PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

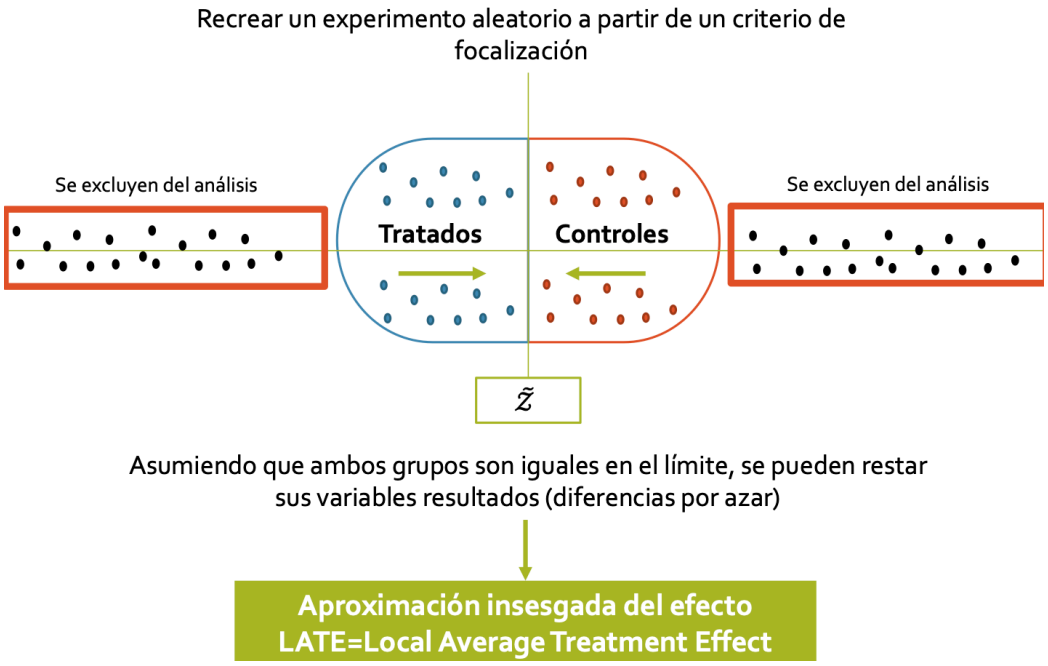
Taller 5: Regresión Discontinua

Datos para la evaluación de impacto en políticas públicas
Diplomado en Ciencia de Datos para Políticas Públicas

22 de Octubre 2022

Análisis de Regresión Discontinua

El método de regresión discontinua (RD) está basado en la idea de que, en contextos altamente reglamentados, existen algunas reglas arbitrarias que pueden ofrecer buenos experimentos. En un diseño de RD todas las unidades tienen un puntaje y un criterio de asignación a un tratamiento según un score de corte. Los controles son aquellos cuyos puntajes están bajo el puntaje de corte.

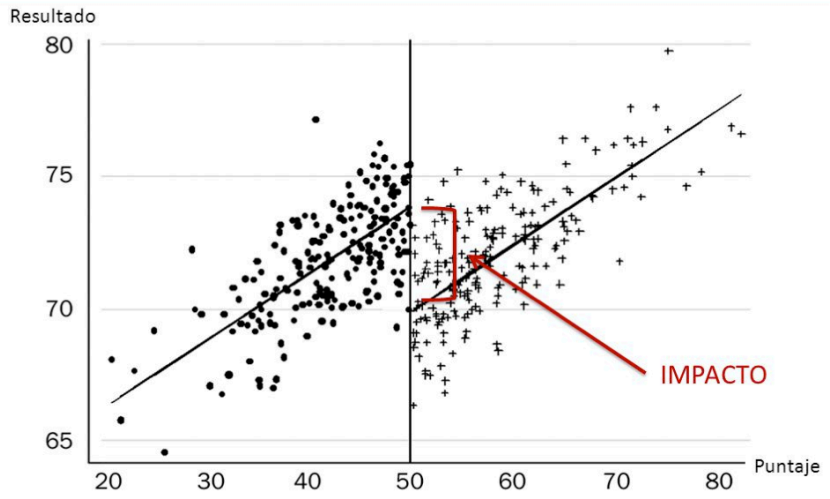


La probabilidad de recibir el tratamiento cambia abruptamente en el puntaje de corte. Este cambio abrupto se llama discontinuidad. La RD se compone de tres elementos fundamentales:

