sum is empty, so

$$1 - \sum_{k=0}^{r-1} \binom{n+k-1}{n-1} (1-x)^n x^k = 1 = (x + (1-x))^{n-1}.$$

Then, as the Binomial Theorem states that for variables x and y and positive integer n, $(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j$ [4, pp. 12-13], we have

$$(x+(1-x))^{n-1} = \sum_{k=0}^{n-1} \binom{r+n-1}{k} (1-x)^k x^{r+n-1-k}.$$

and **Lemma 2.8** holds in the base case.

Now, suppose **Lemma 2.8** holds when $r = j$, for some non-negative integer j. We want to show that **Lemma 2.8** holds when $r = j+1$.

Since $\binom{j+n}{k} = \binom{j+n-1}{k-1} + \binom{j+n-1}{k}$, for $k \geq 1$, we have (starting from the right side)

$$\sum_{k=0}^{n-1} \binom{j+n}{k} (1-x)^k x^{j+n-k}$$

$$= x^{j+n} + \sum_{k=1}^{n-1} \binom{j+n-1}{k-1} (1-x)^k x^{j+n-k} + \sum_{k=1}^{n-1} \binom{j+n-1}{k} (1-x)^k x^{j+n-k}. \tag{2}$$

With a change of bounds, we can see that the first sum is

$$\sum_{k=1}^{n-1} \binom{j+n-1}{k-1} (1-x)^k x^{j+n-k} = \sum_{k=0}^{n-2} \binom{j+n-1}{k} (1-x)^{k+1} x^{j+n-(k+1)}$$

$$= (1-x) \left(\sum_{k=0}^{n-1} \binom{j+n-1}{k} (1-x)^k x^{j+n-(k+1)} - \binom{j+n-1}{n-1} (1-x)^k x^j \right).$$

Then, since **Lemma 2.8** holds when $r = j$, we find the above is equal to

$$(1-x) \left(1 - \sum_{k=0}^{j-1} \binom{n+k-1}{n-1} (1-x)^n x^k - \binom{j+n-1}{n-1} (1-x)^k x^j \right).$$

The second sum, meanwhile, is

$$\sum_{k=1}^{n-1} \binom{j+n-1}{k} (1-x)^k x^{j+n-k} = \sum_{k=0}^{n-1} \binom{j+n-1}{k} (1-x)^k x^{j+n-k} - x^{j+n}$$

$$= x \left(\sum_{k=0}^{n-1} \binom{j+n-1}{k} (1-x)^k x^{j+n-k-1} \right) - x^{j+n}.$$

And since **Lemma 2.8** holds when $r = j$, we find the above is equal to

$$x \left(1 - \sum_{k=0}^{j-1} \binom{n+k-1}{n-1} (1-x)^n x^k \right) - x^{j+n}.$$

Returning to Equation (2), we see

$$x^{j+n}+\sum_{k=1}^{n-1}\binom{j+n-1}{k-1}(1-x)^k x^{j+n-k}+\sum_{k=1}^{n-1}\binom{j+n-1}{k}(1-x)^k x^{j+n-k}$$

$$\begin{aligned}&=(1-x)\left(1-\sum_{k=0}^{j-1}\binom{n+k-1}{n-1}(1-x)^n x^k-\binom{j+n-1}{n-1}(1-x)^k x^j\right)\\&\quad+x\left(1-\sum_{k=0}^{j-1}\binom{n+k-1}{n-1}(1-x)^n x^k\right)\\&=1-\sum_{k=0}^{j-1}\binom{n+k-1}{n-1}(1-x)^n x^k-(1-x)\binom{j+n-1}{n-1}(1-x)^k x^j\\&=1-\sum_{k=0}^{j}\binom{n+k-1}{n-1}(1-x)^n x^k\end{aligned}$$

and we have shown that **Lemma 2.8** holds for $r=j$ implies that **Lemma 2.8** holds for $r=j+1$.

