Mediciones estadísticas de enteros suaves, y su relación con metodo de Stein

Arturo Jaramillo Gil

Centro de Investigación en Matemáticas (CIMAT)

Motivación

Sea $\mathcal{P}$ el conjunto de primos positivos y $[n] := \{1, \ldots, n\}$.

Motivación

Sea $\mathcal{P}$ el conjunto de primos positivos y $[n] := \{1, \dots, n\}$. Sea $\psi : \mathbb{N} \to \mathbb{N}$, la función definida por

$$\psi(k) := \max\{p \in \mathcal{P} \ ; \ p \text{ divide } k\}.$$

Motivación

Sea $\mathcal{P}$ el conjunto de primos positivos y $[n] := \{1, \ldots, n\}$. Sea $\psi : \mathbb{N} \to \mathbb{N}$, la función definida por

$$\psi(k) := \max\{p \in \mathcal{P} \ ; \ p \text{ divide } k\}.$$

Objetivo:

Estimar la magnitud/contenido de

$$\{k \in \mathbb{N} \ ; \ \psi(k) \leq m\}$$

Cantidades de interés

Definition
Decimos que un entero $k \in \mathbb{N}$ es m-suave si $\psi(k) \leq m$. En lo sucesivo, denotaremos por $\mathcal{Z}[n, m]$ a los enteros m-suaves acotados por n.

Definition
Decimos que un entero $k \in \mathbb{N}$ es m-suave si $\psi(k) \leq m$. En lo sucesivo, denotaremos por $\mathcal{Z}[n, m]$ a los enteros m-suaves acotados por n.

La magnitud de $\mathcal{Z}[n, m]$,

$$\Psi(n, m) := |\mathcal{Z}[n, m]|,$$

tiene una importante relación con la hipótesis de Riemann

La aproximación de Dickman

Definition
Sea $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ la solución a

$$u\rho'(u) = -\rho(u-1) \qquad \text{y} \qquad \rho(u) = 1 \ \ \text{para } u \in [0,1].$$

La aproximación de Dickman

Definition
Sea $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ la solución a

$$u\rho'(u) = -\rho(u-1) \qquad \text{y} \qquad \rho(u) = 1 \;\text{ para } u \in [0,1].$$

Theorem
Dickman (1930) muestra que para todo $x \in \mathbb{R}_+$,

$$\lim_{n\to\infty} \frac{1}{n}\Psi(n, n^{1/x}) = \rho(x).$$

La aproximación de Dickman

Definition
Sea $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ la solución a

$$u\rho'(u) = -\rho(u-1) \qquad \text{y} \qquad \rho(u) = 1 \ \ \text{para } u \in [0,1].$$

Theorem
Dickman (1930) muestra que para todo $x \in \mathbb{R}_+$,

$$\lim_{n\to\infty} \frac{1}{n}\Psi(n, n^{1/x}) = \rho(x).$$

Ramaswami (1949) prueba el refinamiento

$$|\frac{1}{n}\Psi(n, n^{1/x}) - \rho(x)| \leq \frac{C_x}{\log(n)}.$$

Uniformidad en la aproximación de Dickman

Definimos, para un nivel de suavidad m, dado,

$$\Upsilon(n, m) := \frac{\log(n)}{\log(m)}.$$

Uniformidad en la aproximación de Dickman

Definimos, para un nivel de suavidad m, dado,

$$\Upsilon(n,m) := \frac{\log(n)}{\log(m)}.$$

Theorem (De Bruijn, 1951)
Para $\varepsilon > 0$, defnimos

$$\mathcal{S}_\varepsilon = \{(n,m) \in \mathbb{N}^2 \ |\Upsilon(n,m) \leq \log(m)^{3/5-\varepsilon}\}.$$

Existe $C_\varepsilon > 0$, tal que

$$\Psi(n,m) = n\rho \circ \Upsilon(n,m)\left(1 + O\left(\frac{\log(\Upsilon(n,m)+1)}{\log(m)}\right)\right)$$

Aproximaciones de Hildebrand

Para $\varepsilon > 0$, definimos

$$\tilde{\mathcal{S}}_\varepsilon = \{(n,m) \in \mathbb{N}^2 \mid m \neq 1 \text{ and } \Upsilon(n,m) \leq \exp\{\log(m)^{3/5-\varepsilon}\}\}$$
$$\hat{\mathcal{S}}_\varepsilon := \{(n,m) \in \mathbb{N}^2 \mid m \neq 1 \text{ and } \Upsilon(n,m) \leq m^{1/2-\varepsilon}\}.$$

Theorem
En 1986, Hildebrand demuestra que la aproximación de De Bruijn se cumple para $(n,m) \in \tilde{\mathcal{S}}_\varepsilon$.

