ECONOMETRÍA - BOOTSTRAP - ECONOMETRÍA ESPECIAL

David Barrera Ojeda

Objetivo

La idea del presente trabajo, es un intento de mostrar los pasos a seguir antes de llegar a aplicar econometría espacial, cuando las variables son georeferenciadas , es importante señalar que se debe agotar las otras que nos brinda la econometría clásica, es importante señalar que la mayoría de las pruebas de la econometría clásica no se replican en econometría espacial, porque en el fondo es la aplicación de MCP (mínimos cuadrados ponderados).

En algunos trabajos es inevitable, su aplicación, si así fuese existen varias técnicas para determinar la matriz de contigüidad, en el presente trabajo se emplea la más simple.

Introducción

Uno de los importantes avances en la ciencia económica de la ultima década ha sido la reincorporación explícita del efecto del espacio geográfico en el análisis de los problemas económicos. A partir de los trabajos de Krugman (1991a y 1991b, et al 1998) sobre lo que se ha llamado la “nueva geografía económica”, resaltando el papel de las externalidades espaciales en los modelos de comercio internacional y crecimiento, se han multiplicado los modelos que estudian la influencia del espacio sobre la localización de empresas, desarrollo de complejos industriales, desarrollo de las inversiones sociales, difusión del conocimiento y la tecnología, etc.

Econometría espacial

Anselin (1988), es probablemente la referencia más citada en los trabajos de econometría espacial, la define como “la colección de técnicas que lidian con las peculiaridades causadas por el espacio en el análisis estadístico de los modelos de la ciencia regional”. Once años más tarde, Anselin (1999) extiende la definición diciendo que “la econometría espacial es una rama de la econometría que se preocupa del tratamiento adecuado de la interacción espacial (autocorrelación espacial) y la estructura espacial (heterogeneidad espacial) en modelos de regresión con datos de corte transversal y de panel de datos". Para más detalle, Paelinck y Klaassen (1979).

Ejemplo Práctico

Se estudia, 15 municipios de la ciudad de La Paz, ¿por qué solamente 15?, debido a las siguientes razones:

1. La mayoría de los municipios no poseen informes, de las cuatro variables del modelo. No hay lugar para los missing.
2. Se reduce a las observaciones colindantes, puesto que para aplicar econometría espacial, necesariamente los municipios deben ser colindantes.

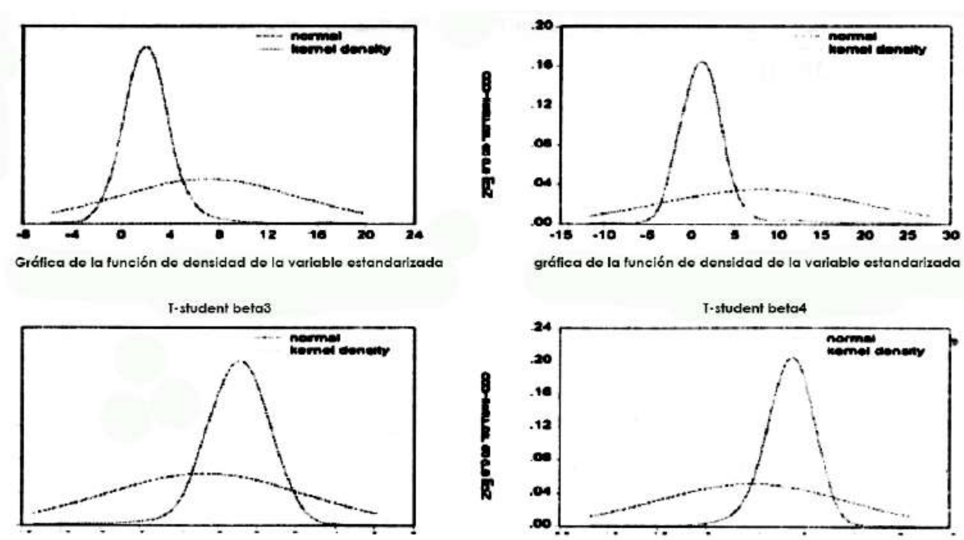
Las variables a estudiar son:

- INVSM: Inversión Social Municipal, son los recursos que el FIS desembolsa en los municipios (año 1997).
- RECOP: Recursos Propios, son los recursos propios con que cuenta cada municipio (año 1997).
- PAUS: Es el porcentaje de ausentismo electoral, en las elecciones municipales de 1995.
- IDH: Es el índice de Desarrollo Humano, alcanzado en los 15 municipios (año 1997).

