

Tópicos de Estadística Espacial Aplicada

Proyecto 4

Asociación espacial en el nivel de progreso social de los cantones de Costa Rica

Fabian Brenes

Osvaldo Ureña

Introducción	3
Desarrollo	4
Datos	4
Análisis Estadístico	6
Conclusiones	11
Bibliografía	12
Anexos	13
Anexo 1: Matriz de pesos	13
Anexo 2: Salidas Modelos Espaciales	14

Introducción

Estudiar los factores que afectan la desigualdad socioeconómica de un país es siempre importante, Rafael Arias y otros autores al respecto mencionan que “el estudio y análisis de los factores que explican los crecientes niveles de desigualdad socioeconómica de un país son indispensables para la implementación de políticas y arreglos institucionales capaces de promover círculos virtuosos entre el crecimiento económico, la equidad social y la gobernabilidad democrática ” (Arias et al, pág:61).

Los autores anteriormente citados mencionan que la desigualdad socioeconómica no solo es un problema de distribución de la riqueza sino que también está relacionada con la exclusión o discriminación por razones de género, étnicas, etarias y espaciales (Arias et al, pág:64). Lo anterior motiva aún más el hecho de analizar la existencia o no de asociación espacial en el nivel de progreso social de los cantones de Costa Rica.

Teniendo en cuenta la importancia del factor espacial en la desigualdad socioeconómica el siguiente estudio pretende responder la siguiente pregunta de investigación **¿Existe asociación espacial en el nivel de progreso social de los cantones de Costa Rica?** Para responder lo anterior se llevará a cabo un modelo de estadística espacial de áreas donde la variable independiente será el “índice de progreso social de los cantones”.

El documento se estructura de la siguiente manera, primero se hace una breve explicación de los datos a utilizar así como una explicación de lo que se entiende como “índice de progreso social” que es la variable medular del presente estudio, seguidamente se elabora el desarrollo metodológico y se finaliza con un apartado de conclusiones.

Desarrollo

Datos

Las variables asociadas a cada cantón son las siguientes:

- Tasa de robos, asaltos y hurtos
- Habitantes por EBAIS
- Cobertura del servicio de recolección de residuos
- Acceso a energía eléctrica
- Acceso a infraestructura cultural y recreativa
- Tasa de empleo formal
- Acceso a conocimientos básicos
- Índice de progreso social

Las variables para construir los habitantes por EBAIS así como la tasa de empleo formal fueron tomadas de la Caja Costarricense del Seguro Social. Por su parte las otras variables fueron tomadas del INCAE.

La variable dependiente es el índice de progreso social (IPS), a continuación se explica brevemente qué se entiende por progreso social. Bajo el esquema del índice de progreso social el “progreso social” se explica como la capacidad de una sociedad para dar a sus ciudadanos las necesidades básicas y poder establecer los lazos que permite a sus ciudadanos mejorar y mantener su calidad de vida, además de crear las condiciones para que todos los individuos alcancen su máximo potencial.

De la definición anterior se puede denotar que hay tres pilares básicos del progreso social que son: Necesidades Básicas, Fundamentos de Bienestar y Oportunidades; cada pilar tiene cuatro componentes, podemos verlo en la siguiente figura.

Figura 1: Componentes del Índice de Progreso Social



Fuente: Stern y Epner, pág. 4

Este índice se calcula en 5 pasos principales, primero imputan los valores perdidos, el segundo paso es estandarizar los indicadores para tenerlos en una escala comparable, luego mediante el método de Análisis de Componentes Principales se agregan los indicadores a un componente, y luego finalmente se calculan las dimensiones y en general el índice de progreso social promediando los componentes y dimensiones respectivamente (Stern y Epner, pág:11).

Es importante observar la correlación de las variables a utilizar con el índice de progreso social. En el cuadro 1 se muestra la matriz de correlaciones de las variables asociadas a los cantones.