By induction, therefore, we have proved the lemma. □

3 The Infinite Case

Theorem 3.1. *Suppose* $d=\prod_{i=1}^{\ell}p_i^{r_i}$. *Then,*

$$\lim_{N\to\infty}P_N(d|X_n)=\prod_{i=1}^{\ell}\left(1-\sum_{k=0}^{r-1}\binom{n+k-1}{n-1}(1-\frac{1}{p})^n(\frac{1}{p})^k\right).$$

Proof. Let x_i be uniformly chosen from $\{1,\dots,N\}$ for all $i\in\{1,\dots,n\}$, and let $X_n=\prod_{i=1}^{n}x_i$.

Let S_N be the set of all possible ordered tuples of values of $(x_1,\dots,x_n)$ such that $d|X_n$. Note that each tuple in S_N is of equal likelihood, as each x_i is chosen uniformly at random. Note further that N^n total tuples of values of $(x_1,\dots,x_n)$ are possible, as there are n elements in the tuple, each of which can take on any of N distinct values. So, clearly,

$$P_N(d|X_n)=\frac{|S_N|}{N^n}.$$

Now, let $c=\lfloor\frac{N}{d}\rfloor$, so $cd\le N<(c+1)d$.

Note that $|S_{cd}|\le|S_N|\le|S_{(c+1)d}|$; this follows immediately from the definition of the sets.

Now,

$$\frac{|S_{cd}|}{((c+1)d)^n}\le\frac{|S_N|}{N^n}\le\frac{|S_{(c+1)d}|}{(cd)^n}.$$

Equivalently,

$$\left(\frac{c}{c+1}\right)^n \frac{|\mathrm{S}_{cd}|}{(cd)^n} \le \frac{|\mathrm{S}_\mathrm{N}|}{\mathrm{N}^n} \le \left(\frac{c+1}{c}\right)^n \frac{|\mathrm{S}_{(c+1)d}|}{((c+1)d)^n}.$$

Using $\mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) = \frac{|\mathrm{S}_\mathrm{N}|}{\mathrm{N}^n}$, we have

$$\left(\frac{c}{c+1}\right)^n \mathrm{P}_{cd}(d|\mathrm{X}_n) \le \frac{|\mathrm{S}_\mathrm{N}|}{\mathrm{N}^n} \le \left(\frac{c+1}{c}\right)^n \mathrm{P}_{(c+1)d}(d|\mathrm{X}_n).$$

Note that the right-hand side in **Theorem 2.6** is independent of c. Therefore, $\mathrm{P}_{cd}(d|\mathrm{X}_n) = \mathrm{P}_d(d|\mathrm{X}_n)$, and we obtain

$$\left(\frac{c}{c+1}\right)^n \mathrm{P}_d(d|\mathrm{X}_n) \le \frac{|\mathrm{S}_\mathrm{N}|}{\mathrm{N}^n} \le \left(\frac{c+1}{c}\right)^n \mathrm{P}_d(d|\mathrm{X}_n).$$

Let's now consider the limit case for the left-hand side; recall that $c = \lfloor \frac{\mathrm{N}}{d} \rfloor$:

$$\lim_{\mathrm{N}\to\infty} \left(\frac{c}{c+1}\right)^n \mathrm{P}_d(d|\mathrm{X}_n) = \lim_{\mathrm{N}\to\infty} \left(\frac{\lfloor \frac{\mathrm{N}}{d} \rfloor}{\lfloor \frac{\mathrm{N}}{d} \rfloor + 1}\right)^n \mathrm{P}_d(d|\mathrm{X}_n) = \mathrm{P}_d(d|\mathrm{X}_n)$$

since $\lim_{\mathrm{N}\to\infty} \left(\frac{\lfloor \frac{\mathrm{N}}{d} \rfloor}{\lfloor \frac{\mathrm{N}}{d} \rfloor+1}\right)^n = 1$.