Aproximaciones de Hildebrand

Para $\varepsilon > 0$, definimos

$$\tilde{\mathcal{S}}_\varepsilon = \{(n,m) \in \mathbb{N}^2 \mid m \neq 1 \text{ and } \Upsilon(n,m) \leq \exp\{\log(m)^{3/5-\varepsilon}\}\}$$
$$\hat{\mathcal{S}}_\varepsilon := \{(n,m) \in \mathbb{N}^2 \mid m \neq 1 \text{ and } \Upsilon(n,m) \leq m^{1/2-\varepsilon}\}.$$

Theorem
En 1986, Hildebrand demuestra que la aproximación de De Bruijn se cumple para $(n,m) \in \tilde{\mathcal{S}}_\varepsilon$. Más aún, si la aproximación se cumple para $(n,m) \in \hat{\mathcal{S}}_\varepsilon$, entonces la hipótesis de Riemann hypothesis es cierta.

Otras medidas de contenido para $\mathscr{Z}[n, m]$

La función $\Psi(n, m)$ puede escribirse como

$$\Psi(n, m) = n\mathbb{P}[J_n \in \mathscr{Z}[n, m]],$$

donde J_n son variables con ley uniforme en $[n]$.

Otras medidas de contenido para $\mathcal{Z}[n, m]$

La función $\Psi(n, m)$ puede escribirse como

$$\Psi(n, m) = n\mathbb{P}[J_n \in \mathcal{Z}[n, m]],$$

donde J_n son variables con ley uniforme en $[n]$. Se puede considerar la modificación

$$\Psi_H(n, m) := n\mathbb{P}[H_n \in \mathcal{Z}[n, m]],$$

donde H_n es una variable tal que $\mathbb{P}[H_n = k] = k^{-1}/L_n$ y

$$L_n = \sum_{j=1}^{n} 1/j$$

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Encontrar un principio de transferencia de estimaciones de Ψ_H a Ψ es un problema abierto.

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Encontrar un principio de transferencia de estimaciones de Ψ_H a Ψ es un problema abierto.
Sea $\pi : \mathbb{R}_+ \to \mathbb{R}$ la función cuenta primos

$$\pi(x) := |\mathcal{P} \cap [1, x]|.$$

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Encontrar un principio de transferencia de estimaciones de Ψ_H a Ψ es un problema abierto.
Sea $\pi : \mathbb{R}_+ \to \mathbb{R}$ la función cuenta primos

$$\pi(x) := |\mathcal{P} \cap [1, x]|.$$

Si Q_n es un elemento uniforme de $\mathcal{P} \cup \{1\}$ en $[0, n/H_n]$,

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Encontrar un principio de transferencia de estimaciones de Ψ_H a Ψ es un problema abierto.

Sea $\pi : \mathbb{R}_+ \to \mathbb{R}$ la función cuenta primos

$$\pi(x) := |\mathcal{P} \cap [1, x]|.$$

Si Q_n es un elemento uniforme de $\mathcal{P} \cup \{1\}$ en $[0, n/H_n]$,

Lemma (Chen, Jaramillo, Yang (2021))
Para $n \geq 21$,

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

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Encontrar un principio de transferencia de estimaciones de Ψ_H a Ψ es un problema abierto.
Sea $\pi : \mathbb{R}_+ \to \mathbb{R}$ la función cuenta primos

$$\pi(x) := |\mathcal{P} \cap [1, x]|.$$

Si Q_n es un elemento uniforme de $\mathcal{P} \cup \{1\}$ en $[0, n/H_n]$,

Lemma (Chen, Jaramillo, Yang (2021))
Para $n \geq 21$,

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

Moraleja
Basta analizar $H_n Q_n$.

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Recordemos que Q_n es un uniforme de $\mathcal{P} \cup \{1\}$ sobre $[0, n/H_n]$.

Relación entre $\Psi(n, m)$ y $\Psi_H(n, m)$

Recordemos que Q_n es un uniforme de $\mathcal{P} \cup \{1\}$ sobre $[0, n/H_n]$.
Particionando sobre el evento $H_n \leq n/m$, y haciendo probabilidad total,

$$\begin{aligned}
&\mathbb{P}[\psi(H_n Q_n) \leq m] \\
&\qquad = \mathbb{P}[\psi(H_n) \leq m]\mathbb{P}[n/m < H_n \mid \psi(H_n) \leq m] \\
&\qquad + \mathbb{P}[\psi(H_n) \leq m]\mathbb{E}\left[\frac{\pi(m)+1}{\pi(n/H_n)+1}\mathbb{1}_{\{\{H_n \leq n/m\}\}} \mid \psi(H_n) \leq m\right].
\end{aligned}$$

