Detectando Relaciones Causales en Procesos Espaciales.

Marcos Herrera[1]

[1]CONICET-IELDE
Universidad Nacional de Salta (Argentina)

Seminario de Investigación

25 de Abril del 2013 (Salta, Argentina)

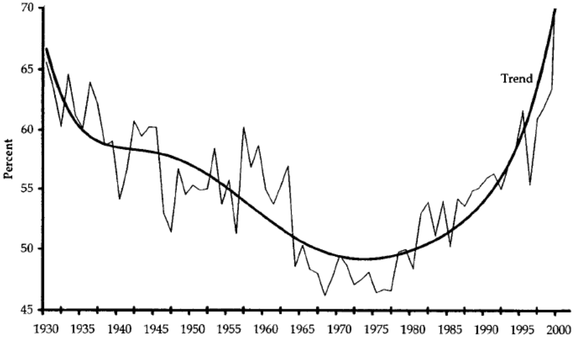
INSTITUTO de ESTUDIOS LABORALES y del DESARROLLO ECONÓMICO

Esquema de Presentación

Esquema de Presentación

Evolución de la Investigación sobre Causalidad

Uso de palabras de la *"familia causal"* en *Economía* (1930-2001)



Hoover (2004). Gráfico muestra el porcentaje de artículos que usan palabras causales ("cause, causes, causality, causation") en comparación a todos los artículos en economía (archivo JSTOR).

Causalidad de Granger y Causalidad Espacial

1. Series de Tiempo: *Causalidad de Granger*
   - **Granger (1969)**: La variable $\{x_t\}_{t\in T}$ *causa a* $\{y_t\}_{t\in T}$ si los valores pasados de x_t mejoran las predicciones de y_t (en términos de varianza condicional).
   - **Granger (1980)** destaca una condición fundamental: la variable causal **debe contener información única** sobre la variable efecto.
2. Corte Transversal: *Causalidad Espacial (Causalidad en Información)*
   - *Diremos que* $\{x_s\}_{s\in S}$ *causa a* $\{y_s\}_{s\in S}$ *si* x_s *brinda* **información incremental** *sobre* y_s *bajo una estructura espacial* W.
     - Enfoque No-Paramétrico: la información es captura por la entropía simbólica condicional.

Diferencias entre Causalidad Temporal y Espacial

Causalidad Temporal: x_t no causa a y_t si $\rho_{yx} = 0$

$$y_t = \rho_{yy} L y_t + \rho_{yx} L x_t + u_y$$

$$x_t = \rho_{xy} L y_t + \rho_{xx} L x_t + u_x$$

Causalidad Espacial: x_s no causa a y_s si $\rho_{yx} = 0$

$$y_s = \rho_{yy} W y_s + \rho_{yx} W x_s + u_y$$

$$x_s = \rho_{xy} W y_s + \rho_{xx} W x_s + u_x$$

Se sustituye pasado/futuro por cercano/distante, pero $L \neq W$. El tiempo es uni-direccional pero el espacio es multi-direccional. Necesitamos definir quién es vecino de quién. Sin embargo, en ambos casos podemos decir que x no causa a y si $\rho_{yx} = 0$.

Esquema de Presentación

Dinámica Simbólica

- *Dinámica simbólica* se basa en la transformación de una serie dentro de una secuencia de símbolos que capturan información que no puede ser observada directamente.
  - Dado un proceso $\{x_s\}_{s \in S}$, definimos una función $f : \{x_s\}_{s \in S} \to \Gamma_n$, donde f es la *función simbolizadora y* Γ_n *es el conjunto de símbolos* $(\sigma^x\,'s)$.
- Conceptos Clave:
  - m es la *dimensión de encaje*: nos brinda la longitud del vecindario.
  - $m -$ *vecindario*: es el conjunto formado por la información de cada observación y sus $(m-1)$ vecinos. *El* $m -$ *vecindario es la información que usamos para construir el símbolo* σ^x.

