

*Test de aleatoriedad para procesos puntuales espaciales  
basado en el cálculo de la dimensión fractal*

YOLANDA BEATRIZ CABALLERO PÉREZ



UNIVERSIDAD NACIONAL DE COLOMBIA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE ESTADÍSTICA  
BOGOTÁ, D.C.  
MAYO DE 2017

*Test de aleatoriedad para procesos puntuales espaciales  
basado en el cálculo de la dimensión fractal*

YOLANDA BEATRIZ CABALLERO PÉREZ

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE  
MAGÍSTER EN CIENCIAS-ESTADÍSTICA

DIRECTOR  
RAMÓN GIRALDO HENAO, PH.D.  
DOCTOR EN ESTADÍSTICA



UNIVERSIDAD NACIONAL DE COLOMBIA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE ESTADÍSTICA  
BOGOTÁ, D.C.  
MAYO DE 2017

**Título en español**

Test de aleatoriedad para procesos puntuales espaciales basado en el cálculo de la dimensión fractal

**Title in English**

*Randomness test for spatial point processes based on the fractal dimension*

**Resumen:** El estudio de procesos puntuales espaciales un área de la estadística espacial permite modelar y analizar la distribución de eventos que ocurren en una región. En la práctica, el primer paso para su análisis consiste en evaluar la hipótesis de aleatoriedad espacial completa (AEC). En este trabajo se propone una prueba basada en la dimensión fractal por el método de *Box Counting* para evaluar dicha hipótesis. Los resultados encontrados, con datos simulados y reales indican que la prueba tiene un buen desempeño.

**Abstract:** *The study of spatial point patterns, as an area of spatial statistics, allows modelling and analyzing the distribution of events recorded in a region of interest. In practical applications, the first step for analysis is evaluate the hypothesis of complete spatial randomness (CSR). In this work a test based on fractal dimension by Box Counting method is proposed in order to evaluate that hypothesis. The results shows an good performance with real and simulated datasets.*

**Palabras clave:** Patrones puntuales, dimensión fractal, aleatoriedad, estadística espacial.

**Keywords:** *Point patterns, fractal dimension, randomness, spatial statistics.*

# Nota de aceptación

---

Jurado  
Martha Patricia Bohorquez Castañeda, Ph.D

---

Jurado  
Carlos Eduardo Alonso Malaver, Ph.D

---

Director  
Ramón Giraldo Henao, Ph.D

Bogotá, D.C., Octubre de 2017

---

---

## Dedicado a

---

---

A la memoria de Luis Fernando Rozo.

A Juan Camilo, Constanza y Michel.

---

---

## Agradecimientos

---

---

A Ramón Giraldo por tanto tiempo dedicado a este documento.

A los amigos que conocí en el camino y a Sergio. Sin todos ellos este viaje habría sido imposible.

---

---

# Índice general

---

---

|   |            |
|---|------------|
| <b>Índice general</b>   | <b>I</b>   |
| <b>Índice de tablas</b>                                       | <b>III</b> |
| <b>Índice de figuras</b>                                      | <b>IV</b>  |
| <b>Introducción</b>   | <b>IX</b>  |
| <b>1. Marco Teórico</b>                                       | <b>1</b>   |
| 1.1. Patrones Puntuales Espaciales . . . . .                  | 1          |
| 1.1.1. Generalidades . . . . .                                | 1          |
| 1.1.2. Patrones Puntuales Espaciales . . . . .                | 3          |
| 1.1.2.1. Intensidad homogénea . . . . .                       | 4          |
| 1.1.2.2. Estacionariedad, isotropía y ergodicidad . . . . .   | 5          |
| 1.1.3. Proceso homogéneo de Poisson y AEC . . . . .           | 5          |
| 1.1.4. Pruebas de Monte Carlo en patrones puntuales . . . . . | 6          |
| 1.1.5. Pruebas para evaluar la hipótesis de AEC . . . . .     | 7          |
| 1.1.5.1. La función $K$ -Ripley . . . . .                     | 8          |
| 1.1.5.2. Función F: Distancia de espacio vacío . . . . .      | 11         |
| 1.1.5.3. Función G: Distancia al vecino más cercano . . . . . | 13         |
| 1.1.6. El efecto de borde . . . . .                           | 14         |
| 1.2. Fractales . . . . .                                      | 16         |
| 1.2.1. Generalidades . . . . .                                | 16         |
| 1.2.2. Ejemplos clásicos . . . . .                            | 17         |
| 1.2.2.1. El conjunto de Cantor . . . . .                      | 17         |
| 1.2.2.2. La curva de Koch . . . . .                           | 18         |
| 1.2.2.3. Triángulo de Sierpinski . . . . .                    | 18         |

---

|   |           |
|---|-----------|
| 1.2.3. Autosimilaridad e invarianza de escala . . . . .   | 20        |
| 1.2.4. Dimensión de similaridad . . . . .   | 21        |
| 1.2.4.1. Cálculo de la dimensión de similaridad para algunos con-<br>juntos fractales . . . . .                         | 22        |
| 1.2.5. Dimensión <i>Box Counting</i> . . . . .  | 23        |
| <b>2. Prueba de AEC basada en dimensión <i>Box Counting</i></b>   | <b>28</b> |
| 2.1. Desarrollo teórico de la prueba de AEC basada en la dimensión <i>Box Counting</i>                                  | 28        |
| 2.1.1. Cálculo de la dimensión fractal por el método <i>Box Counting</i> bajo la<br>hipótesis de AEC . . . . .          | 33        |
| 2.2. Dimensión <i>Box Counting</i> para un patrón puntual espacial regular . . . . .                                    | 35        |
| 2.3. Comparación del comportamiento de la curva bajo AEC con la recta teórica<br>de un patrón puntual regular . . . . . | 37        |
| 2.4. Código e interpretación de la prueba gráfica basada en dimensión <i>Box Coun-<br/>        ting</i> . . . . .       | 37        |
| 2.4.1. Prueba de AEC basada en dimensión <i>Box Counting</i> y el efecto de<br>borde . . . . .                          | 39        |
| <b>3. Estudio de Simulación</b>   | <b>40</b> |
| 3.1. Patrón puntual homogéneo de <i>Poisson</i> . . . . .   | 40        |
| 3.2. Simulación de un patrón de inhibición . . . . .  | 47        |
| 3.3. Simulación de patrón agregado . . . . .  | 54        |
| 3.3.1. Patrón Puntual <i>Thomas</i> . . . . .   | 54        |
| <b>4. Aplicaciones</b>  | <b>58</b> |
| 4.1. Base de datos <i>Cells</i> . . . . .   | 58        |
| 4.2. Base de datos <i>Reedwood</i> . . . . .  | 59        |
| 4.3. Base de datos <i>Longleaf Pines</i> . . . . .  | 59        |
| <b>Conclusiones</b>   | <b>63</b> |
| <b>Trabajo futuro</b>   | <b>64</b> |
| <b>Función <i>BoxCountingDim</i></b>  | <b>65</b> |
| <b>Bibliografía</b>   | <b>67</b> |

---

---

## Índice de tablas

---

---

- 2.1. Variación de la primera dimensión fractal  $D_{B1}$  (2.13), para diferentes tamaños de  $n$ . . . . . 34
- 2.2. Valores resultado de la prueba de AEC basada en dimensión *Box Counting*, para el caso de un patrón homogéneo de Poisson con valores  $n = 100$  y  $k = 1$ . 39
- 3.1. Valores del resultado de la prueba basada en *Box Counting*, para patrones bajo AEC, con diferentes tamaños de muestra,  $n = 39,466,4895$  y  $498953$ , se puede apreciar el aumento de la primera dimensión fractal ( $D_{B1}$ ) a medida que el tamaño de muestra es mayor. . . . . 42

---



---

## Índice de figuras

---



---

|  |    |
|--|----|
| 1.1. Los conjuntos de datos “Cells” y “Reedwood”, contenidos en la librería <i>Spatstats</i> [9]. . . . .  | 2  |
| 1.2. Conjunto de datos <i>Shapley</i> , contenido en la librería <i>Spatstat</i> [9]. Representa la posición (ascensión derecha, declinación) de 4215 galaxias observadas en la concentración <i>Shapley</i> . . . . .   | 3  |
| 1.3. Conjunto de datos <i>Longleaf Pines</i> [23], consiste en la información de coordenadas y diámetros a la altura del pecho (representado por el tamaño del círculo en la gráfica) de 584 pinos en un área cuadrada de 200m x 200m de bosque en el condado Thomas (Georgia, USA). . . . .   | 4  |
| 1.4. Simulación de un patrón puntual bajo AEC con $n = 100$ . . . . .  | 7  |
| 1.5. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función $K$ para cuatro conjuntos de datos. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo $[\hat{K}_{lo}(r), \hat{K}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado $\hat{K}_{obs}(r)$ y en color rojo la curva teórica bajo AEC $\hat{K}_{theo}(r)$ . . . . . | 10 |
| 1.6. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función $F$ para cuatro conjuntos de datos. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo $[\hat{F}_{lo}(r), \hat{F}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado $\hat{F}_{obs}(r)$ y en color rojo la curva teórica bajo AEC $\hat{F}_{theo}(r)$ . . . . . | 12 |
| 1.7. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función $G$ para cuatro conjuntos de datos. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo $[\hat{G}_{lo}(r), \hat{G}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado $\hat{G}_{obs}(r)$ y en color rojo la curva teórica bajo AEC $\hat{G}_{theo}(r)$ . . . . . | 15 |
| 1.8. En la figura se muestra un elemento del patrón puntual cuyos vecinos más cercanos se encuentran fuera de la ventana de observación $A$ . Las funciones estimadas basadas en distancia no serían insesgadas (Elaboración propia). . . . .  | 16 |
| 1.9. Primeras cinco iteraciones del Conjunto de Cantor. . . . .  | 17 |
| 1.10. Primeras cuatro iteraciones en la construcción de la Curva de Koch. . . . .  | 19 |
| 1.11. Construcción del Copo de Nieve de Koch. . . . .  | 20 |

|  |    |
|--|----|
| 1.12. Primeras cinco iteraciones del fractal <i>Sierpinski Gasket</i> . . . . .  | 21 |
| 1.13. División de segmento de línea, unidad de área y volumen en elementos autosimilares de lado $\epsilon$ , tomada de [1]. . . . .   | 22 |
| 1.14. Lake Powell desde la estación espacial EarthKAM Imagen: NASA/EarthKAM.org (Agosto 4, 2017) . . . . .   | 24 |
| 1.15. Aplicación de la metodología <i>Box Counting</i> para la curva de Koch, se varían los valores de $\delta$ , el valor $\Gamma(\delta)$ corresponde al conteo de las celdas de lado $\delta$ que contienen algún segmento de la curva (Elaboración propia). . . . .  | 26 |
| 1.16. Gráfico <i>log-log</i> para el cálculo de $D_B$ en el caso de la curva de Koch, donde ( $D_B = 1.255$ ) es la pendiente de la recta, ajustada por MCO (Elaboración propia) . . . . .   | 27 |
|  |    |
| 2.1. Se puede dividir la ventana de observación $A$ de lado $k$ en $\beta_i$ cuadrados de lado $\delta_i$ que no se traslapan. . . . .   | 29 |
| 2.2. Gráfico <i>log-log</i> bajo AEC obtenido a través de la ecuación (2.8), se realizan cuatro curvas manteniendo constante $n = 100$ y con valores $k=1,10,100$ y $1000$ . . . . .   | 31 |
| 2.3. La curva representa el gráfico <i>log-log</i> obtenido a través de la ecuación (2.8), para un patrón puntual bajo AEC con $n = 100$ y $k = 1$ . El punto representa los valores cuando $\tau_i = 1$ , calculados según (2.11) y (2.12), el valor de la pendiente de la recta desde el origen hasta $\tau_i = 1$ , es la primera dimensión <i>Box Counting</i> ( $D_{B1} = 1.808$ ), calculada según (2.13). . . . . | 32 |
| 2.4. Gráfico <i>log-log</i> bajo AEC obtenido a través de la ecuación (2.8), se dibujaron cuatro curvas manteniendo constante $k = 1$ y con diferentes tamaños de muestra $n=10,100,1000$ y $10000$ . . . . .  | 33 |
| 2.5. Simulación de un patrón puntual regular, en una ventana de $1 \times 1$ con $n = 100$ , el comportamiento se asemeja a un patrón de inhibición. . . . .   | 36 |
| 2.6. Gráfico <i>log-log</i> obtenido a través de la ecuación (2.17) para un patrón puntual totalmente regular, se dibujan cuatro curvas manteniendo constante $n = 100$ y con diferentes valores $k=1,10,100$ y $1000$ . . . . .   | 37 |
| 2.7. El gráfico muestra la comparación entre la recta <i>log-log</i> para un patrón regular, calculada a través de la ecuación (2.17) y la curva bajo AEC obtenida a través de la ecuación (2.8), ambos casos con $k = 1$ y $n = 100$ . . . . .  | 38 |
| 2.8. Resultado de la prueba gráfica para evaluar la hipótesis de AEA, para el caso de un patrón regular con $k = 1$ y $n = 100$ , la curva observada se ubica por encima de las bandas de confianza, por tanto se rechaza la hipótesis nula. 39  |    |
|  |    |
| 3.1. Simulaciones de patrón puntual aleatorio (i.e. homogéneo de <i>Poisson</i> ), con diferentes tamaños de muestra $n=39,466,4895$ y $498953$ . . . . .  | 41 |
| 3.2. Gráfica de la funciones de intensidad para los casos de patrones aleatorios simulados en la Figura 3.1, con diferentes tamaños de muestra $n=39,466,4895$ y $498953$ . . . . .  | 42 |

3.3. Prueba gráfica para evaluar la hipótesis de AEC a través de la dimensión *Box Counting*, para los casos simulados en la Figura 3.1, con  $n=39,466,4895$  y  $498953$ . En todos los casos la curva observada se mantiene al interior de las bandas de confianza, por tanto, se concluye que no hay evidencia para rechazar la hipótesis nula. . . . . 43

3.4. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para cuatro simulaciones de patrón puntual de *Poisson*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{F}_{lo}(r), \hat{F}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ . . . . . 44

3.5. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para cuatro simulaciones de patrón puntual de *Poisson*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{G}_{lo}(r), \hat{G}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ . . . . . 45

3.6. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para cuatro simulaciones de patrón puntual de *Poisson*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{K}_{lo}(r), \hat{K}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ . 46

3.7. Simulaciones de patrón puntual de inhibición utilizando el Modelo II de Matérn con diferentes parámetros de intensidad ( $\lambda$ ) y distancia de inhibición ( $r_i$ ). . . . . 48

3.8. Gráficas de la función de intensidad estimada para las simulaciones de patrón de inhibición de la Figura 3.7. . . . . 49

3.9. Resultados de la prueba gráfica para evaluar la hipótesis de AEC, a través de dimensión *Box Counting*, para los casos de patrones de inhibición simulados en la Figura 3.7, con diferentes tamaños de muestra. En todos los casos se rechaza la hipótesis nula, teniendo en cuenta que la curva observada se ubica por encima de las bandas de confianza. . . . . 50

3.10. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para cuatro simulaciones de patrón puntual de inhibición *Matérn II*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{F}_{lo}(r), \hat{F}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ . . . . . 51

3.11. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para cuatro simulaciones de patrón puntual de inhibición *Matérn II*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{G}_{lo}(r), \hat{G}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ . . . . . 52

3.12. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para cuatro simulaciones de patrón puntual de inhibición *Matérn II*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{K}_{lo}(r), \hat{K}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ . . . . . 53

3.13. Las Figuras 3.13A y 3.13C representan dos simulaciones de patrón agregado *Thomas* con diferentes parámetros y su función de intensidad en las Figuras 3.13B y 3.13D respectivamente. . . . . 55

3.14. Prueba de aleatoriedad espacial para dos simulaciones de patrón puntual *cluster*,  $Thomas(\lambda, \sigma, \mu)$  con diferentes parámetros, mediante el uso de la dimensión *Box Counting*. . . . . 56

3.15. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para dos simulaciones de patrón puntual agregado *Thomas*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{F}_{lo}(r), \hat{F}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ . . . . . 56

3.16. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para dos simulaciones de patrón puntual agregado *Thomas*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{G}_{lo}(r), \hat{G}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ . 57

3.17. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para dos simulaciones de patrón puntual agregado *Thomas*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{K}_{lo}(r), \hat{K}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ . 57

4.1. Gráfica del patrón y función de intensidad para el conjunto de datos *Cells*, contenido en el paquete *Spatstat* ([7] y [55]), representa los centros de 42 células biológicas observadas mediante un microscopio en un corte histológico. 58

4.2. Contraste gráfico para evaluar la hipótesis de *AEA* a través de dimensión *Box Counting*, al conjunto de datos *Cells* [9]. La curva observada se ubica por encima de las bandas de confianza, razón por la cual se rechaza la hipótesis nula y se concluye que se trata de un patrón de inhibición. . . . . 59

4.3. Distribución espacial y función de intensidad del conjunto de datos *Redwood* contenido en el paquete *Spatstat* [7]. Representa 62 secoyas en un cuadrado de 23 metros de lado. Los colores claros muestran agrupamientos de árboles. 60

4.4. Contraste gráfico para evaluar la hipótesis de *AEA* a través de dimensión *Box Counting*, al conjunto de datos *Redwood* [9]. La curva observada se ubica por debajo de las bandas de confianza, razón por la cual se rechaza la hipótesis nula y se concluye que se trata de un patrón agregado o *cluster*. 61

- 
- 4.5. Distribución espacial y función de intensidad del conjunto de datos *Longleaf Pines* contenido en el paquete *Spatstat* [7]. Representa las coordenadas de ubicación 584 pinos en un área cuadrada de doscientos metros de lado (4 hectáreas) de bosque en el condado Thomas (Georgia, USA) [23]. . . . . 61
- 4.6. Contraste gráfico para evaluar la hipótesis de *AEA* a través de dimensión *Box Counting*, al conjunto de datos *Longleaf Pines* [9]. La curva observada se ubica por debajo de las bandas de confianza, razón por la cual se rechaza la hipótesis nula y se concluye que se trata de un patrón agregado o *cluster*. 62

---

---

## Introducción

---

---

*«El Universo está escrito en el lenguaje de las matemáticas y sus caracteres son triángulos, círculos y otras figuras geométricas, sin las cuales es humanamente imposible entender una sola de sus palabras. Sin ese lenguaje, navegamos en un oscuro laberinto.»*

Galileo

El concepto de fractal fue formalmente acuñado por primera vez en 1967 [47]. En el siglo XIX varios autores, entre ellos Poincaré, Julia, Liapunov y Cantor, definieron conjuntos fractales, siendo éste último quien propuso en 1883 el llamado “Conjunto de Cantor” o “Polvo de Cantor”.

Muchas de las formas del universo poseen geometría fractal, v.g. galaxias, neuronas, nubes, ríos, formaciones montañosas, árboles y helechos. También, muchas de las creaciones del ser humano tienen comportamiento fractal, por ejemplo la música clásica occidental tiene fluctuaciones de tono que siguen una ley de potencia [45].

Los fundamentos matemáticos de la teoría fractal han sido tratados extensamente en la literatura ([1],[10],[30],[44],[46],[57] y [58]). El tema ha sido abordado también desde la teoría de la medida ([27],[28] y [29]).

Los fractales brindan una alternativa para el análisis de datos que exhiben comportamientos aparentemente desordenados o de los que no es posible obtener información en la práctica. Por ejemplo, se podrían utilizar para caracterizar sismos de magnitudes pequeñas que no pueden ser detectados por acelerógrafos de gran precisión [62]. Las leyes de potencia y la invarianza de escala o autosimilaridad estadística pueden considerarse las propiedades más valiosas de la geometría fractal. La teoría de fractales ha sido aplicada en diversos campos del conocimiento. Por ejemplo, en las ciencias de la tierra Mandelbrot (1982) propuso técnicas para medir la dimensión fractal de objetos geográficos [48].

El comportamiento fractal se ha extendido al estudio de aglomeraciones urbanas [11]. El concepto de ciudades fractales consiste en el análisis de estructuras autosimilares al interior de las ciudades, vecindades, sectores y redes de transporte. Además de introducir una nueva visión de elementos como la expansión urbana, se proponen metodologías alternativas para realizar estimaciones de densidad poblacional en términos de análisis fractal. Dauphiné (2012) sugirió por primera vez el término Geografía Fractal, en su obra

“*Fractal Geography*” [24].

En áreas como geología y geofísica también ha sido un tema abordado. [62] se enfoca en aplicaciones de teoría del caos, análisis fractal y multifractal en ciencias de la tierra.

Una aplicación de teoría de fractales en estadística se relaciona con series de tiempo. Existen varios trabajos principalmente con aplicaciones en finanzas e hidrología ([33],[42],[43], [49] y [51]).

El análisis multifractal de datos se encuentra estrechamente ligado a la estadística espacial. Varios documentos teóricos y prácticos abordan el tema de multifractalidad en estadística espacial ([13],[14],[15],[16],[17],[18], [19] y [21]). Algunos trabajos incluyen además aplicaciones basadas en Sistemas de Información Geográfica (SIG) ([2],[20],[65] y [66]). [52] realiza una aplicación específica de multifractalidad y estadística clásica para un distrito de oro en Zarshuran (Iran). Literatura más reciente trata el análisis multifractal en geoestadística como una herramienta para determinar anomalías de precio en aglomeraciones urbanas ([39] y [40]).

Con relación a la estadística espacial, los fractales se han empleado principalmente en geoestadística con el objetivo de determinar anomalías en los datos para separarlas de la línea base.

Por otro lado, el análisis de procesos puntuales espaciales es el área de la estadística espacial que se dedica a estudiar las características de eventos que pueden ser representados de manera puntual en el espacio, así como su distribución espacial. Sus antecedentes teóricos provienen desde Poisson (1837), Erlang(1909) y Neyman(1939) en el contexto de redes telefónicas y procesos espaciales agregados o *cluster*, Janossy (1948), Khinchin (1960), Matheron (1975), Matthes, Kerstan, Mecke (1978) y Srinivasan (1974) [35]. La literatura al respecto es extensa ([41],[55] y [56]). [25] trata también el análisis de procesos puntuales espacio temporales. Por su parte, otros se enfocan además del análisis de patrones puntuales, en el uso de la librería especializada *Spatstat* ([5],[6],[9] y [7]).

Un proceso puntual espacial es un proceso estocástico, donde cada realización consiste en un conjunto de puntos en  $\mathbb{R}^d$  (generalmente en  $\mathbb{R}^2$  o  $\mathbb{R}^3$ ). Son útiles en el estudio de eventos naturales como sismos y terremotos, poblaciones de animales o plantas, partículas, células, distribuciones de galaxias o estrellas, ocurrencia de enfermedades, epidemiología y fenómenos sociales [6]. En general la distribución de eventos en una región puede ser estudiada empleando teoría de procesos puntuales espaciales.

El primer paso en el análisis de procesos puntuales espaciales, consiste en evaluar la hipótesis de Aleatoriedad Espacial Completa (AEC). Para ello existen entre otras, pruebas gráficas basadas en distancia al evento más cercano (función  $F$ ), la distancia al vecino más cercano (función  $G$ ), distancias entre eventos (función  $H$ ) y la distancia media entre eventos (función  $K$ ) [25].

---

En este trabajo se propone una prueba de AEC para patrones puntuales, basada en la dimensión fractal, particularmente en la metodología *Box Counting*.

El documento está organizado de la siguiente forma: en el capítulo 1, inicialmente se presentan conceptos básicos de la teoría de patrones puntuales espaciales (tales como estacionariedad, ergodicidad y homogeneidad) y se hace la definición de las pruebas de aleatoriedad basadas en las funciones  $F$ ,  $G$  y  $K$ . Posteriormente, se enuncian las definiciones de geometría fractal, autosimilaridad, autosemejanza, invarianza de escala, algunos ejemplos de fractales matemáticos y dimensión *Box Counting*.

En el capítulo 2 se expone la prueba propuesta en el trabajo que permite evaluar la hipótesis de AEC a través del cálculo de la dimensión fractal, aplicando la metodología *Box Counting*. También se desarrolla un algoritmo computacional para este fin.

En el capítulo 3 se lleva a cabo un estudio de simulación para evaluar el comportamiento de la prueba y comparar con las pruebas tradicionales basadas en las estadísticas  $F$ ,  $G$  y  $K$ .

En el capítulo 4 se aplica la prueba propuesta para evaluar la hipótesis de AEC en tres conjuntos de datos reales ampliamente reportados en la literatura.

El escrito finaliza con conclusiones y trabajo futuro.

---

---

# Marco Teórico

---

---

## 1.1. Patrones Puntuales Espaciales

### 1.1.1. Generalidades

El análisis de procesos puntuales espaciales es una de las tres principales ramas de la estadística espacial, junto a la geoestadística (i.e. análisis de procesos continuos espaciales) y el estudio de procesos espaciales discretos o análisis de áreas [23]. Existen muchas referencias acerca de la teoría, análisis y aplicaciones de patrones puntuales espaciales y espacio temporales ([25],[35],[41] y [55]).

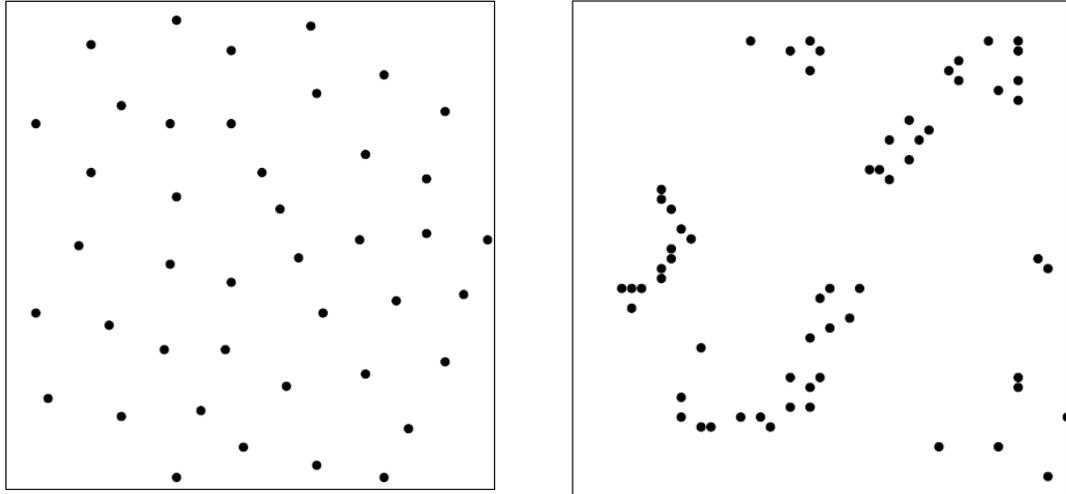
La importancia de los patrones puntuales radica en su utilidad al estudiar la distribución espacial de fenómenos que ocurren en coordenadas específicas, v.g. sismos, deslizamientos, poblaciones animales o vegetales, información epidemiológica y datos de asentamientos humanos.

Algunos ejemplos de patrones puntuales que se mencionan ampliamente en la literatura se encuentran contenidos en la librería *Spatstat* del software *R* [9]. A continuación se describen tres conjuntos de datos que muestran características esenciales.

La Figura 1.1A muestra el conjunto de datos “*Cells*” [9], que corresponde a los centros de 42 células biológicas observadas mediante un microscopio en un corte histológico. Se aprecia como la distancia entre los centros es regular, debido a que el tamaño de las células es regular. Este es un ejemplo típico de lo que se denomina patrón puntual de inhibición.

El conjunto de datos “*Reedwood*” [9], en la Figura 1.1B, representa la ubicación de 65 secoyas (i.e. especie de conífera) en un área de 23 metros de lado. En este caso es posible apreciar agrupamiento de árboles. Es un ejemplo de lo que se conoce como patrón puntual agregado o *cluster*.

La Figura 1.2 muestra el conjunto de datos *Shapley* [9]. Se refiere a la posición (ascensión derecha, declinación) de 4215 galaxias observadas en la concentración *Shapley*. En este caso no resulta trivial definir qué tipo de distribución espacial siguen las observaciones.



(A) Conjunto de datos “*Cells*” [9], representa los centros de 42 células biológicas observadas mediante un microscopio en un corte histológico.

(B) Conjunto de datos “*Reedwood*” [9], representa la ubicación de 65 secoyas en un área de 23 metros de lado.

FIGURA 1.1. Los conjuntos de datos “*Cells*” y “*Reedwood*”, contenidos en la librería *Spatstats* [9].

Cuando un patrón puntual cuenta con información adicional a la localización del evento se denomina “Patrón Puntual Marcado”, éstos son generalizaciones de procesos puntuales y son altamente relevantes en aplicaciones prácticas. A cada punto (localización)  $x_i$  se le asigna una cantidad  $m(x_i)$  (puede ser numérica o categórica) [41]. Por ejemplo, en el caso de un conjunto de sismos cada uno con localización  $x, y, z$  en  $\mathbb{R}^3$  ( $z$  representa la profundidad del epicentro), se tiene un atributo que indica su magnitud (variable de marcado). Existen también patrones puntuales espacio temporales, estos proveen además de la localización, el tiempo de ocurrencia, (v.g. localización y fechas de 100 casos consecutivos de síntomas gastrointestinales no específicos) [25] y [41].

