A Central Limit Theorem for Random Euler-Product Surrogates of $\log|\zeta(1/2+it)|$

Iz Tecúm

Columbia University

Spring 2026

What this talk does

We build a clean probabilistic surrogate for the Euler product of $\zeta(1/2 + it)$ and prove a one-point CLT

$$\frac{X_y(t)}{\sigma_y} \Rightarrow N(0,1) \qquad (y \to \infty),$$

where $X_y(t) = \operatorname{Re} \log Z_y(t)$ is the real part of a *random* truncated Euler product, and

$$\sigma_y^2 = \frac{1}{2} \sum_{p \le y} \frac{1}{p} \sim \frac{1}{2} \log \log y.$$

What this talk does

We build a clean probabilistic surrogate for the Euler product of $\zeta(1/2+it)$ and prove a one-point CLT

$$\frac{X_y(t)}{\sigma_y} \Rightarrow N(0,1) \qquad (y \to \infty),$$

where $X_y(t) = \operatorname{Re}\log Z_y(t)$ is the real part of a *random* truncated Euler product, and

$$\sigma_y^2 = \frac{1}{2}\sum_{p\le y}\frac{1}{p} \sim \frac{1}{2}\log\log y.$$

The proof isolates two mechanisms.

- The **main layer** is a sum of independent bounded increments at scale $p^{-1/2}$, so Lindeberg–Feller applies.
- The **higher Euler-power layers** are square-summable in p and stay $O_{L^2}(1)$, hence negligible after dividing by $\sigma_y \to \infty$.

Statment

Theorem 1 (CLT for the random Euler-product log)

Let $(\theta_p)_{p\le y}$ be i.i.d. uniform on $[0,2\pi)$ and define

$$Z_y(t)=\prod_{p\le y}\left(1-e^{i\theta_p}p^{-1/2-it}\right)^{-1},\qquad X_y(t)=\mathrm{Re}\log Z_y(t).$$

Let

$$\sigma_y^2=\frac{1}{2}\sum_{p\le y}\frac{1}{p}.$$

Then as $y\to\infty$,

$$\frac{X_y(t)}{\sigma_y}\Rightarrow N(0,1),$$

and the limiting law does not depend on t.

Statment

Theorem 1 (CLT for the random Euler-product log)

Let $(\theta_p)_{p\le y}$ *be i.i.d. uniform on* $[0,2\pi)$ *and define*

$$Z_y(t)=\prod_{p\le y}\left(1-e^{i\theta_p}p^{-1/2-it}\right)^{-1},\qquad X_y(t)=\operatorname{Re}\log Z_y(t).$$

Let

$$\sigma_y^2=\frac{1}{2}\sum_{p\le y}\frac{1}{p}.$$

Then as $y\to\infty$*,*

$$\frac{X_y(t)}{\sigma_y}\Rightarrow N(0,1),$$

and the limiting law does not depend on t*.*

Talk goal: derive this from a precise expansion of $\log Z_y(t)$, a $k=1$ vs. $k\ge 2$ split, and standard CLT tools.

Roadmap of the proof

1. Define the model and justify the logarithmic expansion.
2. Prove a symmetry: t is inessential in distribution.
3. Expand $X_y(t)$ into a prime-by-prime trigonometric series.
4. Split $X_y(t) = S_y(t) + R_y(t)$ (main layer + remainder).
5. Compute $\mathrm{Var}(S_y(t)) = \sigma_y^2$ and show $\sigma_y \to \infty$.
6. Prove $S_y(t)/\sigma_y \Rightarrow N(0,1)$ by Lindeberg–Feller.
7. Prove $R_y(t) = O_{L^2}(1)$, hence $R_y(t)/\sigma_y \to 0$ in probability.
8. Apply Slutsky: $X_y(t)/\sigma_y \Rightarrow N(0,1)$.

Outline

Random Euler product model

Fix $y \geq 2$. Sample i.i.d. phases $\theta_p \sim \mathrm{Unif}[0, 2\pi)$ for primes $p \leq y$ and define

$$\xi_p := e^{i\theta_p} \qquad (p \leq y).$$

Random Euler product model

Fix $y \geq 2$. Sample i.i.d. phases $\theta_p \sim \mathrm{Unif}[0, 2\pi)$ for primes $p \leq y$ and define

$$\xi_p := e^{i\theta_p} \qquad (p \leq y).$$

For $t \in \mathbb{R}$, define

$$Z_y(t) = \prod_{p \leq y} \left(1 - \xi_p \, p^{-1/2-it}\right)^{-1}.$$

Random Euler product model

Fix $y \geq 2$. Sample i.i.d. phases $\theta_p \sim \mathrm{Unif}[0, 2\pi)$ for primes $p \leq y$ and define

$$\xi_p := e^{i\theta_p} \qquad (p \leq y).$$

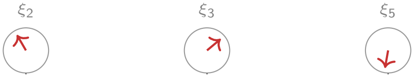
For $t \in \mathbb{R}$, define

$$Z_y(t) = \prod_{p \leq y} \left(1 - \xi_p\, p^{-1/2-it}\right)^{-1}.$$

Interpretation:

- Keep deterministic oscillation $p^{-it} = e^{-it \log p}$ (number-theory "frequency").
- Replace arithmetic dependence by independent random rotations ξ_p.

