Fundamentos Causalidad II

Alvaro J. Riascos Villegas

Julio de 2020

Contenido

Ejemplo: Helados y Crimen

- Considere la representación gráfica:

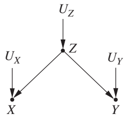


Figure 3.1 A graphical model representing the relationship between temperature (Z), ice cream sales (X), and crime rates (Y)

- Este gráfico refleja la relación probabilística entre las variables aleatoria.
- Condicionar a una variable es observar las demás solo cuando esa esta fija en cierto valor. No cambia el gráfico (i.e. cambiar la perspectiva para observar el mundo).

Ejemplo: Helados y Crimen

- Considere la representación gráfica:

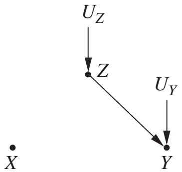


Figure 3.2 A graphical model representing an intervention on the model in Figure 3.1 that lowers ice cream sales

- Intervenir X consiste en fijar su valor independiente de todo lo que pueda afectarla. Cambia el gráfico (cambia el mundo).
- De este nuevo gráfico se deduce que intervenir X no tiene ningún efecto sobre Y.

Condicionar vs. Intervenir

- Condicionar es restringirse al estudio de los ejemplos o la parte de la población que tiene ciertas características.
- Intervenir es cambiar un valor de alguna variable para todo los ejemplos o población (i.e., una manipulación de la dependencia entre las variables).
- Para representar una condicional y diferenciarla de una intervención usamos la siguiente notación: $P(Y \mid X = x)$ y $P(Y \mid do(X = x))$.
- Suponemos que $P(Y \mid do(X = x)) = P_m(Y \mid X = x)$ donde P_m es la distribución de probababilidad del modelo gráfico manipulado.
- En muchas ocasiones (cuando X es binaria) estamos interesados en el efecto casual promedio:

$$P(Y = y \mid do(X = 1)) - P(Y = y \mid do(X = 0)) \qquad (1)$$

- Consideremos el caso de la Paradoja de Simpson (i.e., primera versión).

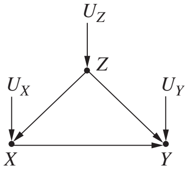


Figure 3.3 A graphical model representing the effects of a new drug, with Z representing gender, X standing for drug usage, and Y standing for recovery

- Si intervenimos $X = x$ se obtiene el nuevo diagrama.

Cálculo de Efectos Causales

- Intervención $X = x$.

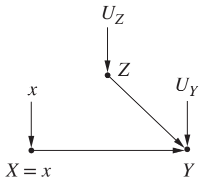


Figure 3.4 A modified graphical model representing an intervention on the model in Figure 3.3 that sets drug usage in the population, and results in the manipulated probability P_m

- Si suponemos que no existen efectos secundarios de la intervención sobre otras variables, $P_m(Y = y \mid Z = z, X = x) = P(Y = y \mid Z = z, X = x)$ y $P_m(Z = z) = P(Z = z)$ entonces podemos calcular el efecto causal $P(Y = y \mid do(X = x))$.

Cálculo de Efectos Causales

- $P(Y = y \mid do(X = x)) = P_m(Y = y \mid X = x)$ por definición.

$$
\begin{aligned}
P_m(Y = y \mid X = x) = & \qquad (2)\\
\sum_z P_m(Y = y \mid X = x, Z = z) P_m(Z = z \mid X = x) & \qquad (3)\\
= \sum_z P_m(Y = y \mid X = x, Z = z) P_m(Z = z) & \qquad (4)\\
= \sum_z P(Y = y \mid X = x, Z = z) P(Z = z) & \qquad (5)
\end{aligned}
$$

- Esto se conoce como el ajuste por Z.

Example (Paradoja de Simpson I)

Sea $X = 1$ tomar la droga, $Y = 1$ recuperarse y $Z = 1$ ser hombre. Utiliando la tabla de frecuencias observadas:

$$P(Y = 1 \mid do(X = 1)) = 0{,}832 \tag{6}$$
$$P(Y = 1 \mid do(X = 0)) = 0{,}7818 \tag{7}$$

Luego $P(Y = 1 \mid do(X = 1)) - P(Y = 1 \mid do(X = 0)) = 0{,}0502$. Esto lo podemos intepretar como la diferencia en la fracción de las personas que se recuperan si todos toman la droga menos la fracción de los que se recuperan si nadie toma la droga.

- Obsérvese que de haberse conducido un experimento aleatorio controlado para conocer el efecto de la droga, el diagrama resultante sería como el diagrama intervenido.

Example (Paradoja de Simpson II)

En este caso una intervención no cambia el grafo (i.e., el grafo tendría la misma forma que si se hubiera hecho un experimento aleatorio controlado).

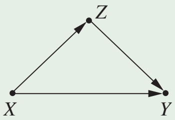


Figure 3.5 A graphical model representing the effects of a new drug, with X representing drug usage, Y representing recovery, and Z representing blood pressure (measured at the end of the study). Exogenous variables are not shown in the graph, implying that they are mutually independent

Como el grafo no cambia: $P(Y = y \mid do(X = x)) = P_m(Y = y \mid X = x) = P(Y = y \mid X = x)$, lo cual explica que se use las frecuencias codicionales (el efecto agregado observado).

