

Introduzione dell'approccio delle risposte potenziali alla valutazione

Fulvia Pennoni

Univerisità degli Studi di Milano-Bicocca

Dipartimento di Statistica e

CRISP (Centro di ricerca interuniversitario

per i servizi di pubblica utilità)

Indice

- Introduzione
- Ricerca sperimentale e randomizzazione
- Effetto causale: notazione e terminologia
- Identificazione dell'effetto causale
- Assunzioni in un contesto non sperimentale
- Caratteristiche e vantaggi dell'approccio
- Bibliografia e suggerimenti bibliografici

Introduzione

- L'approccio delle risposte potenziali è stato proposto negli anni settanta nella valutazione quantitativa di una politica;
- l'approccio classico definito "strutturale" è basato su un'ampia classe di modelli che tengono in considerazione le scelte individuali e vengono utilizzati per simulare la politica sia *ex post* che *ex ante*; è stato criticato principalmente sulle seguenti caratteristiche:
 - le forme funzionali assunte dal modello non sono verificabili;
 - i modelli sono complessi, poco trasparenti e non permettono di essere replicati;
 - non presta la dovuta attenzione alla *validità interna* (il fatto di ricavare interpretazioni causali consistenti) a vantaggio della *validità esterna* (il fatto che i risultati ottenuti possono essere estesi in altri contesti).

Introduzione

Implica le seguenti definizioni:

- effetto (*impact evaluation*);
- risultato *controfattuale*: l'effetto è valutato come differenza tra ciò che si osserva relativamente a una variabile di interesse in presenza della politica e ciò che si sarebbe osservato in assenza dell'intervento;
- tale idea nasce nell'ambito nei lavori riferiti ai disegni sperimentali (si veda Neyman (1923, 1935) e Fisher, 1935);
- utilizzato anche nella letteratura econometrica in altra forma attraverso il lavoro di Roy (1951) e Quandt (1972);
- solo recentemente formalizzato da Rubin (1974, 1977) e Holland (1986) per studi non sperimentali.

Ricerca sperimentale

- La *ricerca sperimentale* è volta a verificare se al variare di certe condizioni (fattori sperimentali o trattamenti) i risultati sperimentali si modificano sistematicamente;
- Cox e Reid (2000) danno la seguente definizione di *esperimento*:
La parola esperimento viene utilizzata in un senso preciso per significare una ricerca dove il sistema di studio è sotto controllo da parte dell'investigatore. Ovvero i soggetti o gli oggetti dell'analisi ed i metodi di misurazione sono selezionati dal ricercatore almeno per ciò che concerne le caratteristiche principali. Al contrario in uno studio osservazionale alcune caratteristiche in particolare l'allocazione degli individui ai gruppi di trattamento, non sono sotto controllo del ricercatore.
- il *trattamento* è tale che può essere manipolato a livello individuale ed assegnato sotto controllo del ricercatore.

Randomizzazione

- Lo studio randomizzato risulta essere il punto di confronto anche negli studi osservazionali in quanto consente di identificare in modo chiaro che cosa si vuole misurare;
- si ha randomizzazione quando si determina un'associazione casuale tra unità sperimentali ed i trattamenti;
- si ritiene che gli studi più influenti sono quelli contraddistinti da assegnazione casuale al trattamento ad esempio il Perry preschool project (1972):
 - ▷ disegno sperimentale volto a valutare l'efficacia della formazione prescolare effettuato su 123 persone di colore nel Michigan;
 - ▷ alcune di queste persone sono state casualmente assegnate ad un trattamento che comprendeva delle lezioni prescolari a casa;
 - ▷ i soggetti sono stati osservati fino all'età di 27 anni (1993).

Randomizzazione

- ▷ Dallo studio emerge che c'è un effetto positivo generalizzato di questo programma;
- ▷ in un recente studio che ha rianalizzato i dati (Anderson, 2008) è stato confermato l'effetto positivo del programma ma solo per le femmine e non per i maschi;
- ▷ lo studio è stato uno dei riferimenti per ideare la politica dello stato del Michigan denominato Head Start Preschool Program attivo dal 1964 tutt'ora in vigore nello stato.

