Teoremas límite, método de Stein y teoría de números.

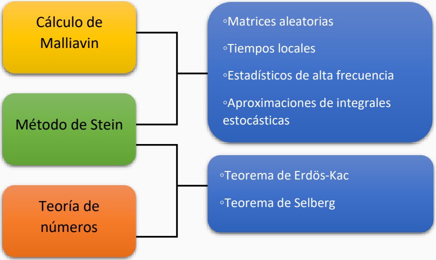
Trabajo conjunto con X. Yang y L. Chen

Arturo Jaramillo Gil

Université du Luxembourg
National University of Singapore

Temas relacionados

A continuación se presentan algunos temas relacionados con teoremas límite



Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos.

Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos. Sea $\omega : \mathbb{N} \to \mathbb{N}$ la función

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

Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos. Sea $\omega : \mathbb{N} \to \mathbb{N}$ la función

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

Por ejemplo, $\omega(54) = \omega(2 \times 3^2) = 2$.

Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos. Sea $\omega : \mathbb{N} \to \mathbb{N}$ la función

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

Por ejemplo, $\omega(54) = \omega(2 \times 3^2) = 2$. Sea J_n una variable aleatoria con distribución uniforme en $\{1, \ldots, n\}$.

Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos. Sea $\omega : \mathbb{N} \to \mathbb{N}$ la función

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

Por ejemplo, $\omega(54) = \omega(2 \times 3^2) = 2$. Sea J_n una variable aleatoria con distribución uniforme en $\{1, \ldots, n\}$.

Objetivo

- Estudiar $\omega(J_n)$.

Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos. Sea $\omega : \mathbb{N} \to \mathbb{N}$ la función

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

Por ejemplo, $\omega(54) = \omega(2 \times 3^2) = 2$. Sea J_n una variable aleatoria con distribución uniforme en $\{1, \ldots, n\}$.

Objetivo

- Estudiar $\omega(J_n)$.
  Describir con la mayor precisión posible el comportamiento asintótico de $\frac{\omega(J_n)-\mu_n}{\sigma_n}$, para μ_n y σ_n apropiados.

Objetivo de ésta plática

Denotemos por $\mathcal{P}$ al conjunto de números primos. Sea $\omega : \mathbb{N} \to \mathbb{N}$ la función

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

Por ejemplo, $\omega(54) = \omega(2 \times 3^2) = 2$. Sea J_n una variable aleatoria con distribución uniforme en $\{1, \ldots, n\}$.

Objetivo

- Estudiar $\omega(J_n)$.
  Describir con la mayor precisión posible el comportamiento asintótico de $\frac{\omega(J_n)-\mu_n}{\sigma_n}$, para μ_n y σ_n apropiados.
- ¿Qué se puede decir en el caso en que ω es reemplazada por una función general $\psi : \mathbb{N} \to \mathbb{N}$ que únicamente satisfaga $\psi(ab) = \psi(a) + \psi(b)$ para $a, b \in \mathbb{N}$ coprimos?

Plan

1. Contexto histórico
2. Resultados principales
3. Ideas de las pruebas
   - Simplificación del modelo
   - Método de Stein

Contexto histórico

Teorema de Erdös-Kac clásico (1940)

Punto de inicio: Paul Erdös y Mark Kac probaron que

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

converge en distribución hacia una variable gaussiana estándar $\mathcal{N}$.

Teorema de Erdös-Kac clásico (1940)

Punto de inicio: Paul Erdös y Mark Kac probaron que

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

converge en distribución hacia una variable gaussiana estándar $\mathcal{N}$.

Intuición: Definamos $\mathcal{P}_n := \mathcal{P} \cap [1, n]$.

Teorema de Erdös-Kac clásico (1940)

Punto de inicio: Paul Erdös y Mark Kac probaron que

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

converge en distribución hacia una variable gaussiana estándar $\mathcal{N}$.

