

Análisis de presencias con procesos de puntos

Generalidades

Gerardo Martín

2022-06-29

Introducción

¿Qué es un patrón de puntos?

- Base de datos de cosas o eventos en espacio

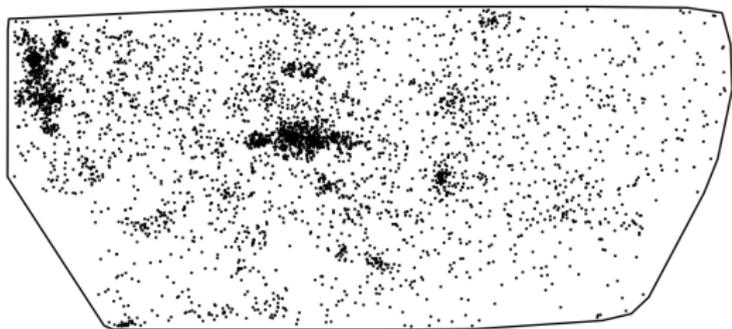
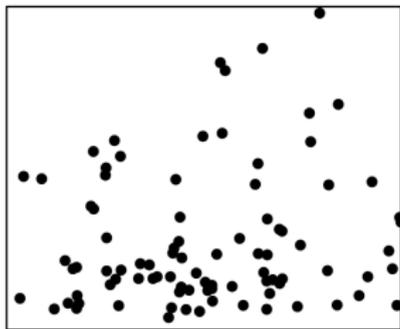


Figure 1: Patrones de puntos de densidad variable. A la izquierda células de mucosa gástrica en corte histológico. A la derecha, cúmulos de galaxias (Baddeley et al. 2016).

- Densidad \rightarrow conteos/unidad espacial

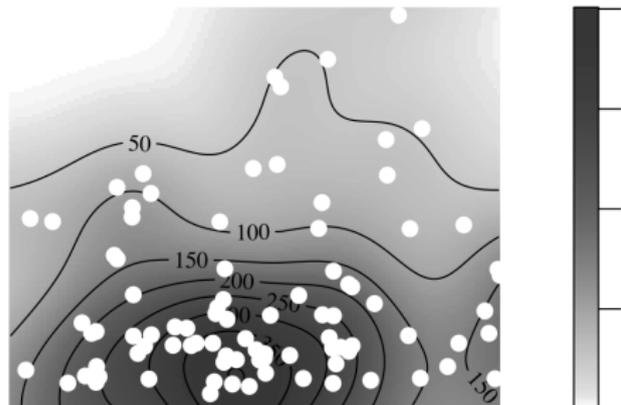
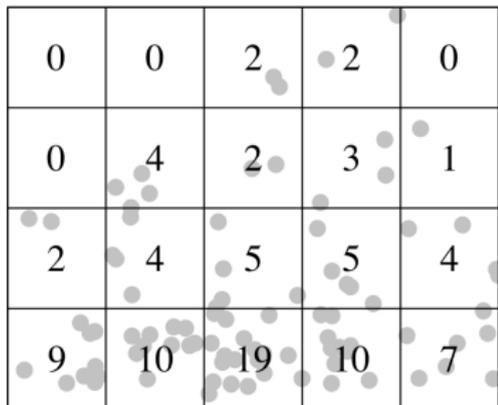