Picture: product $\to$ log $\to$ prime sum



$$Z_y(t) = \underbrace{\boxed{\left(1-\frac{\xi_2}{2^s}\right)^{-1}}\;\boxed{\left(1-\frac{\xi_3}{3^s}\right)^{-1}}\;\boxed{\left(1-\frac{\xi_5}{5^s}\right)^{-1}}\;\cdots}$$

Apply log then $\Re$

$$\boxed{X_y(t) = \Re \sum_{p\le y}\sum_{k\ge 1} \frac{\xi_p^k\, p^{-ikt}}{k\, p^{k/2}}}$$

This "log Euler product $\Rightarrow$ prime sum" step is the entire engine of the proof.

Defining the logarithm cleanly

We study

$$X_y(t) := \operatorname{Re} \log Z_y(t).$$

Key point: each Euler factor is strictly inside the unit disk since

$$\left| \xi_p \, p^{-1/2-it} \right| = p^{-1/2} < 1.$$

Defining the logarithm cleanly

We study

$$X_y(t) := \operatorname{Re}\log Z_y(t).$$

Key point: each Euler factor is strictly inside the unit disk since

$$\left|\xi_p\, p^{-1/2-it}\right| = p^{-1/2} < 1.$$

Therefore we can use the absolutely convergent series identity

$$-\log(1-z) = \sum_{k\geq 1} \frac{z^k}{k} \qquad (|z|<1),$$

with $z = \xi_p\, p^{-1/2-it}$ *for each prime.*

Lemma: absolute convergence of the logarithmic expansion

Lemma 2 (Log expansion for $\log Z_y(t)$)

For fixed y and t,

$$\log Z_y(t) = \sum_{p \le y} \sum_{k \ge 1} \frac{e^{ik(\theta_p - t \log p)}}{k\, p^{k/2}},$$

with absolute convergence (so the order of summation is justified).

Lemma: absolute convergence of the logarithmic expansion

Lemma 2 (Log expansion for $\log Z_y(t)$)

For fixed y and t,

$$\log Z_y(t) = \sum_{p\le y}\sum_{k\ge 1} \frac{e^{ik(\theta_p - t\log p)}}{k\,p^{k/2}},$$

with absolute convergence (so the order of summation is justified).

Why absolute convergence is easy here: for each fixed p,

$$\sum_{k\ge 1}\left|\frac{(\xi_p\,p^{-1/2-it})^k}{k}\right| \le \sum_{k\ge 1} p^{-k/2} = \frac{p^{-1/2}}{1-p^{-1/2}} < \infty,$$

and the prime product is finite.

Cosine expansion for $X_y(t)$

Taking real parts term-by-term yields the explicit trigonometric series

$$X_y(t) = \sum_{p \le y} \sum_{k \ge 1} \frac{\cos(k(\theta_p - t \log p))}{k\, p^{k/2}}.$$

Cosine expansion for $X_y(t)$

Taking real parts term-by-term yields the explicit trigonometric series

$$X_y(t) = \sum_{p \le y} \sum_{k \ge 1} \frac{\cos(k(\theta_p - t \log p))}{k\, p^{k/2}}.$$

This already shows the two scales:

- $k = 1$ terms contribute $\asymp p^{-1/2}$ increments.
- $k \ge 2$ terms have extra decay $p^{-k/2}$ and will be square-summable in p.

Outline

Why t is inessential in distribution

Define shifted phases

$$U_p(t) := \theta_p - t \log p \pmod{2\pi}.$$

Why t is inessential in distribution

Define shifted phases

$$U_p(t) := \theta_p - t \log p \pmod{2\pi}.$$

Because θ_p is uniform on the circle, subtracting a deterministic angle keeps it uniform.

Lemma 3 (Uniform shift symmetry)

If $\Theta \sim \mathrm{Unif}[0, 2\pi)$ and $\alpha \in \mathbb{R}$, then $\Theta - \alpha \pmod{2\pi}$ is also uniform.

Why t is inessential in distribution

Define shifted phases

$$U_p(t) := \theta_p - t \log p \pmod{2\pi}.$$

Because θ_p is uniform on the circle, subtracting a deterministic angle keeps it uniform.

Lemma 3 (Uniform shift symmetry)

If $\Theta \sim \mathrm{Unif}[0, 2\pi)$ and $\alpha \in \mathbb{R}$, then $\Theta - \alpha \pmod{2\pi}$ is also uniform.

Thus $(U_p(t))_{p \le y}$ are i.i.d. uniform for every fixed t.

