Evaluación de impacto:

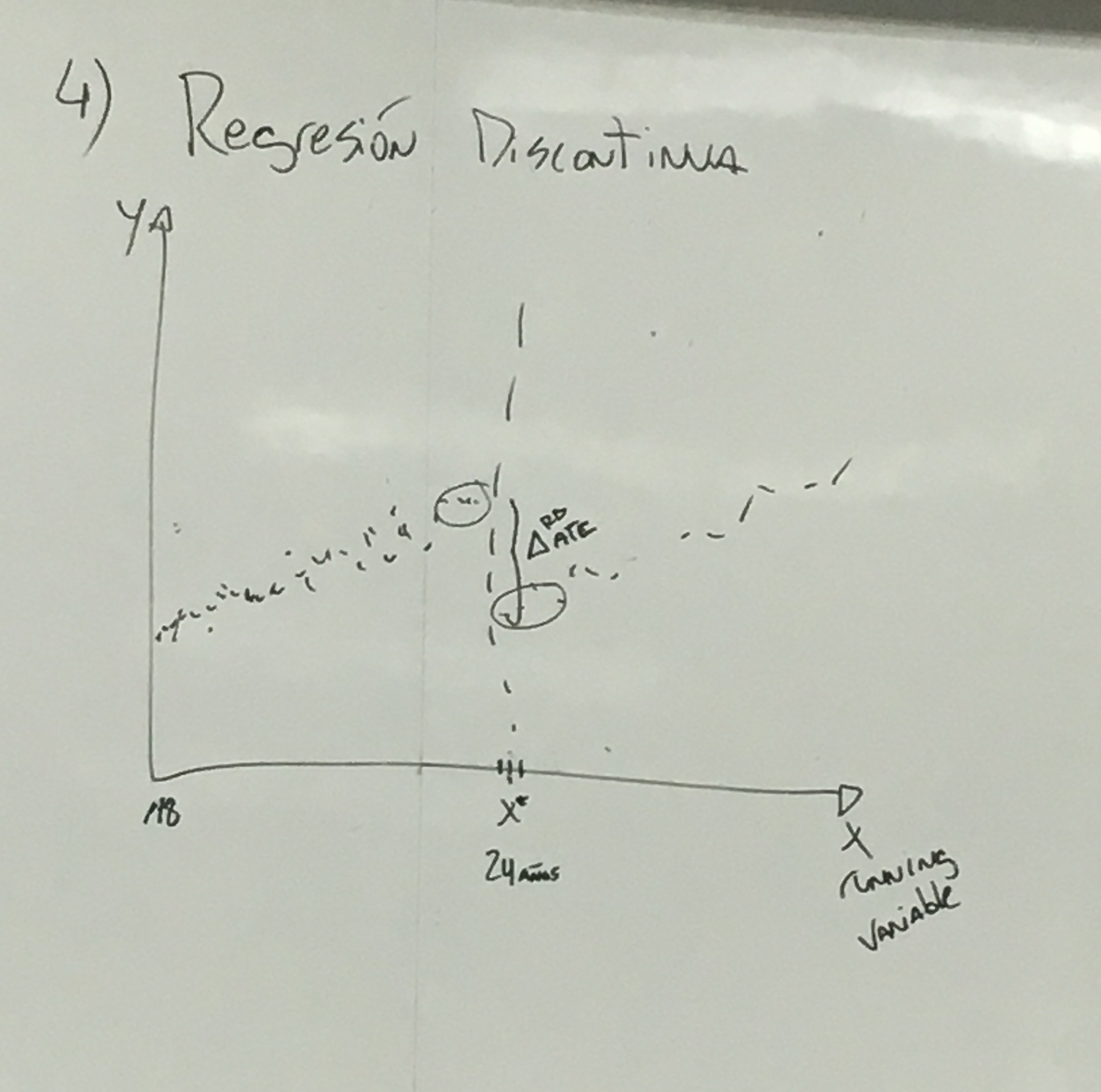
Outcomes o resultados:} ingreso probabilidad de empleo, pruebas estandarizadas.. etc..