Risposte potenziali

- Il termine *potential outcome* è stato proposto da Rubin (1974) per indicare che per ogni individuo della popolazione ci sono due stati alternativi di una stessa causa;
- gli stati alternativi indicano dei trattamenti alternativi detti anche trattamento e controllo;
- lo stato è caratterizzato da una serie di condizioni che possono influenzare la *risposta di interesse*: caratteristica osservabile nelle unità della popolazione;
- **Esempio:** stati alternativi: ricovero e non ricovero ospedaliero;
variabile risposta: stato di salute
- **Esempio:** stati alternativi: fumare sigarette e non fumare sigarette;
variabile risposta: tumore al polmone.

Effetto Causale

- Si parla di effetto causale quando il trattamento determina la risposta ovvero cambiando il valore del trattamento cambia la distribuzione della variabile risposta;
- la presenza di associazione tra trattamento e variabile risposta non comporta necessariamente una relazione causale.
- l'effetto è riferito ad una determinata popolazione sulla quale il trattamento può o non può avere effetto in un certo istante temporale;
- **Esempio:** valutazione dell'effetto di un corso di formazione volto al reinserimento lavorativo dei disoccupati iscritti all'ufficio di collocamento della Regione Lombardia da oltre un anno dopo sei mesi dall'erogazione del corso.

Notazione

Per semplicità di notazione si considera sia la variabile risposta che la variabile trattamento come variabili binarie;

- si considera una popolazione costituita da N unità indicizzate da $i = 1, \dots, N$;
- la variabile binaria che rappresenta la variabile risposta realizzata è Y_i : vale 0 se l'individuo non ha trovato lavoro sei mesi dopo il corso e vale 1 se ha trovato lavoro;
- la variabile binaria che rappresenta la variabile trattamento è C_i : vale 0 se l'individuo i non è trattato: non partecipa al corso di formazione e vale 1 se vi partecipa;
- Y_{0i} è la risposta che si sarebbe osservata se l'individuo fosse stato assegnato al trattamento, se avesse frequentato il corso;
- Y_{1i} è la risposta che si sarebbe osservata se all'individuo non fosse stato assegnato al trattamento, se non avesse frequentato il corso.

Risposte potenziali

- Le risposte potenziali

$$\textit{potential outcome} = \begin{cases} Y_{1i} & C = 1 \\ Y_{0i} & C = 0 \end{cases}$$

indicano lo stato lavorativo dell'individuo se questo non avesse seguito il corso e lo stato lavorativo dello stesso individuo se questo lo avesse seguito a prescindere dai fatti che si sono effettivamente realizzati;

- la variabile risposta realizzata è funzione delle risposte potenziali:

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})C_i$$

se $C_i = 0$ la risposta osservata coincide con Y_{0i} ; se $C_i = 1$ la risposta osservata coincide con Y_{1i} .

Risposte potenziali

La seguente tabella sintetizza le informazioni precedenti:

risposta	fattuale	controfattuale
partecipanti ($C = 1$)	Y_1	Y_0
non partecipanti ($C = 0$)	Y_0	Y_1

- i partecipanti ed i non partecipanti hanno entrambi una risposta fattuale e una controfattuale;
- se un individuo partecipa alla politica viene osservata Y_1 mentre Y_0 non viene mai osservata;
- se un individuo non partecipa alla politica viene osservata Y_0 mentre Y_1 non viene mai osservata.

