

L'approccio dei risultati potenziali all'inferenza causale

Leonardo Grilli & Fabrizia Mealli

Outline

- Concetti di base
- Tipi di inferenza causale
- Esperimenti con noncompliance: approccio dei risultati potenziali e confronto con il metodo delle variabili strumentali e con le equazioni strutturali

Concetti di base

- Risultati potenziali corrispondenti ai livelli di trattamento
- Relazioni fra le unità statistiche
- Meccanismo di assegnazione al trattamento

Concetti di base

- **Unità i** : oggetto/persona o collezione di oggetti/persono *in un certo istante temporale*
 - Stesso oggetto in due istanti temporali \Rightarrow due unità distinte
- **Trattamento z_i** : *azione* applicata ad una unità \Rightarrow difficoltà di includere alcuni concetti di causalità e.g. "he was fired because he was too old"

"No causation without manipulation"

Concetti di base

- Due o più tipi di trattamento
- (unità i , trattamento z_i) \rightarrow risultato potenziale $Y_i(z_i)$
- **Risultato potenziale**: variabile post-trattamento riferita ad un certo istante o intervallo temporale
- **Effetto causale per la i -ma unità**: confronto (tramite differenze, rapporti o altro) fra i risultati potenziali della i -ma unità
es. con due trattamenti $Y_i(1)-Y_i(0)$

Nozione di effetto causale

La definizione di **effetto causale** non dipende da quale risultato viene osservato, quindi non dipende dal meccanismo di assegnazione al trattamento (**disegno**)



Separazione fra **scienza** e **disegno**
(differenza rispetto ai modelli grafici)

Inferenza causale

- **Problema fondamentale dell'inferenza causale:** ogni unità viene assegnata ad un solo trattamento*, per cui
 - il risultato associato al trattamento assegnato viene osservato
 - i risultati associati agli altri trattamenti non sono osservati (dati mancanti)

* Nota bene: l'unità è definita con riferimento ad un certo istante temporale per cui due trattamenti non sono possibili

Controfattualità

- *Prima dell'assegnazione al trattamento* tutti i risultati sono potenzialmente osservabili (purché il meccanismo di assegnazione sia probabilistico, cioè la probabilità di assegnazione sia strettamente positiva)
- *Dopo l'assegnazione al trattamento* il corrispondente risultato è osservato mentre gli altri divengono **controfattuali**

Relazioni fra unità

- Per definizione l'effetto causale relativo ad una singola unità non è indentificabile \Rightarrow servono più unità
- Due possibilità fondamentali:
 - confrontare stesso oggetto in due istanti temporali (analisi longitudinali o panel)
 - confrontare diversi oggetti nello stesso istante temporale

In ogni caso è necessaria qualche assunzione di stabilità

Relazioni fra unità

- Con n unità, il risultato dell' i -ma unità dipende da z vettore n -dimensionale delle assegnazioni al trattamento: $Y_i(z)$
- Servono dei vincoli di esclusione per limitare i casi possibili
- I vincoli di esclusione sono basati sulla conoscenza a priori del fenomeno

SUTVA

- **No interference:** $Y_i(z) = Y_i(z')$ per ogni z e z' che assegnano l' i -ma unità allo stesso trattamento, per cui $Y_i(z) = Y_i(z_i)$

Problematica in caso di: malattie contagiose, effetti di sostituzione nei mercati (spiazzamento)

- **No versions of treatment:** per evitare la proliferazione dei livelli di trattamento
Problematica in caso di partial compliance

SUTVA

- Se SUTVA non è plausibile
 - Ridefinire l'unità di analisi \Rightarrow adottare la prima unità di livello superiore per la quale la SUTVA è plausibile
 - Ridefinire il trattamento
 - Costruire un modello per le interazioni
 - Distinguere alcune versioni del trattamento

Quantità di interesse (causal estimands)