Similarly,

$$\lim_{\mathrm{N}\to\infty} \left(\frac{c+1}{c}\right)^n \mathrm{P}_d(d|\mathrm{X}_n) = \lim_{\mathrm{N}\to\infty} \left(\frac{\lfloor \frac{\mathrm{N}}{d} \rfloor + 1}{\lfloor \frac{\mathrm{N}}{d} \rfloor}\right)^n \mathrm{P}_d(d|\mathrm{X}_n) = \mathrm{P}_d(d|\mathrm{X}_n).$$

Since $\mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) = \frac{|\mathrm{S}_\mathrm{N}|}{\mathrm{N}^n}$ is bounded above and below by values for which the limit exists and is the same, the Squeeze Theorem says that

$$\lim_{\mathrm{N}\to\infty} \mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) = \mathrm{P}_d(d|\mathrm{X}_n).$$

So, by **Theorem 2.6**, we see

$$\lim_{\mathrm{N}\to\infty} \mathrm{P}_\mathrm{N}(d|\mathrm{X}_n) = \prod_{i=1}^{\ell} \left(1 - \sum_{k=0}^{r_i-1} \binom{n+k-1}{n-1} \left(1-\frac{1}{p_i}\right)^n \left(\frac{1}{p_i}\right)^k\right)$$

which is the desired formula. □

4 Special Cases

First, we can find a unique property with powers of 2: namely, the probability is $\frac{1}{2}$ that the product of n integers uniformly randomly chosen from $\{1,\dots,2^n\}$ is divisible by 2^n.

Corollary 4.1. *For a positive integer n,*

$$\mathrm{P}_{2^n}(2^n|\mathrm{X}_n) = \frac{1}{2}.$$

Proof. By **Corollary 2.7**,

$$\begin{aligned}
\mathrm{P}_{2^n}(2^n|\mathrm{X}_n) &= \sum_{k=0}^{n-1} \binom{n+n-1}{k} \left(1-\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n+n-1-k} \\
&= \sum_{k=0}^{n-1} \binom{2n-1}{k} \left(\frac{1}{2}\right)^{2n-1} \\
&= \left(\frac{1}{2}\right)^{2n-1} \sum_{k=0}^{n-1} \binom{2n-1}{k} \\
&= \left(\frac{1}{2}\right)^{2n} \left(\sum_{k=0}^{n-1} \binom{2n-1}{k} + \sum_{k=0}^{n-1} \binom{2n-1}{k}\right).
\end{aligned}$$

Since $\binom{2n-1}{k} = \binom{2n-1}{2n-1-k}$, this is simply the sum of the entries in the $(2n-1)^{\text{th}}$ row of Pascal's Triangle. Famously, the sum of all terms in row R of Pascal's Triangle is 2^{R}, so we obtain

$$\left(\frac{1}{2}\right)^{2n} 2^{2n-1} = \frac{1}{2}$$

as desired. □

Next, we can find a unique property pertinent to primes. Let p be a prime. The probability that the product of p integers uniformly randomly chosen from $\{1,\ldots,p\}$ divides p approaches $1-\frac{1}{e}$ as p approaches ∞.

Corollary 4.2. *Suppose we have p_j, the j^{th} prime. Then,*

$$\lim_{j\to\infty} \mathrm{P}_{p_j}(p_j|\mathrm{X}_{p_j}) = 1-\frac{1}{e}.$$

Proof. By **Theorem 2.6**,

$$\begin{aligned}
\mathrm{P}_{p_j}(p_j|\mathrm{X}_{p_j}) &= 1-\sum_{k=0}^{1-1} \binom{p_j+k-1}{p_j-1} \left(1-\frac{1}{p_j}\right)^{p_j} \left(\frac{1}{p_j}\right)^k \\
&= 1-\left(1-\frac{1}{p_j}\right)^{p_j}.
\end{aligned}$$

Then,

$$\lim_{j\to\infty} \mathrm{P}_{p_j}(p_j|\mathrm{X}_{p_j}) = \lim_{j\to\infty}\left(1-\left(1-\frac{1}{p_j}\right)^{p_j}\right).$$

Recall the well-known formula $\lim_{n\to\infty}(1-\frac{1}{n})^n = \frac{1}{e}$ [2, pp. 77, 376]. Since $j \to \infty$ implies $p_j \to \infty$, we have $\lim_{j\to\infty}\left(1-\frac{1}{p_j}\right)^{p_j} = e^{-1}$, so

$$\lim_{j\to\infty} \mathrm{P}_{p_j}(p_j|\mathrm{X}_{p_j}) = 1-\frac{1}{e}.$$

This, of course, could also be proved by observing that for a prime p,

$$\mathrm{P}_p(p\nmid \mathrm{X}_p) = \prod_{i=1}^{p} \mathrm{P}_p(p\nmid x_i) = \left(1-\frac{1}{p}\right)^p.$$