Proposition: t-invariance of the law

Proposition 4 (The law does not depend on t)

For each fixed y, $X_y(t)$ has the same distribution for all $t \in \mathbb{R}$.

Proposition: t-invariance of the law

Proposition 4 (The law does not depend on t)

For each fixed y, $X_y(t)$ has the same distribution for all $t \in \mathbb{R}$.

Proof idea:

- Write $X_y(t)$ as a measurable function of the family $(U_p(t))_{p \leq y}$ through the cosine expansion.
- For each t, $(U_p(t))_{p \leq y}$ is i.i.d. uniform with the same joint law as $(\theta_p)_{p \leq y}$.
- Therefore the induced law of $X_y(t)$ is independent of t.

Outline

The structural split $X_y = S_y + R_y$

Define the main layer (the $k = 1$ terms)

$$S_y(t) := \sum_{p \le y} \frac{\cos(\theta_p - t \log p)}{\sqrt{p}},$$

and the higher-power remainder

$$R_y(t) := \sum_{p \le y} \sum_{k \ge 2} \frac{\cos(k(\theta_p - t \log p))}{k\, p^{k/2}}.$$

Then

$$X_y(t) = S_y(t) + R_y(t).$$

The structural split $X_y = S_y + R_y$

Define the main layer (the $k = 1$ terms)

$$S_y(t) := \sum_{p \le y} \frac{\cos(\theta_p - t \log p)}{\sqrt{p}},$$

and the higher-power remainder

$$R_y(t) := \sum_{p \le y} \sum_{k \ge 2} \frac{\cos\big(k(\theta_p - t \log p)\big)}{k\, p^{k/2}}.$$

Then

$$X_y(t) = S_y(t) + R_y(t).$$

This split is designed so that:

- $S_y(t)$ has slowly diverging variance $\sim \frac{1}{2} \log \log y$.
- $R_y(t)$ has bounded variance (square-summable in p).

Geometric intuition: a prime-indexed random walk

Define the complex random walk

$$V_y(t) := \sum_{p \le y} \frac{e^{i(\theta_p - t \log p)}}{\sqrt{p}}, \qquad S_y(t) = \mathrm{Re}\ V_y(t).$$

Geometric intuition: a prime-indexed random walk

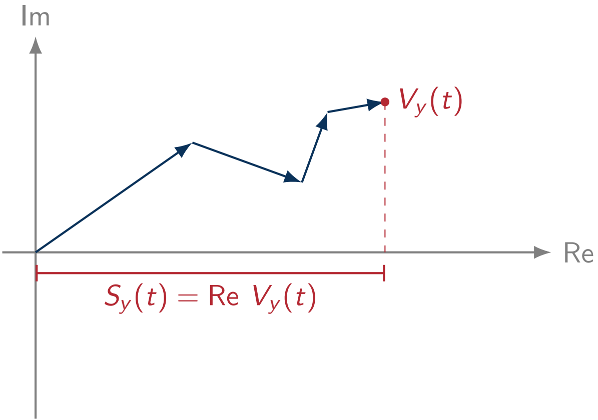
Define the complex random walk

$$V_y(t) := \sum_{p \le y} \frac{e^{i(\theta_p - t \log p)}}{\sqrt{p}}, \qquad S_y(t) = \mathrm{Re}\ V_y(t).$$

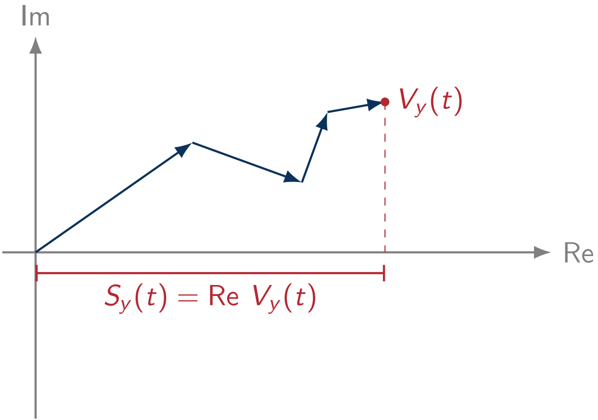
Heuristic:

- Each prime contributes a random vector of length $p^{-1/2}$.
- Directions are independent and uniform.
- Variance accumulates like $\sum_{p \le y} p^{-1} \sim \log\log y$.

A quick picture of the walk



A quick picture of the walk



The CLT is a formal version of "many small independent vectors $\Rightarrow$ Gaussian projection."