Otro ejemplo de patrón puntual marcado es el conjunto de datos *Longleaf Pines* [23] (Figura 1.3). En éste se da la información de coordenadas y diámetros a la altura del pecho (DAP) de 584 pinos en un área cuadrada de 200m x 200m (4 hectáreas) de bosque en el condado Thomas (Georgia, USA). Las localizaciones del patrón observado se notan como  $M = x_1, \dots, x_{584}$ , y sus DAP se notan como  $m(x_1), \dots, m(x_{584})$ . El diámetro del círculo en la gráfica simboliza la variable de marcado  $m(x_i)$ .

En aplicaciones reales los patrones puntuales analizados estadísticamente son finitos. Sin embargo, tiene sentido emplear modelos que corresponden a procesos puntuales infinitos, la elección depende de la situación. En algunos casos los patrones describen fenómenos que ocurren en un área estrictamente delimitada (v.g. centros de marcas de balas en un objetivo circular, centros de los poros en la superficie de un ladrillo, centros de poros de aire en un queso Suizo). En otros casos el patrón puntual puede ser considerado como parte de un patrón mucho mayor y se puede asumir que éste se encuentra distribuido de la misma manera que la ventana de observación ([41] y [64]).

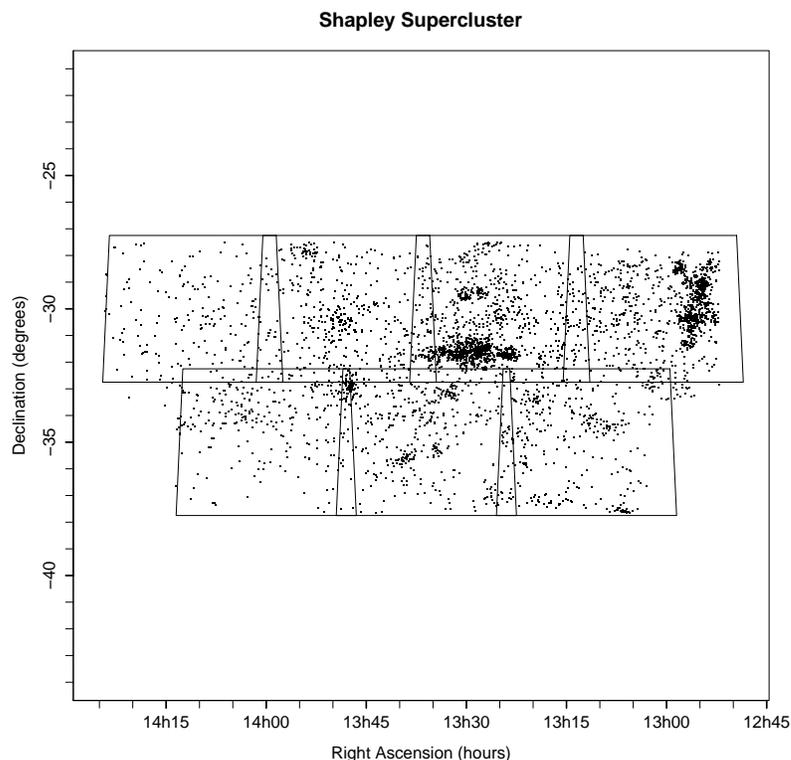


FIGURA 1.2. Conjunto de datos *Shapley*, contenido en la librería *Spatstat* [9]. Representa la posición (ascensión derecha, declinación) de 4215 galaxias observadas en la concentración *Shapley*.

Usualmente, cuando se decide analizar un patrón puntual asumiendo que es una realización de un proceso puntual infinito, se asume homogeneidad e isotropía (propiedades que serán definidas en la sección 1.1.2.2) [64].

El objetivo fundamental del estudio de patrones puntuales espaciales es identificar y modelar la distribución espacial de un conjunto de localizaciones. Es de interés particular la presencia de tendencia en las localizaciones de los puntos. Para concluir que existe agrupamiento (i.e. *clustering*), es necesario en primera medida descartar su ausencia [64].

### 1.1.2. Patrones Puntuales Espaciales

El proceso puntual es un proceso estocástico, cuya realización es un patrón puntual espacial. Este puede ser definido y estudiado en espacios abstractos, pero en general se consideran dos o tres dimensiones (por simplicidad en la presentación, sólo se asume  $\mathbb{R}^2$ ). El patrón puntual espacial es un conjunto de datos finito (colección de localizaciones) ([35],[41],[60] y [64]). También se puede definir como datos en forma de conjunto de puntos con localización irregularmente distribuida en una región espacial ([25] y [35]).

Sea un proceso puntual  $\mathbf{X}$ , donde cada realización es un patrón puntual  $\mathbf{x}$  con número finito de puntos. Se asume que  $\mathbf{X}$  se extiende a través de  $\mathbb{R}^2$  pero es observado solamente

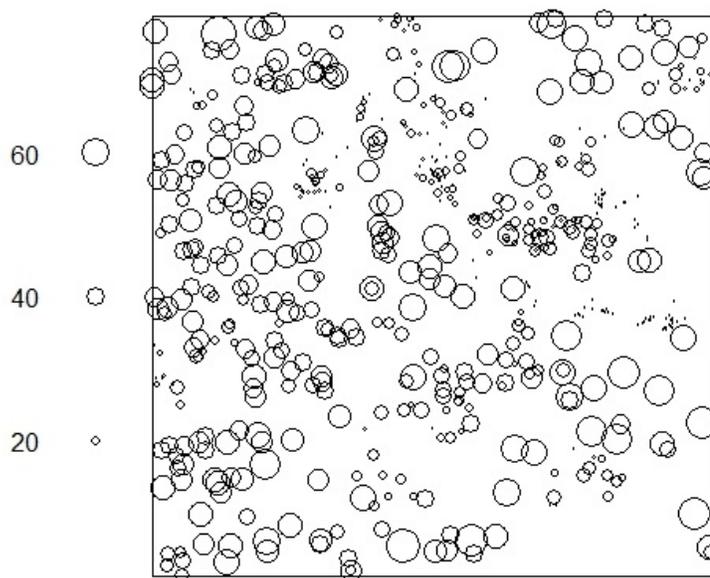


FIGURA 1.3. Conjunto de datos *Longleaf Pines* [23], consiste en la información de coordenadas y diámetros a la altura del pecho (representado por el tamaño del círculo en la gráfica) de 584 pinos en un área cuadrada de  $200m \times 200m$  de bosque en el condado Thomas (Georgia, USA).

al interior de una ventana de observación  $A$ . Se define  $\mathbf{x}$  como una muestra finita de  $\mathbf{X}$ , tal que [6]

$$\mathbf{x} = \{x_1, x_2, \dots, x_n\}, x_i \in A, n \geq 0.$$

La ventana de observación  $A$  es fija y conocida. Para cada región  $A$ , se define  $N(A)$  como la variable aleatoria del conteo de puntos de  $\mathbf{x}$  en  $A$  [6].

Un concepto importante dentro del análisis de patrones puntuales es la **intensidad**, que se define como el valor esperado del número de puntos por unidad de área. Es una medida de abundancia de los eventos. La intensidad puede ser constante a lo largo de la ventana de observación (homogénea) o puede variar de una localización a otra (no homogénea) [6].

### 1.1.2.1. Intensidad homogénea

Sea un proceso puntual  $\mathbf{X}$ , se dice que su intensidad es homogénea si, para cada sub-región  $A$  en  $\mathbb{R}^2$  el número esperado de puntos en  $A$  es proporcional al área de  $A$  denotada como  $|A|$  [6]

$$E[N(\mathbf{X} \cap A)] = \lambda|A|,$$

$\lambda$  se conoce como intensidad del proceso puntual homogéneo, es constante y corresponde al número promedio de puntos por unidad de área y  $N(\mathbf{X})$  es la variable aleatoria

que representa el número de puntos del proceso puntual en la región  $A$ . Un estimador insesgado de  $\lambda$  para un patrón puntual homogéneo ( $\mathbf{x}$ ) es la densidad empírica de puntos, definida como [6]

$$\hat{\lambda} = \frac{N(\mathbf{x})}{|A|}.$$

### 1.1.2.2. Estacionariedad, isotropía y ergodicidad

A continuación se definen tres propiedades importantes en el análisis de patrones puntuales espaciales.

#### Estacionariedad

Un proceso puntual  $\mathbf{X}$  en  $\mathbb{R}^d$ , es llamado estacionario (homogéneo) si las propiedades estadísticas de  $\mathbf{X}$  y  $\mathbf{X} + v$  son idénticas para cualquier vector  $v$  [6]. Sea  $\mathbf{x}$  un patrón puntual y  $\mathbf{x} + v = \{x_i + v : i = 1, \dots, n\}$  el resultado de trasladar cada punto de  $\mathbf{x}$  en un vector  $v$ , se puede afirmar que  $\mathbf{x}$  y  $\mathbf{x} + v$  tienen la misma distribución [41], por tanto

$$N(\mathbf{x}) \stackrel{d}{=} N(\mathbf{x} + v), \text{ para todo } \mathbf{x} \text{ y } v.$$

#### Isotropía

Un proceso puntual es isotrópico cuando es invariante a la rotación con respecto al origen [64].

#### Ergodicidad

Además de la estacionariedad e isotropía, la ergodicidad también se asume con frecuencia en estadística de procesos puntuales espaciales. Si un proceso puntual es ergódico, esto supone que analizar una muestra (un patrón puntual) de un tamaño apropiado permitirá obtener resultados estadísticamente significativos,

$$\lim_{A \uparrow \mathbb{R}^d} \frac{N(A)}{|A|} = \lambda \text{ con probabilidad 1,}$$

donde  $\lambda$  es la intensidad del proceso,  $A$  es una región tal que  $A \subset \mathbb{R}^d$ , de modo que la expresión  $A \uparrow \mathbb{R}^d$  significa que la ventana  $A$  converge al espacio completo  $\mathbb{R}^d$ . La característica numérica más importante para un proceso puntual estacionario es la intensidad  $\lambda$ , que denota la media de puntos por unidad de área o volumen ([60] y [41]).

### 1.1.3. Proceso homogéneo de Poisson y AEC

En un patrón puntual  $\mathbf{x}$  bajo AEC la variable aleatoria  $N(A)$  sigue una distribución de *Poisson*, por tanto se asimila a un proceso homogéneo de Poisson que cumple con dos propiedades fundamentales: Es homogéneo e independiente ([25] y [64]).

Por una parte, la **homogeneidad** implica que el valor esperado del número de puntos del patrón puntual que cae en una región  $A$  debe ser proporcional a su área  $|A|$ , como

$$E[N(A)] = \lambda|A|. \tag{1.1}$$

Con  $\lambda$  constante [6].

La variable aleatoria  $N(A)$  (número de puntos del patrón puntual en una región  $A$ ) sigue una distribución de *Poisson* con media  $\lambda|A|$  como en (1.2), para  $\lambda$  constante, positiva y finita (1.1),

$$P(N(A) = n) = \frac{\lambda^n |A|^n e^{-\lambda|A|}}{n!}, \text{ para } n = 0, 1, \dots \quad (1.2)$$

Por otra parte, la **independencia** implica que si  $A_1$  y  $A_2$  son dos subregiones disyuntas del área de interés  $A$ , entonces, las variables aleatorias  $N(A_1)$  y  $N(A_2)$  son independientes [6].

Dado que bajo AEC los eventos en  $A$  son i.i.d; entonces para el vector de eventos  $(x_1, \dots, x_n)$  en  $A$  satisface qué,

$$P(x_1 \in A_1, \dots, x_n \in A_n) = \prod_{i=1}^n (|A_i|/|A|), \quad A_1, \dots, A_n \subset A. \quad (1.3)$$

Esto quiere decir que cada evento tiene la misma probabilidad de ocurrir en cualquier parte contenida en el espacio  $A$ . Los eventos no se atraen (*cluster*) ni se repelen (inhibición) [23].

El proceso homogéneo de Poisson se toma como referencia y representa un estándar idealizado que es empleado como primera aproximación, por tanto es el punto de partida en el estudio de cualquier proceso puntual espacial.

La mayoría de los análisis comienzan con una prueba para evaluar la hipótesis de AEC. Si esta no se rechaza apenas merece la pena realizar otros análisis estadísticos al patrón puntual. Por otra parte, las pruebas que permiten evaluar la hipótesis de AEC, son usadas también como parte del análisis exploratorio del conjunto de datos ([25] y [41]). La Figura 1.4, muestra una simulación de un patrón homogéneo de *Poisson* en  $\mathbb{R}^2$  (AEC) con  $n = 100$ .

#### 1.1.4. Pruebas de Monte Carlo en patrones puntuales

La hipótesis de AEC puede ser evaluada usando métodos de Monte Carlo como se muestra a continuación.

Suponga una hipótesis simple  $H_0$ . Sea  $t_1$  el valor observado de un estadístico  $T$  para un conjunto de datos y sea  $\{t_2, \dots, t_s\}$  los valores de  $T$  observados para  $s - 1$ , simulaciones independientes de la distribución de  $T$  bajo  $H_0$  (se asume por conveniencia que la distribución de  $T$  es continua). Si  $H_0$  es cierta, los valores  $\{t_1, t_2, \dots, t_s\}$  son intercambiables e igualmente probables. De aquí que si  $R$  denota el número de  $t_i$  observados, tal que  $t_i > t_1$ , entonces el valor  $p$  de Monte Carlo se determina como  $P(R \leq r) = \frac{r + 1}{s}$ , esto implica que para un nivel de significancia predeterminado  $\alpha = \frac{k + 1}{s}$ , la prueba

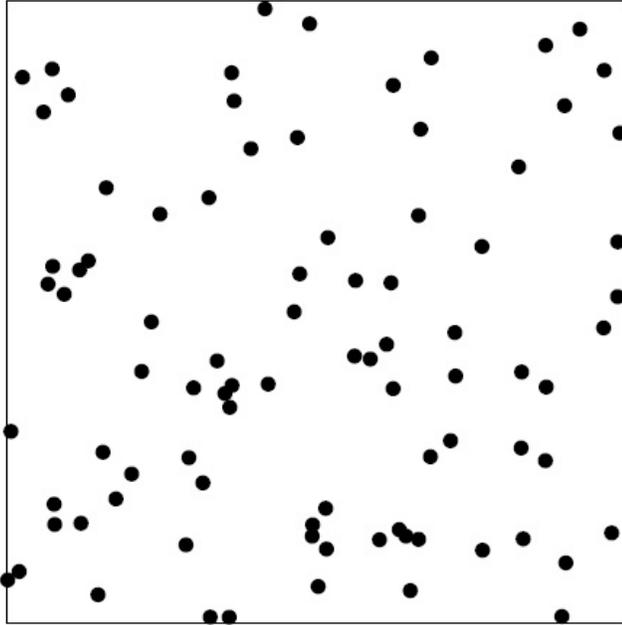


FIGURA 1.4. Simulación de un patrón puntual bajo AEC con  $n = 100$ .

rechaza  $H_0$  si y solo si  $r \leq k$  donde  $r \in \mathbb{Z}^+$  ([25] y [35]).

La pérdida de potencia resultante de una implementación de Monte Carlo es leve comparada con una prueba clásica. Según Marriot (1979), la cantidad de simulaciones  $s$  no necesita ser muy grande; para una prueba al 5% es suficiente con  $s = 100$  y para una prueba al 1% sería suficiente  $s = 500$  ([25] y [35]).

Las pruebas de Monte Carlo son basadas en simulación. Para el caso de las pruebas de AEC, se calcula la función correspondiente para  $s$  simulaciones de patrones puntuales bajo  $H_0$ : AEC, estableciéndose “bandas de confianza”. Posteriormente se calcula la estadística con base en el patrón observado. Si esta se encuentra por fuera de las bandas de confianza se rechaza  $H_0$ .

### 1.1.5. Pruebas para evaluar la hipótesis de AEC

Una prueba de hipótesis típica, compara el valor de un resumen cuantitativo de los datos observados con el calculado bajo el supuesto de que la hipótesis nula es cierta, en este caso la de AEC. En esta sección, se describen tres de las pruebas más empleadas para evaluar la hipótesis de AEC.

### 1.1.5.1. La función $K$ -Ripley

Sea  $d_{ij} = \|x_i - x_j\|$  la distancia entre dos puntos  $x_i, x_j$  del patrón puntual  $\mathbf{x}$ , donde ( $i \neq j$ ). Si el patrón puntual es agregado muchas de las distancias  $d_{ij}$  serán pequeñas. Análogamente, si se trata de un patrón puntual de inhibición pocas de las distancias  $d_{ij}$  serán pequeñas [6], [64] y [25].

Suponga un proceso puntual estacionario  $\mathbf{X}$  y defina  $t(u, r, \mathbf{x})$  como el número de puntos en el patrón  $\mathbf{x}$  que se encuentran a una distancia menor o igual a  $r$  de la localización  $u$ , es decir

$$t(u, r, \mathbf{x}) = \sum_{j=1}^n I_j, \quad I_j = \begin{cases} 1 & \text{si } 0 < \|u - x_j\| \leq r \\ 0 & \text{e.o.c.} \end{cases}. \quad (1.4)$$

La función  $K$  está definida como

$$K(r) = \frac{E[t(u, r, \mathbf{x}) | u \in \mathbf{X}]}{\lambda}.$$

La definición de la función  $K$  no depende de  $u$ , ya que se asume que  $\mathbf{X}$  es estacionario e independiente [6], por lo tanto

$$K(r) = \frac{E[t(u, r, \mathbf{X})]}{\lambda}.$$

La variable aleatoria  $t(u, r, \mathbf{X})$  es el número de puntos de  $\mathbf{X}$  que se localizan al interior del disco de radio  $r$  con centro en  $u$ . Para un proceso homogéneo de *Poisson* se puede establecer que  $E[t(u, r, \mathbf{X})] = \lambda\pi r^2$ , y la función  $K$  bajo AEC está dada por [6],[64] y [25]

$$K_{AEC}(r) = \pi r^2.$$

Para un patrón de inhibición, se esperarían menos eventos dentro de la distancia  $r$  ( $K(r) < \pi r^2$ ). Análogamente se esperaría mayor número de eventos para el caso de un proceso puntual agregado,  $K(r) > \pi r^2$  [59].

A continuación se define la función  $K$  empírica. Sea  $\hat{H}(r)$  la función de distribución empírica definida como la cantidad de valores de  $d_{ij}$  menores que  $r$ , de modo que,

$$\hat{H}(r) = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n I_{ij}, \quad \text{donde } I_{ij} = \begin{cases} 1 & \text{si } d_{ij} \leq r \\ 0 & \text{e.o.c.} \end{cases}. \quad (1.5)$$

El denominador  $n(n-1)$  es el total de pares de puntos distintos.  $\hat{H}(r)$  es el número de puntos que se encuentran dentro de un círculo de radio  $r$ . La contribución de cada evento  $x_i$  a la suma en (1.5) [6] es

$$t_i(r) = \sum_{j \neq i} I_{ij}. \quad (1.6)$$

Reemplazando (1.6) en (1.5) se obtiene

$$\hat{H}(r) = \frac{1}{n(n-1)} \sum_{i=1}^n t_i(r) = \frac{1}{n-1} \bar{t}(r). \quad (1.7)$$

El valor  $\bar{t}(r)$  es el número medio de vecinos dentro de una distancia  $r$  a cada punto. En un patrón bajo AEC se espera que  $\bar{t}(r) = \lambda\pi r^2$  (intensidad por el área del círculo de radio  $r$ ). Dado que el número máximo de vecinos que podría tener cualquier punto es  $n - 1$ , tiene sentido definir  $\lambda = (n - 1)/|A|$ . Multiplicando ambos lados de (1.7) por  $|A|$ , se tiene [6], [64] y [25]

$$|A|\hat{H}(r) = \frac{|A|}{(n - 1)}\bar{t}(r). \quad (1.8)$$

Se define la función empírica  $K$  reemplazando (1.5) en (1.8)

$$\hat{K}(r) = |A|\hat{H}(r) = \frac{|A|}{n(n - 1)} \sum_{i=1}^n \sum_{j \neq i}^n I_{ij}e_{ij}(r),$$

dónde  $e_{ij}(r)$  es una corrección del efecto de borde (sección 1.1.6).

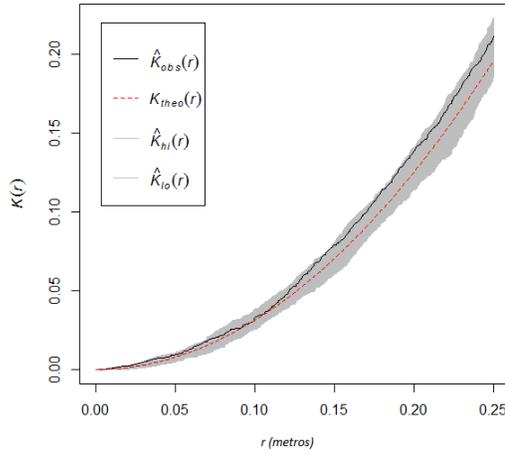
En resumen, la función  $K$  es el promedio acumulativo de puntos que caen dentro de una distancia  $r$  de un punto, corregido el efecto de borde y dividido por la intensidad [6], [64] y [25].

La gráfica de  $r$  contra  $\hat{K}(r)$  permite evaluar empíricamente la hipótesis de AEC. La Figura 1.5 muestra el resultado de la prueba para cuatro conjuntos de datos. La zona en gris determina las bandas de confianza definidas mediante simulaciones de Monte Carlo bajo AEC [ $\hat{K}_{lo}(r)$ ,  $\hat{K}_{hi}(r)$ ]. Se rechaza la hipótesis nula cuando la curva para el patrón observado  $\hat{K}_{obs}(r)$  se ubica fuera de dichas bandas. Si se ubica por debajo indica que se trata de un patrón de inhibición, mientras que, si la curva se encuentra en la parte superior de las bandas calculadas se trata de un patrón agregado. Particularmente la Figura 1.5A muestra el resultado de la prueba gráfica para el caso de una simulación de patrón puntual de *Poisson* con  $n = 103$ , se aprecia como la curva observada se encuentra al interior de las bandas de confianza, por tanto no hay evidencia que permita rechazar la hipótesis nula.

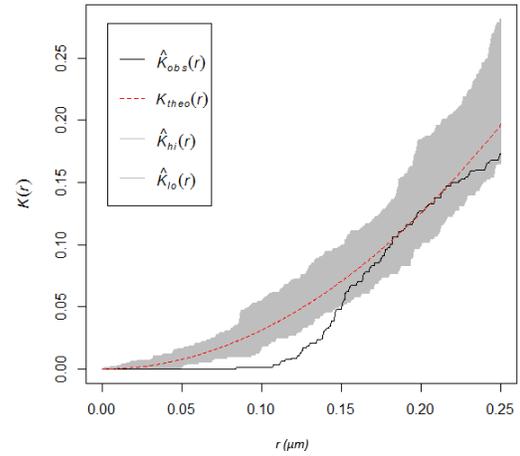
La Figura 1.5B, describe el resultado de la prueba gráfica de AEC basada en la función  $K$  para el conjunto de datos *Cells* (Figura 1.1A), siendo éste un caso típico de patrón puntual de inhibición. Como se puede comprobar en la gráfica, la curva observada de la función se encuentra fuera de las bandas de confianza, por tanto, se debe rechazar la hipótesis nula, concluyendo que se trata de un proceso de inhibición.

De manera similar la Figura 1.5C muestra el resultado de la función  $K$  para el conjunto de datos *Reedwood* (Figura 1.1B), que es un caso de patrón puntual agregado (i.e. *cluster*). La curva observada se encuentra por encima de las bandas de confianza determinadas, por lo tanto, se rechaza la hipótesis de AEC, en favor de que el patrón es agregado.

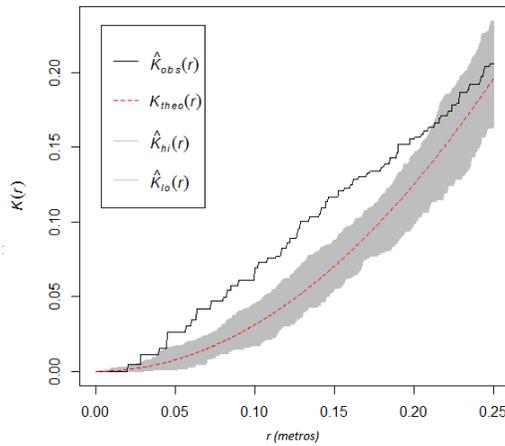
Un caso similar ocurre con el patrón puntual correspondiente a *Longleaf Pines* (Figura 1.3). En la Figura 1.5D se puede observar el resultado de la función  $K$  para este conjunto de datos, de esta se puede inferir que se debe rechazar la hipótesis nula y concluir que las observaciones corresponden a una realización de un patrón agregado.



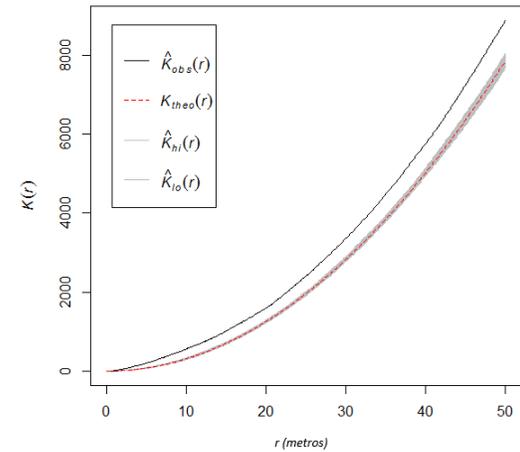
(A) Función  $K$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 103$ .



(B) Función  $K$  con bandas de confianza para el conjunto de datos *Cells* [9].



(C) Función  $K$  con bandas de confianza para el conjunto de datos *Reedwood* [9]



(D) Función  $K$  con bandas de confianza para el conjunto de datos *Longleaf Pines* [9]

FIGURA 1.5. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para cuatro conjuntos de datos. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{K}_{lo}(r)$ ,  $\hat{K}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ .

### 1.1.5.2. Función F: Distancia de espacio vacío

Sea un proceso puntual  $\mathbf{X}$ , la distancia desde una localización fija arbitraria  $u \in \mathbb{R}^2$  al punto más cercano del proceso es llamada distancia de espacio vacío, y se denota como [6]

$$d(u, \mathbf{X}) = \min \{ \|u - x_i\| : x_i \in \mathbf{X} \}.$$

Para un proceso estacionario la función de distribución acumulada de la distancia de espacio vacío se define como [6]

$$F(r) = P \{ d(u, \mathbf{X}) \leq r \}.$$

Se asume que  $\mathbf{X}$  es estacionario, por lo tanto F no depende de  $u$ . La distancia al vecino más cercano de  $\mathbf{X}$  es mayor que  $r$ , si y solo si el disco de radio  $r$  no contiene puntos de  $\mathbf{X}$ ,

$$d(u, \mathbf{X}) > r \text{ sii } N(\mathbf{X} \cap b(u, r)) = 0,$$

donde  $b(u, r)$  es el disco de radio  $r$ , con centro en  $u$  [6].

Si  $\mathbf{X}$  es un proceso homogéneo de *Poisson* en  $\mathbb{R}^2$  de intensidad  $\lambda$ , el número de puntos que ocurren en un área  $A$  tiene distribución de *Poisson* con media  $\lambda|A|$ . Si  $B$  es un disco de radio  $r$ , entonces  $|A| = \pi r^2$ . La función F bajo AEC se define como la probabilidad de que no caigan puntos en  $B$  [6]

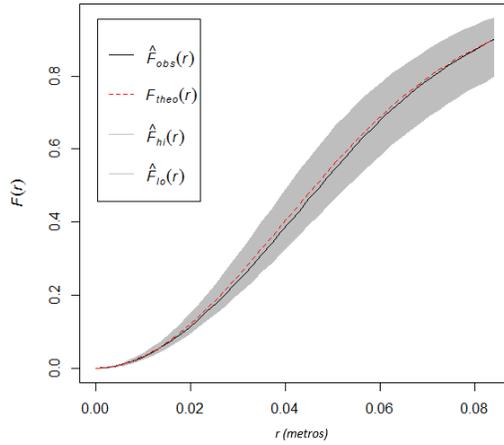
$$\begin{aligned} F_{AEC}(r) &= 1 - P \{ d(u, \mathbf{X}) \} = P \{ N(\mathbf{X} \cap b(u, r)) = 0 \} \\ &= 1 - e^{-\lambda\pi r^2}. \end{aligned}$$

Bajo AEC, se define el valor esperado de la distancia de espacio vacío  $E[d(u, \mathbf{X})] = 1/(2\sqrt{\lambda})$  y su varianza  $Var[d(u, \mathbf{X})] = (1/\pi - 1/4)/\lambda$  [6].