Problemas fundamentales en el principio de transferencia

El manejo de los términos en azul, requieren de la resolución de

Problema I

Estimar con precisión

$$\mathbb{P}[\psi(H_n) \leq m]$$

Problemas fundamentales en el principio de transferencia

El manejo de los términos en azul, requieren de la resolución de

Problema I
Estimar con precisión

$$\mathbb{P}[\psi(H_n) \leq m]$$

El manejo de los términos en rojo, requieren
Problema II
Mediante estimaciones clásicas de π, el manejo del término en rojo, requiere de estimaciones de

$$\mathbb{E}\left[g(H_n/n) \mid \psi(H_n) \leq \kappa(n,m)\right],$$

con $\kappa(n,m)$ escogida adecuadamente.

Estimación de $\mathbb{P}[\psi(H_n) \leq m]$: preliminares

Sean $\{\xi_p\}_{p\in\mathcal{P}}$ variables geométricas con éxito $1/p$.

Estimación de $\mathbb{P}[\psi(H_n) \leq m]$: preliminares

Sean $\{\xi_p\}_{p\in\mathcal{P}}$ variables geométricas con éxito $1/p$.

Proposition (Chen, Jaramillo, Yang, (2021))
La ley de H_n satisface

$$\mathcal{L}((\alpha_p(H_n) \ : \ p \in \mathcal{P}_n)) = \mathcal{L}((\xi_p \ : \ p \in \mathcal{P} \cap [n]) \mid A_n),$$

con

$$A_n := \{ \prod_{p\in\mathcal{P}\cap[n]} p^{\xi_p} \leq n\}. \tag{1}$$

Moreover, if $m \geq 21$, then $\mathbb{P}[A_n] \geq 1/2$.

Observación

Introdujimos variables independientes!

Estimación de $\mathbb{P}[\psi(H_n) \leq m]$: simplificación clave

Podemos escribir

$$\mathbb{P}[\psi(H_n) \leq m] = \left(\frac{1}{L_n} \prod_{p \in \mathcal{P}_m} (1 - 1/p)^{-1} \right) \mathbb{P}\left[S_m \leq \frac{\log(n)}{\mathbb{E}[Z_m]} \right],$$

donde $S_m := Z_m/\lambda_m$, con $\lambda_m := \mathbb{E}[Z_m]$ y

$$Z_m := \sum_{p \in \mathcal{P} \cap [m]} \log(p) \xi_p,$$

Estimación de $\mathbb{P}[\psi(H_n) \leq m]$: simplificación clave

Podemos escribir

$$\mathbb{P}[\psi(H_n) \leq m] = \left(\frac{1}{L_n} \prod_{p \in \mathcal{P}_m} (1 - 1/p)^{-1} \right) \mathbb{P}\left[S_m \leq \frac{\log(n)}{\mathbb{E}[Z_m]} \right],$$

donde $S_m := Z_m / \lambda_m$, con $\lambda_m := \mathbb{E}[Z_m]$ y

$$Z_m := \sum_{p \in \mathcal{P} \cap [m]} \log(p) \xi_p,$$

Estimar la parte azul es un problema bien conocido en teoría de números (fórmula de Mertens) y es aproximadamente e^γ, donde $\gamma =$ constante de Euler.

Estimación de $\mathbb{P}[\psi(H_n) \leq m]$: simplificación clave

Podemos escribir

$$\mathbb{P}[\psi(H_n) \leq m] = \left(\frac{1}{L_n} \prod_{p \in \mathcal{P}_m} (1 - 1/p)^{-1} \right) \mathbb{P}\left[S_m \leq \frac{\log(n)}{\mathbb{E}[Z_m]} \right],$$

donde $S_m := Z_m / \lambda_m$, con $\lambda_m := \mathbb{E}[Z_m]$ y

$$Z_m := \sum_{p \in \mathcal{P} \cap [m]} \log(p) \xi_p,$$

Estimar la parte azul es un problema bien conocido en teoría de números (fórmula de Mertens) y es aproximadamente e^{γ}, donde $\gamma =$ constante de Euler. La parte roja es completamente probabilista

Problema simplificado

Entender la ley asintótica de S_n.

Formulación probabilista de $\mathbb{P}[S_n \leq m] \approx e^{-\gamma}\rho \circ \Upsilon(n, m)$

Se sabe que $e^{-\gamma}\rho(u) = 1$ es una densidad de probabilidad.

Formulación probabilista de $\mathbb{P}[S_n \leq m] \approx e^{-\gamma}\rho \circ \Upsilon(n, m)$

Se sabe que $e^{-\gamma}\rho(u) = 1$ es una densidad de probabilidad.

