Quantitative Erdös-Kac theorem for additive functions, a self-contained probabilistic approach

Joint work with X. Yang and Louis H. Y. Chen

Arturo Jaramillo Gil

Université du Luxembourg
National University of Singapore

Goal

Denote the set of primes by $\mathcal{P}$.

Goal

Denote the set of primes by $\mathcal{P}$. Let $\omega : \mathbb{N} \to \mathbb{N}$ denote the prime factor counting function,

$$\omega(n) := |\{p \in \mathcal{P};\ p \text{ divides } n\}|.$$

Goal

Denote the set of primes by $\mathcal{P}$. Let $\omega : \mathbb{N} \to \mathbb{N}$ denote the prime factor counting function,

$$\omega(n) := |\{p \in \mathcal{P};\ p \text{ divides } n\}|.$$

For instance, $\omega(54) = \omega(2 \times 3^2) = 2$.

Goal

Denote the set of primes by $\mathcal{P}$. Let $\omega : \mathbb{N} \to \mathbb{N}$ denote the prime factor counting function,

$$\omega(n) := |\{p \in \mathcal{P};\ p \text{ divides } n\}|.$$

For instance, $\omega(54) = \omega(2 \times 3^2) = 2$. Let J_n be a random variable with uniform distribution over $\{1, \dots, n\}$.

Goal

Denote the set of primes by $\mathcal{P}$. Let $\omega : \mathbb{N} \to \mathbb{N}$ denote the prime factor counting function,

$$\omega(n) := |\{p \in \mathcal{P};\ p \text{ divides } n\}|.$$

For instance, $\omega(54) = \omega(2 \times 3^2) = 2$. Let J_n be a random variable with uniform distribution over $\{1, \ldots, n\}$.

Goal

- Study $\omega(J_n)$.

Goal

Denote the set of primes by $\mathcal{P}$. Let $\omega : \mathbb{N} \to \mathbb{N}$ denote the prime factor counting function,

$$\omega(n) := |\{p \in \mathcal{P};\ p \text{ divides } n\}|.$$

For instance, $\omega(54) = \omega(2 \times 3^2) = 2$. Let J_n be a random variable with uniform distribution over $\{1, \ldots, n\}$.

Goal

- Study $\omega(J_n)$. Describe as accurately as possible the asymptotic behavior of $\frac{\omega(J_n) - \mu_n}{\sigma_n}$, for suitable chosen μ_n and σ_n.

Goal

Denote the set of primes by $\mathcal{P}$. Let $\omega : \mathbb{N} \to \mathbb{N}$ denote the prime factor counting function,

$$\omega(n) := |\{p \in \mathcal{P};\ p \text{ divides } n\}|.$$

For instance, $\omega(54) = \omega(2 \times 3^2) = 2$. Let J_n be a random variable with uniform distribution over $\{1, \ldots, n\}$.

Goal

- Study $\omega(J_n)$. Describe as accurately as possible the asymptotic behavior of $\frac{\omega(J_n) - \mu_n}{\sigma_n}$, for suitable chosen μ_n and σ_n.
- What can be said when ω is replaced by a general function $\psi : \mathbb{N} \to \mathbb{N}$ only satisfying $\psi(ab) = \psi(a) + \psi(b)$ for $a, b \in \mathbb{N}$ coprime?

Plan

1. Historical context
2. Main results
3. Ideas behind the proofs
    - Simplifying the model
    - Stein's method

Historical context

Classical Erdös-Kac theorem (1940)

Starting point: Paul Erdös and Mark Kac, proved that

$$Z_n := \frac{\omega(J_n) - \log\log(n)}{\sqrt{\log\log(n)}} \tag{1}$$

converge in distribution towards a standard Gaussian random variable $\mathcal{N}$.

Classical Erdös-Kac theorem (1940)

Starting point: Paul Erdös and Mark Kac, proved that

$$Z_n := \frac{\omega(J_n) - \log\log(n)}{\sqrt{\log\log(n)}} \tag{1}$$

converge in distribution towards a standard Gaussian random variable $\mathcal{N}$.

Some intuition: Denote $\mathcal{P}_n := \mathcal{P} \cap [1, n]$.

Classical Erdös-Kac theorem (1940)

Starting point: Paul Erdös and Mark Kac, proved that

$$Z_n := \frac{\omega(J_n) - \log\log(n)}{\sqrt{\log\log(n)}} \tag{1}$$

converge in distribution towards a standard Gaussian random variable $\mathcal{N}$.

Some intuition: Denote $\mathcal{P}_n := \mathcal{P} \cap [1, n]$. The convergence in (1) is hinted by the decomposition

$$\omega(J_n) = \sum_{p \in \mathcal{P}_n} \mathbb{1}_{\{p \text{ divides } J_n\}}, \tag{2}$$

Intuition about Erdös-Kac theorem

One guesses that $\mathbb{1}_{\{p \text{ divides } J_n\}}$ are weakly dependent since for $d \in \mathbb{N}$,

$$\mathbb{P}[d \text{ divides } J_n] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{d \text{ divides } k\}} = \frac{1}{n}\left\lfloor \frac{n}{d} \right\rfloor \approx \frac{1}{d}. \tag{3}$$

Intuition about Erdös-Kac theorem

One guesses that $\mathbb{1}_{\{p \text{ divides } J_n\}}$ are weakly dependent since for $d \in \mathbb{N}$,

$$\mathbb{P}[d \text{ divides } J_n] = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{d \quad \text{divides} \quad k\}} = \frac{1}{n} \left\lfloor \frac{n}{d} \right\rfloor \approx \frac{1}{d}. \tag{3}$$