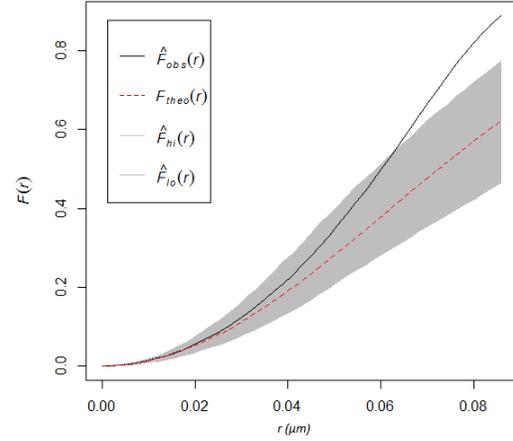
La gráfica de  $r$  contra  $F(r) = 1 - e^{-\lambda\pi r^2}$  permite evaluar empíricamente la hipótesis de AEC. La Figura 1.6 ilustra el resultado para diferentes conjuntos de datos. Cuando se trata de un patrón puntual bajo AEC la curva observada se ubica al interior de las bandas de confianza en color gris (Figura 1.6A). Para el caso de un patrón agregado, los espacios vacíos son mayores que en para el caso de un patrón bajo AEC y de inhibición. La curva por encima de las bandas de confianza definidas mediante una prueba de Monte Carlo permite concluir que se trata de un patrón inhibición y una curva por debajo de las bandas de confianza corresponde a un patrón puntual agregado (i.e. *cluster*) [25].

La gráfica de  $r$  contra  $\hat{F}(r)$  permite evaluar empíricamente la hipótesis de AEC. La Figura 1.6B muestra el resultado de la prueba gráfica de AEC basada en la función  $F$  para el conjunto de datos *Cells*. La parte superior de la curva observada se encuentra por encima de las bandas de confianza. Se rechaza la hipótesis nula, se trata de un patrón puntual de inhibición.

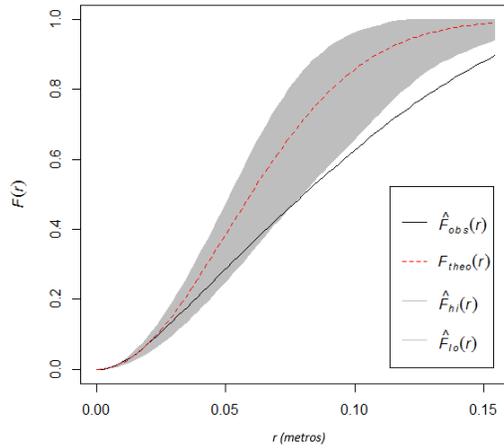
En el caso del conjunto de datos *Reedwood*, la curva observada se encuentra por debajo de las bandas de confianza, como se puede apreciar en la Figura 1.6C. Se rechaza



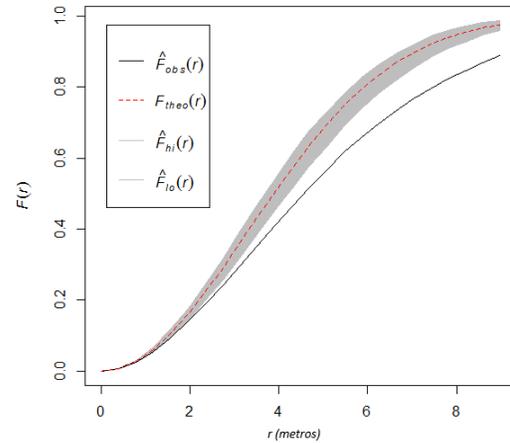
(A) Función  $F$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 103$ .



(B) Función  $F$  con bandas de confianza para el conjunto de datos *Cells* [9].



(C) Función  $F$  con bandas de confianza para el conjunto de datos *Reedwood* [9]



(D) Función  $F$  con bandas de confianza para el conjunto de datos *Longleaf Pines* [9]

FIGURA 1.6. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para cuatro conjuntos de datos. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{F}_{lo}(r)$ ,  $\hat{F}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ .

también la hipótesis de AEC y se concluye que el patrón es agregado.

Finalmente, en el caso del conjunto de datos *Longleaf Pines*; de lo mostrado en la Figura 1.6D se concluye que las observaciones corresponden a una realización de un patrón agregado.

### 1.1.5.3. Función G: Distancia al vecino más cercano

Si  $x_i$  es uno de los puntos en un patrón puntual  $\mathbf{x}$ , la distancia al vecino más cercano  $d_i = \min_{i \neq j} \|x_j - x_i\|$ , también puede escribirse como [6]

$$d_i = d(x_i, \mathbf{x}/x_i).$$

Se define como la distancia más corta desde  $x_i$  hasta el patrón  $\mathbf{x}/x_i$ , que consiste en todos los puntos de  $\mathbf{x}$  excepto  $x_i$  [6].

Para un proceso estacionario  $\mathbf{X}$  la función de distribución de la distancia al vecino más cercano se define como [6]

$$G(r) = P \{d(u, \mathbf{x}/u) \leq r | \mathbf{X} \text{ tiene un punto en } u\}; \text{ para } r \geq 0.$$

El valor  $u$  es una localización arbitraria.

Suponga  $\mathbf{X}$  un proceso uniforme de *Poisson* en  $\mathbb{R}^2$  con intensidad  $\lambda$ . Teniendo en cuenta que los eventos son independientes, la localización de  $u$  no afecta la distribución de probabilidad. Para AEC [6]

$$G(r) = P \{d(u, \mathbf{x}/u \leq r)\}.$$

Para un patrón bajo AEC la distribución de G es igual a la distribución de F y se define[6]

$$G_{AEC}(r) = 1 - e^{-\lambda\pi r^2}.$$

Si el patrón puntual es agregado muchas de las distancias  $d_i$  serán pequeñas. Análogamente, si se trata de un patrón puntual de inhibición pocas de las distancias  $d_i$  serán pequeñas [6].

Es posible definir la función  $\hat{G}(r)$  como:

$$\hat{G}(r) = \frac{\sum_{i=1}^n I_i}{n} \quad \text{donde: } I_i = \begin{cases} 1 & \text{si } d_i \in d_i : d_i \leq r, \forall_i \\ 0 & \text{e.o.c.} \end{cases}, \quad (1.9)$$

donde,  $d_i = \min d_{ij}, \forall_j \neq i \in S, i = 1, 2, \dots, n$ .

La función  $G$  representa el número de elementos del patrón con distancias menores a  $r$ , normalizadas por el total de puntos  $n$  del patrón puntual. Mediante pruebas de Monte

Carlo se genera un intervalo de confianza que permite aceptar o rechazar la hipótesis de AEC [25].

La Figura 1.7 muestra el resultado de la prueba gráfica para diferentes conjuntos de datos. En el caso de un patrón puntual bajo AEC la curva observada se ubica al interior de las bandas de confianza en color gris (Figura 1.7A). En el caso de un patrón puntual agregado, las distancias al vecino más cercano deben ser más pequeñas que lo esperado bajo condiciones de AEC (curva por encima de las bandas de confianza calculadas). Análogamente, un patrón puntual regular o de inhibición debería tener distancias mayores hasta su vecino más cercano que lo esperado para un patrón puntual bajo AEC (curva por debajo de las bandas de confianza).

La gráfica de  $r$  contra  $G(r)$  permite evaluar empíricamente la hipótesis de AEC. La Figura 1.7B, describe el resultado de la prueba gráfica de AEC basada en la función  $G$  para el conjunto de datos *Cells*, siendo éste un caso típico de patrón puntual de inhibición. La zona en gris determina las bandas de confianza definidas mediante simulaciones de Monte Carlo bajo la hipótesis de AEC. Como se puede comprobar en la gráfica, la curva observada de la función se encuentra por debajo de las bandas de confianza, por tanto, se debe rechazar la hipótesis de AEC, confirmando que se trata de un proceso de inhibición.

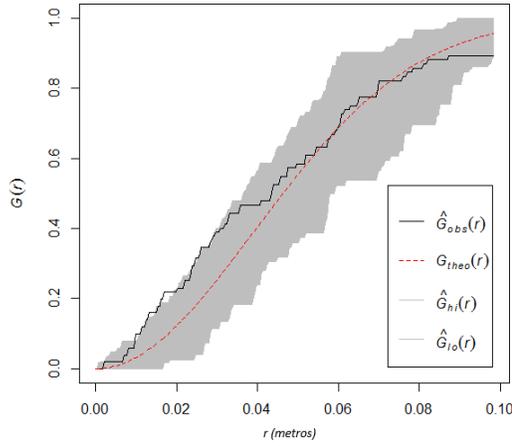
La Figura 1.7C muestra el resultado de la función  $G$  para el conjunto de datos *Reedwood*. La curva observada se encuentra por encima de las bandas de confianza determinadas. Se rechaza la hipótesis de AEC, en favor de que el patrón es agregado.

Para el caso *Longleaf Pines*. La Figura ??) muestra el resultado de la función  $G$  de este patrón. Se debe rechazar la hipótesis nula y concluir que las observaciones corresponden a una realización de un patrón agregado.

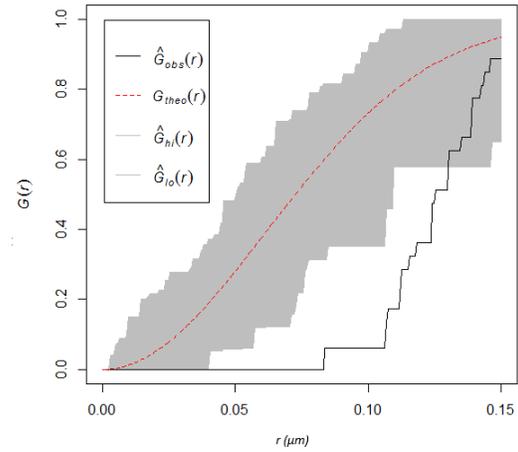
### 1.1.6. El efecto de borde

A pesar de que se asume que el patrón puntual estacionario es infinito, en la mayoría de los casos se cuenta únicamente con información al interior de una ventana de observación  $A$ . Si el conjunto de datos solo provee los puntos al interior de la ventana, la información completa del vecindario no está disponible para los puntos ubicados cerca del borde de  $A$ . Dado que las funciones  $F$ ,  $G$  y  $K$  se basan en distancias, para su estimación se necesita información afuera de la ventana de observación. La Figura 1.8 muestra un elemento del patrón puntual cuyos vecinos más cercanos se encuentran fuera de la ventana de observación  $A$ . En este caso las funciones estimadas basadas en distancia no serían insesgadas [41].

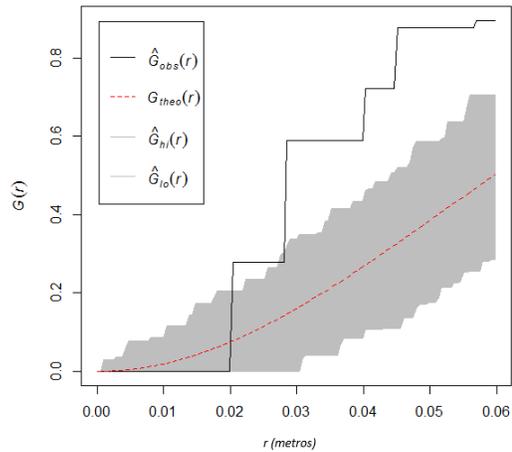
En resumen, el efecto de borde se refiere al error relacionado con la información que no es observable cuando se calcula un estadístico cercano al límite del área de interés. Existen varias metodologías para corregir el efecto de borde en los estimadores de las funciones  $F$ ,  $G$  y  $K$ . Este tema puede ser profundizado en la literatura ([41] y [6]).



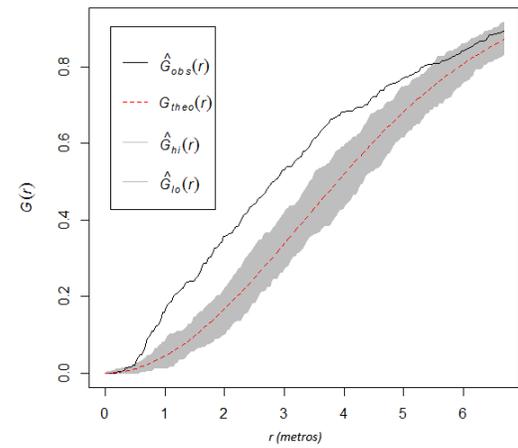
(A) Función  $G$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 103$ .



(B) Función  $G$  con bandas de confianza para el conjunto de datos *Cells* [9].



(C) Función  $G$  con bandas de confianza para el conjunto de datos *Reedwood* [9]



(D) Función  $G$  con bandas de confianza para el conjunto de datos *Longleaf Pines* [9]

FIGURA 1.7. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para cuatro conjuntos de datos. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{G}_{to}(r), \hat{G}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ .

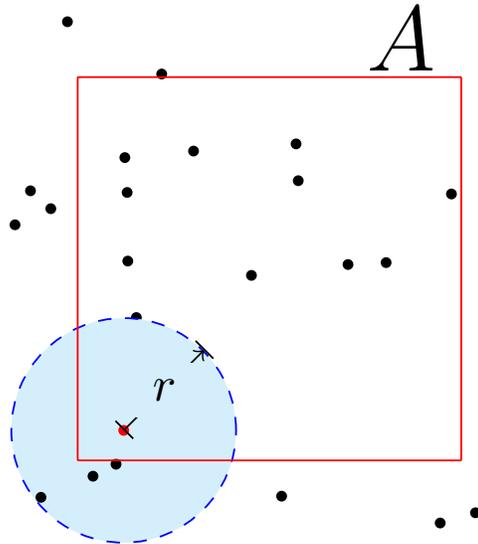


FIGURA 1.8. En la figura se muestra un elemento del patrón puntual cuyos vecinos más cercanos se encuentran fuera de la ventana de observación  $A$ . Las funciones estimadas basadas en distancia no serían insesgadas (Elaboración propia).

## 1.2. Fractales

### 1.2.1. Generalidades

*«Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line.»*

B. Mandelbrot [48]

Si bien la geometría Euclidiana puede simplificar el mundo que nos rodea no es la manera más exacta de representar los objetos naturales, la geometría de la naturaleza es sumamente compleja y pareciera en ocasiones comportarse desordenadamente.

Las figuras geométricas tradicionales tienden a simplificarse a medida que la escala aumenta. Por ejemplo, el resultado de ampliar un círculo es una curva, al acercarse más es una recta. No obstante, la naturaleza no se comporta de este modo, una montaña que de lejos podría representarse mediante un triángulo no se simplifica a medida que nos acercamos.

Mandelbrot (1967), pretendía medir con cierto grado de exactitud la longitud de la costa de Gran Bretaña, concluyendo que el resultado depende de la escala de medición. Esto significa que la longitud de la costa es menor si se mide empleando una imagen de satélite, que si se mide estando de pie sobre ella. ¿Qué pasaría si el objetivo es medir la longitud de la costa empleando una regla o un microscopio?. El resultado puede ser resumido en una ley de potencia: **a mayor escala, mayor longitud** [47].

En términos simples un fractal es un objeto cuya estructura se replica (de manera exacta o similar) a medida que se amplía, v.g. la hoja de un helecho tiene la misma forma

que el helecho completo. Para poder afirmar que un objeto tiene comportamiento fractal, éste debe cumplir con dos propiedades: ser autosimilar (autoafín) y poseer dimensión fraccionaria, estos conceptos serán definidos adelante.

### 1.2.2. Ejemplos clásicos

En esta sección se explican algunos de los clásicos ejemplos de fractales matemáticos.

#### 1.2.2.1. El conjunto de Cantor

Cantor, propuso uno de los más importantes conjuntos fractales. Su desarrollo teórico es definido de manera simple en [28]:

En primer lugar, se parte del segmento de recta constituido por el intervalo cerrado  $C_0 = [0, 1]$ .

Posteriormente, se obtiene el conjunto  $C_1$  al remover de  $C_0$  la tercera parte central, es decir:  $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ .

El siguiente conjunto  $C_2$ , se obtiene al remover la tercera parte central de los dos intervalos en  $C_1$ , es decir:  $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ .

En cada iteración el conjunto decrece, tal que:  $C_0 \supseteq C_1 \supseteq C_2 \supseteq \dots$ . Por lo tanto el Conjunto de Cantor es el límite  $C$  de la intersección de los conjuntos  $C_n$ ; es decir

$$C = \bigcap_{k \in \mathbb{N}} C_k.$$

En la Figura 1.9 se representan las primeras cinco iteraciones del Conjunto de Cantor, se aprecia la disminución de la longitud del conjunto, para valores grandes de  $k$  el conjunto se convierte en una serie de puntos, por lo que también recibe el nombre de *Polvo de Cantor*.

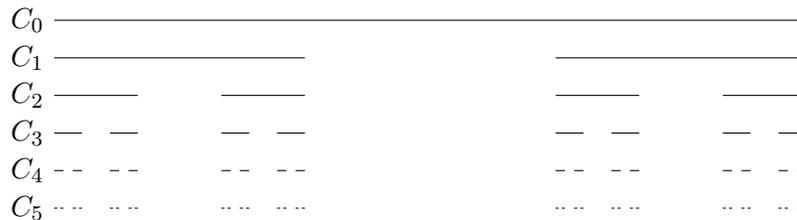


FIGURA 1.9. Primeras cinco iteraciones del Conjunto de Cantor.

Para la  $k$ -ésima iteración, el conjunto  $C_k$  está compuesto por  $2^k$  intervalos disyuntos, cada uno con longitud  $(1/3)^k$ . Luego, la longitud total de  $C_k$ , se puede escribir como

$$L(C_k) = 2^k \left(\frac{1}{3}\right)^k = \left(\frac{2}{3}\right)^k.$$

Cuando  $k \rightarrow \infty$ , se tiene

$$L(C) = \lim_{k \rightarrow \infty} \left(\frac{2}{3}\right)^k = 0, \quad (1.10)$$

y

$$\lim_{k \rightarrow \infty} 2^k = \infty. \quad (1.11)$$

De (1.10) se concluye que el *Polvo de Cantor* es un conjunto de longitud total igual a cero. Está compuesto por infinitos intervalos (1.11).

### 1.2.2.2. La curva de Koch

Su construcción se describe con detalle en [28]. Se desarrolla a continuación.

Se parte del segmento de recta constituido por el intervalo cerrado  $C_0 = [0, 1]$ , con longitud  $L(C_0) = 1$ . A continuación, se obtiene el conjunto  $C_1$ , al remover de  $C_0$  la tercera parte central, y reemplazarla por un triángulo equilátero con lados de longitud  $\frac{L(C_0)}{3} = \frac{1}{3}$  al cual se le suprime la base,  $C_1$  está compuesto por 4 segmentos iguales y su longitud es:  $L(C_1) = \frac{4}{3}L(C_0)$ . El mismo procedimiento se repite en las iteraciones siguientes, de modo que para la  $k$ -ésima iteración  $L(C_k) = \frac{4}{3}L(C_{k-1})$ . En la Figura 1.10 se representan las primeras cuatro iteraciones en la construcción de la curva de Koch.

Cuando  $k \rightarrow \infty$  la longitud de la curva  $L(C_k) \rightarrow \infty$ . De aquí se concluye que la *curva de Koch* tiene longitud infinita aunque puede ser dibujada en un segmento finito, es continua en todos sus puntos sin embargo no es derivable en ninguno [48].

Una variación de la *Curva de Koch* es el *Copo de Nieve de Koch*. Este se construye partiendo de un triángulo equilátero y siguiendo el mismo procedimiento que para la *Curva de Koch*. El resultado de sus primeras tres iteraciones se aprecia en la Figura 1.11.

### 1.2.2.3. Triángulo de Sierpinski

La construcción de este conjunto fractal se describe a continuación [28].

Iniciando con la superficie de un triángulo equilátero de lado  $l = 1$ ; su área se puede escribir como,  $A(S_0) = \frac{\sqrt{3}}{4} \cdot l^2 = \frac{\sqrt{3}}{4}$  y su perímetro  $P(S_0) = 3 \cdot l = 3$ .

Posteriormente, se trazan líneas que unen los puntos medios de cada lado de  $S_0$ .  $S_1$  es el resultado luego de remover el triángulo central (no se remueve su perímetro), se obtienen en total 3 triángulos equiláteros, cada uno de lado  $l_1 = \frac{1}{2}$ , por tanto se define el área y

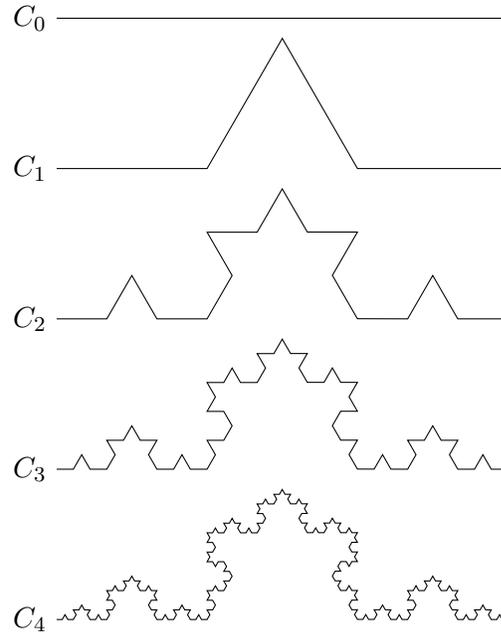


FIGURA 1.10. Primeras cuatro iteraciones en la construcción de la Curva de Koch.

perímetro de  $S_1$  como

$$A(S_1) = 3 \frac{\sqrt{3}}{4} \cdot \left(\frac{1}{2}\right)^2,$$

$$P(S_1) = 3 \cdot \frac{1}{2},$$

donde  $S_1$  es un subconjunto de  $S_0$ .

Para obtener  $S_2$  se repite el procedimiento anterior en cada uno de los 3 triángulos que componen  $S_1$ , el resultado son  $3^2 = 9$  triángulos equiláteros de lado  $l_2 = \left(\frac{1}{2}\right)^2$ .

El resultado final se puede ver en la Figura 1.12; se cumple que  $S_0 \supseteq S_1 \supseteq S_2 \supseteq \dots$ , el triángulo de Sierpinski se define como

$$S = \bigcap_{k \in \mathbb{N}} S_k.$$

El conjunto  $S_k$  para la  $k$ -ésima iteración está compuesto por  $3^k$  triángulos equiláteros de lado  $l_k = 2^{-k}$ . El área de cada triángulo resultante es  $A(S_k) = l_k^2 \cdot \frac{\sqrt{3}}{4} = 2^{-2k} \cdot \frac{\sqrt{3}}{4}$ . Por tanto, el área total del triángulo en la iteración  $k$  es

$$A(S_k) = 3^k \cdot l_k^2 \cdot \frac{\sqrt{3}}{4} = 3^k \cdot 2^{-2k} \cdot \frac{\sqrt{3}}{4}$$

$$= \left(\frac{3}{4}\right)^k \frac{\sqrt{3}}{4} = \frac{3^{k+\frac{1}{2}}}{4^{k+1}},$$

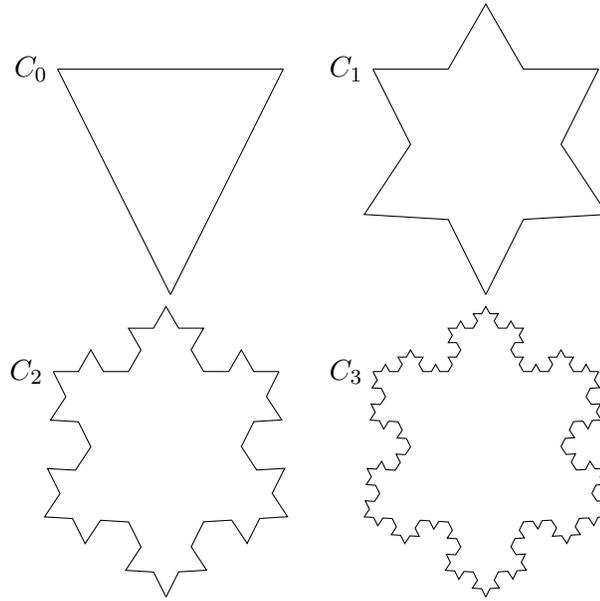


FIGURA 1.11. Construcción del Copo de Nieve de Koch.

cuando  $k \rightarrow \infty$ , se tiene

$$\lim_{k \rightarrow \infty} A(S_k) = \lim_{k \rightarrow \infty} \frac{3^{k+\frac{1}{2}}}{4^{k+1}} = 0. \quad (1.12)$$

El segmento de línea que dibuja los bordes del triángulo  $S_k$  contiene los bordes de todos los triángulos obtenidos en iteraciones anteriores, dado que  $S_k$  está compuesto por  $3^k$  triángulos equiláteros, cada uno con lado de longitud  $l_k = 2^{-k}$ , el perímetro total de  $S_k$  se puede escribir:

$$P(S_k) = 3 \cdot 3^k \cdot 2^{-k} = 3 \left( \frac{3}{2} \right)^k,$$

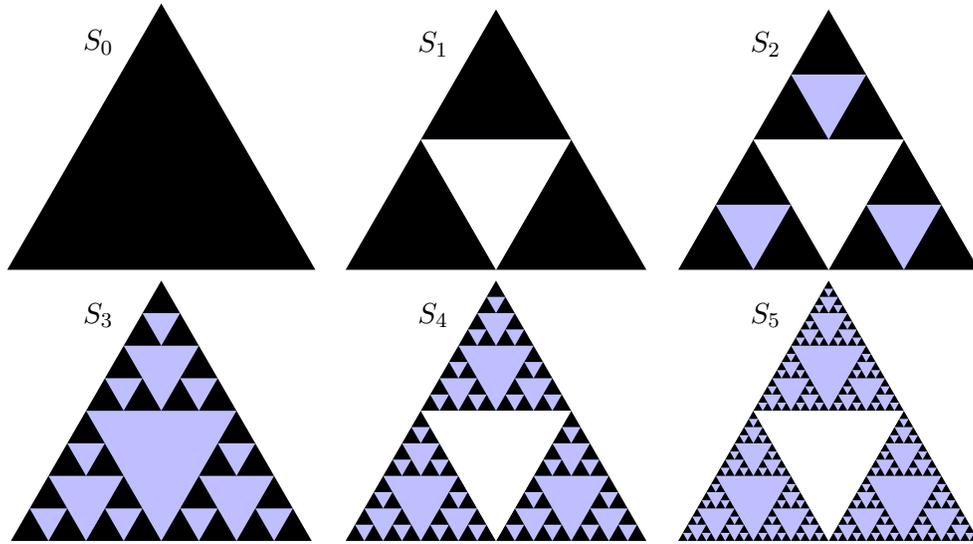
cuando  $k \rightarrow \infty$  se tiene

$$\lim_{k \rightarrow \infty} P(S_k) = \lim_{k \rightarrow \infty} 3 \cdot \left( \frac{3}{2} \right)^k = \infty. \quad (1.13)$$

Se puede concluir que cuando  $k \rightarrow \infty$  el área del triángulo converge a cero (1.12) y su perímetro tiende a infinito (1.13). El *Triángulo de Sierpinski* es un objeto geométrico con perímetro infinito pero sin área.

### 1.2.3. Autosimilaridad e invarianza de escala

“Al ampliar partes de un fractal, emergen estructuras que parecen ser idénticas al original” [58]. Ésta es una definición simple y acertada de autosimilaridad. Cuando un objeto puede ser expresado como la unión de objetos o patrones, donde cada uno de ellos es una copia reducida del objeto completo, se habla de autosimilaridad exacta. Cuando las partes conservan algunas propiedades estadísticas del original, se habla de

FIGURA 1.12. Primeras cinco iteraciones del fractal *Sierpinski Gasket*.

autosimilaridad estadística o autoafinidad (i.e. las partes son parecidas al original).

Los fractales matemáticos como los vistos en la sección anterior son exactamente autosimilares a diferencia de los fractales que se encuentran en la naturaleza, en estos casos sus partes son parecidas al todo, v.g. helecho, brócoli. Derivada de esta propiedad fundamental (autosimilaridad) se puede afirmar que los fractales presentan invarianza de escala, esto significa que tienen la capacidad de mantener su geometría (o partes de ella) a diferentes escalas. Por esta razón no es posible determinar su escala basándose únicamente en la forma del objeto. El Polvo de Cantor, la curva de Koch y el triángulo de Sierpinski son exactamente autosimilares y presentan invarianza de escala.

#### 1.2.4. Dimensión de similaridad

En esta sección se expone una definición de dimensión por similaridad (i.e. dimensión fractal) descrita por Mandelbrot (1967) ([1],[47] y [48]).

Una línea de longitud  $L = 1$  puede ser dividida en  $M$  segmentos autosimilares más pequeños, cada uno con longitud  $\epsilon$ , entonces

$$L = M \cdot \epsilon = 1. \quad (1.14)$$

La línea se compone por  $M$  segmentos autosimilares con escala  $\epsilon = 1/M$ . Ahora, considerando el caso de una unidad de área  $A = 1$ , ésta puede ser dividida en  $M$  unidades de área autosimilares más pequeñas, cada una con área  $\epsilon^2$ , como

$$A = M \cdot \epsilon^2 = 1. \quad (1.15)$$

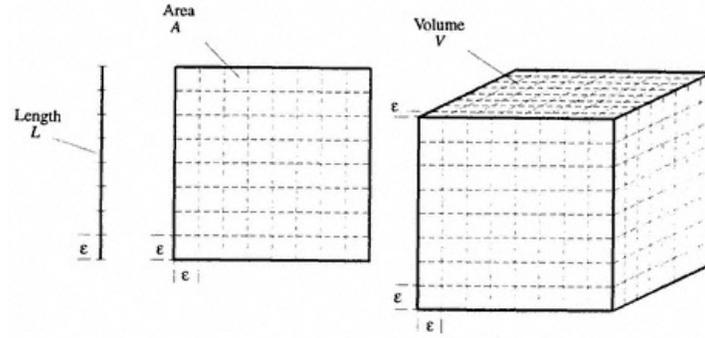


FIGURA 1.13. División de segmento de línea, unidad de área y volumen en elementos autosimilares de lado  $\epsilon$ , tomada de [1].