Se puede resaltar que todas las variables se correlacionan positivamente con el IPS excepto la variable “tasa de robos, asaltos y hurtos” que se correlaciona negativamente, por lo tanto todas las variables tienen la dirección de correlación esperada. Por otro lado las dos variables que tienen la correlación positiva más fuerte con el IPS son “cobertura del servicio de recolección de basura” y “acceso a energía eléctrica”

Cuadro 1

Matriz de correlación de las variables

	IPS	Tasa de robos asaltos y hurtos	Habitantes por EBAIS	Cobertura del servicio de recolección de residuos	Acceso a energía eléctrica	tasa empleo formal	Acceso a Conocimientos Básicos
IPS	1,000	-0,329	0,434	0,511	0,491	0,240	0,301
Tasa de robos. asaltos y hurtos	-0,329	1,000	-0,015	-0,021	-0,144	0,455	0,036
Habitantes por EBAIS	0,434	-0,015	1,000	0,361	0,303	0,146	-0,097
Cobertura del servicio de recolección de residuos	0,511	-0,021	0,361	1,000	0,349	0,242	0,003
Acceso a energía eléctrica	0,491	-0,144	0,303	0,349	1,000	0,160	0,225
tasa empleo formal	0,240	0,455	0,146	0,242	0,160	1,000	0,229
Acceso a Conocimientos Básicos	0,301	0,036	-0,097	0,003	0,225	0,229	1,000

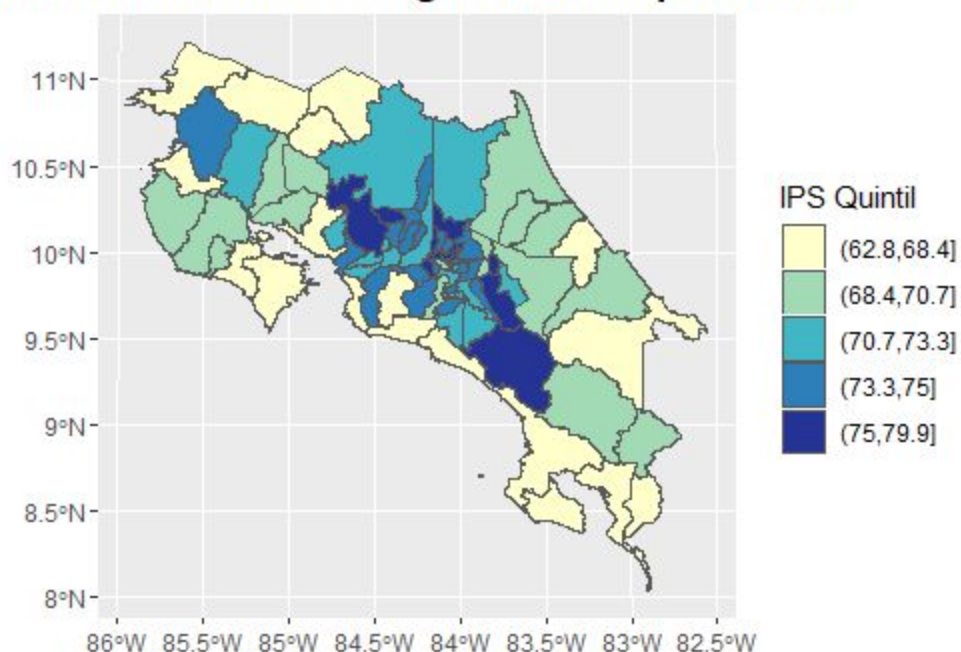
Fuente: Elaboración Propia

Análisis Estadístico

A partir de lo mencionado anteriormente, se procede a evaluar la dependencia espacial del Índice de Progreso Social (IPS) variable que representa el nivel de desarrollo social

de los cantones, seguidamente se estima una regresión lineal del IPS en función de la tasa de robos y hurtos, habitantes por EBAIS, Cobertura del servicio de recolección de residuos, acceso a energía eléctrica, tasa de empleo formal (empleados adscritos al seguro de salud entre la población económicamente activa), acceso a conocimientos básicos, y luego se evalúa la dependencia espacial tanto de la variable dependiente y el residuo de la regresión lineal.