□

5 A Primality Test

From the above theorems, we obtain an amusing but extremely impractical primality test. To avoid weaker results for small numbers, we restrict to integers greater than or equal to 10. To implement this test, we must first define an upper bound for $\mathrm{P}_d(d|\mathrm{X}_n)$ when d is prime, and a lower bound when d is not prime.

Proposition 5.1. *If $d \geq 10$ is prime,*

$$\mathrm{P}_d(d|\mathrm{X}_d) < 0.65.$$

Proof. First, note that since $d \geq 10$ is prime, $d \neq 10$. So, $d \geq 11$.

Since d is prime, by **Theorem 2.6**,

$$\mathrm{P}_d(d|\mathrm{X}_d) = 1-\sum_{k=0}^{1-1}\binom{d+k-1}{d-1}\left(1-\frac{1}{d}\right)^d\left(\frac{1}{d}\right)^k$$

$$= 1-\left(1-\frac{1}{d}\right)^d.$$

We need the following lemma.

Lemma 5.2. *Let $0<x<1$. Then, $(1-x)^{\frac{1}{x}} < e^{-1}$, and $(1-x)^{\frac{1}{x}}$ is decreasing as x increases.*

Proof. It is well known that $\lim_{x\to 0^+}(1-x)^{\frac{1}{x}} = e^{-1}$. Now, it suffices to show that $(1-x)^{\frac{1}{x}}$ is decreasing as x increases.

Further, it suffices to show $g(x) = \log\left((1-x)^{\frac{1}{x}}\right) = \frac{\log(1-x)}{x}$ is decreasing. By the quotient rule,

$$g'(x) = \frac{\frac{-x}{1-x}-\log(1-x)}{x^2}.$$

To show $g'(x)$ is negative, it now suffices to show that its numerator, $n(x) = \frac{-x}{1-x} - \log(1-x)$, is negative.

First, note $n'(x) = \frac{-1}{(1-x)^2} + \frac{1}{1-x} = \frac{-x}{(1-x)^2} < 0$ for $0 < x < 1$. So, $n(x)$ is decreasing.

Then, observe $n(0) = 0$. Since $n(x)$ is decreasing, $n(x)$ is negative for all $0 < x < 1$, and we are done. □

We restate the above lemma as $(1-x)^{\frac{1}{x}}$ is increasing as x decreases. So, we find that $(1-\frac{1}{d})^d$ is increasing as $d > 1$ increases. Equivalently, $1 - \left(1-\frac{1}{d}\right)^d$ is decreasing as d increases. Therefore, when $d \geq 11$,

$$1 - \left(1 - \frac{1}{d}\right)^d \leq 1 - \left(1 - \frac{1}{11}\right)^{11} < 0.65.$$

□

Proposition 5.3. *If* $d \geq 10$ *is composite,*

$$\mathrm{P}_d(d|\mathrm{X}_d) > 0.85$$

Proof. Let's first treat the case where d is a prime power: $d = p^r$ with $r \geq 2$. Within this case, we will further subdivide into three cases: $p = 2$, $r \geq 4$; $p \geq 5$, $r = 2$; and $p \geq 3$, $r \geq 3$. These three cases will cover all prime powers $d \geq 10$.

In the first case, $d = 2^r$, $r \geq 2$. We observe

$$\mathrm{P}_d(d|\mathrm{X}_d) = \mathrm{P}_{2^r}(2^r|\mathrm{X}_{2^r}) = 1 - \mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_{2^r})$$

Note that $r \geq 4$ implies $2^r \geq 4r$.