1. Un puntaje (running variable)
2. Puntaje de corte (cutoff)
3. Un tratamiento

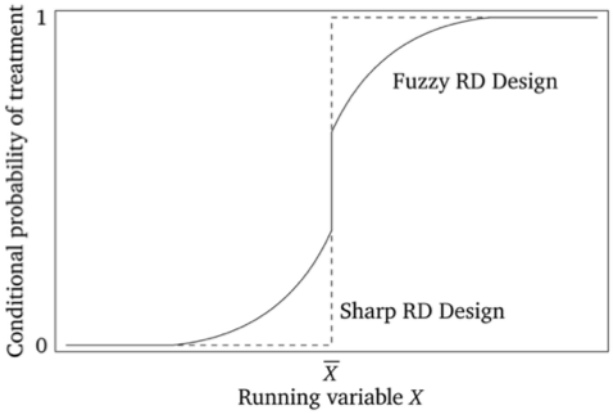
Tipos RD

Hay mucho contextos en que se utiliza algún tipo de regla para asignar un programa. Cuando existen discontinuidades en la asignación podemos implementar un RD **¿Se les ocurre alguno?**



$$\tilde{\tau}_{RD}(\tilde{Z}) = (\lim_{Z \to Z_0^-} \mathbf{E}\left[Y_i \,|Z_i = Z\right]) \; - (\lim_{Z \to Z_0^+} \mathbf{E}\left[Y_i \,|Z_i = Z\right])$$

La probabilidad de ser tratado es una función de cambios discontinuos de los puntajes de las observaciones en el puntaje de corte. Podemos explotar este aspecto de RD con dos tipos de diseños.



- **Sharp RD**: El puntaje determina completamente la asignación a los grupos de tratamiento y de control.

RD lo interpretamos como el efecto del tratamiento cerca del punto de corte, por esta razón es un LATE:

$$\delta_{SRD} = E[Y_i^1 - Y_i^0 | X_i = C_0]$$

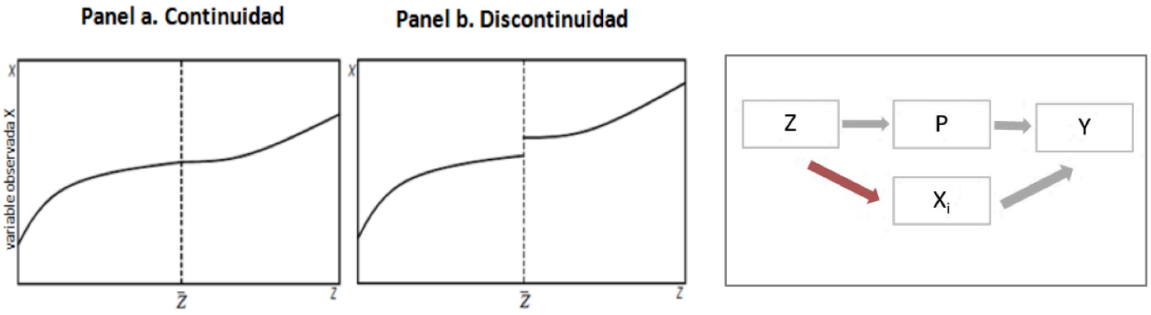
- **Fuzzy RD**: El puntaje no determina completamente la asignación del tratamiento (uso de variables instrumentales aplicando sus mismos supuestos).