Una unidad de área se compone por  $M$  unidades de área autosimilares, con escala  $\epsilon = 1/(M^{\frac{1}{2}})$ .

Sea, una unidad de volumen  $V = 1$ , compuesta por  $M$  unidades autosimilares más pequeñas, cada una con volumen  $\epsilon^3$ , entonces;

$$V = M \cdot \epsilon^3 = 1. \quad (1.16)$$

Una unidad de volumen se compone por  $M$  unidades autosimilares más pequeñas, con escala  $\epsilon = 1/(M^{\frac{1}{3}})$ .

La Figura 1.13 representa la división de objetos geométricos autosimilares. Las ecuaciones (1.14), (1.15) y (1.16) se pueden generalizar como

$$\begin{aligned} M\epsilon^{D_s} &= 1, \\ \epsilon^{D_s} &= \frac{1}{M}, \\ \log(\epsilon^{D_s}) &= \log\left(\frac{1}{M}\right), \\ D_s \log(\epsilon) &= \log\left(\frac{1}{M}\right) = -\log(M), \\ D_s &= \frac{-\log(M)}{\log(\epsilon)} = \frac{\log(M)}{-\log(\epsilon)} = \frac{\log(M)}{\log\left(\frac{1}{\epsilon}\right)}. \end{aligned} \quad (1.17)$$

$D_s$  es la dimensión de similaridad [1].

#### 1.2.4.1. Cálculo de la dimensión de similaridad para algunos conjuntos fractales

A continuación se realiza el cálculo de la dimensión de similaridad  $D_s$  para tres conjuntos fractales descritos en la sección 1.2.2.

##### Conjunto de Cantor

Al tomar el segmento de línea inicial  $[0, 1]$ , éste se divide en 3 segmentos iguales ( $\epsilon = 1/3$ ),

de los que se conservan  $M = 2$ , reemplazando en 1.17;

$$D_s = \frac{\log(2)}{\log\left(\frac{1}{\frac{1}{3}}\right)} = \frac{\log(2)}{\log(3)} = 0.6309.$$

El Polvo de Cantor, tiene dimensión fraccionaria entre 0 y 1.

### Curva de Koch

Al tomar el segmento de línea inicial  $[0, 1]$  de la curva, éste se divide en 3 segmentos iguales ( $\epsilon = 1/3$ ), de los que se conservan  $M = 4$ , reemplazando en (1.17);

$$D_s = \frac{\log(4)}{\log\left(\frac{1}{\frac{1}{3}}\right)} = \frac{\log(4)}{\log(3)} = 1.2618. \quad (1.18)$$

La Curva de Koch, tiene dimensión fraccionaria, muy cercana a 1.

### Triángulo de Sierpinski

Al tomar el triángulo equilátero inicial de lado  $l_0 = 1$ , éste se divide en 4 triángulos más pequeños de lado  $l_1 = 1/2$  y ( $\epsilon = 1/2$ ), de ellos se conservan  $M = 3$  triángulos, reemplazando en (1.17);

$$D_s = \frac{\log(3)}{\log\left(\frac{1}{\frac{1}{2}}\right)} = \frac{\log(3)}{\log(2)} = 1.5849.$$

Los fractales vistos hasta ahora son exactamente autosimilares. Otro tipo de fractales son los aleatorios, que son estadísticamente autosimilares. Cada parte de un fractal tiene las mismas propiedades estadísticas que el conjunto completo, más no es una copia exacta, esto ocurre con la mayoría de elementos fractales en la naturaleza [1].

### 1.2.5. Dimensión *Box Counting*

La Figura 1.14 muestra una imagen satelital de un cuerpo de agua, a simple vista se puede afirmar que tiene comportamiento fractal. En este caso como en la mayoría de los fractales naturales (v.g. río, helecho, costa de una isla) no es trivial el cálculo de la dimensión de similaridad  $D_s$  empleando la ecuación (1.17), sin embargo, existe una metodología empírica que permite determinar la dimensión fractal en estos casos: *Box Counting*. Se puede emplear para analizar cualquier estructura (hasta  $\mathbb{R}^3$ ), además es muy conveniente en el análisis de imágenes ya que su estructura asemeja los pixeles [60].

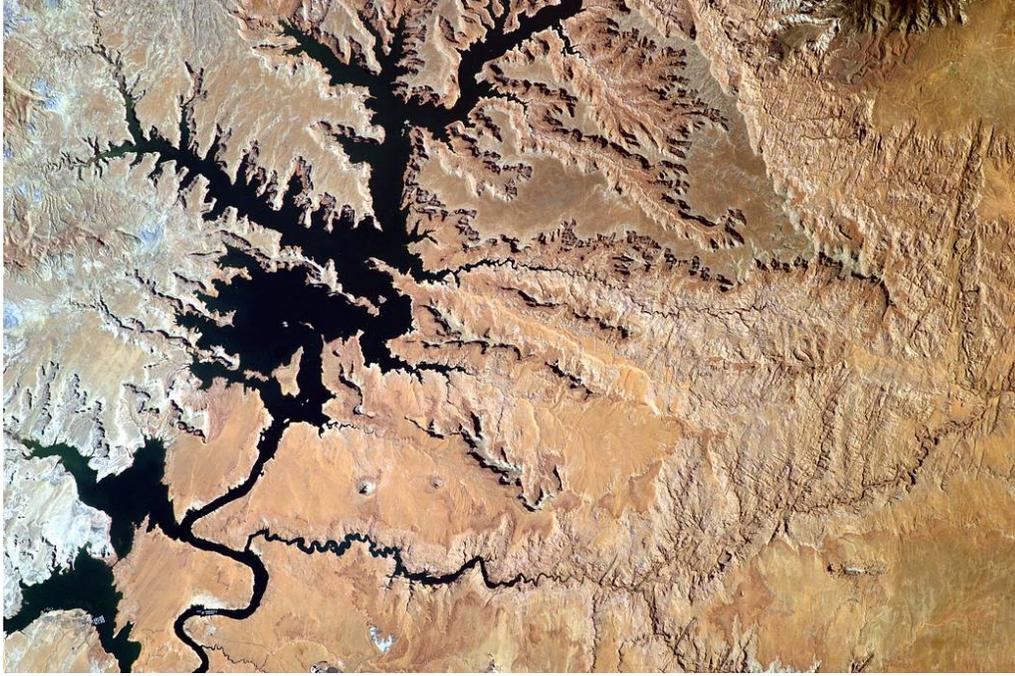


FIGURA 1.14. Lake Powell desde la estación espacial EarthKAM Imagen: NASA/EarthKAM.org (Agosto 4, 2017)

Para comenzar, se cubre el objeto con cuadros de lado  $\delta$  que no se traslapan y se cuenta la cantidad de cuadros necesarios para cubrir el fractal ( $\Gamma(\delta)$ ). Similarmente se pueden usar segmentos de línea de longitud  $\delta$  o cubos de lado  $\delta$ .

En general se necesitarán  $\Gamma(\delta) = \frac{1}{\delta^{D_B}} = \delta^{-D_B} = \left(\frac{1}{\delta}\right)^{D_B}$  cuadros para cubrir el objeto, donde  $D_B$  es la dimensión *Box Counting* [1].

La dimensión *Box Counting* se obtiene como el límite de  $D_B$  cuando  $\delta \rightarrow 0$ ,

$$D_B = \lim_{\delta \rightarrow 0} \frac{\log(\Gamma(\delta))}{\log\left(\frac{1}{\delta}\right)}, \quad (1.19)$$

lo cual según [1] corresponde a

$$\log(\Gamma(\delta)) = D_B \log\left(\frac{1}{\delta}\right), \quad (1.20)$$

y

$$D_B = \frac{\log(\Gamma(\delta))}{\log\left(\frac{1}{\delta}\right)}. \quad (1.21)$$

El resultado (1.20) es de la forma  $y = a + xb$ , por tanto, al realizar la gráfica  $\log\left(\frac{1}{\delta}\right)$  vs  $\log(\Gamma(\delta))$ , para diferentes valores de  $\delta$  y ajustar una recta por mínimos cuadrados ordinarios (MCO) se puede estimar la dimensión fractal como

$$\widehat{D}_B = b. \quad (1.22)$$

En algunos trabajos la recta se dibuja de la forma  $\log(\Gamma(\delta)) = -D_B \log(\delta)$  de modo que  $\widehat{D}_B = -b$  [60].

La Figura 1.15 muestra la aplicación de la metodología para el caso de la curva de Koch, así como el conteo de cajas  $\Gamma(\delta)$  para diferentes valores de  $\delta$ . Se puede calcular  $D_B$  empleado un solo valor de  $\delta$ , teniendo en cuenta el resultado que se observa en la última parte de la Figura 1.15. Reemplazando  $[(\delta = 1/12), (\Gamma(\delta) = 24)]$  en (1.21);

$$D_B = \frac{\log(\Gamma(\delta))}{\log\left(\frac{1}{\delta}\right)} = \frac{\log(24)}{\log\left(\frac{1}{12}\right)} = 1.2789. \quad (1.23)$$

La Figura 1.16 muestra la construcción del gráfico  $\log\text{-}\log$  ( $\log(1/\delta)$ ,  $\log(\Gamma(\delta))$ ), los puntos representan los resultados obtenidos en la Figura 1.15. Al ajustar una recta por MCO se obtiene  $D_B = 1.2545$ , éste valor es muy similar al obtenido en en (1.18).

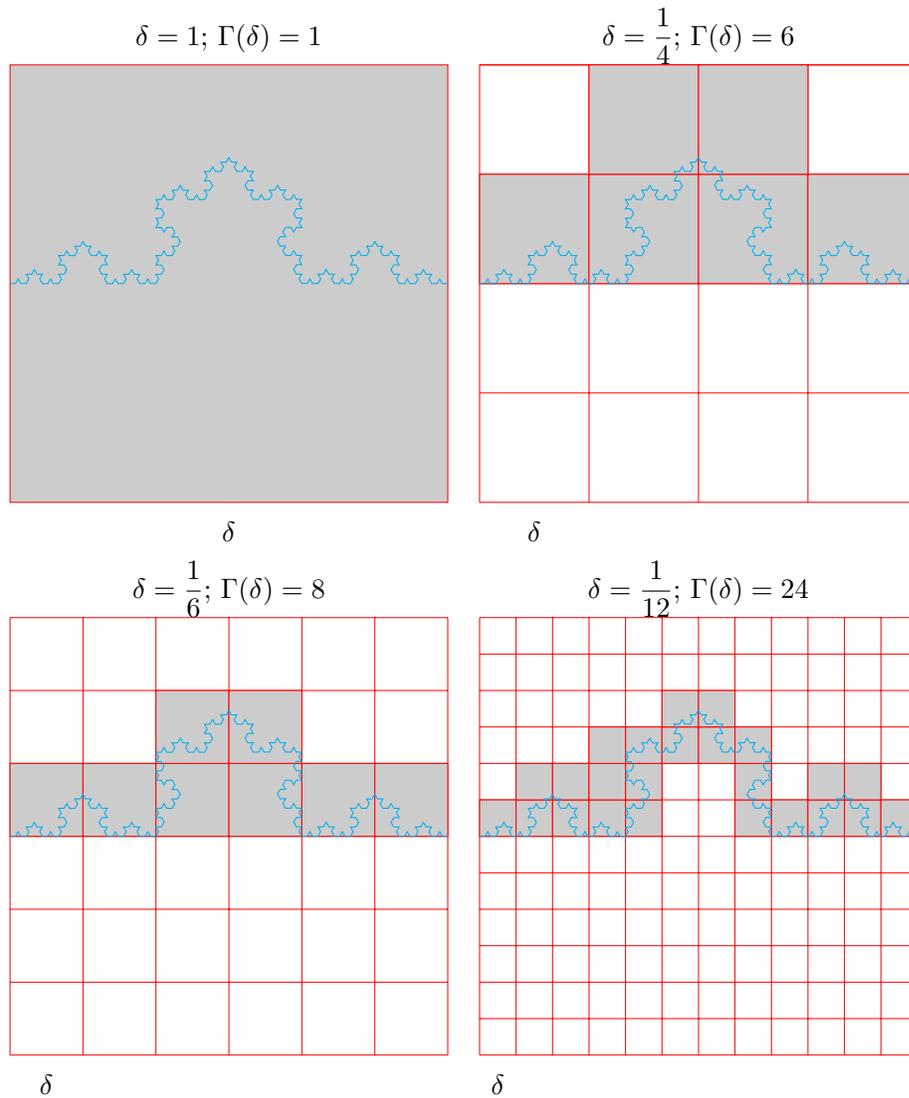


FIGURA 1.15. Aplicación de la metodología *Box Counting* para la curva de Koch, se varían los valores de  $\delta$ , el valor  $\Gamma(\delta)$  corresponde al conteo de las celdas de lado  $\delta$  que contienen algún segmento de la curva (Elaboración propia).

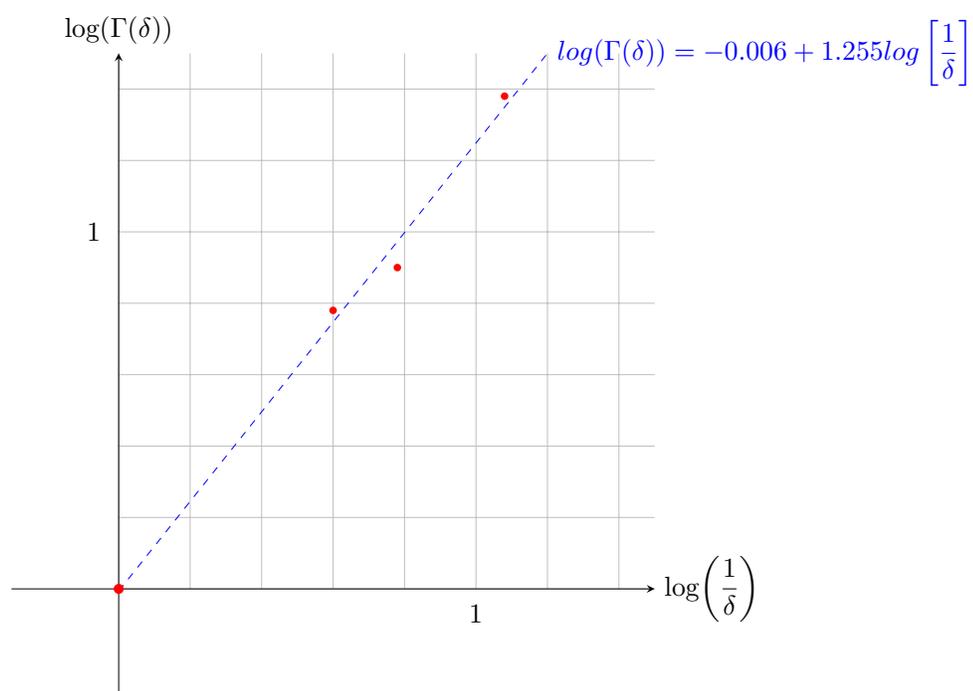


FIGURA 1.16. Gráfico *log-log* para el cálculo de  $D_B$  en el caso de la curva de Koch, donde ( $D_B = 1.255$ ) es la pendiente de la recta, ajustada por MCO (Elaboración propia)

---



---

## Prueba de AEC basada en dimensión *Box Counting*

---



---

En esta sección se expone en primera instancia la prueba gráfica y cálculo de la dimensión fractal empleando la metodología *Box Counting* bajo la hipótesis de AEC. Posteriormente, se realiza el desarrollo teórico de la prueba y cálculo de la dimensión fractal para el caso de un patrón puntual regular. Se compara el resultado de la prueba para un patrón puntual regular con uno que presenta AEC con tamaños de muestra  $n$  iguales.

Finalmente se describe la interpretación de la prueba gráfica propuesta, a través de una función realizada en el programa estadístico R [54].

### 2.1. Desarrollo teórico de la prueba de AEC basada en la dimensión Box Counting

Suponga un patrón puntual  $\mathbf{x}$  bajo AEC, con una muestra  $n$  de eventos en  $A \in R^2$ , donde  $A$  es un cuadrado de lado  $k$ . Sea  $N(A)$  la variable aleatoria que cuenta el número de eventos de  $\mathbf{x}$  que ocurren en  $A$  de manera que  $N(A) \sim Poisson(\lambda)$ , la probabilidad de que ocurran exactamente  $y$  eventos en un cuadrado de área  $|A|$  se puede escribir cómo:

$$P(N(A) = y) = \frac{e^{-\lambda} \lambda^y}{y!}. \quad (2.1)$$

De (2.1) se puede calcular la probabilidad de que ocurra al menos un evento en un cuadrado de área  $|A|$  como

$$P(N(A) > 0) = 1 - P(N(A) = 0) = 1 - \frac{e^{-\lambda} (\lambda)^0}{0!} = 1 - e^{-\lambda}. \quad (2.2)$$

Ahora, teniendo en cuenta el procedimiento para el cálculo de  $D_B$  descrito en la sección 1.2.5, es posible dividir el cuadrado original  $A$  de lado  $k$  en  $\beta_i$  cuadrados ( $A_i$ ) de lado  $\delta_i$  que no se traslapan (Figura 2.1), con área  $|A_i| = \delta_i^2$ , donde  $\delta_i$  es una fracción de  $k$ , tal que  $\delta_i = \frac{k}{i}$ , para  $i = 1, 2, \dots, \infty$ .

El área de  $A$  se escribe cómo:

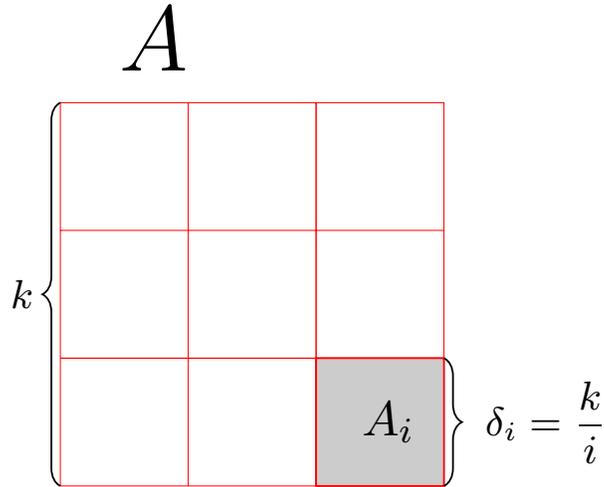


FIGURA 2.1. Se puede dividir la ventana de observación  $A$  de lado  $k$  en  $\beta_i$  cuadrados de lado  $\delta_i$  que no se traslapan.

$$|A| = \beta_i |A_i| = \beta_i \delta_i^2 = \beta_i \frac{k^2}{i^2}.$$

Se puede definir el total de cuadrados  $\beta_i$  de lado  $\delta_i$  que no se traslapan como

$$\beta_i = i^2 = \left(\frac{k}{\delta_i}\right)^2 = \frac{k^2}{\delta_i^2}. \quad (2.3)$$

Para cada tamaño de la unidad de área es posible definir una intensidad homogénea  $\tau_i = \frac{n}{\beta_i}$  como el número de puntos por unidad de área  $|A_i|$ ,  $\tau_i$ , a pesar de que este valor es diferente para cada  $|A_i|$  es constante a lo largo de la ventana de observación  $A$ .

Reemplazando (2.3), en  $\tau_i = \frac{n}{\beta_i}$  se tiene;

$$\tau_i = \frac{n}{\beta_i} = \frac{n}{\frac{k^2}{\delta_i^2}} = \frac{n\delta_i^2}{k^2}. \quad (2.4)$$

Reemplazando (2.4) en (2.2), es posible determinar la probabilidad de que ocurra por lo menos un evento en un cuadrado de área  $|A_i|$ , de modo que

$$P(N(A_i) > 0) = 1 - e^{-\tau_i} = 1 - e^{-\frac{n\delta_i^2}{k^2}}. \quad (2.5)$$

Por otra parte, sea  $\Gamma(\delta_i)$  la variable aleatoria que cuenta el número de celdas de lado  $\delta_i$  que contienen al menos un evento del patrón puntual:

$$\Gamma(\delta_i) = \sum_{i=1}^{\beta_i} z_i, \quad \text{donde: } z_i = I_{N(A_i)} = \begin{cases} 1 & \text{si } N(A_i) > 0 \\ 0 & \text{si } N(A_i) = 0 \end{cases}. \quad (2.6)$$

De (2.5) y (2.6), se puede definir el valor esperado de  $\Gamma(\delta_i)$ , como

$$\begin{aligned}
E[\Gamma(\delta_i)] &= E\left[\sum_{i=1}^{\beta_i} z_i\right] = \sum_{i=1}^{\beta_i} E[z_i] \\
&= \sum_{i=1}^{\beta_i} [(1)P(N(A_i) > 0) + (0)P(N(A_i) = 0)] \\
&= \sum_{i=1}^{\beta_i} \left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] \\
&= \beta_i \left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] = \frac{k^2}{\delta_i^2} \left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]. \tag{2.7}
\end{aligned}$$

Aplicando  $\log$  a ambos lados de (2.7) se obtiene

$$\begin{aligned}
\log(E[\Gamma(\delta_i)]) &= \log\left(\frac{k^2}{\delta_i^2} \left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]\right) \\
&= \log\left[\frac{k^2}{\delta_i^2}\right] + \log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] \\
&= 2\log(k) - 2\log(\delta_i) + \log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] \\
&= 2\log(k) + 2\log\left(\frac{1}{\delta_i}\right) + \log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] \blacksquare \tag{2.8}
\end{aligned}$$

A partir de la ecuación (2.8) se puede construir la gráfica  $\log\text{-}\log$  teórica para el cálculo de la dimensión fractal por el método *Box Counting*.

El valor de  $\log(E[\Gamma(\delta_i)])$  depende de la longitud de lado del cuadrado inicial ( $k$ ) y del número de eventos del patrón puntual ( $n$ ).

En la primera partición, cuando  $i = 1$ , existe un único cuadrado inicial que contiene todos los eventos, por tanto  $\Gamma(\delta_1) = \delta_1 = \beta_1 = 1$  y por tanto  $\log(\Gamma(\delta_1)) = \log(1) = 0$ . Reemplazado en (2.8) y despejando  $\log(1/(\delta_1))$  se obtiene

$$\begin{aligned}
\log(\Gamma(\delta_1)) = \log(1) &= 2\log(k) + 2\log\left(\frac{1}{\delta_1}\right) + \log\left[1 - e^{-\frac{nk^2}{\delta_1^2}}\right] \\
0 &= \log(k) - \log\left(\frac{1}{\delta_1}\right) + \log\left[1 - e^{-\frac{nk^2}{k^2}}\right] \\
\log\left(\frac{1}{\delta_1}\right) &= -\log(k) - \frac{1}{2}\log[1 - e^{-n}]. \tag{2.9}
\end{aligned}$$

El resultado obtenido en (2.9) es el cruce de la curva con el eje  $x$ .

Nótese que  $\lim_{n \rightarrow \infty} \log(1 - e^{-n}) = \log(1) = 0$ . Esto quiere decir que para valores muy grandes de muestra  $n$ , el cruce de la curva con el eje  $x$  está determinado por el valor de  $k$  y se puede calcular como  $-\log(k)$ .

En la Figura 2.2 se puede observar la comparación de gráficas teóricas resultantes de (2.8) para diferentes valores de  $k$ , manteniendo constante  $n$ . Se evidencia en dicha figura que el aumento de  $k$  se traduce en un desplazamiento de la curva a la izquierda.

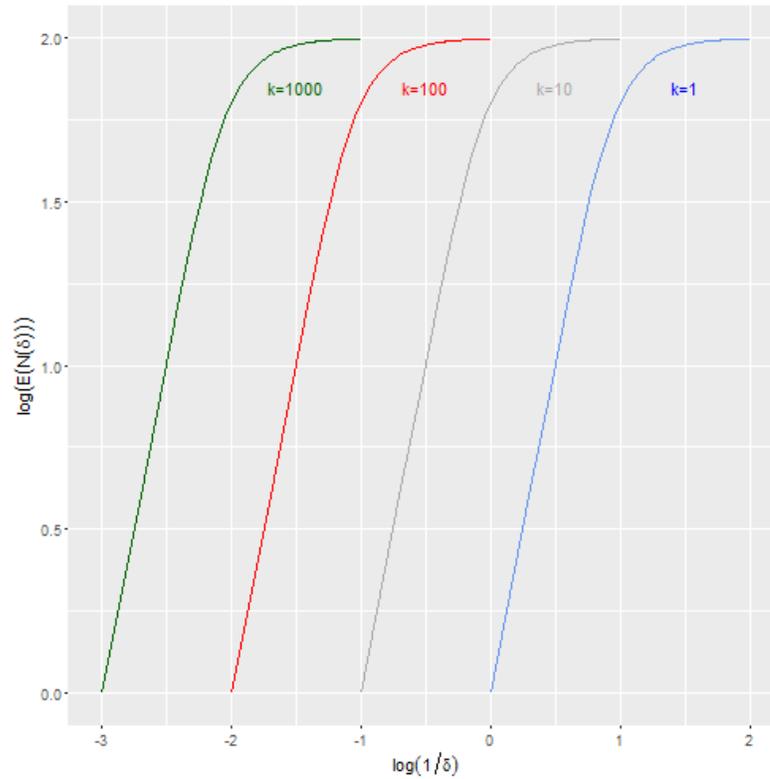


FIGURA 2.2. Gráfico *log-log* bajo AEC obtenido a través de la ecuación (2.8), se realizan cuatro curvas manteniendo constante  $n = 100$  y con valores  $k=1,10,100$  y  $1000$ .

Dado que los patrones puntuales que se analizan en la práctica son finitos, no tiene sentido continuar realizando particiones una vez que el valor esperado de puntos por unidad de área es menor a uno ( $\tau_i < 1$ ) esto quiere decir que hay menor cantidad de puntos que cuadrados. Esto ocurre con  $\delta\sqrt{n}$  como se muestra en (2.10).

$$\begin{aligned} \tau_i &= \frac{n}{\beta_i} = \frac{n}{i^2} = \frac{n\delta_i^2}{k^2} = 1, \\ i &= \sqrt{n}, \\ \delta\sqrt{n} &= \frac{k}{\sqrt{n}}, \end{aligned} \tag{2.10}$$

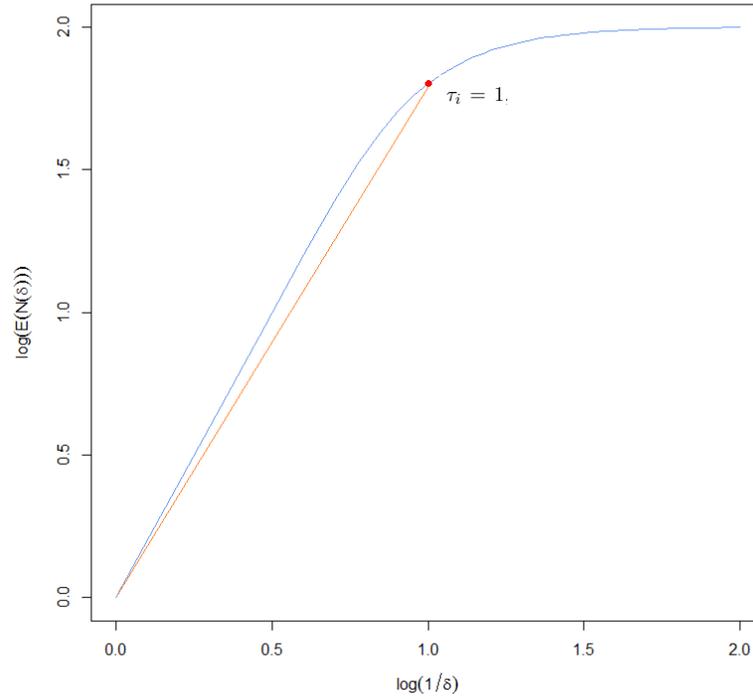


FIGURA 2.3. La curva representa el gráfico *log-log* obtenido a través de la ecuación (2.8), para un patrón puntual bajo AEC con  $n = 100$  y  $k = 1$ . El punto representa los valores cuando  $\tau_i = 1$ , calculados según (2.11) y (2.12), el valor de la pendiente de la recta desde el origen hasta  $\tau_i = 1$ , es la primera dimensión *Box Counting* ( $D_{B1} = 1.808$ ), calculada según (2.13).