Thus, if $p_1, \ldots, p_r \in \mathcal{P}_n$ are different primes,

$$\begin{aligned} \mathbb{P}[\mathbb{1}_{\{p_1 \text{ divides } J_n\}} = 1, \ldots, \mathbb{1}_{\{p_r \text{ divides } J_n\}} = 1] &\approx \frac{1}{p_1 \cdots p_r} \\ &\approx \mathbb{P}[\mathbb{1}_{\{p_1 \text{ divides } J_n\}} = 1] \cdots \mathbb{P}[\mathbb{1}_{\{p_r \text{ divides } J_n\}} = 1] \end{aligned}$$

Intuition about Erdös-Kac theorem

One guesses that $\mathbb{1}_{\{p \text{ divides } J_n\}}$ are weakly dependent since for $d \in \mathbb{N}$,

$$\mathbb{P}[d \text{ divides } J_n] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{d \text{ divides } k\}} = \frac{1}{n}\left\lfloor \frac{n}{d} \right\rfloor \approx \frac{1}{d}. \tag{3}$$

Thus, if $p_1, \dots, p_r \in \mathcal{P}_n$ are different primes,

$$\begin{aligned}\mathbb{P}[\mathbb{1}_{\{p_1 \text{ divides } J_n\}} = 1, \dots, \mathbb{1}_{\{p_r \text{ divides } J_n\}} = 1] &\approx \frac{1}{p_1 \cdots p_r} \\ &\approx \mathbb{P}[\mathbb{1}_{\{p_1 \text{ divides } J_n\}} = 1] \cdots \mathbb{P}[\mathbb{1}_{\{p_r \text{ divides } J_n\}} = 1]\end{aligned}$$

Warning: nowadays it is known that the r.v. $\mathbb{1}_{\{p \text{ divides } J_n\}}$, for $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ are approximately independent if $\alpha_n \to \infty$ is suitably chosen (example: $\alpha_n := 3 \log\log(n)^2$).

Intuition about Erdös-Kac theorem

One guesses that $\mathbb{1}_{\{p \text{ divides } J_n\}}$ are weakly dependent since for $d \in \mathbb{N}$,

$$\mathbb{P}[d \text{ divides } J_n] = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{d \text{ divides } k\}} = \frac{1}{n} \left\lfloor \frac{n}{d} \right\rfloor \approx \frac{1}{d}. \tag{3}$$

Thus, if $p_1, \ldots, p_r \in \mathcal{P}_n$ are different primes,

$$\begin{aligned} \mathbb{P}[\mathbb{1}_{\{p_1 \text{ divides } J_n\}} = 1, \ldots, \mathbb{1}_{\{p_r \text{ divides } J_n\}} = 1] &\approx \frac{1}{p_1 \cdots p_r} \\ &\approx \mathbb{P}[\mathbb{1}_{\{p_1 \text{ divides } J_n\}} = 1] \cdots \mathbb{P}[\mathbb{1}_{\{p_r \text{ divides } J_n\}} = 1] \end{aligned}$$

Warning: nowadays it is known that the r.v. $\mathbb{1}_{\{p \text{ divides } J_n\}}$, for $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ are approximately independent if $\alpha_n \to \infty$ is suitably chosen (example: $\alpha_n := 3 \log\log(n)^2$). *But α_n cannot be equal to one!*

Question

Can the asymptotic Gaussianity of Z_n be quantitatively assessed with respect to a suitable probability metric? such as distance d_{K}, defined as

$$d_{\mathrm{K}}(X, Y) = \sup_{z \in \mathbb{R}} |\mathbb{P}[X \leq z] - \mathbb{P}[Y \leq z]|$$

Question

Can the asymptotic Gaussianity of Z_n be quantitatively assessed with respect to a suitable probability metric? such as distance d_{K}, defined as

$$d_{\mathrm{K}}(X, Y) = \sup_{z \in \mathbb{R}} |\mathbb{P}[X \leq z] - \mathbb{P}[Y \leq z]|$$

or

$$d_1(X, Y) = \sup_{h \in \mathrm{Lip}_1} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|,$$

where Lip_1 is the family of Lipschitz functions with Lipschitz constant at most one.

Question

Can the asymptotic Gaussianity of Z_n be quantitatively assessed with respect to a suitable probability metric? such as distance d_{K}, defined as

$$d_{\mathrm{K}}(X, Y) = \sup_{z \in \mathbb{R}} |\mathbb{P}[X \leq z] - \mathbb{P}[Y \leq z]|$$

or

$$d_1(X, Y) = \sup_{h \in \mathrm{Lip}_1} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|,$$

where Lip_1 is the family of Lipschitz functions with Lipschitz constant at most one. We define as well

$$d_{TV}(X, Y) = \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathbb{P}[X \in A] - \mathbb{P}[Y \in A]|.$$

LeVeque's conjecture (1949)

LeVeque, showed that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \frac{\log \log \log(n)}{\log \log(n)^{\frac{1}{4}}},$$

for some constant $C > 0$ independent of n.

LeVeque's conjecture (1949)

LeVeque, showed that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \frac{\log \log \log(n)}{\log \log(n)^{\frac{1}{4}}},$$

for some constant $C > 0$ independent of n. He also conjectured that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \log \log(n)^{-\frac{1}{2}}.$$

LeVeque's conjecture (1949)

LeVeque, showed that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \frac{\log\log\log(n)}{\log\log(n)^{\frac{1}{4}}},$$

for some constant $C > 0$ independent of n. He also conjectured that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \log\log(n)^{-\frac{1}{2}}.$$

This was shown to be true later by Rényi and Turán (1958).

LeVeque's conjecture (1949)

LeVeque, showed that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \le C \frac{\log\log\log(n)}{\log\log(n)^{\frac{1}{4}}},$$

for some constant $C > 0$ independent of n. He also conjectured that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \le C \log\log(n)^{-\frac{1}{2}}.$$

This was shown to be true later by Rényi and Turán (1958). The approach consisted on approximating $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$.