Data la SUTVA, si definiscono delle quantità di interesse a livello aggregato

- Media della differenza tra i risultati potenziali:

$$E_p [Y_i(1) - Y_i(0)] = E_p [Y_i(1)] - E_p [Y_i(0)]$$

$$P = \begin{cases} \text{tutta la popolazione} \Rightarrow \text{ATE} \\ \text{un suo sottoinsieme, es. } P = \{i : Z_i = 1\} \Rightarrow \text{ATT} \end{cases}$$

La popolazione P può essere finita o infinita (superpopolazione)

Quantità di interesse (causal estimands)

- La media della differenza (o in generale di una combinazione lineare) tra i risultati potenziali dipende solo dalle **distribuzioni marginali** di $Y(0)$ e $Y(1)$

$$E_p [Y_i(1) - Y_i(0)] = E_p [Y_i(1)] - E_p [Y_i(0)]$$

- In particolare questo estimand non dipende da $\rho = \text{corr}(Y(0), Y(1))$, che è **non identificabile**
- Tuttavia se P è una **popolazione finita** la distribuzione campionaria dello *stimatore* dipende da ρ (cfr. approccio di Neyman)

Quantità di interesse (causal estimands)

- Medie di funzioni non lineari dei risultati potenziali, ad es.
 - ARR $E_p [Y_i(1)/Y_i(0)] \neq E_p [Y_i(1)] / E_p [Y_i(0)]$
 - Proporzione di beneficiati dal trattamento

$$E_p [I \{Y_i(1) > Y_i(0)\}]$$

Questi estimand dipendono dalla **distribuzione congiunta** di $Y(0)$ e $Y(1)$ e quindi da ρ

\Rightarrow *analisi di sensitività oppure bounds*

Identificabilità e stima

- L'assegnazione al trattamento non modifica i risultati potenziali $Y_i(0)$ e $Y_i(1)$ e quindi l'effetto causale; tuttavia determina il risultato osservato

$$Y_i^{obs} \equiv Y_i(Z_i) = Y_i(1) \cdot Z_i + Y_i(0) \cdot (1 - Z_i)$$

- In generale

$$f(Y_i^{obs} | Z_i = 1) = f(Y_i(1) | Z_i = 1) \neq f(Y_i(1))$$

$$f(Y_i^{obs} | Z_i = 0) = f(Y_i(0) | Z_i = 0) \neq f(Y_i(0))$$

Identificabilità e stima

Pertanto l'ATE (causal estimand)

$$E(Y_i(1)) - E(Y_i(0))$$

è diverso dalla differenza osservata fra trattati e non trattati (che è un parametro di associazione)

$$E(Y_i^{obs} | Z_i = 1) - E(Y_i^{obs} | Z_i = 0)$$

salvo assumere che il trattamento sia **unconfounded**, cioè $Z_i \perp Y_i(0), Y_i(1)$

Identificabilità e stima

- L'assunzione di **unconfoundedness**

$$Z_i \perp Y_i(0), Y_i(1)$$

non è testabile. In quali casi è plausibile?

- E' plausibile negli **esperimenti randomizzati** poiché Z_i è indipendente da tutte le variabili pre-trattamento, inclusi i risultati potenziali
- Negli **studi osservazionali** occorre affidarsi alle covariate, assumendo che l'**unconfoundedness** valga condizionatamente alle covariate osservate

Tipi di inferenza causale

Meccanismo di assegnazione al trattamento	Natura degli outcome $Y_i(z_i)$	
	Quantità fisse	Variabili aleatorie (modello)
Noto (esperim. randomizzati classici)	Fisher Neyman	Model-based analysis (likelihood or Bayesian)
Ignoto ma ignorabile	Propensity score Matching	
Ignoto e non ignorabile	(inusuale)	