Figure 2: Ejemplo de medición de intensidad de puntos

Puntos pueden representar tipos de objetos

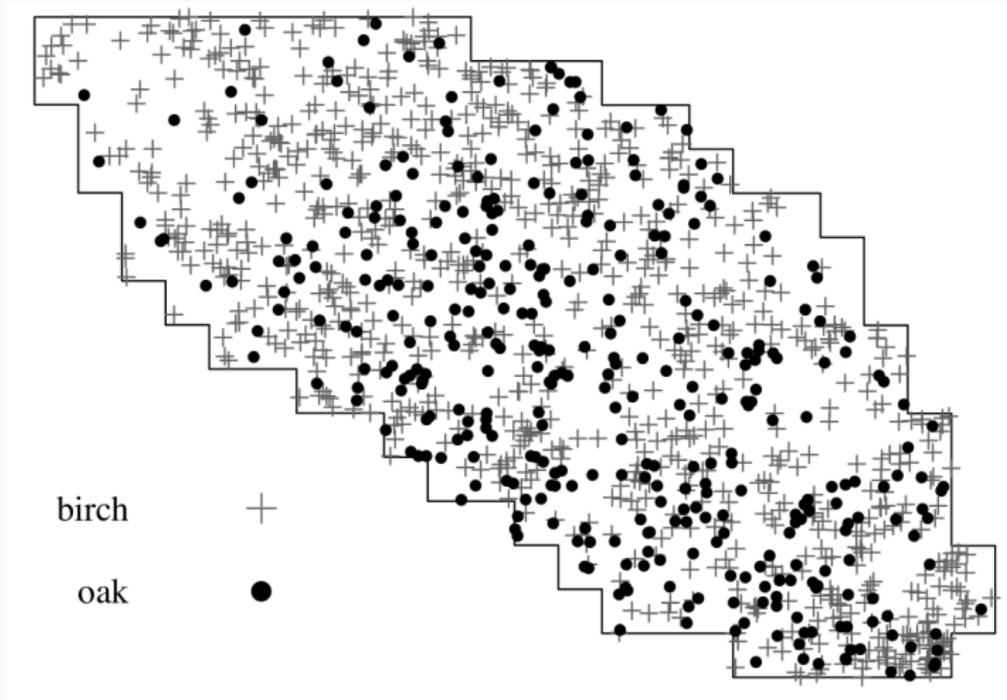


Figure 3: Ubicaciones de dos especie de árbol, abeto y roble, en la misma parcela.

Puntos pueden representar mediciones

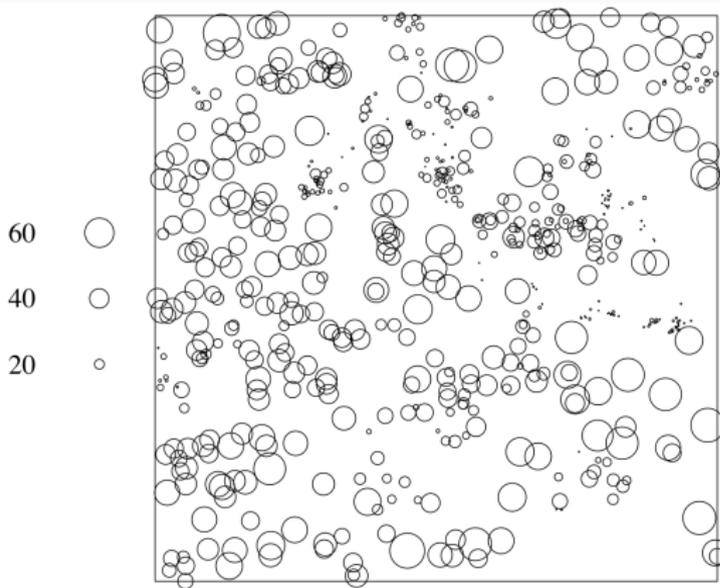


Figure 4: Ubicaciones de árboles con mediciones de diámetro.

Puntos pueden estar definidos en 1-4 dimensiones



Figure 5: Ejemplos de procesos de puntos en 1 y 3 dimensiones

Los procesos de puntos pueden estar definidos en relación a covariables.