In order for $2^r \nmid \prod_{j=1}^{2^r} x_j$ to be true, we need

$$2^r \nmid \prod_{j=1}^{r} x_j,\ 2^r \nmid \prod_{j=r+1}^{2r} x_j,\ 2^r \nmid \prod_{j=2r+1}^{3r} x_j, \text{ and } 2^r \nmid \prod_{j=3r+1}^{4r} x_j$$

to all be true. This means $\mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_{2^r}) \leq \mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_r)^4$. So,

$$1 - \mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_{2^r}) \geq 1 - \mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_r)^4.$$

Using **Corollary 4.1**, we see $\mathrm{P}_{2^r}(2^r|\mathrm{X}_r) = \frac{1}{2}$, so $\mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_r) = \frac{1}{2}$. Then,

$$1 - \mathrm{P}_{2^r}(2^r \nmid \mathrm{X}_r)^4 = 1 - \left(\frac{1}{2}\right)^4 = \frac{15}{16} > 0.85$$

and we have completed the case where $d = 2^r$, $r \geq 2$.

Let’s now examine the case where $d = p^2$, $p \geq 5$. Here, by **Theorem 2.6**,

$$\mathrm{P}_d(d|\mathrm{X}_d) = 1 - \sum_{k=0}^{1} \binom{p^2+k-1}{p^2-1} \left(1-\frac{1}{p}\right)^{p^2} \left(\frac{1}{p}\right)^k$$

$$= 1 - \binom{p^2-1}{p^2-1}\left(1-\frac{1}{p}\right)^{p^2} - \binom{p^2}{p^2-1}\left(1-\frac{1}{p}\right)^{p^2} \frac{1}{p}$$

$$= 1 - \left(1-\frac{1}{p}\right)^{p^2} (1+p)$$

Then, by **Lemma 5.2**,

$$1 - \left(1-\frac{1}{p}\right)^{p^2} (1+p) = 1 - \left(\left(1-\frac{1}{p}\right)^{p}\right)^{p} (1+p) > 1 - e^{-p}(1+p).$$

The derivative of $1-e^{-p}(1+p)$ is $p^2 e^{-p} > 0$ for all $p > 0$. So, this function is strictly increasing, meaning

$$1 - e^{-p}(1+p) \geq 1 - 6e^{-5} \approx 0.96 > 0.85$$

for all $p \geq 5$. We have completed the case where $d = p^2$, $p \geq 5$.

Finally, we address the case where $d = p^r$, $p \geq 3$, and $r \geq 3$ via induction. While we have a base case for $p \geq 5$ with $r = 2$, we do not have one for $p = 3$. We therefore first show that $\mathrm{P}_d(d|\mathrm{X}_d) > 0.85$ when $d = 3^3$.

By **Theorem 2.6**,

$$\mathrm{P}_{3^3}(3^3|\mathrm{X}_{3^3}) = 1 - \sum_{k=0}^{2} \binom{26+k}{26} \left(1-\frac{1}{3}\right)^{27} \left(\frac{1}{3}\right)^k > 0.999 > 0.85.$$

Now, we can use induction in the following manner: suppose

$$\mathrm{P}_{p^r}(p^r|\mathrm{X}_{p^r}) > 0.85$$

for some $p \geq 3$, $r \geq 2$. We want to show that this is true for p, $r+1$.

Similarly to what we did in the powers of 2 case, observe the following. If $p^r | \prod_{j=1}^{2p^r} x_j$ and $p | \prod_{j=2p^r+1}^{3p^r} x_j$ are both true, then $p^{r+1} | \prod_{j=1}^{p^{r+1}} x_j$. Therefore,

$$\mathrm{P}_{p^{r+1}}(p^{r+1}|\mathrm{X}_{p^{r+1}}) \geq \mathrm{P}_{p^{r+1}}(p^r|\mathrm{X}_{2p^r})\mathrm{P}_{p^{r+1}}(p|\mathrm{X}_{p^r}).$$