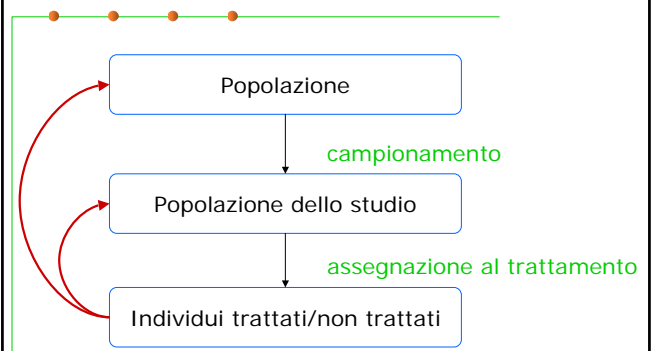
Esperimenti randomizzati

- Approccio di Fisher
 - Unica fonte di variabilità: distribuzione di randomizzazione
 - Interesse: Sharp null hypothesis (tutti gli effetti causali sono nulli)
 - Intervallo di confidenza assumendo effetto del trattamento costante fra individui

Esperimenti randomizzati

- Approccio di Neyman
 - Principale fonte di variabilità: distribuzione di randomizzazione
 - Possibile estensione a superpopolazione
 - Interesse: ATE
 - Stimatore per ATE e sue proprietà

Fonti di variabilità e tipi di inferenza



Analisi model-based

- Generalizzazione / Estensione ad una superpopolazione
- Smoothing (importante quando vi sono molte covariate)
- Due approcci principali:
 - Verosimiglianza \Rightarrow marginalizzazione
 - Bayesiano \Rightarrow imputazione

Impostazione di Maldonado e Greenland (2002)

- Unità \Rightarrow **Target**: popolazione obiettivo durante l'intervallo temporale eziologico (outcome definito rispetto ad un intervallo)
- Trattamento \Rightarrow **Exposure distribution**: distribuzione dell'esposizione nella popolazione (non solo quanti elementi esposti, ma anche quali)
- Effetto causale \Rightarrow **Causal contrast**: confronto fra caratteristiche della popolazione (es. rischio relativo) sotto diverse distribuzioni di esposizione

Esperimenti con noncompliance

- Randomizzazione implica che Z_i (*assegnazione al trattamento*) è indipendente da tutte le variabili **pre-assegnazione** (fra cui i risultati potenziali) e quindi si può assumere unconfoundedness
- Tuttavia, il *trattamento effettivo* D_i è una variabile **post-assegnazione**: il condizionamento a D_i non è innocuo perché può far saltare l'unconfoundedness

Esperimenti con noncompliance

- In ordine temporale $Z_i \rightarrow D_i \rightarrow Y_i$
- Z_i (*assegnazione al trattamento*)
- D_i (*trattamento effettivo*)
 - Risultati potenziali $D_i(0)$ e $D_i(1)$
 - Risultato osservato $D_i^{obs} = D_i(Z_i)$
- Y_i (*outcome*)
 - Risultati potenziali $Y_i(z, d)$, cioè $Y_i(0,0)$, $Y_i(0,1)$, $Y_i(1,0)$, $Y_i(1,1)$
 - Risultato osservato $Y_i^{obs} = Y_i(Z_i, D_i(Z_i))$

Interessa stimare l'effetto causale di D su Y

Esperimenti con noncompliance

- Per una data unità in un certo esperimento tutti gli effetti causali di Z su altre variabili post-assegnazione sono definiti a priori in termini di risultati potenziali
 - Z su D: $D_i(1) - D_i(0)$
 - Z su Y: $Y_i(1, D_i(1)) - Y_i(0, D_i(0))$
- Poiché Z è random entrambi questi effetti sono identificabili

Esperimenti con noncompliance

Se D è binaria ci sono 4 possibili casi (compliance status, ovvero tipo di reazione all'assegnazione al trattamento)

$D_i(1)$	$D_i(0)$	$D_i(1) - D_i(0)$	Tipo	Prob.
1	0	1	Complier (C)	π_C
0	0	0	Never taker (NT)	π_{NT}
1	1	0	Always taker (AT)	π_{AT}
0	1	-1	Defier (D)	π_D