Gráfico 1: Índice de Progreso Social por cantón



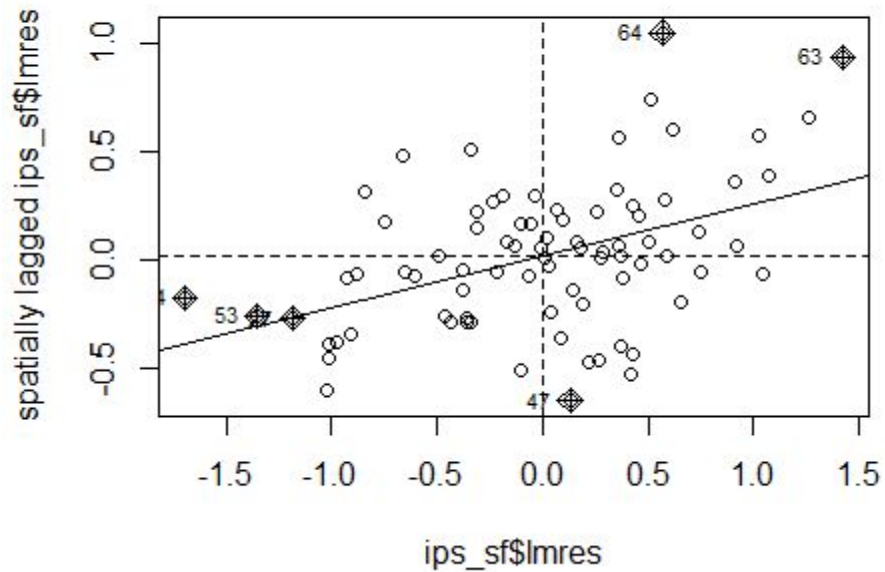
Según se observa en el gráfico 1, el nivel de progreso social es desigual en Costa Rica, donde se denota que los cantones con mejor desempeño se ubican en el centro del país, caso contrario, los cantones con nivel de progreso social bajo, correspondiente al quintil 1 y 2, se concentran en las zonas periféricas. Al efectuar la prueba de Moran, utilizando una matriz de pesos con el movimiento de la reina (ver anexo 1), para evaluar la dependencia espacial del IPS por cantón, se obtiene un valor 0.47, con un valor p de 0.001, lo cual indica que existe evidencia estadística significativa para rechazar la hipótesis nula de que el IPS por cantón no tiene dependencia espacial.

Cuadro 2: Regresión Lineal, variable dependiente IPS

Indicador		Parámetro	Std.Error	t Value	Pr(> t)
Intercepto		0.0000	0.0722	0	1
Tasa de robos, asaltos y hurtos	-	0.4133	0.0845	-4.893	0.00001 ***
Habitantes por EBAIS		0.2648	0.0809	3.274	0.00162 **
Cobertura del servicio de recolección de residuos		0.2956	0.0827	3.574	0.00062 ***
Acceso a energía eléctrica		0.1528	0.0837	1.827	0.07171 .
tasa empleo formal		0.2360	0.0887	2.661	0.00955 **
Acceso a Conocimientos Básicos		0.2523	0.0780	3.234	0.00182 **
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

El cuadro 2 muestra los resultados de la regresión lineal con variables estandarizadas, donde destaca que todas los parámetros son estadísticamente significativos con un nivel de significancia de 10%, adicionalmente tienen el signo esperado. No obstante, al aplicar la prueba de Moran a los residuos, se obtiene un valor de 0.23903 y valor probabilístico de 0.001, por lo cual se rechaza la hipótesis nula de que los residuos no muestran dependencia espacial, con un nivel de significancia de 5%. El gráfico 2 reafirma la conclusión anterior, al evidenciar una relación positiva entre los residuos con sus rezagos espaciales.

Gráfico 2: Residuos vs residuos rezagados espacialmente



Al realizar la regresión lineal del IPS en función de covariables fundamentadas en la teoría, sigue persistiendo la dependencia espacial, por lo cual se procede a estimar un modelo Autorregresivo Simultáneo (SAR), el cual modela el error considerando la dependencia espacial entre las áreas de la siguiente forma (Bivand et al, 2008):

$$e_i = \sum_{i=1}^m b_{ij}e_i + \varepsilon_i$$

Y un modelo Autorregresivo Condicional (CAR), cuya especificación se basa en la distribución condicional del término de error espacial, lo que significa que la distribución del error e_i condicionada a e_{i-1} es dada (Bivand et al, 2008). Adicionalmente, se estiman los modelos espaciales con la matriz de pesos basado en el movimiento de la reina (W Reina) y el movimiento de la torre (W Torre) y se define cual ajusta mejor de acuerdo con el AIC.

