



Universidad de Costa Rica
Sistema de Estudios de Posgrado
Programa de Posgrado en Estadística

SP-1649 Tópicos de Estadística Espacial Aplicada

INFORMACIÓN GENERAL

Plan de estudios al que pertenece el curso:	730501 Maestría Académica en Estadística y 730506 Maestría Profesional en Estadística
Tipo de curso:	Optativo para la Maestría Académica en Estadística y Optativo Maestría Profesional en Estadística
Modalidad:	Teórico
Número de créditos:	4 créditos
Horas presenciales:	4 horas semanales
Horario:	<i>Jueves de 17:15 a 21:00 virtual, Virtual Alto (100%)</i>
Aula:	<i>vía Zoom</i>
Horas de atención de estudiantes:	2 horas semanales, con cita previa
Requisitos:	No tiene
Correquisitos:	R y lectura en inglés
Ciclo lectivo:	<i>II-2020</i>
Profesor(a):	<i>Dra. Marcela Alfaro Córdoba</i>



PROGRAMA

Justificación

La estadística espacial es un área relativamente nueva, y que agrupa diferentes métodos estadísticos para estudiar la localización geográfica de eventos (como, por ejemplo, personas, nitrógeno, árboles, células, pobreza, precipitación). En algunas áreas de aplicación, el simple hecho de visualizar los eventos en un mapa o sistema de coordenadas puede ser de gran utilidad e inclusive puede ayudar a desarrollar hipótesis científicas. Sin embargo, hay casos donde la distribución espacial puede contener aún mayor información. De forma muy simplificada, se puede decir que el objetivo de los métodos estadísticos espaciales es entender los procesos aleatorios detrás de la distribución geográfica de los eventos, o estudiar el efecto que la distribución geográfica tiene sobre otros procesos.

El área de estudio puede dividirse en tres grandes temas, de acuerdo al tipo de datos recolectados: discretos, continuos o agregados área. El tipo de problema que se puede solucionar, el análisis estadístico y la interpretación correspondiente difieren entre estas tres áreas. El curso está diseñado como curso teórico-práctico de 4 créditos y opcional dentro de la maestría profesional en Estadística.

Objetivo general

El objetivo general del curso es que el estudiante se familiarice con el análisis estadístico espacial, los problemas que se pueden estudiar dentro de esta área, y los métodos estadísticos existentes. Además, identificar las tres grandes sub-áreas, y comprender y aplicar los métodos de análisis apropiados para cada una.

Objetivos específicos

Al final de curso, el estudiante será capaz de:

1. Comprender lo que se entiende como análisis estadístico espacial y su relevancia.
2. Identificar el tipo de problemas donde se puede llevar a cabo un análisis estadístico espacial.
3. Identificar el tipo de métodos que se pueden aplicar dependiendo del tipo de los datos recolectados.
4. Aplicar métodos para el análisis estadístico espacial de datos ubicados en localizaciones discretas.



5. Aplicar métodos para el análisis estadístico espacial de datos distribuidos en forma continua en el espacio.
6. Aplicar métodos para el análisis estadístico espacial de datos recolectados en forma agregada por áreas.
7. Interpretar los resultados del análisis estadístico espacial.

Descripción del curso

- I. **Introducción:** Por qué es importante incorporar información espacial; cuál es el valor agregado de un análisis espacial; cuándo y dónde se puede utilizar estadística espacial; cuál es la conexión con los sistemas de información geográfica (SIG).
- II. **Manejo de datos espaciales en R:** Clases de datos, visualización, importación, exportación y proyecciones.
- III. **Procesos puntuales (PP):** Se utiliza en situaciones donde la localización del objeto (por ejemplo, una persona o una medición) es aleatoria; por ejemplo, árboles, viviendas, células. Uno de los desafíos puede ser identificar aglomeraciones entre los puntos. Ejemplos de áreas de aplicación: epidemiología, ciencias forestales, microbiología. Tópicos: uso de las funciones F, G y K para describir PP; interacción entre dos PP; detección de aglomeraciones (*clusters*); casos con una sola fuente puntual; análisis de PP con marcas, donde los puntos contienen información adicional.
- IV. **Geoestadística:** En este caso la información de interés está distribuida en forma continua en diferentes concentraciones sobre un área geográfica. Los datos que se recolectan consisten en mediciones llevadas a cabo en puntos distribuidos sobre el área. Un primer desafío puede ser interpolar los datos para estimar la concentración sobre el área, por ejemplo, los niveles de nitrógeno en la tierra. Áreas de aplicación: medio ambiente, geología, hidrología. Tópicos: autocorrelación espacial; variogramas; *kriging*; análisis basado en modelos lineales generalizados (GLM) y GAMs.
- V. **Estadística de áreas:** En algunos casos la información solo está disponible agregada para áreas, por ejemplo, por municipalidades, países o rejillas. Ejemplos de áreas de aplicación: salud pública, sociología, economía. Suavización (*smoothing*), aglomeraciones, autoregresión espacial, regresión espacial.

Metodología

Dentro de cada tema, el material se impartirá con ayuda de presentaciones, lectura de textos extraídos de libros, y ejemplos prácticos. Además, se distribuirán artículos científicos relacionados con el tema para lectura y análisis; los artículos se escogerán de diferentes áreas de aplicación. Los estudiantes además



realizarán un proyecto final comprensivo. Este curso requiere una moderada cantidad de trabajo práctico por parte del estudiante.

