

Distribución espacial de profesionales en Ciencia y Tecnología de Costa Rica

Pedro Montenegro & Pablo Vivas

10 de octubre, 2019

1. Introducción

Los profesionales en Ciencia y Tecnología (CyT) son vitales para el desarrollo de cualquier nación. Las investigaciones realizadas en este campo del conocimiento y la labor de estos profesionales, indudablemente, afectan en mayor o menor escala el desenvolvimiento de la especie humana. Conocer dónde se ubican los profesionales con estas características es una labor imperante para un país, ya que a raíz de esa información pueden surgir hipótesis, conclusiones y esfuerzos para generar oportunidades donde no las haya. La ubicación de un profesional en ciencia y tecnología puede revelar mucha información, pero la agrupación de profesionales con rasgos similares genera mayores insumos para la toma de decisiones. Debido a lo anterior, surge la siguiente pregunta de investigación: ¿Existen patrones en la distribución espacial de los egresados en ciencia y tecnología de Costa Rica?

El estudio se realiza con datos de profesionales en áreas científico-tecnológicas formados en Costa Rica durante el periodo 2000-2017 y que cuentan con un grado académico que va desde el bachillerato universitario hasta el doctorado. Las carreras consideradas se dividen en 4 grupos : Ciencias Agrícolas, Ciencias Exactas y Naturales, Ciencias Médicas e Ingeniería y Tecnología. Esta información proviene de la base de datos de talento profesional 2000-2017 del portal HIPATIA del Programa Estado de la Nación (PEN, 2019). Esta integra la Base de datos de graduados (BADAGRA) y la de la Oficina de Reconocimiento y Equiparación de Grados (ORE), ambas del Consejo Nacional de Rectores (CONARE) que registran, respectivamente, los diplomas

de universidades públicas y privadas a nivel nacional, así como los graduados en el extranjero que iniciaron un trámite para su equiparación en el país. En esta base de datos la información de los títulos es transformada y depurada para reflejar cada graduado una única vez según su grado académico más alto y más reciente.

Dado que los datos carecen de información espacial sobre la ubicación de los profesionales, se integran con una segunda fuente de información que permite aproximar esta dimensión espacial. Así, para generar una variable que aproxime el lugar de residencia de cada individuo se utiliza el domicilio electoral registrado por el Tribunal Supremo de Elecciones de Costa Rica (TSE) en el padrón electoral del año 2018. La integración entre los datos de talento profesional y los datos del TSE permitió ubicar un total de 106.184 personas agrupadas en 1.734 centros de votación.

En la primera parte de este documento se detallan las características más importantes de la distribución espacial de los egresados en ciencia y tecnología de Costa Rica, aquí se presentan gráficos útiles para entender el fenómeno, al igual que herramientas para decidir si la distribución de puntos es completamente aleatoria o no. Por otro lado, se realiza un análisis de procesos puntuales, con la intención de entender el modelo teórico detrás de la generación de los puntos. En el epílogo del documento se discuten los resultados obtenidos y se destaca las limitaciones del estudio así como las posibles mejoras para un análisis más profundo.

2. Desarrollo

Se presenta, la distribución de los egresados en ciencia y tecnología, para cada uno de los 1734 centros de votación. Cabe destacar que para hacer posible la comparación, se calculó el porcentaje de egresados en ciencia y tecnología, para cada uno de los centros de votación con respecto a todas las personas inscritas en cada uno de los centros de votación. Los resultados se muestran en la figura 1. Una mapa similar se puede apreciar en los anexos, donde se representa la densidad de los puntos sobre todo el territorio nacional, sin embargo, el comportamiento también es el mismo: una fuerte concentración en las zonas céntricas, que se disminuye al alejarse de esta zona.