Intuición: Definamos $\mathcal{P}_n := \mathcal{P} \cap [1, n]$. La convergencia en (1) se intuye de la descomposición

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

Intuición en el teorema de Erdös-Kac

Podemos conjecturar que $\mathbb{1}_{\{p \text{ divide } J_n\}}$ son débilmente dependientes para $d \in \mathbb{N}$, ya que

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

Intuición en el teorema de Erdös-Kac

Podemos conjecturar que $\mathbb{1}_{\{p \text{ divide } J_n\}}$ son débilmente dependientes para $d \in \mathbb{N}$, ya que

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

Por ende, si $p_1, \ldots, p_r \in \mathcal{P}_n$ son primos distintos,

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

Intuición en el teorema de Erdös-Kac

Podemos conjecturar que $\mathbb{1}_{\{p \text{ divide } J_n\}}$ son débilmente dependientes para $d \in \mathbb{N}$, ya que

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

Por ende, si $p_1, \ldots, p_r \in \mathcal{P}_n$ son primos distintos,

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

Nota: actualmente se conoce que las variables $\mathbb{1}_{\{p \text{ divide } J_n\}}$, con $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ son aproximadamente independientes si $\alpha_n \to \infty$ es escogido adecuadamente (por ejemplo: $\alpha_n := 3 \log \log(n)^2$).

Intuición en el teorema de Erdös-Kac

Podemos conjecturar que $\mathbb{1}_{\{p \text{ divide } J_n\}}$ son débilmente dependientes para $d \in \mathbb{N}$, ya que

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

Por ende, si $p_1, \ldots, p_r \in \mathcal{P}_n$ son primos distintos,

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

Nota: actualmente se conoce que las variables $\mathbb{1}_{\{p \text{ divide } J_n\}}$, con $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ son aproximadamente independientes si $\alpha_n \to \infty$ es escogido adecuadamente (por ejemplo: $\alpha_n := 3 \log\log(n)^2$). *¡Sin embargo α_n no puede ser igual a uno!*

Pregunta

¿Podemos estimar el error de la aproximación gaussiana con respecto a una métrica adecuada? como la distancia definida por

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

Pregunta

¿Podemos estimar el error de la aproximación gaussiana con respecto a una métrica adecuada? como la distancia definida por

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

o

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

donde Lip_1 es la familia de funciones Lipschitz con constante de Lipschitz menor o igual a uno.

Pregunta

¿Podemos estimar el error de la aproximación gaussiana con respecto a una métrica adecuada? como la distancia definida por

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

o

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

donde Lip_1 es la familia de funciones Lipschitz con constante de Lipschitz menor o igual a uno. Definimos adicionalmente

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

Conjetura de LeVeque (1949)

LeVeque mostró que

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

para una constante $C > 0$ independiente de n.

Conjetura de LeVeque (1949)

LeVeque mostró que

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

para una constante $C > 0$ independiente de n. También conjeturó que

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

Conjetura de LeVeque (1949)

LeVeque mostró que

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

para una constante $C > 0$ independiente de n. También conjeturó que

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

Ésto fue demostrado posteriormente por Rényi y Turán (1958).

Conjetura de LeVeque (1949)

LeVeque mostró que

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

para una constante $C > 0$ independiente de n. También conjeturó que

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

Ésto fue demostrado posteriormente por Rényi y Turán (1958). La idea central consiste en aproximar $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$.

Conjetura de LeVeque (1949)

LeVeque mostró que

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

para una constante $C > 0$ independiente de n. También conjeturó que

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

Ésto fue demostrado posteriormente por Rényi y Turán (1958). La idea central consiste en aproximar $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$.

Ingredientes principales: fórmula de Perron, series de Dirichlet y algunas estimaciones de la función zeta de Riemann ζ alrededor de la banda $\{z \in \mathbb{C} \ ; \ \Re(z) = 1\}$.

Convergencia mod-ϕ (2014)

Actualmente, se conoce mucho sobre $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$.

Convergencia mod-ϕ (2014)

Actualmente, se conoce mucho sobre $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$. Ello motivó el estudio de las propiedades asintóticas de $\omega(J_n)$ mediante un análisis basado en una aproximación del tipo

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

donde M_n es una variable aleatoria Poisson con parámetro $\log\log(n)$ y $F(\lambda)$ es una función posiblemente no-trivial (ver el trabajo de Barbour, Kowalski y Nikeghbali in 2014).

