

LE METODOLOGIE DI  
CAMPIONAMENTO E  
DI SCOMPOSIZIONE  
DELLA DEVIANZA  
NELLE RILEVAZIONI  
NAZIONALI 2013-2017

In collaborazione con  
Piero Demetrio Falorsi

Le rilevazioni degli  
apprendimenti  
nazionali



## Sommario

PREFAZIONE.....	3
1. INTRODUZIONE.....	4
2. LE RILEVAZIONI INVALSI.....	6
3. TECNICA DI MISURAZIONE.....	6
4. OBIETTIVI DELL'INDAGINE.....	8
4.1 SIMBOLOGIA DI BASE.....	9
4.2 PUNTEGGI MEDI.....	11
4.3 COMPONENTI DELLA DEVIANZA.....	11
5. CAMPIONAMENTO.....	12
5.1 PRIMO STADIO DI CAMPIONAMENTO.....	12
5.2 SECONDO STADIO.....	14
5.3 PROBABILITÀ DI INCLUSIONE FINALE.....	15
5.4 ALLOCAZIONE DEL CAMPIONE.....	16
6. STIME CAMPIONARIE.....	19
6.1 DESCRIZIONE GENERALE.....	19
6.2 PESO DI SECONDO STADIO.....	20
6.3 PESO DI PRIMO STADIO.....	20
6.4 PESO ALUNNO.....	20
6.5 STIMA DEI PUNTEGGI MEDI.....	21
6.6 STIMA DELLE COMPONENTI DELLA DEVIANZA.....	22
7. VARIABILITA' CAMPIONARIA.....	24
7.1 ESPRESSIONE DELLA VARIANZA CAMPIONARIA.....	24
7.2 STIMA DELLA VARIANZA CAMPIONARIA.....	25
7.3 UN MODELLO PER LA PRESENTAZIONE SINTETICA DEGLI ERRORI CAMPIONARI.....	28
8. IL CHEATING E L'AFFIDABILITA' DEI DATI.....	36
8.1 L'INDIVIDUAZIONE E CORREZIONE DEL <i>CHEATING</i> MEDIANTE L'APPROCCIO <i>FUZZY CLUSTERING</i> .....	36
8.2 COME VIENE TRATTATO NEL CAMPIONE: IL LIVELLO 8.....	37
RIFERIMENTI BIBLIOGRAFICI.....	38
APPENDICE - Approfondimenti metodologici.....	39
A.1 RIFORMULAZIONE DEI PUNTEGGI MEDI.....	39
A.2 DIFFERENTI COMPONENTI DELLA DEVIANZA.....	40
A.3 CONDIZIONE DI AUTOPONDERAZIONE.....	41
A.4 ALGORITMO PER IL PESO CALIBRATO.....	42
A.5 STIMA DELLE COMPONENTI DELLA DEVIANZA.....	43
A.6 ESPRESSIONE ALTERNATIVA DELLA STIMA DELLA VARIANZA REGIONALE.....	46
A.7 UN APPROCCIO <i>MULTISTEP</i> PER L'INDIVIDUAZIONE E LA CORREZIONE DEL <i>CHEATING</i> NEI RISULTATI DEL SNV.....	47

## PREFAZIONE

In un'economia matura, basata sui servizi a valore aggiunto, il livello di competenze rappresenta la base essenziale che permette di produrre innovazione e generare nuova ricchezza. In questo quadro diviene essenziale, controllare l'efficacia del sistema educativo, evidenziarne punti di attenzione e sistemi di eccellenza.

Un adeguato sistema di controllo e monitoraggio consente quindi di progettare in modo appropriato le specifiche *policy* per la scuola indicando agli organi di governo le priorità di intervento. Inoltre, facendo emergere le buone pratiche, offre al sistema nel suo complesso la capacità di evolvere automigliorandosi.

La *Rilevazione sugli apprendimenti* costituisce la base primaria per la costruzione del sistema di monitoraggio e controllo e nel corso degli anni ha permesso di accumulare un patrimonio di conoscenze essenziale al sistema paese.

La progettazione e la realizzazione della *Rilevazione*, costituisce una buona pratica di integrazione di conoscenze specialistiche di alto livello, come le competenze tematiche sui sistemi e sugli ambiti dell'apprendimento, le competenze statistiche metodologiche, le competenze organizzative gestionali. La *Rilevazione* costituisce anche un buon esempio di cooperazione istituzionale in cui l'ISTAT e l'INVALSI, i due Enti più rilevanti per la produzione dell'informazione sul sistema scolastico, hanno positivamente cooperato per fornire al paese un prodotto di ottima qualità scientifica e rispondente agli standard più elevati a livello internazionale.

In questo quadro, la piena trasparenza sulle metodologie adottate costituisce un valore aggiunto all'informazione rilasciata. Sono quindi contenta e onorata della pubblicazione del presente volume che descrive in dettaglio tutta la strategia di rilevazione adottata, fornendo a tutti i ricercatori e agli utilizzatori dei dati tutte le informazioni essenziali a certificarne la qualità dei metodi e dei risultati ottenuti.

La Presidente dell'INVALSI

Annamaria Ajello

## 1. INTRODUZIONE

L'INVALSI (Istituto nazionale per la valutazione del sistema educativo di istruzione e formazione, cfr. d. lgs. n. 286/2004), tra i suoi compiti, ha quello di “*attuare verifiche periodiche e sistematiche sulle conoscenze e abilità degli studenti*”.

Tale compito è realizzato effettuando ogni anno una *Rilevazione sugli apprendimenti* che - secondo quanto stabilito dalla Direttiva 85/2012 del Ministero dell'Istruzione, dell'Università e della Ricerca - è rivolta ai seguenti *livelli scolari*<sup>1</sup>:

- le classi II e V della scuola primaria, corrispondenti ai livelli scolari 2° e 5°;
- le classi I<sup>2</sup> e III della scuola secondaria di primo grado corrispondenti ai livelli scolari 6° e 8°;
- la classe II della scuola secondaria di secondo grado<sup>3</sup>, corrispondente al 10° livello scolastico.

La Rilevazione sui livelli suddetti è condotta su tutte le scuole, statali e paritarie, del Paese e su tutti gli studenti delle suddette classi. Le informazioni di interesse sono raccolte mediante la somministrazione agli alunni di prove su ambiti disciplinari specifici (Italiano<sup>4</sup> e Matematica). La rilevazione è costituita dall'unione di due indagini distinte che differiscono per la modalità di somministrazione delle prove che rimangono le medesime in entrambe le rilevazioni:

1. nell'*Indagine Standard* (di seguito denominata *IS*) le prove sono effettuate avvalendosi di somministratori e valutatori appartenenti al corpo docente della scuola. L'indagine è effettuata sulla quasi totalità delle classi per gli anni scolari oggetto di interesse;
2. nell'*Indagine Campionaria di Controllo* (di seguito denominata *ICC*) le prove sono svolte sotto il controllo di osservatori esterni al corpo docente della classe e della scuola<sup>5</sup>. L'*ICC*

<sup>1</sup> Il presente lavoro fa riferimento nello specifico alle Rilevazioni nazionali svolte durante l'anno scolastico 2012-13.

<sup>2</sup> Dall'a.s. 2013-14 la rilevazione nel primo anno della scuola secondaria di primo grado (livello scolastico 6°) non è più stata effettuata grazie all'ampia disponibilità dei codici SIDI che consentono di restituire gli esiti del livello scolastico 5° anche al livello scolastico 6°.

<sup>3</sup> Si noti che al momento la rilevazione non include come livello la classe V della scuola secondaria di secondo grado. Tuttavia, secondo i programmi già definiti, tale livello verrà considerato nelle rilevazioni dei prossimi anni. Inoltre, come previsto dalla legge 176/2007, l'INVALSI ha predisposto la Prova nazionale, inserita all'interno dell'esame di Stato conclusivo del primo ciclo di istruzione.

<sup>4</sup> La prova di Italiano comprende una parte relativa alla comprensione della lettura e una parte relativa alla riflessione sul funzionamento della lingua italiana (Grammatica).

<sup>5</sup> Nel caso della Prova nazionale svolta al termine del III anno della scuola secondaria di primo grado l'osservatore esterno coincide con il Presidente della Commissione di esame.

coinvolge la parte rimanente delle classi non interessate dall'IS. Tale sottoinsieme viene scelto con criteri probabilistici.

Il compito dell'osservatore esterno nell'ICC è quello di monitorare la correttezza della somministrazione, a garanzia del rispetto delle procedure, e di riportare le risposte fornite dagli allievi su apposite schede elettroniche predisposte dall'INVALSI. La scelta di prevedere un'indagine campionaria con osservatori esterni è connessa al fatto che questi ultimi sono meno propensi ad aiutare gli allievi nella risoluzione delle prove (fenomeno noto con il termine *cheating*<sup>6</sup>) e, di conseguenza, le valutazioni che si ottengono con l'ICC, pur se soggette all'errore di campionamento, sono meno esposte all'errore di misura derivante dal *cheating* che tende a sovrastimare i livelli effettivi delle competenze. Il fenomeno del *cheating* viene trattato in modo più approfondito nel capitolo 8 del presente lavoro.

Per l'anno scolastico 2012-2013, la Rilevazione nel suo complesso ha coinvolto 141.784 classi di circa 13.200 Istituti, statali e paritari, e tutti gli studenti delle suddette classi, ossia 2.862.757 alunni. L'ICC ha coinvolto 184.280 allievi in 9.047 classi.