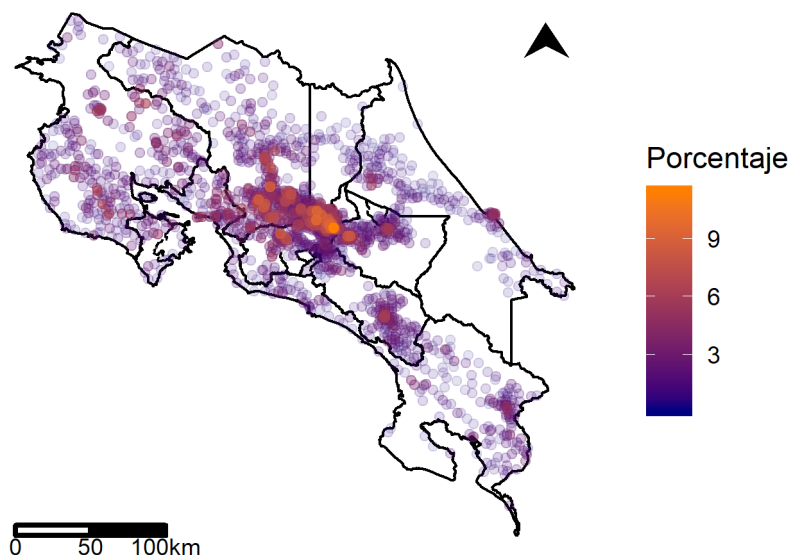


Figura 1: Costa Rica: Distribución porcentual de los egresados en CyT, según centro de votación

A partir de la figura 1, se observa este comportamiento marcado de los profesionales egresados en CyT. Estos profesionales se aglutinan en el centro del país, sin embargo existen otras locaciones donde es posible ver porcentajes elevados. Un ejemplo de esto son en las zona de Limón (cantón), Pérez Zeledón y algunos sectores de Guanacaste.

Con base en este análisis gráfico, se conjetura que la distribución de los puntos no sigue una distribución completamente aleatoria, esto por la fuerte influencia que ejerce el sector central del país para las oportunidades de obtener un título universitario en una carrera referente a ciencia y tecnología. Para obtener una mejor idea del comportamiento de los puntos a lo largo de Costa Rica, se emplearon las funciones F & G . Para cada una de las funciones se realizaron un total de 20 simulaciones, y se eliminaron el 1 % de cada extremo. Los resultados de las simulaciones se presentan en la figura 2.

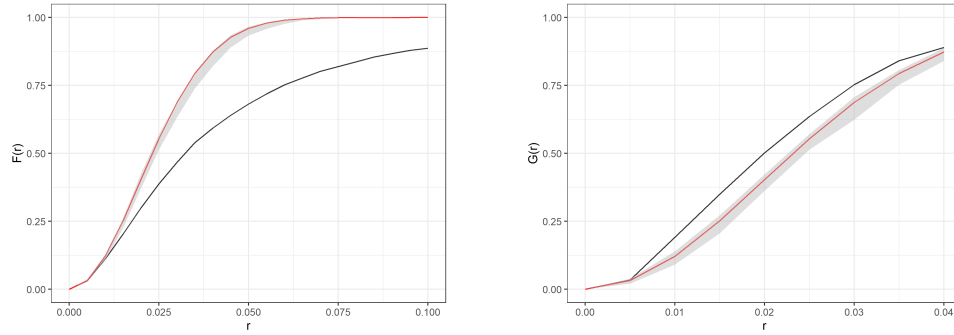


Figura 2: Simulación de funciones F y G

Con base en la figura 2, se observa que el patrón de puntos empíricos difiere de la generación teórica de los mismos. En el primer gráfico, la curva observada está por debajo de los valores teóricos que se observarían si la distribución hubiese sido completamente al azar. Esto complementado con la figura 1 da indicios de agrupaciones. Por otro lado, la función G crece rápidamente al inicio y por ende se llega a la misma conclusión, sin embargo, no está tan marcada la diferencia en comparación a la función F .