Aplicando *log*;

$$\begin{aligned} \log\left(\frac{1}{\delta\sqrt{n}}\right) &= \log\left(\frac{\sqrt{n}}{k}\right) \\ &= \frac{1}{2}\log(n) - \log(k). \end{aligned} \quad (2.11)$$

Reemplazando (2.11) y (2.10) en (2.8), se puede calcular  $\log(E[\Gamma(\delta\sqrt{n})])$  cuando  $\tau_i = 1$ :

$$\begin{aligned} \log(E[\Gamma(\delta\sqrt{n})]) &= 2\log(k) + 2\left(\frac{1}{2}\log(n) - \log(k)\right) + \log\left[1 - e^{-\frac{nk^2}{nk^2}}\right] \\ &= 2\log(k) - 2\log(k) + \log(n) + \log(1 - e^{-1}) \\ &= \log(n) + \log(0.6321206) \\ &= \log(n) - 0.1992001. \end{aligned} \quad (2.12)$$

Los valores definidos en (2.11) y (2.12), cuando  $\tau_i = 1$ , representan el punto en cual la curva cambia de pendiente. El resultado se puede ver en la Figura 2.3.

En la Figura 2.4 se evidencia el efecto de la variación de  $n$  en la curva. Un aumento en el tamaño de la muestra  $n$  se traduce en el aumento de amplitud de la curva.

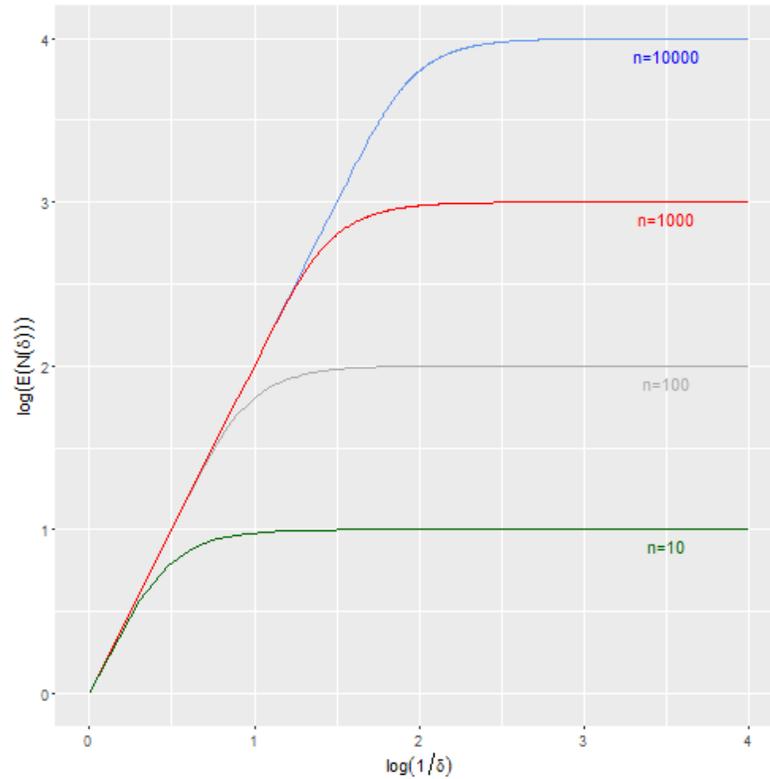


FIGURA 2.4. Gráfico *log-log* bajo AEC obtenido a través de la ecuación (2.8), se dibujaron cuatro curvas manteniendo constante  $k = 1$  y con diferentes tamaños de muestra  $n=10,100,1000$  y  $10000$ .

### 2.1.1. Cálculo de la dimensión fractal por el método *Box Counting* bajo la hipótesis de AEC

La gráfica teórica *log-log* para definir la dimensión fractal por el método *Box Counting* que se obtiene con la ecuación (2.8) (i.e. bajo AEC) muestra un comportamiento multifractal (i.e. tiene varias dimensiones fractales) y cambia de pendiente cuando  $\tau_i = 1$ . En este caso se observan dos dimensiones fractales. La primera dimensión fractal  $D_{B1}$  definida como la pendiente de la curva desde el origen cuando  $\delta_1 = k$  ( $i = 1$ , ver (2.9)), hasta  $\tau_i = 1$  con  $\delta_{\sqrt{n}}$  ( $i = \sqrt{n}$ , ver (2.10)). La segunda dimensión fractal  $D_{B2}$  corresponde a la pendiente de la curva desde  $\tau_i = 1$ , hasta  $1/\delta_i \rightarrow \infty$ . En la Figura 2.3 se muestra un ejemplo de cálculo de la primera dimensión fractal  $D_{B1}$ .

Para determinar la primera dimensión fractal  $D_{B1}$ , en el caso de un patrón puntual completamente aleatorio, se define la pendiente de la recta desde el punto  $[\log(1/(\delta_1)), \log(\Gamma(\delta_1))]$ , hasta el punto  $[\log(1/(\delta_{\sqrt{n}})), \log(\Gamma(\delta_{\sqrt{n}}))]$ .

Reemplazando (2.9), (2.11) y (2.12) en (1.21)

$$\begin{aligned}
D_{B1} &= \frac{\log(\Gamma(\delta_{\sqrt{n}})) - \log(\Gamma(\delta_1))}{\log\left(\frac{1}{\delta_{\sqrt{n}}}\right) - \log\left(\frac{1}{\delta_1}\right)} \\
&= \frac{\log(n) - 0.1992001 - \log(1)}{\frac{1}{2}\log(n) - \log(k) - \left(-\log(k) - \frac{1}{2}\log[1 - e^{-n}]\right)} \\
&= \frac{\log(n) - 0.1992001}{\frac{1}{2}\log(n) - \log(k) + \log(k) + \frac{1}{2}\log[1 - e^{-n}]} \\
&= \frac{\log(n) - 0.1992001}{\frac{1}{2}\log(n) + \frac{1}{2}\log[1 - e^{-n}]} \\
&= 2 \cdot \frac{\log(n) - 0.1992001}{\log(n) + \log[1 - e^{-n}]}. \tag{2.13}
\end{aligned}$$

|          | <b>n</b> |      |      |      |       |        |         |
|----------|----------|------|------|------|-------|--------|---------|
|          | 10       | 100  | 500  | 1000 | 50000 | 100000 | 1000000 |
| $D_{B1}$ | 1.60     | 1.80 | 1.85 | 1.87 | 1.915 | 1.92   | 1.934   |

TABLA 2.1. Variación de la primera dimensión fractal  $D_{B1}$  (2.13), para diferentes tamaños de  $n$ .

$D_{B1}$  depende de  $n$ , la Figura 2.3 muestra el resultado del cálculo de  $D_{B1}$  para el caso  $k = 1$  y  $n = 100$ . La tabla 2.1 muestra como el aumento en el tamaño de muestra  $n$ , resulta en mayor valor calculado  $D_{B1}$ , cuando  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} D_{B1} = \lim_{n \rightarrow \infty} 2 \cdot \frac{\log(n) - 0.1992001}{\log(n) + \log[1 - e^{-n}]} = 2.$$

De lo anterior se puede concluir que la primera dimensión fractal  $D_{B1}$  para un patrón puntual que presenta AEC, está dada por una recta con pendiente que tiende a 2, para valores suficientemente grandes de  $n$ .

En la práctica, la curva se dibuja hasta cuando  $\log(\Gamma(\delta_i)) = \log(n)$ . Reemplazando esta igualdad en (2.8)

$$\begin{aligned}
\log(\Gamma(\delta_i)) = \log(n) &= 2\log(k) + 2\log\left(\frac{1}{\delta_i}\right) + \log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right], \\
\log\left(\frac{1}{\delta_i}\right) &= \frac{1}{2}\log(n) - \log(k) - \frac{1}{2}\log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]. \tag{2.14}
\end{aligned}$$

La segunda dimensión fractal  $D_{B2}$ , corresponde a la pendiente de la recta desde el punto  $\left[\log\left(1/(\delta_{\sqrt{n}})\right), \log(\Gamma(\delta_{\sqrt{n}}))\right]$  hasta el punto  $[\log(1/(\delta_i)), \log(n)]$ .

Reemplazando (2.11), (2.12) y (2.14) en (1.21) se tiene

$$\begin{aligned}
D_{B2} &= \frac{\log(\Gamma(\delta_i)) - \log(\Gamma(\delta_{\sqrt{n}}))}{\log\left(\frac{1}{\delta_i}\right) - \log\left(\frac{1}{\delta_{\sqrt{n}}}\right)} \\
&= \frac{\log(n) - (\log(n) - 0.1992001)}{\left(\frac{1}{2}\log(n) - \log(k) - \frac{1}{2}\log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]\right) - \left(\frac{1}{2}\log(n) - \log(k)\right)} \\
&= \frac{\log(n) - \log(n) + 0.1992001}{\frac{1}{2}\log(n) - \log(k) - \frac{1}{2}\log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] - \frac{1}{2}\log(n) + \log(k)} \\
&= \frac{0.1992001}{-\frac{1}{2}\log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]} \\
&= -\frac{0.3984002}{\log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]}. \tag{2.15}
\end{aligned}$$

Usando (2.15) y (1.19) se puede definir también  $D_{B2}$  como

$$\begin{aligned}
D_{B2} &= \lim_{\delta_i \rightarrow 0} \left[ -\frac{0.3984002}{\log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right]} \right] \\
D_{B2} &= -\frac{0.3984002}{\lim_{\delta_i \rightarrow 0} \left[ \log\left[1 - e^{-\frac{n\delta_i^2}{k^2}}\right] \right]} \\
&= \frac{0.3984002}{-\infty} \\
&= \frac{0.3984002}{\infty} = 0, \tag{2.16}
\end{aligned}$$

es decir, el valor de la segunda dimensión fractal ( $D_{B2}$ ), para un patrón puntual completamente aleatorio, depende del factor  $-(n\delta_i^2)/(k^2)$ .  $D_{B2}$  tiende a cero cuando  $\delta_i \rightarrow 0$ .

## 2.2. Dimensión *Box Counting* para un patrón puntual espacial regular

Suponga un patrón puntual totalmente regular, con una muestra  $n$  de eventos (Figura 2.5); el número de eventos del patrón puntual contenidos en un área  $|A|$  es una constante y se puede definir como,  $\omega_i = \frac{n}{\beta_i}$ .

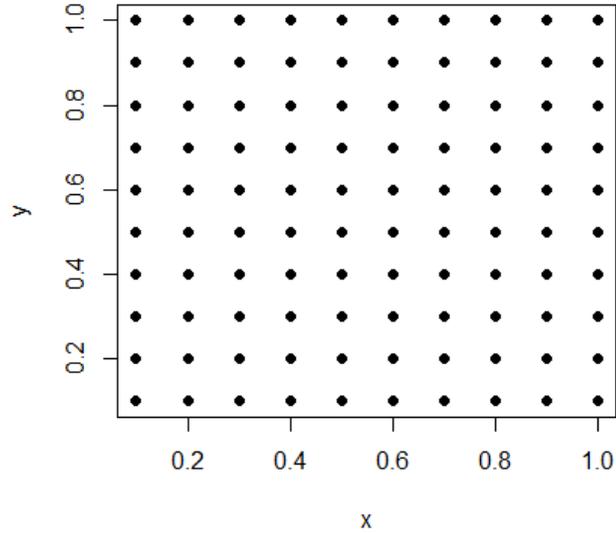


FIGURA 2.5. Simulación de un patrón puntual regular, en una ventana de  $1 \times 1$  con  $n = 100$ , el comportamiento se asemeja a un patrón de inhibición.

Siguiendo el procedimiento para el cálculo de  $D_B$ , descrito en la sección 1.2.5, es posible dividir el cuadrado original de lado  $k$ , en  $\beta_i = i^2$  cuadrados de lado  $\delta_i$ ; donde  $\delta_i$  es una fracción de  $k$ , de modo que  $\delta_i = k/i$ , para  $i = 1, 2, \dots, \infty$ . Mientras exista mayor o igual cantidad de puntos que de cuadrados de lado  $\delta_i$  (i.e.  $n \geq \beta_i$ ) todos los cuadrados contienen elementos del patrón puntual, por lo que  $\Gamma(\delta_i) = \beta_i$ , para todo  $\beta_i \leq n$ .

Sea  $\Gamma(\delta_i)$ , el número de celdas de lado  $\delta_i$ , que contienen al menos un elemento del patrón puntual,

$$\begin{aligned}
 \log(\Gamma(\delta_i)) &= \log(\beta_i) \\
 &= \log\left(\frac{k^2}{\delta_i^2}\right) \\
 &= 2\log(k) + 2\log\left(\frac{1}{\delta_i}\right). \tag{2.17}
 \end{aligned}$$

Note que (2.17) tiene forma similar a (1.20) y de allí se puede deducir  $D_B = 2$ .

$D_B$  para un patrón puntual totalmente regular (2.17) no depende de  $n$ , solamente de  $k$ . La Figura 2.6 evidencia el efecto de la variación de  $k$ . Un aumento en el tamaño inicial del área de interés de lado  $k$ , se traduce en una traslación de la recta a la izquierda.

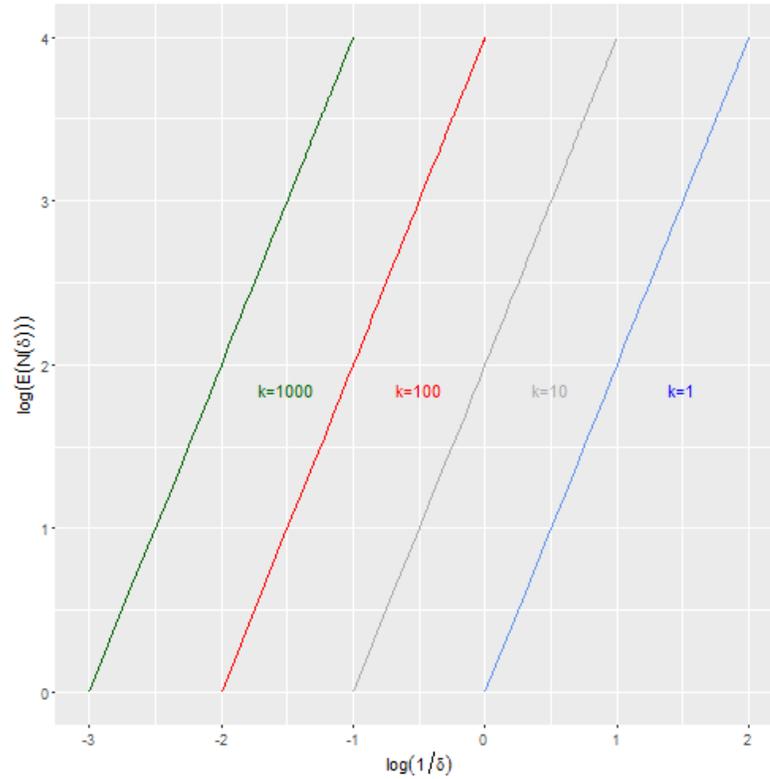


FIGURA 2.6. Gráfico *log-log* obtenido a través de la ecuación (2.17) para un patrón puntual totalmente regular, se dibujan cuatro curvas manteniendo constante  $n = 100$  y con diferentes valores  $k=1,10,100$  y  $1000$ .

### 2.3. Comparación del comportamiento de la curva bajo AEC con la recta teórica de un patrón puntual regular

La Figura 2.7 muestra una comparación entre el gráfico de un patrón regular (2.17) y el de otro aleatorio (2.8) para el mismo tamaño de muestra  $n = 100$  y  $k = 1$ . La curva para los patrones de inhibición se comporta de manera similar al caso regular y se ubica por encima de la curva bajo AEC.

### 2.4. Código e interpretación de la prueba gráfica basada en dimensión *Box Counting*

Para evaluar el comportamiento de la prueba propuesta se creó una función en R [54] (apéndice 4.3). Esta recibe como parámetros dos vectores de coordenadas  $x$  y  $y$  de un patrón puntual observado. En primer lugar, calcula los valores de  $\Gamma(\delta_i)$  observados para cada  $\delta_i = k/i$ ,  $i = 1, 2, \dots, n$ , retorna una matriz de tamaño  $n \times 2$  que contiene en la primera columna los valores calculados  $\log(1/\delta_i)$  y en la segunda columna los valores  $\log(\Gamma(\delta_i))$ . Posteriormente realiza una prueba de Monte Carlo con  $s = 100$  simulaciones de un patrón puntual bajo AEC (i.e. 95% de significancia), con valores  $n$  y  $k$  del patrón pun-

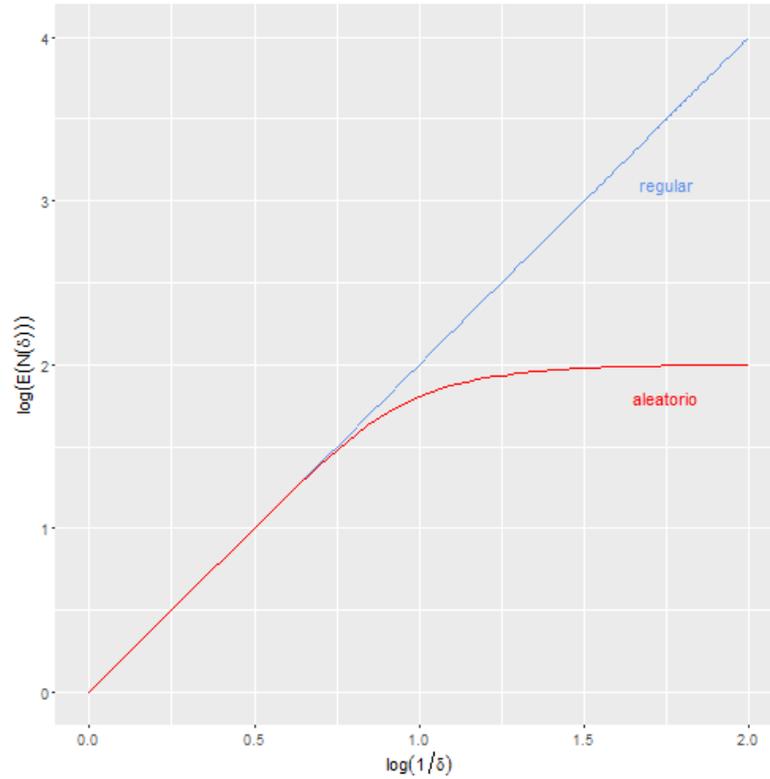


FIGURA 2.7. El gráfico muestra la comparación entre la recta *log-log* para un patrón regular, calculada a través de la ecuación (2.17) y la curva bajo AEC obtenida a través de la ecuación (2.8), ambos casos con  $k = 1$  y  $n = 100$ .

tual observado, para cada una de las simulaciones realiza el cálculo de los valores  $\log(\Gamma(\delta_i))$ .

Para cada valor  $\log(1/\delta_i)$  se calcula  $\log(\Gamma(\delta_i))$  y se construyen las bandas de confianza definiendo como limite inferior el resultado ubicado con  $\alpha = 0.025\%$  y como limite superior el resultado ubicado con  $\alpha = 0.975\%$ . Como resultado la función dibuja las bandas de confianza en gris, calcula y dibuja la gráfica teórica bajo AEC según (2.8), muestra en azul la curva para el patrón puntual observado y calcula los siguientes valores para la curva aleatoria: cruce de la curva teórica con el eje  $x$ , valores de los ejes  $x$  y  $y$  cuando  $\tau_i = 1$  y dimensión *Box Counting*  $D_{B1}$ .

La Figura 2.8 y la Tabla 2.2, muestran el resultado de la aplicación de la prueba propuesta para el caso de un patrón puntual regular con  $n = 100$  y  $k = 1$  (Figura 2.5). En gris las bandas de confianza y en verde la curva bajo AEC. Para que no se rechace la hipótesis nula, la curva del patrón puntual observado (en azul) debe encontrarse al interior de las bandas de confianza. Una curva por encima de las bandas de confianza, como en éste caso, indica que las observaciones siguen un patrón regular o de inhibición que también resulta en una pendiente mayor para la primera parte de la curva y por tanto mayor dimensión *Box Counting*. Una curva por debajo de las bandas de confianza revelaría que los datos corresponden a una realización de un patrón agregado (*cluster*).

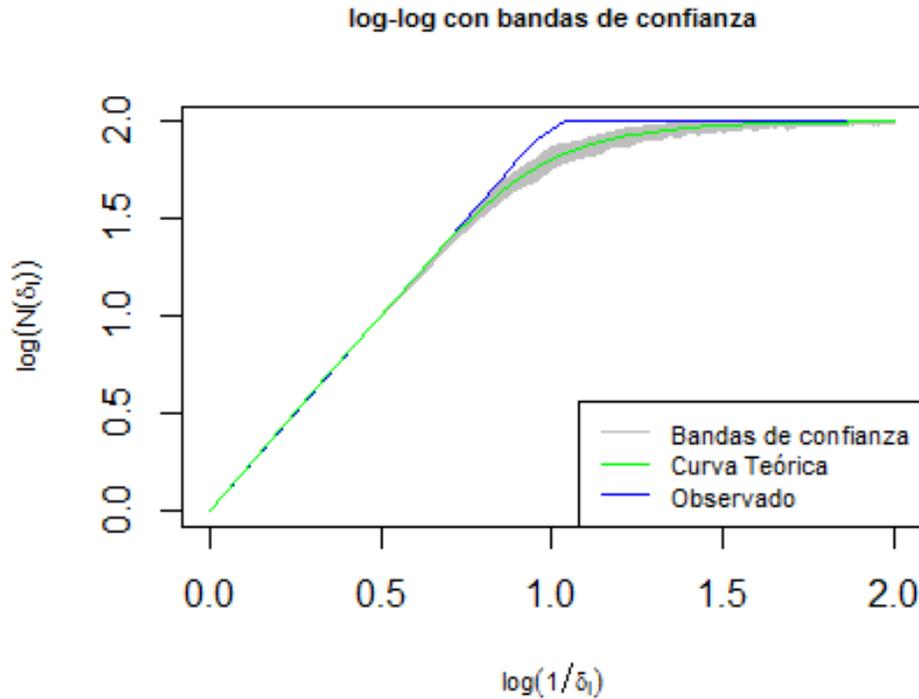


FIGURA 2.8. Resultado de la prueba gráfica para evaluar la hipótesis de AEA, para el caso de un patrón regular con  $k = 1$  y  $n = 100$ , la curva observada se ubica por encima de las bandas de confianza, por tanto se rechaza la hipótesis nula.

#### 2.4.1. Prueba de AEC basada en dimensión *Box Counting* y el efecto de borde

Nótese que la metodología *Box Counting* no se encuentra basada en distancias, sino en particiones de la ventana de observación  $A$ . Por tanto, para la construcción de la prueba no es necesario tener en cuenta información por fuera de  $A$ . De allí que es necesario considerar y realizar correcciones al efecto de borde en la prueba gráfica basada en el método de *Box Counting*.

| Valores bajo AEC                               | Calculado |
|--|-----------|
| Cruce con el eje $x$                           | 0         |
| Valor del eje $x$ cuando $\tau = 1$            | 0.8116246 |
| Valor del eje $y$ cuando $\tau = 1$            | 1.4240492 |
| Primera dimensión <i>Box Counting</i> $D_{B1}$ | 1.7545662 |

TABLA 2.2. Valores resultado de la prueba de AEC basada en dimensión *Box Counting*, para el caso de un patrón homogéneo de Poisson con valores  $n = 100$  y  $k = 1$ .

---



---

## Estudio de Simulación

---



---

Usando la librería *Spatstat* [9] del software R [54], se simulan patrones puntuales homogéneos, regulares y agregados. En cada caso se evalúa la hipótesis de AEC, mediante la prueba propuesta basada en el cálculo de la dimensión *Box Counting*, utilizando la función descrita en el apéndice 4.3. El resultado se compara con los obtenidos a través de pruebas gráficas basadas en las funciones  $F$ ,  $G$  y  $K$ , respectivamente.

### 3.1. Patrón puntual homogéneo de *Poisson*

Se simularon cuatro realizaciones de procesos homogéneos de Poisson (i.e. completamente aleatorios), con diferentes tamaños de muestra  $n = 39, 466, 4895, 498953$ . El resultado se presenta en la Figura 3.1. Para el caso de  $n = 498953$  ya no es posible distinguir los puntos por separado. La función de intensidad estimada de cada uno de estos se muestra en la figura 3.2.

Se realizó la prueba para evaluar la hipótesis de AEC basada en el cálculo de la dimensión *Box Counting*. El resultado en los cuatro casos se puede apreciar en la Figura 3.3 y en la Tabla 3.1. En los cuatro casos simulados la curva observada se encuentra al interior de las bandas de confianza (color gris en la gráfica), por tanto se concluye que no existe evidencia para rechazar la hipótesis de AEC. Para el caso  $n = 39$  en la Figura 3.3A son más evidentes las bandas de confianza ya que el rango de la gráfica es menor. Cuando aumenta el rango de cálculo de la dimensión *Box Counting*, es decir, cuando aumenta el tamaño de la muestra, se hace más difícil la percepción de las bandas, aunque un acercamiento en una ventana particular permitiría corroborar el resultado.

Para cada conjunto de datos simulado se realizan también las pruebas gráficas de AEC con base en las funciones  $F$ ,  $G$  y  $K$ , con el fin de comparar los resultados con los obtenidos mediante *Box Counting*. Las Figuras 3.4 a 3.6 muestran el resultado de las pruebas gráficas basadas en las funciones  $F$ ,  $G$  y  $K$  para los cuatro conjuntos de datos simulados. En todos los casos la curva observada se ubica al interior de las bandas de confianza (color gris en las gráficas). Por tanto, no existe evidencia para rechazar la hipótesis nula y se concluye que el patrón se distribuye aleatoriamente en la ventana de observación.

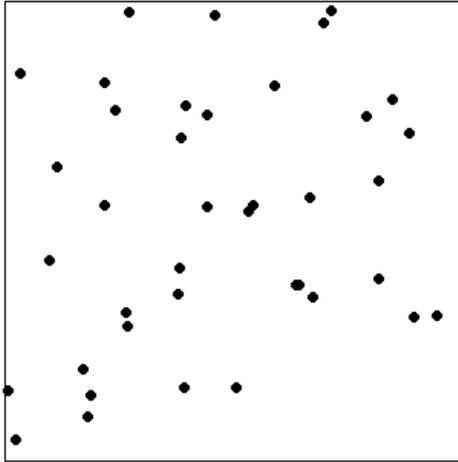
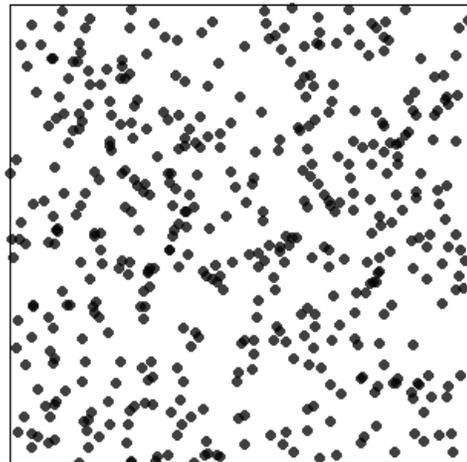
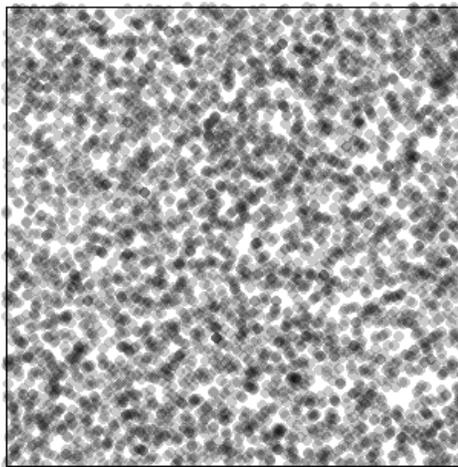
(A) Caso homogéneo de *Poisson*,  $n = 39$ .(B) Caso homogéneo de *Poisson*,  $n = 466$ .(C) Caso homogéneo de *Poisson*,  $n = 4895$ .(D) Caso homogéneo de *Poisson*,  $n = 498953$ .

FIGURA 3.1. Simulaciones de patrón puntual aleatorio (i.e. homogéneo de *Poisson*), con diferentes tamaños de muestra  $n=39,466,4895$  y  $498953$ .

Para el caso  $n = 498953$  el cálculo de las bandas de confianza para la función  $F$  presenta problemas desde el punto de vista computacional. Con tamaños de muestra mayores a 3000, la prueba  $K$  no calcula corrección al efecto borde en su lugar realiza un estimado.

De los resultados anteriores se puede concluir que la prueba gráfica basada en *Box Counting* es satisfactoria, independientemente del tamaño de muestra. En éste último sentido podría ser preferible respecto a la prueba gráfica  $F$  que es computacionalmente costosa (con el software R). Por otra parte, ya que la prueba propuesta no requiere hacer corrección al efecto de borde, para tamaños de muestra mayores a 3000 puede ser más precisa que la prueba gráfica basada en la función  $K$ .