Convergencia mod-ϕ (2014)

Actualmente, se conoce mucho sobre $\mathbb{E}[e^{\mathbf{i}\lambda\omega(J_n)}]$. Ello motivó el estudio de las propiedades asintóticas de $\omega(J_n)$ mediante un análisis basado en una aproximación del tipo

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

donde M_n es una variable aleatoria Poisson con parámetro $\log\log(n)$ y $F(\lambda)$ es una función posiblemente no-trivial (ver el trabajo de Barbour, Kowalski y Nikeghbali in 2014). Concluimos así,

Theorem
Existe una constante $C > 0$ tal que

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

Otras perspectivas (método de Stein)

Recordemos la heurística de que $\mathbb{1}_{\{p \text{ divide } J_n\}}$, para $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ son aproximadamente independientes.

Otras perspectivas (método de Stein)

Recordemos la heurística de que $\mathbb{1}_{\{p \text{ divide } J_n\}}$, para $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ son aproximadamente independientes.

Perspectiva de método de Stein

Variables con dependencias débiles son el objeto central de la teoría de método de Stein.

Otras perspectivas (método de Stein)

Recordemos la heurística de que $\mathbb{1}_{\{p \text{ divide } J_n\}}$, para $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ son aproximadamente independientes.

Perspectiva de método de Stein
Variables con dependencias débiles son el objeto central de la teoría de método de Stein.

Adam Harper (2009) usó dichas herramientas para mostrar que

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

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

Otras perspectivas (método de Stein)

Recordemos la heurística de que $\mathbb{1}_{\{p \text{ divide } J_n\}}$, para $p \in \mathcal{P} \cap [1, n^{\frac{1}{\alpha_n}}]$ son aproximadamente independientes.

Perspectiva de método de Stein
Variables con dependencias débiles son el objeto central de la teoría de método de Stein.

Adam Harper (2009) usó dichas herramientas para mostrar que

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

donde $\alpha_n := 3 \log\log(n)^2$. Consecuentemente,

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

Otras perspectivas (aproximación con variables independientes)

Otra idea consiste en comparar la ley de $(\mathbb{1}_{\{p \text{ divide } J_n\}} \ ; \ p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}})$ con variables aleatorias independientes.

Otras perspectivas (aproximación con variables independientes)

Otra idea consiste en comparar la ley de $(\mathbb{1}_{\{p \text{ divide } J_n\}} \; ; \; p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}})$ con variables aleatorias independientes.

Kubilius (1964) mostró que si $\beta_n \to \infty$, entonces

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

donde B_p son variables Bernoulli independientes, con $\mathbb{P}[B_p] = 1/p$.

Otras perspectivas (aproximación con variables independientes)

Otra idea consiste en comparar la ley de $(\mathbb{1}_{\{p \text{ divide } J_n\}} \ ; \ p \in \mathcal{P} \cap [1, n^{\frac{1}{\beta_n}})$ con variables aleatorias independientes.

Kubilius (1964) mostró que si $\beta_n \to \infty$, entonces

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

donde B_p son variables Bernoulli independientes, con $\mathbb{P}[B_p] = 1/p$. Por lo tanto,

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

Como consecuencia, obtenemos un resultado similar al de Harper.

Perspectiva de permutaciones sesgadas

Arratia (2013) sugiere comparar J_n con el producto de un primo aleatorio Q_n y un producto parcial T_n de permutaciones sesgadas de los factores de J_n. Probó que

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

Perspectiva de permutaciones sesgadas

Arratia (2013) sugiere comparar J_n con el producto de un primo aleatorio Q_n y un producto parcial T_n de permutaciones sesgadas de los factores de J_n. Probó que

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

Ello se utilizó para mostrar que si $d_\Omega : \mathbb{N}^2 \to \mathbb{N}$ denota la distancia de inserción-supresión $d_\Omega(\prod_{p\in\mathcal{P}} p^{\alpha_p}, \prod_{p\in\mathcal{P}} p^{\beta_p}) := \sum_{p\in\mathcal{P}} |\alpha_p - \beta_p|$, y $d_{1,\Omega}$ entonces la distancia de Wasserstein asociada satisface

Perspectiva de permutaciones sesgadas

Arratia (2013) sugiere comparar J_n con el producto de un primo aleatorio Q_n y un producto parcial T_n de permutaciones sesgadas de los factores de J_n. Probó que