Mapa Regular 3x3: Ejemplo

$X_{s_1}=4$	$X_{s_2}=1$	$X_{s_3}=3$
$X_{s_4}=6$	$X_{s_5}=2$	$X_{s_6}=5$
$X_{s_7}=1$	$X_{s_8}=2$	$X_{s_9}=4$

$Y_{s_1}=5$	$Y_{s_2}=2$	$Y_{s_3}=4$
$Y_{s_4}=0$	$Y_{s_5}=2$	$Y_{s_6}=3$
$Y_{s_7}=7$	$Y_{s_8}=9$	$Y_{s_9}=3$

Estimamos la MEDIANA de X e Y: $X < M_e^x$ e $Y < M_e^y$ en gris

$X_{s_1} = 4$	$X_{s_2} = 1$	$X_{s_3} = 3$
$X_{s_4} = 6$	$X_{s_5} = 2$	$X_{s_6} = 5$
$X_{s_7} = 1$	$X_{s_8} = 2$	$X_{s_9} = 4$

$Y_{s_1} = 5$	$Y_{s_2} = 2$	$Y_{s_3} = 4$
$Y_{s_4} = 0$	$Y_{s_5} = 2$	$Y_{s_6} = 3$
$Y_{s_7} = 7$	$Y_{s_8} = 9$	$Y_{s_9} = 3$

Función Indicadora: si $X \geq M_e^x \rightarrow 1$ e $Y \geq M_e^y \rightarrow 1$, en otro caso$\rightarrow 0$

$\tau_{s_1} = 1$	$\tau_{s_2} = 0$	$\tau_{s_3} = 1$
$\tau_{s_4} = 1$	$\tau_{s_5} = 0$	$\tau_{s_6} = 1$
$\tau_{s_7} = 0$	$\tau_{s_8} = 0$	$\tau_{s_9} = 1$

$\tau_{s_1} = 1$	$\tau_{s_2} = 0$	$\tau_{s_3} = 1$
$\tau_{s_4} = 0$	$\tau_{s_5} = 0$	$\tau_{s_6} = 1$
$\tau_{s_7} = 1$	$\tau_{s_8} = 1$	$\tau_{s_9} = 1$

Supongamos $m = 4$, entonces el $4-$*vecindario* para X e Y en s_1 es: $X_4(s_1) = (\tau_{s_1}, \tau_{s_2}, \tau_{s_4}, \tau_{s_5})$ e $Y_4(s_1) = (\tau_{s_1}, \tau_{s_2}, \tau_{s_4}, \tau_{s_5})$

$\tau_{s_1} = 1$	$\tau_{s_2} = 0$	$\tau_{s_3} = 1$
$\tau_{s_4} = 1$	$\tau_{s_5} = 0$	$\tau_{s_6} = 1$
$\tau_{s_7} = 0$	$\tau_{s_8} = 0$	$\tau_{s_9} = 1$

$$X_4(s_1) = (1, 0, 1, 0)$$

$\tau_{s_1} = 1$	$\tau_{s_2} = 0$	$\tau_{s_3} = 1$
$\tau_{s_4} = 0$	$\tau_{s_5} = 0$	$\tau_{s_6} = 1$
$\tau_{s_7} = 1$	$\tau_{s_8} = 1$	$\tau_{s_9} = 1$

$$Y_4(s_1) = (1, 0, 0, 0)$$

Función Simbolizadora: Símbolo (σ_{s_1}) captura el número de vecinos que tienen igual valor al de la localización s_1

$\tau_{s_1} = 1$	$\tau_{s_2} = 0$	$\tau_{s_3} = 1$
$\tau_{s_4} = 1$	$\tau_{s_5} = 0$	$\tau_{s_6} = 1$
$\tau_{s_7} = 0$	$\tau_{s_8} = 0$	$\tau_{s_9} = 1$

$$X_4(s_1) = (1, 0, 1, 0) \to \sigma_{s_1}^{x} = 1$$

$\tau_{s_1} = 1$	$\tau_{s_2} = 0$	$\tau_{s_3} = 1$
$\tau_{s_4} = 0$	$\tau_{s_5} = 0$	$\tau_{s_6} = 1$
$\tau_{s_7} = 1$	$\tau_{s_8} = 1$	$\tau_{s_9} = 1$

$$Y_4(s_1) = (1, 0, 0, 0) \to \sigma_{s_1}^{y} = 0$$

Entropía: Medida de Información

Entropía es una medida de incertidumbre asociada con la distribución de una variable aleatoria. Para una variable aleatoria discreta, como los símbolos, la **entropía simbólica** de $\{x_s\}_{s\in S}$:

$$h_x(m) = -\sum_{\sigma\in\Gamma} p_\sigma ln(p_\sigma).$$

donde p_σ es la probabilidad del símbolo σ. Las probabilidades son aproximadas por la frecuencia relativa de los símbolos ($\hat{p}_\sigma \approx p_\sigma$).
La **entropía simbólica condicional** de x_s dado y_s:

$$h_{x|y}(m) = \sum_{\sigma^y\in\Gamma} p(\sigma^y)\, h_{x|\sigma^y}(m)$$

Esta medida cuantifica la entropía remanente de x_s dado el valor de y_s (en términos simbólicos).

Resumen del Proceso de Simbolización

1. Transformar las variables en una secuencia de símbolos que capturan información espacial: Cada símbolo representa el número de vecinos que se comporta de forma similar a la localización s, en comparación a la mediana del proceso.
2. Los símbolos son variables categóricas con una distribución de probabilidad esperada.
3. Usando la distribución empírica de los símbolos, estimamos una medida de información: entropía condicional.
4. Usando dos medidas de entropía condicional construimos el test de causalidad espacial.

Esquema de Presentación

Test de Causalidad

Sean $\{x_s\}_{s\in S}$ e $\{y_s\}_{s\in S}$ dos procesos espaciales y sea $\mathcal{W}(x,y)$ el conjunto de estructuras de dependencia-espacial entre x e y. Usaremos $\mathcal{X}_{\mathcal{W}} = \{W_i x\}$ e $\mathcal{Y}_{\mathcal{W}} = \{W_i y\}$ con $W_i \in \mathcal{W}(x,y)$ para denotar el conjunto de rezagos espaciales de x e y.

Proponemos la hipótesis nula:

$$\begin{gathered} H_0 : \{x_s\}_{s\in S} \ \textit{no causa a} \ \{y_s\}_{s\in S} \\ \textit{bajo la estrutura espacial } \mathcal{X}_{\mathcal{W}} \text{ e } \mathcal{Y}_{\mathcal{W}} \end{gathered}$$

con el siguiente estadístico:

$$\hat{\delta}(\mathcal{Y}_{\mathcal{W}}, \mathcal{X}_{\mathcal{W}}) = \hat{h}_{y|\mathcal{Y}_{\mathcal{W}}}(m) - \hat{h}_{y|\mathcal{Y}_{\mathcal{W}},\mathcal{X}_{\mathcal{W}}}(m)$$

Si $\mathcal{X}_{\mathcal{W}}$ no contiene información extra sobre y entonces $\hat{\delta}(\mathcal{Y}_{\mathcal{W}}, \mathcal{X}_{\mathcal{W}}) = 0$, en otro caso $\hat{\delta}(\mathcal{Y}_{\mathcal{W}}, \mathcal{X}_{\mathcal{W}}) > 0$.
El valor crítico es obtenido por bootstrap de bloques espaciales.

Experimento Monte Carlo

Los siguientes parámetros globales son utilizados en la generación de procesos:

$$N \in \{400, 800\}\,,\ \rho \in \{0{,}3; 0{,}5; 0{,}7\}\,,\ m \in \{4, 5\}\,,$$

donde N es el tamaño muestral, ρ el parámetro de autocorrelación espacial y m la dimensión de encaje. Usaremos el R^2 esperado para controlar la relación lineal en cada proceso.
Para analiza el tamaño, consideramos los siguientes procesos generadores:

$$\begin{aligned} y &= \rho_y y + \varepsilon, \\ x &= \rho_x x + \nu, \end{aligned}$$

donde $\varepsilon, \nu \sim \mathcal{N}(0,1)$, con: $(\rho_y, \rho_x) \in \{(0,0)\,, (0{,}3, 0{,}7)\,, (0{,}5, 0{,}3)\}$.