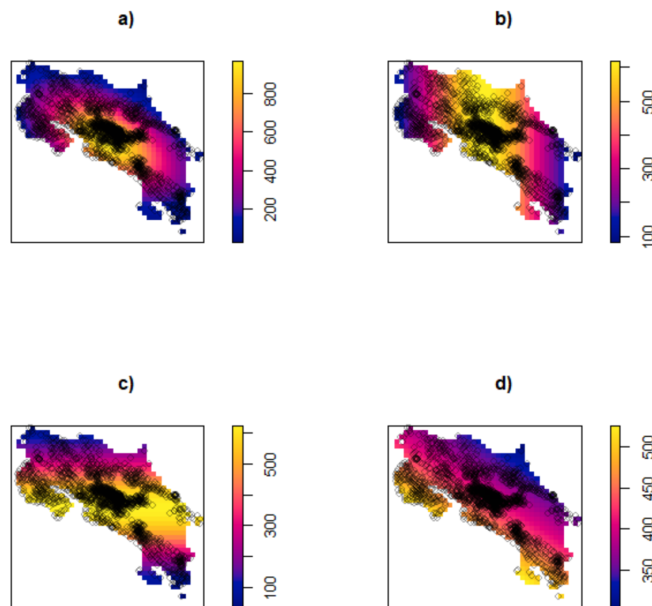


Figura 3: Representación gráfica de 4 modelos teóricos propuestos

Aunque más adelante se describen los modelos utilizados para la generación de la figura 3, la información preliminar da una señal sobre el ajuste del mismo para el fenómeno en cuestión. En primer lugar, el modelo a) genera una densidad curvilínea, donde la mayor densidad se concentra en el corazón del país y conforme se aleja de esa zona, la densidad baja. El modelo b) y c) poseen franjas verticales y horizontales de densidad, respectivamente. Por último el modelo d) parece ser el más alejado de la realidad, pues no logra capturar la gran cantidad de profesionales ubicados en el área metropolitana del país.

Patrones de tendencia

El argumento de tendencia de los modelos PPM describen las posibles relaciones espaciales y los efectos de las covariables construidas a través de las coordenadas de los eventos, para esto utilizamos la densidad de las observaciones y se modela la intensidad a través de las coordenadas para identificar los patrones de intensidad y las interacción entre puntos, entre mayor sea la concentración de puntos mayor intensidad, para ello nos vamos a centrar en los modelos de procesos poisson no homogéneo, de ante mano el gráfico 1 la concentración de puntos en zonas céntricas y en las zonas alejadas se observan pequeñas agrupaciones aisladas que se relacionan con las zonas céntricas de las diferentes provincias.

El primer modelo, se utiliza para observar la relación de los puntos únicamente considerando el intercepto y la longitud. La medida de bondad de ajuste ($AIC = -17411,31$)

Modelo espacial Poisson No homogéneo (x)					
Terminos	Estimate	S.E.	CI95.lo	CI95.hi	Ztest
Intercepto	3,717	2,417	-1,021	8,455	
x	-0,027	0,029	-0,084	0,029	

Figura 4: Elaboración propia

El segundo modelo, muestra la patrón de puntos únicamente con la longitud. Se puede observar que la latitud y su intercepto son significativos, mostrando indicios que la relación se concentra de este a oeste de Costa Rica, principalmente siguiendo el patrón

de la GAM. ($AIC = -17418,2$)

Modelo espacial Poisson No homogeneo (y)					
Terminos	Estimate	S.E.	CI95.lo	CI95.hi	Ztest
Intercepto	7,040	0,363	6,329	7,751	***
y	-0,102	0,036	-0,174	-0,031	**

Figura 5: Elaboración propia

El tercer modelo, utiliza la relación de las covariables longitud y latitud. Las covariables con coeficientes significativos al 5 % son la latitud y la latitud, de esta forma es interesante observar que los gráficos de intensidad explican la relación concentrada en el centro del país. Además, el criterio de información AKaike hace relación que dicho modelo presenta mejor ajuste ($AIC = -18070,57$)

Modelo espacial Poisson No homogeneo (x,y)					
Terminos	Estimate	S.E.	CI95.lo	CI95.hi	Ztest
Intercepto	-4,294	2,957	-10,089	1,501	
x	-0,150	0,039	-0,226	-0,074	***
y	-0,233	0,050	-0,331	-0,135	***

