Análisis espacial avanzado

I de Moran

I de Moran Global

Resume en un coeficiente el grado de asociación entre un nivel dado de la variable considerada en un área geográfica respecto del promedio ponderado de la misma variable en áreas vecinas

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \overline{y})(y_j - \overline{y})}{\left(\sum_{i=1}^{n}(y_i - \overline{y})^2\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right)}$$

Moran, P. A. P. (1950) Notes on continuous stochastic phenomena. Biometrika, 37, 17–23.
Cliff , A. D. and Ord, J. K. (1973) Spatial Autocorrelation. London: Pion.

Autocorrelación espacial

“Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes.”

Ley de Tobler

TOBLER, W. R (1970). *A computer model simulation of urban growth in the Detroit region*. Economic Geography 46(2).

Esto nos permite observar el fenómeno de la dependencia espacial o autocorrelación espacial

Haining, R. (2003) Spatial Data Analysis: Theory and Practice. Cambridge: Cambridge University Press.

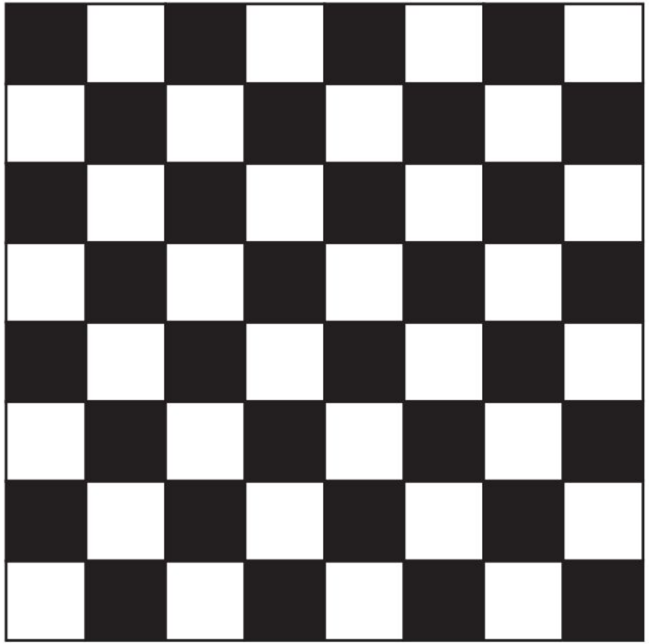
Autocorrelación espacial

El grado en el cual una variable está espacialmente correlacionada con ella misma. Puede ser que todos los valores sean similares (autocorrelación positiva) o totalmente independientes (todos los valores son disímiles entre sí)

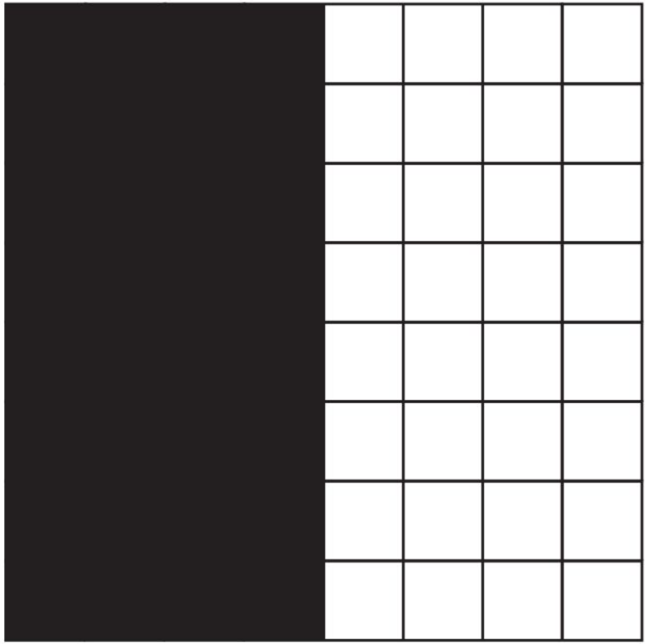
Haining, R. (2003) Spatial Data Analysis: Theory and Practice. Cambridge: Cambridge University Press.