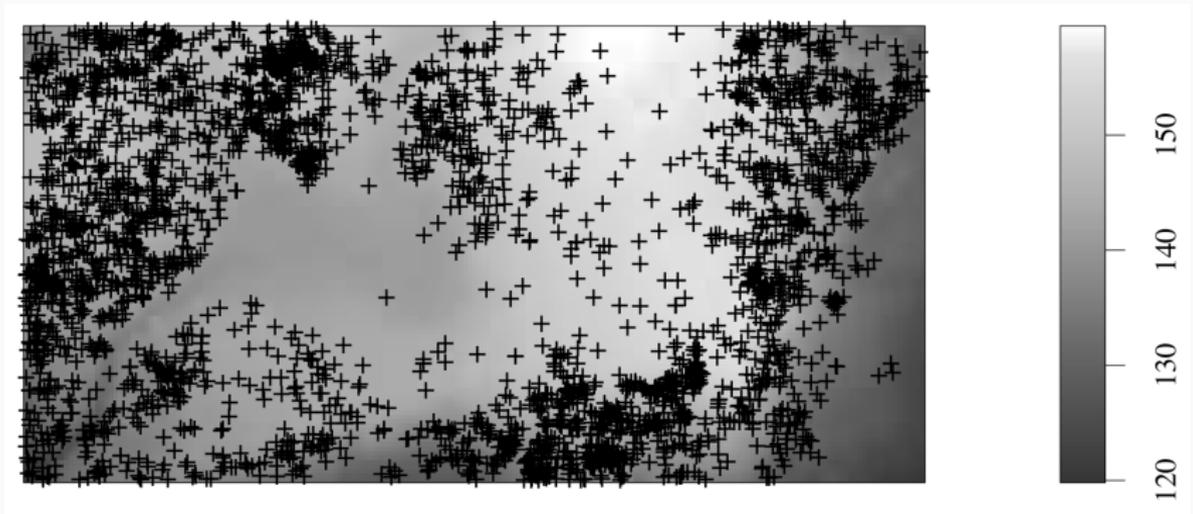


Figure 6: Datos de Beilschmiedia pendula sobre un modelo digital de elevación.

- Completamente aleatorios (homogéneos)
- *Heterogéneos*
 - Procesos de puntos generalmente producidos por covariables
- Agregados

El modelado de procesos de puntos

- Estimar variación de densidad
- Densidad = No. puntos / unidad de área

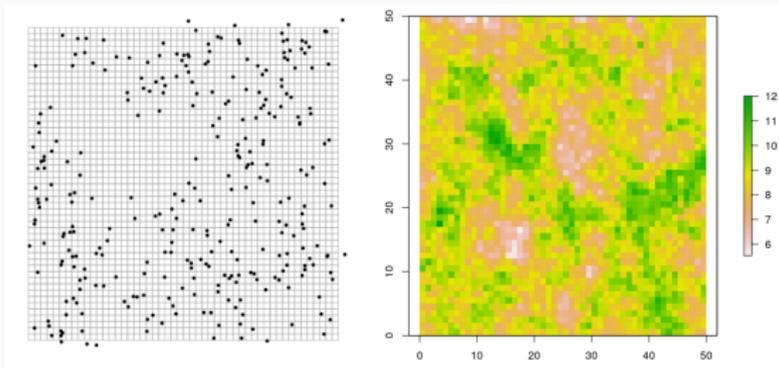


Figure 7: Se analiza un patrón para predecir variación continua.

- Estadísticamente, densidad puede tratarse con la distribución Poisson

Procesos de puntos en ecología

- Datos más comunes → sólo presencia
- Colecciones de patrones de puntos



Análisis de procesos de puntos

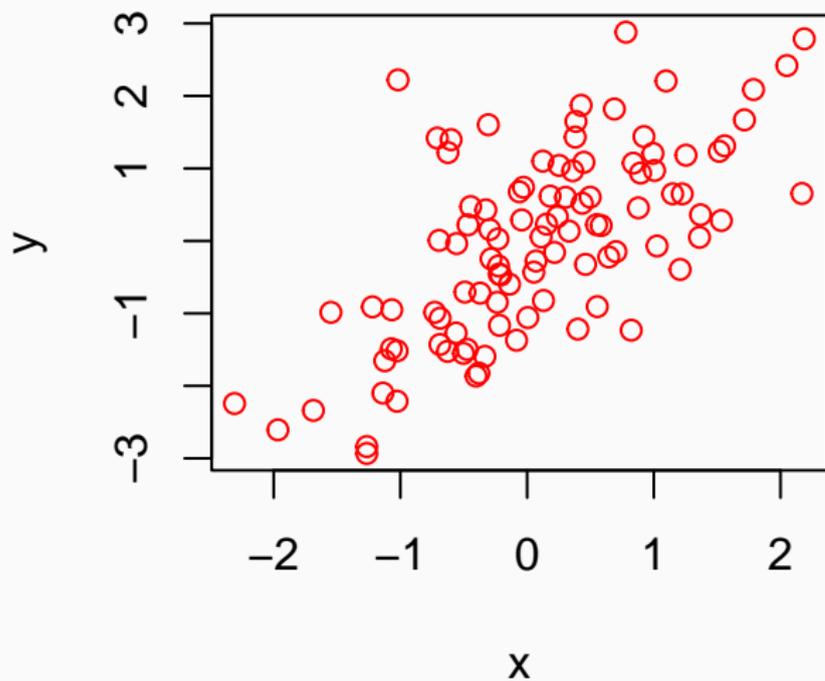
- Medir relación entre x y y
 - ¿Cómo afecta x al promedio de y ?
- x produce a y
 - x variable independiente
 - y variable dependiente de x