Esperimenti con noncompliance

- La caratterizzazione delle unità come C, NT, AT e D in base alla reazione all'assegnazione al trattamento è
 - non osservabile**, cioè i gruppi C, NT, AT e D sono **gruppi latenti**
 - Relativa al comportamento delle unità nell'**esperimento attuale** e non in assoluto
 - non è influenzata dall'effettiva assegnazione per cui si può considerare una variabile **pre-assegnazione** (questo è utile perché essendo Z random, la composizione della popolazione in termini di compliance status è identica per ogni livello di Z)
- Per gli AT e i NT non c'è alcuna informazione per stimare l'effetto causale di D su Y in quanto Z non riesce a muovere D

Gruppi osservabili e gruppi latenti

Generale Monotonicità Randomized eligibility

Generale			Monotonicità			Randomized eligibility		
D^{obs}	0	1	D^{obs}	0	1	D^{obs}	0	1
Z			Z			Z		
0	NT/C	AT/D	0	NT/C	AT	0	NT/C	.
1	NT/D	AT/C	1	NT	AT/C	1	NT	C

L'appartenenza ai gruppi latenti in generale non è determinabile (salvo per certe combinazioni di Z e D^{obs} a seconda dello schema).

Tuttavia se Z è random e i gruppi latenti non sono più di 3 si possono stimare le probabilità di appartenenza ai gruppi latenti.

Stima del LATE

Assumiamo **SUTVA**

Scomposizione dell'ITT (Intention-to-treat effect)

$$ITT = ITT_C \cdot \pi_C + ITT_{NT} \cdot \pi_{NT} + ITT_{AT} \cdot \pi_{AT} + ITT_D \cdot \pi_D$$

$$ITT_{group} = E[Y_i(1, D_i(1)) - Y_i(0, D_i(0)) | i \in group]$$

$$\pi_{group} = P[i \in group]$$

Stima del LATE

Assumiamo **Vincolo di Esclusione per NT e AT:**

$$ITT_{NT} = ITT_{AT} = 0$$

(riguarda la relazione fra Z e Y)

$$ITT = ITT_C \cdot \pi_C + ITT_D \cdot \pi_D$$

Stima del LATE

Assumiamo **Monotonicità forte:**

1) Effetto medio non nullo di Z su D

$$E[D_i(1) - D_i(0)] = \pi_C - \pi_D \neq 0$$

2) Monotonicità (No defiers): $\pi_D = 0$

OSS: 1 & 2 implicano $\pi_C > 0$

(riguardano la relazione fra Z e D)

$$ITT = ITT_C \cdot \pi_C \quad \text{con } \pi_C > 0$$

Stima del LATE

IV estimand:

$$ITT_C = \frac{ITT}{\pi_C} = \frac{E[Y_i(1, D_i(1)) - Y_i(0, D_i(0))]}{E[D_i(1) - D_i(0)]}$$

OSS: l'ITT per i complier è il rapporto fra due ITT, quello su Y fratto quello su D

Se il trattamento è random o comunque ignorabile ognuno dei due ITT è stimabile come differenza fra le medie dei gruppi di trattati e di non trattati

Stima del LATE

Interpretazione dell'IV estimand:

Assumiamo **Vincolo di Esclusione per i Complier:**

$$Y_i(0, D_i(0)) = Y_i(1, D_i(0))$$

$$Y_i(1, D_i(1)) = Y_i(0, D_i(1))$$

$$ITT_C = E[Y_i(1, D_i(1)) - Y_i(0, D_i(0)) | i \in complier]$$

$$= E[Y_i(1) - Y_i(0) | i \in complier]$$

così ITT per i complier è interpretabile come effetto causale di D su Y

Due diversi estimand

■ Posto che Z sia ignorabile si possono stimare senza distorsione

- ITT (Z su Y)
 - Intera popolazione
 - Richiede solo SUTVA
- LATE (D su Y)
 - Complier (sottoinsieme non osservabile)
 - Richiede SUTVA, Exclusion restriction e Strong Monotonicity

Di solito interessa l'effetto di D su Y per cui il LATE, sebbene più complesso, è più rilevante dell'ITT

Altri estimand (piu interessanti?)

