

Le metodologie di campionamento e scomposizione della devianza nelle rilevazioni nazionali dell'Invalsi

Piero Demetrio Falorsi, Patrizia Falzetti, Roberto Ricci

Discussione

M. Giovanna Ranalli¹

¹Dip. Scienze Politiche, Università di Perugia
giovanna.ranalli@unipg.it

Servizio Statistico Invalsi — ISTAT
19 Novembre 2020, 15:30

Agenda

- 1 Sintesi
- 2 Disegno di campionamento
- 3 Fase di stima
- 4 Cheating

Sintesi 1

Rilevazione sugli apprendimenti rivolta a tutte le scuole, statali e paritarie, su diversi gradi scolari

- gradi 2° e 5° → fascicoli cartacei
- grado 8° → CBT
- gradi 10° e 13° → CBT

Due indagini distinte che differiscono per il soggetto che somministra e supervisiona la prova

ICC Indagine Campionaria di Controllo: prova svolta sotto il controllo di personale esterno al personale docente su un campione probabilistico di classi.

IS Indagine Standard: prova svolta sotto il controllo del personale docente su tutte le altre classi.

Sintesi 1

Rilevazione sugli apprendimenti rivolta a tutte le scuole, statali e paritarie, su diversi gradi scolari

- gradi 2° e 5° → fascicoli cartacei
- grado 8° → CBT
- gradi 10° e 13° → CBT

Due indagini distinte che differiscono per il soggetto che somministra e supervisiona la prova

ICC Indagine Campionaria di Controllo: prova svolta sotto il controllo di personale esterno al personale docente su un campione probabilistico di classi.

IS Indagine Standard: prova svolta sotto il controllo del personale docente su tutte le altre classi.

Sintesi 2

Cheating motivazione principale dell'ICC

- Sovrastima sistematica
- Errore di misura non a media nulla
- Valutazione qualità della rilevazione IS e aggiustamento delle stime finali

Discussione piuttosto tecnica

- Disegno di campionamento
- Fase di stima
- Aggiustamento della sovrastima da cheating

Notazione per decodificare alcuni commenti

- ? Minor issue
- ?? Major comment

Sintesi 2

Cheating motivazione principale dell'ICC

- Sovrastima sistematica
- Errore di misura non a media nulla
- Valutazione qualità della rilevazione IS e aggiustamento delle stime finali

Discussione piuttosto tecnica

- Disegno di campionamento
- Fase di stima
- Aggiustamento della sovrastima da cheating

Notazione per decodificare alcuni commenti

? Minor issue

?? Major comment

Sintesi 2

Cheating motivazione principale dell'ICC

- Sovrastima sistematica
- Errore di misura non a media nulla
- Valutazione qualità della rilevazione IS e aggiustamento delle stime finali

Discussione piuttosto tecnica

- Disegno di campionamento
- Fase di stima
- Aggiustamento della sovrastima da cheating

Notazione per decodificare alcuni commenti

- ? Minor issue
- ?? Major comment

Agenda

- 1 Sintesi
- 2 Disegno di campionamento
- 3 Fase di stima
- 4 Cheating

Parametri oggetto di interesse

- Valori medi dei punteggi y
- Livello di dettaglio: $22 \times 13 = 286$ domini di stima
 - 21 'regioni' + Italia = 22 domini territoriali
 - 2 primaria + 1 secondaria primo grado + 2 secondaria secondo grado \times 5 tipologie di scuole = 13 sottopopolazioni

Parametri oggetto di interesse

- **Devianza** tra i punteggi D e sue componenti
 - $D1$: devianza fra le scuole
 - $D2$: devianza fra le classi entro la scuola
 - $D3$: devianza fra gli alunni entro la classe

? $D1$ cambia secondo il livello territoriale?

Regioni $D1$ è la devianza fra le scuole **entro le Regioni**

Italia $D1$ è la devianza fra le scuole **entro il territorio nazionale** o si scompone ulteriormente?