**Tavola 1.1 - Totale delle classi e degli studenti per livello coinvolti nella Rilevazione per l'anno scolastico 2012-2013**

LIVELLO SCOLARE	TOTALE CLASSI	TOTALE CLASSI CAMPIONE	TOTALE STUDENTI	TOTALE STUDENTI CAMPIONE INDAGINE ICC
2°: II PRIMARIA	29.391	1.679	560.140	31.999
5°: V PRIMARIA	29.726	1.662	557.995	31.198
6°: I SECONDARIA PRIMO GRADO	27.082	1.711	590.728	37.321
8°: III SECONDARIA PRIMO GRADO (Prova nazionale)	29.385	1.420	593.407	28.676
10°: II SECONDARIA SECONDO GRADO	26.200	2.575	560.487	55.086
<b>TOTALE</b>	<b>141.784</b>	<b>9.047</b>	<b>2.862.757</b>	<b>184.280</b>

Il presente rapporto è finalizzato a illustrare le principali caratteristiche dell'Indagine Campionaria di Controllo ed è articolato nel modo seguente: i capitoli 2, 3 e 4 descrivono rispettivamente il sistema delle indagini, le tecniche di misurazione utilizzate e gli obiettivi statistici che vengono perseguiti. I capitoli 5, 6 e 7 approfondiscono le parti relative al campionamento, alla produzione delle stime e al calcolo degli errori campionari. Infine, il capitolo 8 esamina la problematica del *cheating*.

<sup>6</sup> Con il termine *cheating* in ambito educativo si indicano le pratiche messe in atto dagli studenti (*student cheating*) o dagli insegnanti (*teacher cheating*) per alterare, sovrastimandoli, i risultati delle prove standardizzate di valutazione.

## 2. LE RILEVAZIONI INVALSI

Analogamente a quanto è avvenuto in molti paesi industrializzati, anche l'Italia si è dotata abbastanza recentemente di un Servizio Nazionale di Valutazione. Dopo una fase sperimentale iniziata nei primi anni del 2000, a partire dal 2008 la rilevazione standardizzata di alcuni apprendimenti fondamentali, la comprensione della lettura e la matematica, ha assunto carattere di stabilità. L'attuale assetto del Servizio Nazionale di Valutazione ha avuto inizio con l'introduzione della Prova Nazionale all'interno dell'esame di Stato conclusivo del primo ciclo, passando poi alla somministrazione delle prove standardizzate in tutto il Paese, prima nelle classi II e V della scuola primaria, poi anche nella classe I della scuola secondaria di primo grado (fino all'a.s. 2012-13) per poi andare a pieno regime a partire dall'anno scolastico 2010-2011 con il coinvolgimento della scuola secondaria di secondo grado.

L'INVALSI è stato chiamato a effettuare precise scelte tecniche e strategiche per garantire la realizzazione di un Servizio Nazionale di Valutazione in grado di conseguire l'importante obiettivo della misurazione dei livelli di apprendimento in alcune aree disciplinari cruciali per la crescita individuale degli studenti e generale del Paese. Analogamente alle più importanti esperienze internazionali (OCSE-PISA, IEA-TIMSS, IEA-PIRLS, ecc.) e nazionali il Servizio Nazionale di Valutazione prende avvio dalla definizione dei quadri di riferimento per la valutazione (QdR), resi pubblici diverso tempo prima delle rilevazioni. I QdR rivestono un ruolo fondamentale per la costruzione delle prove standardizzate da somministrare agli allievi. Mediante il quadro di riferimento per la valutazione vengono definiti gli ambiti, i processi cognitivi e i compiti oggetto di rilevazione, delimitando quindi il campo rispetto al quale sono costruite le prove. I QdR, inoltre, sono molto importanti per la Scuola poiché essi permettono a ciascuna istituzione scolastica di mettere in relazione ciò che viene rilevato mediante le prove standardizzate con quanto viene insegnato secondo modi e metodi autonomamente scelti.

## 3. TECNICA DI MISURAZIONE

La somministrazione delle prove INVALSI viene effettuata in genere nell'arco temporale compreso tra maggio e giugno. Per l'anno scolastico 2012-13 le prove sono iniziate il 7 maggio e sono terminate il 17 giugno con la Prova Nazionale. Il tempo previsto per ciascuna prova è differenziato in base al livello scolare. Di seguito si riportano i tempi di somministrazione di ciascuna prova e del questionario-studente, ove previsto.

**Tavola 3.1 - Tempi di somministrazione**

Prova	II primaria	V primaria	I sec. I gr.	III sec. I gr.	II sec. II gr.
Prova preliminare di lettura	2 min.	-	-	-	-
Prova di Italiano	45 min.	75 min.	75 min.	75 min.	90 min.
Prova di Matematica	45 min.	75 min.	75 min.	75 min.	90 min.
Questionario studente	-	30 min.	30 min.	-	30 min.

Come già avvenuto nelle rilevazioni precedenti, nella scuola primaria, vista l'età degli alunni coinvolti, le prove si sono svolte in due giornate distinte al fine di evitare l'effetto affaticamento che si sarebbe potuto verificare se le prove, come negli altri livelli scolastici, fossero state somministrate nello stesso giorno, una di seguito all'altra.

Alle prove hanno partecipato tutti gli studenti delle classi interessate<sup>7</sup>. Per quanto riguarda, in particolare, gli alunni con bisogni educativi speciali, le modalità di partecipazione alle prove variano a seconda che si tratti delle prove del Servizio Nazionale di Valutazione o della Prova Nazionale (PN, relativa alla classe terza della scuola secondaria di primo grado). Per quest'ultima, infatti, trovano applicazione le norme vigenti, mentre per le prove del Servizio Nazionale di Valutazione<sup>8</sup>, per le quali non è prevista l'attribuzione di alcun voto, non avendo esse lo scopo di valutare i singoli alunni, la decisione se e come far partecipare gli allievi con particolari bisogni educativi è demandata al Dirigente Scolastico e al consiglio di classe che, alla luce della reale situazione dello studente e da quanto indicato da una specifica nota congiunta MIUR-INVALSI, possono adottare tutte le misure idonee per tutelare sia le esigenze dell'allievo sia il regolare svolgimento delle prove per gli altri studenti, senza che venga modificato il protocollo di somministrazione, il cui rispetto è essenziale trattandosi di prove standardizzate.

Al termine di ciascuna giornata di somministrazione, l'INVALSI invia, tramite *e-mail*, a tutte le scuole le griglie di correzione delle prove. Per la Prova Nazionale che, così come richiamato dalla C.M. 48 del 31 maggio 2012, concorre alla definizione del voto finale dell'esame di Stato, la griglia di correzione, unitamente ai criteri per l'attribuzione del voto, viene spedita alle ore 12.00 del giorno della PN, sia alle scuole, sia agli Uffici scolastici regionali e territoriali, nonché pubblicata sul sito INVALSI.

<sup>7</sup> A causa di un concomitante sciopero indetto da alcuni sigle sindacali, in pochi casi, un numero molto limitato di classi, mai superiore all'1,5% delle classi campione, non ha preso parte alle rilevazioni.

<sup>8</sup> Cfr. [http://www.invalsi.it/snvpn2013/documenti/materiale\\_info/Nota\\_alunni\\_con\\_particolari\\_bisogni\\_educativi\\_2013.pdf](http://www.invalsi.it/snvpn2013/documenti/materiale_info/Nota_alunni_con_particolari_bisogni_educativi_2013.pdf).

## 4. OBIETTIVI DELL'INDAGINE

L'ICC ha la finalità di fornire stime di:

- valori medi dei punteggi riportati alle prove;
- differenti componenti della devianza tra i punteggi, distinguendo la variabilità individuale all'interno della classe, da quella tra classi dello stesso Istituto e quella tra Istituti differenti.

Le stime di cui sopra devono essere fornite separatamente per ciascuna tipologia di prova (*Italiano e Matematica*) e per *domini di stima*<sup>9</sup>.

Nella rilevazione in oggetto, i domini di stima sono individuati dall'incrocio tra le diverse aree territoriali e le sottopopolazioni di alunni definite dalla *tipologia di scuola* e dal livello scolare di interesse (vedi Tavola 4.1).

I principali domini di studio sono i 22 domini territoriali:

- l'Italia;
- le province autonome di Trento e Bolzano (lingua italiana)<sup>10</sup>, le 20 regioni e l'Italia.

Si noti che le stime possono essere anche prodotte per ripartizione geografica. Tuttavia il disegno di campionamento di seguito illustrato è realizzato considerando come prioritari i 22 domini sopra citati.

Si distinguono 5 *tipologie di scuola*:

1. Scuola primaria
2. Scuola secondaria di primo grado
3. Licei
4. Istituti Tecnici
5. Istituti Professionali.

I livelli scolari oggetto di rilevazione sono:

- il secondo e il quinto anno di corso della scuola primaria;

<sup>9</sup> I domini di stima sono sotto-insiemi della popolazione oggetto d'indagine (gli alunni della scuola primaria e secondaria in Italia) per i quali si vuole fissare a priori la dimensione del campione, ossia quanti alunni appartenenti a ciascun specifico dominio saranno rilevati nell'indagine campionaria.

<sup>10</sup> Nel seguito per semplicità espositiva, il termine regione sarà usato in modo generico, indicando anche la provincia autonoma.