En regresión lineal simple, tendríamos dos variables continuas:

x	y
-0.5604756	-1.2708822
-0.2301775	0.0267062
1.5587083	1.3120164
0.0705084	-0.2770342
0.1292877	-0.8223308
1.7150650	1.6700373

Ejemplo - Gráfica de dispersión

A las cuales realizamos un análisis exploratorio:



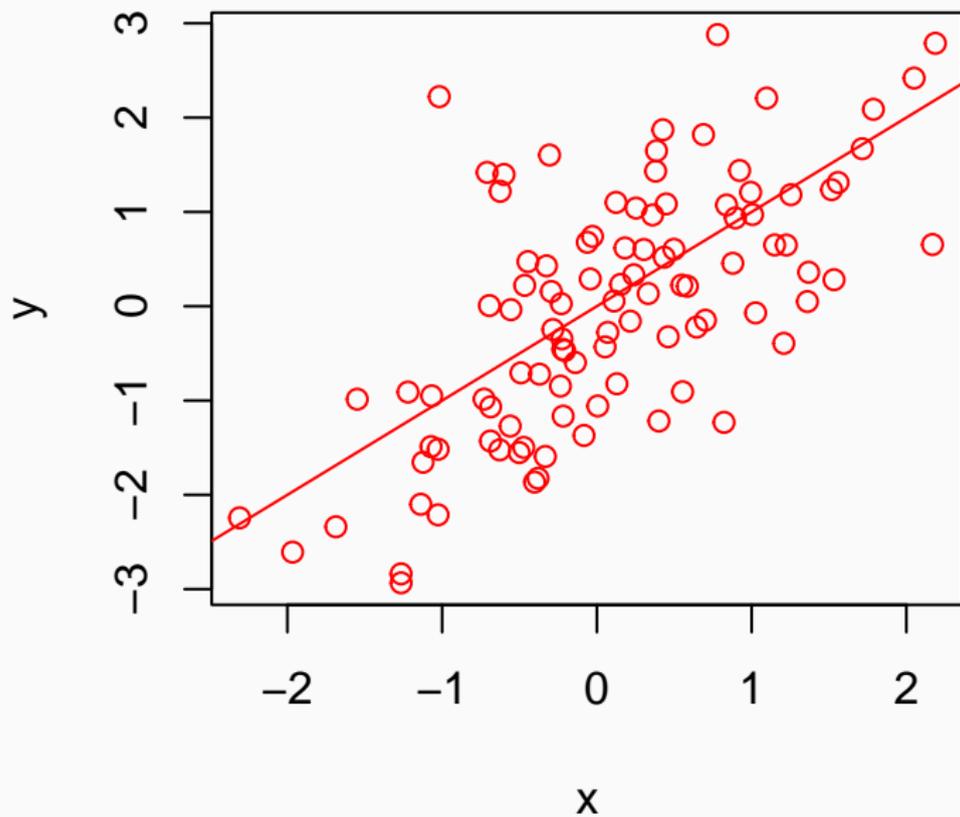
- Siendo un poco más generales que arriba y asumiendo que y es un patrón de puntos, podemos tener modelos de regresión así:

$$y(x) = \alpha + \beta_1 x_1 + \dots + \beta_n x_n + \varepsilon$$

- x son las variables ambientales

- y es la intensidad por unidad de área
- α es el intercepto, y β_i son los efectos de x sobre y
 - Cuánto va a cambiar y si “aumento” el valor de x en una unidad
- ε es el error, varianza de y que x no explica

Ejemplo - La línea de regresión



En el caso simple anterior:

- $y = \alpha + \beta \times x$

- $\alpha = 0$

- $\beta = 1$

Regresión consiste en estimar todos los coeficientes para las variables x .