Outline

Fourier orthogonality on the circle

Let $U \sim \mathrm{Unif}[0, 2\pi)$. Then basic Fourier orthogonality gives:

$$\mathbb{E}[\cos(kU)] = 0, \qquad \mathbb{E}[\cos(kU)\cos(\ell U)] = \begin{cases} \frac{1}{2}, & k = \ell, \\ 0, & k \neq \ell. \end{cases}$$

Fourier orthogonality on the circle

Let $U \sim \mathrm{Unif}[0, 2\pi)$. Then basic Fourier orthogonality gives:

$$\mathbb{E}[\cos(kU)] = 0, \qquad \mathbb{E}[\cos(kU)\cos(\ell U)] = \begin{cases} \frac{1}{2}, & k = \ell, \\ 0, & k \neq \ell. \end{cases}$$

This lemma is the engine behind:

- mean-zero of each prime contribution,
- diagonalization of second moments,
- clean variance formulas.

Proof of orthogonality (micro-steps)

Step 1. $\mathbb{E}[\cos(kU)] = \frac{1}{2\pi}\int_0^{2\pi} \cos(ku)\, du = 0.$

Proof of orthogonality (micro-steps)

Step 1. $\mathbb{E}[\cos(kU)] = \frac{1}{2\pi}\int_0^{2\pi} \cos(ku)\,du = 0.$
Step 2. Use product-to-sum:

$$\cos(ku)\cos(\ell u) = \frac{1}{2}\cos((k-\ell)u) + \frac{1}{2}\cos((k+\ell)u).$$

Proof of orthogonality (micro-steps)

Step 1. $\mathbb{E}[\cos(kU)] = \frac{1}{2\pi}\int_0^{2\pi} \cos(ku)\, du = 0.$
Step 2. Use product-to-sum:

$$\cos(ku)\cos(\ell u) = \frac{1}{2}\cos((k-\ell)u) + \frac{1}{2}\cos((k+\ell)u).$$

Step 3. Integrate over $[0, 2\pi]$. If $k \neq \ell$, both frequencies are nonzero, so both integrals vanish.

Proof of orthogonality (micro-steps)

Step 1. $\mathbb{E}[\cos(kU)] = \frac{1}{2\pi}\int_0^{2\pi} \cos(ku)\, du = 0.$
Step 2. Use product-to-sum:

$$\cos(ku)\cos(\ell u) = \frac{1}{2}\cos((k-\ell)u) + \frac{1}{2}\cos((k+\ell)u).$$

Step 3. Integrate over $[0, 2\pi]$. If $k \neq \ell$, both frequencies are nonzero, so both integrals vanish.
Step 4. If $k = \ell$, then $\cos((k-\ell)u) = \cos(0) = 1$ and

$$\mathbb{E}[\cos^2(kU)] = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 0 = \frac{1}{2}.$$

Mean and variance of the main layer

Write $U_p(t) = \theta_p - t \log p \bmod 2\pi$. Then $U_p(t)$ are i.i.d. uniform, so

$$\mathbb{E}\left[\frac{\cos(U_p(t))}{\sqrt{p}}\right] = 0.$$

Mean and variance of the main layer

Write $U_p(t) = \theta_p - t \log p \bmod 2\pi$. Then $U_p(t)$ are i.i.d. uniform, so

$$\mathbb{E}\left[\frac{\cos(U_p(t))}{\sqrt{p}}\right] = 0.$$

Independence across primes yields

$$\mathrm{Var}(S_y(t)) = \sum_{p \le y} \mathrm{Var}\left(\frac{\cos(U_p(t))}{\sqrt{p}}\right) = \sum_{p \le y} \frac{1}{p} \mathrm{Var}(\cos(U)).$$

Mean and variance of the main layer

Write $U_p(t) = \theta_p - t \log p \bmod 2\pi$. Then $U_p(t)$ are i.i.d. uniform, so

$$\mathbb{E}\left[\frac{\cos(U_p(t))}{\sqrt{p}}\right] = 0.$$

Independence across primes yields

$$\mathrm{Var}(S_y(t)) = \sum_{p \le y} \mathrm{Var}\left(\frac{\cos(U_p(t))}{\sqrt{p}}\right) = \sum_{p \le y} \frac{1}{p} \mathrm{Var}(\cos(U)).$$

Since $\mathrm{Var}(\cos(U)) = \mathbb{E}[\cos^2(U)] = \frac{1}{2}$, we obtain the explicit variance formula

$$\sigma_y^2 := \mathrm{Var}(S_y(t)) = \frac{1}{2} \sum_{p \le y} \frac{1}{p}.$$

Why $\sigma_y \to \infty$ (Mertens scale)

A standard prime harmonic sum asymptotic states

$$\sum_{p \le y} \frac{1}{p} = \log\log y + B_1 + o(1) \qquad (y \to \infty),$$

so

$$\sigma_y^2 = \frac{1}{2} \log\log y + O(1).$$

Why $\sigma_y \to \infty$ (Mertens scale)

A standard prime harmonic sum asymptotic states

$$\sum_{p \le y} \frac{1}{p} = \log\log y + B_1 + o(1) \qquad (y \to \infty),$$

so

$$\sigma_y^2 = \frac{1}{2} \log\log y + O(1).$$

For the CLT itself, only the divergence $\sum_{p \le y} \frac{1}{p} \to \infty$ is essential.

