Central Limit Theorems for Random Multiplicative Functions

K. Soundararajan (joint work with Max Xu)

Random multiplicative function:

<u>Steinhaus</u> (Complex valued)

$$f(n) = \prod f(p_i)^{\alpha_i} \qquad n = \prod p_i^{\alpha_i}.$$

$f(p)$ chosen independently & uniformly on unit circle $\mathbb{T} = \{|z| = 1\}$.

<u>Rademacher</u> (real valued, $\pm 1, 0$)

$f(n) = 0$ if n is not squarefree

$f(n) = \prod f(p_i)$ if $n = \prod p_i$ is squarefree

$$f(p) = \begin{cases} 1 & \text{prob } 1/2 \\ -1 & \text{prob } 1/2 \end{cases}$$

independently for different p.

Model for arithmetic functions:

$\lambda(n)$, $\mu(n)$, $\chi(n)$, n^{it}.

Fundamental question: Distribution of $\sum_{n \le x} f(n)$

for a random multiplicative f.

Recent progress on such questions by HARPER.

Harper's result proving HELSON'S Conjecture:

Steinhaus case:

$$\mathbb{E}\left(\left| \sum_{n \le x} f(n) \right|^2 \right) = \sum_{m,n \le x} \mathbb{E}\left(f(m)\overline{f(n)}\right)$$

$$= x$$

But

$$\mathbb{E}\left(\left| \sum_{n \le x} f(n) \right| \right) \asymp \frac{\sqrt{x}}{(\log \log x)^{1/4}}.$$

Harper: Towards a law of the iterated logarithm.

f = random Steinhaus mult. function.

Almost surely: $\exists$ arbitrarily large x

with

$$\left| \sum_{n \le x} f(n) \right| \gtrsim \sqrt{x} \, (\log \log x)^{1/4 - \varepsilon}.$$

(Law of iterated log for sums of independent variables: $\sqrt{x} \, (\log \log x)^{1/2}$).

Upper bound: (Lau, Tenenbaum & Wu)

almost surely:

$$\left| \sum_{n \le x} f(n) \right| = O\left(\sqrt{x} \, (\log \log x)^{2+\varepsilon}\right).$$

Answered a question of Halasz:

Related Results:

Model problem: $X(k)$ = independent standard complex gaussians mean 0, var 1.

(Real & Imaginary parts are indep. real gaussians with mean 0 & variance $1/2$).

$$\sum_{n=0}^{\infty} A(n) z^n = \exp\left(\sum_{k=1}^{\infty} \frac{X(k)}{\sqrt{k}} z^k \right)$$

S. & Zaman: $\mathbb{E}(|A(N)|^2) = 1$

but $\mathbb{E}(|A(N)|) \asymp \frac{1}{(\log N)^{1/4}}$.

Gerspach: Almost surely, $\exists$ arbitrarily large N with $|A(N)| \geq (\log N)^{1/4 - \varepsilon}$.

Connections with the Riemann zeta function.

Fyodorov - Hiary - Keating Conjecture: For almost all $t \in [T, 2T]$

$$\max_{|u-t|\leq 1} |\zeta(\tfrac{1}{2}+iu)| \asymp \frac{(\log T)}{(\log\log T)^{3/4}}.$$

Progress by Harper; Arguin, Bourgade & Radziwill

Harper:

$$\frac{1}{T}\left(\int_T^{2T} \left(\frac{1}{\log T}\int_t^{t+1} |\zeta(\tfrac{1}{2}+iu)|^2\,du\right)^{\frac{1}{2}} dt\right.$$

$$\ll \frac{1}{(\log\log T)^{1/4}}.$$

Focus of this talk: Central limit theorems

for $\sum_{\substack{n \le x \\ n \in A}} f(n)$ or $\sum_{n \le x} a_n f(n)$.

Examples:

$\sum_{p \le x} f(p)$ — sum of independent random variables.

Hough: For any fixed k,

$\sum_{\substack{n \le x \\ \omega(n) = k}} f(n)$ is Gaussian.

Harper: $k = o(\log \log x)$ is okay.

But not Gaussian when k is of size $\log \log x$.

Chatterjee & S. (Rademacher case).

Gaussian in short intervals $[x, x+y]$:

Provided $\begin{cases} y = o(x/\log x) \\ y \geq x^{1/5+\varepsilon} \end{cases}$ (so enough squarefrees).

Harper's work: NOT Gaussian for all $n \leq x$

In fact not Gaussian if $y \geq x/\exp(\sqrt{\log\log x})$.

Question: (open) Precisely in what short intervals does there exist a CLT?

Other interesting subsets on which we have CLT?

Theorem: (S. & Xu). $A \subseteq [1, N]$ with

$$|A| \geq N / \exp(\sqrt{\log N \log\log N}).$$

Suppose there is a subset S of A with

$$|S| = (1 + o(1))|A|$$

and such that

$$\#\{ s_1 s_2 = s_3 s_4 : s_i \in S, \text{ non-trivial solution} \} = o(|S|^2).$$

Then as f ranges over Steinhaus random multiplicative functions

$$\frac{1}{\sqrt{|A|}} \sum_{n \in A} f(n)$$

is a standard complex Gaussian with mean 0 and variance 1.