Looking at $\mathrm{P}_{p^{r+1}}(p^r|\mathrm{X}_{2p^r}) = 1 - \mathrm{P}_{p^{r+1}}(p^r \nmid \mathrm{X}_{2p^r})$, we similarly observe that $p^r \nmid \mathrm{X}_{2p^r}$ can only be true if $p^r \nmid \prod_{j=1}^{p^r} x_j$ and $p^r \nmid \prod_{j=p^r+1}^{2p^r} x_j$. So, $\mathrm{P}_{p^{r+1}}(p^r \nmid \mathrm{X}_{2p^r}) \leq \mathrm{P}_{p^{r+1}}(p^r \nmid \mathrm{X}_{p^r})^2$. By our inductive hypothesis, $\mathrm{P}_{p^{r+1}}(p^r \nmid \mathrm{X}_{p^r})^2 < 0.15^2$, and

$$1 - \mathrm{P}_{p^{r+1}}(p^r \nmid \mathrm{X}_{2p^r}) > 1 - 0.15^2.$$

Further, we observe by **Theorem 2.6** that $\mathrm{P}_{p^{r+1}}(p|\mathrm{X}_{p^r}) = 1-\left(1-\frac{1}{p}\right)^{p^r} = 1-\left(\left(1-\frac{1}{p}\right)^{p}\right)^{p^{r-1}}$. By **Lemma 5.2**, we see $1-\left(\left(1-\frac{1}{p}\right)^{p}\right)^{p^{r-1}} > 1-e^{-p^{r-1}}$. Since $1-e^{-p^{r-1}}$ is increasing as p and r increase, and $p \geq 3$ and $r \geq 2$, we have $1-e^{-p^{r-1}} > 1-e^{-3}$. So,

$$\mathrm{P}_{p^{r+1}}(p|\mathrm{X}_{p^r}) > 1-e^{-3}$$

Returning to our original equation, we have

$$\begin{aligned}\mathrm{P}_{p^{r+1}}(p^{r+1}|\mathrm{X}_{p^{r+1}}) &\geq \mathrm{P}_{p^{r+1}}(p^r|\mathrm{X}_{2p^r})\mathrm{P}_{p^{r+1}}(p|\mathrm{X}_{p^r})\\ &= (1-\mathrm{P}_{p^{r+1}}(p^r \nmid \mathrm{X}_{2p^r}))\mathrm{P}_{p^{r+1}}(p|\mathrm{X}_{p^r})\\ &= (1-0.15^2)(1-e^{-3}) \approx 0.93 > 0.85.\end{aligned}$$

Therefore we have shown $\mathrm{P}(d|\mathrm{X}_d) > 0.85$ when d is a prime power that is not a prime.

Now assume d is not a prime power, so we can write $d = \prod_{i=1}^{\ell} p_i^{r_i}$ with $\ell \geq 2$. Then

$$\mathrm{P}_d(d|\mathrm{X}_d) = \prod_{i=1}^{\ell}\left(1-\mathrm{P}_d(p_i^{r_i} \nmid \mathrm{X}_d)\right).$$

We need to show $\mathrm{P}_d(p^r \nmid \mathrm{X}_d)$ is small.

If $p^r \nmid \mathrm{X}_d$, then p^r doesn't divide any of the individual factors X_d that are chosen independently from $\{1,2,\ldots,d\}$. Therefore,

$$\mathrm{P}_d(p^r \nmid \mathrm{X}_d) \leq \mathrm{P}_d(p^r \nmid x_1)^d = \left(1-p^{-r}\right)^d.$$

Then, by **Lemma 5.2**,

$$\left(1-p^{-r}\right)^d = \left(1-p^{-r}\right)^{p^r\cdot\frac{d}{p^r}} < e^{\frac{-d}{p^r}}.$$

The case $\ell = 2$ is now straightforward: we have $d = p_1^{r_1}p_2^{r_2}$. Without loss of generality, say $p_1 < p_2$; since $d \geq 10$, we have $p_1 \geq 2$, $p_2 \geq 5$. Then,

$$\begin{aligned}\mathrm{P}_d(d|\mathrm{X}_d) &= \left(1-\mathrm{P}_d(p_1^{r_1} \nmid \mathrm{X}_d)\right)\left(1-\mathrm{P}_d(p_2^{r_2} \nmid \mathrm{X}_d)\right)\\ &\geq (1-e^{-p_2^{r_2}})(1-e^{-p_1^{r_1}}) \geq (1-e^{-5})(1-e^{-2}) > 0.85.\end{aligned}$$

Henceforth, we may assume $\ell \geq 3$.