Cuadro 3: Comparativo entre modelos

Indicador	Mod. Lineal	SAR W Reina	SAR W Torre	CAR W Reina	CAR W Torre
Intercepto	0.000	-0.038	-0.029	-0.238	-0.200
Tasa de robos, asaltos y hurtos	-0.413 ***	-0.365 ***	-0.368 ***	-0.391 ***	-0.386 ***
Habitantes por EBAIS	0.265 **	0.202 *	0.199 *	0.165 *	0.155 *
Cobertura del servicio de recolección de residuos	0.296 ***	0.111	0.116 .	0.207 **	0.209 **
Acceso a energía eléctrica	0.153 .	0.159 **	0.164 **	0.168 *	0.182 *
tasa empleo formal	0.236 **	0.135	0.123	0.215 *	0.221 *
Acceso a Conocimientos Básicos	0.252 **	0.299 ***	0.305 ***	0.260 ***	0.253 ***
AIC	168.6	152.6	151.2	156.6	155.0

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

El menor AIC corresponde al modelo SAR con la matriz de pesos W Reina con un valor de 151.2, en comparación a 168.6 de la regresión lineal, lo que sugiere que incorporar el componente espacial mejora los resultados obtenidos, no obstante, el lambda del modelo SAR corresponde a 0.69 y un valor $p < 0.001$ lo que indica que los residuos continúan presentando dependencia espacial.

Los parámetros obtenidos (salida completa del modelo en el anexo 2) indican que al aumentar la tasa de robos, asaltos y hurtos en una desviación estándar, el IPS disminuye en 0.368 desviaciones estándar. De igual manera, al aumentar en una desviación estándar los EBAIS por habitante, el IPS aumenta en 0.199 desviaciones estándar. El acceso a energía eléctrica muestra una asociación positiva con el IPS de 0.164, manteniendo las demás variables constantes, mientras que el parámetro de acceso a conocimientos básicos es de 0.305. En estas cuatro variables, los parámetros son estadísticamente significativos con un nivel alfa de 5%.

Por su parte, las variables Cobertura del Servicio de Recolección de Residuos y la Tasa de Empleo Formal no son estadísticamente significativos si se selecciona un alfa de 5%, no obstante, muestran una magnitud considerable (0.116 y 0.123) por lo cual se mantienen en el modelo.

Conclusiones

Es importante resaltar que de manera gráfica se puede observar como los cantones que se encuentran en el centro del país tienen un mayor índice de progreso social que aquellos cantones que se ubican en la periferia.

Luego se pudo constatar que las covariables elegidas para realizar la regresión con el índice de progreso social se correlacionan con este de la manera que se espera, es decir, todas las variables tienen una correlación positiva excepto la tasa de robos, asaltos y hurtos que se correlaciona de manera negativa.

Luego aplicando la prueba de Moran utilizando la matriz de pesos definida con el movimiento de la reina se demostró que existe evidencia estadística significativa para rechazar la hipótesis nula de que el IPS por cantón no tiene dependencia espacial.

Aplicando la regresión lineal con las variables estandarizadas se obtuvieron coeficientes con el signo esperado y significativos con un nivel de confianza del 10%, esta regresión tiene la debilidad de que los residuos muestran dependencia espacial.

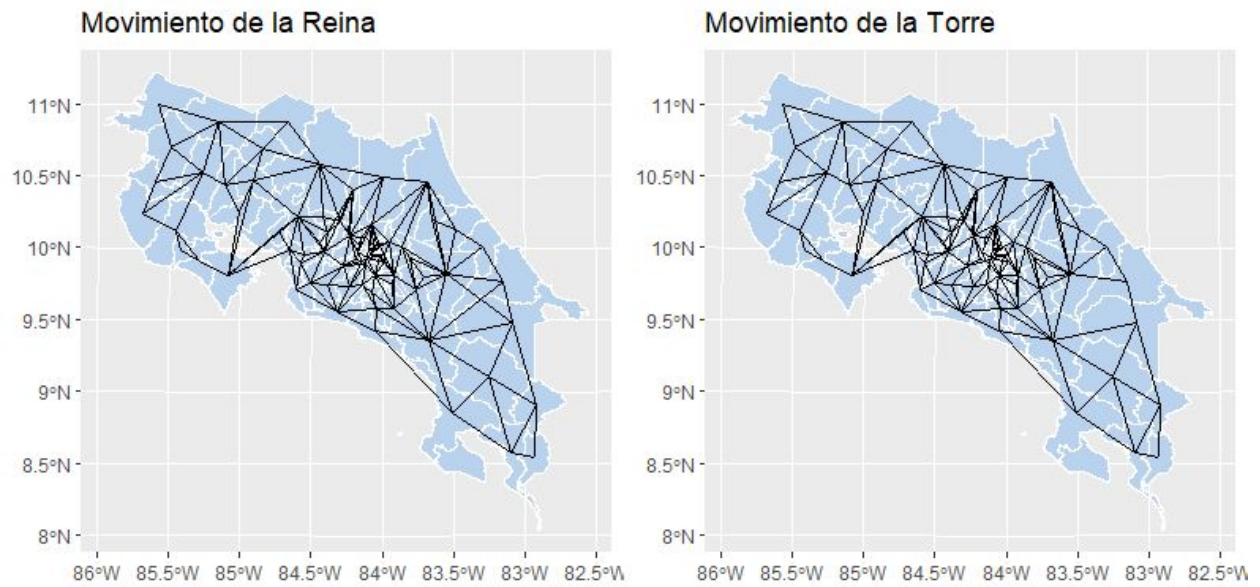
Al aplicar también el modelo SAR se obtuvo un AIC menor al obtenido al realizar la regresión lineal lo que sugiere que incorporar el componente espacial mejora respecto a los resultados obtenidos con la regresión lineal, la debilidad del modelo SAR es que también presenta dependencia espacial de los residuos.

