Who am I?

José Antonio Perusquía Cortés

12 de junio 2023

Agenda

- Formación Académica
- Experiencia Profesional
- Investigación
- Formación de Recursos Humanos

Formación Académica

- Licenciatura en Actuaría
- Titulación por alto rendimiento académico

Formación Académica

- Maestría en Ciencias Matemáticas
- Intercambio académico en la Universidad de Bath
- Titulación por tesis bajo la supervisión de Dr Ramsés H. Mena

Formación Académica

CIENCIAS

iimas

UNIVERSITY OF BATH

University of Kent

- Doctorado en Estadística
- Supervisores: Dr Cristiano Villa y Dr Jim Griffin

Experiencia Profesional

- Ayudante de Profesor de Asignatura con Dra Ruth Selene Fuentes de
  - Inferencia Estadística
  - Series de Tiempo y Análisis de Supervivencia
  - Análisis Multivariado

Experiencia Profesional

- Graduate Teaching Assistant en materias de
  - Estadística, Probabilidad y Finanzas

- Financial Mathematics Tutor

- PGR Seminar Organiser

Experiencia Profesional

University of Kent → Imperial College London

- Investigador Postdoctoral Asociado en el Departamento de Cómputo
- Supervisor: Dr Giuliano Casale
- Proyecto de detección de anomalías en sistemas de cómputo distribuidos

Experiencia Profesional (Actual)

CIENCIAS

University of Kent

Imperial College London

iimas

- Investigador Postdoctoral y Candidato a Investigador Nacional (SNI)
- Supervisor: Dr Ramsés H. Mena
- Proyecto de análisis de sensibilidad de procesos stick-breaking

- Profesor de Asignatura A (Facultad de Ciencias)
  - Análisis Multivariado

Investigación

iimas

- Análisis de sensibilidad de procesos stick-breaking usando la divergencia Kullback-Leibler
- Colaboradores: Dr Ramsés H. Mena y Dr Mario Diaz Torres
- Un preprint

Procesos stick-breaking

- Muchas y muy variadas formas de definirlos a través de particulares elecciones de la colección v_i, e.g.:

  - Procesos independientes (proceso Dirichlet)
  - Procesos dependientes (proceso geométrico) donde $v_i = v \sim Be(a, b)$[5].
  - Procesos intercambiables, e.g.

$$v_i \mid \nu \sim \nu$$

$$\nu \sim Dir(\beta, \nu_0)^6$$

[5] Mena R. & Ruggiero, M. & Walker S.(2011). Geometric stick-breaking processes for continuous-time Bayesian nonparametric modeling. Jr. of Statistical Planning and Inference.

[6] Gil-Leyva, M.F. & Mena R.H. (2021). Stick-Breaking Processes With Exchangeable Length Variables. JASA.

Análisis de Sensibilidad I

- **Idea:** Si tenemos dos MPA's P y P' queremos estudiar que tan diferentes son.

- Para hacerlo recurrimos a la divergencia Kullback-Leibler ya que:

1. Proporciona una forma "manejable" de hacer las matemáticas

2. Teorema (Desigualdad de Pinsker)

Para dos medidas P y P' definidas en un espacio medible $(\Omega, \mathscr{F})$ se cumple

