

Ejercicio 3 : Regresión de Poisson

Antes del **Miércoles a las 16.00** horas se deben subir dos ficheros:

- un fichero script de `gretl` que compute todo lo necesario para responder las preguntas del ejercicio. El nombre del fichero debería ser

E2QM_tugrupo_tunombre.inp.

- un fichero de texto con las respuestas a las preguntas. El nombre del fichero de texto debería ser

E2QM_tugrupo_tunombre.txt.

Los ficheros se pueden subir solo una vez.

Las respuestas se tendrán que imprimir y entregar en clase también. Quien no pueda asistir a clase, tendrá que comunicarlo con antelación—o, caso de que la ausencia sea imprevista, en la siguiente clase— y deberá entregar siempre un justificante en la clase siguiente.

Los scripts que no corran correctamente no serán evaluados.

-
1. **Infraestructura de Transporte por Carretera:** En este ejercicio vamos a aplicar la estimación de Poisson para probar los efectos que la infraestructura de transporte tienen en la actividad económica. La investigación original sobre la que se basa este ejercicio fue publicada en *Regional Science and Urban Economics* en 2004 por Adelheid Holl. El trabajo analiza el impacto de la infraestructura vial de la ubicación de nuevos establecimientos manufactureros en los municipios españoles desde 1980 a 1994, un período en el que la mayor parte de la red de autovías y autopistas se construyó. El autor sostiene que mejores conexiones de transporte pueden hacer que las áreas de menor actividad económica ganen atractivo a medida que adquieren un mejor acceso a los mercados en las áreas centrales. Pero, al mismo tiempo, la competencia proveniente de las empresas existentes en las aglomeraciones económicas puede aumentar. Con la movilidad de las empresas, la distribución de los beneficios de la inversión en infraestructura no está a priori clara. El archivo `Holl2004.gdt` contiene una muestra aleatoria de 1540 Municipios españoles en 1991, que supone una submuestra de los datos utilizados en el artículo original. El conjunto de datos contiene información a nivel de municipio sobre el número de nuevas plantas en manufacturas (*niman0*), la distancia en km del centro geográfico del municipio a la autopista más cercana (*dis12*), el área en kilómetros cuadrados (*area*), la proporción de la población ocupada que trabaja en el sector de las manufacturas (*pmanu*), la proporción de la población con un título universitario (*puni*), y los códigos INE de provincia y municipio.

- (a) Tabule $niman0$ y proporcione estimaciones de $\Pr(niman0 = 0)$, $E(niman0)$ y también de $E(niman0|niman0 > 0)$. Comente los resultados.
- (b) Proporcione estimaciones del siguiente modelo:

$$\Pr(niman0 > 0|x) = \Phi(\beta_0 + x'\beta)$$

donde $x = \{dis12, area, pmanu, puni\}$ y $\Phi(\cdot)$ representa la función de densidad acumulativa de la normal estándar. ¿La presencia de una autopista cerca del municipio aumenta la probabilidad de nuevas plantas en el sector de las manufacturas?

- (c) Usted quiere evaluar cuántas nuevas plantas se espera que aparezcan en un municipio cuando una nueva autovía se construye cerca. Usted estima el modelo de regresión lineal:

$$niman0 = \beta_0 + x'\beta + \epsilon$$

donde supone que $E(\epsilon|x) = 0$. ¿Cuál es el cambio estimado en nuevas plantas cuando la distancia de la autopista más cercana disminuye de 50 km (el promedio más una des.típ.) a 5 km? ¿Es éste un efecto significativo? Teniendo en cuenta sus respuestas en la parte (a), ¿Cree que el supuesto $cov(\epsilon|x) = 0$ es razonable?

- (d) Teniendo en cuenta que sólo números no negativos y discretos son posibles para $niman0$ y que muchos municipios tienen cero nuevas plantas, mientras que otros municipios pueden tener más de diez nuevas plantas, se tiene en cuenta el modelo de Poisson:

$$\Pr(niman0 = y|x) = \frac{e^{y(\beta_0 + x'\beta)} e^{-\beta_0 + x'\beta}}{y!}.$$

¿Cómo calcularía en el modelo de Poisson el cambio esperado en nuevas plantas cuando la distancia de la autopista más cercana disminuye de 50 a 5 km?

- (e) Estime el modelo de la parte (d). ¿Cuál es, de acuerdo con el modelo de Poisson, el cambio estimado en nuevas plantas cuando la distancia más cercana a la autopista disminuye desde 50 a 5 km? ¿Es éste un efecto significativo? ¿Hay evidencia de sobre-dispersión? Calcule los errores estándar correctos.
- (f) Ud. considera una especificación más flexible para la distancia, incluyendo término cuadrático en la distancia. ¿Sus conclusiones de la parte (e) cambian? Interprete ahora los efectos marginales de construir una nueva autopista.
- (g) En las clases teóricas hemos declarado que hay otro modelo (el modelo binomial negativo) que es apropiado cuando se tiene exceso de dispersión. Estime el modelo binomial negativo (con el comando `negbin`). Teniendo en cuenta que los coeficientes β tienen la misma interpretación que en el modelo de Poisson, explique cómo cambian sus conclusiones cuando resuelve el problema de sobre-dispersión.