

# Bayesian causality

Simone Gervasoni Maria

10 marzo 2024

## Indice

<b>1</b>	<b>La storia della causalità</b>	<b>3</b>
<b>2</b>	<b>Potential Outcome model</b>	<b>3</b>
2.1	Studi randomizzati: Ruolo della randomizzazione . . . . .	5
2.2	Studi osservazionali . . . . .	6

# 1 La storia della causalità

Il problema della causalità è di fondamentale interesse per capire meglio il mondo che ci circonda, perciò la domanda è stata affrontata da filosofi e scienziati. La sua definizione rigorosa nel campo statistico è stata solo affrontata molto recentemente, [proprio perché estremamente vago], da Neyman (1932) e poi sviluppata da Rubin negli anni '70, i quali hanno fatto riferimento al modello dei potential outcome spiegato nel capitolo 2. Questo modello diverge concettualmente da alcune definizioni portate avanti da filosofi come John Stuart Mill che definisce la causalità come “the antecedent, or the concurrence of antecedents, on which [a given phenomenon] is invariably and unconditionally consequent”, per Mill dunque possiamo dire che A causa B se e solo se ogni volta che succede A succede anche B . Questo modello di causalità è molto riduttivo e ignora la aleatorietà degli eventi, per questo dobbiamo introdurre il potential outcome model.

## 2 Potential Outcome model

Il potential outcome model definisce l'effetto causale di un evento A come la differenza tra i due "stati" del mondo, cioè il mondo dove accade A e il mondo dove non accade A.

Poniamo ad esempio di voler capire se veramente un medicinale possa migliorare il mal di testa. Formalizziamo il problema ponendo X come l'insieme di covariate dei pazienti, D il regime di trattamento (che assume il valore 1 se viene dato il medicinale mentre assume valore 0 quando al paziente viene somministrato il placebo) e  $Y_i^{obs}$  come il numero di minuti per cui persiste il mal di testa. Dunque se potessimo conoscere contemporaneamente  $Y_i^{obs}|D = 1$ , che chiameremo  $Y_i^1$ , e  $Y_i^{obs}|D = 0$ , che chiameremo  $Y_i^0$ , allora calcolare se il medicinale causa un miglioramento risulterà molto semplice. Introduciamo un esempio numerico :

Age	Sex	$Y_i^0$	$Y_i^1$	$\delta_i$
20	M	20	21	-1
20	F	15	3	12
20	M	8	10	-2
20	F	16	15	1
30	M	12	13	-1
30	F	8	5	3
30	M	2	11	-9
30	F	15	26	11

Tabella 1: Tabella esperimento

Vediamo che il  $\delta$  non è sempre positivo dunque la medicina non ha migliorato la situazione per tutti, notiamo però che in generale ha ridotto il tempo del

malessere, questo però non basta dobbiamo essere più precisi e quantificare esattamente la tipologia e la dimensione del effetto.

### Definizione parametri

Definiamo quindi alcune quantità che ci risulteranno utili :

- CATE o Conditional Average Treatment Effect è definito come  $E[Y_i^1 - Y_i^0|X] = E[\delta_i|X]$ , quindi  $CATE_{(M,20)} = E[\delta_i|X = (M, 20)] \approx \frac{18-2}{2} = 8$  , possiamo quindi dire che tra i maschi ventenni la medicina causa in media una riduzione di 8 minuti nella durata del mal di testa .
- ATE o average treatment effect come  $E[Y_i^1 - Y_i^0] = E[\delta_i]$ , quindi  $ATE = E[\delta_i] \approx \frac{18+12-2+18+3-9+11}{8}$ .
- ATT o average treatment on the treated  $E[Y_i^1 - Y_i^0|D = 1] = E[\delta_i|D = 1]$
- ATU o average treatment on the untreated  $E[Y_i^1 - Y_i^0|D = 0] = E[\delta_i|D = 0]$

