Correlations of Fractional Parts of Dilated Harmonic Sequences

Jeff Lagarias
University of Michigan

David Montague
Stanford University

Joint Math Meetings-San Diego
Special Session on Arithmetic Statistics,
January 10, 2013

Credits

- Work of Jeff Lagarias was partially supported by NSF grants DMS-0801029 and DMS-1101373.

- Some work of David Montague was done as part of an REU at the University of Michigan in 2009, supported by NSF.

Table of Contents

Dilated Harmonic Sequences-1

The harmonic sequence is $1, \frac{1}{2}, \frac{1}{3}...,$

The dilated harmonic sequence with integer dilation factor n is

$$n, \frac{n}{2}, \frac{n}{3}, ...$$

Its fractional parts are

$$x_k(n) := \{\frac{n}{k}\},$$

for $k = 1, 2, 3, ...$

More generally, one could allow a real dilation factor $y > 0$. Then there is a decoupling:

$$x_k(y) := \{\frac{y}{n}\} = x_k([y]) + \frac{y - [y]}{n}.$$

Dilated Harmonic Sequences-2

Problem 1. What is the distribution of the numbers

$$x_k(n) = \{\frac{n}{k}\},$$

for $1 \le k \le f(n)$, as $n \to \infty$?

Answer: It will depend on $f(n)$.

When $f(n)$ is small compared to n, we might expect that the successive fractional parts might be "random".

For large $f(n)$ we will get a lot of fractional parts very near 0. The distribution must approach a delta function supported at $x = 0$.

Dilated Harmonic Sequences-3

Consider the special case $f(n) = n$.

This was investigated by Dirichlet, while studying the divisor problem. He showed:

Theorem. (Dirichlet 1849)

$$\sum_{k=1}^{n} \{\frac{n}{k}\} = (1-\gamma)n + O(\sqrt{n}),$$

where $\gamma = 0.57721\cdots$ is Euler's constant.

The expected value of the fractional parts is $1 - \gamma = 0.42278\cdots$, and the fractional parts cannot be uniformly distributed in this range, since the mean is not $1/2$.

Dilated Harmonic Sequences-4

Dirichlet guessed from this that there are more fractional parts in $[0, 1/2]$ than in $[1/2, 1]$. He then determined that the number of large fractional parts is asymptotic to $(\log 4 - 1)n$ and the smaller ones to $(2 - \log 4)n$. Here

$$2 - \log 4 = 0.613705.$$

Dilated Harmonic Sequences-5

Consider next the special case $f(n) = \sqrt{n}$.

Dirichlet also encountered this case in connection with the divisor problem, which is that of estimating

$$\sum_{1 \le n \le x} d(n) = x \log x + (2\gamma - 1)x + \Delta(n).$$

His hyperbola method showed that

$$\Delta(n) = -2\Big(\sum_{k=1}^{\sqrt{n}} \{\frac{x}{k}\} - \frac{1}{2} \Big).$$

This immediately gives

$$|\Delta(n)| = O(\sqrt{n}),$$

but one should expect further cancellation in the sum.

Dilated Harmonic Sequences-6

Suppose that the x_k were independent, identically distributed. random variables.

Theorem. (Law of iterated logarithm) For a sequence of independent, identically distributed (iid) uniformly distributed random variables on $[0, 1]$. Set

$$Y_N := \sum_{i=1}^{N} (x_k - \frac{1}{2})$$

Drawing $Y_1, Y_2, Y_3, \ldots$with new x_k each time, then with probability one

$$\limsup_{N \to \infty} \frac{Y_N}{\sqrt{N} \log \log N} = 1/\sqrt{3}.$$

Dilated Harmonic Sequences-7

- This random model predicts:

$$|\Delta(n)| \stackrel{?}{=} O(n^{1/4} \log\log n)$$

.

- But are the $x_k(n) = \{n/k\}$ approximately i.i.d. uniform on $[0, 1]$, as k varies, for $f(n) = \sqrt{n}$?

Dilated Harmonic Sequences-8

Continuing with $\Delta(n)$, estimating its size is called the Dirichlet Divisor problem.

van der Corput(1923) improved Dirichlet's bound to $O(n^{1/3})$, and the current record is $O(n^{0.31490})$ of Huxley (2003). In the other direction:

Theorem. (Hardy and Landau 1916)

$$|\Delta(n)| = \Omega(n^{1/4}).$$

This bound improve by Soundararajan (2003) to:

$$\Omega(n^{1/4}(\log n)^{1/4}(\log\log n)^{b}(\log\log\log n)^{-5/8})$$

where $b = 3/4(2^{4/3} - 1)$.