Lemma 5.4. *If $x \geq 3$, then $e^{-x} \leq 1.35/x^3$.*

Proof. x^3e^{-x} takes its maximum value of $27/e^3 \approx 1.34$ when $x = 3$. □

Lemma 5.5. *Suppose $x_i \geq 0$ for $1 \leq i \leq m$. Then*

$$\prod_{i=1}^{m}(1-x_i) \geq 1-\sum_{i=1}^{m}x_i.$$

Proof. The case $m = 1$ is trivial and the case $m = 2$ is the inequality

$$(1 - x_1)(1 - x_2) = 1 - (x_1 + x_2) + x_1 x_2 \geq 1 - (x_1 + x_2).$$

The general statement follows by induction. □

Suppose that $\frac{d}{p^r} = 2$ for some p. The only way this can happen is when $d = 2p^r$ for some odd prime p, which means $\ell = 2$, which is the case we just treated. Therefore, we assume $\frac{d}{p^r} \geq 3$ for all p. The first two lemmas yield

$$\mathrm{P}_d(p^r \nmid \mathrm{X}_d) \leq \left(1 - p^{-r}\right)^d \leq e^{\frac{-d}{p^r}} \leq 1.35\left(\frac{p^r}{d}\right)^3.$$

Therefore,

$$\mathrm{P}_d(d|\mathrm{X}_d) \geq \prod_{i=1}^{\ell}\left(1 - 1.35\left(\frac{p^r}{d}\right)^3\right) \geq 1 - 1.35\sum_{i=1}^{\ell}\frac{p_i^{3r_i}}{d^3}.$$

This last sum is obtained by taking the cubes of ℓ distinct integers, taking the products of all but one of them, and summing the reciprocals of these numbers. If we replace these numbers by the smaller numbers $2, 3, 4, \ldots, \ell + 1$, the reciprocals get larger, so

$$\sum_{i=1}^{\ell}\frac{p_i^{3r_i}}{d^3} \leq \sum_{j=2}^{\ell+1}\frac{j^3}{(\ell+1)!^3} \leq \sum_{j=1}^{\ell+1}\frac{j^3}{(\ell+1)!^3} = \frac{(\ell+1)^2(\ell+2)^2}{4(\ell+1)!^3},$$

where we have used the formula $\sum_{j=1}^{n} j^3 = \frac{n^2(n+1)^2}{4}$.