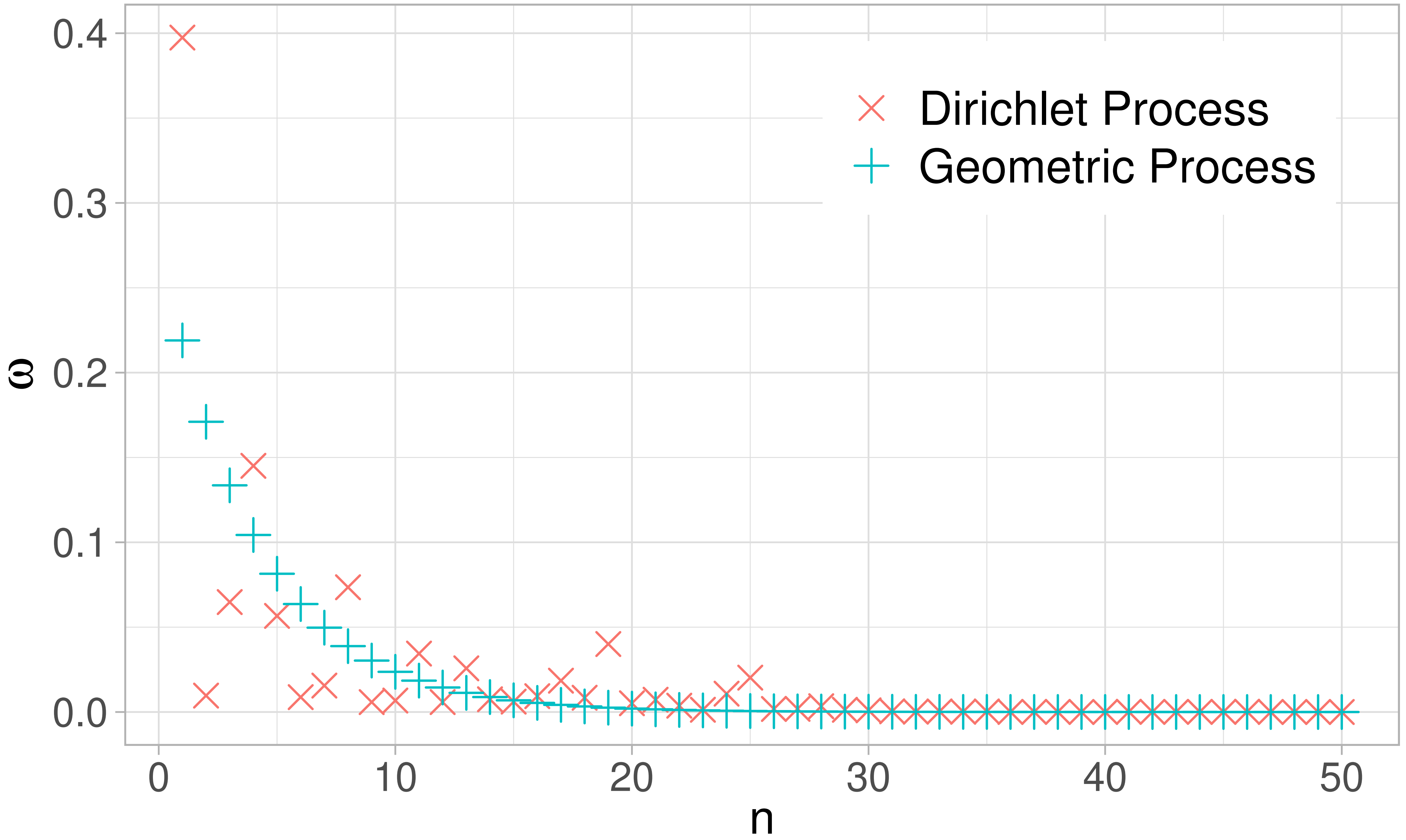
$$\mathrm{TV}(P, P') \leq \sqrt{\frac{D_{KL}(P\,|\,P')}{2}}$$

Análisis de Sensibilidad I

- $P \sim Dir\left(\theta, P_0\right)$ y $P' \sim Geo\left(a, b, P_0\right)$

¿Cómo usar la divergencia Kullback-Leibler?

- Suponer que tienen las mismas localizaciones y sólo difieren en los pesos



Análisis de Sensibilidad I

Consecuencias

- La divergencia Kullback-Leibler solo depende de las variables $\{v_i\}_{i\geq 1}, v$

$$D_{KL}(P\,||\,P') = \sum_{n\geq 1} \left[\prod_{j<n} (1 - v_j) \right] d(v_n\,||\,v)$$

donde

$$d(v_n\,||\,v) = v_n \log\left(\frac{v_n}{v}\right) + (1 - v_n)\log\left(\frac{1 - v_n}{1 - v}\right)$$

- La divergencia es una variable aleatoria!