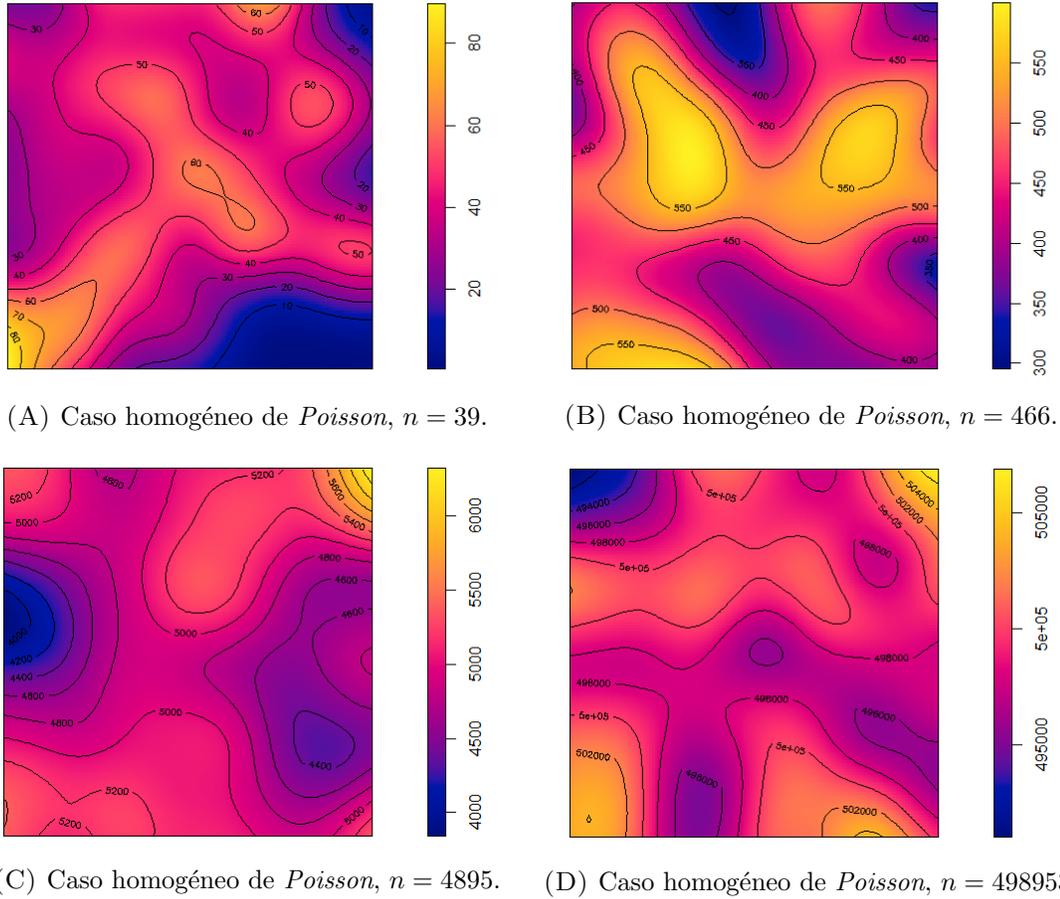


FIGURA 3.2. Gráfica de la funciones de intensidad para los casos de patrones aleatorios simulados en la Figura 3.1, con diferentes tamaños de muestra  $n=39,466,4895$  y  $498953$ .

| $n$    | Cruce de la curva $\log\text{-}\log$ con el eje $x$ | Valor en el eje $x$ cuando $\lambda = 1$ | Valor en el eje $y$ cuando $\lambda = 1$ | Primera dimensión fractal $D_{B1}$ |
|--------|---|--|--|------------------------------------|
| 39     | 0   | 0.795532                                 | 1.391865                                 | 1.749606                           |
| 466    | 0   | 1.334193                                 | 2.469186                                 | 1.850696                           |
| 4895   | 0   | 1.844885                                 | 3.490553                                 | 1.892025                           |
| 498953 | 0   | 2.849485                                 | 5.499770                                 | 1.930093                           |

TABLA 3.1. Valores del resultado de la prueba basada en *Box Counting*, para patrones bajo AEC, con diferentes tamaños de muestra,  $n= 39,466,4895$  y  $498953$ , se puede apreciar el aumento de la primera dimensión fractal ( $D_{B1}$ ) a medida que el tamaño de muestra es mayor.

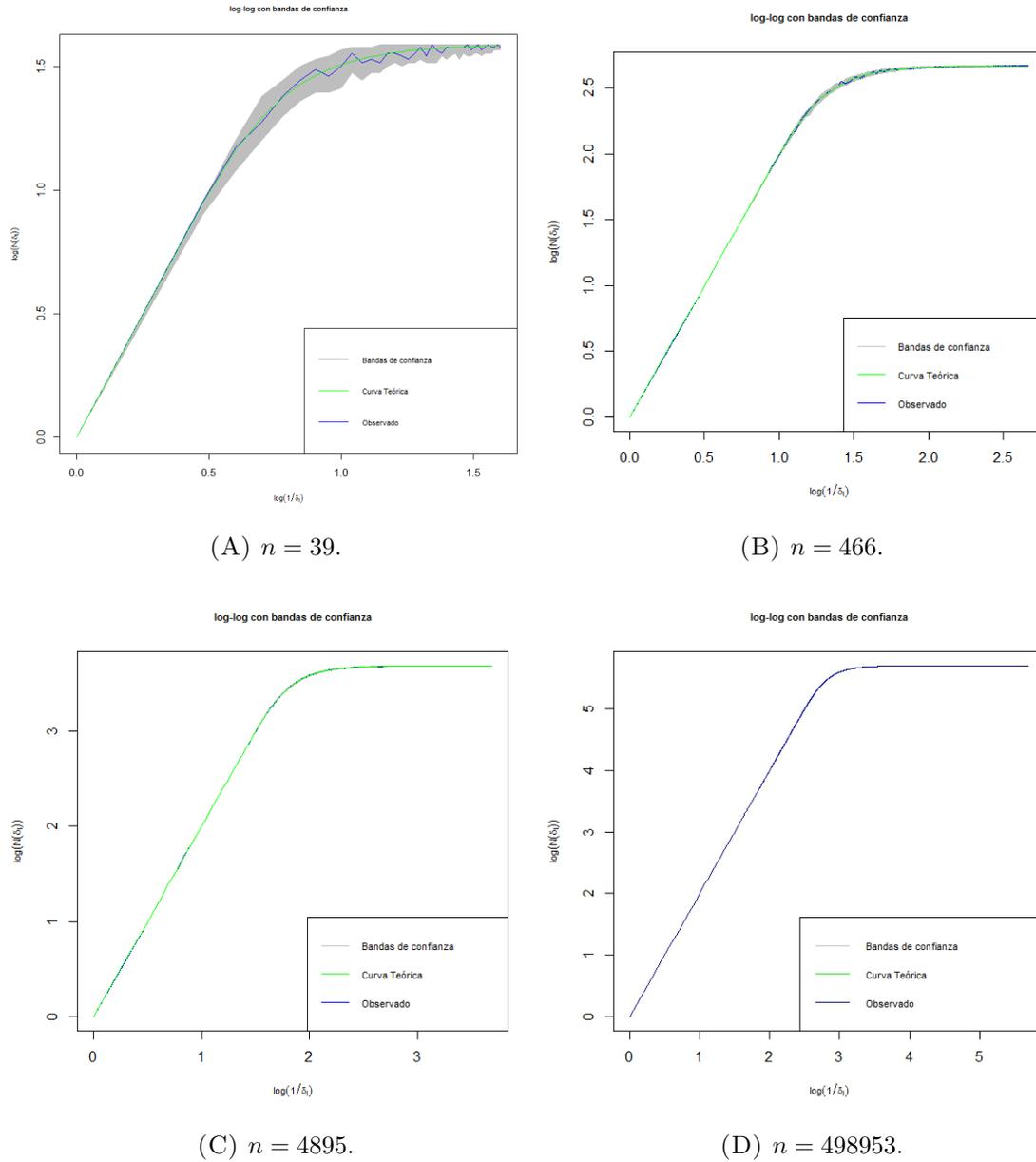
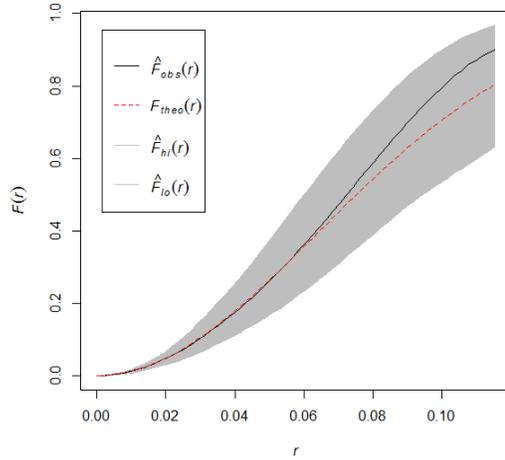
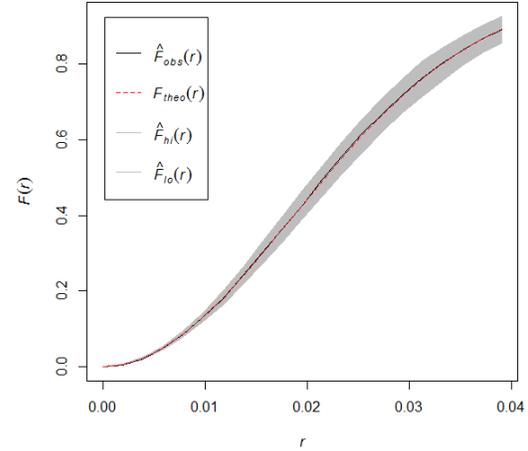


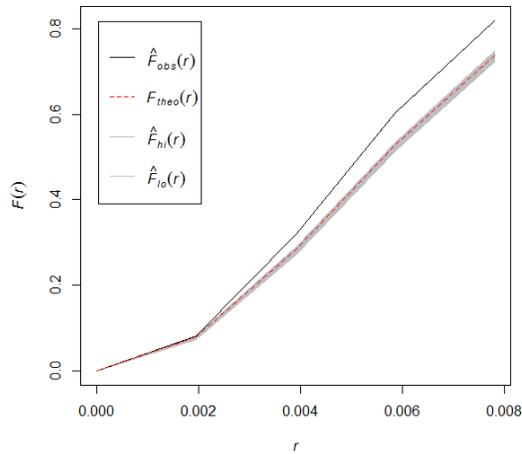
FIGURA 3.3. Prueba gráfica para evaluar la hipótesis de AEC a través de la dimensión *Box Counting*, para los casos simulados en la Figura 3.1, con  $n=39,466,4895$  y  $498953$ . En todos los casos la curva observada se mantiene al interior de las bandas de confianza, por tanto, se concluye que no hay evidencia para rechazar la hipótesis nula.



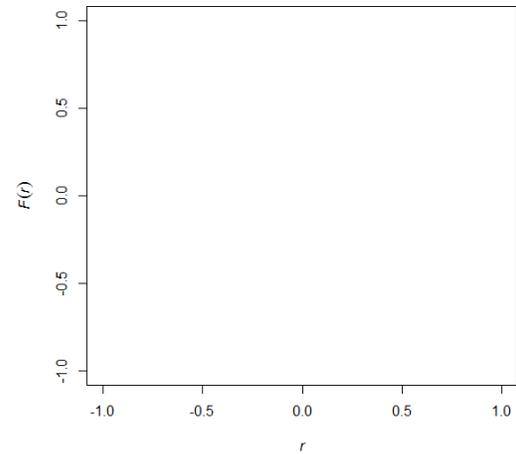
(A) Función  $F$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 39$ .



(B) Función  $F$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 466$ .

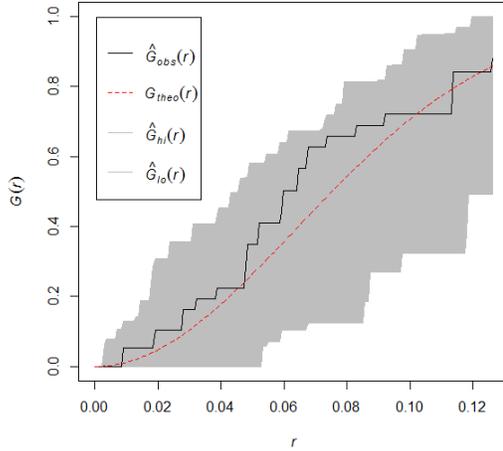


(C) Función  $F$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 4895$ .

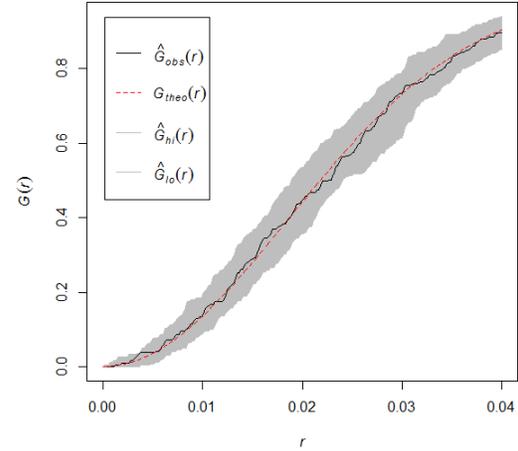


(D) Función  $F$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 498953$ .

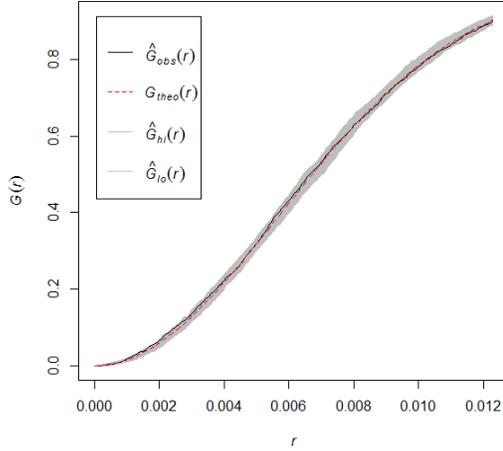
FIGURA 3.4. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para cuatro simulaciones de patrón puntual de *Poisson*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{F}_{lo}(r)$ ,  $\hat{F}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ .



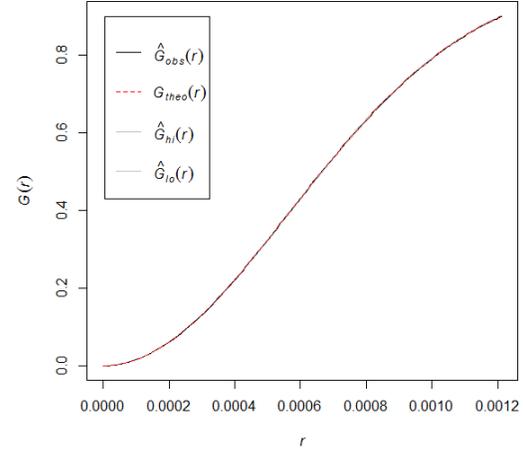
(A) Función  $G$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 39$ .



(B) Función  $G$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 466$ .

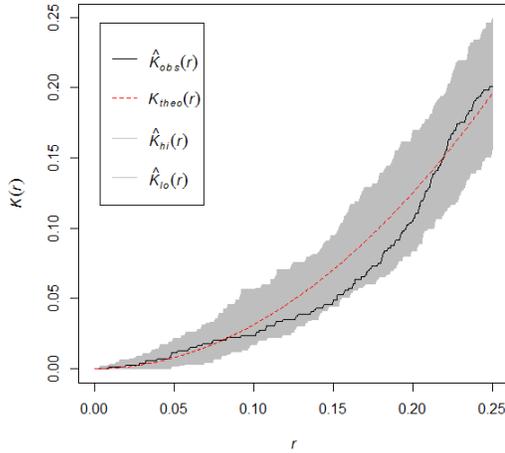


(C) Función  $G$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 4895$ .

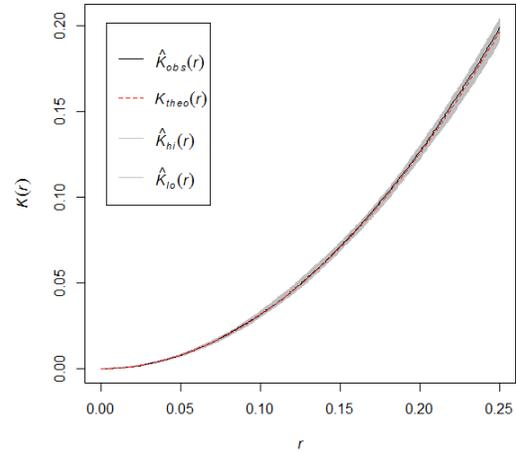


(D) Función  $G$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 498953$ .

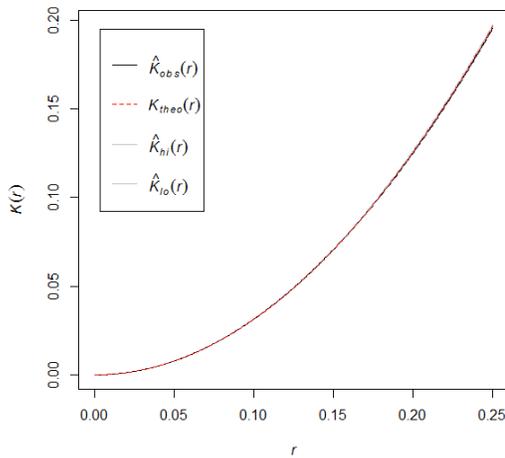
FIGURA 3.5. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para cuatro simulaciones de patrón puntual de *Poisson*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{G}_{l_0}(r)$ ,  $\hat{G}_{h_i}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ .



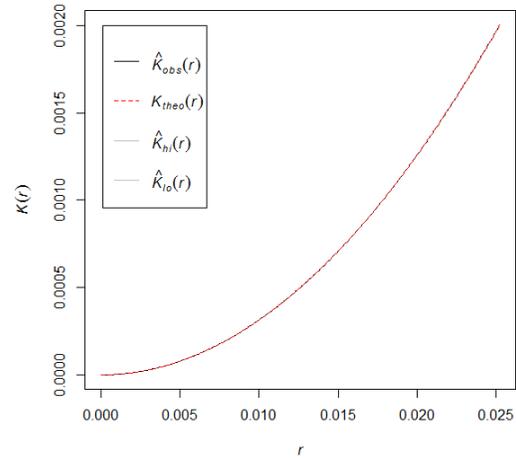
(A) Función  $K$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 39$ .



(B) Función  $K$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 466$ .



(C) Función  $K$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 4895$ .



(D) Función  $K$  con bandas de confianza para una simulación de patrón puntual de *Poisson* con  $n = 498953$ .

FIGURA 3.6. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para cuatro simulaciones de patrón puntual de *Poisson*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{K}_{lo}(r)$ ,  $\hat{K}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ .

## 3.2. Simulación de un patrón de inhibición

Se simularon cuatro realizaciones de proceso de inhibición estacionario utilizando el Modelo II de Matérn en una ventana de observación de una unidad de área. Se procuró que los tamaños de muestra resultantes para cada caso sean similares a los obtenidos en la sección 3.1, en este caso los tamaños de muestra obtenidos son  $n=50, 441, 6952$  y  $456038$ .

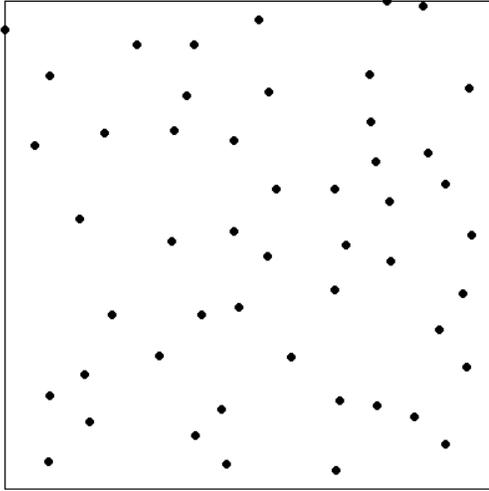
Este proceso se genera partiendo de la simulación de un proceso uniforme de *Poisson* con intensidad ( $\lambda$ ). Para cada punto se define un tiempo de llegada que corresponde a un número uniformemente distribuido entre  $[0, 1]$ ; un punto es borrado del conjunto de datos si se encuentra a una distancia menor de la distancia de inhibición  $r_i$  (parámetro que se determina durante la simulación) de otro punto con tiempo de llegada anterior [9].

La Figura 3.7 muestra la distribución espacial de los patrones simulados. Las respectivas gráficas de la función de intensidad estimada se dan en la Figura 3.8.

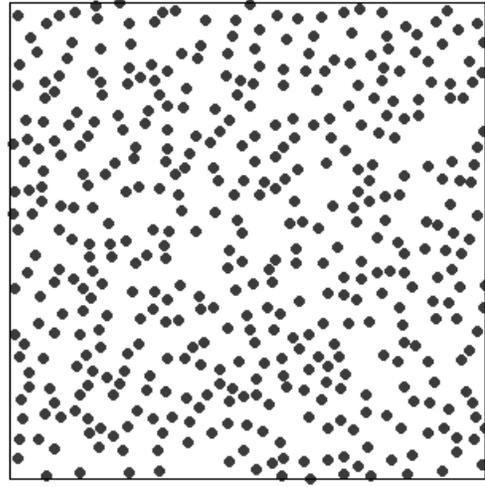
En primer lugar, se realizó en cada caso la prueba gráfica para evaluar la hipótesis de AEC mediante dimensión *Box Counting*. Su resultado se puede evidenciar en la Figura 3.9. Se aprecia, como es de esperarse, que la curva (con los cuatro tamaños de muestra) se ubica por encima de las bandas de confianza. Como conclusión se rechaza la hipótesis nula en favor de la hipótesis de inhibición.

Con el fin de verificar los resultados se realizaron las pruebas gráficas para evaluar la hipótesis de AEC basadas en las funciones  $F$ ,  $G$  y  $K$ . La gráfica resultante para el caso de la función  $F$  con tamaño de muestra  $n = 50$  no aportó evidencia suficiente para rechazar la hipótesis nula (Figura 3.10A), para los tamaños de muestra  $n = 441$  y  $6952$  (Figuras 3.10B y 3.10C) los resultados permiten rechazar la hipótesis nula y asumir que los datos siguen un patrón de inhibición. Por otra parte, en el caso  $n = 456038$  el cálculo de las bandas de confianza para la función  $F$  presenta problemas desde el punto de vista computacional (Figura 3.10D).

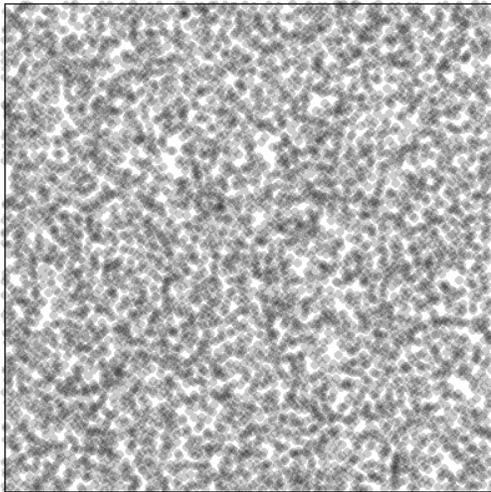
Por otra parte, la prueba gráfica basada en la función  $G$  permite rechazar la hipótesis nula en todos los casos (Figuras 3.11A a 3.11D) en favor de que los datos siguen un patrón de inhibición. Finalmente la prueba gráfica basada en la función  $K$  (Figuras 3.12A a 3.12D) permite rechazar la hipótesis de AEA para los tamaños de muestra  $n = 50$  y  $n = 441$  pero no aporta evidencia suficiente para rechazar la hipótesis nula en los casos  $n = 6952$  y  $n = 456038$ . Con tamaños de muestra mayores a 3000, la prueba  $K$  no calcula corrección al efecto borde en su lugar realiza un estimado.



(A) Simulación de patrón de inhibición, *Matérn II* ( $\lambda = 100, r_i = 0.07$ ), tamaño de muestra  $n = 50$ .



(B) Simulación de patrón de inhibición, *Matérn II* ( $\lambda = 1000, r_i = 0.025$ ), tamaño de muestra  $n = 441$ .

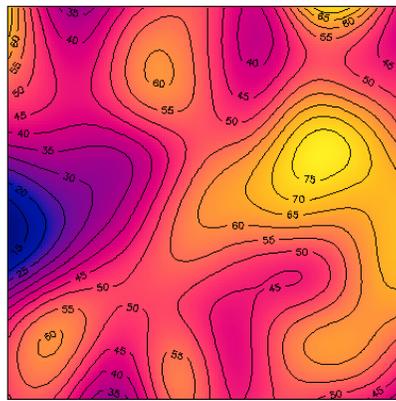


(C) Simulación de patrón de inhibición, *Matérn II* ( $\lambda = 10000, r_i = 0.005$ ), tamaño de muestra  $n = 6952$ .

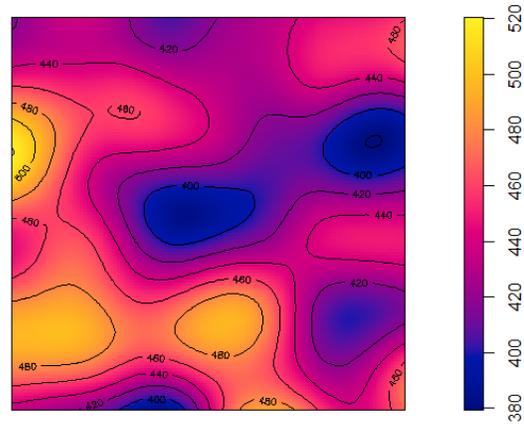


(D) Simulación de patrón de inhibición, *Matérn II* ( $\lambda = 1000000, r_i = 0.0005$ ), tamaño de muestra  $n = 456038$ .

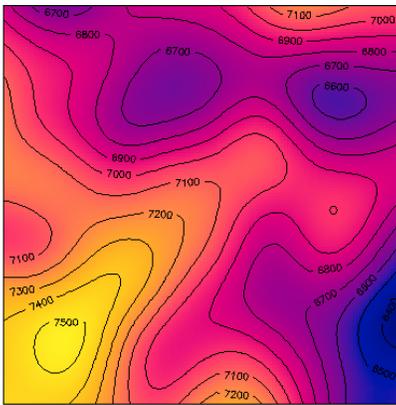
FIGURA 3.7. Simulaciones de patrón puntual de inhibición utilizando el Modelo II de Matérn con diferentes parámetros de intensidad ( $\lambda$ ) y distancia de inhibición ( $r_i$ ).



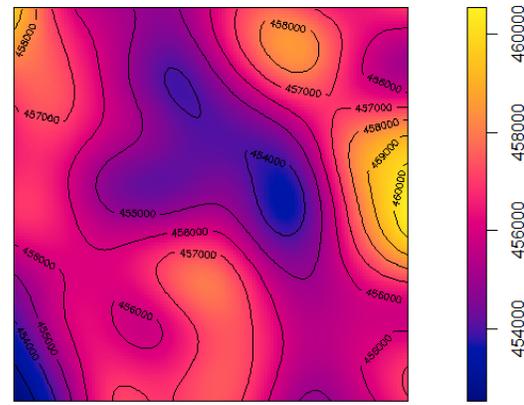
(A) Función de intensidad patrón de inhibición simulado en la Figura 3.8A,  $n = 50$ .



(B) Función de intensidad patrón de inhibición simulado en la Figura 3.7B,  $n = 441$ .

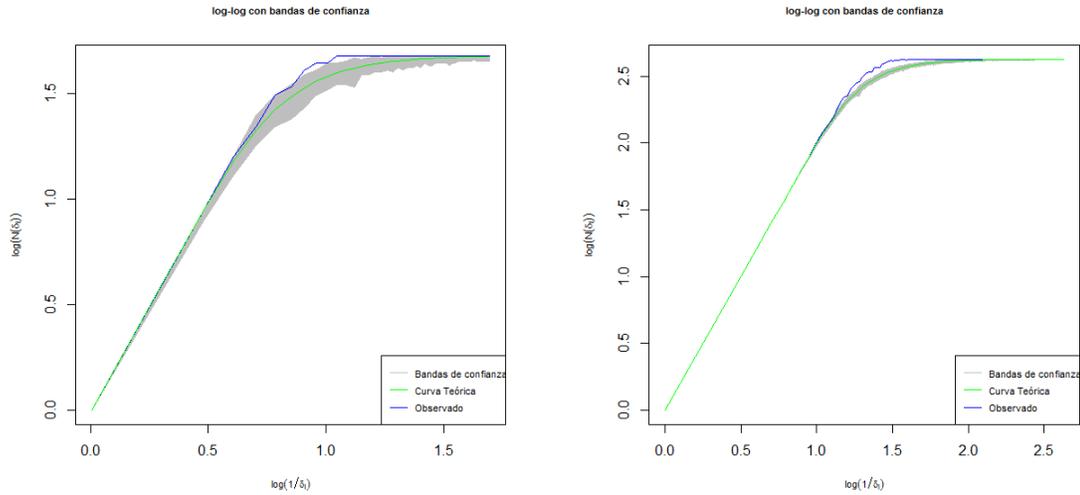


(C) Función de intensidad patrón de inhibición simulado en la Figura 3.7C,  $n = 6952$ .