Simulaciones Monte Carlo: Casos de Potencia

PGD1: **Modelo Lineal**

$$y = (I - \rho W)^{-1} (\beta x + \theta W x + \varepsilon) .$$

PGD2: **Modelo No Lineal 1**

$$y = \left[(I - \rho W)^{-1} (\beta x + \theta W x + \varepsilon) \right]^{-1} .$$

PGD3: **Modelo No Lineal 2**

$$y = \exp \left\{ \left[(I - \rho W)^{-1} (\beta x + \theta W x + \varepsilon) \right]^{3} \right\} .$$

siendo $x \sim \mathcal{N}(0,1)$, $\varepsilon \sim \mathcal{N}(0,1)$ y $cov(x, \varepsilon) = 0$.

Las tablas muestran la potencia global estimadas del estadístico. Potencia Global hace referencia al caso donde se rechaza H_0 de no causalidad de x a y, pero no podemos rechazar simultáneamente H_0 de no causalidad de y a x. Recordar que la idea es capturar unidireccionalidad en información.

Tamaño Muestral Empírico

Cuadro: Tamaño Muestral del Test $\hat{\delta}\left(\mathcal{Y}_{\mathcal{W}}, \mathcal{X}_{\mathcal{W}}\right)$ al nivel del 5 %

N	ρ_y	ρ_x	m	
			4	5
400	0,0	0,0	5,3	4,5
	0,3	0,7	2,6*	1,2*
	0,5	0,3	3,6*	1,8*
800	0,0	0,0	5,6	4,6
	0,3	0,7	3,4*	1,3*
	0,5	0,3	5,4	2,5*

Nota: Boots: 99. Bloques: 8.

Replicaciones: 1000.

* Significativo al nivel crítico del 5 %.

Potencia Empírica Global al 5 %

$DGP1: \ LINEAL$		$N = 400$		$N = 800$	
m		4	5	4	5
$R^2_{y/x} = 0{,}4$	$\rho = 0{,}3$	93,0	99,0	77,0	96,0
	$\rho = 0{,}5$	93,0	100,0	90,0	99,0
	$\rho = 0{,}7$	87,0	89,0	98,0	98,0
$R^2_{y/x} = 0{,}6$	$\rho = 0{,}3$	88,0	100,0	62,0	95,0
	$\rho = 0{,}5$	100,0	100,0	90,0	95,0
	$\rho = 0{,}7$	99,0	96,0	95,0	100,0
$R^2_{y/x} = 0{,}8$	$\rho = 0{,}3$	89,0	100,0	55,0	86,0
	$\rho = 0{,}5$	97,0	99,0	87,0	99,0
	$\rho = 0{,}7$	98,0	98,0	97,0	100,0

Nota: Boots: 99. Bloques: 8. Replicaciones: 100.

Potencia Empírica Global al 5 %

$DGP2: \ NO - LINEAL\,1$		$N=400$		$N=800$	
m		4	5	4	5
$R^2_{y/x}=0{,}4$	$\rho=0{,}3$	87,0	93,0	91,0	97,0
	$\rho=0{,}5$	89,0	90,0	89,0	93,0
	$\rho=0{,}7$	60,0	73,0	84,0	96,0
$R^2_{y/x}=0{,}6$	$\rho=0{,}3$	91,0	98,0	92,0	94,0
	$\rho=0{,}5$	95,0	90,0	92,0	94,0
	$\rho=0{,}7$	78,0	71,0	93,0	97,0
$R^2_{y/x}=0{,}8$	$\rho=0{,}3$	95,0	95,0	86,0	96,0
	$\rho=0{,}5$	97,0	94,0	93,0	96,0
	$\rho=0{,}7$	86,0	78,0	90,0	93,0

Nota: Boots: 99. Bloques: 8. Replicaciones: 100.

Potencia Empírica Global al 5 %

$DGP3: \ NON - LINEAL\,2$		$N = 400$		$N = 800$	
m		4	5	4	5
$R^2_{y/x} = 0{,}4$	$\rho = 0{,}3$	97,0	100,0	80,0	89,0
	$\rho = 0{,}5$	97,0	99,0	93,0	97,0
	$\rho = 0{,}7$	89,0	93,0	96,0	100,0
$R^2_{y/x} = 0{,}6$	$\rho = 0{,}3$	87,0	97,0	74,0	81,0
	$\rho = 0{,}5$	95,0	96,0	84,0	95,0
	$\rho = 0{,}7$	98,0	92,0	95,0	98,0
$R^2_{y/x} = 0{,}8$	$\rho = 0{,}3$	85,0	97,0	69,0	83,0
	$\rho = 0{,}5$	91,0	99,0	89,0	90,0
	$\rho = 0{,}7$	99,0	97,0	96,0	98,0

Nota: Boots: 99. Bloques: 8. Replicaciones: 100.

