Uniform Distribution and Rigidity

Atharv Gudi, Yi Jin, Atharva Naik, Yue Su, Ying Zhao
Sujeet Bhalerao and Joseph Rosenblatt

Introduction

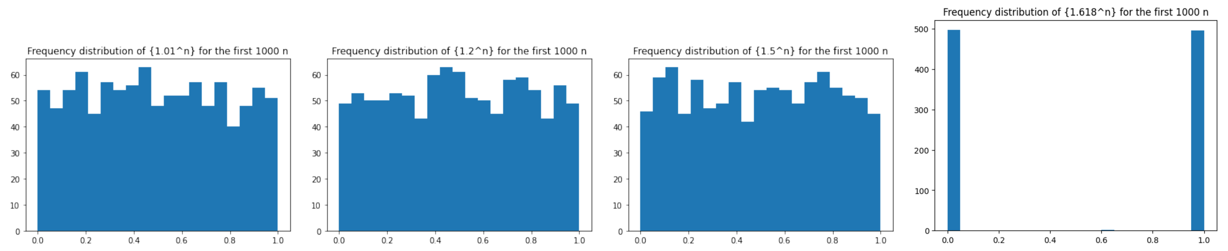
Definition. *A sequence* (x_n) *is uniformly distributed if for all continuous functions* $f : [0,1] \to \mathbb{R}$, *the averages*

$$A_N f = \frac{1}{N}\sum_{n=1}^{N} f(x_n) \to \int_0^1 f(t)\,dt = I.$$

For example, if θ is an irrational real number, the sequence of fractional parts $(\{n\theta\} : n \geq 1)$ is **uniformly distributed**.

Uniform Distribution of Powers of Rationals

To explore the density of $\{\theta^n\}$ in $[0,1)$, we created a program to find the number of iterations n for the fractional part of the powers of any real number θ^n to enter an epsilon ball around a target between $[0,1)$. We graph the fractional part of the first 1000 powers of various $\theta \in \mathbb{R}$. We also graph the powers of the irrational golden ratio, ϕ.



It is known that the sequence (θ^n) is uniformly distributed mod 1 for almost all θ (except for a set of Lebesgue measure 0). We try to quantify the **probability that a sequence will be uniformly distributed** by an analysis of the distributions of its fractional parts.

In particular, for n values of a uniformly distributed sequence and m equally spaced bins of the unit interval, we expect that the proportion of elements in each bin is roughly $\frac{1}{m}$. We measure the deviation from this expectation as a probability. For the same sequences as above, we find that:

with probability 0.9879, (1.05^n) is uniformly distributed
with probability 0.9879, (1.2^n) is uniformly distributed
with probability 0.9879, (1.5^n) is uniformly distributed
with probability 0.5525, (ϕ^n) is uniformly distributed

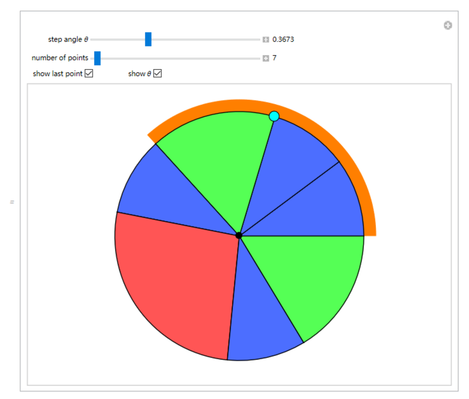
Three Gaps Theorem

Definition. *A simple continued fraction is an expression of the form*

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}$$

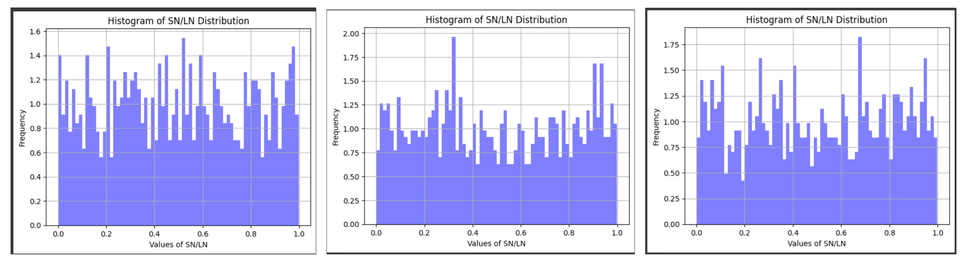
with $a_i \in \mathbb{Z}_{>0}$, *which is abbreviated as* $x = [a_0; a_1, a_2, ...]$. *The* n^{th} *convergent of* $[a_0; a_1, a_2, \cdots]$ *is the finite continued fraction* $[a_0; a_1, a_2, ..., a_n]$, *denoted* $\frac{p_n}{q_n}$.