Interpretation: the Gaussian normalization grows extremely slowly (double logarithm).

Outline

Triangular array setup

Define a prime-indexed triangular array

$$X_{p,y}(t) := \frac{\cos(\theta_p - t\log p)}{\sqrt{p}} \qquad (p \le y),$$

so

$$S_y(t) = \sum_{p \le y} X_{p,y}(t), \qquad \mathbb{E}[X_{p,y}(t)] = 0, \qquad \sum_{p \le y} \operatorname{Var}(X_{p,y}(t)) = \sigma_y^2.$$

Triangular array setup

Define a prime-indexed triangular array

$$X_{p,y}(t) := \frac{\cos(\theta_p - t \log p)}{\sqrt{p}} \qquad (p \le y),$$

so

$$S_y(t) = \sum_{p \le y} X_{p,y}(t), \qquad \mathbb{E}[X_{p,y}(t)] = 0, \qquad \sum_{p \le y} \operatorname{Var}(X_{p,y}(t)) = \sigma_y^2.$$

Uniform bound:

$$|X_{p,y}(t)| \le \frac{1}{\sqrt{p}} \le \frac{1}{\sqrt{2}}.$$

Lindeberg condition becomes trivial here

Fix $\varepsilon > 0$. We need

$$\frac{1}{\sigma_y^2} \sum_{p \le y} \mathbb{E}\left[X_{p,y}(t)^2 \, \mathbf{1}\{|X_{p,y}(t)| > \varepsilon \sigma_y\}\right] \to 0.$$

Lindeberg condition becomes trivial here

Fix $\varepsilon > 0$. We need

$$\frac{1}{\sigma_y^2} \sum_{p \le y} \mathbb{E}\big[X_{p,y}(t)^2 \, \mathbf{1}\{|X_{p,y}(t)| > \varepsilon \sigma_y\}\big] \to 0.$$

But

$$\max_{p \le y} \frac{|X_{p,y}(t)|}{\sigma_y} \le \frac{1}{\sigma_y \sqrt{2}} \to 0,$$

so for large y, $|X_{p,y}(t)| \le \varepsilon \sigma_y$ for every prime $p \le y$.

Lindeberg condition becomes trivial here

Fix $\varepsilon > 0$. We need

$$\frac{1}{\sigma_y^2}\sum_{p\le y}\mathbb{E}\big[X_{p,y}(t)^2\,\mathbf{1}\{|X_{p,y}(t)| > \varepsilon\sigma_y\}\big] \to 0.$$

But

$$\max_{p\le y}\frac{|X_{p,y}(t)|}{\sigma_y} \le \frac{1}{\sigma_y\sqrt{2}} \to 0,$$

so for large y, $|X_{p,y}(t)| \le \varepsilon\sigma_y$ for every prime $p \le y$.
Therefore the indicator is identically 0 for large y and the Lindeberg sum vanishes.

The CLT for $S_y(t)$

Theorem 5 (CLT for the main layer)

With $S_y(t) = \sum_{p\le y} \cos(\theta_p - t\log p)/\sqrt{p}$ *and* $\sigma_y^2 = \frac{1}{2}\sum_{p\le y}\frac{1}{p}$,

$$\frac{S_y(t)}{\sigma_y} \Rightarrow N(0,1) \qquad (y\to\infty).$$

The CLT for $S_y(t)$

Theorem 5 (CLT for the main layer)

With $S_y(t) = \sum_{p\le y} \cos(\theta_p - t\log p)/\sqrt{p}$ *and* $\sigma_y^2 = \frac{1}{2}\sum_{p\le y}\frac{1}{p}$,

$$\frac{S_y(t)}{\sigma_y} \Rightarrow N(0,1) \qquad (y\to\infty).$$

Proof:

- independent triangular array $(X_{p,y}(t))$,
- total variance $\sigma_y^2 \to \infty$,
- Lindeberg condition holds (previous slide),
- apply Lindeberg–Feller.

A useful way to phrase the conclusion

For any bounded continuous test function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}\left[\varphi\left(\frac{S_y(t)}{\sigma_y}\right)\right] \longrightarrow \mathbb{E}\left[\varphi(G)\right], \qquad G \sim N(0,1).$$

A useful way to phrase the conclusion

For any bounded continuous test function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}\left[\varphi\left(\frac{S_y(t)}{\sigma_y}\right)\right] \longrightarrow \mathbb{E}\left[\varphi(G)\right], \qquad G \sim N(0,1).$$

And because t-invariance holds at finite y, the limit is automatically independent of t.