También es utilizado (en su variante Local) para detectar clusters espaciales. En sus 4 variantes. Más adelante veremos eso...

Resultados del índice de Moran

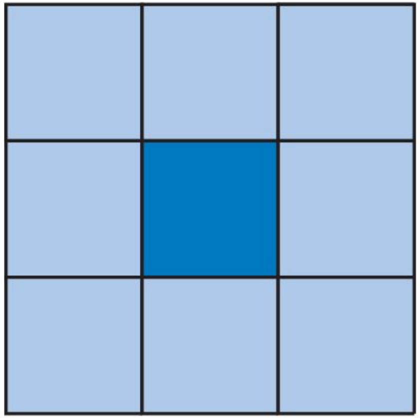


Negative spatial autocorrelation
$I = -1.000$

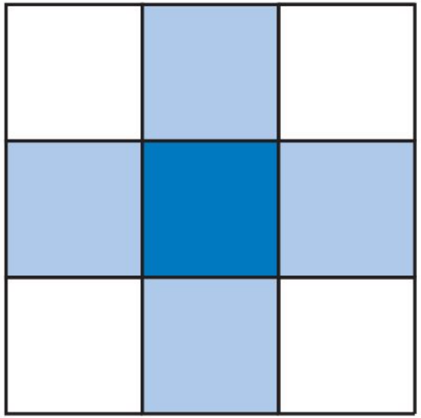


Positive spatial autocorrelation
$I = 0.865$

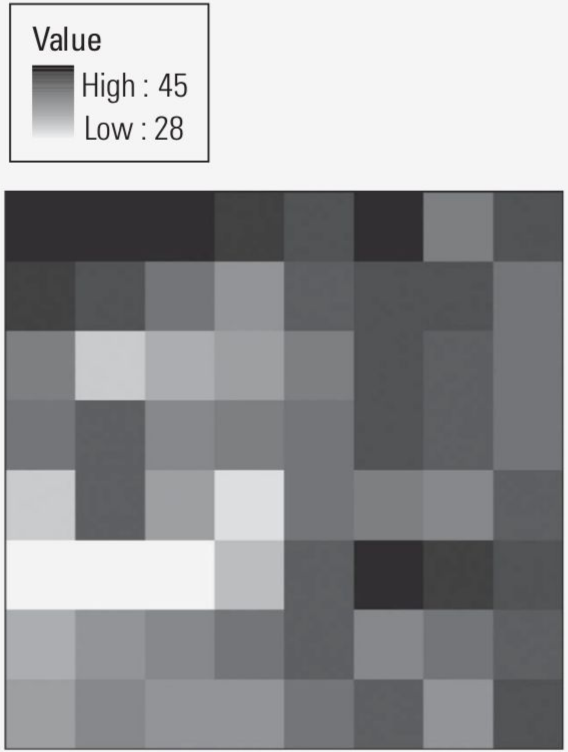
Adyacencias tipo Torre o tipo Reina



Queen's case contiguity



Rook's case contiguity



A

45	44	44	43	42	44	38	42
43	42	34	35	41	42	42	39
38	32	34	35	38	42	41	39
39	41	37	38	39	42	41	39
32	41	35	31	39	38	37	41
29	29	28	33	41	44	43	42
34	36	37	39	41	37	39	41
35	37	36	36	39	41	36	42