Three gaps theorem given $x_n = n\theta$, the fractional parts $\{x_n\}, n = 1, \ldots, N$ will be at most three gaps between $[0,1]$.

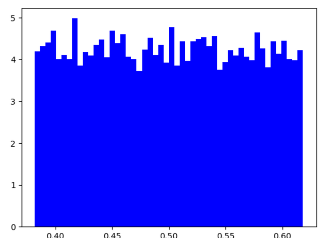


To estimate the ratio between smallest gap and largest gap, we use continued fraction on $x = [x_0 : x_1, x_2, ...]$. The convergences p_n/q_n are obtained from the continued fraction and use the tail ξ_m. It says that as N increases from $q_m + 1$ to q_{m+1}, the ratio s_N/L_N increases monotonically from $\frac{\xi_m}{1+\xi_m}$ to $\frac{1}{1+\xi_{m+1}}$

We suppose the data are in an interval $[0,1]$. You take a large L and divide into equal intervals $J_j = [j/L, (j+1)/L$ with $j = 0, \ldots, L-1$. We count the number D_j of data points d_n in each J_j. We graph it with the base of the bar on the interval J_j and the height equal to D_j.



There are some interesting examples. When we choose $x = \tau = (\sqrt{5}-1)/2$, then we get $\xi_m = \tau$ for all m. The fraction sequence will be $[0, 1, 1, ...]$. If we fixed N between $q_m + 1 \leq N \leq q_{m+1}$, we have $\frac{\tau}{1+\tau} \leq s_N/L_N \leq \frac{1}{1+\tau}$. Now s_N/L_N is not constant, but it is bounded above and below by constants. $\frac{\tau}{1+\tau}$ and $\frac{1}{1+\tau}$. This distribution of the fractional parts is close to a regular partition, which makes the rate of uniform distribution, and the value of the discrepancy D_N in cases like this optimal.



Farey Sequences, Diophantine Approximation, Rigidity Sequences

Farey sequences and Farey approximation

The Farey sequence is defined as F_n, written in increasing order, of all the rational numbers between 0 and 1 that have only the numbers $1, 2, 3, ..., n$ as denominators. We can do the same thing for rational numbers between any two positive numbers. For example we can consider sequences between 1 and 2 where we have $F_1 = \frac{1}{1}, \frac{2}{1}, F_2 = \frac{1}{1}, \frac{3}{2}, \frac{2}{1}$
For the two positive rational numbers $\frac{b}{d}$ and $\frac{a}{c}$ the mediant is defined as $\frac{a+b}{c+d}$ with the nice property that if $0 < \frac{b}{d} < \frac{a}{c}$, then $\frac{b}{d} < \frac{a+b}{c+d} < \frac{a}{c}$.
Farey sequences give an interesting way of approximating rational numbers.
To find a rational approximation to an irrational number using Farey fractions, one can pick the interval between Farey fractions that contains the target number and narrow the interval at each step. If the target number is between $\frac{b}{d}, \frac{a}{c}$, then at the next step you have to decide which of the two intervals below contains the target number: $[\frac{b}{d}, \frac{a+b}{c+d}], [\frac{a+b}{c+d}, \frac{a}{c}]$.

Diophantine approximation and rigidity

Diophantine approximation concerns how well we can approximate a given irrational number x by rational numbers. In fact, there is a constant C such that for any irrational x, there are infinitely many integer pairs (p,q) such that $|x - \frac{p}{q}| \leq \frac{C}{q^2}$.
Hurwitz's theorem says $C = \frac{1}{\sqrt{5}}$ is the smallest value of C. If the golden ratio and all the numbers "equivalent" to it are deleted, the smallest value of C can be $\frac{1}{\sqrt{8}}$.
Diophantine approximation of x gives us rigidity sequences $(q_n : n \geq 1)$: sequences such that the fractional parts $\{q_n x\}$ converge to 0 modulo one. That is, there are fractions f_n such that $q_n x = p_n + f_n$, and f_n goes to zero.

Farey Fractions and the Riemann Hypothesis

Uniform distribution in Farey sequence as $n \to \infty$

The number of elements of $\mathcal{F}_n$ is $N_n = \sum_{k=1}^{n} \varphi(q) = \frac{3n^2}{\pi^2} + O(n \log n)$. For each interval I $\subseteq (0,1]$, the size of $\mathcal{F}_n(I)$ is $N_n = \frac{3|I|n^2}{\pi^2}(1 + \frac{\log n}{n}) + O_\epsilon(n^{-1+\epsilon})$. As $n \to \infty$, we have $\frac{N_n(I)}{N_n} = |I|(1 + O(\frac{\log n}{n})) + O_\epsilon(n^{-1+\epsilon}) \to 0$. This shows that Farey fractions are uniformly distributed as $n \to \infty$. It is worth observing that The sum of a function at the points in the Farey sequence is given by $\sum_{\alpha \in \mathcal{F}_n(I)}^{L(n)} f(\alpha) = \sum_{k=1}^{\infty} \sum_{j=1}^{k} f(\frac{j}{k}) f(\frac{n}{k})$.