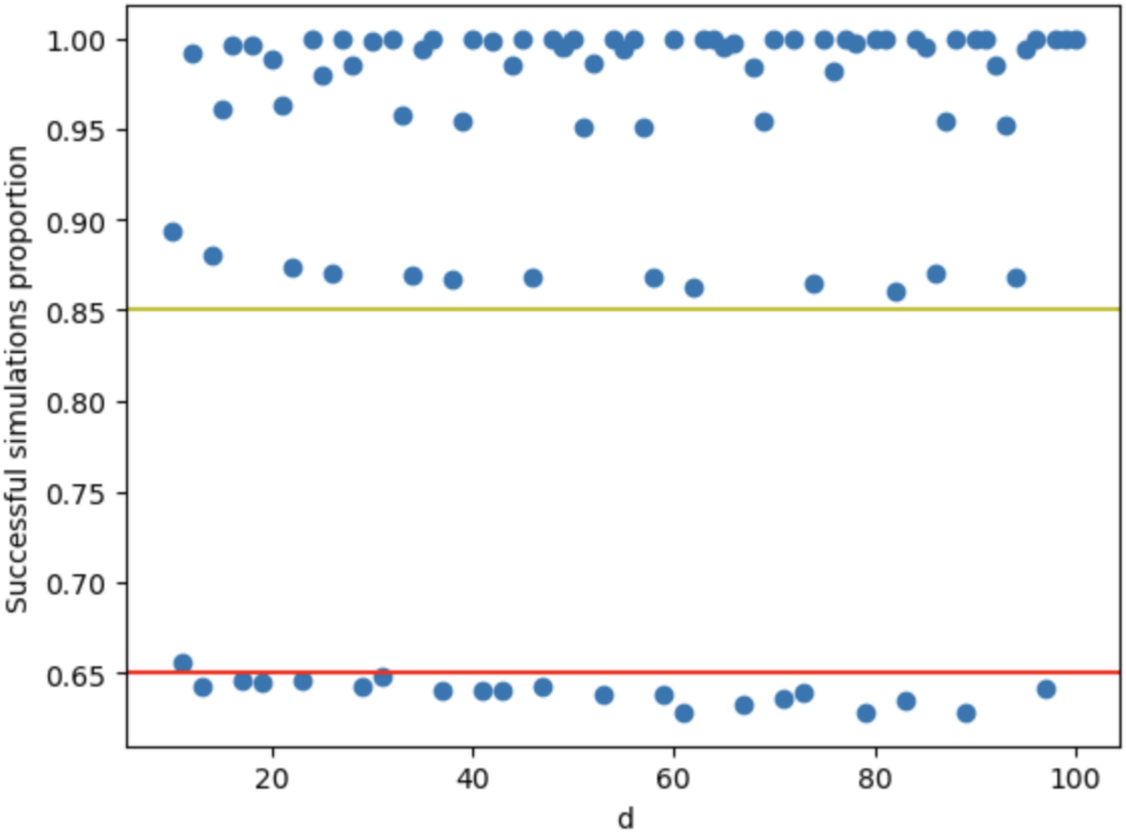
The sequence $\frac{(\ell+1)^2(\ell+2)^2}{4(\ell+1)!^3}$ is decreasing for $\ell \geq 2$ since the ratio of two successive terms is less than 1. Therefore, we can bound it by the $\ell = 3$ term, which is $\frac{25}{3456}$, and obtain

$$\mathrm{P}r(d|\mathrm{X}_d) \geq 1 - 1.35\frac{25}{3456} \approx 0.99 > 0.85.$$

This completes the proof of **Proposition** 5.3, though it should be noted that by restricting d even further, we can significantly raise the lower bound of $\mathrm{P}_d(d|\mathrm{X}_d)$. □

Remark 5.6. From **Proposition** 5.1 and **Proposition** 5.3, we obtain a probabilistic primality test. The test works in the following manner: choose a positive integer m which represents the number of trials to run. Then, for each trial, select d integers from the set $\{1, \ldots, d\}$ uniformly at random. If their product is divisible by d, call the trial a success; otherwise, call it a failure. After all m trials are run, observe the proportion which are a success. If this proportion is above 0.75, then we classify the number as non-prime. Otherwise, we classify it as prime.

Though such a primality test is extremely slow relative to other primality tests, it is a fun consequence of the main theorems. Below, we see the results of running 10,000 simulations for each number from 10 to 100 to test its primality; numbers near the red line (y = 0.65) appear to be prime, and numbers above the yellow line (y = 0.85) appear to be composite.



Note the near-straight line just above 0.85: these are $2 \cdot p$, for primes p.

Acknowledgements

I am incredibly grateful to Professor Lawrence C. Washington from the University of Maryland Mathematics Department and to Ashvin Rajan for all of their advice and assistance in completing this paper. I would also like to thank the referee for carefully reading the paper and providing many helpful comments.

References

[1] "525,600 Minutes Of Math." Edited by Oliver Roeder, *FiveThirtyEight*, FiveThirtyEight, 8 Feb. 2019, https://fivethirtyeight.com/features/525600-minutes-of-math/. Accessed 5 May 2023.

[2] Bartle, Robert G. and Sherbert, Donald R. *Introduction to Real Analysis.* 4th ed., Wiley, Hoboken, New Jersey, 2011.

[3] "Come On Down And Escape The Maze." Edited by Oliver Roeder, *FiveThirtyEight*, FiveThirtyEight, 15 Feb. 2019, https://fivethirtyeight.com/features/come-on-down-and-escape-the-maze/. Accessed 5 May 2023.

[4] Rosen, Kenneth H. *Elementary Number Theory and Its Applications*, Addison-Wesley, Reading, Massachusetts, 1986.

Noah Y. Fine
University of Maryland
noahyfine@gmail.com