- il sesto e l'ottavo, relativi al primo e al terzo anno di corso della scuola secondaria di primo grado;
- il decimo anno scolastico, relativo al secondo anno di corso della scuola secondaria di secondo grado.

L'incrocio tra livelli scolari e tipologia di scuola porta a definire 7 distinte sottopopolazioni.

**Tavola 4.1 - Sottopopolazioni individuate dal livello scolastico e dalla tipologia di scuola**

Tipologia di scuola	Livello scolastico	Sottopopolazione
Scuola primaria	2°	1
	5°	2
Scuola secondaria di primo grado	6°	3
	8°	4
Licei	10°	5
Istituti Tecnici	10°	6
Istituti Professionali	10°	7

In totale quindi, l'ICC deve produrre stime per 154 domini di stima, dove tale numero è ottenuto dal prodotto di 22 (domini territoriali) per 7 (sottopopolazioni).

#### 4.1 Notazione di base

Al fine di descrivere formalmente i parametri di interesse è utile considerare la notazione sotto riportata. Si indichi con:

- $t$  la tipologia di scuola:  $t=1, \dots, 5$ ;
- $r$  il singolo anno scolastico:  $r=1, \dots, 13$ ; con  $r=1, \dots, 5$  per  $t=1$ ;  $r=6, 7, 8$  per  $t=2$ ;  $r=9, \dots, 13$  per  $t>2$ ;
- $h$  la regione o la provincia autonoma;
- $i$  il plesso scolastico (in seguito indicato in modo generico anche come scuola);
- $j$  la sezione nell'ambito dell'Istituto;
- $k$  il singolo alunno;
- $A$  l'insieme degli alunni e il numero di alunni;
- $S$  l'insieme delle strutture scolastiche. A livello Italia o a livello regionale,  $S$  denota l'insieme delle scuole; quando si considera una specifica scuola, indica i plessi o le sezioni nell'ambito della stessa e anche il numero delle strutture stesse.

$A$  e  $S$  sono ulteriormente denotati mediante i pedici  $t$  e  $r$  (nell'ordine) alla loro sinistra e con gli indici  $h$ ,  $i$  e  $j$  (sempre nell'ordine) alla loro destra. La mancanza di un indice denota che la quantità rappresenta il totale rispetto all'indice mancante. Al fine di rendere più chiaro l'utilizzo della notazione introdotta, ne riportiamo alcuni esempi nel riquadro seguente.

**Box 4.1 - Esempi di utilizzo della notazione**

- ✓  ${}_{tr}A$  indica il numero (e la popolazione corrispondente) degli scolari che in Italia frequentano l'anno  $r$  nella tipologia di scuola  $t$ ;
- ✓  ${}_{tr}S$  denota il numero (e la popolazione corrispondente) dei plessi scolastici per l'anno  $r$  nella tipologia di scuola  $t$ ;
- ✓  ${}_{tr}A_{(h)}$  rappresenta il numero (e la popolazione corrispondente) degli scolari che in Italia frequentano l'anno  $r$  nella tipologia di scuola  $t$  nella regione o provincia autonoma ( $h$ );
- ✓  ${}_tA_{(h)ij}$  indica il numero degli alunni che frequentano la sezione  $j$  nella scuola  $i$  di tipologia  $t$  nella regione  $h$ ;
- ✓  ${}_tS_{(h)i}$  indica il numero delle sezioni nella scuola  $i$  di tipologia  $t$  nella regione  $h$ .

Dato  ${}_{tr}y_{(h)ijk}$  il punteggio (o votazione alla prova)  $y$  ottenuto dall'alunno  $k$  che frequenta l'anno  $r$  nella sezione  $j$  della scuola  $i$  di tipologia  $t$  nella regione  $h$ , si denota con

$${}_{tr}\bar{y}_{(h)ij} = \frac{{}_{tr}Y_{(h)ij}}{{}_{tr}A_{(h)ij}}, \quad {}_{tr}\bar{y}_{(h)i} = \frac{{}_{tr}Y_{(h)i}}{{}_{tr}A_{(h)i}},$$

rispettivamente il valore medio del punteggio della prova  $y$  per l'anno scolastico  $r$  per gli scolari della sezione  $j$  della scuola  $i$  e il valore medio per la scuola, dove

$${}_{tr}Y_{(h)ij} = \sum_{k=1}^{{}_{tr}A_{(h)ij}} {}_{tr}y_{(h)ijk}, \quad {}_{tr}Y_{(h)i} = \sum_{k=1}^{{}_{tr}S_{(h)i}} {}_{tr}y_{(h)ijk},$$

rispettivi totali del punteggio nella sezione  $j$  e nella scuola  $i$ .

## 4.2 Punteggi medi

A partire dalla notazione precedentemente introdotta, i parametri d'interesse dell'ICC relativi ai valori medi dei punteggi regionali e nazionali alle prove sono espressi rispettivamente come:

$$\bullet \quad {}_{tr}\bar{y}(h) = \frac{1}{{}_{tr}A(h)} \sum_{i=1}^{{}_{tr}S(h)} \sum_{j=1}^{{}_{tr}S(h)i} \sum_{k=1}^{{}_{tr}A(h)ij} {}_{tr}y(h)ijk, \quad t = 1, \dots, 5; r \in {}_tR; h=1, \dots, 21 \quad (4.1)$$

$$\bullet \quad {}_{tr}\bar{y} = \frac{1}{{}_{tr}A} \sum_{h=1}^{21} \sum_{i=1}^{{}_{tr}S(h)} \sum_{j=1}^{{}_{tr}S(h)i} \sum_{k=1}^{{}_{tr}A(h)ij} {}_{tr}y(h)ijk \quad t = 1, \dots, 5; r \in {}_tR; \quad (4.2)$$

con  ${}_tR: \{r=2, 5 \text{ per } t=1; r=6, 8 \text{ per } t=2; r=10 \text{ per } t>2\}$ .

In appendice A.1 sono riportate alcune utili riformulazioni dei parametri precedenti.

## 4.3 Componenti della devianza

La devianza complessiva,  ${}_{tr}D$  tra i punteggi  ${}_{tr}y(h)ijk$  degli allievi che frequentano l'anno scolastico  $r$  nella tipologia di scuola  $t$  è data da:

$${}_{tr}D = \sum_{h=1}^{21} \sum_{i=1}^{{}_{tr}S(h)} \sum_{j=1}^{{}_{tr}S(h)i} \sum_{k=1}^{{}_{tr}A(h)ij} ({}_{tr}y(h)ijk - {}_{tr}\bar{y})^2 \quad t = 1, \dots, 5; r \in {}_tR \quad (4.3)$$

Aggiungendo e sottraendo all'interno della differenza quadratica i valori medi  ${}_{tr}\bar{y}(h)ij$  e  ${}_{tr}\bar{y}(h)i$ , la devianza  ${}_{tr}D$  può essere espressa come somma di tre componenti (vedi appendice A.2):

$${}_{tr}D = {}_{tr}D1 + {}_{tr}D2 + {}_{tr}D3$$

che denotano rispettivamente la devianza tra scuole ( ${}_{tr}D1$ ), quella tra classi della medesima scuola ( ${}_{tr}D2$ ) e quella tra gli alunni di una stessa classe ( ${}_{tr}D3$ ), dove

$${}_{tr}D1 = \sum_{h=1}^{21} \sum_{i=1}^{{}_{tr}S(h)} {}_{tr}A(h)i ({}_{tr}\bar{y}(h)i - {}_{tr}\bar{y})^2$$

$${}_{tr}D2 = \sum_{h=1}^{21} \sum_{i=1}^{{}_{tr}S(h)} \sum_{j=1}^{{}_{tr}S(h)i} {}_{tr}A(h)ij ({}_{tr}\bar{y}(h)ij - {}_{tr}\bar{y}(h)i)^2$$

$${}_{tr}D3 = \sum_{h=1}^{21} \sum_{i=1}^{{}_{tr}S_{(h)}} \sum_{j=1}^{{}_{tr}S_{(h)i}} \sum_{k=1}^{{}_{tr}A_{(h)ij}} ({}_{tr}y_{(h)ijk} - {}_{tr}\bar{y}_{(h)ij})^2 .$$

Le espressioni della devianza e delle relative componenti a livello regionale sono determinate mediante le espressioni precedenti, limitando le sommatorie alla specifica regione.

L'analisi delle componenti della devianza fornisce una chiave di interpretazione degli esiti delle prove standardizzate di considerevole rilevanza. In un sistema scolastico "ideale", la devianza totale  ${}_{tr}D$  dovrebbe essere spiegata dalle differenze osservate tra gli studenti ( ${}_{tr}D3$ ) e non dalla loro suddivisione nelle classi ( ${}_{tr}D2$ ) o nelle scuole ( ${}_{tr}D1$ ). Infatti, tanto più un sistema è caratterizzato da bassa devianza tra le classi e tra le scuole, tanto maggiore è l'equità di quel sistema, ipotizzando che i risultati medi generali siano elevati. Infatti, in questo caso l'incidenza dell'inserimento in una data classe o in una data scuola ha poco peso e buoni risultati sono garantiti a tutti e a ciascuno, sostanzialmente in ragione delle caratteristiche individuali dell'allievo, indipendentemente dalla particolare scuola o classe frequentata.