1. Econometría Clásica

1.a. Pruebas Bootstrap

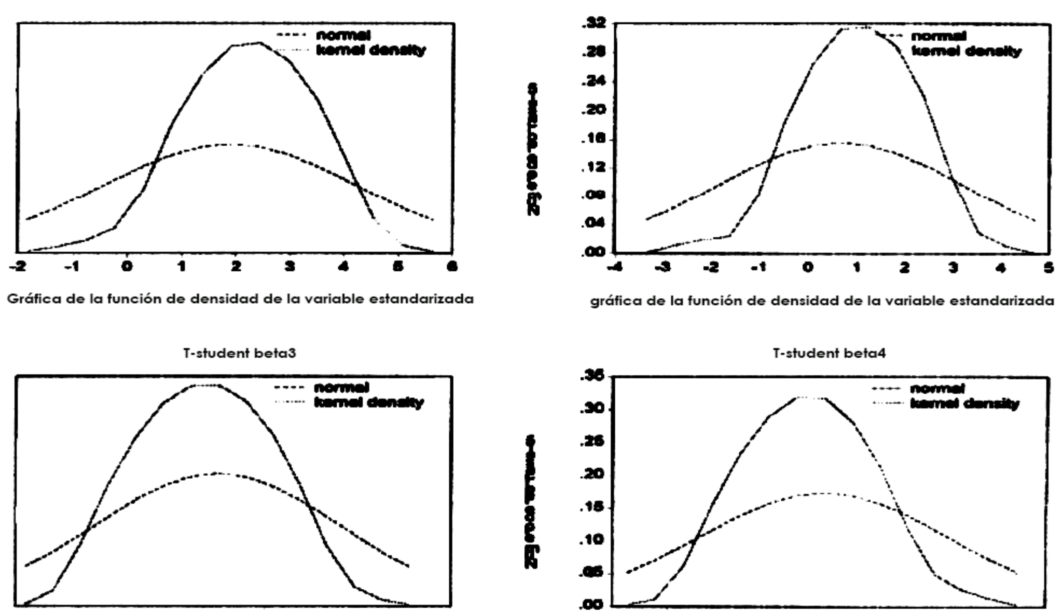
Número de réplicas=1000				
Estadísticos de los T-student				
Coeficientes	**cons**	**recop**	**paus**	**idh**
Mean	2.142	1.694	-0.962	-1.164
Median	2.060	1.353	-0.992	-1.131
Skewness	0.749	7.654	-0.231	-0.686
Kurtosis	5.354	118.719	8.763	5.442
Test de Normalidad				
Shapiro-Francia	8.102	9.217	7.429	7.182
Shapiro-Wilk	9.917	11.704	8.913	8.495
Sesgo	-0.248	0.510	0.049	0.353



La presencia de los valores outliers, ocasiona problemas serios, como por ejemplo: Los T-stat son bastantes inestables, la significación recop (recursos propios) es la más afectada, en el sesgo se debe tener cuidado, puesto que se está tomando la media, puede ocasionar inconsistencias por el problema de la poca normalidad de los T-student.

Prueba Jackknife

Pruebas Jackknife				
Número de obs.= 15				
Jackknife de los T-student				
Estadísticos	**C**	**RECOP**	**PAUS**	**IDH**
Media	2.237	1.016	-0.963	-1.384
Error.est	2.183	2.529	1.280	2.179
Sesgo	-2.632	-0.171	1.012	2.090



Con la técnica de remuestreo jackknife, ocurre que el sesgo es mínimo, en la significación de recop, sin embargo es la variable que más varía.

1.b. Estimadores Robustos

Para subsanar ese problema: una alternativa es obtener estimadores robustos:

Regressión Robusta				
Método=Huber				
Variables	Coeficientes	Erorr.Est.	T-stat	Prob
C	21.45	12.979	1.653	0.127
RECOP	-0.095	0.127	-0.747	0.471
Gft	-11.524	11.139	-1.035	0.323
IDH	-8.976	26.548	-0.338	0.742
Suma de los residuos al cuadrado=404.318904				
Error Estandar de la regresión=6.062694451				
Valor de Máximo Verosimilitud= -45.99023099				