$D1A$ devianza fra le scuole entro le Regioni

$D1B$ devianza fra le Regioni entro il territorio nazionale

Parametri oggetto di interesse

- **Devianza** tra i punteggi D e sue componenti
 - $D1$: devianza fra le scuole
 - $D2$: devianza fra le classi entro la scuola
 - $D3$: devianza fra gli alunni entro la classe

? $D1$ cambia secondo il livello territoriale?

Regioni $D1$ è la devianza fra le scuole **entro le Regioni**

Italia $D1$ è la devianza fra le scuole **entro il territorio nazionale** o si scompone ulteriormente?

$D1A$ devianza fra le scuole entro le Regioni

$D1B$ devianza fra le Regioni entro il territorio nazionale

Piano di campionamento

- Piano di campionamento a due stadi stratificato
 - 1° stadio Scuole – campionamento con probabilità proporzionali al numero di alunni nella scuola
 - 2° stadio Sezioni – campionamento casuale semplice delle sezioni nelle scuole estratte
- Stratificazione delle scuole secondo i domini di interesse
- Allocazione di compromesso del campione di 1° stadio: scelta del coefficiente β
 - ? Quali valori soglia sono stati scelti per il CV?
 - ? $\beta \approx 0.3$: quota relativamente minore rispetto all'allocazione proporzionale, per alcune regioni la variabilità era troppo alta (typo eq 5.6)
 - ?? Quale β per D e le sue componenti?

Piano di campionamento

- Piano di campionamento a due stadi stratificato
 - 1° stadio Scuole – campionamento con probabilità proporzionali al numero di alunni nella scuola
 - 2° stadio Sezioni – campionamento casuale semplice delle sezioni nelle scuole estratte
- Stratificazione delle scuole secondo i domini di interesse
- Allocazione di compromesso del campione di 1° stadio: scelta del coefficiente β
 - ? Quali valori soglia sono stati scelti per il CV?
 - ? $\beta \approx 0.3$: quota relativamente minore rispetto all'allocazione proporzionale, per alcune regioni la variabilità era troppo alta (typo eq 5.6)
 - ?? Quale β per D e le sue componenti?

Campionamento al primo stadio

- Il piano di campionamento cerca di assecondare il modello che spiega la relazione della variabile di interesse con le variabili a disposizione note per tutte le scuole
- ?? Dai dati disponibili, studio della relazione del punteggio y con alcune caratteristiche oltre a quelle usate per la stratificazione per un possibile bilanciamento: punteggio precedente, componenti di devianza precedenti, scuola statale/paritaria ...
- ?? Complicazione (vedi stima varianza) della selezione con prob. variabile è supportata dai dati?
- ?? Utilizzare come variabile di interesse per la definizione del piano di campionamento l'ammontare di cheating
- ?? Possibile componente panel o rotazione del campione ICC

Campionamento al primo stadio

- Il piano di campionamento cerca di assecondare il modello che spiega la relazione della variabile di interesse con le variabili a disposizione note per tutte le scuole
- ?? Dai dati disponibili, studio della relazione del punteggio y con alcune caratteristiche oltre a quelle usate per la stratificazione per un possibile bilanciamento: punteggio precedente, componenti di devianza precedenti, scuola statale/paritaria ...
- ?? Complicazione (vedi stima varianza) della selezione con prob. variabile è supportata dai dati?
- ?? Utilizzare come variabile di interesse per la definizione del piano di campionamento l'ammontare di cheating
- ?? Possibile componente panel o rotazione del campione ICC