B Moran's *I*: Rook's case: 0.505
Queen's case: 0.400

Ecuación de I de Moran Global

7	8	11
11	9	10
11	12	9

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \overline{y})(y_j - \overline{y})}{\left(\sum_{i=1}^{n}(y_i - \overline{y})^2\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right)}$$

$$I = \frac{9 \times 3.975}{21.556 \times 40} = \frac{35.778}{862.222} = 0.041$$

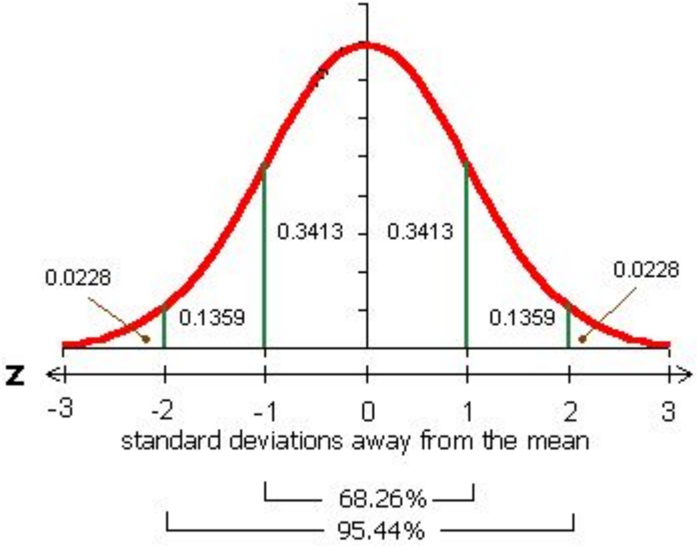
Ejemplo

Imaginemos una clase. Estas son las notas de un examen.

6	6	8	9	9
9	6	8	10	8
5	4	9	8	9
3	6	6	7	6
10	3	4	5	4

Normalizemos

-0.34	-0.34	0.60	1.07	1.07
1.07	-0.34	0.60	1.54	0.60
-0.81	-1.28	1.07	0.60	1.07
-1.75	-0.34	-0.34	0.13	-0.34
1.54	-1.75	-1.28	-0.81	-1.28



$$z = \frac{x - \mu}{\sigma}$$

Moran Local LISA

Si bien Moran Global nace para analizar autocorrelación espacial, su variante local puede utilizarse para detectar cluster espaciales. Estos se denominan LISA o Local Indicators of Spatial Association.

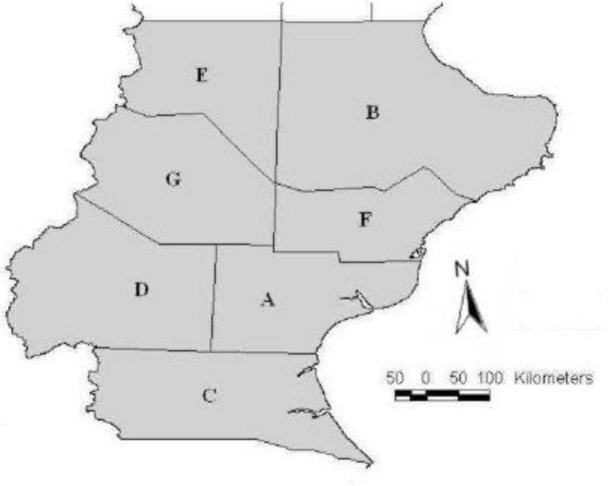
Calculado para celda del medio (42) con filas estandarizadas y pesos torre (rook)

45	44	44
43	42	39
38	32	34

$$I_i = \left[\frac{z_i}{s^2}\right]\sum_{j=1}^{n} w_{ij} z_j,\ j \neq i$$

y_i	z_i	w_{ij}	$w_{ij}z_j$
45	4.889	0.000	0.000
43	2.889	0.250	0.722
38	–2.111	0.000	0.000
44	3.889	0.250	0.972
42	1.889	0.000	0.000
32	–8.111	0.250	–2.028
44	3.889	0.000	0.000
39	–1.111	0.250	–0.278
34	–6.111	0.000	0.000
Sum		1.000	–0.611

$$I_i = \frac{(P_{xi} - \overline{\mathrm{P}}_x)\sum_j c_{ij}(P_{xj} - \overline{\mathrm{P}}_x)}{\sum_i \frac{(P_{xi} - \overline{\mathrm{P}}_x)^2}{N}}$$