Supuestos de la RD

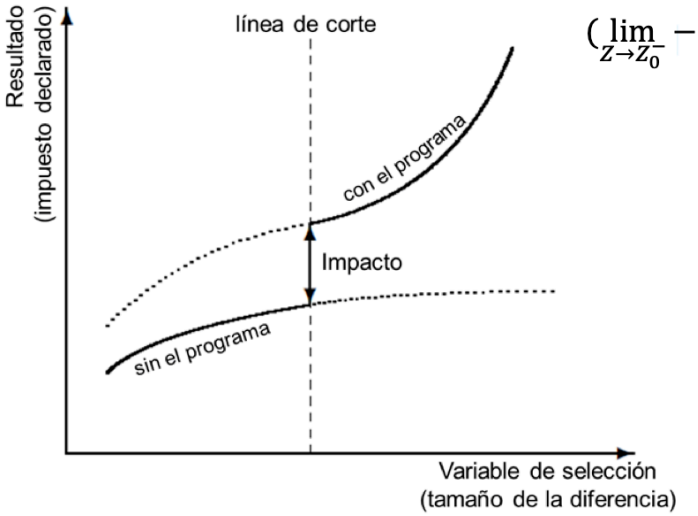
Para que δ_{SRD} sea la estimación insesgada del efecto debemos considerar los siguientes supuestos:

Continuidad: $E[Y_i^0|X_i = C_0]$ y $ E[Y_{i}^1 | X_i=C_0]$ son funciones continuas de X. En otras palabras, en ausencia del tratamiento, los resultados potenciales no habrían saltado en c_0. Este es un supuesto difícil de testear, **¿en que casos cree que este supuesto puede ser vulnerado?**

Variables observadas no deben ser función de la asignación (Zi)

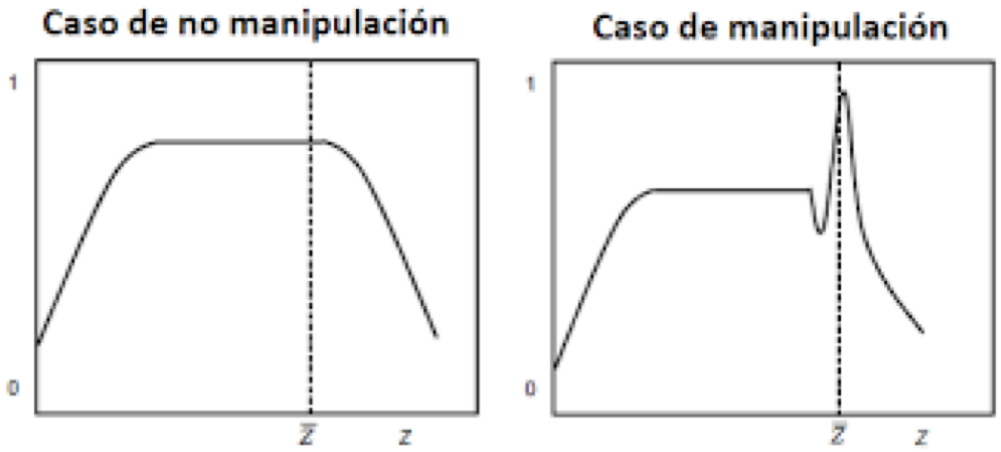


Discontinuidad existe: debemos observar un salto en torno a la discontinuidad para evaluar un efecto. Esto se puede obtener gráficamente.



$$(\lim_{Z \to Z_0^-} - \boldsymbol{P(Z = z)} \neq (\lim_{Z \to Z_0^+} + \boldsymbol{P(Z = z)}$$

Tratamiento no manipulable: no hay manipulación en torno al instrumento de selección.