Campionamento al primo stadio

- Il piano di campionamento cerca di assecondare il modello che spiega la relazione della variabile di interesse con le variabili a disposizione note per tutte le scuole
- ?? Dai dati disponibili, studio della relazione del punteggio y con alcune caratteristiche oltre a quelle usate per la stratificazione per un possibile bilanciamento: punteggio precedente, componenti di devianza precedenti, scuola statale/paritaria ...
- ?? Complicazione (vedi stima varianza) della selezione con prob. variabile è supportata dai dati?
- ?? Utilizzare come variabile di interesse per la definizione del piano di campionamento l'ammontare di cheating
- ?? Possibile componente panel o rotazione del campione ICC

Campionamento al primo stadio

- Il piano di campionamento cerca di assecondare il modello che spiega la relazione della variabile di interesse con le variabili a disposizione note per tutte le scuole
- ?? Dai dati disponibili, studio della relazione del punteggio y con alcune caratteristiche oltre a quelle usate per la stratificazione per un possibile bilanciamento: punteggio precedente, componenti di devianza precedenti, scuola statale/paritaria ...
- ?? Complicazione (vedi stima varianza) della selezione con prob. variabile è supportata dai dati?
- ?? Utilizzare come variabile di interesse per la definizione del piano di campionamento l'ammontare di cheating
- ?? Possibile componente panel o rotazione del campione ICC

Campionamento al primo stadio

- Il piano di campionamento cerca di assecondare il modello che spiega la relazione della variabile di interesse con le variabili a disposizione note per tutte le scuole
- ?? Dai dati disponibili, studio della relazione del punteggio y con alcune caratteristiche oltre a quelle usate per la stratificazione per un possibile bilanciamento: punteggio precedente, componenti di devianza precedenti, scuola statale/paritaria ...
- ?? Complicazione (vedi stima varianza) della selezione con prob. variabile è supportata dai dati?
- ?? Utilizzare come variabile di interesse per la definizione del piano di campionamento l'ammontare di cheating
- ?? Possibile componente panel o rotazione del campione ICC

Agenda

- 1 Sintesi
- 2 Disegno di campionamento
- 3 Fase di stima**
- 4 Cheating

Fase di stima

- Pesì di riporto all'universo: peso di primo stadio è calibrato per il numero e la dimensione delle scuole per tipologia e grado.
- ?? Altre variabili su cui calibrare o da usare per analisi sensitività (scuola statale/paritaria, punteggio precedente, sesso ...)
- ? Perché stimatore rapporto (eq 6.3)? Il denominatore è calibrato (eq. 6.2).
- ? Perché introdurre la tipologia differenziata t^* (pagg. 35-36)? Non era già un dominio stratificato?
- ? Tabelle 12-18 sono le stime da ICC o ICC+IS?
- ?? Perché proprio $D2$ stimata per differenza? Cosa cambierebbe se si stimassero tutte le componenti indipendentemente?

Fase di stima

- Pesì di riporto all'universo: peso di primo stadio è calibrato per il numero e la dimensione delle scuole per tipologia e grado.
- ?? Altre variabili su cui calibrare o da usare per analisi sensitività (scuola statale/paritaria, punteggio precedente, sesso ...)
- ? Perché stimatore rapporto (eq 6.3)? Il denominatore è calibrato (eq. 6.2).
- ? Perché introdurre la tipologia differenziata t^* (pagg. 35-36)? Non era già un dominio stratificato?
- ? Tabelle 12-18 sono le stime da ICC o ICC+IS?
- ?? Perché proprio $D2$ stimata per differenza? Cosa cambierebbe se si stimassero tutte le componenti indipendentemente?

Stima della varianza

- Formule per il calcolo delle varianze campionarie per chi dispone dei micro-dati
- ?? Stima di varianza per le componenti della devianza?
- ?? Fornire pesi Jackknife o Bootstrap?