LeVeque's conjecture (1949)

LeVeque, showed that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \frac{\log\log\log(n)}{\log\log(n)^{\frac{1}{4}}},$$

for some constant $C > 0$ independent of n. He also conjectured that

$$d_{\mathrm{K}}(Z_n, \mathcal{N}) \leq C \log\log(n)^{-\frac{1}{2}}.$$

This was shown to be true later by Rényi and Turán (1958). The approach consisted on approximating $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$.

Main ingredients: Perron's formula, Dirichlet series and some estimates on the Riemann zeta function ζ around the vertical strip $\{z \in \mathbb{C}\ ;\ \Re(z) = 1\}$.

Mod-ϕ convergence approach (2014)

Nowadays, a lot is known about $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$.

Mod-ϕ convergence approach (2014)

Nowadays, a lot is known about $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$. This motivated studying the asymptotic properties of $\omega(J_n)$ by analyzing a relation of the type

$$\frac{\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]}{\mathbb{E}[e^{\mathbf{i}\lambda M_n}]} \approx F(\lambda),$$

where M_n is a random variable with Poisson distribution of parameter $\log\log(n)$ and $F(\lambda)$ is a (possibly non-trivial) function (work by Barbour, Kowalski and Nikeghbali in 2014).

Mod-ϕ convergence approach (2014)

Nowadays, a lot is known about $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$. This motivated studying the asymptotic properties of $\omega(J_n)$ by analyzing a relation of the type

$$\frac{\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]}{\mathbb{E}[e^{\mathbf{i}\lambda M_n}]} \approx F(\lambda),$$

where M_n is a random variable with Poisson distribution of parameter $\log\log(n)$ and $F(\lambda)$ is a (possibly non-trivial) function (work by Barbour, Kowalski and Nikeghbali in 2014). This lead to

Theorem
There exists a constant $C > 0$, such that

$$d_{TV}(\omega(J_n), M_n) \leq C \log\log(n)^{-\frac{1}{2}}. \tag{4}$$

Other approaches (Stein's method)

Recall the heuristics that $\mathbb{1}_{\{p \text{ divides } J_n\}}$, for $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ are approximately independent.

Other approaches (Stein's method)

Recall the heuristics that $\mathbb{1}_{\{p \text{ divides } J_n\}}$, for $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ are approximately independent.

Stein's method perspective

Random variables with weak dependence are the main object in the theory of Stein's method.

Other approaches (Stein's method)

Recall the heuristics that $\mathbb{1}_{\{p \text{ divides } J_n\}}$, for $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ are approximately independent.

Stein's method perspective
Random variables with weak dependence are the main object in the theory of Stein's method.

Adam Harper (2009) used this so show

$$d_{TV}\left(\sum_{p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]} \mathbb{1}_{\{p \text{ divides } J_n\}}, M_n \right) \leq \frac{1}{2 \log \log(n)} + \frac{5.2}{\log \log(n)^{\frac{3}{2}}},$$

where $\alpha_n := 3 \log \log(n)^2$.

Other approaches (Stein's method)

Recall the heuristics that $\mathbb{1}_{\{p \text{ divides } J_n\}}$, for $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ are approximately independent.

Stein's method perspective
Random variables with weak dependence are the main object in the theory of Stein's method.

Adam Harper (2009) used this so show

$$d_{TV}\left(\sum_{p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]} \mathbb{1}_{\{p \text{ divides } J_n\}}, M_n \right) \leq \frac{1}{2 \log \log(n)} + \frac{5.2}{\log \log(n)^{\frac{3}{2}}},$$

where $\alpha_n := 3 \log \log(n)^2$. Consequence,

$$d_K\big(\omega(J_n), M_n\big) \leq \frac{C \log \log \log(n)}{\sqrt{\log \log(n)}}.$$

Other approaches (Independence approximation)

Another idea consists on comparing $(\mathbb{1}_{\{p \text{ divides } J_n\}}\ ;\ p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}})$ with independent random variables.

Other approaches (Independence approximation)

Another idea consists on comparing $(\mathbb{1}_{\{p \text{ divides } J_n\}}\ ;\ p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}})$ with independent random variables.

Kubilius (1964) showed that if $\beta_n \to \infty$,

$$d_{TV}\left((\mathbb{1}_{\{p \text{ divides } J_n\}};\ p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}}]), (B_p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}}])\right) \leq e^{-c\beta_n},$$

where B_p are independent Bernoulli r.v. with $\mathbb{P}[B_p] = 1/p$.

Other approaches (Independence approximation)

Another idea consists on comparing $(\mathbb{1}_{\{p \text{ divides } J_n\}} \; ; \; p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}})$ with independent random variables.

Kubilius (1964) showed that if $\beta_n \to \infty$,

$$d_{TV}\left((\mathbb{1}_{\{p \text{ divides } J_n\}}; \; p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}}]), (B_p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}}]) \right) \leq e^{-c\beta_n},$$

where B_p are independent Bernoulli r.v. with $\mathbb{P}[B_p] = 1/p$. Thus,

$$\sum_{p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}}]} \mathbb{1}_{\{p \text{ divides } J_n\}} \approx \sum_{p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}}]} B_p$$

Consequence similar to Harper's result.