En un modelo de regresión lineal, tenemos:

- W son mediciones de peso en Kg, y su modelo es:

$$W = \alpha + \beta Edad + \beta_{Sexo} + \varepsilon$$

- Donde las unidades que produce el modelo son peso en Kg

- Mediciones: coordenadas de presencia
- ¿Qué unidades produce el modelo estadístico para los puntos de presencia?

- Regresión lineal simple
 - $-\infty > y < \infty, y \in \mathbb{R}$
 - $y \approx \mathcal{N}$ (distribución Normal)
- Procesos de puntos
 - $y > 0, y \in \mathbb{Z}$
 - $y \approx \mathcal{P}$ (distribución Poisson)

Para que $y > 0$

- Regresión lineal

- $y(x) = \alpha + \beta_1 x_1 + \dots$

- Regresión log-lineal

- $\log y(x) = \alpha + \beta_1 x_1 + \dots$

- es decir:

$$y(x) = \exp(\alpha + \beta_1 x_1 + \dots)$$

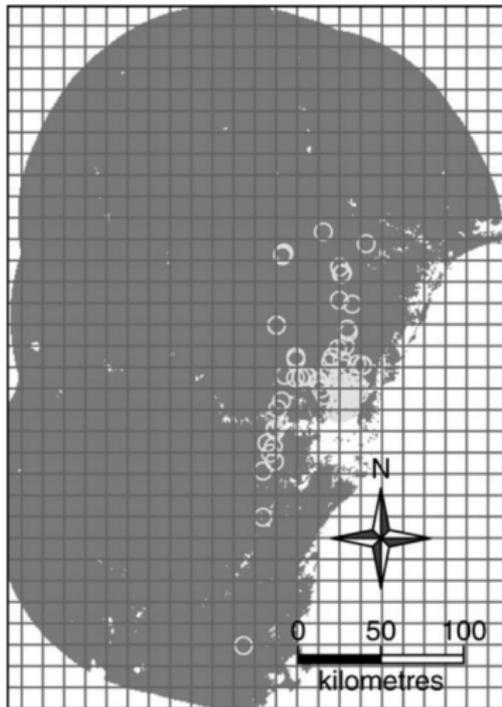
Equivalence of MAXENT and Poisson Point Process Models for Species Distribution Modeling in Ecology

Ian W. Renner* and David I. Warton

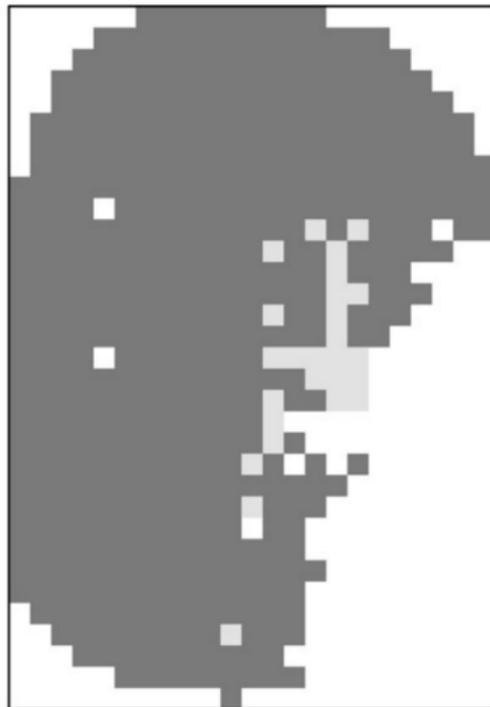
School of Mathematics and Statistics and Evolution & Ecology Research Centre,
The University of New South Wales, NSW 2052, Australia.

**email*: Ian.Renner@unsw.edu.au

Point pr



MAXENT



$$\ln \pi(g_i) = \ln \mu(g_i) = x(g_i)\beta$$

Son funciones log-lineales equivalentes de un conjunto de predictores

- Maxent: maximiza la entropía ($\pi(g_i)$)
- Procesos de puntos en general: maximizan verosimilitud Poisson (conteo de puntos por unidad espacial)