## 5. CAMPIONAMENTO

Al fine di stimare i parametri  ${}_{tr}\bar{y}_{(h)}$  e  ${}_{tr}\bar{y}$ , si estrae un campione probabilistico a due stadi di selezione, stratificato al primo stadio, in cui le unità del primo stadio sono i plessi scolastici (le scuole) e quelle del secondo stadio le sezioni.

### 5.1 Primo stadio di campionamento

Le scuole sono stratificate per *tipologia, t, e regione, h*.

Questo tipo di stratificazione garantisce che i 154 domini di stima corrispondano a uno strato o possano essere ottenuti mediante aggregazioni degli strati dell'indagine. In tal modo è possibile pianificare la numerosità campionaria nei domini di stima e controllare in fase di disegno l'errore di campionamento.

Nell'ambito dello specifico strato,  ${}_tS_{(h)}$ , si seleziona senza reimmissione e con probabilità variabili un campione probabilistico,  ${}_tS_{(h)}$ , di istituti scolastici (tale simbolo denota anche la dimensione del campione stesso).

La  $i$ -esima scuola dello strato viene selezionata con probabilità di inclusione,  ${}_t\pi_{(h)i}$ , proporzionale al numero di alunni della scuola:

$${}_t\pi_{(h)i} = {}_tS_{(h)} {}_t\alpha_{(h)i} \tag{5.1}$$

dove

$${}_t\alpha_{(h)i} = \frac{{}_t\ddot{A}_{(h)i}}{\sum_{i=1} {}_tS_{(h)} {}_t\ddot{A}_{(h)i}} \tag{5.1b}$$

in cui

- ${}_t\ddot{A}_{(h)i} = \left[ \frac{1}{2} ({}_t2A_{(h)i} + {}_t5A_{(h)i}) \right] \frac{2}{{}_tS_{(h)i}}$  per  $t=1$ ,
- ${}_t\ddot{A}_{(h)i} = {}_t6A_{(h)i} \frac{2}{{}_tS_{(h)i}}$  per  $t=2$ ,
- ${}_t\ddot{A}_{(h)i} = {}_t10A_{(h)i} \frac{2}{{}_tS_{(h)i}}$  per  $t > 2$

e

$${}_tS_{(h)i} = \min (2, {}_tS_{(h)i}). \tag{5.2}$$

Il rapporto  $2/{}_tS_{(h)i}$ , pari a 1 per gli istituti scolastici con almeno due sezioni e uguale a 2 per le piccole scuole con una sola sezione, raddoppia la probabilità di inclusione dei piccoli istituti scolastici assicurando una uguale probabilità di inclusione finale a tutti gli alunni e a tutte le sezioni di un medesimo strato (vedi paragrafo 5.3) a prescindere dalla dimensione, in termini di alunni, della scuola.

Il campione è selezionato con il metodo di Pareto (Särndal e Lundström, 2005), mediante una procedura articolata nei seguenti passi:

- a) per ogni scuola  $i$ , si genera la variabile casuale  ${}_t\omega_{(h)i}$  come realizzazione di una distribuzione uniforme  $Unif(0,1)$  e si calcola

$${}_t q_{(h)i} = \frac{{}_t \omega_{(h)i} (1 - {}_t \pi_{(h)i})}{{}_t \pi_{(h)i} (1 - {}_t \omega_{(h)i})}$$

b) il campione è costituito dalle  ${}_t s_{(h)}$  scuole a cui corrispondono i valori più bassi delle quantità  ${}_t q_{(h)i}$ .

Il numero atteso di classi campione in ciascuno strato è pari a

$${}_t \bar{S}_{(h)} = \sum_{i=1}^{{}_t S_{(h)}} {}_t \pi_{(h)i} \cdot {}_t S_{(h)i} .$$

## 5.2 Secondo stadio

Nella scuola campione  $(h)i$  della tipologia  $t$  viene selezionato, a probabilità uguali e senza reimmissione, un campione di  ${}_t s_{(h)i}$  sezioni. In tal modo, in ciascuna scuola si seleziona un campione di 2 sezioni, a eccezione delle scuole con un'unica sezione che viene selezionata con probabilità pari a 1 nel campione di secondo stadio.

Pertanto, la singola sezione viene selezionata con probabilità di secondo stadio pari a:

$${}_t \pi_{(h)j|i} = \frac{{}_t S_{(h)i}}{{}_t S_{(h)i}} \tag{5.3}$$

Nelle sezioni selezionate nel campione vengono sottoposti alla prova gli alunni delle classi i cui livelli scolari sono d'interesse per la rilevazione. Di conseguenza, tutti gli  ${}_{tr} A_{(h)ij}$  alunni appartenenti all'anno di interesse  $r (r \in {}_t R)$  per la tipologia  $t$  della sezione estratta vengono inclusi nel campione.

Il totale degli alunni estratti nel campione nella scuola campione per l'anno scolare  $r$  è

$${}_{tr} a_{(h)i} = \sum_{j \in {}_t s_{(h)i}} {}_{tr} A_{(h)ij} .$$

Il valore atteso di alunni campione a livello regionale e nazionale è pari a:

$${}_{tr}\bar{a}_{(h)} = \sum_{i=1} {}_tS_{(h)} \sum_{j=1} {}_tS_{(h)i} {}_t\pi_{(h)i} {}_t\pi_{(h)j|i} {}_{tr}A_{(h)ij} .$$

### 5.3 Probabilità di inclusione finale

Il singolo alunno  $k$  che frequenta l'anno  $r$  nella scuola  $i$  della tipologia  $t$  dello strato  $h$  viene selezionato con probabilità di inclusione finale pari a:

$${}_{tr}\pi_{(h)ijk} = {}_t\pi_{(h)i} {}_t\pi_{(h)j|i} = {}_tS_{(h)} \frac{{}_t\ddot{A}_{(h)i}}{\sum_{i=1} {}_tS_{(h)} {}_t\ddot{A}_{(h)i}} \frac{{}_tS_{(h)i}}{{}_tS_{(h)}} . \quad (5.4)$$

Nell'ipotesi che la dimensione delle classi in termini di alunni sia pressappoco costante nello strato e che ogni classe sia approssimativamente costituita da  ${}_t\bar{A}_{(h)}$  alunni, si dimostra che vale la seguente condizione

$${}_{tr}\pi_{(h)ijk} \cong \frac{2 {}_tS_{(h)}}{{}_t\bar{S}_{(h)}} \quad (5.5)$$

in cui  ${}_t\bar{S}_{(h)} = {}_{(>1)}{}_t\bar{S}_{(h)} + 2 {}_{(=1)}{}_t\bar{S}_{(h)}$ , dove  ${}_{(>1)}{}_t\bar{S}_{(h)}$  denota il numero delle sezioni appartenenti alle scuole con almeno due sezioni e  ${}_{(=1)}{}_t\bar{S}_{(h)}$  indica il numero delle sezioni nello strato appartenenti alle scuole con un'unica sezione.

La (5.5) è una relazione importante in quanto dimostra che con il criterio di selezione prescelto ciascun alunno e ciascuna classe dello strato hanno la medesima probabilità di essere selezionati nel campione, indipendentemente dalla dimensione della scuola selezionata nel primo stadio.

La condizione di *autoponderazione*, approssimativa, espressa dalla (5.5) rappresenta un fattore di *ottimalità del disegno* in quanto tende ad annullare l'incremento di variabilità derivante dalla variabilità dei pesi finali.

## 5.4 Allocazione del campione

Il numero di scuole a livello nazionale della tipologia  $t$ ,  ${}_t s$ , è definito cercando di individuare la dimensione che preveda costi inferiori al *budget* disponibile e che al contempo sia sostenibile dalla complessa organizzazione logistica che è necessario mettere in atto per tutte le fasi della rilevazione.

Tale analisi ha portato a definire una dimensione campionaria,  ${}_t s$ , pari a 750 scuole per tipologia di scuola.

Questo numero è stato suddiviso tra due sottoinsiemi di regioni:

- l'insieme delle regioni Auto Rappresentative (AR), costituito dalle regioni di piccole dimensioni, meno di 20 Istituti scolastici; in particolare l'insieme AR comprende la Valle d'Aosta e la Provincia autonoma di Bolzano (lingua italiana);
- l'insieme delle regioni Non Auto Rappresentative (NAR), costituito dalle rimanenti 19 regioni/province.

All'insieme AR viene assegnato un campione di  ${}_t s_{AR}$  scuole, pari a

$${}_t s_{AR} = {}_t S_{AR} = \sum_{h \in AR} {}_t s(h).$$

All'insieme NAR, con numero di scuole pari a  ${}_t s_{NAR} = {}_t s - {}_t s_{AR}$ , viene assegnata la parte rimanente del campione nazionale  ${}_t s_{NAR} = {}_t s - {}_t s_{AR}$ .

Il campione  ${}_t s_{NAR}$  viene quindi allocato tra le regioni dell'insieme *NAR* mediante una procedura di compromesso che tenta di individuare una soluzione ragionevole tra l'esigenza di:

- trovare un criterio ottimale che porti a minimizzare la varianza della stima nazionale,  $V({}_{tr} \hat{y})$ ;
- garantire una varianza di campionamento approssimativamente uguale alle stime regionali  $V({}_{tr} \hat{y}(h))$ .