Agenda

- 1 Sintesi
- 2 Disegno di campionamento
- 3 Fase di stima
- 4 Cheating

Cheating

- Classico problema di valutazione (potential outcomes)
- Usare informazione proveniente da ICC per aggiustare IS
- Problema di integrazione di indagini

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | | ✓ |

- y^* variabile osservata con errore

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | ? | ✓ |

Cheating

- Classico problema di valutazione (potential outcomes)
- Usare informazione proveniente da ICC per aggiustare IS
- Problema di integrazione di indagini

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | | ✓ |

- y^* variabile osservata con errore

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | ? | ✓ |

Cheating

- Classico problema di valutazione (potential outcomes)
- Usare informazione proveniente da ICC per aggiustare IS
- Problema di integrazione di indagini

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | | ✓ |

- y^* variabile osservata con errore

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | ? | ✓ |

Un possibile approccio



Park, S., Kim, J. K., & Stukel, D. (2017)

A measurement error model approach to survey data integration: combining information from two surveys.

Metron, 75(3), 345-357.

- Siano $f(y|x, \theta)$ e $f(y^*|x, y; \theta^*)$ le distribuzioni condizionate per y e per y^*
- Si imputano i valori di y in IS tramite le migliori predizioni del controfattuale del valore y , corrette per l'errore di misura presente in y^* . I valori imputati sono generati a partire da $f(y|x, y^*; \theta, \theta^*)$
- La stima (pseudoML via algoritmo EM) di θ e θ^* permette di ottenere

$$\hat{y}_i = E(y_i|x_i, y_i^*; \hat{\theta}, \hat{\theta}^*), \quad \text{per } i \in \text{IS}$$

Un possibile approccio



Park, S., Kim, J. K., & Stukel, D. (2017)

A measurement error model approach to survey data integration: combining information from two surveys.

Metron, 75(3), 345-357.

- Siano $f(y|x, \theta)$ e $f(y^*|x, y; \theta^*)$ le distribuzioni condizionate per y e per y^*
- Si imputano i valori di y in IS tramite le migliori predizioni del controfattuale del valore y , corrette per l'errore di misura presente in y^* . I valori imputati sono generati a partire da $f(y|x, y^*; \theta, \theta^*)$
- La stima (pseudoML via algoritmo EM) di θ e θ^* permette di ottenere

$$\hat{y}_i = E(y_i|x_i, y_i^*; \hat{\theta}, \hat{\theta}^*), \quad \text{per } i \in \text{IS}$$

Mass imputation

| | variabili ausiliarie x | punteggio y | punteggio y^* |
|-----|--------------------------|---------------|-----------------|
| ICC | ✓ | ✓ | |
| IS | ✓ | ✓ | ✓ |

- La stima (pseudoML via algoritmo EM) di θ e θ^* permette di ottenere

$$\hat{y}_i = E(y_i | x_i, y_i^*; \hat{\theta}, \hat{\theta}^*), \quad \text{per } i \in \text{IS}$$

- Un possibile stimatore combinato è dato da

$$\bar{y}_{comb} = \frac{\sum_{ICC} w_{iICC} y_i + \sum_{IS} w_{iIS} \hat{y}_i}{\sum_{ICC} w_{iICC} + \sum_{IS} w_{iIS}}$$

Conclusioni

- Lavoro rigoroso, completo, adeguato...
 - Parametro di interesse *vero* di ICC è la distorsione da cheating
- Valutare se continuare l'indagine su CBT
- Monitoraggio per studio del trend nel tempo degli indicatori di interesse
- Spazio per sperimentazioni di innovazioni metodologiche

Conclusioni

- Lavoro rigoroso, completo, adeguato...
 - Parametro di interesse *vero* di ICC è la distorsione da cheating
- Valutare se continuare l'indagine su CBT
- Monitoraggio per studio del trend nel tempo degli indicatori di interesse
- Spazio per sperimentazioni di innovazioni metodologiche

Conclusioni

- Lavoro rigoroso, completo, adeguato...
 - Parametro di interesse *vero* di ICC è la distorsione da cheating
- Valutare se continuare l'indagine su CBT
- Monitoraggio per studio del trend nel tempo degli indicatori di interesse
- Spazio per sperimentazioni di innovazioni metodologiche