Bibliografía

- Arias et al. (2011). Transformación productiva y desigualdad en Costa Rica. Revista Ciencias Economicas.
- Bivand, Roger, Pebesma, Edzer, Gómez-Rubio, Virgilio (2008). Applied Spatial Data Analysis with R. New York, Springer.
- Stern Scott, Epner Tamar. 2019 social progress index, methodology summary.

Anexos

Anexo 1: Matriz de pesos



Anexo 2: Salidas Modelos Espaciales

CAR W Reina

Indicador	Parámetro	Std. Error	Valor Z	Pr(> z)
Intercepto	- 0.238	0.155	- 1.538	0.124
Tasa_de_robos._asaltos_y_hurtos	- 0.391	0.078	- 5.033	0.000
Habitantes_por_EBAIS	0.165	0.070	2.370	0.018
Cobertura_del_servicio_de_recolección_de_residu	0.207	0.074	2.806	0.005
Acceso_a_energía_eléctrica	0.168	0.075	2.249	0.025
tasa_empleo_formal	0.215	0.090	2.406	0.016
Acceso_a_Conocimientos_Básicos	0.260	0.066	3.950	0.000

CAR W Torre

Indicador	Parámetro	Std. Error	Valor Z	Pr(> z)
Intercepto	- 0.200	0.176	- 1.136	0.256
Tasa_de_robos._asaltos_y_hurtos	- 0.386	0.076	- 5.092	0.000
Habitantes_por_EBAIS	0.155	0.069	2.253	0.024
Cobertura_del_servicio_de_recolección_de_residu	0.209	0.073	2.870	0.004
Acceso_a_energía_eléctrica	0.182	0.074	2.452	0.014
tasa_empleo_formal	0.221	0.089	2.484	0.013
Acceso_a_Conocimientos_Básicos	0.253	0.065	3.882	0.000

SAR W Reina

Indicador	Parámetro	Std. Error	Valor Z	Pr(> z)
Intercepto	- 0.038	0.189	- 0.203	0.839
Tasa de robos. asaltos y hurtos	- 0.365	0.076	- 4.806	0.000
Habitantes por EBAIS	0.202	0.065	3.114	0.002
Cobertura del servicio de recolección de residuos	0.111	0.069	1.603	0.109
Acceso a energía eléctrica	0.159	0.070	2.271	0.023
tasa empleo formal	0.135	0.090	1.505	0.132
Acceso a Conocimientos Básicos	0.299	0.062	4.813	0.000

SAR W Torre

Indicador	Parámetro	Std. Error	Valor Z	Pr(> z)
Intercepto	- 0.029	0.189	- 0.151	0.880
Tasa_de_robos._asaltos_y_hurtos	- 0.368	0.074	- 4.977	0.000
Habitantes_por_EBAIS	0.199	0.065	3.073	0.002
Cobertura_del_servicio_de_recolección_de_residu	0.116	0.069	1.677	0.094
Acceso_a_energía_eléctrica	0.164	0.070	2.349	0.019
tasa_empleo_formal	0.123	0.089	1.395	0.163
Acceso_a_Conocimientos_Básicos	0.305	0.062	4.890	0.000