Figura 6: Elaboración propia

El cuarto modelo, analiza la relación polinomial de segundo orden para las coordenadas cartesianas, sin embargo, la matriz de varianzas y covarianzas no son invertibles computacionalmente, se recomienda utilizar otros métodos de estimación como enfoque Bayesiano para la estimación. ($AIC = -17431,02$)

**Modelo proceso espacial Poisson No homogéneo
segundo orden (x,y)**

Nombres	Coefficientes
(Intercept)	-6.398,84
x	-162,95
y	-93,57
$I(x^2)$	-1,06
$I(x * y)$	-1,58
$I(y^2)$	-2,03

Figura 7: Elaboración propia

Además, para la estimación de los modelos el óptimo ancho de banda es 0.0007777786, es la división más recomendada para disminuir el error cuadrático medio según el método Diggle.

Se concluye esta sección con la generación de un mapa de puntos realizada con el modelo seleccionado. Este mapa se presenta en la figura 8. A partir de este mapa, se puede observar que aunque la distribución de puntos es similar al presentado en la figura 1, existen discrepancias prácticas que supone la existencia de un mejor modelo para este fenómeno.



Figura 8: Simulación modelo elegido

3. Discusión y conclusiones

Las ubicaciones de los profesionales en ciencia y tecnología revelan algo esperado: una mayor concentración porcentual en las zonas céntricas del país, las oportunidades para obtener un título referido a una carrera de ciencia y tecnología son más probables si se vive cerca del área metropolitana. Estos se concentran fundamentalmente en una franja que atraviesa de este a oeste la zona central del país mientras que en las zonas costeras y cercanas a las fronteras, los profesionales son casi extintos.

Estos hechos se ven apoyados tanto por las funciones F y G, como los modelos confeccionados. Una limitante de este estudio referida a este tema, es la poca cantidad de simulaciones realizadas para la construcción de intervalos de confianza empíricos. Aunque el comportamiento de las observaciones está bien definido, en la función G podría quedar inquietud sobre la cantidad de simulaciones realizadas.

Por otro lado, en términos del problema, varios trabajos han advertido sobre la insuficiencia de las políticas de transformación productiva en Costa Rica, que no han logrado aliviar el desacople entre crecimiento económico y social de los territorios fuera

de la GAM y el dinamismo que ha generado la inversión extranjera directa principalmente concentrada en la región central.

El desarrollo autocentrado en la Gran Área Metropolitana, a su vez, no viene exento de limitaciones. La región central del país está llegando a un nivel de saturación que se evidencia en crecientes problemas ambientales, riesgo de desastres por la falta de ordenamiento territorial, ineficiencia de la infraestructura vial y del sistema de transporte, pérdida habitacional, altos costos de almacenamiento, entre otros problemas, que condicionan las posibilidades de alcanzar y sostener la tasa de crecimiento económica que se requiere en las décadas futuras.

4. Apéndice

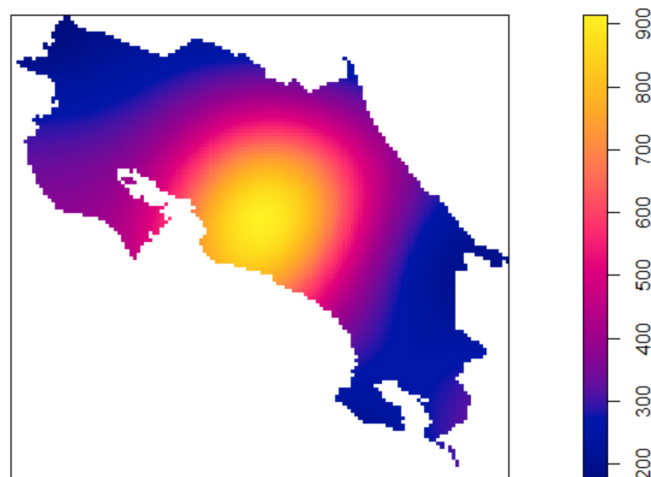


Figura 9: Costa Rica: Densidad de los egresados en CyT