Análisis de Sensibilidad I

- La esperanza:

$$\mathbb{E}(D_{KL}(P\,||\,P')) = (\theta+1)\mathbb{E}(d(v_1\,||\,v))$$

- La varianza:

$$\mathrm{Var}(D_{KL}(P\,||\,P')) = \frac{(\theta+2)}{2}\mathbb{E}(d^2(v_1\,||\,v)) + (\theta+1)(\theta+2)\mathbb{E}[(1-v_1)d(v_1\,||\,v)d(v_2\,||\,v)] - [(\theta+1)\mathbb{E}(d(v_1\,||\,v))]^2$$

Análisis de Sensibilidad I

- La divergencia Kullback-Leibler de P' con respecto a P

$$D_{KL}(P'||P) = \sum_{n\geq 1} (1-v)^{n-1} d(v||v_n)$$

- Su esperanza es finita solo si $a > 1$

$$\mathbb{E}(D_{KL}(P'||P)) = \sum_{n=1}^{\infty} \left(\frac{a\Gamma(b+n-1)\Gamma(a+b)}{\Gamma(b)\Gamma(a+b+n)} \left[\psi(a+1) - \psi(a+b+n) + \gamma + \psi(\theta+1)\right] + \frac{\Gamma(b+n)\Gamma(a+b)}{\Gamma(b)\Gamma(a+b+n)} \left[\psi(b+n) - \psi(a+b+n) + \frac{1}{\theta}\right] \right)$$

Análisis de Sensibilidad I

- Para $a = 1$ y $b = \theta$ se tiene que

$$\mathbb{E}(D_{KL}(P\,||\,P')) = 1$$

- No es ideal entonces necesitamos algo más: $v_1 = v$

$$\mathbb{E}(D_{KL}(P\,||\,P')) = \frac{\theta}{\theta + 1}$$

Análisis de Sensibilidad II

- **Objetivo:** Estudiar la divergencia de P_β con respecto a P' donde P_β está definido por las variables

$$\begin{aligned} v_i \,|\, \nu &\sim \nu \\ \nu &\sim Dir(\beta, \nu_0) \end{aligned}$$

y $\nu_0 = Be(1,\theta)$.

Análisis de Sensibilidad II

Lemma

Sea P_β y P' un proceso stick-breaking intercambiable con un proceso Dirichlet como medida directriz y el proceso geométrico respectivamente, entonces

$$D_\theta(\beta) = \mathbb{E}[D_{KL}(P_\beta || P')] = \sum_{n=1}^{\infty} \sum_{\pi \in \mathscr{P}([n])} F_\theta(\pi) p_\beta(\pi)$$

donde

$$F_\theta(\pi) p_\beta(\pi) = \mathbb{E}\left[\left(\prod_{j=1}^{k-1} (1 - v_j)^{|A_j|}\right) (1 - v_k)^{|A_k| - 1} d(v_k || v)\right] \frac{\beta^k}{(\beta)^{(n)}} \prod_{j=1}^{k} (|A_j| - 1)!$$