Dilated Harmonic Sequences-9

To what extent do the $x_k(n)$ behave like independent, uniformly distributed random variables for $f(n) = \sqrt{n}$?

Answer. The $x_k(n)$ are individually uniformly distributed as $n \to \infty$, with $f(n) = \sqrt{n}$.

Follows from: Isbell-Schanuel (1976), Saffari-Vaughan (1977)

Answer: We will show for $f(n) = \sqrt{n}$ that the joint random variables $(x_k(n), x_{k+1}(n))$ are correlated. They go to a limiting distribution which is a "continuous" distribution on the square $[0,1]^2$, but it is not the uniform distribution on the square.

Disclaimer: We get no results about $\Delta(n)$.

Dilated Harmonic Sequences-10

The distribution of $\{\frac{n}{k}\}$ was investigated in detail by B. Saffari -R. Vaughan (Ann. Inst. Fourier 1976/1977) for a range of $f(n)$. They are obtained general, flexible bounds, as well as rates of convergence. Their results imply uniform distribution at this scale, but their main result does not apply for slow growing $f(n) = O(n^{1/3} \log n)$.

They get an explicit formula for the mass of the distribution on the interval $[0, \alpha)$, for $f(x) = cx$, with $0 < c \leq 1$. which depends on c. For $c = 1$, it is the continuous distribution:

$$Prob[\{n/k\} \in [0, \alpha]] = \sum_{k=1}^{\infty} \frac{\alpha}{k(k+\alpha)}.$$

2. Results-1

For very small $f(n)$ there is no limiting distribution.
(*) If $f(n) < C \log n$ then there is no limit distribution of fractional parts as $n \to \infty$.

Theorem 1. **(Uniform Distribution)**
If $f(n)$ is increasing and if

$$f(n) >> \exp\Big((\log 2 + \epsilon)(\log n / \log\log n)\Big)$$

for some $\epsilon > 0$, and as $n \to \infty$

$$f(n) = o(n),$$

then the limit distribution is the uniform distribution.

This result is in in Schanuel-Isbell (1976) for the range down to $f(n) = n^{\epsilon}$.

3. Results-2

Theorem. *For fixed* $0 < c < \infty$

$$f(n) \sim cn$$

as $n \to \infty$, *then the distribution* $x_k(n)$ *approaches a limiting distribution which appears to be continuous on* $[0, 1]$. *(It may vanish on part of the interval.) The Fourier series of this distribution can be given explicitly.*

Note. If $f(n)/n \to \infty$ then the distribution of $x_k(n)$ has limiting distribution a delta function at $x = 0$.

The case $0 < c \leq 1$ is done in Saffari-Vaughan(1977).

3. Results-3

Our main result extends the result above to multiple correlations. Throughout we assume, for some $\epsilon > 0$,

$$f(n) >> \exp\Big((\log 2 + \epsilon)(\log n / \log\log n)\Big).$$

Theorem **(Joint Pair Distribution)** *One has a trichotomy.*
(1) The joint distribution of pairs $(x_k(n), x_{k+1}(n))$ will be be uniform up to a scale $o(\sqrt{n})$.

(2) The joint distribution will be non-uniform with a distribution "continuous" on the square for $f(n) = c\sqrt{n}$, $0 < c < \infty$;

(3) The joint distribution $(x_k(n), x_{k+1}(n))$ will be totally correlated (i.e. supported on the diagonal of the square) whenever $f(n)/\sqrt{n} \to \infty$ as $n \to \infty$.

3, Results-4

For multiple joint distribution of $(x_k(n), ..., x_{k+j}(n))$ there is also a trichotomy. The threshold for change of behavior is proportional to the scale $n^{1/(j+1)}$.

In the totally-correlated regime $x_{k+j}(n)$ will in the limit $n \to \infty$ be totally determined by the values $(x_k(n), x_{k+1}(n), ..., x_{k+j-1}(n))$ when $f(n)/n^{1/j} \to \infty$.

In the middle range, where the distribution changes, the Fourier coefficients are obtained for the distribution on the unit $(j + 1)$-cube.

3, Methods

The proofs use the van der Corput method.

A key point is to go after the Fourier coefficients of the distribution, rather than after a direct formula for the cumulative density function, as in Saffari-Vaughan.

Work to do: We have not unwound what the Fourier coefficients say about the density function of the distribution.

Thank you for your attention!