Porcentaje de viviendas con baño exclusivo por departamento. Pcia. De Santa Cruz

Departamento	% Baño Exclusivo
A) Corpen Aike	98,92
B) Deseado	95,36
C) Guer Aike	96,35
D) Lago Argentino	97,26
E) Lago Buenos Aires	98,21
F) Magallanes	98,33
G) Río Chico	96,15
Promedio	***97,23***
Varianza	***10,46***

Fuente: Celemin

1) establecer la matriz de pesos a partir de la cantidad de vecinos que tiene cada departamento (contigüidad tipo *Rook*)

Matriz de pesos

	A	B	C	D	E	F	G	Total
A	0	0	*1*	*1*	0	*1*	*1*	*4*
B	0	0	0	0	*1*	*1*	0	2
C	*1*	0	0	*1*	0	0	0	2
D	*1*	0	*1*	0	0	0	*1*	*3*
E	0	*1*	0	0	0	0	*1*	*2*
F	*1*	*1*	0	0	0	0	*1*	*3*
G	*1*	0	0	*1*	*1*	*1*	0	4
Suma								*20*

Matriz de pesos estandarizada

	A	B	C	D	E	F	G	Total
A	0	0	*0,25*	*0,25*	0	*0,25*	*0,25*	*1*
B	0	0	0	0	*0,5*	*0,5*	0	*1*
C	*0,5*	0	0	*0,5*	0	0	0	*1*
D	*0,33*	0	*0,33*	0	0	0	*0,33*	*1*
E	0	*0,5*	0	0	0	0	*0,5*	*1*
F	*0,33*	*0,33*	0	0	0	0	*0,33*	*1*
G	*0,25*	0	0	*0,25*	*0,25*	*0,25*	0	*1*

Departamento	% Baño Exclusivo (x_i)	Promedio ($\bar{x}$)	$x_i - \bar{x}$ (z_i)	Varianza (m_2)	z_i / m_2
A) Corpen Aike	98,92	97,23	1,69	10,46	*0,1616*
B) Deseado	95,36	97,23	-1,87	10,46	*-0,1788*
C) Guer Aike	96,35	97,23	-0,88	10,46	*-0,0841*
D) Lago Argentino	97,26	97,23	0,03	10,46	*0,0029*
E) Lago Buenos Aires	98,21	97,23	0,98	10,46	*0,0937*
F) Magallanes	98,33	97,23	1,10	10,46	*0,1051*
G) Río Chico	96,15	97,23	-1,08	10,46	*-0,1032*

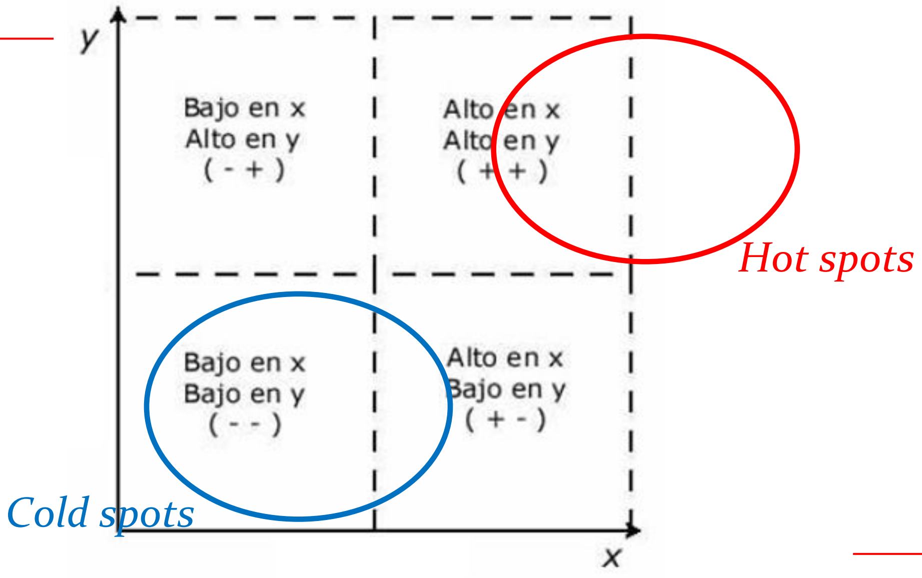
Fuente: Celemin (2009)