- Equivalentes a MaxEnt
 - Sin Regularización
 - *Features* lineal y cuadrática

Methods in Ecology and Evolution



Methods in Ecology and Evolution 2015, **6**, 366–379

doi: 10.1111/2041-210X.12352

SPECIAL FEATURE – REVIEW

NEW OPPORTUNITIES AT THE INTERFACE BETWEEN ECOLOGY AND STATISTICS

Point process models for presence-only analysis

Ian W. Renner^{1*}, Jane Elith², Adrian Baddeley³, William Fithian⁴, Trevor Hastie⁴,
Steven J. Phillips⁵, Gordana Popovic⁶ and David I. Warton⁶

¹School of Mathematical and Physical Sciences, The University of Newcastle, University Drive, Callaghan, NSW 2308, Australia; ²School of BioSciences, The University of Melbourne, Parkville, Vic. 3010, Australia; ³Department of Mathematics & Statistics, Curtin University, GPO Box U1987, Perth, WA 6845, Australia; ⁴Department of Statistics, Stanford University, 390 Serra Mall, Stanford, CA 94303, USA; ⁵2201 4th Street, Boulder, CO 80304, USA; and ⁶School of Mathematics and Statistics and Evolution & Ecology Research Centre, The University of New South Wales, Sydney, NSW 2052, Australia

- Regresión log-lineal
 - Logística
 - Maxent
- Elipsoides (Martín et al. 2022)
 - Centroides existen en espacio
 - Sin colinealidad

Condiciones para equivalencia entre MPPs y envolturas

Discrepancies between point process models and environmental envelopes identify the niche centroid – geography configuration

Gerardo Martín^{1,*}, Carlos Yáñez-Arenas², Xavier Chiappa-Carrara^{1,2}

¹ Departamento de Sistemas y Procesos Naturales, Escuela Nacional de Estudios Superiores unidad Mérida, Universidad Nacional Autónoma de México, Ucu, Yucatán 97357, México

² Laboratorio de Ecología Geográfica, Unidad de Conservación de la Biodiversidad, UMDI-Sisal, Facultad de Ciencias, Universidad Nacional Autónoma de México, Sierra Papacal, Yucatán 97302, México

Statistical density

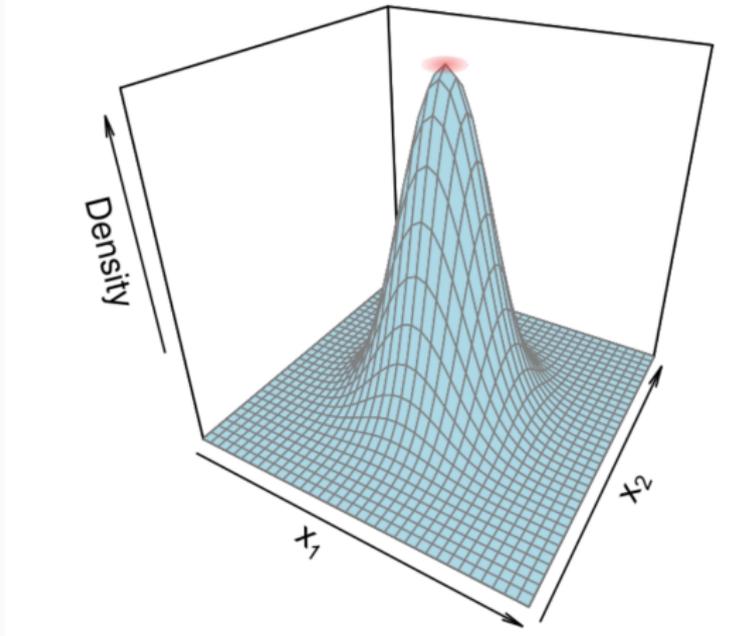


Figure 9: Frecuencia de condiciones ambientales en sitios de presencia para dos variables no correlacionadas.