Análisis de Sensibilidad II

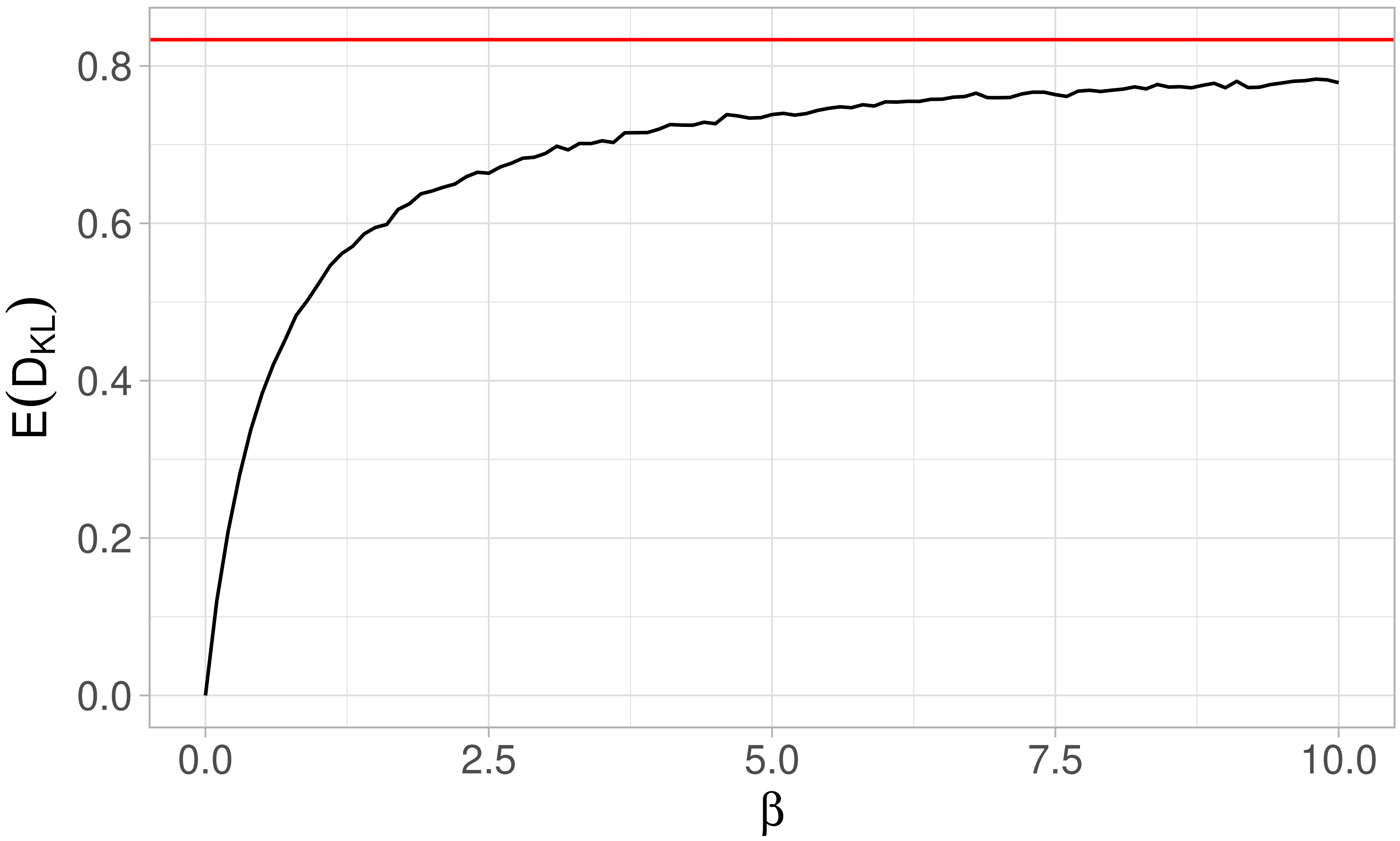
- No hay expresión cerrada pero …

- Para $\beta = 0$

$$D_\theta(0) = \sum_{n=1}^{\infty} \mathbb{E}[(1 - v_1)^{n-1} d(v_1 \,||\, v)]$$