Regressión Robusta				
Método=Ramsay				
Variables	Coeficientes	Erorr.Est.	T-stat	Prob
C	11.293	6.552	1.724	0.113
RECOP	-0.183	0.062	-2.96	0.013
PAUS	-9.871	5.416	-1.822	0.096
IDH	12.46	13.162	0.947	0.364
Suma de los residuos al cuadrado=114.6177953				
Error Estandar de la regresión=3.227971433				
Valor de Máximo Verosimilitud=-36.53572455				

Regressión Robusta				
Método=Andrew				
Variables	Coeficientes	Erorr.Est.	T-stat	Prob
C	17.753	11.261	1.576	0.143
RECOP	-0.143	0.111	-1.29	0.223
PAUS	-10.839	9.541	-1.136	0.28
IDH	-0.77	23.085	-0.033	0.974
Suma de los residuos al cuadrado=320.7480712				
Error Estandar de la regresión=5.399899589				
Valor de Máximo Verosimilitud= -44.25362142				

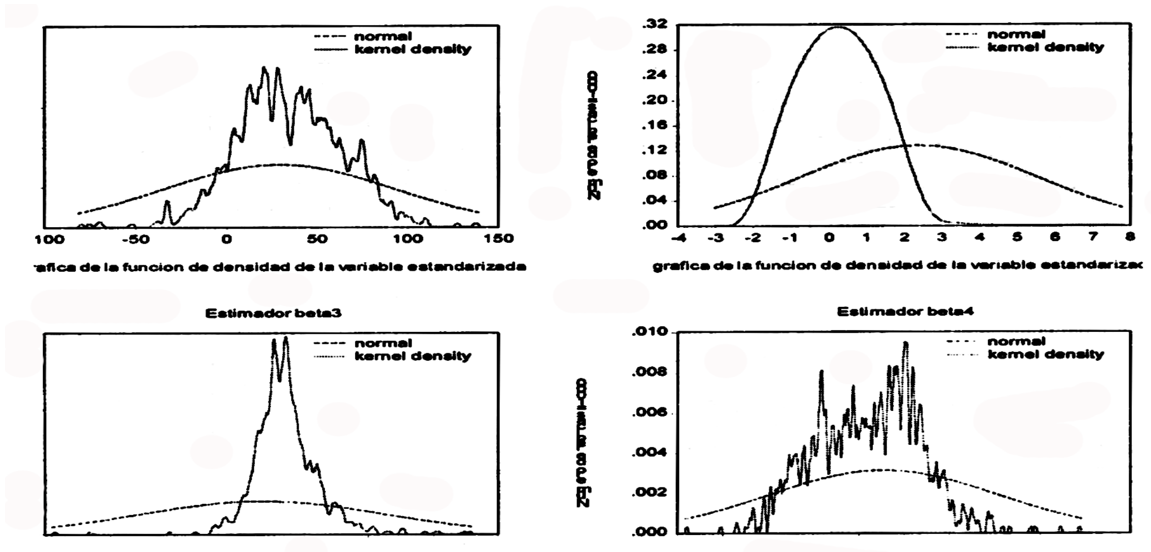
Regressión Robusta				
Método=Tuckey				
Variables	Coeficientes	Erorr.Est.	T-stat	Prob
C	26.152	5.224	5.006	0.000
RECOP	1.697	0.208	8.172	0.000
PAUS	-20.442	3.424	-5.969	0.000
IDH	-22.763	9.794	-2.324	0.040
Suma de los residuos al cuadrado=40.11289226				
Error Estandar de la regresión=1.909614246				
Valor de Máximo Verosimilitud=-28.66143488				

El único método de Tukey arroja los mejores T-stat.

Nota.- Con este resultado hay que asumir con bastante cuidado, puesto que los valores outliers pueden ocasionar un efecto devastador, para comprobar, vamos a bootstropear los coeficientes.

