

Análisis de Regresión y Causalidad 7

Estadística Urbana

Sebastián Espinoza Rojas

Universidad de Chile
Primavera 2017

- En ausencia de tratamientos asignados aleatoriamente se desarrollaron otros métodos que permiten encontrar el efecto.
- Entre estos métodos se encuentran Matching Propensity Score y Regresión Discontinua.
- La idea principal de estos métodos es tratar de encontrar grupos que sean iguales en todas sus características salvo en el programa.

Matching Propensity Score

Introducción

- El método de Matching o Pareo consiste en encontrar en la muestra un *clon estadístico* de la persona tratada de tal forma que sólo sean distintos en términos del tratamiento.
- Si se cumple lo anterior, cualquier diferencia en términos de la variable de resultado Y se deberá al tratamiento.
- En STATA se ocupa el comando *psmatch2*.

Matching Propensity Score

Ejemplo

- Si se quiere determinar el efecto en términos de gentrificación de una liberalización de las normas de construcción en la región metropolitana se tendría que no necesariamente las comunas desreguladas son iguales que las reguladas (condición para calcular la diferencia simple).
- Dado eso, mediante el método de matching podríamos buscar comunas con iguales características observables (ingreso promedio, cantidad de habitantes, etc) a las de las desreguladas, de tal forma que su única diferencia sea la desregulación.
- Si la única diferencia entre ambas es la desregulación, la diferencia de gentrificación entre las comunas será sólo debido a la desregulación.

Matching Propensity Score

Metodología

- Lo que hace el método es considerar un conjunto de características que puede determinar una probabilidad de ser tratado. Luego, para a todas aquellas personas que tienen idéntica probabilidad de ser tratado.
- Dado eso, mediante el método de matching podríamos buscar comunas con iguales características observables (ingreso promedio, cantidad de habitantes, etc) a las de las desreguladas, de tal forma que su única diferencia sea la desregulación.
- Si la única diferencia entre ambas es la desregulación, la diferencia de gentrificación entre las comunas será sólo debido a la desregulación.

Matching Propensity Score

Supuestos

- Existen dos elementos relevantes que mencionar: el supuesto de igualdad en no observables y la existencia de soporte común.
- Si queremos que la causa de la diferencia entre ambos grupos sea el tratamiento no pueden existir características no observables que contaminen los resultados, ya que alguna de ellas podría ser la causante de las diferencias.
- Además de lo anterior, debe existir soporte común, esto es, gente que se pueda parrear. Esto último, si bien es necesario que exista, se puede testear, por lo que no constituye el principal supuesto de la metodología.

Regresión Discontinua

Introducción

- La idea principal de esta metodología es explotar políticas tales que la participación del programa depende de un punto de corte. Ej: Entrar a la u depende del puntaje ponderado del estudiante.
- La metodología explota la discontinuidad en torno al punto de corte. Se asumirá que la gente a cada lado del punto de corte es estadísticamente igual entre sí, de tal forma que diferencias en su variable de resultado se van a deber exclusivamente al tratamiento.
- En STATA se ocupa el comando *rdrobust*.

Regresión Discontinua

Ejemplo

- Asuma que volvemos a analizar el impacto que tiene en términos de salud la entrega de recursos en invierno a familias vulnerables. Suponga que ahora, el alcalde de la comuna analizada decide entregarle los recursos a aquellas familias que tengan un puntaje de ficha de protección social igual o menor a 7.000 puntos.
- Posiblemente, una familia que tiene 6.900 puntos es estadísticamente igual a otra que tiene 7.100, no obstante una recibe el beneficio y la otra no. Luego, cualquier diferencia en términos de visitas al doctor entre ellas se deberá exclusivamente al tratamiento.

Regresión Discontinua

Supuestos

- El supuesto de identificación relevante es el de igualdad en No Observables (igual que en matching). Luego, si queremos acusar causalidad entre ambos grupos, necesariamente la única diferencia debe ser la asistencia al programa. Por construcción, ambos grupos son iguales en observables, pero no en No Observables.
- Se debe tener en consideración que aquí sólo se encontrará un efecto local válido sólo para el grupo de interés (en torno al puntaje de corte), y no para toda la población. *Se debe notar que una persona con 18.000 y otra con 1.000 son muy distintas, por lo que sería difícil compararlas.