Fissata la dimensione del campione di scuole a livello nazionale,  ${}_t S_{NAR}$

- il criterio al punto (a) porta a definire un campione regionale proporzionale al numero di scuole della regione:

$${}_t S_{(h)} = {}_t S_{NAR} \frac{{}_t S_{(h)}}{{}_t S_{NAR}},$$

- il criterio al punto (b) porta a definire un campione approssimativamente uguale a livello di regione:

$${}_t S_{(h)} = \frac{{}_t S_{NAR}}{19}.$$

Il criterio di compromesso si fonda sulla definizione di un valore  ${}_t \beta$  compreso tra 0 e 1 che definisce la dimensione del campione mediante la seguente combinazione lineare convessa:

$${}_t S_{(h)} = \beta_t \left( {}_t S_{NAR} \frac{{}_t S_{(h)}}{{}_t S_{NAR}} \right) + (1 - \beta_t) {}_t S_{NAR} \frac{{}_t S_{(h)}}{{}_t S_{NAR}} \quad (5.6)$$

Il valore di  ${}_t \beta$  viene definito in modo tale che la dimensione campionaria assicuri che l'errore di campionamento a livello nazionale e regionale rimanga sotto determinate soglie ritenute accettabili per la pubblicazione dei dati:

$${}_t \beta : \begin{cases} CV({}_{tr} \hat{y}(h)) \leq {}_t \overline{CVH} & h \in NAR \\ CV({}_{tr} \hat{y}) \leq {}_t \overline{CV} \end{cases}$$

in cui  $CV({}_{tr} \hat{y}(h))$  e  $CV({}_{tr} \hat{y})$  denotano i coefficienti di variazione della stima a livello regionale e nazionale e  ${}_t \overline{CVH}$  e  ${}_t \overline{CV}$  denotano i valori soglia di tali coefficienti ritenuti accettabili per l'allocazione, essendo

$$CV({}_{tr} \hat{y}(h)) = \sqrt{V({}_{tr} \hat{y}(h))} / {}_{tr} \hat{y}(h) \quad (5.7)$$

e

$$CV({}_{tr}\hat{y}) = \sqrt{V({}_{tr}\hat{y})} / {}_{tr}\hat{y}. \quad (5.8)$$

Dal punto di vista computazionale, i valori per il calcolo dei coefficienti di variazione sono desunti dalla rilevazione dell'anno precedente. Le formule di calcolo adottate sono illustrate nel seguente paragrafo 6. La variabile  $y$  presa a riferimento per il calcolo delle numerosità campionarie è la media dei punteggi complessivi. In genere, il valore di  ${}_t\beta$  prescelto si attesta intorno a 0,3.

Le allocazioni campionarie per l'anno 2013 per le varie tipologie di scuole evidenziano una similarità di fondo; pertanto, ci limiteremo alla discussione di quella relativa alla scuola secondaria di primo grado di seguito riportata.

**Tavola 5.1 - Allocazione del campione nella scuola secondaria di primo grado classi I: anno 2013**

Regione e/o Provincia autonoma	Totale Scuole	Totale Alunni	Campione						
			Allocazione proporzionale tra le regioni		Allocazione uguale tra le regioni		Allocazione di compromesso ${}_t\beta = 0,3$		
			Scuole	CV%	Scuole	CV%	Scuole	Classi	CV%
Valle d'Aosta	17	1.143	17	1,98	17	1,98	17	33	1,98
Piemonte	388	35.010	50	2,19	38	2,55	41	82	2,43
Liguria	146	11.224	15	3,90	38	2,33	31	61	2,63
Lombardia	957	84.150	123	1,44	38	2,70	63	125	2,06
Prov. Aut. di Bolzano (l. it.)	18	1.038	18	1,38	18	1,38	18	31	1,38
Prov. Aut. di Trento	63	4.993	5	5,20	38	1,52	28	56	1,94
Veneto	486	41.891	48	1,83	38	2,07	41	82	1,99
Friuli-Venezia Giulia	121	9.624	15	4,29	38	2,45	31	62	2,80
Emilia-Romagna	348	35.518	40	2,00	38	2,05	38	76	2,04
Toscana	322	27.047	32	2,37	38	2,16	36	72	2,22
Umbria	79	7.239	8	4,37	38	1,74	29	67	2,08
Marche	160	12.740	16	3,46	38	2,10	31	60	2,36
Lazio	586	45.628	73	2,06	38	2,97	48	93	2,60
Abruzzo	144	11.263	13	3,65	38	1,94	30	59	2,23
Molise	47	2.686	5	9,35	38	1,22	28	95	2,45
Campania	643	63.586	100	1,89	38	3,18	56	112	2,58
Puglia	377	41.498	65	2,22	38	3,01	46	92	2,70
Basilicata	90	5.491	7	5,82	38	2,20	29	64	2,66
Calabria	276	19.147	25	3,27	38	2,60	34	66	2,76
Sicilia	579	48.495	66	2,16	38	2,91	46	91	2,61
Sardegna	195	10.963	10	3,56	38	1,74	29	58	2,01
<b>Italia</b>	<b>6.042</b>	<b>520.374</b>	<b>750</b>	<b>0,59</b>	<b>750</b>	<b>0,80</b>	<b>750</b>	<b>1.537</b>	<b>0,69</b>

Si notano i seguenti aspetti:

- l'allocazione proporzionale, porta al valore minimo del coefficiente di variazione della stima nazionale (pari allo 0,59%) privilegiando le grandi regioni rispetto a regioni di

piccole e medie dimensioni. Vi sono cinque regioni con valori del CV% superiore al 3% (celle in giallo) e cinque regioni hanno un CV superiore al 4 e al 5% (celle in rosso);

- l'allocazione uguale, conduce a un notevole peggioramento della stima nazionale (pari allo 0,80%). D'altra parte, vi è un maggiore equilibrio negli errori delle stime regionali. Due sole regioni presentano un CV superiore al 3% (celle gialle) e vi sono 12 regioni (celle celesti) con CV compreso tra il 2 e il 3%;
- l'allocazione di compromesso è un'ottima soluzione, in quanto viene controllato il CV della stima nazionale e, al contempo, presenta il medesimo profilo dell'allocazione equiripartita;
- in alcune regioni AR per diminuire ulteriormente l'errore di campionamento si sono selezionate più di due classi per istituto. (tale caso, è capitato raramente e solo nel campionamento delle scuole superiori).

## 6. STIME CAMPIONARIE

### 6.1 Descrizione generale

A ciascun alunno viene attribuito un peso campionario che indica quanti scolari, non coinvolti nell'ICC sono rappresentati dall'alunno incluso nel campione. Ad esempio, un peso pari a 50, indica che l'alunno rappresenta sé stesso più altri 49 scolari non coinvolti nell'indagine.

Il peso campionario (o semplicemente peso) di un alunno viene ottenuto moltiplicando:

- il peso della scuola (peso di primo stadio);
- il peso della sezione all'interno della scuola (peso di secondo stadio).

La somma del peso di secondo stadio su tutti gli alunni campione della scuola fa ritornare il numero di alunni della scuola.

Il peso di primo stadio viene ottenuto con una procedura di calibrazione che assicura che:

- la somma del peso di primo stadio su tutti i plessi campione di una determinata tipologia riproduce il numero dei plessi scolastici di quella determinata tipologia nella regione;

- la somma su tutti i plessi campione di una determinata tipologia dei prodotti del peso di primo stadio per il numero di scolari del plesso che frequentano il livello scolare  $r$  coincide il numero di alunni nello strato che frequentano il livello scolare  $r$ .

## 6.2 Peso di secondo stadio

Il peso di secondo stadio per la sezione  $j$  selezionata nella scuola  $i$  è dato da

$${}_{tr}w_{(h)j|i} = \frac{1}{{}_t\pi_{(h)j|i}} \frac{{}_{tr}A_{(h)i}}{\sum_{j=1}^{{}_tS_{(h)i}} \frac{1}{{}_t\pi_{(h)j|i}} {}_{tr}A_{(h)ij}^*} \quad (6.1)$$

in cui  ${}_{tr}A_{(h)ij}^*$  denota il numero di alunni rispondenti al campione per l'anno scolare nella sezione  $j$  della scuola  $i$  di tipologia  $t$  nella regione  $h$ .

## 6.3 Peso di primo stadio

Il peso finale di primo stadio,  ${}_{tr}w_{(h)i}$ , viene ottenuto risolvendo il seguente problema di minimo vincolato

$${}_{tr}w_{(h)i} = \begin{cases} \sum_{i=1}^{{}_tS_{(h)}} D\left(\frac{1}{{}_t\pi_{(h)i}}, {}_{tr}w_{(h)i}\right) = \min \\ \sum_{i=1}^{{}_tS_{(h)}} {}_{tr}w_{(h)i} = {}_tS_{(h)} \\ \sum_{i=1}^{{}_tS_{(h)}} {}_{tr}w_{(h)i} {}_{tr}A_{(h)i} = {}_{tr}A_{(h)} \end{cases} \quad (6.2)$$

in cui  $D\left(\frac{1}{{}_t\pi_{(h)i}}, {}_{tr}w_{(h)i}\right) = {}_{tr}w_{(h)i} \ln\left(\frac{{}_{tr}w_{(h)i}}{{}_t\pi_{(h)i}}\right) - {}_{tr}w_{(h)i} + \frac{1}{{}_t\pi_{(h)i}}$  è la distanza logaritmica

tra il peso iniziale,  $\frac{1}{{}_t\pi_{(h)i}}$ , e quello finale  ${}_{tr}w_{(h)i}$ .

L'algoritmo per la soluzione del problema (6.2) è illustrato in appendice A.4.