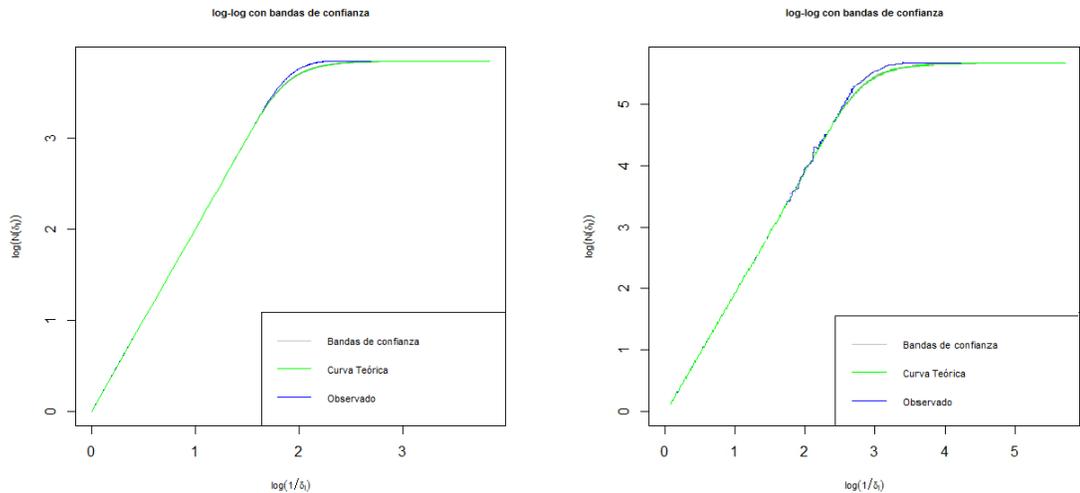
(D) Función de intensidad patrón de inhibición simulado en la Figura 3.7D,  $n = 456038$ .

FIGURA 3.8. Gráficas de la función de intensidad estimada para las simulaciones de patrón de inhibición de la Figura 3.7.



(A) Prueba gráfica para evaluar la hipótesis de AEC, a través de la dimensión *Box Counting*, caso de patrón de inhibición simulado con tamaño de muestra  $n = 50$ .

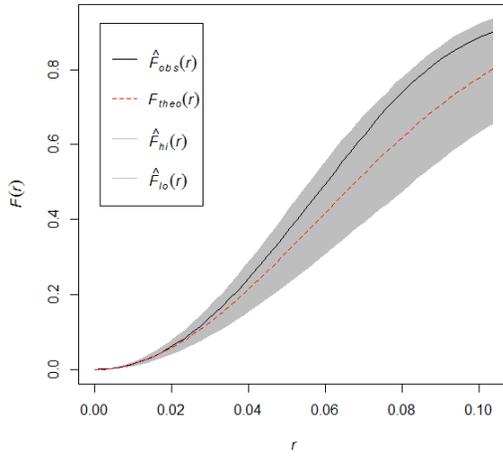
(B) Prueba gráfica para evaluar la hipótesis de AEC, a través de la dimensión *Box Counting*, caso de patrón de inhibición simulado con tamaño de muestra  $n = 441$ .



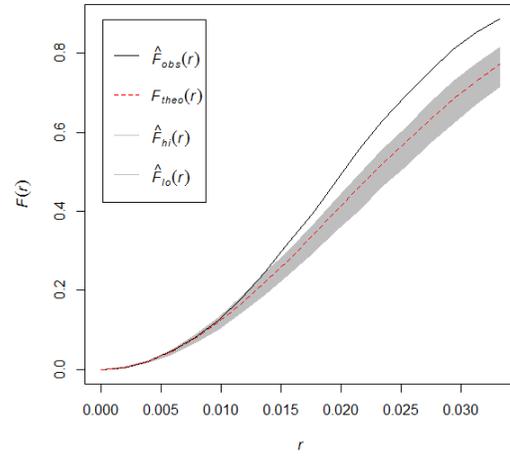
(C) Prueba gráfica para evaluar la hipótesis de AEC, a través de la dimensión *Box Counting*, caso de patrón de inhibición simulado con tamaño de muestra  $n = 6952$ .

(D) Prueba gráfica para evaluar la hipótesis de AEC, a través de la dimensión *Box Counting*, caso de patrón de inhibición simulado con tamaño de muestra  $n = 456038$ .

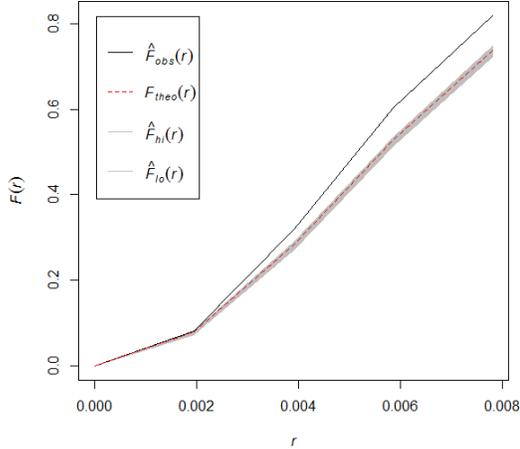
FIGURA 3.9. Resultados de la prueba gráfica para evaluar la hipótesis de AEC, a través de dimensión *Box Counting*, para los casos de patrones de inhibición simulados en la Figura 3.7, con diferentes tamaños de muestra. En todos los casos se rechaza la hipótesis nula, teniendo en cuenta que la curva observada se ubica por encima de las bandas de confianza.



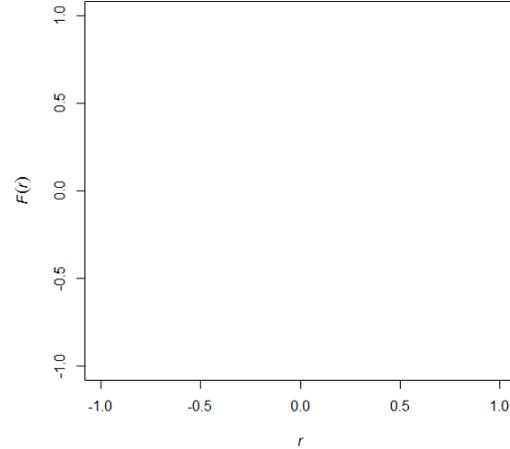
(A) Función  $F$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 50$ .



(B) Función  $F$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 441$ .

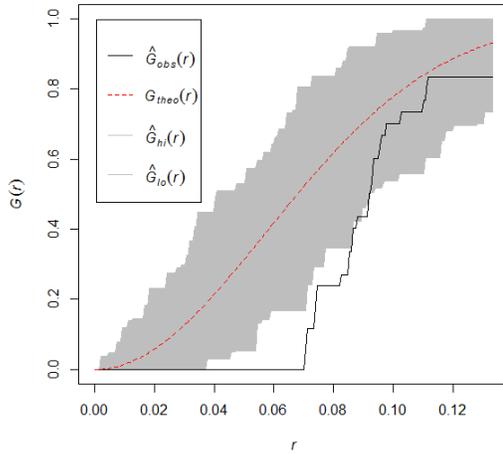


(C) Función  $F$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 6952$ .

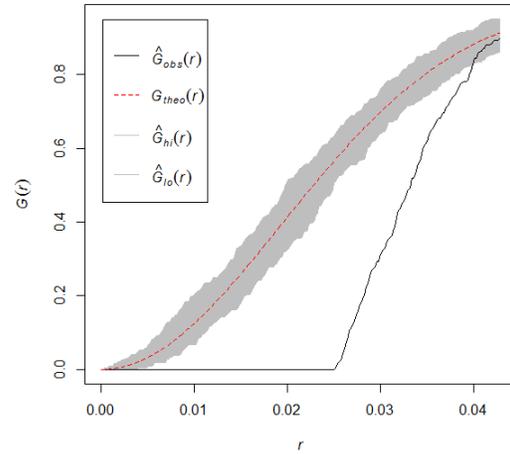


(D) Función  $F$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 456038$ .

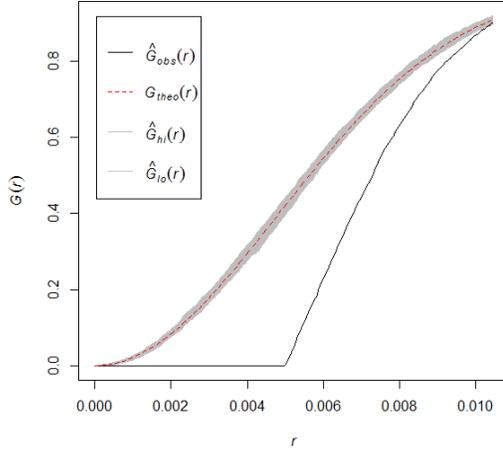
FIGURA 3.10. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para cuatro simulaciones de patrón puntual de inhibición *Matérn II*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{F}_{lo}(r)$ ,  $\hat{F}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ .



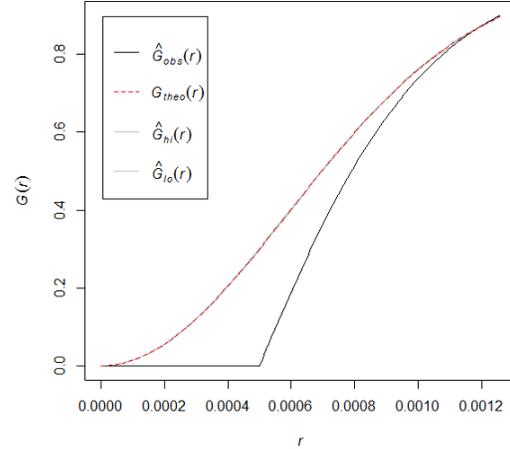
(A) Función  $G$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 50$ .



(B) Función  $G$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 441$ .

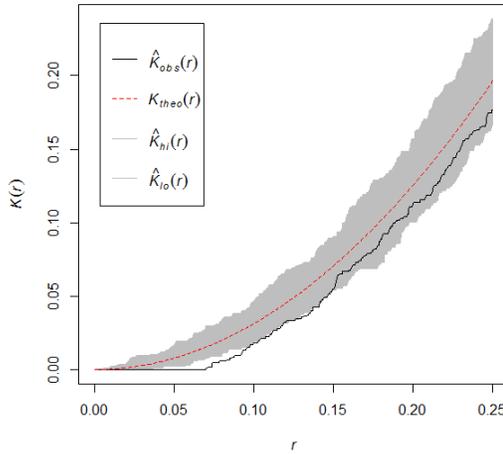


(C) Función  $G$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 6952$ .

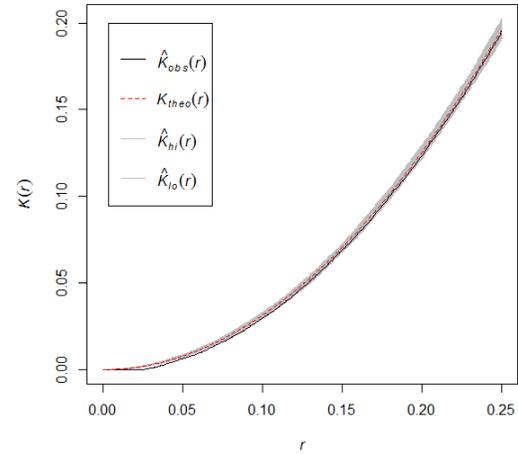


(D) Función  $G$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 456038$ .

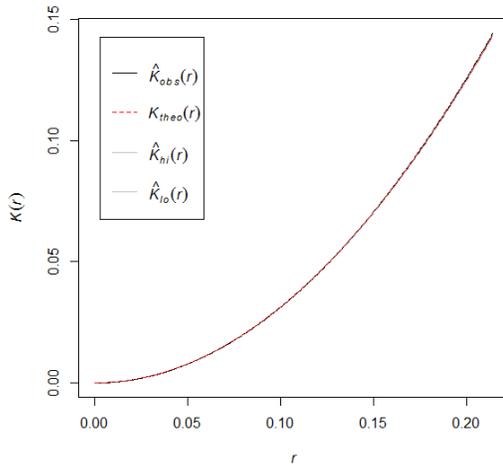
FIGURA 3.11. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para cuatro simulaciones de patrón puntual de inhibición *Matérn II*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{G}_{lo}(r), \hat{G}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ .



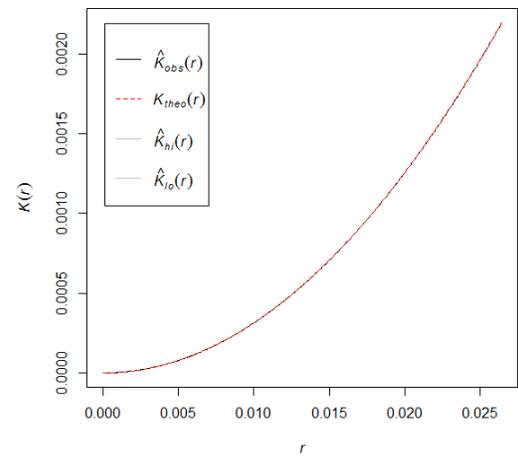
(A) Función  $K$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 50$ .



(B) Función  $K$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 441$ .



(C) Función  $K$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 6952$ .



(D) Función  $K$  con bandas de confianza para una simulación de patrón puntual de inhibición *Matérn II*, tamaño de muestra  $n = 456038$ .

FIGURA 3.12. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para cuatro simulaciones de patrón puntual de inhibición *Matérn II*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo  $[\hat{K}_{lo}(r), \hat{K}_{hi}(r)]$ . En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ .

### 3.3. Simulación de patrón agregado

#### 3.3.1. Patrón Puntual *Thomas*

Para evaluar el desempeño de la prueba propuesta bajo agregación espacial y comparar está misma con las clásicas (basadas en funciones  $F$ ,  $G$  y  $K$ ) se simularon dos casos particulares de proceso no homogéneo.

Neyman (1939), propuso un modelo basado en larvas de insectos. Que consiste en un proceso de *Poisson* para grupos de larvas (hijos) alrededor de cada insecto (padre) [25].

Supóngase que ocurren  $m$  eventos padres, con una distribución  $Poisson(\lambda)$  y que cada padre independientemente produce un número aleatorio de eventos hijos  $S$ , todos ellos ocupando la misma posición que su padre. El número de padres  $m$  en una región  $A$  sigue una distribución de Poisson con media  $\lambda|A|$ . El número de hijos en  $A$ ,  $N(A)$  es igual a  $S_1 + \dots + S_M$ , y si la función generadora de probabilidad de  $S$  es  $f_S(z)$ , entonces la función generadora de probabilidad de  $N(A)$  está dada por [25]:

$$f(z; A) = e^{-\lambda|A|\{1-f_S(z)\}}.$$

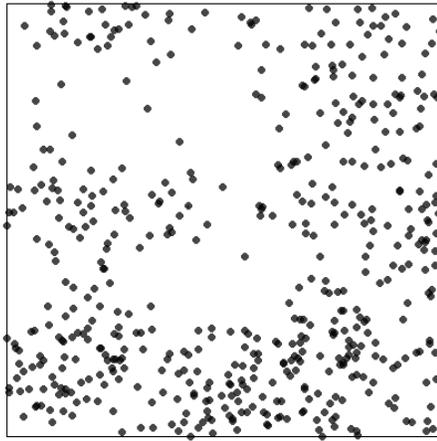
El proceso de *Thomas* es una variación del proceso *Neyman* que consiste en agregaciones de patrones distribuidos aleatoriamente, cada hijo desplazado de su padre una distancia isotrópica Gaussiana  $N(0, \sigma^2 I)$  [6] e incluye los padres en el patrón final, esto evita agregaciones de tamaño cero [25].

Se simuló un proceso *Thomas* con parámetros  $\lambda = 30$ ,  $\sigma = 0.08$  y  $\mu = 20$  ( $n = 524$ ) y otro con  $\lambda = 200$  y  $\mu = 200$  manteniendo  $\sigma$  ( $n = 41832$ ). Los patrones resultantes se presentan en la Figuras 3.13A y 3.13C y sus correspondientes funciones de intensidad estimadas en las Figuras 3.13B y 3.13D, respectivamente. Tanto la inspección de los patrones como los gráficos de las funciones de intensidad estimada señalan que existe agregación espacial.

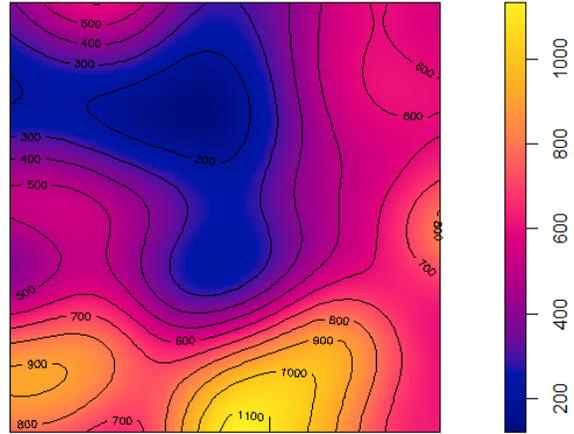
Con ambas simulaciones se obtuvo el gráfico basado en *Box Counting* (Figuras 3.14A y 3.14B). Estos indican que debe rechazarse la hipótesis de aleatoriedad en favor de la agregación (puesto que las curvas están por debajo de las bandas de confianza).

Los gráficos correspondientes a las funciones  $F$ ,  $G$  y  $K$  (Figuras 3.15A a 3.17B) permiten llegar a la misma conclusión.

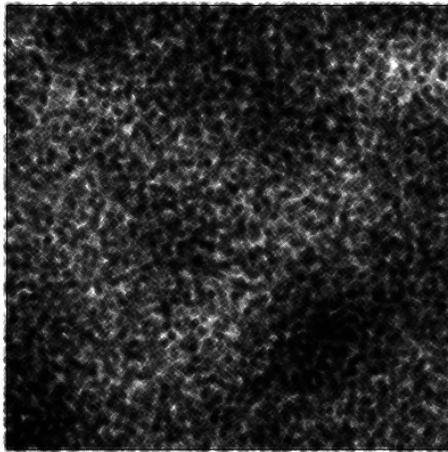
La conclusión general de la simulación en el caso del patrón agregado es igual a los casos anteriores (aleatorio e inhibición). Es decir, la prueba basada en Box Counting permite obtener resultados análogos a los que arrojan las funciones  $F$ ,  $G$  y  $K$ . En casos en que  $n$  aumenta la prueba propuesta es preferible a la  $F$ .



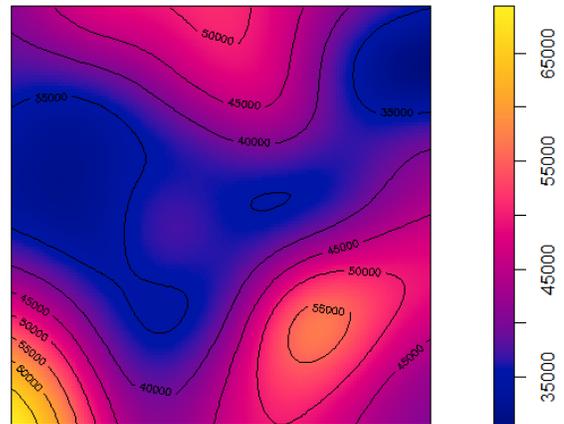
(A) Simulación de patrón agregado,  $Thomas(\lambda = 30, \sigma = 0.08, \mu = 20)$ , tamaño de muestra  $n = 524$ .



(B) Función de intensidad para el patrón simulado en la Figura 3.13A.

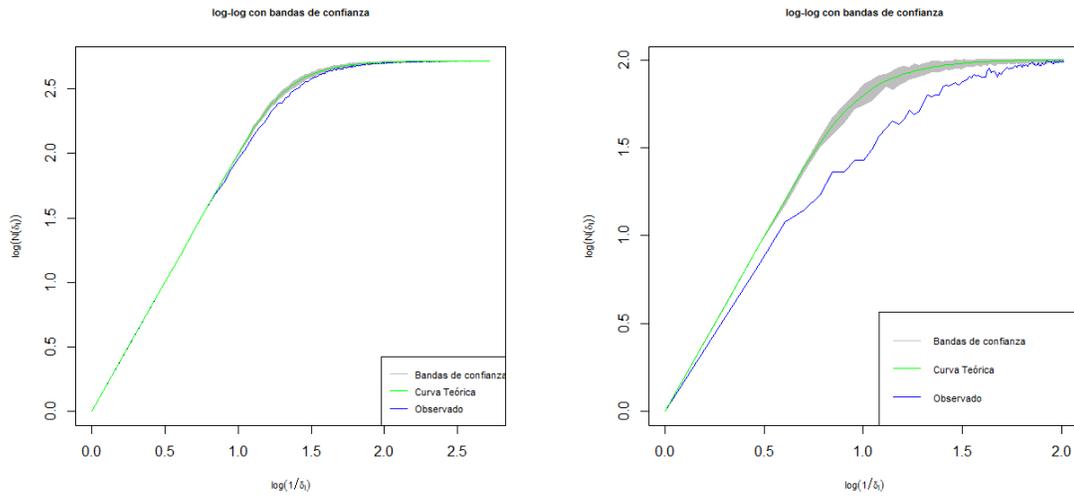


(C) Simulación de patrón agregado,  $Thomas(\lambda = 200, \sigma = 0.08, \mu = 200)$ , tamaño de muestra  $n = 41832$ .



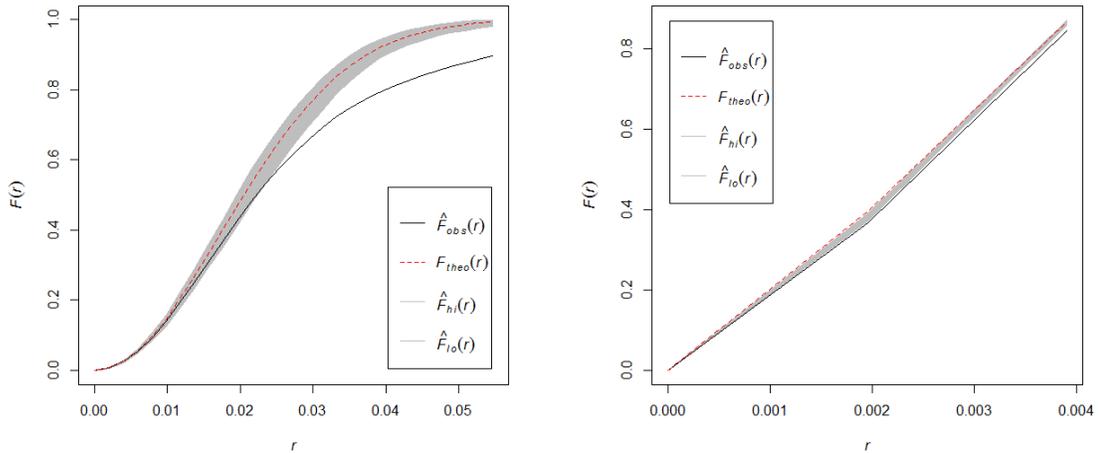
(D) Función de intensidad para el patrón simulado en la Figura 3.13C.

FIGURA 3.13. Las Figuras 3.13A y 3.13C representan dos simulaciones de patrón agregado  $Thomas$  con diferentes parámetros y su función de intensidad en las Figuras 3.13B y 3.13D respectivamente.



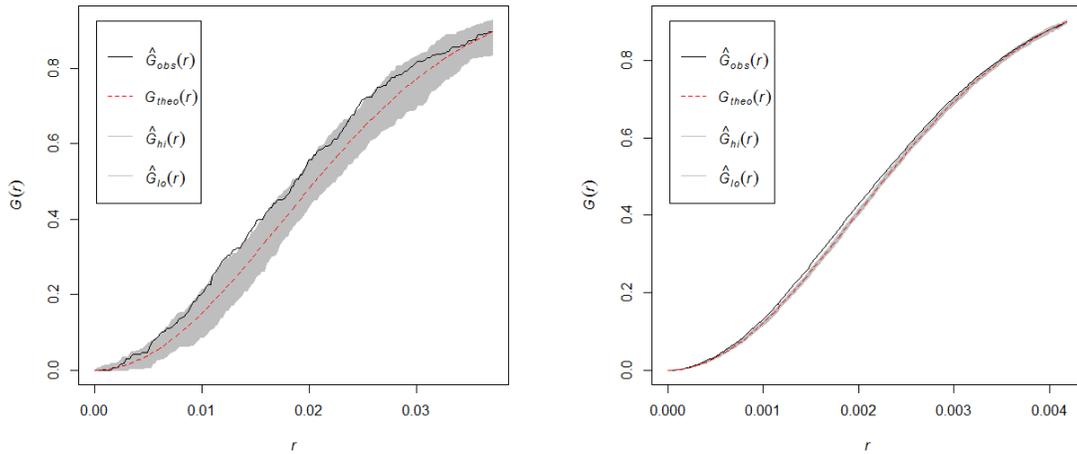
(A)  $Thomas(\lambda = 30, \sigma = 0.08, \mu = 20)$ , tamaño de muestra  $n = 524$ . (B)  $Thomas(\lambda = 200, \sigma = 0.08, \mu = 200)$ , tamaño de muestra  $n = 41832$ .

FIGURA 3.14. Prueba de aleatoriedad espacial para dos simulaciones de patrón puntual *cluster*,  $Thomas(\lambda, \sigma, \mu)$  con diferentes parámetros, mediante el uso de la dimensión *Box Counting*.



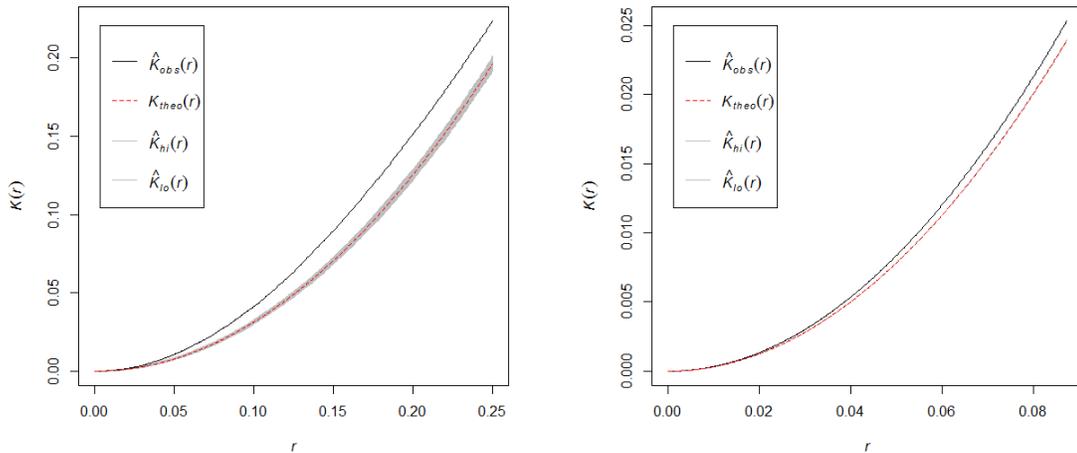
(A) Función  $F$  con bandas de confianza para una simulación de patrón puntual agregado *Thomas*, tamaño de muestra  $n = 524$ . (B) Función  $F$  con bandas de confianza para una simulación de patrón agregado *Thomas*, tamaño de muestra  $n = 41832$ .

FIGURA 3.15. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $F$  para dos simulaciones de patrón puntual agregado *Thomas*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{F}_{lo}(r), \hat{F}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{F}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{F}_{theo}(r)$ .



(A) Función  $G$  con bandas de confianza para una simulación de patrón puntual agregado *Thomas*, tamaño de muestra  $n = 524$ . (B) Función  $G$  con bandas de confianza para una simulación de patrón agregado *Thomas*, tamaño de muestra  $n = 41832$ .

FIGURA 3.16. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $G$  para dos simulaciones de patrón puntual agregado *Thomas*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{G}_{lo}(r)$ ,  $\hat{G}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{G}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{G}_{theo}(r)$ .



(A) Función  $K$  con bandas de confianza para una simulación de patrón puntual agregado *Thomas*, tamaño de muestra  $n = 524$ . (B) Función  $K$  con bandas de confianza para una simulación de patrón agregado *Thomas*, tamaño de muestra  $n = 41832$ .

FIGURA 3.17. Pruebas gráficas para evaluar la hipótesis de AEC con base en la función  $K$  para dos simulaciones de patrón puntual agregado *Thomas*. Se muestran en color gris las bandas de confianza calculadas mediante simulaciones de Monte Carlo [ $\hat{K}_{lo}(r)$ ,  $\hat{K}_{hi}(r)$ ]. En color negro la curva calculada para el patrón observado  $\hat{K}_{obs}(r)$  y en color rojo la curva teórica bajo AEC  $\hat{K}_{theo}(r)$ .

---



---

## Aplicaciones

---



---

### 4.1. Base de datos *Cells*

En la Figura 4.1 se presenta el patrón espacial *Cells* (descrito anteriormente en la sección 1.1.1) y su función de intensidad estimada. Es un ejemplo de patrón de inhibición.