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

Ello se utilizó para mostrar que si $d_\Omega : \mathbb{N}^2 \to \mathbb{N}$ denota la distancia de inserción-supresión $d_\Omega(\prod_{p\in\mathcal{P}} p^{\alpha_p}, \prod_{p\in\mathcal{P}} p^{\beta_p}) := \sum_{p\in\mathcal{P}} |\alpha_p - \beta_p|$, y $d_{1,\Omega}$ entonces la distancia de Wasserstein asociada satisface

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

donde

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

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

Resultados principales

Teorema del límite central para funciones aditivas

Sea $\psi : \mathbb{N} \to \mathbb{N}$ tal que $\psi(ab) = \psi(a) + \psi(b)$ para a, b co-primos.

(H1) Tenemos que

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

(H2) Existe una función (posiblemente no acotada) $\Psi : \mathcal{P} \to \mathbb{R}_+$ que satisface

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

y tal que para todo $p \in \mathcal{P}_n$,

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

Resultado principal para la distancia de Kolmogorov

Sean μ_n y $\sigma_n > 0$ dados por

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

Resultado principal para la distancia de Kolmogorov

Sean μ_n y $\sigma_n > 0$ dados por

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

Theorem (Chen, Jaramillo, Yang)

*Supongamos que ψ satisface **(H1)** y **(H2)**. Entonces, si $X_p := \sigma_n^{-1}\psi(p^{\xi_p})$ y $\sigma_n^2 \geq 3(\|\psi\|_{\mathcal{P}}^2 + \|\Psi\|_{\mathcal{P}}^2)$,*

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

donde

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

Main result for Wasserstein distance

Theorem (Chen, Jaramillo, Yang)

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

donde

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

Ideas de las pruebas

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Para todo $i \in \mathbb{N}$, $k_1, ..., k_i \in \mathbb{N}_0$ y $p_1, \ldots, p_i \in \mathcal{P}$ primos distintos,

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

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Para todo $i \in \mathbb{N}$, $k_1, ..., k_i \in \mathbb{N}_0$ y $p_1, \ldots, p_i \in \mathcal{P}$ primos distintos,

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

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

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Para todo $i \in \mathbb{N}$, $k_1, ..., k_i \in \mathbb{N}_0$ y $p_1, \ldots, p_i \in \mathcal{P}$ primos distintos,

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

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

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

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Para todo $i \in \mathbb{N}$, $k_1, ..., k_i \in \mathbb{N}_0$ y $p_1, \ldots, p_i \in \mathcal{P}$ primos distintos,

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

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

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

Pregunta: ¿Podemos utilizar las variables ξ_p para reconstruir por completo la variable J_n?

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Para todo $i \in \mathbb{N}$, $k_1, ..., k_i \in \mathbb{N}_0$ y $p_1, \ldots, p_i \in \mathcal{P}$ primos distintos,

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

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

$$(\alpha_{p_1}(J_n), \ldots, \alpha_{p_i}(J_n)) \stackrel{Ley}{\to} (\xi_{p_1}, \ldots, \xi_{p_i})$$

Pregunta: ¿Podemos utilizar las variables ξ_p para reconstruir por completo la variable J_n?

Respuesta: no fácilmente

Multiplicidades de los factores primos

Para $p \in \mathcal{P}$ dado, definimos $\alpha_p : \mathbb{N} \to \mathbb{N}$ por

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

Para todo $i \in \mathbb{N}$, $k_1, ..., k_i \in \mathbb{N}_0$ y $p_1, \ldots, p_i \in \mathcal{P}$ primos distintos,

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

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

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

Pregunta: ¿Podemos utilizar las variables ξ_p para reconstruir por completo la variable J_n?
Respuesta: no fácilmente... pero...

Modelo simplificado: la distribución harmónica H_n

Sea H_n una v.a. con $\mathbb{P}[H_n = k] = \frac{1}{L_n k}$ para $k \leq n$, donde $L_n := \sum_{k=1}^{n} \frac{1}{k}$.