Applications:

(1) If $y \le \frac{x}{(\log x)^{\alpha-\varepsilon}}$, $\alpha = 2\log 2 - 1$

then $\frac{1}{\sqrt{y}} \sum_{x \le n \le x+y} f(n)$ is Gaussian.

Key idea: restrict to $S \subseteq [x, x+y]$ of typical integers with $(1+o(1)) \log\log x$ prime factors.

(2) Shifted primes: $A = \{p+2 \le N\}$.

$\frac{1}{\sqrt{|A|}} \sum_{n \in A} f(n)$ is Gaussian.

(3) Sums of Two squares in short intervals.

$A = \{n \in [x, x+y] : n = \square + \square\}$.

Gaussian if $y = o(x)$.

(4) There exists a subset A of $[1, N]$ with

$$|A| \geq \frac{N}{(\log N)^{\theta} (\log\log N)^{6}}$$

with $\theta = 1 - \frac{(1+\log\log 4)}{\log 4} = 0.043\cdots$

such that $\frac{1}{\sqrt{|A|}} \sum_{n \in A} f(n)$ is Gaussian

Ford's work on such a set A where the product set $A \cdot A$ has size $\sim |A|^2/2$.

Related to the multiplication table problem.

Is there a larger subset of $[1, N]$ on which CLT holds.

⑤ More general sums $\sum_{n \le x} a_n f(n)$.

Complex Gaussian with mean 0 and variance

$$V = \sum_{n \le x} |a_n|^2$$

Provided:

$$\sum_{\substack{n_1, n_2, n_3, n_4 \le x \\ n_1 n_2 = n_3 n_4 \\ P(n_1) = P(n_3) \,;\, P(n_2) = P(n_4) \\ n_1 \ne n_3,\ n_2 \ne n_4}} a_{n_1} a_{n_2} \overline{a_{n_3}}\, \overline{a_{n_4}} = o(V^2)$$

+ similar condition with $P(n_1) = P(n_2) = P(n_3) = P(n_4)$ terms.

Application:

If θ is a real number with

$$\|q\theta\| \gg \exp(-q^{1/3})$$

then $\sum_{n \leq x} f(n) e(n\theta)$ is Complex Gaussian.

Includes all algebraic irrationals θ, π etc.

Benatar, Nishry & Rodgers had shown such a result for almost all θ.

Idea of Proof.

Moments blow up.

Based on martingale central limit theorem.

$X_1, \dots, X_N$ real valued random variables.

Martingale difference sequence:

$$\mathbb{E}[X_1] = 0, \quad \mathbb{E}[X_{n+1} \mid X_1, \dots, X_n] = 0.$$

Compute Fourier transform $\mathbb{E}[e^{itS_N}]$

$(S_N = X_1 + \dots + X_n)$.

$$\mathbb{E}[e^{itS_N}] = e^{-t^2/2} \qquad \text{(Gaussian)}$$

$$+ O\left(e^{t^2} \mathbb{E}[\max |X_n|]\right)$$

$$+ O\left(e^{t^2} \mathbb{E}\left[\min\left(1, \left|\sum_{n=1}^{N} X_n^2 - 1\right|\right)\right]\right)$$

Weaker versions (easier to compute)

$$\mathbb{E}\left[\max |X_n|\right] \leq \left(\sum_{n=1}^{N} \mathbb{E}[X_n^4]\right)^{1/4}$$

$$\mathbb{E}\left[\max\left(1, \left|\sum_{n=1}^{N} X_n^2 - 1\right|\right)\right] \leq \left\{\mathbb{E}\left[\left(\sum_{n=1}^{N} X_n^2 - 1\right)^2\right]\right\}^{1/2}.$$

(Basically fourth moment conditions).

<u>Idea of proof:</u>

$$e^{it(X_1 + \cdots + X_N)} \approx \prod_{n=1}^{N}\left((1+itX_n)\, e^{-\frac{t^2 X_n^2}{2}}\right)$$

$$\approx \left(\prod_{n=1}^{N} (1+itX_n)\right) e^{-t^2/2}$$

$$\mathbb{E}\left[\prod_{n=1}^{N} (1+itX_n)\right] = 1$$ by martingale property.

Application to multiplicative functions. (pioneered by Harper)

$$\sum_{n \in A} f(n) = \sum_{p} \Big(\sum_{\substack{n \in A \\ P(n) = p}} f(n) \Big)$$

$$X_p = \frac{1}{\sqrt{|A|}} \sum_{\substack{n \in A \\ P(n) = p}} f(n).$$

Martingale difference sequence.

Need to understand

$$\sum_{p} \mathbb{E}\Big[\frac{1}{|A|^2} \Big| \sum_{\substack{n \in A \\ P(n) = p}} f(n) \Big|^4 \Big]$$

and

$$\mathbb{E}\Big[\Big(\frac{1}{|A|} \sum_{p} \Big| \sum_{\substack{n \in A \\ P(n) = p}} f(n) \Big|^2 - 1 \Big)^2 \Big].$$