Definition
Una variable D tiene ley Dickman si su densidad es $e^{-\gamma}\rho(x)$, para $x \geq 0$.

Formulación probabilista de $\mathbb{P}[S_n \leq m] \approx e^{-\gamma}\rho \circ \Upsilon(n, m)$

Se sabe que $e^{-\gamma}\rho(u) = 1$ es una densidad de probabilidad.

Definition
Una variable D tiene ley Dickman si su densidad es $e^{-\gamma}\rho(x)$, para $x \geq 0$.

Lemma (Pinsky (2016))
D tiene ley Dickman si para toda función de prueba $f : \mathbb{R}_+ \to \mathbb{R}$,

$$\mathbb{E}[Dg(D)] - \int_0^1 \mathbb{E}[g(D+t)]dt = 0. \tag{2}$$

Formulación probabilista de $\mathbb{P}[S_n \leq m] \approx e^{-\gamma}\rho \circ \Upsilon(n, m)$

Se sabe que $e^{-\gamma}\rho(u) = 1$ es una densidad de probabilidad.

Definition
Una variable D tiene ley Dickman si su densidad es $e^{-\gamma}\rho(x)$, para $x \geq 0$.

Lemma (Pinsky (2016))
D tiene ley Dickman si para toda función de prueba $f : \mathbb{R}_+ \to \mathbb{R}$,

$$\mathbb{E}[Dg(D)] - \int_0^1 \mathbb{E}[g(D+t)]dt = 0. \tag{2}$$

Alternativamente, si $f = g'$,

$$\mathbb{E}[Df'(D)] + f(D) - f(D+1) = 0. \tag{3}$$

Heurística de Stein

Si la izquierda de (2) es approximadamente cero, D es aproximadamente Dickman

Método de Stein para leyes Dickman

La heurística se formaliza considerando la ecuación de Stein

$$xf_z'(x) + f_z(x) - f_z(x+1) = h_z(x) - \mathbb{E}[h_z(D)], \tag{4}$$

donde $h_z(x) := 1_{[0,z]}(x)$.

Método de Stein para leyes Dickman

La heurística se formaliza considerando la ecuación de Stein

$$xf_z'(x) + f_z(x) - f_z(x+1) = h_z(x) - \mathbb{E}[h_z(D)], \tag{4}$$

donde $h_z(x) := 1_{[0,z]}(x)$. Por cálculos elementales, para toda variable X con valores en $\mathbb{R}_+$,

$$\mathbb{P}[X \leq u] - \mathbb{P}[D \leq u] = \mathbb{E}[Xf_z'(X) + f_z(X) - f_z(X+1)],$$

Moraleja

Estimar proximidad entre X y D se reduce a estimar

$$\mathbb{E}[Xf_z'(X) + f_z(X) - f_z(X+1)].$$

Implementación técnica I

La clave es entender a la cantidad $\mathbb{E}[(Z_m/\lambda_m)f_z'(Z_n)$.

Implementación técnica I

La clave es entender a la cantidad $\mathbb{E}[(Z_m/\lambda_m)f_z'(Z_n)$. El argumento fundamental se basa en escribir a esta cantidad como

$$\frac{1}{\lambda_m}\sum_{p\in\mathcal{P}\cap[m]}\log(p)\mathbb{E}\left[\xi_p f_z'\left(\frac{1}{\lambda_m}\sum_{p\in\mathcal{P}\cap[m]}\log(p)\xi_p f_z'\right)\right],$$

y cálculos explícitos que involucran a variables aleatorias geométricas.

Resultado principal

Theorem (Jaramillo, Yang (2024))
Se tiene que

$$\Psi_H(n,m) = n\Upsilon(n,m)\left(\mathcal{I}[\rho]\circ\Upsilon(n,m) + O(\Upsilon(n,m)\log\log(m)/\log(m))\right),$$

uniformemente sobre $16 \leq m \leq n$, *donde*

$$\mathcal{I}[\rho](x) = \int_0^x \rho(t)dt.$$

Preguntas abiertas

Algunas preguntas abiertas

- Se puede decir algo sobre el comportamiento de H_n condicional a $\psi(H_n) \leq \kappa_{m,n}$, para alguna barrera $\kappa_{m,n}$ dada?
- Se puede decir algo sobre la convergencia hacia la función ρ, al estilo de 'convergencia de densidades'?
- Se puede hacer una formulación del problema en extensiones de $\mathbb{Q}$?

Gracias!!!

References

Jaramillo Yang (Arxiv). Approximation of smooth numbers for Harmonic samples.