Effetto causale individuale

- Dalla relazione

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})C_i$$

si determina *l'effetto causale* definito come

$$(Y_{1i} - Y_{0i})$$

differenza lineare tra risposte potenziali dell'*i*-esimo individuo se trattato e se non trattato;

- In pratica si tratta di valutare il cambiamento tra due variabili risposta a parità di altre condizioni per un soggetto tra due diversi stati.
- il corso ha un effetto causale individuale se

$$Y_{1i} \neq Y_{0i}$$

Effetto causale medio

- Il *problema fondamentale dell'inferenza causale*: si può osservare solo una delle due risposte potenziali per ogni unità a seconda dello stato del trattamento di cui l'unità ha fatto esperienza;
- invece dell'effetto individuale si fa riferimento alle caratteristiche della *distribuzione* dell'effetto causale nella popolazione come ad esempio la media, allo scopo di ricavare informazioni sulla risposta potenziale che non è mai osservata;
- l'effetto causale medio *Average Treatment Effect, ATE* è definito come

$$ATE = E(Y_1) - E(Y_0)$$

la media delle differenze delle risposte a livello individuale per i soggetti trattati e non trattati.

Esempio teorico

- Supponiamo che la popolazione precedente riferita ai disoccupati sia costituita da 20 individui e di conoscere le risposte potenziali per ogni soggetto;
- Y^1 è pari a 1 se il soggetto lavora un sei mesi e ha fatto il corso e paria 0 se non lavora, mentre Y^0 vale 1 se il soggetto lavora e non ha fatto il corso e vale 0 se non lavora e non ha fatto il corso;

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Y^0	0	1	0	0	0	1	0	0	1	1	0	1	1	0	0	0	1	1	1	1
Y^1	1	0	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	0	0	0

- l'effetto causale medio si può determinare con le rispettive probabilità essendo variabili binarie

$$\theta = E(Y^1) - E(Y^0) = P(Y^1 = 1) - P(Y^0 = 1)$$

- essendo $P(Y^1 = 1) = 10/20 = 0.5$ la probabilità di trovare lavoro se non è stato fatto il corso è $P(Y^0 = 1) = 10/20 = 0.5$;
- l'effetto causale medio è pari a zero.

Esempio teorico

- Le risposte potenziali presentano dei valori mancanti
- si osserva la variabile C che indica se il soggetto ha seguito il corso di formazione e la risposta Y che indica lo stato lavorativo nell'istante considerato;

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Y^0	0	1	0	0	-	-	-	-	1	1	0	-	-	-	-	-	-	-	-	-
Y^1	-	-	-	-	0	0	0	1	-	-	-	1	1	1	1	1	1	0	0	0
C	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
Y	0	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	0	0	0

- L'associazione che esiste tra il fatto di aver fatto il corso e il fatto di lavorare nel periodo considerato si esprime come una probabilità condizionata

$$D = E[Y|C = 1] - E[Y|C = 0] = P(Y = 1|C = 1) - P(Y = 1|C = 0);$$

- risulta che tra i 13 soggetti che hanno fatto il corso ci sono 7 persone che lavorano;
- su 7 persone che non hanno fatto il corso ci sono 3 persone che lavorano;
- esiste un'associazione tra il fatto di aver fatto il corso ed il fatto di aver trovato un lavoro: $53.8\% - 42.8\% = 11.0\%$.

Associazione e randomizzazione

- L'associazione statistica non equivale alla relazione causale: l'effetto causale si basa sul confronto degli stessi soggetti dopo che lo stesso soggetto è stato sottoposto a condizioni alternative;
- la misura di associazione confronta la probabilità di risposta per gli individui che nella popolazione sono risultati trattati e non trattati.
- la randomizzazione garantisce che l'associazione osservata possa essere interpretata in senso causale;
- i valori mancanti risultano essere mancanti completamente a caso;
- consente di stimare in modo consistente l'effetto pur non conoscendo entrambe le risposte potenziali;
- garantisce l'indipendenza del trattamento dalle risposte potenziali

$$C \perp\!\!\!\perp (Y^0, Y^1)$$

Randomizzazione

- La randomizzazione risulta essere il punto di riferimento dell'approccio basato sulle risposte potenziali;
- **Esempio:** dallo studio randomizzato condotto da Hsia *et al.* (2006) per valutare l'effetto degli ormoni prescritti alle donne in menopausa è emerso un effetto nefasto degli stessi contrariamente ai risultati di precedenti studi osservazionali;
- se il soggetto non è esposto in modo casuale al trattamento l'effetto causale non può essere stimato dai dati campionari in modo consistente a meno di introdurre ipotesi più forti che devono essere verificate.