Variables, que apuede observar, no son impactos

Impacto es la diferencia en outcomes. No puede ser ingreso, es cuanto aumento el ingreso, si es significativo o no. 🡪 Δi = Y1i – Yoi. No existe la misma persona i tratada o no tratada, ese es el problema. El Yoi 🡪 es el que llamamos contra factual. Y el problema de esto es que el contra factual no existe.

Soluciones:

1. Científica: es observar a la misma persona a lo largo del tiempo
   1. Δi = Y1,t,i – Y1,t-1,i 🡪 el supuesto detrás de esto: Y1,t-1,i ≈ Yo,t,i
   2. La diferencia entre estas dos es que y1 y yo es que y1 es los que ya recibieron la clase de computación y la comparo con la que tenían antes. Y lo otro es imposible porque son los que recibieron y su misma versión si no hubieran recibido.
   3. Problema: es asumir que todo el cambio es por el “1” ósea el programa o política que se dio.
2. Estadística: me aporvecho de la estadística para que el grupo de control sea gente que no fue estudiada, pero se parece mucho a las tratadas
   1. ΔATE: average treatment effect: E(Y1,i) – E(Yo,i)
   2. Δ ATE (naive = ingenuo) = y1(barra) – yo(barra) (los promedios) 🡪 el problema de este estimador naive, es la selección, osea ver que sean comparables. (la forma naive es llegar y agarrar todos y sacar los promedios y esta forma esta mala)
      1. por ejemplo, ir a AAA tienen más reincidencia en que los que no. Problema: los que van a AAA tienen peor nivel. Entonces no es que AAA sea malo. Resumen si hay problema de selección, es naive y esta sesgado.
   3. Solución 1) para que funcione Naive:
      1. Selección aleatoria: la idea de esto es que si tengo que calcular promedios, entre 1 y 0 lo hago aleatoriamente, por ejemplo si regalo cursos de manejo, los que postulen, vamos a elegir al a mitad en un método de selección aleatoria. Entonces no sabes si manejaban bien o mal. Si es selección aleatoria se pude hacer de dos maneras:
         1. Hacer una regresión : Y = Bo + BiD + B2D\*edad🡪 BI ≈ al ATE. Y porqué , porque se pueden agregar más variables. Cuidado con: bien seleccionados, y cumplimientos del programa.
         2. DIf –Dif: si no hubiese existido el tratamiento, hubiese sido la línea punteada. Condición : los otros países son no tratadas. Supuesto los dos mantienen la misma tendencia, y el conrafactual es la línea punteada.
         3. VI: Y = α0 + α1\*D, buscamos variables instrumentales que cumplan dos condiciones:
            1. Cov (Z,ε) = 0
            2. Cov (Z,D) ≠ 0

Por ejemplo que les envíen publicidad del tratamiento AAA, porque esto se correlaciona con D (si fue o no fue al tratamiento) y no correlaciona con el error porque lo seleccione aleatoriamente.

Eso de variables instrumentales se hace en 2 etapas. Primero d = e0 + e1Z, obtengo un D gorro y después corro la regresión, de Y = α0 + α1\*D(gorro)

* + - 1. Regresión Discontinua:
         1. Supuesto que si estoy justo antes o justo después de los 24 es el hecho de haber cumplido 24 y que ya no recibe subsidio. Sólo la diferencia entre los que están en los círculos es comparable, el resto son cualquiera.
      2. Matching: gran ventaja con los otros es que no es regresiones complejas etc.. es una simple diferencia.

comparar cada uno de los verdes con cada uno de los rojos más parecidos, y acá hay 4 métodos:

* + - * 1. Nearest Neigbor: elijo a los que yo quiera a 1 o 2 más cercanos en x
        2. Strata : separo en 5 grupos y cada rojo con los verdes de su grupo
        3. Bandwidth: cada verde manda y tiene su vecindad, entonces no es divisiones en X sino que según la vecindad del verde.
        4. Propensity Score: usa muchas x`s y les da un puntaje