Regla de Efectos Causales

- Sea G un grafo y PA_X los parientes de la variable X. Entonces el efecto causal de X en Y es:

$$P(Y = y \mid do(X = x)) = \tag{8}$$

$$\sum_z P(Y = y \mid X = x, PA_X = z) P(PA_X = z) \tag{9}$$

- Alternativamente:

$$P(Y = y \mid do(X = x)) = \tag{10}$$

$$\sum_z \frac{P(Y = y, X = x, PA_X = z)}{P(X = x \mid PA_X = z)} \tag{11}$$

$P(X = x \mid PA_X = z)$ se conoce como el *propensity score*.

- En resumen, el modelo grafico permite identificar los parientes de X que bajo condiciones no experimentales permitirían identificar el valor de X o su distribución de probabilidad.

Contenido

Representación Potencialmente Insuficiente: Variables Ocultas

- Cuándo la estructura de un grafo es suficiente para determinar un efecto causal solo de los datos observados?
- Criterio de la Puerta Trasera: G un grafo (dirigido acíclico). Decimos que un conjunto de variables Z satisface el criterio de la puerta trasera relativo a las variables X, Y si:
  1. Ningun nodo Z es un descendiente de X.
  2. Z bloquea cualquier camino entre X y Y que tenga un enlace que apunta hacia X.
- En ese caso el efecto causal se calcula como:

$$P(Y = y \mid do(X = x)) = \sum_{z} P(Y = y \mid X = x, Z = z) P(Z = z) \tag{12}$$

Representación Potencalmente Insuficiente: Variables Ocultas

- Obsérvese que PA_X siempre satisface el criterio de la puerta trasera relativo a X y Y.

Backdoor Criterion Intuitivamente

1. No hay relaciones espureas entre X y Y.
2. Se mantiene todos los caminos dirigidos entre X y Y sin modificación.
3. No se crean nuevos caminos espurios.

Backdoor Criterion: Ejemplo

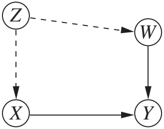


Figure 3.6 A graphical model representing the relationship between a new drug (X), recovery (Y), weight (W), and an unmeasured variable Z (socioeconomic status)

- La estrategia es buscar una variable(s) que satisfaga el criterio de BC relativo a X, Y.
- W no es un descendiente de X y bloque el camino (backdoor path) $X \leftarrow Z \rightarrow W \rightarrow Y$.
- Luego, si el modelo es una descripcón correcta de las relaciones aun si Z no se observa se puede calcular el efecto causal de X en Y ajustando por W.

Contenido

Importancia de Variables

- Existen muchas formas de estudiar el papel de las variables explicativas en modelos de aprendizaje de máquinas.
- Estas dependen del objetio del análisis. Cuando el énfasis es en predicción existen varios conceptos de importacia de las variables: Análisis de varianza, importancia de variables (i.e., árboles), efectos marginales, esperanza condicional, dependencia parcial, etc.

Importancia de Variables

- Cuando el objetivo es analizar el efecto causal el problema es más difícil y para esto hemos estudiado toda esta formalización en grafos del concepto de causalidad.
- Vamos a mostrar que una forma de estimar la importancia de variables en modelos arbirarios (cajas negras) de aprendizaje de máquinas, los gráficos de dependencia parcial de Friedman, pueden tambien bajo ciertas circunstancias utilizarse para estimar efectos causales.

Gráficos de Dependencia Parcial

- Sea X_S y $Z = X_C$ un conjunto de variables y su complemento respectivamente. Sea h una función de aprendizaje: $Y = h(X_S, Z)$. Definimos la dependencia parcial $h_{X_S}(x_S)$, de Y en X_S en el punto x_S como:

$$E_Z[h(x_S, Z)] = \int h(x_S, Z) dF(z)$$

- Obsérvese que aquí se está integrando sobre la distribución marginal de Z.
- Esto es diferente a la esperanza condicional: $E[h(X_S, Z) \mid X_S = x]$, en donde se integra con respecto a la distribuón condicional de Z a $X = x$.
- Un estimador de la dependencia parcial es:

$$E_Z[h(x_S, Z)] = \frac{1}{n} \sum_{i=1}^{n} h(x_S, z_i) \tag{13}$$

Gráficos de Dependencia Parcial

- PDP permite estimar la relación de causalidad de X_S en Y controlando por Z.
- Para ver esto recordemos la ecuación de ajuste:

$$F(y \mid do(X_S = x_S)) = \int F(y \mid X_S = x_S, Z = z) dF(z)$$

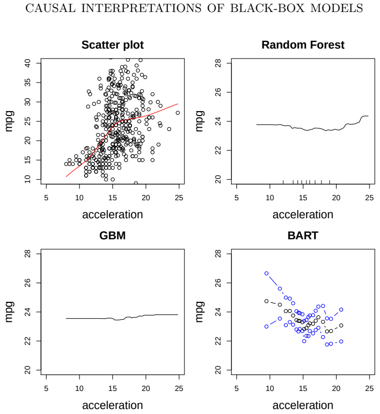
- Si calculamos el valor esperado en ambos lados de la variable aleatoria $F(Y \mid do(X_S = x_S))$:

$$E[Y \mid do(X_S = x_S)] = \int E[Y \mid X_S = x_S, Z = z] dF(z) \quad (14)$$

- Obsérvese que la función de aprendizaje óptimo para problemas de regresión es: $h(x_S, z) = E[Y \mid X_S = x_S, Z = z]$

Gráficos de Dependencia Parcial y Causalidad

- Si Z satisface la propiedad de la puerta trasera con respecto a X_S, Y, entonces la dependencia parcial Y en X_S estima el efecto causal de X_S en Y.



(A) Scatter plot and partial dependence plots using different black-box algorithms.