The size biased permutation approach

Arratia (2013) suggests comparing J_n with a partial product of a biased permutation of factors T_n and a random prime P_n. He proves that

$$d_{TV}(J_n, T_nQ_n) \leq C\frac{\log\log(n)}{\log(n)}.$$

The size biased permutation approach

Arratia (2013) suggests comparing J_n with a partial product of a biased permutation of factors T_n and a random prime P_n. He proves that

$$d_{TV}(J_n, T_n Q_n) \leq C \frac{\log\log(n)}{\log(n)}.$$

This is used to show that if $d_\Omega : \mathbb{N}^2 \to \mathbb{N}$ denotes the insertion deletion distance $d_\Omega(\prod_{p\in\mathcal{P}} p^{\alpha_p}, \prod_{p\in\mathcal{P}} p^{\beta_p}) := \sum_{p\in\mathcal{P}} |\alpha_p - \beta_p|$, and $d_{1,\Omega}$ the associated Wasserstein distance,

The size biased permutation approach

Arratia (2013) suggests comparing J_n with a partial product of a biased permutation of factors T_n and a random prime P_n. He proves that

$$d_{TV}(J_n, T_n Q_n) \leq C \frac{\log \log(n)}{\log(n)}.$$

This is used to show that if $d_\Omega : \mathbb{N}^2 \to \mathbb{N}$ denotes the insertion deletion distance $d_\Omega(\prod_{p \in \mathcal{P}} p^{\alpha_p}, \prod_{p \in \mathcal{P}} p^{\beta_p}) := \sum_{p \in \mathcal{P}} |\alpha_p - \beta_p|$, and $d_{1,\Omega}$ the associated Wasserstein distance,

$$\lim_{n \to \infty} d_{1,\Omega}(J_n, \prod_{p \in \mathcal{P}_n} p^{\xi_p}) = 2,$$

where

$$\mathbb{P}[\xi_p = k] = p^{-k}(1 - 1/p),$$

for $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Main results

CLT for additive functions

Let $\psi : \mathbb{N} \to \mathbb{N}$ be such that $\psi(ab) = \psi(a) + \psi(b)$ for a, b co-prime.

(H1) We have that

$$\|\psi\|_{\mathcal{P}} := \sup_{p \in \mathcal{P}} |\psi(p)| < \infty.$$

(H2) There exists a (possibly unbounded) function $\Psi : \mathcal{P} \to \mathbb{R}_+$ satisfying

$$|||\Psi|||_{\mathcal{P}} := \left(\sum_{p \in \mathcal{P}} \frac{\Psi(p)^2}{p^2} \right)^{1/2} < \infty,$$

and such that for all $p \in \mathcal{P}_n$,

$$\|\psi(p^{\xi_p + 2})\|_{L^2(\Omega)} \leq \Psi(p).$$

Main result for Kolmogorov distance

Let μ_n and $\sigma_n > 0$ be given by

$$\mu_n = \sum_{p \in \mathcal{P}_n} \mathbb{E}[\psi(p^{\xi_p})] \qquad \text{and} \qquad \sigma_n^2 = \sum_{p \in \mathcal{P}_n} \mathbb{V}\mathrm{ar}[\psi(p^{\xi_p})]. \tag{5}$$

Main result for Kolmogorov distance

Let μ_n and $\sigma_n > 0$ be given by

$$\mu_n = \sum_{p \in \mathcal{P}_n} \mathbb{E}[\psi(p^{\xi_p})] \qquad \text{and} \qquad \sigma_n^2 = \sum_{p \in \mathcal{P}_n} \mathbb{V}\mathrm{ar}[\psi(p^{\xi_p})]. \tag{5}$$

Theorem (Chen, Jaramillo, Yang)

*Suppose that ψ satisfies **(H1)** and **(H2)**. Then, if $X_p := \sigma_n^{-1}\psi(p^{\xi_p})$, and provided that $\sigma_n^2 \geq 3(\|\psi\|_{\mathcal{P}}^2 + \|\!|\Psi|\!\|_{\mathcal{P}}^2)$,*

$$d_{\mathrm{K}}\left(\frac{\psi(J_n) - \mu_n}{\sigma_n}, \mathcal{N}\right) \leq \frac{\kappa_1}{\sigma_n} + \kappa_2 \sum_{p \in \mathcal{P}_n} \mathbb{E}[|X_p|^3] + \frac{\kappa_3 \log\log(n)}{\log(n)},$$

where

$$\kappa_1 := 29.2\|\psi\|_{\mathcal{P}} + 34.8\|\!|\Psi|\!\|_{\mathcal{P}} \quad \kappa_2 := 97.2 \quad \kappa_3 := 61. \tag{6}$$

Main result for Wasserstein distance

Theorem (Chen, Jaramillo, Yang)

$$d_1\left(\frac{\psi(J_n)-\mu_n}{\sigma_n}, \mathcal{N}\right) \leq \frac{\kappa_4}{\sigma_n} + \kappa_5 \sum_{p\in\mathcal{P}_n} \mathbb{E}[|X_p|^3] + \kappa_6 \frac{\log\log(n)^{\frac{3}{2}}}{\log(n)^{\frac{1}{2}}}, \qquad (7)$$

where

$$\kappa_4 := 16.6\|\psi\|_{\mathcal{P}} + 11.3|||\Psi|||_{\mathcal{P}} \quad \kappa_5 := 24 \quad \kappa_6 := 21\|\psi\|_{\mathcal{P}} + 45.$$

Ideas behind the proofs

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

For any $i \in \mathbb{N}$ and $k_1, ..., k_i \in \mathbb{N}_0$,

$$\bigcap_{j=1}^{i} \{\alpha_{p_j}(J_n) \geq k_j\} = \bigcap_{j=1}^{i} \{p_j^{k_j} \text{ divides } J_n\} = \left\{\prod_{j=1}^{i} p_j^{k_j} \text{ divides } J_n\right\},$$

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

For any $i \in \mathbb{N}$ and $k_1, ..., k_i \in \mathbb{N}_0$,

$$\bigcap_{j=1}^{i} \{\alpha_{p_j}(J_n) \geq k_j\} = \bigcap_{j=1}^{i} \{p_j^{k_j} \text{ divides } J_n\} = \left\{ \prod_{j=1}^{i} p_j^{k_j} \text{ divides } J_n \right\},$$

so $\lim_{n\to\infty} \mathbb{P}[\alpha_{p_j}(J_n) \geq k_j \ \text{ for all } 1 \leq j \leq i] = p_1^{-k_1} \cdots p_i^{-k_i}$.