Esempio

- The Tennessee Student/Teacher Achievement Ratio è uno studio randomizzato volto a valutare l'effetto dell'ampiezza delle classi nell'apprendimento alla scuola elementare (Finn e Achilles, 1990);
- i risultati di numerosi studi osservazionali hanno dimostrato che esiste un debole legame tra l'ampiezza della classe e l'apprendimento;
- nello stato dove la ricerca è stata condotta si riscontra che spesso gli studenti più in difficoltà chiedono di essere allocati alle classi più piccole;
- attraverso la randomizzazione gli studenti assegnati alle classi di dimensioni ridotte e a quelle più ampie sono ugualmente confrontabili;
- dai risultati è emerso un rilevante e durevole effetto sull'apprendimento delle classi meno ampie.

Esempio

- The Tennessee Student/Teacher Achievement Ratio ha avuto un costo pari a 12 milioni di dollari
- è stato implementato per la coorte di bambini nati nel 1985-86;
- ha coinvolto in totale 11600 bambini che sono stati seguiti dalla prima alla terza elementare;
- dai risultati è emerso un rilevante e durevole effetto delle classi di dimensioni ridotte sull'apprendimento;
- l'ampiezza media delle classi nello stato è di 22.3 bambini;
- gli studenti sono stati assegnati a caso ad una delle seguenti classi:
 - una classe dai 15 ai 17 bambini;
 - una classe dei 22 ai 25 bambini con un insegnante aggiuntiva part time;
 - una classe con circa 22 bambini con l'ausilio di un'insegnante a tempo pieno.

Disegno quasi-sperimentale

- I disegni quasi sperimentali costituiscono un'alternativa alla randomizzazione quando questa presenta elevati costi sociali e difficoltà logistiche;
- Angrist e Lavy (1999) hanno stimato l'effetto della numerosità della classe sul livello di apprendimento senza dover applicare l'esperimento controllato;
 - in Israele si prevedono 40 alunni per classe, altrimenti si formano due classi da 20 e da 21 alunni rispettivamente. Gli studenti nelle classi di diversa dimensione sono supposti essere simili per quanto riguarda l'abilità ed il contesto familiare;
 - l'assegnazione rappresenta un trattamento assegnato in modo casuale ed è possibile confrontare le risposte di interesse;
 - è emerso uno stretto legame tra la numerosità di classe e l'apprendimento.

Effetto medio del trattamento sui trattati

- L'effetto medio del trattamento sui trattati (*average treated effect on the treated, ATT*) è definito come

$$E[Y_1 - Y_0|C = 1] = E[Y_1|C = 1] - E[Y_0|C = 1]$$

- esprime la media delle differenze tra lo stato lavorativo risultante per coloro che hanno seguito il corso e lo stato lavorativo che avrebbero avuto se non avessero seguito il corso;
- misura i "benefici" ottenuti in media dagli individui che hanno usufruito della politica;
- diversamente da ATE si riferisce solo ai partecipanti alla politica;
- come ATE non può essere direttamente identificato.

Effetto di selezione

La differenza media osservata sulla risposta condizionata al trattamento è legata all'effetto causale dalla seguente relazione:

$$E[Y|C = 1] - E[Y|C = 0] = E[Y_1|C = 1] - E[Y_0|C = 1] + E[Y_0|C = 1] - E[Y_0|C = 0]$$

- dove $E[Y_1|C = 1] - E[Y_0|C = 1]$ è ATT
- mentre $E[Y_0|C = 1] - E[Y_0|C = 0]$ è il termine che identifica l'effetto di selezione che comporta una distorsione.