Outline

What $R_y(t)$ looks like prime-by-prime

Recall

$$R_y(t) = \sum_{p \le y} \sum_{k \ge 2} \frac{\cos(k(\theta_p - t \log p))}{k\, p^{k/2}}.$$

What $R_y(t)$ looks like prime-by-prime

Recall

$$R_y(t) = \sum_{p \le y} \sum_{k \ge 2} \frac{\cos(k(\theta_p - t \log p))}{k\, p^{k/2}}.$$

Group by primes:

$$R_y(t) = \sum_{p \le y} R_p(t), \qquad R_p(t) := \sum_{k \ge 2} \frac{\cos(kU_p(t))}{k\, p^{k/2}}.$$

Primes remain independent, so the variance of R_y is a sum of one-prime variances.

One-prime variance computation (orthogonality kills cross-terms)

Since $\mathbb{E}[\cos(kU)] = 0$ and cosine modes are orthogonal,

$$\mathrm{Var}(R_p(t)) = \mathbb{E}[R_p(t)^2] = \mathbb{E}\left[\sum_{k,\ell\geq 2} \frac{\cos(kU)\cos(\ell U)}{k\ell\, p^{(k+\ell)/2}}\right].$$

One-prime variance computation (orthogonality kills cross-terms)

Since $\mathbb{E}[\cos(kU)] = 0$ and cosine modes are orthogonal,

$$\mathrm{Var}(R_p(t)) = \mathbb{E}[R_p(t)^2] = \mathbb{E}\left[\sum_{k,\ell \geq 2} \frac{\cos(kU)\cos(\ell U)}{k\ell\, p^{(k+\ell)/2}}\right].$$

Only diagonal terms contribute:

$$\mathbb{E}[\cos(kU)\cos(\ell U)] = 0 \ (k \neq \ell), \qquad \mathbb{E}[\cos^2(kU)] = \frac{1}{2}.$$

One-prime variance computation (orthogonality kills cross-terms)

Since $\mathbb{E}[\cos(kU)] = 0$ and cosine modes are orthogonal,

$$\mathrm{Var}(R_p(t)) = \mathbb{E}[R_p(t)^2] = \mathbb{E}\left[\sum_{k,\ell\geq 2} \frac{\cos(kU)\cos(\ell U)}{k\ell\, p^{(k+\ell)/2}}\right].$$

Only diagonal terms contribute:

$$\mathbb{E}[\cos(kU)\cos(\ell U)] = 0 \ (k \neq \ell), \qquad \mathbb{E}[\cos^2(kU)] = \frac{1}{2}.$$

Therefore

$$\mathrm{Var}(R_p(t)) = \frac{1}{2}\sum_{k\geq 2} \frac{1}{k^2\, p^k} \leq \frac{1}{2}\sum_{k\geq 2} \frac{1}{p^k} = \frac{1}{2} \cdot \frac{p^{-2}}{1-p^{-1}} \ll \frac{1}{p^2}.$$

Summing over primes: bounded L^2 size

Independence across primes gives

$$\mathrm{Var}(R_y(t)) = \sum_{p \le y} \mathrm{Var}(R_p(t)) \ll \sum_{p \le y} \frac{1}{p^2} \le \sum_{p} \frac{1}{p^2} < \infty.$$

Summing over primes: bounded L^2 size

Independence across primes gives

$$\mathrm{Var}(R_y(t)) = \sum_{p \le y} \mathrm{Var}(R_p(t)) \ll \sum_{p \le y} \frac{1}{p^2} \le \sum_{p} \frac{1}{p^2} < \infty.$$

Hence $\sup_{y,t} \mathbb{E}[R_y(t)^2] \le C$ for an absolute constant C, i.e.

$$R_y(t) = O_{L^2}(1) \quad \text{uniformly in } y, t.$$

Summing over primes: bounded L^2 size

Independence across primes gives

$$\mathrm{Var}(R_y(t)) = \sum_{p \le y} \mathrm{Var}(R_p(t)) \ll \sum_{p \le y} \frac{1}{p^2} \le \sum_{p} \frac{1}{p^2} < \infty.$$

Hence $\sup_{y,t} \mathbb{E}[R_y(t)^2] \le C$ for an absolute constant C, i.e.

$$R_y(t) = O_{L^2}(1) \quad \text{uniformly in } y, t.$$

Since $\sigma_y \to \infty$,

$$\frac{R_y(t)}{\sigma_y} \to 0 \quad \text{in } L^2 \text{ and hence in probability.}$$

Outline

Slutsky

We have two limits:

$$\frac{S_y(t)}{\sigma_y} \Rightarrow N(0,1), \qquad \frac{R_y(t)}{\sigma_y} \to 0 \text{ in probability.}$$

Slutsky

We have two limits:

$$\frac{S_y(t)}{\sigma_y} \Rightarrow N(0,1), \qquad \frac{R_y(t)}{\sigma_y} \to 0 \text{ in probability.}$$

Thus by Slutsky's theorem,

$$\frac{X_y(t)}{\sigma_y} = \frac{S_y(t)}{\sigma_y} + \frac{R_y(t)}{\sigma_y} \Rightarrow N(0,1).$$

Slutsky

We have two limits:

$$\frac{S_y(t)}{\sigma_y} \Rightarrow N(0,1), \qquad \frac{R_y(t)}{\sigma_y} \to 0 \text{ in probability.}$$

Thus by Slutsky's theorem,

$$\frac{X_y(t)}{\sigma_y} = \frac{S_y(t)}{\sigma_y} + \frac{R_y(t)}{\sigma_y} \Rightarrow N(0,1).$$

This is the main theorem.