Modelo simplificado: la distribución harmónica H_n

Sea H_n una v.a. con $\mathbb{P}[H_n = k] = \frac{1}{L_n k}$ para $k \leq n$, donde $L_n := \sum_{k=1}^{n} \frac{1}{k}$. Entonces,

Proposition

Supongamos que $n \geq 21$. Definamos el evento

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

y el vector aleatorio $\vec{C}(n) := (\alpha_p(H_n); p \in \mathcal{P}_n)$. Entonces las variables $Y_p := \psi(p^{\xi_p})$, indexadas por $p \in \mathcal{P}_n$, satisfacen

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

Relación con la distribución harmónica

Sea $\{Q(k)\}_{k\geq 1}$ una sucesión de variables aleatorias independientes, e independientes de (J_n, H_n), donde $Q(k)$ se distribuye uniforme sobre

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

Relación con la distribución harmónica

Sea $\{Q(k)\}_{k\geq 1}$ una sucesión de variables aleatorias independientes, e independientes de (J_n, H_n), donde $Q(k)$ se distribuye uniforme sobre

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

Sea $\pi(n) := |\mathcal{P} \cap [1, n]|$. Usando el hecho de que para $n \geq 229$,

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

Relación con la distribución harmónica

Sea $\{Q(k)\}_{k\geq 1}$ una sucesión de variables aleatorias independientes, e independientes de (J_n, H_n), donde $Q(k)$ se distribuye uniforme sobre

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

Sea $\pi(n) := |\mathcal{P} \cap [1, n]|$. Usando el hecho de que para $n \geq 229$,

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

Lemma (Chen, Jaramillo y Yang)

La siguiente cota se cumple para $n \geq 21$

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

Relación entre $\psi(J_n)$ y $\psi(H_n)$

Usando el hecho de que $|\psi(H_n Q(n/H_n)) - \psi(H_n)| \leq \|\psi\|_{\mathcal{P}}$,

Relación entre $\psi(J_n)$ y $\psi(H_n)$

Usando el hecho de que $|\psi(H_n Q(n/H_n)) - \psi(H_n)| \leq \|\psi\|_{\mathcal{P}}$, podemos probar fácilmente que

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

Relación entre $\psi(J_n)$ y $\psi(H_n)$

Usando el hecho de que $|\psi(H_n Q(n/H_n)) - \psi(H_n)| \leq \|\psi\|_{\mathcal{P}}$, podemos probar fácilmente que

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

Nuevo objetivo: acotar $d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n}, \mathcal{N}\right)$

Relación entre $\psi(J_n)$ y $\psi(H_n)$

Usando el hecho de que $|\psi(H_nQ(n/H_n)) - \psi(H_n)| \le \|\psi\|_{\mathcal{P}}$, podemos probar fácilmente que

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

Nuevo objetivo: acotar $d_{\mathrm{K}}\left(\frac{\psi(H_n)-\mu_n}{\sigma_n}, \mathcal{N}\right)$

Metodología utilizada

Ya que $\psi(H_n)$ es condicionalmente igual en ley a $\sum_{p\in\mathcal{P}_n}\psi(p^{\xi_p})$, utilizaremos *método de Stein.*

Método de Stein

Lemma (Lema de Stein)
Para toda función suave $f : \mathbb{R} \to \mathbb{R}$,

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

Método de Stein

Lemma (Lema de Stein)
Para toda función suave $f : \mathbb{R} \to \mathbb{R}$,

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

Heurística de Stein: si X es una variable $\mathbb{R}$-valuada, tal que

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

para una clase suficientemente extensa de funciones f, entonces Z se aproxima a $\mathcal{N}$.

Método de Stein

Lemma
Sea $h_r : \mathbb{R} \to \mathbb{R}$ *la función dada por* $h_r(x) := \mathbb{1}_{(-\infty,r]}(x)$*, para* $r \in \mathbb{R}$*. Entonces, la ecuación*

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

tiene una solución única $f = f_r$*, que satisface*

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

Lemma
Sea $h_r : \mathbb{R} \to \mathbb{R}$ la función dada por $h_r(x) := \mathbb{1}_{(-\infty,r]}(x)$, para $r \in \mathbb{R}$. Entonces, la ecuación

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

tiene una solución única $f = f_r$, que satisface

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

Por lo tanto, si X es una variable aleatoria,

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

donde f pertenece a la familia de funciones que satisfacen (12)

Método de Stein para $\psi(H_n)$

Como antes, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r es la solución a la ecuación de Stein asociada y $Y_p := \psi(p^{\xi_p})$.