- Para β grande deberíamos recuperar la $D_{KL}(P \,||\, P')$

Análisis de Sensibilidad II



Análisis de Sensibilidad II

Teorema

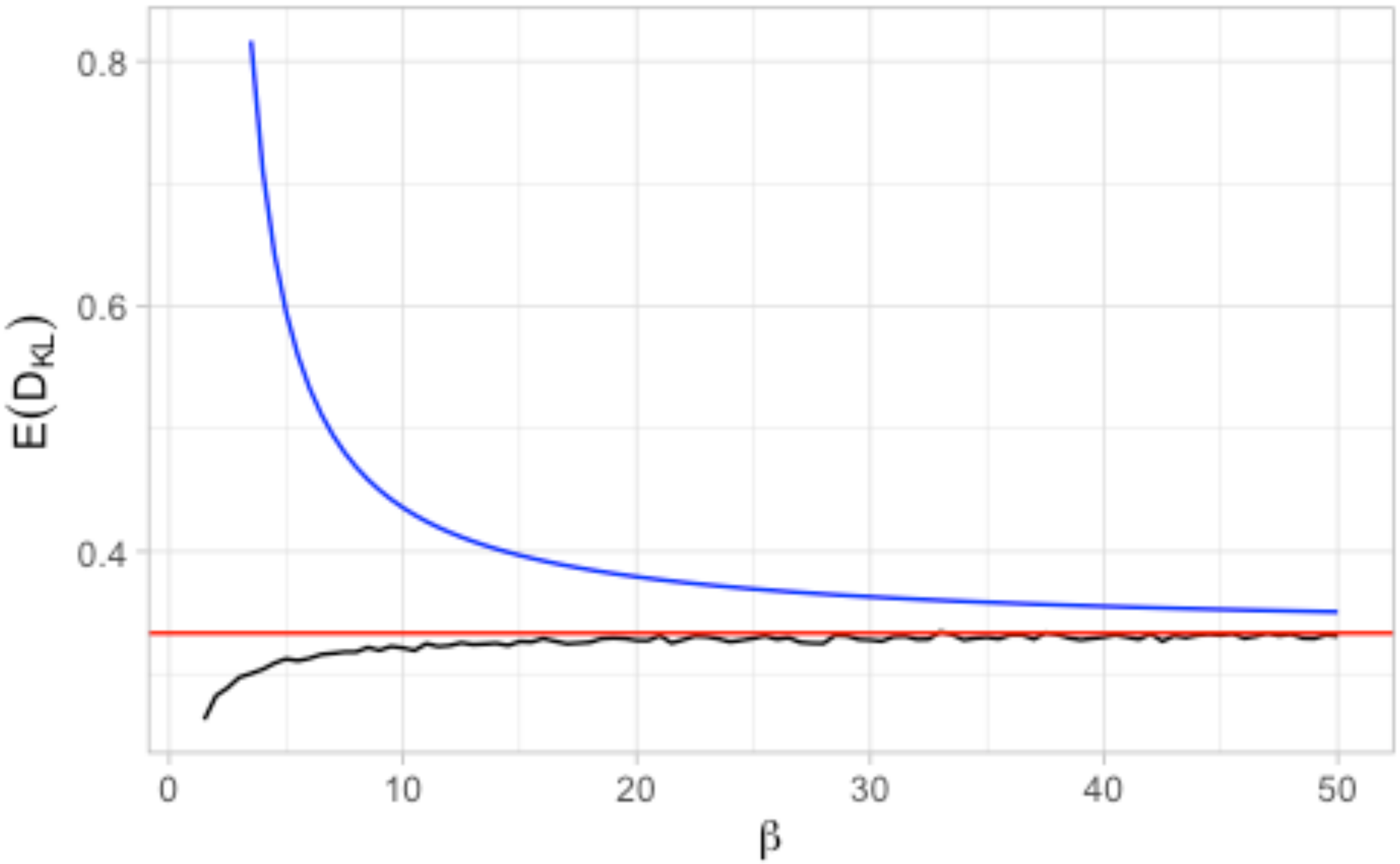
Sea P_β y P' un proceso stick-breaking intercambiable y el proceso geométrico respectivamente, entonces para $\beta > \theta + 1$:

- $D_\theta(\beta)$ es continua
- $D_\theta(\beta) \leq \frac{\theta}{\theta + 1} \frac{\beta^2}{(\beta - 1)(\beta - (\theta + 1))}$

Y así:

$$\lim_{\beta \to \infty} \mathbb{E}(D_{KL}(P_\beta || P')) = \mathbb{E}(D_{KL}(P || P'))$$

Análisis de Sensibilidad II



- **Mensaje:** Para θ pequeño y β grande no hay beneficio

Otras Líneas de Investigación

University of Kent

- Métodos bayesianos para clasificación supervisada de datos binario con aplicaciones a ciberseguridad
- Colaboradores: Dr Jim Griffin y Dr Cristiano Villa
- Un artículo publicado y un preprint

Otros Líneas de Investigación

Imperial College London

- Detección de anomalías en sistemas de cómputo distribuidos
- Colaboradores: Dr Giuliano Casale y Dr Alim Ul Gias
- Un artículo aceptado y uno sometido

Formación de Recursos Humanos

- Servicio Social

  Co-supervisión de 7 alumnos con Dr Ramsés H. Mena

  Supervisión de Martín Guzmán Ruelas

- Alumnos

  Sinodal de Mariana Guadalupe Lagunas Alvarado (tesis licenciatura)

  Director de tesis de licenciatura de Andrés Uriel Ortiz Covarrubias (por inscribir)

¡Gracias por su atención!