Estimates equivalent to the Riemann Hypothesis, a couple using the Farey Sequence

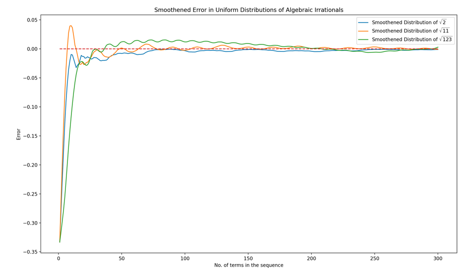
There are many equivalent forms of the Riemann Hypothesis. For example, consider the growth of the function $M(x) := \sum_{j \leq x} \mu(j)$, where $\mu(j)$ is a Möbius function. According to Littlewood(1912), $M(n) = O_\epsilon(n^{\frac{1}{2}+\epsilon}) \forall \epsilon > 0$ is equivalent to Riemann Hypothesis.

Two estimates equivalent to the Riemann Hypothesis, using the Farey sequence: $\sum_{j=1}^{N_n} (\alpha_j - \frac{j}{N_n})^2 = O_\epsilon(n^{-1+\epsilon}), \forall \epsilon > 0$ given by Franel in 1924 and $\sum_{j=1}^{N_n} |\alpha_j - \frac{j}{N_n}| = O_\epsilon(n^{\frac{1}{2}+\epsilon}), \forall \epsilon > 0$ given by Landau given in 1924.

Numerical Integration

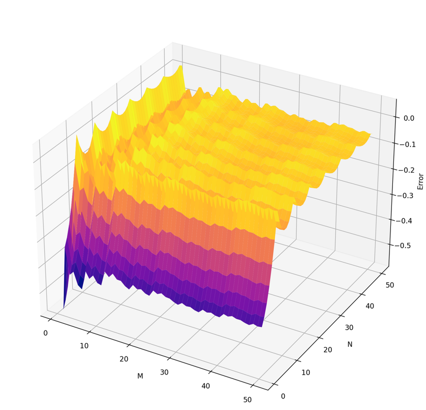
We have discovered that the convergence of averages in the definition of uniform distribution always includes **oscillation**: for any non-constant f, we have infinitely often $A_N f < A_{N+1} f$ and also infinitely often $A_{N+1} f < A_N f$. What about **oscillation around the mean**: for infinitely many N, $A_N f < I$, and also for infinitely many N, $A_N f > I$? We have found that this is not always the case, but it is **typical**.

The same oscillation behavior occurs in ergodic theory for the time averages $A_N^T f(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^n x)$ of an ergodic transformation T. Discovering oscillation for averaging in uniform distribution and in ergodic theory has led us to **expand** our study of oscillation to other averaging processes and to multi-variable cases of uniform distribution averaging and for averaging in ergodic theory.



Oscillation of error in 1D distributing $\{n\theta\}, \theta \in \mathbb{Q}$

The plots depict the error between an integral $\int_0^1 f$ and its approximation $A_N f$ using some finite terms of the series against the number of terms (N) taken in both 1 and 2 dimensions. The oscillations about the zero-error mark can be observed.



Oscillation of error in 2D of $(\{M\alpha\}, \{N\beta\}), \alpha, \beta \in \mathbb{Q}$

Such uniform distributions with irrationals are considered low discrepancy **Sobol sequences**, a class of sequences whose terms are spaced out specifically to be as equidistributed as possible to reduce the discrepancy.

The usage of uniform distributions as Sobol sequences is computationally inexpensive in the context of quantitative finance. Although it is copious in lower dimensions (where Riemann integration is the easiest and most accurate method of integrating), not only does it have a smaller Big-O, but it also results in a smaller error when taking a slightly larger number of terms.

References

[1] Boca, F. P., Zaharescu, A. (2006). Farey fractions and two-dimensional tori. In Noncommutative geometry and number theory (pp. 57-77). Aspects Math., E37, Vieweg.
[2] Niven, I. (2013). Diophantine Approximations. Dover Books on Mathematics. Courier Corporation.
[3] Niven, I., Zuckerman, H. S., and Montgomery, H. L. (1991). An Introduction to the Theory of Numbers (5th ed.). John Wiley and Sons.

Support for this project was provided by the Illinois Geometry Lab and the Department of Mathematics at the University of Illinois at Urbana-Champaign.