1.c Bootstrap de los Coeficientes

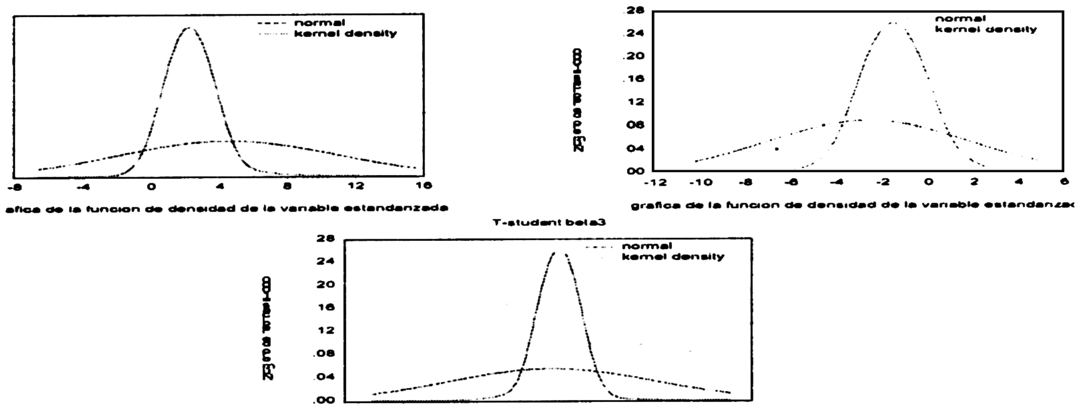
Resample Bootstrap				
Regresando=invs.				
Regresores=recop paus idh				
Número de réplicas=1000				
Estadísticos de los Coeficientes				
Coeficientes	cons	recop	paus	Idh
Mean	34.514	0.273	-11.334	-39.942
Median	33.157	0.257	-12.821	-36.502
Skewness	-0.056	2.850	0.404	0.015
Kurtosis	3.296	27.035	10.010	2.968
Test de Normalidad				
Shapiro-F	2.684	9.078	7.963	3.794
Shapiro-Wilk	2.588	11.46	9.679	4.002
Sesgo	-12.271	0.092	7.545	20.626



Definitivamente los valores outliers, generan inestabilidad de los estimadores de los coeficientes, ninguno es significativo, por esta razón vamos a excluir la variable recop, pero no podemos reducir los municipios, puesto que como son datos georeferenciales esto ocasionaría sesgos impredecibles, por esto excluimos recop.

1.d Bootstrap de los T-Student

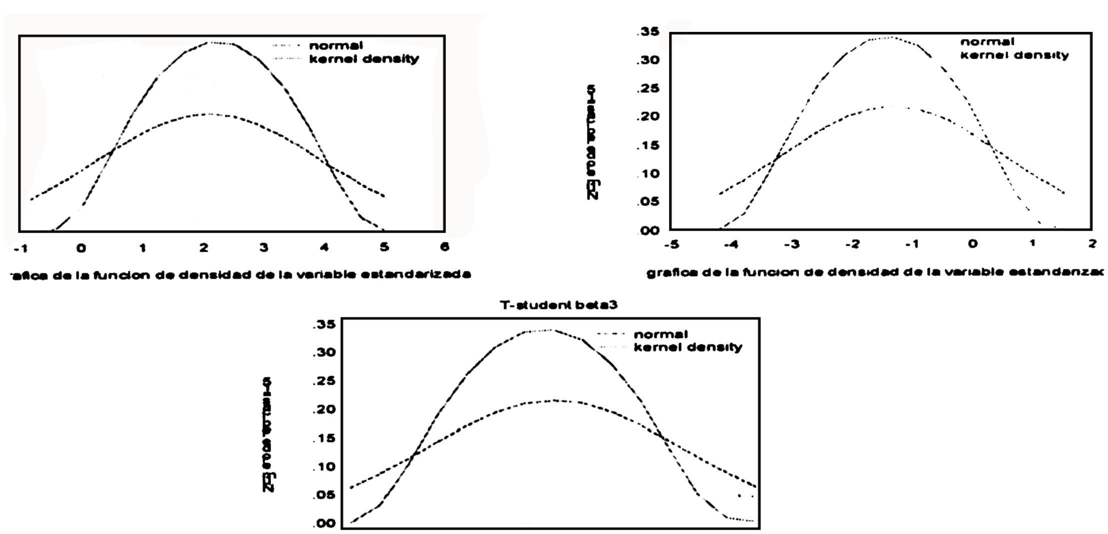
Regresando=invsm			
Regresores=recop paus			
Número de réplicas=1000			
Estadísticos de los T-Student			
Coeficientes	**cons**	**paus**	**Idh**
Mean	2.391	-1.531	-1.244
Median	2.250	-1.529	-1.241
Skewness	1.643	-0.510	-0.197
Kurtosis	12.976	6.406	13.763
Test de Normalidad			
Shapiro-Francia	8.7	7.063	8.629
Shapiro-Wilk	10.854	8.391	10.718
Sesgo	0.055	-0.045	-0.001



Los T-stat no mejoran, debido al coeficiente de curtosis, la constante es una excepción pero no interesa en este caso.