El curso se administrará mediante Github (<https://malfaro2.github.io/SP1649-II20/>) y la aplicación de mensajería Slack. Los y las estudiantes se inscribirán en el curso e interactuarán con la profesora mediante estas plataformas para desarrollar sus trabajos, allí ubicarán sus entregas y tendrán disponible el material del curso. Resulta imprescindible disponer de un acceso internet al que puedan acceder los estudiantes, hacer las consultas respectivas, comunicarse con la profesora, recibir y enviar los materiales del curso. Se espera que cada estudiante tenga acceso a una computadora personal para desarrollar las prácticas en su hogar.

Las áreas de aplicación se adaptarán a los intereses de los estudiantes, que además tendrán la oportunidad de trabajar con datos propios. Las cuatro horas de lecciones semanales se dividirán en 2: 2 horas de clase presencial y 2 horas de laboratorio virtual, más tareas y el proyecto final asignado.

Cronograma

I(1), II(3), III(2), IV(5), V(3). El número de semanas de cada tema (entre paréntesis) es un valor estimado.

Bibliografía

Libro de texto:

Bivand, R.S., Pebesma, E.J. Gómez-Rubio, V. (2008) Applied Spatial Analysis with R. Springer Science.

Otras referencias:

Banerjee, S., Carlin, B., Gelfand, A.E. (2015). Hierarchical Modeling and Analysis of Spatial Data. CRC Press.

Chun, Y. And Griffith, D. (2013). Spatial Statistics & Geostatistics. GIST.

Cressie, N.A.C. (1993). Statistics for Spatial Data. Wiley.

Diggle, P.J. (2003). Statistical Analysis of Spatial Point Patterns. (2nd edition). Edward Arnold.

Gelfand A.E., Diggle, P.J, Fuentes, M., Guttorp, P. (editors) (2010) Handbook of Spatial Statistics. Chapman & Hall/CRC.

Lawson, A.B., Biggeri, A., Böhning, D., Lesaffre, E., Viel, J-F, Bertolinni, R. (editors). (1999) Disease Mapping and Risk Assessment for Public Health. Wiley.



Petris, G., Petrone, S., Campagnoli, P. (2009). Dynamic Linear Models with R. Springer Science.

Waller, L.A. and Gotway, C.A. (2004). Applied Spatial Statistics for Public Health Data. Wiley.

Referencias en línea:

<http://happygitwithr.com/>

<https://www4.stat.ncsu.edu/~reich/st733/Syllabus/>

<https://bookdown.org/robinlovelace/geocompr/>

<https://www.springer.com/gp/book/9781461476177>

<http://pakillo.github.io/R-GIS-tutorial/>

<http://www.rspatial.org/cases/index.html>

<https://labscn-unalmed.github.io/estadistica-espacial/index.html>

Evaluación

La evaluación será mediante tareas y proyectos. Los proyectos serán definidos por la profesora en correspondencia a las propuestas de los y las estudiantes. Para este propósito la comunicación en clase y los correos electrónicos son muy relevantes para concretar los proyectos, su alcance y desarrollo. El curso se plantea evaluar mediante la aplicación práctica de los conocimientos adquiridos a situaciones particulares.

Rubro	Porcentaje
Laboratorio 1	12.5%
Laboratorio 2	12.5%
Laboratorio 3	12.5%
Laboratorio 4	12.5%
Laboratorio 5	12.5%
Proyecto Final	30%
Avances del proyecto final (3)	7.5%
TOTAL	100%



No.	Semana	Tema	Evaluaciones	Observaciones
1	11 agosto	Introducción – Teoría Explicar el proyecto final del curso		
2	18 agosto	Manejo de Datos Espaciales en R Flujos de Trabajo	Lab 1	
3	25 de agosto	Manejo de Datos Espaciales en R Flujos de Trabajo	Lab 1 + Avance 1	Entrega 1 del proyecto: propuesta de tema
4	1 de setiembre	Manejo de Datos Espaciales en R Flujos de Trabajo	Lab 1	
5	8 setiembre	Procesos Puntuales	Lab 2	Entrega laboratorio 1
6	15 setiembre	Procesos Puntuales	Lab 2	
7	22 setiembre	Geoestadística	Lab 3	Entrega laboratorio 2
8	29 setiembre	Geoestadística	Lab 3 + Avance 2	Entrega 2 del proyecto: descripción de los datos
9	6 octubre	Geoestadística	Lab 3	
10	13 octubre	GAMs	Lab 4	Entrega laboratorio 3
11	20 octubre	GAMS	Lab 4 + Avance 3	Entrega 3 del proyecto: al menos dos métodos bien definidos
12	27 octubre	Estadística de Áreas Asignar laboratorio 5	Lab 5	Entrega laboratorio 4
13	3 noviembre	Estadística de Áreas	Lab 5	
14	10 noviembre	Estadística de Áreas	Lab 5	
15	17 noviembre	Preguntas y respuestas del proyecto		Entrega laboratorio 5
16	24 de noviembre	Presentación de Proyectos	Proyecto Final	Entrega del proyecto y presentación