Effetto di selezione

- Rappresenta la differenza nella risposta che si sarebbe osservata tra i partecipanti al corso e i non partecipanti se il corso non fosse stato erogato;
- dipende dalle differenze esistenti precedenti al trattamento tra i due gruppi;
- la differenza media osservata coincide con l'ATT solo nel caso in cui l'effetto di selezione è nullo;
- coinvolge il termine che è osservato $E[Y_0|C = 1]$ ed il termine che non è osservabile $E[Y_0|C = 0]$ riferito alla risposta potenziale per i non partecipanti al corso; pertanto non può essere misurato.

Assunzioni

- Negli studi osservazionali esiste un meccanismo di assegnazione (o processo di selezione) delle unità: il processo per cui alcune unità scelgono di sottoporsi al trattamento;
- l'assunzione di indipendenza condizionata detta *conditional independence assumption (CIA)* (o selezione sulle osservate) se verificata consente di stimare l'effetto medio causale;
- si basa sull'esistenza di una variabile (o più variabili) concomitante X che non è influenzata dal trattamento e tale che

$$(Y_1, Y_0) \perp\!\!\!\perp C | X$$

condizionando rispetto ad questa il trattamento è indipendente dalle risposte potenziali;

- ipotesi spesso utilizzata in modo implicito quando si interpretano in senso causale i coefficienti di regressione stimati.

Assunzioni

- ovvero il meccanismo che ha portato a scegliere di frequentare il corso è assunto essere irrilevante sul valore realizzato della variabile risposta;
- nell'esempio del trapianto di cuore potrebbe essere il sesso;
- se è plausibile la CIA il confronto tra la probabilità di trovare lavoro dopo il corso per coloro che lo hanno fatto e coloro che non lo hanno fatto diventa una stima consistente dell'effetto causale

$$E[Y|X, C = 1] - E[Y|X, C = 0] = E[Y_1 - Y_0|X]$$

- un'attenta analisi del processo di selezione può portare a riprodurre in un contesto osservazionale le stesse proprietà di uno studio randomizzato in modo da garantire la validità interna.

Assunzioni

- L'assunzione di stabilità detta *Stable Unit Treatment-Value Assumption (SUTVA)* (detta anche di equilibrio parziale) implica che:
 - ogni unità sia assegnata esclusivamente o al trattamento o al controllo;
 - che non vi siano interazioni tra le unità della popolazione: che la risposta degli individui non sia influenzata da cambiamenti potenziali nell'esposizione al trattamento di altri individui: poco plausibile in alcuni contesti come ad esempio in epidemiologia.

Assunzioni

- L'assunzione di sovrapposizione detta *full common support* (anche replicabilità) implica che

$$0 < P(C_i = 1 | X_i = x) < 1$$

per ogni possibile valore delle covariate esistono delle unità che risultano trattate e non trattate;

- garantisce che vi sia una certa variabilità del trattamento rispetto alle unità.

Suggerimenti di Rubin

- Secondo Rubin per riferirsi ad uno studio osservazionale come fosse un disegno sperimentale occorre:
 - tenere in considerazione gli agenti e le covariate che questi utilizzano per prendere la decisione di esporsi al trattamento;
 - il dati a disposizione devono comprendere le covariate di cui sopra e se queste non risultano disponibili occorre cercare altre fonti informative;
 - occorre creare dei sottogruppi o sottoclassi omogenee rispetto alle covariate osservate entro le quali i controlli ed i trattati siano bilanciati;
 - fare inferenza limitamente a tali sottoclassi;
- tale procedimento comporta la creazione di un disegno oggettivo in grado di approssimare un disegno randomizzato in ogni sottoclasse.