1.e Jacknife de los T-student

Pruebas Jackknife			
Número de obs.=15			
Jackknife de los T-student			
Estadísticos	**C**	**PAUS**	**IDH**
Media	2.236	-1.417	-1.190
Error.est	1.082	1.030	0.993
Sesgo	-1.408	0.971	0.741



Resulta que ninguno es significativo, otra alternativa es aplicar econometria espacial.

2. Econometría espacial

2.a. Matriz de Contigüidad

Uno de los elementos fundamentales de la econometria espacial es su forma de utilizar la información geográfica contenida en las observaciones de procesos que ocurren espacialmente. En este sentido, muchas de las técnicas desarrolladas en la geoestadística y la estadística espacial han sido adaptadas para capturar los efectos espaciales en la estimación de modelos económicos.

Una de las formas más comunes de representar la ubicación geográfica de un conjunto de polígonos es a través de una Matriz de Conectividad o de Contigüidad. Esta es una matriz cuadrada que tiene el mismo número de filas (municipios) o columnas (municipios) que el número de polígonos independientes del mapa en estudio y que por convención se la denomina por W.

Los valores utilizados para representar vecindad son variados. La formulación más simple es una matriz de contigüidad binaria, es decir, los elementos de W serán igual a 1 si dos polígonos (municipios) son vecinos y cero en otro caso.

Esta matriz de contigüidad tiene ceros en la diagonal principal porque se asume que un polígono (municipio) no puede ser vecino consigo mismo. Adicionalmente, en la práctica esta matriz se estandariza por filas, es decir, se divide cada componente de la fila de la matriz por la suma de todos los elementos de esa fila de modo que la suma de cada fila es igual a uno, esta forma es muy útil para crear los rezagos espaciales.

2.b. Dependencia o autocorrelación espacial

Para poder aplicar, un modelo de econometria espacial, debe existir autocorrelación espacial, esto depende de la técnica que se aplique, entre ellas Moran, a la vez también depende como definamos la matriz de contigüidad (W), por simplicidad en el ejemplo usaremos, la contigüidad binaría luego estandarizada.

meth: 'moran'	istat: 1.3706
nobs: 15	imean: -0.0525
nvar: 3	ivar: 0.0185
morani: 0.1339	prob_0.1560

A un nivel de significación del 5%, no existe correlación espacial

Sólo con fines ilustrativos, vamos a suponer que efectivamente existe autocorrelación espacial, tendríamos los siguientes resultados:

Estimadores de los coeficientes:

Cte : 34.2688	Paus:-0.8285
Paus:-17.9821	Idh :2.0323
Idh >26.3695	R^2: 0.3198
Los T-student:	Rho:0.495
Cte:-1.3544	

Nota.- Observemos que los T-Student, mejora la variable IDH, si empleamos otras técnicas para determinar W, probablemente mejoren los T-Student, y esto es explicable, porque conociendo la ubicación geográfica de los 15 municipios, donde se aplicó no sea el más recomendable, pero el trabajo no es concluyente, sino más ilustrativo, mostrando las ventajas que posee frente a la econometría clásica.

Nota.- Para determinar la matriz W se debe contar con una información de primera, totalmente actualizada de los 15 municipios.

Nota.- Para dar validez a los estadísticos T-Student, es recomendable aplicar las pruebas de remuestreo a la econometría espacial, para esto se debe contar con un equipo de software de alta resolución.

Nota.- El coeficiente de autocorrelación espacial señala que existe una correlación de los residuos del 0.495 (autocorrelación de primer orden), es decir, si realmente se desea subsanar éste problema, se debe usar procesos ARMA en econometría espacial, esto complica ligeramente su tratamiento, es un poco menos simple.

Bibliografía:

1.- RYAN, D.L., VON HOHENBALKEN, B. y WEST, D.S. (1990), “An Econometric-Spatial Analysis of the Growth and Decline of Shopping Centers”. Regional Science and Urban Economics, n°20 pp. 313-326.

2.- SPATIAL ECONOMETRICS USING MATLAB, JAMES P. LESAGE Department of Economics, University of Toledo, August, 1999

3.- METODO DE REMUESTREO EN SERIES TEMPORALES, DR. ANDRES ALONSO FERNÁNDEZ, Departamento de matemáticas, Universidad Autónoma de Madrid, 2002.

Software utilizado: MATLAB, SPACETATE y MACROEVIEWS

Índice