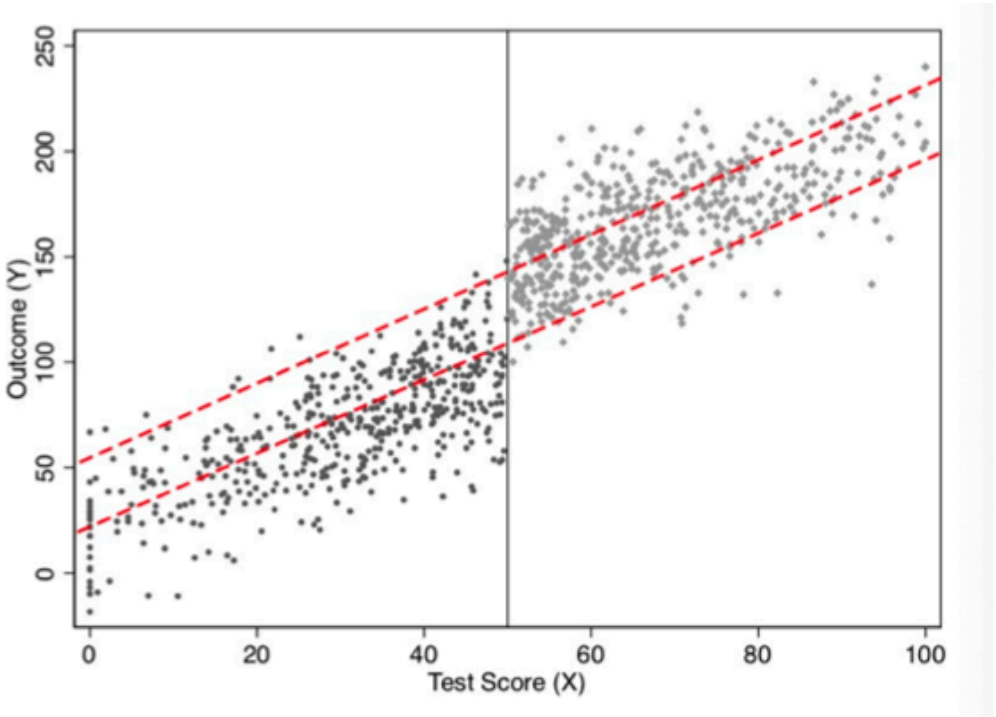


Estimación vía OLS (RD Sharp)

Típicamente se estandariza la running variable $(X_i - c_0)$, luego podemos ajustar la siguiente regresión:

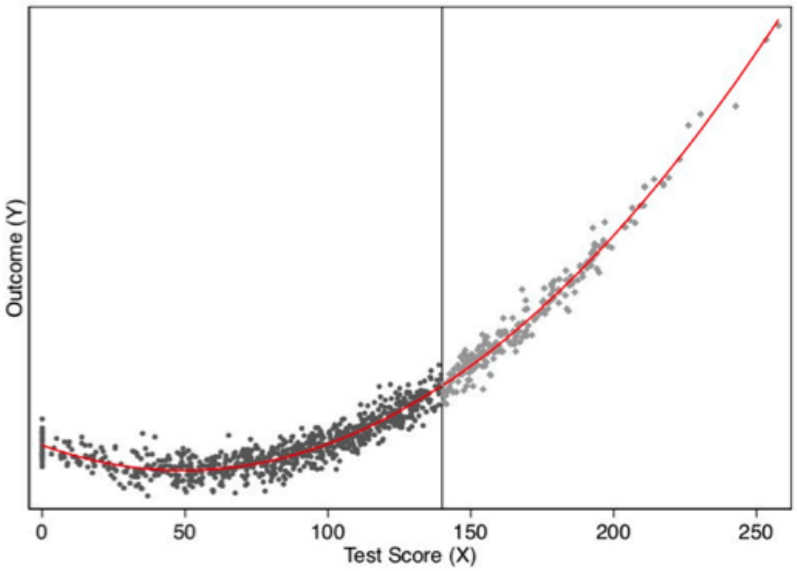
$$Y_i = \alpha + \beta_0(X_i - c_0) + \delta D_i + \beta_1 Di(X_i - c_0) + \epsilon_i$$

¡Ojo! Que aplica esta regresión asume que X_i e Y_i se relacionan de forma lineal, lo que no necesariamente es así.