## 6.4 Peso alunno

Il peso finale dell'alunno  $ijk$  è dato da:

$$tr^{W(h)ijk} = tr^{W(h)i} tr^{W(h)j|i}.$$

### 6.5 Stima dei punteggi medi

Le stime sono ottenute mediante lo stimatore *rapporto separato* per strato. I valori medi regionali sono calcolati come

$$tr \hat{y}(h) = \frac{tr \hat{Y}(h)}{tr \hat{A}(h)} \tag{6.3}$$

dove

$$tr \hat{Y}(h) = \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A^*(h)ij} tr Y(h)ijk tr W(h)ijk$$

$$tr \hat{A}(h) = \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A^*(h)ij} tr W(h)ijk .$$

I valori medi nazionali sono quindi stimati:

$$tr \hat{y} = \frac{tr \hat{Y}}{tr \hat{A}} \tag{6.4}$$

essendo

$$tr \hat{Y} = \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A^*(h)ij} tr Y(h)ijk tr W(h)ijk$$

$$tr \hat{A} = \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A^*(h)ij} tr W(h)ijk .$$

Sotto la (6.2) si ha  $tr \hat{A}(h) = tr A(h)$  e  $tr \hat{A} = tr A$ .

Le stime dei punteggi medi relative all'anno 2013 per ripartizione geografica, livello scolastico e materia indagata sono riportate nella tavola 6.1.

Tavola 6.1 – Punteggi medi per Regione, Livello scolastico e materia indagata

Ripartizione geografica	Livello 2		Livello 5		Livello 6		Livello 8		Licei		Tecnici		Professionali		Livello 10	
	Ita	Mat	Ita	Mat	Ita	Mat	Ita	Mat	Ita	Mat	Ita	Mat	Ita	Mat	Ita	Mat
Nord Ovest	204	203	206	206	208	211	205	210	230	228	204	210	178	180	212	214
Valle d'Aosta	206	204	209	208	204	205	189	210	229	218	206	208	188	189	214	209
Piemonte	207	206	208	211	207	209	203	205	230	236	201	206	176	178	210	215
Liguria	203	201	207	204	203	205	201	209	222	214	191	198	172	172	205	202
Lombardia	203	201	205	204	209	212	206	212	231	228	207	213	180	182	214	215
Nord Est	203	202	204	205	204	207	210	210	228	224	206	217	174	181	210	213
Prov. Aut. Bolzano (l. it.)	201	204	197	200	205	209	201	205	221	222	196	208	181	181	207	211
Prov. Aut. Trento	211	210	207	211	203	210	212	213	226	226	210	228	183	175	218	226
Veneto	202	202	205	205	202	205	211	211	226	224	209	217	182	185	211	214
Friuli-Venezia Giulia	207	208	206	206	209	212	206	214	226	231	206	220	170	182	209	219
Emilia-Romagna	202	201	203	203	203	207	209	207	231	222	201	213	167	176	207	209
Centro	201	203	202	202	203	202	200	203	213	211	193	198	170	173	199	201
Toscana	198	201	204	204	200	203	203	209	212	215	199	204	164	171	198	203
Umbria	202	205	204	203	202	203	208	206	220	219	200	204	169	175	204	206
Marche	207	206	206	206	209	210	209	207	225	221	201	208	171	176	206	207
Lazio	201	204	200	200	203	200	195	199	209	206	184	190	176	174	197	196
Sud	196	196	194	195	196	192	197	188	211	201	182	184	162	167	192	189
Abruzzo	201	200	205	205	201	198	206	205	207	197	187	189	162	173	194	191
Molise	201	199	200	200	200	197	206	184	210	207	180	185	160	174	192	195
Campania	193	193	192	193	192	186	195	182	211	200	177	179	158	164	190	186
Puglia	199	199	194	196	200	199	199	194	212	203	189	189	168	169	195	191
Sud e Isole	194	195	191	190	185	183	186	187	203	192	180	182	153	163	185	183
Basilicata	198	201	197	203	198	197	193	189	210	204	188	193	167	173	194	193
Calabria	192	197	191	189	185	182	191	180	205	199	180	182	157	168	187	187
Sicilia	192	191	188	187	183	182	181	186	202	190	179	180	148	162	183	181
Sardegna	204	205	198	196	192	185	197	197	202	183	182	184	156	160	186	178
Italia	200	200	200	200	200	200	200	200	217	212	194	199	167	172	200	200

## 6.6 Stima delle componenti della devianza

### Devianza totale D

Una stima approssimativamente corretta della devianza totale  $tr \hat{D}_{cor}$  è data da (vedi appendice A.5):

$$tr \hat{D}_{cor} = \sum_{h=1}^{21} \sum_{i=1}^{tr S^{(h)}} tr W_{(h)ijk} \sum_{j=1}^{tr S^{(h)i}} \sum_{k=1}^{tr A_{ij}} y_{ijk}^2 - tr \hat{A} (tr \hat{y}^2 - \hat{V}(tr \hat{y})), \quad (6.5)$$

in cui la stima della varianza campionaria  $\hat{V}(tr \hat{y})$  è definita dall'espressione (7.7).

Componente D3: Devianza tra gli scolari della classe

È immediato dimostrare che una stima corretta di della devianza tra gli alunni della classe è data da

$$tr \hat{D}3 = \sum_{h=1}^{21} \sum_{i=1}^{S(h)} \sum_{j=1}^{S(h)i} \sum_{k=1}^{tr A_{ij}} tr W(h)_{ijk} (tr Y(h)_{ijk} - tr \bar{Y}(h)_{ij})^2 .$$

Componente D1: Devianza tra scuole

La stima approssimata della devianza tra scuole è data da (vedi appendice A.5)

$$tr \hat{D}1_{cor} = tr \hat{D}1 - \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} tr \ddot{V} (tr \hat{Y}(h)_i) + tr A \hat{V} (tr \hat{Y}) , \tag{6.6}$$

dove

$$tr \hat{D}1 = \sum_{h=1}^{21} \sum_{i=1}^{S(h)} tr W(h)_i (tr \hat{A}(h)_i (tr \hat{Y}(h)_i - tr \hat{Y}))^2 ,$$

in cui  $\ddot{V}(\hat{y}_i)$  indica la stima della variabilità a livello d'istituto la cui espressione è riportata nell'appendice A.5.

Componente D2: Devianza tra le classi

La componente D2 può essere stimata per differenza:

$$tr \hat{D}2 = tr \hat{D}_{cor} - (tr \hat{D}1_{cor} + tr \hat{D}3).$$

Le stime delle componenti della devianza relative all'anno 2013 per ripartizione geografica, livello scolare e materia sono riportate nelle seguenti tavole 6.2 e 6.3.

Stime regionali

Le stime della devianza e delle relative componenti a livello regionale sono determinate mediante le espressioni precedenti, limitando le sommatorie alla specifica regione.

**Tavola 6.2 - Componenti della devianza dei Punteggi medi complessivi di Italiano per Ripartizione geografica e Livello scolare<sup>11</sup>**

Ripartizione	Livello 2			Livello 5			Livello 6			Licei			Professionali			Tecnici			Livello 10		
	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3
Nord Ovest	7,8	4,4	87,9	9,5	5,5	85,0	11,7	6,9	81,4	14,1	44,4	41,5	12,9	34,0	53,1	18,0	5,8	76,2	37,9	11,4	50,6
Nord Est	9,2	3,6	87,2	8,5	4,3	87,1	8,2	4,9	86,9	15,4	43,7	40,9	11,6	38,7	49,7	16,8	7,8	75,4	39,2	10,2	50,6
Centro	9,1	4,5	86,3	9,3	5,5	85,3	12,4	2,3	85,3	21,2	14,9	63,9	14,7	47,8	37,5	20,2	10,9	68,9	35,2	6,2	58,7
Sud	16,8	9,0	74,2	16,5	7,8	75,7	13,7	7,1	79,2	18,9	16,5	64,6	9,6	61,9	28,5	21,0	27,6	51,4	41,6	9,1	49,3
Sud e isole	17,1	7,9	74,9	13,2	11,9	74,9	12,5	16,4	71,2	30,3	6,9	62,8	7,4	73,3	19,4	12,7	36,4	50,9	40,1	18,0	41,8
<b>Italia</b>	<b>12,6</b>	<b>5,0</b>	<b>82,4</b>	<b>13,4</b>	<b>4,8</b>	<b>81,8</b>	<b>15,3</b>	<b>3,9</b>	<b>80,8</b>	<b>24,6</b>	<b>22,7</b>	<b>52,7</b>	<b>14,1</b>	<b>51,4</b>	<b>34,5</b>	<b>27,4</b>	<b>8,6</b>	<b>64,0</b>	<b>45,1</b>	<b>4,5</b>	<b>50,4</b>

**Tavola 6.3 - Componenti della devianza dei Punteggi medi complessivi di Matematica per Ripartizione geografica e Livello scolare**