Non possiamo calcolare gli ultimi due perché siamo ancora nel caso ipotetico dove conosciamo i due potential outcomes. Ovviamente questa tabella non potrà mai essere riempita come abbiamo mostrato sopra perché possiamo veramente conoscere una sola quantità tra  $Y_i^0$  e  $Y_i^1$ , sarà quindi impossibile avere certezza sulle quantità definite prima ma bisognerà stimarle. Dunque è utile fare la distinzione tra i valori *fattuali* cioè cosa è veramente successo e *controfattuale* cioè cosa sarebbe accaduto se il regime di trattamento fosse stato diverso. Possiamo capire meglio la relazione tra potential outcome e observed outcome attraverso la "switching equation":

$$Y_i^{obs} = D_i \cdot Y_i^1 + (1 - D_i) \cdot Y_i^0 \quad (1)$$

Vediamo quindi che  $Y_i^{obs} = Y_i^1$  quando  $D_i = 1$ , mentre  $Y_i^{obs} = Y_i^0$  quando  $D_i = 0$ . Bisogna precisare che la differenza tra  $Y_i^{obs}$  e  $Y_i^0, Y_i^1$  è che i primi sono valori effettivi, empirici che si possono osservare mentre gli ultimi sono valori ex-ante e dunque esistono prima che la  $D_i$  venga somministrata.

Quello che invece ci ritroveremo nella realtà sarà una tabella più simile a questa

Age	Sex	$Y_i^0$	$Y_i^1$
20	M	20	?
20	F	15	?
20	M	?	10
20	F	?	15
30	M	12	?
30	F	8	?
30	M	?	11
30	F	?	26

Tabella 2: Tabella esperimento

Abbiamo dunque ridotto il problema della causalità ad un problema di *Missing Data*, questa è la forza del modello usato

## 2.1 Studi randomizzati: Ruolo della randomizzazione

Spesso si sente parlare di studi clinici randomizzati controllati, cioè studi in cui a metà dei partecipanti viene dato un placebo e all'altra metà il medicinale che è oggetto di studio, ovviamente né i medici né i pazienti sono a conoscenza di quali persone hanno assunto una o l'altra cosa. La situazione può essere espressa con un DAG in questa maniera:

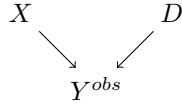


Figura 1: Dag per studio randomizzato

Ma perché questo procedimento è ormai la norma per verificare l'efficacia di un farmaco? E cosa ci può dire veramente un esperimento del genere? Sotto il potential outcome model questo procedimento può essere formalizzato tramite questa indipendenza

$$D \perp\!\!\!\perp (Y^0, Y^1) \quad (2)$$

ciò significa che il trattamento non è amministrato in base agli potential outcome, ma a caso. Non bisogna però confondersi e pensare che ciò implichi che  $D \perp\!\!\!\perp Y^{obs}$ , infatti ciò è espressamente falso ogni volta che il medicinale ha effetto su infatti noi speriamo che il nostro trattamento abbia un effetto su  $Y^{obs}$ . Grazie alla equazione 2 abbiamo che  $E[Y_i^1] = E[Y_i^1|D_i = 1]$  e la stessa cosa per  $E[Y_i^0] = E[Y_i^0|D_i = 0]$ . Possiamo quindi dire che :

$$ATE = E[Y_i^1 - Y_i^0] \quad (3)$$

$$= E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 0] \quad (4)$$

$$= E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 1] \quad (5)$$

$$= E[Y_i^1|D_i = 0] - E[Y_i^0|D_i = 0] \quad (6)$$

Nell'equazione 4 entrambi gli addendi sono quantità *fattuali*, grazie alla legge dei grandi numeri stimiamo queste quantità con delle semplici medie, chiameremo quindi SDO la semplice differenza delle media osservate per i gruppi:

$$SDO = \frac{1}{N_1} \sum_i (y_i|d_i = 1) - \frac{1}{N_2} \sum_i (y_i|d_i = 0) \stackrel{n \rightarrow \infty}{\rightleftharpoons} ATE$$

Dalla equazione 5 e 6 possiamo concludere che nel caso randomizzato avremo che  $ATE = ATT = ATU$ .

## 2.2 Studi osservazionali

Gli studi osservazionali (?) sono invece strutturati in modo diverso, i ricercatori osservano e raccolgono dati su individui o gruppi senza intervenire o manipolare alcun aspetto dell'ambiente di studio, per 2 motivi principali

1. Potrebbe essere non etico o pratico, dividere la popolazione in maniera casuale (immaginiamo di volere studiare se fumare causa un aumento nel rischio del cancro)
2. Gli eventi sono già accaduti (immaginiamo studio su come i tassi di interesse influenzino il consumo)

Dunque la cosa principale che cambia dal caso precedente è che  $D \not\perp (Y^0, Y^1)$ , possiamo dunque rappresentare la situazione con il seguente DAG:

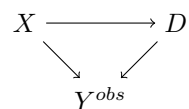


Figura 2: Dag per studio osservazionale

Una possibile soluzione è stata proposta da Rubin con il propensity score matching .