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

For any $i \in \mathbb{N}$ and $k_1, ..., k_i \in \mathbb{N}_0$,

$$\bigcap_{j=1}^{i} \{\alpha_{p_j}(J_n) \geq k_j\} = \bigcap_{j=1}^{i} \{p_j^{k_j} \text{ divides } J_n\} = \left\{ \prod_{j=1}^{i} p_j^{k_j} \text{ divides } J_n \right\},$$

so $\lim_{n\to\infty} \mathbb{P}[\alpha_{p_j}(J_n) \geq k_j \text{ for all } 1 \leq j \leq i] = p_1^{-k_1} \cdots p_i^{-k_i}$. This gives

$$(\alpha_{p_1}(J_n), \ldots, \alpha_{p_i}(J_n)) \overset{Law}{\to} (\xi_{p_1}, \ldots, \xi_{p_i})$$

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

For any $i \in \mathbb{N}$ and $k_1, ..., k_i \in \mathbb{N}_0$,

$$\bigcap_{j=1}^{i} \{\alpha_{p_j}(J_n) \geq k_j\} = \bigcap_{j=1}^{i} \{p_j^{k_j} \text{ divides } J_n\} = \left\{ \prod_{j=1}^{i} p_j^{k_j} \text{ divides } J_n \right\},$$

so $\lim_{n\to\infty} \mathbb{P}[\alpha_{p_j}(J_n) \geq k_j \;\text{ for all } 1 \leq j \leq i] = p_1^{-k_1} \cdots p_i^{-k_i}$. This gives

$$(\alpha_{p_1}(J_n), \ldots, \alpha_{p_i}(J_n)) \overset{Law}{\to} (\xi_{p_1}, \ldots, \xi_{p_i})$$

Question: can we use the ξ_p to construct a r.v. equal in law to J_n?

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

For any $i \in \mathbb{N}$ and $k_1, ..., k_i \in \mathbb{N}_0$,

$$\bigcap_{j=1}^{i} \{\alpha_{p_j}(J_n) \geq k_j\} = \bigcap_{j=1}^{i} \{p_j^{k_j} \text{ divides } J_n\} = \left\{\prod_{j=1}^{i} p_j^{k_j} \text{ divides } J_n\right\},$$

so $\lim_{n\to\infty} \mathbb{P}[\alpha_{p_j}(J_n) \geq k_j \text{ for all } 1 \leq j \leq i] = p_1^{-k_1} \cdots p_i^{-k_i}$. This gives

$$(\alpha_{p_1}(J_n), \dots, \alpha_{p_i}(J_n)) \overset{Law}{\to} (\xi_{p_1}, \dots, \xi_{p_i})$$

Question: can we use the ξ_p to construct a r.v. equal in law to J_n?
Answer: not easily

Multiplicities of prime factors

For a given $p \in \mathcal{P}$, define $\alpha_p : \mathbb{N} \to \mathbb{N}$, by

$$k = \prod_{p \in \mathcal{P}} p^{\alpha_p(k)}. \tag{8}$$

For any $i \in \mathbb{N}$ and $k_1, ..., k_i \in \mathbb{N}_0$,

$$\bigcap_{j=1}^{i} \{\alpha_{p_j}(J_n) \geq k_j\} = \bigcap_{j=1}^{i} \{p_j^{k_j} \text{ divides } J_n\} = \left\{ \prod_{j=1}^{i} p_j^{k_j} \text{ divides } J_n \right\},$$

so $\lim_{n\to\infty} \mathbb{P}[\alpha_{p_j}(J_n) \geq k_j \;\text{ for all } 1 \leq j \leq i] = p_1^{-k_1} \cdots p_i^{-k_i}$. This gives

$$(\alpha_{p_1}(J_n), \dots, \alpha_{p_i}(J_n)) \overset{Law}{\to} (\xi_{p_1}, \dots, \xi_{p_i})$$

Question: can we use the ξ_p to construct a r.v. equal in law to J_n?
Answer: not easily... but...

A simplified model: Harmonic distribution H_n

Let H_n be a r.v. with $\mathbb{P}[H_n = k] = \frac{1}{L_n k}$, where $L_n := \sum_{k=1}^{n} \frac{1}{k}$.

A simplified model: Harmonic distribution H_n

Let H_n be a r.v. with $\mathbb{P}[H_n = k] = \frac{1}{L_n k}$, where $L_n := \sum_{k=1}^{n} \frac{1}{k}$. Then,

Proposition

Suppose that $n \geq 21$. Define the event

$$A_n := \Big\{ \prod_{p \in \mathcal{P}_n} p^{\xi_p} \leq n \Big\}, \tag{9}$$

as well as the random vector $\vec{C}(n) := (\alpha_p(H_n); p \in \mathcal{P}_n)$. Then the random variables $Y_p := \psi(p^{\xi_p})$, indexed by $p \in \mathcal{P}_n$, satisfy

$$\mathcal{L}(\psi(H(n))) = \mathcal{L}(\sum_{p \in \mathcal{P}_n} Y_p | A_n). \tag{10}$$

Link to the Harmonic distribution

Let $\{Q(k)\}_{k\geq 1}$ be independent r.v. independent of (J_n, H_n) with $Q(k)$ uniformly distributed over

$$\mathcal{P}_k^* := \{1\} \cup \mathcal{P}_k.$$

Link to the Harmonic distribution

Let $\{Q(k)\}_{k\geq 1}$ be independent r.v. independent of (J_n, H_n) with $Q(k)$ uniformly distributed over

$$\mathcal{P}_k^* := \{1\} \cup \mathcal{P}_k.$$

Let $\pi(n) := |\mathcal{P} \cap [1, n]|$. Using the fact that for $n \geq 229$,

$$\left|\pi(n) - \int_0^n \frac{1}{\log(t)} dt\right| \leq \frac{181n}{\log(n)^3}, \tag{11}$$