Tamaño Empírico bajo Especificación Errónea

N	ρ_y	ρ_x	m	
			3	5
400	0,0	0,0	4,9	3,7*
	0,3	0,7	4,0	1,0*
	0,5	0,3	4,9	1,7*
800	0,0	0,0	4,7	5,9
	0,3	0,7	5,4	1,9*
	0,5	0,3	4,3	2,7*

Nota: Boots: 99. Bloques: 8.

Replicaciones: 1000.

* Significancia al nivel crítico del 5 %.

$P.G.D.$ fue generado usando $m_{óptimo} = 4$.

Potencia Empírica Global bajo Especificación Errónea

$P.G.D,3$		$N = 400$		$N = 800$	
m		3	5	3	5
$R^2_{y/x} = 0{,}4$	$\rho = 0{,}3$	66,0	95,0	82,0	92,0
	$\rho = 0{,}5$	55,0	94,0	80,0	94,0
	$\rho = 0{,}7$	35,0	67,0	68,0	100,0
$R^2_{y/x} = 0{,}6$	$\rho = 0{,}3$	76,0	97,0	70,0	91,0
	$\rho = 0{,}5$	62,0	93,0	66,0	96,0
	$\rho = 0{,}7$	46,0	84,0	77,0	99,0
$R^2_{y/x} = 0{,}8$	$\rho = 0{,}3$	76,0	98,0	75,0	89,0
	$\rho = 0{,}5$	74,0	98,0	75,0	95,0
	$\rho = 0{,}7$	53,0	88,0	78,0	99,0

Nota: Boots: 99. Bloques: 8. Replicaciones: 100.

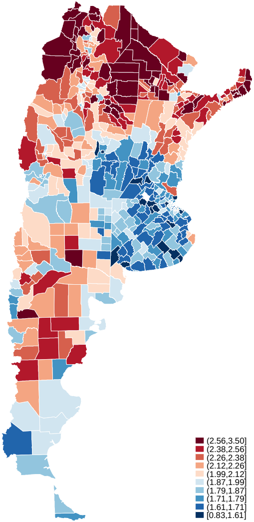
$P.G.D.$ generado bajo $m_{óptimo} = 4$. Val. Crit. 5 %

Esquema de Presentación

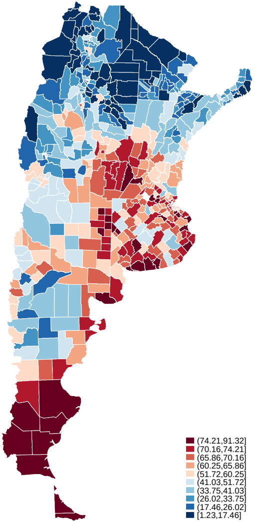
Utilidad Empírica del Test de Causalidad Espacial

- Estudio de causalidad entre fecundidad y el nivel de ingreso.
- Existe evidencia empírica no concluyente a la hora de establecer causalidad (de Granger):
  - Evidencia favorable a que el *INGRESO* causa a la *FECUNDIDAD* (Bonitsis y Geithman, 1987).
  - Otros estudios encuentran evidencia empírica de bidireccionalidad (ambas se determinan simultáneamente).
- Problema: para Argentina NO existe la posibilidad de contrastar esta hipótesis dada la falta de series temporales extensas.
- Posibilidad: puede aplicarse el test de causalidad $\hat{\delta}\left(\mathcal{Y}_{\mathcal{W}}, \mathcal{X}_{\mathcal{W}}\right)$ siempre que el espacio geográfico sea relevante.
- Información disponible por departamento: *% de hogares no pobres* (proxy de ingreso), *tasa de fecundidad promedio*.