Ripartizione	Livello 2			Livello 5			Livello 6			Licei			Professionali			Tecnici			Livello 10		
	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3	% D1	% D2	% D3
Nord Ovest	8,4	4,1	87,5	12,4	6,0	81,6	12,5	9,2	78,3	29,2	36,3	34,6	10,4	42,8	46,8	22,9	13,8	63,4	43,1	15,4	41,5
Nord Est	10,6	2,7	86,7	9,9	5,5	84,7	7,8	6,5	85,7	28,2	32,1	39,7	9,4	38,4	52,1	19,0	21,7	59,3	37,9	15,2	47,0
Centro	12,3	7,8	79,8	11,1	7,6	81,3	13,1	2,9	84,0	40,9	13,7	45,4	10,6	52,1	37,3	25,5	5,3	69,2	44,8	6,6	48,6
Sud	24,4	8,7	67,0	24,1	9,4	66,5	16,3	9,5	74,2	32,3	10,9	56,8	8,4	66,7	24,9	17,6	31,6	50,8	36,1	16,9	47,1
Sud e isole	25,9	7,9	66,2	22,9	13,2	63,9	12,8	23,5	63,7	38,4	11,0	50,6	8,2	71,8	20,0	12,8	35,9	51,3	31,9	26,6	41,5
<b>Italia</b>	<b>16,6</b>	<b>5,5</b>	<b>77,9</b>	<b>17,6</b>	<b>6,0</b>	<b>76,3</b>	<b>18,3</b>	<b>4,4</b>	<b>77,3</b>	<b>42,0</b>	<b>14,5</b>	<b>43,5</b>	<b>12,2</b>	<b>55,0</b>	<b>32,8</b>	<b>36,0</b>	<b>4,9</b>	<b>59,1</b>	<b>48,7</b>	<b>6,1</b>	<b>45,1</b>

## 7. VARIABILITA' CAMPIONARIA

### 7.1 Espressione della varianza campionaria

Un'espressione approssimata della varianza campionaria dello stimatore definito nel capitolo 6 può essere derivata sulla base dei seguenti risultati

- il teorema 11.2 in Cochran (1977),
- il metodo di approssimazione lineare (Woodrooff, 1966),
- il criterio d'approssimazione proposto da Deville e Tillè (2005), che permette di trascurare le probabilità di inclusione congiunte.

Su tali basi la varianza della stima regionale può essere espressa come:

$$V_{(tr)}(\hat{y}_{(h)}) = \frac{1}{tr A_{(h)}^2} V_{(tr)}(\hat{Y}_{(h)})$$

<sup>11</sup> % D1=Percentuale Devianza tra scuole su Devianza Totale, % D2= Percentuale Devianza tra classi su Devianza Totale, % D3= Percentuale Devianza dentro le classi su Devianza Totale

$$\cong \frac{1}{tr A^2} \left[ \sum_{i=1}^{tS(h)} \left( \frac{1}{tr \pi_{(h)i}} - 1 \right) tr Z_{(h)i}^2 + \sum_{i=1}^{tS(h)} \frac{tr S_{(h)i} (tr S_{(h)i} - tr S_{(h)i})}{tr \pi_{(h)i} tr S_{(h)i}} tr \sigma_{z(h)i}^2 \right], \quad (7.1)$$

essendo

$$tr Z_{(h)i} = \sum_{j=1}^{tr S_{(h)i}} tr Z_{(h)ij}, \quad tr \sigma_{z(h)i}^2 = \frac{1}{tr S_{(h)i} - 1} \sum_{j=1}^{tr S_{(h)i}} \left( tr Z_{(h)ij} - \frac{tr Z_{(h)i}}{tr S_{(h)i}} \right)^2,$$

$$tr Z_{(h)ij} = \sum_{k=1}^{tr A_{(h)ij}} tr z_{(h)ijk}$$

in cui

$$tr z_{(h)ijk} = tr y_{(h)ijk} - tr \bar{y}_{(h)}$$

rappresenta la variabile linearizzata riferita al singolo alunno.

La variabilità della stima nazionale può essere ottenuta come

$$V(tr \hat{y}) \cong \frac{1}{tr A^2} \sum_{h=1}^{21} V(tr \hat{Y}_{(h)}) \quad (7.2)$$

In appendice A.6 è illustrata una riscrittura della (7.1) utile al metterne in evidenza il termine  $tr S_{(h)}$ , semplificando in tal modo i calcoli necessari per la determinazione della numerosità campionaria.

## 7.2 Stima della varianza campionaria

La stima della varianza campionaria può essere determinata in modo semplice, adottando l'ipotesi che in primo stadio si sia effettuato un campionamento senza reimmissione.

### Regioni NAR

Nel contesto appena introdotto, la stima della varianza campionaria della stima regionale,

$tr \hat{y}_{(h)}$ , è data da:

$$\hat{V}(\text{tr} \hat{Y}(h)) = \frac{1}{\text{tr} A^2(h)} \hat{V}(\text{tr} \hat{Y}(h)), \quad (7.3)$$

essendo

$$\hat{V}(\text{tr} \hat{Y}(h)) = \frac{t^{S(h)}}{t^{S(h)} - 1} \sum_{i=1}^{t^{S(h)}} (\text{tr} \hat{Z}(h)_i - \text{tr} \hat{\bar{Z}}(h))^2, \quad (7.4)$$

in cui

$$\text{tr} \hat{Z}(h)_i = \sum_{j=1}^{t^{S(h)i}} \sum_{k=1}^{\text{tr} A^*(h)ij} \text{tr} z_{(h)ijk} \text{tr} w_{(h)ijk} \quad \text{tr} \hat{\bar{Z}}(h) = \frac{1}{t^{S(h)}} \sum_{i=1}^{t^{S(h)}} \text{tr} \hat{Z}(h)_i. \quad (7.5)$$

Regioni AR

Nelle regioni di tipo Auto Rappresentativo, la formula da applicare é la seguente:

$$\hat{V}(\text{tr} \hat{Y}(h)) = \frac{1}{\text{tr} A^2(h)} \sum_{i=1}^{t^S} \hat{V}(\text{tr} \hat{Y}(hi)), \quad (7.6)$$

essendo

$$\hat{V}(\text{tr} \hat{Y}(hi)) = \frac{t^{S(h)i}}{t^{S(h)i} - 1} \sum_{j=1}^{t^{S(h)i}} (\text{tr} \hat{Z}(h)_{ij} - \text{tr} \hat{\bar{Z}}(h)_i)^2,$$

in cui

$$\text{tr} \hat{Z}(h)_{ij} = \sum_{k=1}^{\text{tr} A^*(h)ij} \text{tr} z_{(h)ijk} \text{tr} w_{(h)ijk}, \quad \text{tr} \hat{\bar{Z}}(h)_i = \frac{1}{t^{S(h)i}} \sum_{i=1}^{t^{S(h)i}} \text{tr} \hat{Z}(h)_{ij}.$$

Stima Nazionale

La variabilità della stima nazionale può essere ottenuta come

$$\hat{V}(\text{tr} \hat{y}) \cong \frac{1}{\text{tr} A^2} \sum_{h=1}^{21} \hat{V}(\text{tr} \hat{Y}(h)). \quad (7.7)$$

La seguente tavola 7.1 riporta gli errori di campionamento stimati, espressi nei termini di coefficiente di variazione percentuale, definiti da:

$$\hat{CV}_{(tr \hat{y}_{(h)})} = 100 \sqrt{\hat{V}_{(tr \hat{y}_{(h)})}} / tr \hat{y}_{(h)}$$

e

$$\hat{CV}_{(tr \hat{y})} = 100 \sqrt{\hat{V}_{(tr \hat{y})}} / tr \hat{y}.$$

La stima dei coefficienti di variazione costituisce la base per la determinazione degli intervalli di confidenza. Con probabilità approssimativamente pari a 0,95 i valori incogniti  $tr \bar{y}$  e  $tr \bar{y}_{(h)}$  dei parametri di interesse sono compresi nei seguenti intervalli:

- $\text{Prob} \left\{ tr \hat{y} - 2_{tr \hat{y}_{(h)}} [\hat{CV}\%_{(tr \hat{y})} / 100] \leq tr \bar{y} \leq tr \hat{y} + 2_{tr \hat{y}_{(h)}} [\hat{CV}\%_{(tr \hat{y})} / 100] \right\} = 0,95 ;$
- $\text{Prob} \left\{ tr \hat{y}_{(h)} - 2_{tr \hat{y}_{(h)}} [\hat{CV}\%_{(tr \hat{y}_{(h)})} / 100] \leq tr \bar{y}_{(h)} \leq tr \hat{y}_{(h)} + 2_{tr \hat{y}_{(h)}} [\hat{CV}\%_{(tr \hat{y}_{(h)})} / 100] \right\} = 0,95.$