Link to the Harmonic distribution

Let $\{Q(k)\}_{k\geq 1}$ be independent r.v. independent of (J_n, H_n) with $Q(k)$ uniformly distributed over

$$\mathcal{P}_k^* := \{1\} \cup \mathcal{P}_k.$$

Let $\pi(n) := |\mathcal{P} \cap [1, n]|$. Using the fact that for $n \geq 229$,

$$\left|\pi(n) - \int_0^n \frac{1}{\log(t)} dt\right| \leq \frac{181n}{\log(n)^3}, \tag{11}$$

Lemma (Chen, Jaramillo and Yang)

The following bound (analogous to the one by Arratia) holds for $n \geq 21$

$$d_{\mathrm{TV}}(J_n, H_n Q(n/H_n)) \leq 61 \frac{\log \log n}{\log n}.$$

Simplifying $\omega(J_n)$ to $\omega(H_n)$

We can easily show that

$$\begin{aligned} d_{\mathrm{K}}\left(\frac{\psi(J_n)-\mu_n}{\sigma_n},\mathcal{N}\right) &\le d_{TV}\left(J_n, H_n Q(n/H_n)\right) \\ &+ d_{TV}\left(\psi(H_n Q(n/H_n)), \psi(H_n)+\psi(Q(n/H_n))\right) \\ &+ d_{\mathrm{K}}\left(\frac{\psi(H_n)+\psi(Q(n/H_n)-\mu_n}{\sigma_n}, \frac{\psi(H_n)-\mu_n}{\sigma_n}\right) \\ &+ d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n},\mathcal{N}\right). \end{aligned}$$

Simplifying $\omega(J_n)$ to $\omega(H_n)$

We can easily show that

$$\begin{aligned} d_{\mathrm{K}}\left(\frac{\psi(J_n)-\mu_n}{\sigma_n},\mathcal{N}\right) &\leq d_{TV}\left(J_n, H_n Q(n/H_n)\right) \\ &+ d_{TV}\left(\psi(H_n Q(n/H_n)), \psi(H_n)+\psi(Q(n/H_n))\right) \\ &+ d_{\mathrm{K}}\left(\frac{\psi(H_n)+\psi(Q(n/H_n)-\mu_n}{\sigma_n}, \frac{\psi(H_n)-\mu_n}{\sigma_n}\right) \\ &+ d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n},\mathcal{N}\right). \end{aligned}$$

New goal: bound $d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n},\mathcal{N}\right)$

Simplifying $\omega(J_n)$ to $\omega(H_n)$

We can easily show that

$$d_{\mathrm{K}}\left(\frac{\psi(J_n)-\mu_n}{\sigma_n},\mathcal{N}\right) \le d_{TV}\left(J_n, H_nQ(n/H_n)\right)$$
$$+ d_{TV}\left(\psi(H_nQ(n/H_n)), \psi(H_n)+\psi(Q(n/H_n))\right)$$
$$+ d_{\mathrm{K}}\left(\frac{\psi(H_n)+\psi(Q(n/H_n)-\mu_n}{\sigma_n}, \frac{\psi(H_n)-\mu_n}{\sigma_n}\right)$$
$$+ d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n},\mathcal{N}\right).$$

New goal: bound $d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n},\mathcal{N}\right)$

Methodology used

Since $\psi(H_n)$ is conditionally equal to $\sum_{p\in\mathcal{P}_n}\psi(p^{\xi_p})$, we use *Stein's method.*

Stein's method

Lemma (Stein's lemma)
For every smooth $f : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}[f'(\mathcal{N})] = \mathbb{E}[\mathcal{N} f(\mathcal{N})]$$

Stein's method

Lemma (Stein's lemma)
For every smooth $f : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}[f'(\mathcal{N})] = \mathbb{E}[\mathcal{N} f(\mathcal{N})]$$

Stein's heuristics: if X is an $\mathbb{R}$-valued random variable such that

$$\mathbb{E}[f'(X)] \approx \mathbb{E}[X f(X)],$$

for a large class of functions f, then Z is close to $\mathcal{N}$ in some meaningul sense.

Stein's method

Lemma
Let $h_r : \mathbb{R} \to \mathbb{R}$ be given by $h_r(x) := \mathbb{1}_{(-\infty,r]}(x)$, for some $r \in \mathbb{R}$. Then, the equation

$$f'(x) - xf(x) = h_r(x) - \mathbb{E}[h_r(\mathcal{N})]$$

has a unique solution $f = f_r$, satisfying

$$\sup_{w \in \mathbb{R}} |f_r'(w)| \leq 2 \qquad \text{and} \qquad f_r(w) \leq \sqrt{\pi/2} \tag{12}$$

Lemma
Let $h_r : \mathbb{R} \to \mathbb{R}$ be given by $h_r(x) := \mathbb{1}_{(-\infty, r]}(x)$, for some $r \in \mathbb{R}$. Then, the equation

$$f'(x) - xf(x) = h_r(x) - \mathbb{E}[h_r(\mathcal{N})]$$

has a unique solution $f = f_r$, satisfying

$$\sup_{w \in \mathbb{R}} |f_r'(w)| \leq 2 \qquad \text{and} \qquad f_r(w) \leq \sqrt{\pi/2} \tag{12}$$

Thus, if X is some r.v.

$$d_K(X, \mathcal{N}) \leq \sup_f |\mathbb{E}[f'(X) - Xf(X)]|$$

where f ranges over the functions satisfying (12)

Stein's method for $\psi(H_n)$

As before, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r is Stein's solution and $Y_p := \psi(p^{\xi_p})$.