Caratteristiche dell'approccio delle risposte potenziali

- Tale approccio presuppone di dover specificare attentamente il modo in cui il trattamento influenza il valore atteso della variabile risposta per la popolazione di riferimento;
- presuppone che si identifichi accuratamente lo scenario in cui la risposta viene osservata;
- non richiede delle specifiche assunzioni circa la forma funzionale o sulla distribuzione di probabilità della variabile risposta;
- non richiede di modellizzare il modo in cui l'individuo sceglie di essere trattato oppure subisce il trattamento;
- separa la risposta potenziale dal modo in cui viene assegnato il trattamento;

Caratteristiche dell'approccio delle risposte potenziali

- Richiede di verificare le assunzioni necessarie per stimare in modo consistente l'effetto causale medio a cui si è interessati;
- consente di identificare in senso pieno i parametri d'interesse e di individuare alcune forme di eterogeneità nelle risposte rispetto al trattamento;
- risulta essere un metodo che gode di una robusta validità interna;
- la metodologia è piuttosto semplice.

Riferimenti bibliografici

- Agrist J. D., Lavy V. (1999). Using Maimonides' Rule To Estimate The Effect Of Class Size On Scholastic Achievement. *The quarterly Journal of Economics*, **114**, 533-575.
- Cox, D.R., Reid, N. (2000). *The theory of the design of experiment*, Chapman & Hall/CRC.
- Finn J. D., Achilles C. M. (1990). Answers and Questions about class size: A statewide experiment. *American Educational Research Journal*, **28**, 557–577.
- Fisher R. A. (1935). Contribution to a discussion on J. Neyman's paper on statistical problems in agricultural experimentation. *J. Roy. Statist. Soc., Suppl.*, **2**, 154-157.
- Hsia J., Langer R.D., Manson J.E., Kuller L., Johnson K.C., Hendrix S.L., Pettinger M., Heckbert S.R., Greep N., Crawford S., Eaton C.B., Kostis J.B., Caralis P., Prentice R. (2006). Conjugated equine estrogens and coronary heart disease: the Women's Health Initiative.

Riferimenti bibliografici

- Neyman J. S. (1935). Statistical problems in agricultural experimentation (with Discussion). *J. R. Statist. Soc. series B*, **2**, 107-80.
- Neyman J. S. (1990)[1923]. On the application of probability theory to agricultural experiment. Essay on principles. *Stat. Science*, **5**, 465-80.
- Quandt, R. E. (1972). A new approach to estimating switching regression, *Journal of the American Statistical Association*, **67**, 306-310.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, **3**, 135-146.
- Rubin D. B. (1974). Estimating causal effect of treatments in randomized and non randomized studies. *Journal of Educational Psychology*, **66**, 688–701.
- Rubin D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Edu. Stat.*, **2**, 1–26.

Indicazioni Bibliografiche

Angrist J. and Pischke J.S. (2009). *Mostly harmless econometrics*.
Oxford: Princeton University press.

Angrist, J., A. Krueger (2000). *Empirical Strategies in Labor Economics*,
Handbook of Labor Economics, O. Ashenfelter and D. Card (eds.),
North Holland, 1277-1366.

Heckman, J. (1981). Heterogeneity and state dependence. In *Studies of
Labor Markets*, Ed. S. Rosen. University of Chicago Press, pp.
91–139.

Heckman, J. J. (2001). Micro Data, Heterogeneity, and the Evaluation of
Public policy: Nobel Lecture, *Journal of Political Economy*, **109**,
673-748.

Indicazioni Bibliografiche

Holland, P. W. (1986). Statistics and Causal Inference, *Journal of the American Statistical Association*, 81, 945-970.

Meyer, B. (1995). Natural and Quasi-experiments in Economics, *Journal of Business and Economic Statistics*, **13**, 151-161.

The Econometrics of Panel Data Fundamentals and Recent Developments in Theory and Practice Series: Advanced Studies in Theoretical and Applied Econometrics , Vol. 46 Mátyás, L., Sevestre P. (Eds.). Springer-Verlag (2008).

Bartolucci F., Pennoni F. (2009). Impact evaluation of job training programs by a latent variable model. *Sottoposto*). arXiv:0909.10965.