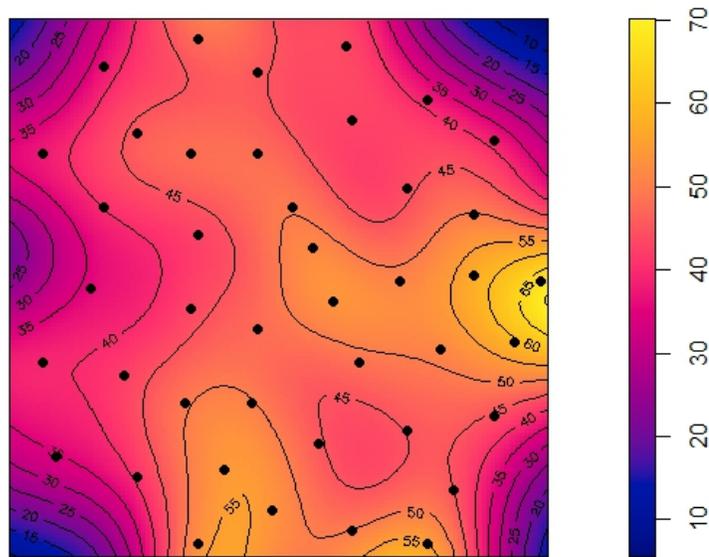


FIGURA 4.1. Gráfica del patrón y función de intensidad para el conjunto de datos *Cells*, contenido en el paquete *Spatstat* ([7] y [55]), representa los centros de 42 células biológicas observadas mediante un microscopio en un corte histológico.

En la sección 1.1.5, se evaluó la hipótesis de AEC para este conjunto de datos a través de pruebas basadas en las funciones  $F$ ,  $G$  y  $K$  (Figuras 1.6B, 1.7B) y 1.5B). Como resultado, en todos los casos se rechazó la hipótesis nula y se concluyó que se trata de un proceso de inhibición.

En la Figura 4.2 se evalúa ahora la hipótesis de AEC empleando la prueba propuesta basada en dimensión *Box Counting*. La curva calculada se ubica por encima de las bandas de confianza, lo que indica que se rechaza la hipótesis nula. La prueba confirma los

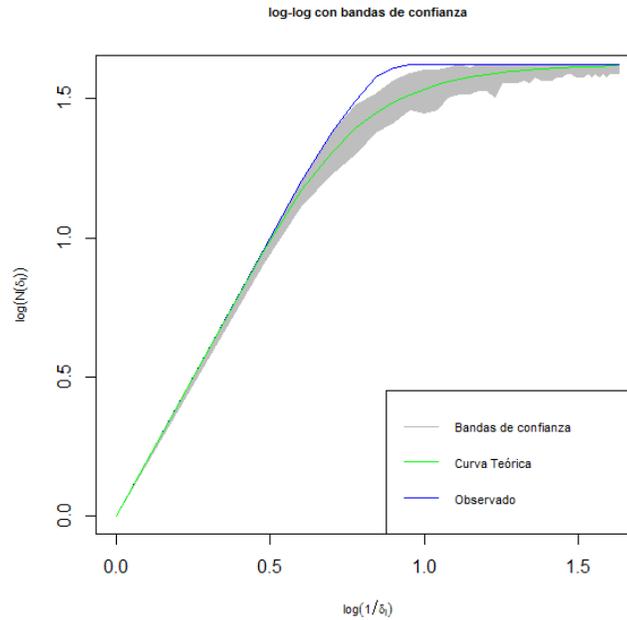


FIGURA 4.2. Contraste gráfico para evaluar la hipótesis de *AEA* a través de dimensión *Box Counting*, al conjunto de datos *Cells* [9]. La curva observada se ubica por encima de las bandas de confianza, razón por la cual se rechaza la hipótesis nula y se concluye que se trata de un patrón de inhibición.

resultados ya conocidos para este conjunto de datos, es decir que se trata de un patrón de inhibición.

## 4.2. Base de datos *Reedwood*

Se trata de un ejemplo típico de patrón agregado (descrito anteriormente en la sección 1.1.1). Su distribución espacial y función de intensidad estimada se observa en la Figura 4.3.

Se evalúa la hipótesis de AEC mediante la prueba propuesta basada en dimensión *Box Counting* (Figura 4.4). La curva calculada se ubica por debajo de las bandas de confianza, por tanto se rechaza la hipótesis nula. En la sección 1.1.5, se evaluó la hipótesis de AEC para este conjunto de datos a través de pruebas basadas en las funciones  $F$ ,  $G$  y  $K$  (Figuras 1.6C, 1.7C y 1.5C). En todos los casos se rechazó la hipótesis nula y se concluyó que se trata de un proceso de agregado.

## 4.3. Base de datos *Longleaf Pines*

Otro ejemplo descrito anteriormente en la sección 1.1.1 es el patrón *Longleaf Pines* la Figura 4.5 muestra su distribución espacial y función de intensidad estimada. No es

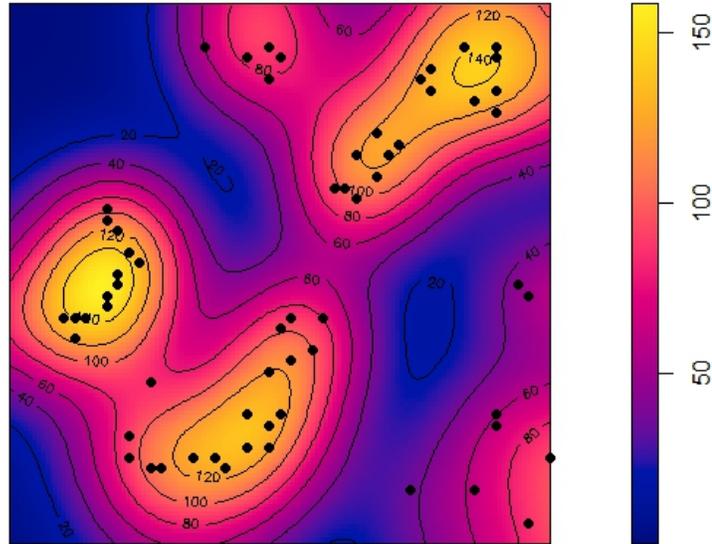


FIGURA 4.3. Distribución espacial y función de intensidad del conjunto de datos *Redwood* contenido en el paquete *Spatstat* [7]. Representa 62 secoyas en un cuadrado de 23 metros de lado. Los colores claros muestran agrupamientos de árboles.

evidente asumir de qué tipo de patrón se trata.

La hipótesis de AEC para este conjunto de datos usando pruebas basadas en las funciones  $F$ ,  $G$  y  $K$  fué evaluada en la sección 1.1.5 (Figuras 1.6D, 1.7D y 1.5D). En todos los casos se rechazó la hipótesis nula y se concluyó que se trata de un patrón puntual agregado.

Finalmente, en la Figura 4.6 se evalúa la hipótesis nula empleando la prueba propuesta basada en dimensión *Box Counting*. La curva calculada se ubica por debajo de las bandas de confianza, lo que es evidencia para rechazar ésta hipótesis. La prueba confirma los resultados ya conocidos para este conjunto de datos, concluyendo que se trata de un patrón agregado.

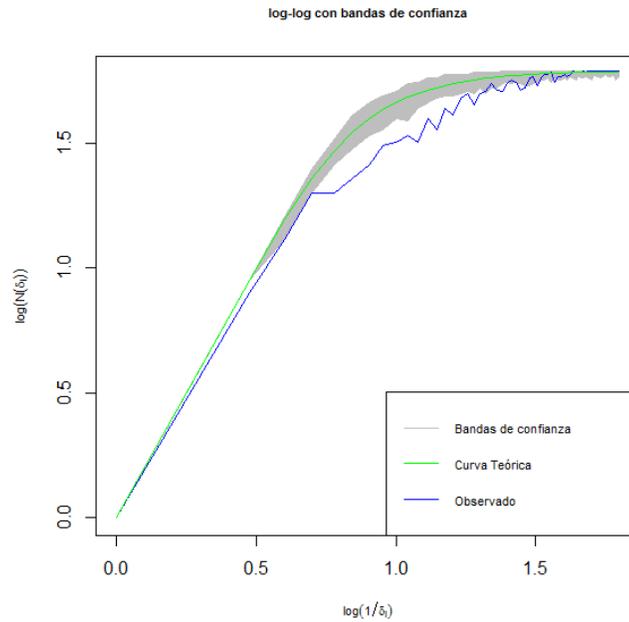


FIGURA 4.4. Contraste gráfico para evaluar la hipótesis de *AEA* a través de dimensión *Box Counting*, al conjunto de datos *Redwood* [9]. La curva observada se ubica por debajo de las bandas de confianza, razón por la cual se rechaza la hipótesis nula y se concluye que se trata de un patrón agregado o *cluster*.

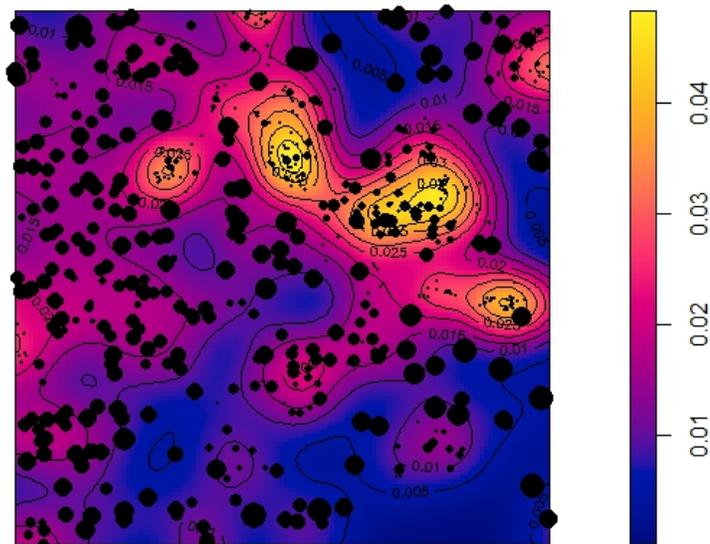


FIGURA 4.5. Distribución espacial y función de intensidad del conjunto de datos *Longleaf Pines* contenido en el paquete *Spatstat* [7]. Representa las coordenadas de ubicación 584 pinos en un área cuadrada de doscientos metros de lado (4 hectáreas) de bosque en el condado Thomas (Georgia, USA) [23].

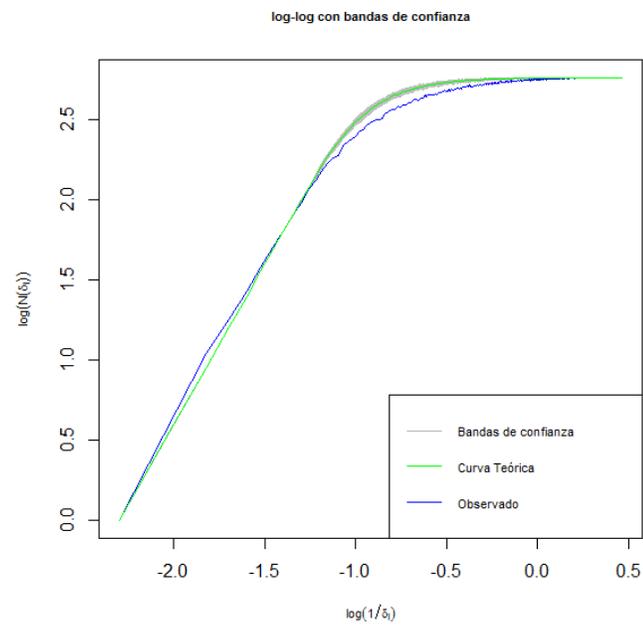


FIGURA 4.6. Contraste gráfico para evaluar la hipótesis de *AEA* a través de dimensión *Box Counting*, al conjunto de datos *Longleaf Pines* [9]. La curva observada se ubica por debajo de las bandas de confianza, razón por la cual se rechaza la hipótesis nula y se concluye que se trata de un patrón agregado o *cluster*.

---

---

## Conclusiones

---

---

- Se propone y desarrolla teóricamente una prueba para evaluar la hipótesis de AEC basada en el cálculo de la dimensión fractal, mediante la metodología *Box Counting*. Bajo condiciones de AEC se definen dos dimensiones fractales. La primera  $D_{B1}$ , definida desde  $(1/\delta_i) = 1$  hasta  $\tau_i = 1$ , depende del tamaño de muestra  $n$ . La segunda  $D_{B2}$  desde  $\tau_i = 1$  hasta cuando  $(1/\delta_i) \rightarrow \infty$ , tiende a cero cuando  $(1/\delta_i) \rightarrow \infty$ .

Para realizar la prueba gráfica se realizan simulaciones de Monte Carlo y se obtienen bandas de confianza bajo AEC. La prueba consiste en establecer si la curva obtenida con los datos se ubica en el interior de las bandas de confianza calculadas bajo aleatoriedad.

Los resultados teóricos y los análisis de datos reales y simulados permiten concluir que curvas *Box Counting* por encima de las bandas de confianza corresponden a patrones regulares, así mismo, las que están por debajo hacen referencia patrones agregados.

- La prueba propuesta mostró buen desempeño en los tres escenarios (aleatoriedad, inhibición y agregación), presentando incluso mejor comportamiento que las funciones  $F$  y  $K$ . El resultado también fué el esperado en las aplicaciones con datos reales. En ningún caso se presentaron inconsistencias respecto a lo obtenido con las pruebas tradicionales basadas en las funciones  $F$ ,  $G$  y  $K$ . Una de las ventajas de la prueba es que no es necesario considerar efecto de borde, ya que se basa únicamente en particiones de la ventana de observación.
- En casos de patrones puntuales espaciales con valores grandes de  $n$ , la función  $F$  presenta inconvenientes en la generación de las bandas de confianza. La función  $K$  muestra inconvenientes para descartar la hipótesis nula en el caso de patrones de inhibición y no calcula el efecto de borde, sino que realiza una estimación. En estos casos la prueba propuesta podría ser mejor alternativa.

---

---

## Trabajo futuro

---

---

- El presente trabajo deja abierta la exploración de la teoría de fractales en el estudio de patrones puntuales espaciales. Una tarea es estudiar el comportamiento de la prueba de *AEC* basada en dimensión *Box Counting* para casos particulares en que las pruebas gráficas tradicionales presenten inconvenientes, como en el caso de los patrones puntuales híbridos [8].
- Hacer la implementación de la prueba propuesta en una librería del software R, mejorando sus tiempos de respuesta.
- Profundizar en el significado de la dimensión fractal que resulta de la aplicación de la prueba propuesta así como las aplicaciones de la teoría de fractales al modelar patrones puntuales espaciales.
- Realizar el desarrollo teórico de la prueba propuesta para el caso de patrones puntuales en  $\mathbb{R}^3$ .
- Extender la teoría considerada al caso de patrones puntuales espacio-temporales.
- Para el caso de patrones marcados espacio-temporales plantear un modelo dinámico entre la dimensión *Box Counting* y las covariables.

---



---

## Función BoxCountingDim

---



---

La función **BoxCountingDim** escrita en *R* [54], realiza el contraste gráfico que permite evaluar la hipótesis de AEC, mediante simulaciones de Monte Carlo. Calcula y grafica las bandas de confianza y curva teórica bajo la hipótesis nula, dibuja la curva *log-log* observada, lo que permite comparar su comportamiento y rechazar  $\mathcal{H}_0$  cuando ésta se ubica fuera de las bandas de confianza. A continuación se muestra el código fuente, que si bien es susceptible de mejoras, presenta buen desempeño.

```

1 #####Función que realiza el test de aleatoriedad propuesto#####
2 BoxCountingDim=function(X1,Y1){
3 k=max(X1,Y1)
4 BoxCount1 = function(X,Y)
5 {
6 n=length(X)
7 Nd=1
8 del=k
9 for (i in 2:n){
10 i=1+i
11 k0=k/i
12 (XY=unique(data.frame(ceiling(X/k0),ceiling(Y/k0))))
13 (Na=length(XY$ceiling.X.k0.))
14 Nd=c(Nd,Na)
15 del=c(del,k0)
16 }
17 m=matrix(c(log10(1/del),log10(Nd)),ncol=2)
18 return(m)
19 }
20 b1=BoxCount1(X1,Y1)
21 z=100
22 n=length(b1[,1])
23 nn=subset(b1,b1[,2]==log10(n))
24 nmin=min(nn)
25 del=b1[,1]
26 todos=matrix(nrow=n,ncol = z)
27 for (i in 1:z){
28 x=runif(n,0,k)
29 y=runif(n,0,k)
30 as.vector(b<-BoxCount1(x,y)[,2])
31 todos[,i]=b
32 }

```

```

33 bandas=matrix(nrow=n,ncol=2)
34 for (i in 1:n){
35 banda=range(todos[i,])
36 bandas[i,]=banda
37 }
38 p1=data.frame(del,bandas[,1])
39 p2=data.frame(del,bandas[,2])
40 names(p2)=c("px","py")
41 names(p1)=c("px","py")
42 p1 <- p1[order(-del),]
43 pp=rbind(p2,p1)
44 plot(del,bandas[,1],type="l",col="gray",main="log-log con bandas de
confianza",cex.main=0.7,cex.lab=0.7,xlab=expression(log(1/delta[i])),ylab
=expression(log(N(delta[i])))
45 legend("bottomright",c("Bandas de confianza","Curva Teórica","Observado"),
col=c("gray","Green","Blue"),lty = 1, border="gray",cex = 0.7)
46 polygon(pp$px,pp$py,col="gray",border = "gray")
47 points(del,b1[,2],type="l",col="Blue")
48 delta=k/seq(1,n)
49 logN=2*log10(k)-2*log10(delta)+log10(1-exp(-n*delta^2/k^2))
50 points(log10(1/delta),logN,type = "l",col="Green")
51 teorica<-function(n,k){
52 delta=k/seq(1,n)
53 logN=2*log10(k)-2*log10(delta)+log10(1-exp(-n*delta^2/k^2))
54 crucex=-1*log10(k)-1/2*(log10(1-exp(-1*n)))
55 infley=log10(n)-0.1992001
56 inflex=1/2*log10(n)-log10(k)
57 DB1=2*(log10(n)-0.1992001)/(log10(n)-log10(1-exp(-1*n)))
58 res=c(crucex, inflex, infley, DB1)
59 nombres=c("Cruce x", "inflex.x", "inflex.y", "DB1")
60 f=data.frame(nombres, res)
61 return(f)
62 print(f)
63 }
64 teorica(n,k)
65 }

```

---

---

## Bibliografía

---

---

- [1] Paul S Addison, *Fractals and chaos: an illustrated course*, CRC Press, 1997.
- [2] Peyman Afzal, Ahmad Khakzad, Parviz Moarefvand, N Rashidnejad Omran, Bijan Esfandiari, and Younes Fadakar Alghalandis, *Geochemical anomaly separation by multifractal modeling in kahang (gor gor) porphyry system, central iran*, Journal of Geochemical Exploration **104** (2010), no. 1, 34–46.
- [3] Frederik P Agterberg, *Fractals, multifractals, and change of support*, Geostatistics for the next century, Springer, 1994, pp. 223–234.
- [4] ———, *Multifractals and geostatistics*, Journal of Geochemical Exploration **122** (2012), 113–122.
- [5] Adrian Baddeley, *Analysis of spatial point processes*, Spatstat package. CRAN (2003).
- [6] Adrian Baddeley et al., *Analysing spatial point patterns with r*, Citeseer, 2015.
- [7] Adrian Baddeley, Rolf Turner, et al., *Spatstat: An r package for analyzing spatial point patterns*, 2004.
- [8] Adrian Baddeley, Rolf Turner, Jorge Mateu, and Andrew Bevan, *Hybrids of gibbs point process models and their implementation*, Journal of Statistical Software **55** (2013), no. 11, 1–43.
- [9] Adrian Baddeley, Rolf Turner, Ege Rubak, Kasper Klitgaard Berthelsen, Ute Hahn, Abdollah Jalilian, Marie-Colette van Lieshout, Tuomas Rajala, Dominic Schuhmacher, Rasmus Waagepetersen, et al., *Package spatstat*, (2015).
- [10] Michael F Barnsley, *Fractals everywhere*, Academic press, 2014.
- [11] Michael Batty and Paul A Longley, *Fractal cities: a geometry of form and function*, Academic Press, 1994.
- [12] FJ Caniego, R Espejo, MA Martín, and F San José, *Multifractal scaling of soil spatial variability*, Ecological Modelling **182** (2005), no. 3, 291–303.
- [13] Emmanuel John Muico Carranza, *Geochemical anomaly and mineral prospectivity mapping in gis*, vol. 11, Elsevier, 2008.
- [14] Qiuming Cheng, *Multifractal modelling and spatial analysis with gis: Gold potential estimation in the mitchell-sulphurets area, northwestern british columbia.*, University of Ottawa (Canada)., 1995.

- 
- [15] ———, *Discrete multifractals*, *Mathematical Geology* **29** (1997), no. 2, 245–266.
- [16] ———, *Multifractal modeling and lacunarity analysis*, *Mathematical Geology* **29** (1997), no. 7, 919–932.
- [17] ———, *Multifractality and spatial statistics*, *Computers & Geosciences* **25** (1999), no. 9, 949–961.
- [18] Qiuming Cheng, FP Agterberg, and SB Ballantyne, *The separation of geochemical anomalies from background by fractal methods*, *Journal of Geochemical Exploration* **51** (1994), no. 2, 109–130.
- [19] Qiuming Cheng and Frederik P Agterberg, *Multifractal modeling and spatial statistics*, *Mathematical Geology* **28** (1996), no. 1, 1–16.
- [20] Qiuming Cheng, H Russell, D Sharpe, Frank Kenny, and Ping Qin, *Gis-based statistical and fractal/multifractal analysis of surface stream patterns in the oak ridges moraine*, *Computers & geosciences* **27** (2001), no. 5, 513–526.
- [21] Quiming Cheng and Frederik P Agterberg, *Multifractal modeling and spatial point processes*, *Mathematical Geology* **27** (1995), no. 7, 831–845.
- [22] Ashvin Chhabra and Roderick V Jensen, *Direct determination of the  $f(\alpha)$  singularity spectrum*, *Physical Review Letters* **62** (1989), no. 12, 1327.
- [23] Noel Cressie, *Statistics for spatial data*, John Wiley & Sons, 1991.
- [24] Andre Dauphine, *Fractal geography*, Wiley, 2012.
- [25] Peter J Diggle et al., *Statistical analysis of spatial point patterns*, Academic press, 1983.
- [26] Roussos Dimitrakopoulos, *Geostatistics for the next century: An international forum in honour of michel david's contribution to geostatistics, montreal, 1993*, vol. 6, Springer Science & Business Media, 2012.
- [27] Gerald Edgar, {*Classics on Fractals*}, (1993).
- [28] ———, *Measure, topology, and fractal geometry*, Springer Science & Business Media, 2007.
- [29] ———, *Integral, probability, and fractal measures*, Springer Science & Business Media, 2013.
- [30] Kenneth Falconer, *Fractal geometry: mathematical foundations and applications*, John Wiley & Sons, 2004.
- [31] Pierre Frankhauser, *The fractal approach. a new tool for the spatial analysis of urban agglomerations*, *Population* **10** (1998), no. 1, 205–240.
- [32] Carlo Gaetan, Xavier Guyon, and Kevin Bleakley, *Spatial statistics and modeling*, vol. 81, Springer, 2010.
- [33] Juan Pablo Gallego Valencia et al., *Aplicación de la teoría de caos para el análisis y pronóstico de series de tiempo financieras en colombia*, Ph.D. thesis, Universidad Nacional de Colombia, 2010.

- 
- [34] Oscar Javier García Cabrejo, Giovanni Moreno Sánchez, et al., *Herramientas de análisis estadístico y espacial aplicadas en la separación línea base-anomalía geoquímica. implicaciones en exploración de recursos naturales. parte 1: revisión metodológica*, Geología Colombiana-An International Journal on Geosciences; Vol. 31 (2006); 27-38 Geología Colombiana; Vol. 31 (2006); 27-38 2357-3767 0072-0992.
- [35] Alan E Gelfand, Peter Diggle, Peter Guttorp, andMontserrat Fuentes, *Handbook of spatial statistics*, CRC press, 2010.
- [36] Mario A Goncalves, Antonio Mateus, and Vitor Oliveira, *Geochemical anomaly separation by multifractal modelling*, Journal of Geochemical Exploration **72** (2001), no. 2, 91–114.
- [37] Mario A Goncalves, M Vairinho, and Vitor Oliveira, *Study of geochemical anomalies in mombeja area using a multifractal methodology and geostatistics*, IV IAMG **98** (1998), 590–595.
- [38] Derek Gregory, *Geographical imaginations*, Blackwell Oxford, 1994.
- [39] Shougeng Hu, Qiuming Cheng, Le Wang, and Shuyun Xie, *Multifractal characterization of urban residential land price in space and time*, Applied Geography **34** (2012), 161–170.
- [40] Shougeng Hu, Qiuming Cheng, Le Wang, and Deyi Xu, *Modeling land price distribution using multifractal idw interpolation and fractal filtering method*, Landscape and Urban Planning **110** (2013), 25–35.
- [41] Janine Illian, Antti Penttinen, Helga Stoyan, and Dietrich Stoyan, *Statistical analysis and modelling of spatial point patterns*, vol. 70, John Wiley & Sons, 2008.
- [42] Anna Batlle Joher and Sandra Grebol Montoro, *Fractales en los mercados financieros*, (2008).
- [43] S Khan, AR Ganguly, and S Saigal, *Detection and predictive modeling of chaos in finite hydrological time series*, Nonlinear Processes in Geophysics **12** (2005), no. 1, 41–53.
- [44] Andrzej Lasota and Michael C Mackey, *Chaos, fractals, and noise: stochastic aspects of dynamics*, vol. 97, Springer Science & Business Media, 2013.
- [45] Daniel J Levitin, Parag Chordia, and Vinod Menon, *Musical rhythm spectra from bach to joplin obey a 1/f power law*, Proceedings of the National Academy of Sciences **109** (2012), no. 10, 3716–3720.
- [46] Renaud Lopes and Nacim Betrouni, *Fractal and multifractal analysis: a review*, Medical image analysis **13** (2009), no. 4, 634–649.
- [47] Benoit Mandelbrot, *How long is the coast of britain? statistical self-similarity and fractional dimension*, Science (1967).
- [48] ———, *The fractal geometry of nature*, Freeman, 1982.
- [49] Benoit Mandelbrot and Ian Stewart, *Fractals and scaling in finance*, Nature **391** (1998), no. 6669, 758–758.

- 
- [50] Jesper Møller and Rasmus P Waagepetersen, *Modern statistics for spatial point processes*, Scandinavian Journal of Statistics **34** (2007), no. 4, 643–684.
- [51] JS Murgua, E Campos-Cantón, E Jiménez, and A Corpus-Cordero, *Análisis de series de tiempo caóticas*, (2007).
- [52] Ahad Nazarpour, Nematolah Rashidnejad Omran, Ghodratolal Rostami Paydar, Behnam Sadeghi, Fatemeh Matroud, and Ali Mehrabi Nejad, *Application of classical statistics, logratio transformation and multifractal approaches to delineate geochemical anomalies in the zarshuran gold district, nw iran*, Chemie der Erde-Geochemistry **75** (2015), no. 1, 117–132.
- [53] Assela Pathirana and Srikantha Herath, *Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity*, Hydrology and Earth System Sciences Discussions **6** (2002), no. 4, 695–708.
- [54] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [55] Brian D Ripley, *Modelling spatial patterns*, Journal of the Royal Statistical Society. Series B (Methodological) (1977), 172–212.
- [56] ———, *Spatial statistics*, vol. 575, John Wiley & Sons, 2005.
- [57] Gustavo Rubiano, *Fractales para profanos*, 2002.
- [58] ———, *Iteración y fractales con mathematica*, Universidad Nacional de Colombia, 2009.
- [59] Oliver Schabenberger and Carol A Gotway, *Statistical methods for spatial data analysis*, CRC press, 2004.
- [60] Dietrich Stoyan and Helga Stoyan, *Fractals, random shapes and point fields: methods of geometrical statistics*, 1994.
- [61] Till Tantau, *The tikz and pgf packages, manual for version 3.0. 0*, 2013.
- [62] Donald L Turcotte, *Fractals and chaos in geology and geophysics*, Cambridge university press, 1997.
- [63] Hans Wackernagel, *Multivariate geostatistics*, Springer Science & Business Media, 2003.
- [64] Lance A Waller and Carol A Gotway, *Applied spatial statistics for public health data*, vol. 368, John Wiley & Sons, 2004.
- [65] Gongwen Wang, Emmanuel John M Carranza, Renguang Zuo, Yinglong Hao, Yangsong Du, Zhenshan Pang, Yue Sun, and Jianan Qu, *Mapping of district-scale potential targets using fractal models*, Journal of Geochemical Exploration **122** (2012), 34–46.
- [66] Haicheng Wang, Qiuming Cheng, and Renguang Zuo, *Quantifying the spatial characteristics of geochemical patterns via gis-based geographically weighted statistics*, Journal of Geochemical Exploration **157** (2015), 110–119.
- [67] Thorsten Wiegand and Kirk A Moloney, *Handbook of spatial point-pattern analysis in ecology*, CRC Press, 2013.

- 
- [68] Renguang Zuo, Emmanuel John M Carranza, and Qiuming Cheng, *Fractal/multifractal modelling of geochemical exploration data*, Journal of Geochemical Exploration **122** (2012), 1–3.