Stein's method for $\psi(H_n)$

As before, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r is Stein's solution and $Y_p := \psi(p^{\xi_p})$. By the identity

$$\mathcal{L}(\psi(H(n))) = \mathcal{L}(\sum_{p\in\mathcal{P}_n} Y_p | \prod_{p\in\mathcal{P}_n} p^{\xi_p} \leq n),$$

Stein's method for $\psi(H_n)$

As before, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r is Stein's solution and $Y_p := \psi(p^{\xi_p})$. By the identity

$$\mathcal{L}(\psi(H(n))) = \mathcal{L}(\sum_{p\in\mathcal{P}_n} Y_p | \prod_{p\in\mathcal{P}_n} p^{\xi_p} \leq n),$$

we have,

$$\mathbb{E}\left[h_r\left(\frac{\psi(H_n)-\mu_n}{\sigma_n}\right) - \mathbb{E}[h_r(\mathcal{N})]\right] = \frac{\mathbb{E}[(f_r'(W) - Wf_r(W))I]}{\mathbb{P}[\prod_{p\in\mathcal{P}_n} p^{\xi_p} \leq n]},$$

where

$$W = W_n := \sigma_n^{-1}(\sum_{p\in\mathcal{P}_n} \psi(p^{\xi_p}) - \mu_n)$$

$$I = I_n := \mathbb{1}_{\{\prod_{p\in\mathcal{P}_n} p^{\xi_p} \leq n\}}.$$

Stein's method for $\psi(H_n)$

As before, $h_r = \mathbb{1}_{\{(-\infty, r]\}}$, f_r is Stein's solution and $Y_p := \psi(p^{\xi_p})$. By the identity

$$\mathcal{L}(\psi(H(n))) = \mathcal{L}(\sum_{p \in \mathcal{P}_n} Y_p | \prod_{p \in \mathcal{P}_n} p^{\xi_p} \leq n),$$

we have,

$$\mathbb{E}\left[h_r\left(\frac{\psi(H_n) - \mu_n}{\sigma_n}\right) - \mathbb{E}[h_r(\mathcal{N})]\right] = \frac{\mathbb{E}[(f_r'(W) - Wf_r(W))I]}{\mathbb{P}[\prod_{p \in \mathcal{P}_n} p^{\xi_p} \leq n]},$$

where

$$W = W_n := \sigma_n^{-1}(\sum_{p \in \mathcal{P}_n} \psi(p^{\xi_p}) - \mu_n)$$

$$I = I_n := \mathbb{1}_{\{\prod_{p \in \mathcal{P}_n} p^{\xi_p} \leq n\}}.$$

New goal: estimate

$$\mathbb{E}[(f_r'(W) - Wf_r(W))I].$$

Bounding $\mathbb{E}[(f_r'(W) - Wf_r(W))I]$

Let $\{\xi_p'\}_{p\in\mathcal{P}}$ be an independent copy of $\{\xi_p\}_{p\in\mathcal{P}}$, and Θ a random variable uniformly distributed over $\mathcal{P}_n$ and independent of $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$.

Bounding $\mathbb{E}[(f_r'(W) - Wf_r(W))I]$

Let $\{\xi_p'\}_{p\in\mathcal{P}}$ be an independent copy of $\{\xi_p\}_{p\in\mathcal{P}}$, and Θ a random variable uniformly distributed over $\mathcal{P}_n$ and independent of $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$. For each $n \in \mathbb{N}$, set

$$W' = \sigma_n^{-1}(\psi(\Theta^{\xi_\Theta'}) + \sum_{p\in\mathcal{P}_n\setminus\{\Theta\}} \psi(p^{\xi_p}) - \mu_n)$$

$$I' = \mathbb{1}_{\{\theta^{\xi_\theta'} \prod_{p\in\mathcal{P}_n\setminus\{\theta\}} p^{\xi_p} \le n\}}.$$

Then $((W, I), (W', I')) \stackrel{Law}{=} ((W', I'), (W, I))$.

Bounding $\mathbb{E}[(f_r'(W) - Wf_r(W))I]$

Let $\{\xi_p'\}_{p\in\mathcal{P}}$ be an independent copy of $\{\xi_p\}_{p\in\mathcal{P}}$, and Θ a random variable uniformly distributed over $\mathcal{P}_n$ and independent of $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$. For each $n \in \mathbb{N}$, set

$$W' = \sigma_n^{-1}(\psi(\Theta^{\xi_\Theta'}) + \sum_{p\in\mathcal{P}_n\setminus\{\Theta\}} \psi(p^{\xi_p}) - \mu_n)$$

$$I' = \mathbb{1}_{\{\theta^{\xi_\theta'} \prod_{p\in\mathcal{P}_n\setminus\{\theta\}} p^{\xi_p} \leq n\}}.$$

Then $((W, I), (W', I')) \stackrel{Law}{=} ((W', I'), (W, I))$. By exchangeability,

$$\mathbb{E}[(W' - W)(f_r(W)I - f_r(W')I')] = 0.$$

Bounding $\mathbb{E}[(f_r'(W) - Wf_r(W))I]$

Let $\{\xi_p'\}_{p\in\mathcal{P}}$ be an independent copy of $\{\xi_p\}_{p\in\mathcal{P}}$, and Θ a random variable uniformly distributed over $\mathcal{P}_n$ and independent of $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$. For each $n \in \mathbb{N}$, set

$$W' = \sigma_n^{-1}(\psi(\Theta^{\xi_\Theta'}) + \sum_{p\in\mathcal{P}_n\setminus\{\Theta\}} \psi(p^{\xi_p}) - \mu_n)$$

$$I' = \mathbb{1}_{\{\theta^{\xi_\theta'} \prod_{p\in\mathcal{P}_n\setminus\{\theta\}} p^{\xi_p} \leq n\}}.$$