Método de Stein para $\psi(H_n)$

Como antes, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r es la solución a la ecuación de Stein asociada y $Y_p := \psi(p^{\xi_p})$. Gracias a la identidad

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

Como antes, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r es la solución a la ecuación de Stein asociada y $Y_p := \psi(p^{\xi_p})$. Gracias a la identidad

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

se tiene que,

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

donde

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

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

Método de Stein para $\psi(H_n)$

Como antes, $h_r = \mathbb{1}_{\{(-\infty,r]\}}$, f_r es la solución a la ecuación de Stein asociada y $Y_p := \psi(p^{\xi_p})$. Gracias a la identidad

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

se tiene que,

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

donde

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

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

Nuevo objetivo: estimar

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

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

Sea $\{\xi_p'\}_{p\in\mathcal{P}}$ una copia independiente de $\{\xi_p\}_{p\in\mathcal{P}}$, y Θ una variable aleatoria con distribución uniforme sobre $\mathcal{P}_n$ e independiente de $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$.

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

Sea $\{\xi_p'\}_{p\in\mathcal{P}}$ una copia independiente de $\{\xi_p\}_{p\in\mathcal{P}}$, y Θ una variable aleatoria con distribución uniforme sobre $\mathcal{P}_n$ e independiente de $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$. Para cada $n \in \mathbb{N}$, definimos

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

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

Entonces $((W, I), (W', I')) \overset{Ley}{=} ((W', I'), (W, I))$.

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

Sea $\{\xi_p'\}_{p\in\mathcal{P}}$ una copia independiente de $\{\xi_p\}_{p\in\mathcal{P}}$, y Θ una variable aleatoria con distribución uniforme sobre $\mathcal{P}_n$ e independiente de $\{(\xi_p', \xi_p)\}_{p\in\mathcal{P}}$. Para cada $n \in \mathbb{N}$, definimos

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

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

Entonces $((W, I), (W', I')) \overset{Ley}{=} ((W', I'), (W, I))$. Por intercambiabilidad,

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

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

Se tiene que $LHS := -2\mathbb{E}[(W' - W)f_r(W)I]$ satisface

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

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

Se tiene que $LHS := -2\mathbb{E}[(W' - W)f_r(W)I]$ satisface

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

de manera que

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

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

Definamos $X_p := \sigma_n^{-1} Y_p$ y

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

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

Definamos $X_p := \sigma_n^{-1} Y_p$ y

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

Para estimar RHS formalizamos la aproximación

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

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

Definamos $X_p := \sigma_n^{-1} Y_p$ y

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

Para estimar RHS formalizamos la aproximación

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

para obtener

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

Conclusión

Concluimos que

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

Por lo tanto, el resultado se sigue de un análisis adecuado de las aproximaciones.

Caso Poisson

Theorem (Chen, Jaramillo y Yang)

Sea M_n una variable aleatoria Poisson con parámetro $\log\log(n)$ *y definamos* $\Omega : \mathbb{N} \to \mathbb{N}$ *por* $\Omega(m) := \sup_{p \in \mathcal{P}_n} \alpha_p(m)$.

Caso Poisson

Theorem (Chen, Jaramillo y Yang)

Sea M_n una variable aleatoria Poisson con parámetro $\log\log(n)$ y definamos $\Omega : \mathbb{N} \to \mathbb{N}$ por $\Omega(m) := \sup_{p \in \mathcal{P}_n} \alpha_p(m)$. Entonces se tiene que

$$
\begin{aligned}
d_{\mathrm{TV}}(\omega(J_n), M_n) &\leq \frac{7.2}{\sqrt{\log\log(n)}} + 67.4 \frac{\log\log(n)}{\log(n)} \\
d_{\mathrm{TV}}(\Omega(J_n), M_n) &\leq \frac{14}{\sqrt{\log\log(n)}}.
\end{aligned}
$$

Referencias

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.