- Il LATE, cioè l'effetto sui Complier, non è certo la quantità più interessante, ma è l'unica stimabile senza ulteriori assunzioni non testabili
- In molte applicazioni sono interessanti
 - ATT $E[Y_i(1) - Y_i(0) | D_i^{obs} = 1]$ Sottoinsieme osservabile
 - ATE $E[Y_i(1) - Y_i(0)]$ Intera popolazione

Per stimare l'ATT o l'ATE servono ulteriori assunzioni non testabili

Stima di ATT e ATE

Ass. I: $E[Y_i(1) - Y_i(0) | Z_i = 1, D_i^{obs} = 1]$ **C o AT**
 $= E[Y_i(1) - Y_i(0) | Z_i = 0, D_i^{obs} = 1]$ **D o AT**

Ass. II: $E[Y_i(1) - Y_i(0) | Z_i = z, D_i^{obs} = 1]$
 $= E[Y_i(1) - Y_i(0) | Z_i = z, D_i^{obs} = 0] \quad \forall z$

Robins (1989): sotto le assunzioni di AIR (1996), eccetto la monotonicità,

Ass. I \Rightarrow IV estimand = ATT

Ass. I & Ass. II \Rightarrow IV estimand = ATE

LIMITI PER ATE

Sotto le assunzioni che consentono la stima del LATE

$$ATE = \pi_C \cdot E[Y_i(1) - Y_i(0) | i \in C] + \pi_{AT} \cdot E[Y_i(1) - Y_i(0) | i \in AT] + \pi_{NT} \cdot E[Y_i(1) - Y_i(0) | i \in NT]$$

$E[Y_i(0) | i \in AT]$ non stimabile
 $E[Y_i(1) | i \in NT]$ non stimabile

- Le probabilità $\pi_C, \pi_{AT}, \pi_{NT}$ sono stimabili
- L'effetto causale è stimabile per C, non per AT e NT
- Facendo variare le due medie non stimabili nell'insieme dei valori ammissibili si ottengono dei limiti non parametrici per ATE

Analisi di sensitività

- Teniamo ferma la SUTVA e l'assunzione di effetto medio non nullo di Z su D
- L'ignorabilità di Z non è rilevante per la definizione degli estimand
- Valutiamo la rimozione alternativa di
 - Exclusion restriction
 - Per i noncomplier
 - Per i complier
 - Monotonicità

Analisi di sensitività

Monotonicità **Si**

Vincolo di esclusione per i noncomplier **No**

$$ITT = ITT_C \cdot \pi_C + ITT_{NT} \cdot \pi_{NT} + ITT_{AT} \cdot \pi_{AT}$$

$$\frac{ITT}{\pi_C} = ITT_C + ITT_{NT} \cdot \frac{\pi_{NT}}{\pi_C} + ITT_{AT} \cdot \frac{\pi_{AT}}{\pi_C}$$

$$= ITT_C + ITT_{AT\&NT} \cdot \frac{\pi_{AT} + \pi_{NT}}{\pi_C}$$

$$ITT_{AT\&NT} = \frac{ITT_{AT} \cdot \pi_{AT} + ITT_{NT} \cdot \pi_{NT}}{\pi_{AT} + \pi_{NT}}$$

Analisi di sensitività

Monotonicità **Si**

Vincolo di esclusione per i complier **No**

Assumendo additività degli effetti, per ogni $i \in C$

$$Y_i(1, D_i(1)) - Y_i(0, D_i(0)) = Y_i(1, 1) - Y_i(0, 0) = [Y_i(1, 1) - Y_i(0, 1)] + [Y_i(0, 1) - Y_i(0, 0)]$$

H_i eff. Causale di Z su Y a parità di D G_i eff. Causale di D su Y a parità di Z

$$ITT_C = E[Y_i(1, D_i(1)) - Y_i(0, D_i(0)) | i \in C] = E[H_i | i \in C] + E[G_i | i \in C]$$