Variance bookkeeping (sanity)

Because $R_y(t) = O_{L^2}(1)$ and $S_y(t)$ has variance $\sigma_y^2 \to \infty$,

$$\mathrm{Var}(X_y(t)) = \mathrm{Var}(S_y(t)) + O(1) = \sigma_y^2 + O(1).$$

Variance bookkeeping (sanity)

Because $R_y(t) = O_{L^2}(1)$ and $S_y(t)$ has variance $\sigma_y^2 \to \infty$,

$$\mathrm{Var}(X_y(t)) = \mathrm{Var}(S_y(t)) + O(1) = \sigma_y^2 + O(1).$$

So the normalization σ_y is correct for X_y as well.

Outline

Why a characteristic-function route is natural

Independence across primes makes the characteristic function factorize. For $\lambda \in \mathbb{R}$,

$$\mathbb{E}\left[\exp\left(i\lambda \frac{S_y(t)}{\sigma_y}\right)\right] = \prod_{p \leq y} \mathbb{E}\left[\exp\left(i\lambda \frac{\cos(U_p(t))}{\sigma_y \sqrt{p}}\right)\right].$$

Why a characteristic-function route is natural

Independence across primes makes the characteristic function factorize. For $\lambda \in \mathbb{R}$,

$$\mathbb{E}\left[\exp\left(i\lambda\frac{S_y(t)}{\sigma_y}\right)\right] = \prod_{p\le y}\mathbb{E}\left[\exp\left(i\lambda\frac{\cos(U_p(t))}{\sigma_y\sqrt{p}}\right)\right].$$

So everything reduces to understanding

$$\mathbb{E}[e^{ia\cos U}] \quad \text{for small } a.$$

One-prime factor: a Bessel function J_0

Lemma 6

If $U \sim \mathrm{Unif}[0, 2\pi)$, *then*

$$\mathbb{E}[e^{ia\cos U}] = J_0(a),$$

the Bessel function of the first kind. Moreover, as $a \to 0$,

$$\log J_0(a) = -\frac{a^2}{4} + O(a^4).$$

One-prime factor: a Bessel function J_0

Lemma 6

If $U \sim \mathrm{Unif}[0, 2\pi)$, then

$$\mathbb{E}[e^{ia\cos U}] = J_0(a),$$

the Bessel function of the first kind. Moreover, as $a \to 0$,

$$\log J_0(a) = -\frac{a^2}{4} + O(a^4).$$

Heuristic consequence:

$$\log \mathbb{E}\left[e^{i\lambda S_y/\sigma_y}\right] = \sum_{p \le y} \log J_0\left(\frac{\lambda}{\sigma_y \sqrt{p}}\right) \approx -\frac{\lambda^2}{4} \sum_{p \le y} \frac{1}{\sigma_y^2 p} = -\frac{\lambda^2}{2},$$

matching the characteristic function of $N(0, 1)$.

Where the Gaussian term wins

The quadratic term sums like

$$\sum_{p \le y} \frac{1}{p} \sim \log\log y \quad \Rightarrow \quad \text{diverges.}$$

Where the Gaussian term wins

The quadratic term sums like

$$\sum_{p \le y} \frac{1}{p} \sim \log\log y \quad \Rightarrow \quad \text{diverges.}$$

The next error term is governed by

$$\sum_{p \le y} \frac{1}{p^2} \quad \Rightarrow \quad \text{converges.}$$

Where the Gaussian term wins

The quadratic term sums like

$$\sum_{p\leq y}\frac{1}{p}\sim \log\log y \quad \Rightarrow \quad \text{diverges.}$$

The next error term is governed by

$$\sum_{p\leq y}\frac{1}{p^2} \quad \Rightarrow \quad \text{converges.}$$

So the prime-by-prime cumulant tail is summable, and the quadratic term dominates, yielding a CLT (and often a route to rates).