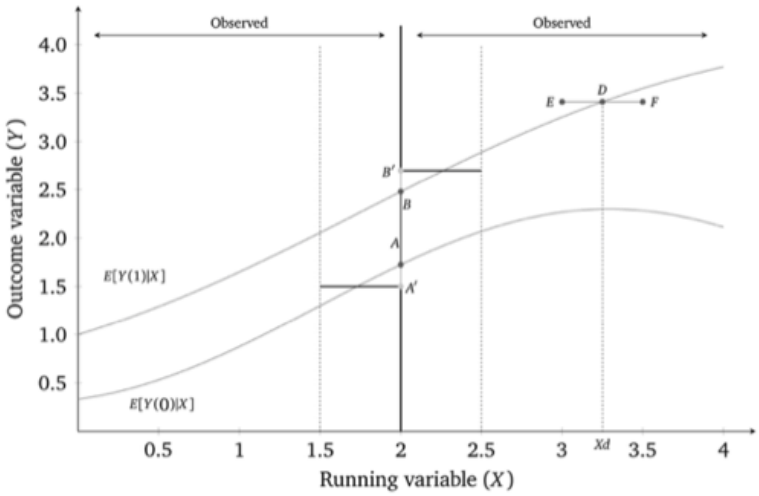


En escenarios de no linealidad podemos usar:

- Polinomios:



- Local linear nonparametric regression y uso de kernels: Esto es básicamente un regresión ponderada restringida a una "ventana". Un kernel triangular le da más peso a las observaciones que están cerca del cutoff.



Para estimar una RD en la práctica:

- Elegir kernel (uniforme está bien para un RD en que todas las observaciones pesen lo mismo)
- Elegir un bandwidth (para esto existen programas que lo estiman de manera óptima)
- Aplicar la regresión:

$$Y_i = \alpha + \beta_0(X_i - c_0) + \delta D_i + \beta_1 Di(X_i - c_0) + \epsilon_i$$

Material extra: DAGS (Directed Acyclic Graphs)

Los modelos gráficos son inmensamente útiles para diseñar una estrategia de identificación creíble. Para esto debemos considerar que **la causalidad corre en una dirección**. Específicamente, avanza en el tiempo. No hay ciclos en un DAG. A su vez, los DAG **explican la causalidad en términos de contrafactuales**. Es decir, un efecto causal se define como una comparación entre dos estados del mundo: un estado que realmente sucedió cuando alguna intervención adquirió algún valor y otro estado que no sucedió (el "contrafactual") bajo alguna otra intervención.

Los efectos causales pueden ocurrir de dos formas:

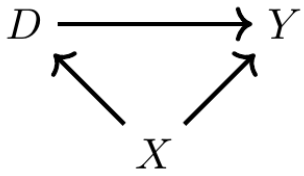
Directos: $D \to Y$

Mediados por una tercera variable: $D \to X \to Y$

A pesar de que los DAGS son fundamentales para comprender el papel que desempeña el conocimiento previo en la identificación de los efectos causales, también nos sirven para:

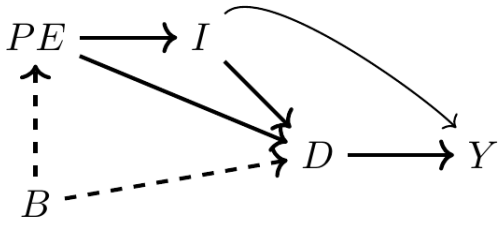
- Para comunicar diseños de investigación y estimadores
- Ayudar a desarrollar un diseño de investigación creíble para identificar los efectos causales de alguna intervención cuando tenemos sesgos por variables omitidas.

¿Qué representa el siguiente DAG?



Veamos un ejemplo aplicado: Una pregunta clásica en economía laboral es si la educación universitaria aumenta los ingresos. Según el modelo de capital humano de Becker (1994), la educación aumenta el producto marginal de uno y, dado que los trabajadores reciben su producto marginal en mercados competitivos, la educación también aumenta sus ingresos. Pero la educación universitaria no es aleatoria; se elige de manera óptima dadas las preferencias subjetivas de un individuo y las limitaciones de recursos.

Tomemos el siguiente modelo:



Donde:

D: Años de educación; Y: Ingresos; PE: Educación de los padres; I: Ingresos familiares; B: Matriz de covariables no observadas como la genética, el entorno familiar y habilidades cognitivas.