Distance to centroid

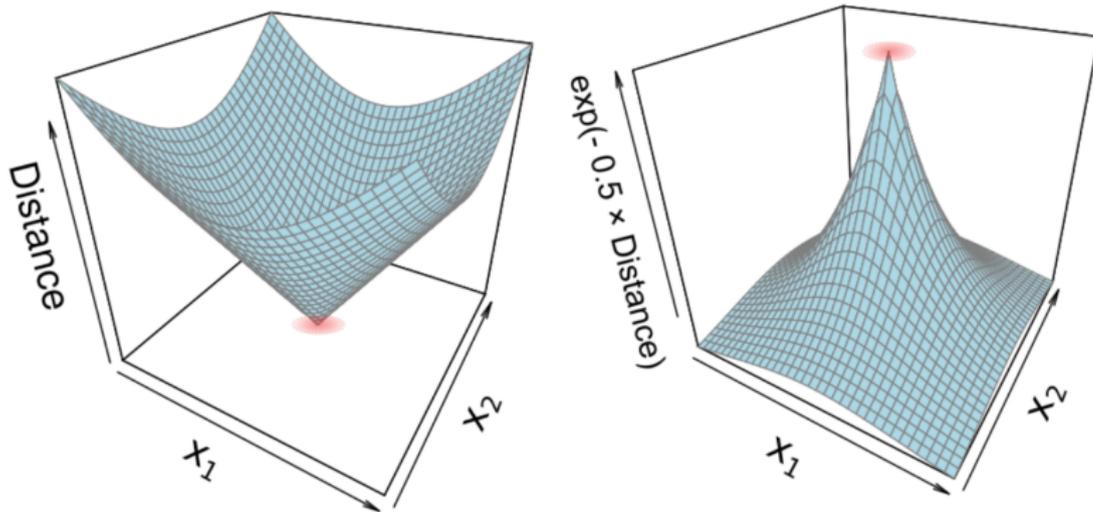


Figure 10: Distancia al centroide en dos dimensiones ambientales no correlacionadas

Point intensity

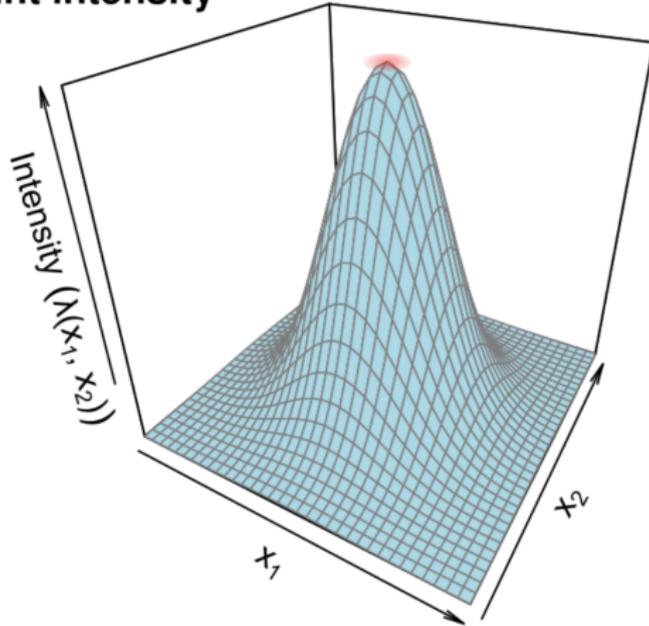


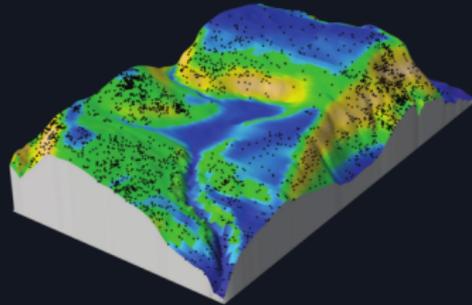
Figure 11: Intensidad de puntos en dos dimensiones ambientales no correlacionadas para el mismo proceso de puntos.

- Procesos de puntos son la herramienta *ad-hoc* para puntos
- Maxent es una herramienta estadística de regresión, per se, no genera distribuciones
 - Favorabilidad, idoneidad son interpretaciones de los valores producidos
 - Los valores producidos representan el número esperado de puntos de presencia (exceptuando la salida logística)
- Los PPMs son más transparentes, y por medio de **spatstat**:
 - Herramientas exploratorias, estimar y comparar efectos, diagnosticar modelos
 - Modelar estructura espacial

- Formateo
- Difícil automatizar
- Más programación
- Selección de modelo laboriosa
- Optimización puede ser difícil
- Poco práctico si analizaremos muchas especies

Spatial Point Patterns

Methodology and Applications with R



Adrian Baddeley • Ege Rubak • Rolf Turner