Outline

What we check

We simulate the normalized main term

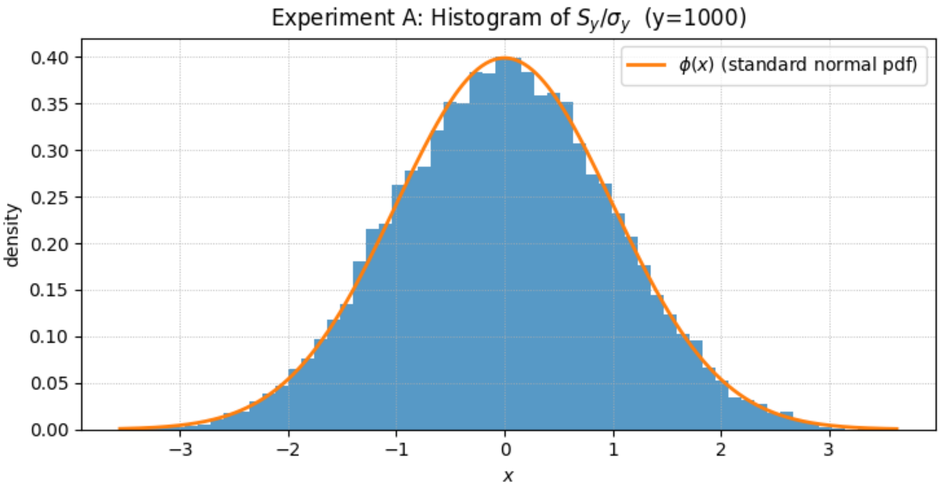
$$\frac{S_y}{\sigma_y}, \qquad S_y = \sum_{p \le y} \frac{\cos \theta_p}{\sqrt{p}}, \qquad \sigma_y^2 = \frac{1}{2} \sum_{p \le y} \frac{1}{p},$$

at a fixed cutoff $y = 1000$ (many independent trials) and compare to $N(0, 1)$:

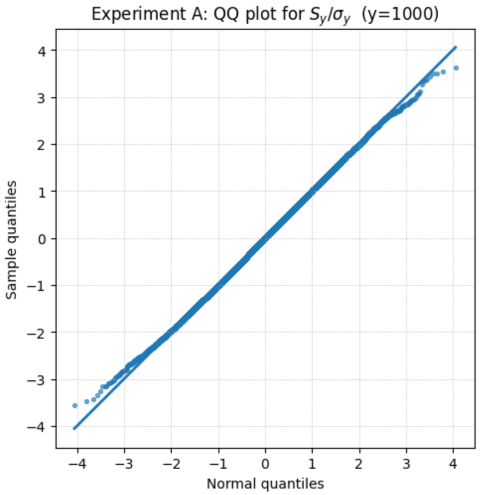
- histogram with standard normal density overlay,
- QQ plot versus standard normal quantiles.

This directly probes the CLT prediction for S_y/σ_y.

QQ plot at $y = 1000$



Numerical experiment: histogram at $y = 1000$



Outline

What the model captures

This model isolates a mechanism:

- prime-indexed independent phases $\Rightarrow$ sum of small independent increments,
- variance scale $\sigma_y^2 = \frac{1}{2}\sum_{p\le y}\frac{1}{p} \sim \frac{1}{2}\log\log y$,
- higher Euler-power layers square-summable $\Rightarrow$ $O_{L^2}(1)$ remainder,
- therefore X_y/σ_y is asymptotically Gaussian.

What the model omits

In the true zeta setting:

- phases $p^{-it} = e^{-it \log p}$ are deterministic and coupled through a single parameter t,
- one studies distribution as t varies, not literal independence across primes,
- rare resonant events and multi-t constraints are beyond this mean-field surrogate.

What the model omits

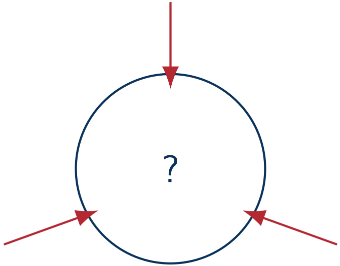
In the true zeta setting:

- phases $p^{-it} = e^{-it \log p}$ are deterministic and coupled through a single parameter t,
- one studies distribution as t varies, not literal independence across primes,
- rare resonant events and multi-t constraints are beyond this mean-field surrogate.

So this is best read as a clean "variance-scale + typical fluctuation" model.

Thank you!

Questions?



(ilt2109@columbia.edu / https://iz-tecum.github.io/)

Outline

References I

- H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Univ. Press, 2007.
- G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, 3rd ed., American Mathematical Society, 2015.
- P. Billingsley, *Probability and Measure*, 3rd ed., Wiley, 1995.
- R. Durrett, *Probability: Theory and Examples*, 5th ed., Cambridge Univ. Press, 2019.
- W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. II*, 2nd ed., Wiley, 1971.
- A. Gut, *Probability: A Graduate Course*, 2nd ed., Springer, 2013.
- V. V. Petrov, *Limit Theorems of Probability Theory*, Oxford Univ. Press, 1995.
- Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3rd ed., Springer, 1997.
- G. Harman, *Prime-Detecting Sieves*, Princeton Univ. Press, 2007.

References II

- A. Ivić, *The Riemann Zeta-Function: Theory and Applications*, Dover, 2003.
- W. Narkiewicz, *The Development of Prime Number Theory*, Springer, 2000.
- E. Kowalski, *An Introduction to Probabilistic Number Theory*, lecture notes / manuscript.