$$I_A = (98{,}92 - 97{,}23) / 10{,}46 \text{ x } [(0{,}25 \text{ x } {-0{,}88}) + (0{,}25 \text{ x } 0{,}03) + (0{,}25 \text{ x } 1{,}10) + (0{,}25 \text{ x } {-1{,}08})]$$
$$0{,}1616 \text{ x } [-0{,}22 + 0{,}0075 + 0{,}275 + (-0{,}27)] = \mathit{-0{,}0335}$$
$$I_B = (95{,}36 - 97{,}23) / 10{,}46 \text{ x } [(0{,}5 \text{ x } 0{,}98) + (0{,}5 \text{ x } 1{,}10)]$$
$$-0{,}1788 \text{ x } [0{,}49 + 0{,}55] = \mathit{-0{,}1859}$$
$$I_C = (96{,}35 - 97{,}23) / 10{,}46 \text{ x } [(0{,}5 \text{ x } 1{,}69) + (0{,}5 \text{ x } 0{,}03)]$$
$$-0{,}0841 \text{ x } [0{,}845 + 0{,}015)] = \mathit{-0{,}0723}$$
$$I_D = (97{,}26 - 97{,}23) / 10{,}46 \text{ x } [(0{,}33 \text{ x } 1{,}69) + (0{,}33 \text{ x } {-0{,}88}) + (0{,}33 \text{ x } {-1{,}08})]$$
$$0{,}0029 \text{ x } [0{,}5577 + (-0{,}2904) + (-0{,}3564)] = \mathit{-0{,}0003}$$
$$I_E = (98{,}21 - 97{,}23) / 10{,}46 \text{ x } [(0{,}5 \text{ x } {-1{,}87}) + (0{,}5 \text{ x } {-1{,}08})]$$
$$0{,}0937 \text{ x } [-0{,}935 + (-0{,}54)] = \mathit{-0{,}1382}$$
$$I_F = (98{,}33 - 97{,}23) / 10{,}46 \text{ x } [(0{,}33 \text{ x } 1{,}69) + (0{,}33 \text{ x } {-1{,}87}) + (0{,}33 \text{ x } {-1{,}08})]$$
$$0{,}1051 \text{ x } [(0{,}5577 + (-0{,}6171) + (-0{,}3564)] = \mathit{-0{,}0437}$$
$$I_G = (96{,}15 - 97{,}23) / 10{,}46 \text{ x } [(0{,}25 \text{ x } 1{,}69) + (0{,}25 \text{ x } 0{,}03) + (0{,}25 \text{ x } 0{,}98) + (0{,}25 \text{ x } 1{,}10)]$$
$$-0{,}1032 \text{ x } [0{,}4225 + 0{,}0075 + 0{,}245 + 0{,}275] = \mathit{-0{,}0980}$$

La suma de los valores locales ($I_A + I_B + I_C + I_D + I_F + I_G$) conforman el valor global de I de Moran = 0,5720

El *Moran Scatterplot* es una herramienta de análisis que permite en un gráfico de dispersión el comportamiento de cada unidad espacial

Valores estandarizados del promedio de los valores en unidades vecinas



Valores estandarizados de una variable para cada unidad espacial

Fuente: Buzai (2005)