Then $((W, I), (W', I')) \overset{Law}{=} ((W', I'), (W, I))$. By exchangeability,

$$\mathbb{E}[(W' - W)(f_r(W)I - f_r(W')I')] = 0.$$

so

$$-2\mathbb{E}[(W' - W)f_r(W)I] = \mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)].$$

Handling $-2\mathbb{E}[(W' - W)f_r(W)I]$

We observe that $LHS := -2\mathbb{E}[(W' - W)f_r(W)I]$ satisfies

$$\begin{aligned} LHS &= -\frac{2}{\pi(n)}\mathbb{E}\left[\frac{(\sum_{\theta\in\mathcal{P}_n} Y'_\theta - \mu_n) - (\sum_{\theta\in\mathcal{P}_n} Y_\theta - \mu_n)}{\sigma_n} f_r(W)I\right] \\ &= \frac{2}{\pi(n)}\mathbb{E}[Wf_r(W)I] - \frac{2}{\pi(n)}\mathbb{E}[W]\mathbb{E}[f_r(W)I], \end{aligned}$$

Handling $-2\mathbb{E}[(W' - W)f_r(W)I]$

We observe that $LHS := -2\mathbb{E}[(W' - W)f_r(W)I]$ satisfies

$$
\begin{aligned}
LHS &= -\frac{2}{\pi(n)}\mathbb{E}\left[\frac{(\sum_{\theta\in\mathcal{P}_n} Y'_\theta - \mu_n) - (\sum_{\theta\in\mathcal{P}_n} Y_\theta - \mu_n)}{\sigma_n} f_r(W)I\right] \\
&= \frac{2}{\pi(n)}\mathbb{E}[Wf_r(W)I] - \frac{2}{\pi(n)}\mathbb{E}[W]\mathbb{E}[f_r(W)I],
\end{aligned}
$$

so

$$
LHS = \frac{2}{\pi(n)}\mathbb{E}[Wf_r(W)I],
$$

Handling $\mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)]$

Define $X_p := \sigma_n^{-1} Y_p$ and

$$RHS := \mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)],$$

Handling $\mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)]$

Define $X_p := \sigma_n^{-1} Y_p$ and

$$RHS := \mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)],$$

To estimate RHS we formalize the approximation

$$\begin{aligned} RHS &\approx \frac{1}{\pi(n)} \sum_{p \in \mathcal{P}_n} \mathbb{E}[(X'_p - X_p)^2 f'_r(W)I] \\ &\approx \frac{1}{\pi(n)} \sum_{p \in \mathcal{P}_n} \mathbb{E}[(X'_p - X_p)^2]\mathbb{E}[f'_r(W)I] \\ &= \frac{2\mathbb{V}\mathrm{ar}(W)}{\pi(n)} \mathbb{E}[f'_r(W)I], \end{aligned}$$

Handling $\mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)]$

Define $X_p := \sigma_n^{-1} Y_p$ and

$$RHS := \mathbb{E}[(W' - W)(f_r(W')I' - f_r(W)I)],$$

To estimate RHS we formalize the approximation

$$\begin{aligned} RHS &\approx \frac{1}{\pi(n)} \sum_{p \in \mathcal{P}_n} \mathbb{E}[(X_p' - X_p)^2 f_r'(W)I] \\ &\approx \frac{1}{\pi(n)} \sum_{p \in \mathcal{P}_n} \mathbb{E}[(X_p' - X_p)^2]\mathbb{E}[f_r'(W)I] \\ &= \frac{2\mathbb{V}\mathrm{ar}(W)}{\pi(n)} \mathbb{E}[f_r'(W)I], \end{aligned}$$

to obtain

$$RHS \approx \frac{2}{\pi(n)} \mathbb{E}[f_r'(W)I],$$

Conclusion

We conclude that

$$0 = |RHS - LHS| \approx |\frac{2}{\pi(n)}(\mathbb{E}[Wf_r(W)I] - \mathbb{E}[f_r'(W)I])|.$$

Thus, the result follows by a careful analysis of the approximations.

Poisson case

Theorem (Chen, Jaramillo and Yang)

Let M_n be a Poisson distribution with parameter $\log\log(n)$ *and define* $\Omega : \mathbb{N} \to \mathbb{N}$ *by* $\Omega(m) := \sup_{p \in \mathcal{P}_n} \alpha_p(m)$.

Poisson case

Theorem (Chen, Jaramillo and Yang)

Let M_n be a Poisson distribution with parameter $\log\log(n)$ *and define* $\Omega : \mathbb{N} \to \mathbb{N}$ *by* $\Omega(m) := \sup_{p \in \mathcal{P}_n} \alpha_p(m)$*. Then we have*

$$d_{\mathrm{TV}}(\omega(J_n), M_n) \le \frac{7.2}{\sqrt{\log\log(n)}} + 67.4 \frac{\log\log(n)}{\log(n)}$$
$$d_{\mathrm{TV}}(\Omega(J_n), M_n) \le \frac{14}{\sqrt{\log\log(n)}}.$$

References

- Chen L., Jaramillo A., Yang X. A probabilistic approach to the Erdös-Kac theorem for additive functions. Soon in Arxiv.
- R. Arratia. On the amount of dependence in the prime factorization of a uniform random integer. In Contemporary combinatorics, volume 10 of Bolyai Soc. Math. Stud., pages 29–91. János Bolyai Math. Soc., Budapest, 2002.
- A. D. Barbour, E. Kowalski, and A. Nikeghbali. Mod-discrete expansions. Probab. Theory Related Fields, 158(3-4):859–893, 2014.
- Adam J. Harper. Two new proofs of the Erdös-Kac theorem, with bound on the rate of convergence, by Stein's method for distributional approximations. Math. Proc. Cambridge Philos. Soc., 147(1):95–114, 2009.