Analisi di sensitività

Monotonicità **No**

Vincolo di esclusione (per tutti, anche defier) **Si**

$$ITT = ITT_C \cdot \pi_C + ITT_D \cdot \pi_D$$

$$\frac{ITT}{\pi_C - \pi_D} = ITT_C + \frac{\pi_D}{\pi_C - \pi_D} [ITT_C - (-ITT_D)]$$

IV
estimand

LATE

Effetto causale di
D su Y per i defier

$$-ITT_D = -E[Y_i(1,0) - Y_i(0,1) | i \in D] \\ = E[Y_i(1) - Y_i(0) | i \in D]$$

Verificabilità delle assunzioni

Vincolo di esclusione

- Per gli AT e i NT: se i gruppi fossero identificabili, il vincolo di esclusione sarebbe testabile confrontando la media di Y per livello di Z
- Per i complier: riguarda un confronto fra risultati non osservabili (controfattuali a priori)
 $Y_i(0, D_i(0)) = Y_i(1, D_i(0)) \quad Y_i(1, D_i(1)) = Y_i(0, D_i(1))$
 si può rendere plausibile tramite accorgimenti quali l'uso di placebo, blinding e double-blinding

- Monotonicità: non verificabile, spesso plausibile, sicuramente vera negli esperimenti in cui gli assegnati al controllo non hanno accesso al trattamento

Formulazione delle assunzioni

- In molte applicazioni l'approccio dei risultati potenziali può risultare preferibile per il modo in cui le assunzioni vengono formulate
 - Separazione fra assunzioni sul fenomeno di interesse e assunzioni sul meccanismo di assegnazione al trattamento
 - Formulazione a livello individuale anziché aggregato (es. interpretazione in termini di modelli di comportamento degli agenti economici basati sulla massimizzazione dell'utilità)

Modelli ad equazioni strutturali

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$$

$$D_i^* = \alpha_0 + \alpha_1 Z_i + v_i$$

$$D_i = I(D_i^* > 0)$$

Dummy endogenous
variable model

β_1 rappresenta l'effetto
causale di D su Y

In questa versione l'effetto è costante per ogni individuo (ma si può generalizzare assumendo che β_1 sia una variabile aleatoria

$$\text{corr}(\varepsilon_i, v_i) \neq 0 \Rightarrow \text{corr}(\varepsilon_i, D_i) \neq 0 \Rightarrow \hat{\beta}_1^{OLS} \text{ distorto}$$

Modelli ad equazioni strutturali

Tipiche assunzioni per la stima IV di β_1 riguardano i termini di disturbo

$$E[Z_i \varepsilon_i] = 0$$

Vincolo di esclusione

$$E[Z_i v_i] = 0$$

Assegnazione random

Non strettamente necessaria per la stima IV

Inoltre si assume che $\text{cov}(D_i, Z_i) \neq 0$

Modelli ad equazioni strutturali

$\alpha_1 > 0 \Rightarrow$ monotonicità (no defiers)

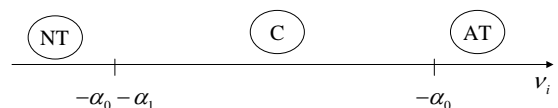
$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$$

$$D_i^* = \alpha_0 + \alpha_1 Z_i + v_i$$

$$D_i = I(D_i^* > 0)$$

$$\{i : i \in AT\} = \{i : v_i > -\alpha_0\}$$

$$\{i : i \in NT\} = \{i : v_i \leq -\alpha_0 - \alpha_1\}$$



Modelli ad equazioni strutturali

Il modello assume lo stesso effetto causale β_1 per tutti gli individui

Tuttavia solo per gli individui

$$\{i : -\alpha_0 - \alpha_1 < v_i \leq -\alpha_0\}$$

sono potenzialmente disponibili informazioni per la stima di β_1

Per gli altri individui l'uso del modello per stimare β_1 comporta una estrapolazione

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$$

$$D_i^* = \alpha_0 + \alpha_1 Z_i + v_i$$

$$D_i = I(D_i^* > 0)$$