Análisis Exploratorio

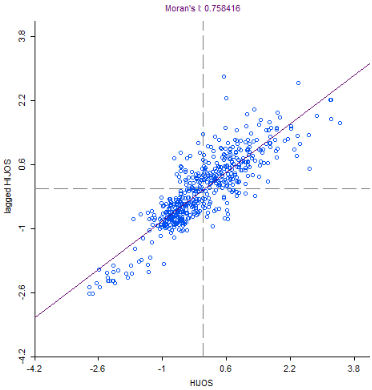


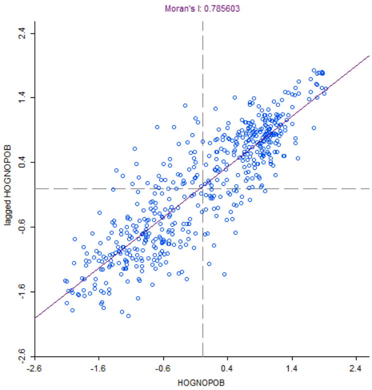
Tasa de Fecundidad (15-49), 2001



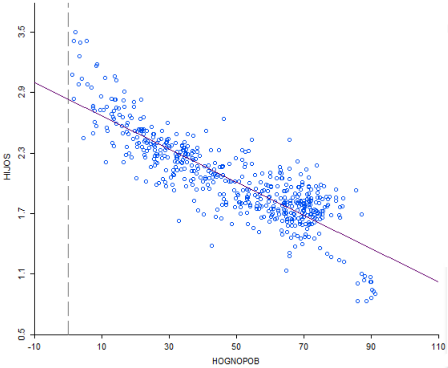
Hogares No Pobres (%), 2001

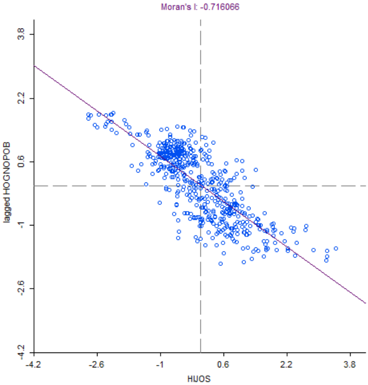
Tests de Moran Univariantes





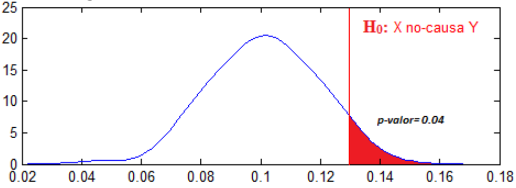
Relación Bivariante y Moran Bivariante

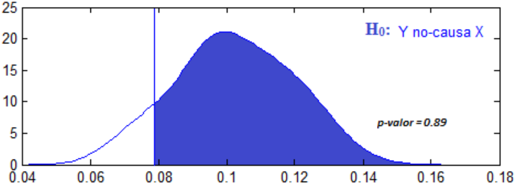




Resultados del Test $\hat{\delta}\left(\mathcal{Y}_W, \mathcal{X}_W\right)$. Distribución Empírica

Rechazamos No-causalidad de HOGNOPOB a HIJOS y no podemos rechazar No-causalidad de Hijos a HOGNOPOB. CONC.: "HOGNOPOB causa a HIJOS"





Nota 1: $Y = HIJOS$; $X = HOGNOPOB$.

Nota 2: 99 boots, 8 bloques, $m = 5$, $N = 528$.

Esquema de Presentación

Resumen

- Se presenta un concepto operativo de causalidad en modelos espaciales: Unidireccionalidad en información.
- El test es simple de obtener, libre de supuestos restrictivos (formas funcionales u otros aspectos de especificación).
- El nivel de significancia es obtenido mediante bootstrap de bloques espaciales.
- El test $\hat{\delta}\left(\mathcal{Y}_{\mathcal{W}}, \mathcal{X}_{\mathcal{W}}\right)$ tiene buen comportamiento respecto al tamaño empírico y potencia global bajo procesos lineales y no lineales.
- El test tiene buen comportamiento aún cuando se elige una dimensión de encaje superior a la óptima $(m > m_{optimal})$.
- La aplicación empírica muestra la utilidad del test ante la imposibilidad de aplicar Granger.
    - La causalidad espacial va desde HOGNOPOB a HIJOS.

Gracias por su Atención