**Tavola 7.1 - Errori standard per Ripartizione geografica, Livello scolastico e materia indagata**

Ripartizione geografica	Livello 2		Livello 5		Livello 6		Livello 8		Licei		Tecnici		Professionali	
	Ita	Mat	Ita	Mat										
<b>Nord Ovest</b>	1,2	1,2	1,2	1,4	1,2	1,3	1,6	0,9	1,9	3,1	1,7	2,0	1,7	1,4
<b>Valle d'Aosta</b>	2,8	2,9	2,5	2,6	2,1	2,1	14,7	3,0	6,5	8,9	6,5	7,3	5,7	4,1
<b>Piemonte</b>	1,4	1,5	1,8	2,7	2,0	1,9	2,6	2,0	2,7	5,0	3,1	4,1	4,3	3,0
<b>Liguria</b>	1,9	2,5	2,1	2,4	2,5	2,3	3,7	1,5	4,3	5,5	3,6	3,0	3,4	2,2
<b>Lombardia</b>	1,7	1,7	1,7	1,9	1,7	1,9	2,3	1,0	2,8	4,4	2,3	2,5	1,9	1,6
<b>Nord Est</b>	1,1	1,1	1,2	1,3	1,1	1,0	1,2	1,4	2,1	3,1	1,6	2,0	1,9	1,3
<b>Prov. Aut. Bolzano (l. it.)<sup>12</sup></b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>Prov. Aut. Trento</b>	2,7	2,5	1,8	2,5	2,3	2,3	1,4	1,6	4,9	7,5	3,9	6,5	1,1	9,0
<b>Veneto</b>	1,7	1,6	1,8	1,9	1,7	1,6	2,1	2,0	3,4	4,8	2,2	3,1	2,0	2,0
<b>Friuli-Venezia Giulia</b>	2,5	2,8	2,1	2,4	1,5	1,7	3,5	1,6	4,4	6,3	3,2	6,6	3,2	2,5
<b>Emilia-Romagna</b>	1,8	2,0	2,1	2,3	2,0	1,8	1,4	2,7	3,6	5,6	3,1	3,2	3,0	1,4
<b>Centro</b>	1,0	1,3	1,1	1,2	1,1	1,2	2,4	1,7	2,1	3,3	1,9	2,1	2,3	1,7
<b>Toscana</b>	1,8	1,7	1,9	2,0	2,0	1,8	3,9	2,9	4,6	6,8	3,1	3,5	3,8	2,3
<b>Umbria</b>	1,7	3,0	2,8	3,1	1,8	2,0	2,3	2,8	4,6	7,5	3,8	4,2	4,0	2,7
<b>Marche</b>	2,2	2,0	1,7	1,9	1,6	1,4	2,4	2,5	3,2	5,1	2,0	2,8	3,8	2,6
<b>Lazio</b>	1,6	2,3	1,8	2,0	1,9	2,0	4,1	2,9	2,8	4,8	2,8	3,0	3,9	3,3
<b>Sud</b>	1,4	1,8	1,4	1,6	1,3	1,4	1,6	2,3	1,6	2,3	1,9	1,6	1,8	1,4
<b>Abruzzo</b>	1,6	2,6	1,9	2,7	2,2	2,2	2,9	3,5	4,2	5,6	4,7	3,3	3,8	2,9
<b>Molise</b>	2,6	3,0	2,5	3,0	2,2	2,5	4,1	8,3	5,7	8,1	5,9	1,7	8,5	3,1
<b>Campania</b>	2,3	2,8	2,2	2,5	2,0	2,0	2,2	3,7	2,1	3,0	2,7	2,2	3,0	2,1
<b>Puglia</b>	1,8	2,8	1,8	2,2	1,9	2,1	2,9	3,1	3,3	4,4	2,9	2,4	1,6	2,2
<b>Sud e Isole</b>	1,6	1,9	1,3	1,6	1,4	1,4	3,0	2,5	2,3	2,5	1,5	1,3	2,0	1,6
<b>Basilicata</b>	2,9	2,7	2,6	2,8	2,6	2,6	5,5	5,1	2,9	5,3	2,9	4,6	2,4	2,5
<b>Calabria</b>	3,0	3,6	2,9	3,4	2,1	2,1	4,3	5,0	4,2	5,6	3,3	2,1	4,6	4,2
<b>Sicilia</b>	2,3	2,8	1,8	2,3	2,1	2,2	4,8	3,9	3,6	3,4	2,2	2,0	2,9	2,2
<b>Sardegna</b>	2,0	3,2	1,9	2,9	2,2	2,2	4,2	2,5	3,1	5,3	3,1	3,2	2,3	1,7
<b>Italia</b>	<b>0,6</b>	<b>0,7</b>	<b>0,6</b>	<b>0,7</b>	<b>0,7</b>	<b>0,7</b>	<b>0,9</b>	<b>0,9</b>	<b>1,0</b>	<b>1,5</b>	<b>1,0</b>	<b>1,1</b>	<b>1,0</b>	<b>0,8</b>

Come si può notare dai valori riportati nella tavola 7.1, gli errori standard ottenuti sono piuttosto modesti, permettendo quindi di cogliere in modo adeguato le differenze, laddove esse esistano, tra una regione e un'altra. Ovviamente, come noto in letteratura, l'entità dell'errore standard dipende anche dalla consistenza numerica della popolazione di riferimento. Infatti, si nota che quest'ultimo tende a decrescere, all'aumentare della numerosità della popolazione scolastica dell'area territoriale considerata.

### 7.3 Un modello per la presentazione sintetica degli errori campionari

In questo paragrafo riportiamo alcuni modelli per la presentazione sintetica degli errori campionari. Essi consentono di determinare in modo semplice le varianze campionarie delle

<sup>12</sup> Data la specificità della scuola in lingua italiana della Provincia autonoma di Bolzano e la limitata consistenza numerica della popolazione scolastica in essa presente, si è preferito prendere in considerazione l'intera popolazione scolastica e non un campione da essa estratto.

stime dei punteggi medi che possono essere calcolati per specifici sottoinsiemi di studenti, esempio, le femmine o i maschi e gli alunni delle classi paritarie.

I maschi e le femmine costituiscono un esempio di sottoinsiemi che si distribuiscono con proporzioni approssimativamente uguali nel territorio; mentre gli alunni delle scuole paritarie o pubbliche costituiscono un esempio di sottoinsiemi che si distribuiscono in modo diseguale. I modelli utilizzati per la presentazione sintetica degli errori campionari differiscono tra i due tipi di sottoinsiemi in quanto quelli che si distribuiscono in modo diseguale presentano una componente aggiuntiva di varianza (dovuta appunto alla disomogeneità della distribuzione) di cui bisogna tenere conto nella modellizzazione della varianza.

### 7.3.1 Statistica deff e varianza campionaria

Nel caso si voglia conoscere l'errore campionario per una statistica differente da quelle riportate nelle tabelle precedenti, si può fare riferimento ai modelli per la presentazione sintetica degli errori di campionamento basati sulla statistica *deff* (Kish, 1965) espressa come:

$$tr\ deff_{(h)} = \frac{\hat{V}(tr\ \hat{y}_{(h)})}{\hat{V}_{srs}(tr\ \hat{y}_{(h)})}, \quad (7.8)$$

in cui  $\hat{V}_{srs}(tr\ \hat{y}_{(h)})$  denota la varianza della stima ottenuta sotto il campionamento casuale semplice, espressa come,

$$\hat{V}_{srs}(tr\ \hat{y}_{(h)}) = \frac{tr\ \hat{\sigma}_{(h)}^2}{tr\ a_{(h)}},$$

in cui

$$tr\ \hat{\sigma}_{(h)}^2 = \frac{1}{tr\ a_{(h)}} \sum_{i=1}^{tr\ s_{(h)}} \sum_{j=1}^{tr\ s_{(h)i}} \sum_{k=1}^{tr\ A^*_{(h)ij}} (tr\ y_{(h)ijk} - tr\ \bar{y}_{(h)})^2$$

$$tr\ \bar{y}_{(h)} = \frac{1}{tr\ a_{(h)}} \sum_{i=1}^{tr\ s_{(h)}} \sum_{j=1}^{tr\ s_{(h)i}} \sum_{k=1}^{tr\ A^*_{(h)ij}} tr\ y_{(h)ijk}$$

essendo  ${}_{tr}a_{(h)}$  il numero di alunni campione rispondenti nella regione  $h$  per l'anno  $r$  e la tipologia di scuola  $t$ .

La statistica  $deff$  può anche essere espressa come:

$${}_{tr}deff_{(h)} = 1 + {}_{tr}\rho_{(h)} ({}_{tr}\bar{a}_{(h)} - 1),$$

essendo  ${}_{tr}\rho_{(h)}$  il coefficiente di correlazione intraclasse e

$${}_{tr}\bar{a}_{(h)} = \frac{{}_{tr}a_{(h)}}{{}_{tr}s_{(h)}}$$

il numero medio di alunni campione per Istituto.

Disponendo di una stima del  $deff$  e avendo calcolato  ${}_{tr}a_{(h)}$ , la correlazione intraclasse può essere stimata come:

$${}_{tr}\rho_{(h)} = \frac{{}_{tr}deff_{(h)} - 1}{{}_{tr}\bar{a}_{(h)} - 1}.$$

**Box 7.1 – Formula per le regioni AR (Valle d’Aosta e Bolzano)**

La precedente relazione vale per le regioni NAR, nelle rimanenti regioni AR la formula da applicare é la seguente:

$$tr\,deff_{(h)} = 1 + tr\rho_{(h)} (tr\bar{a}_{(h)[ij]} - 1),$$

essendo  $tr\rho_{(h)}$  il coefficiente di correlazione intraclasse e

$$tr\bar{a}_{(h)[ij]} = \frac{tr\,a_{(h)}}{\sum_{i=1}^{tr\,S_{(h)}} tr\,S_{(h)}}$$

il numero medio di alunni campione per classe.

Disponendo di una stima del  $deff$  e avendo calcolato  $tr\bar{a}_{(h)[ij]}$ , la correlazione intraclasse può essere stimata come:

$$tr\rho_{(h)} = \frac{tr\,deff_{(h)} - 1}{tr\bar{a}_{(h)[ij]} - 1}.$$

La varianza della stima può quindi essere espressa come:

$$\hat{V}(tr\hat{y}_{(h)}) = \frac{1}{tr\,a_{(h)}} \frac{(tr\,A_{(h)} - tr\,a_{(h)})}{tr\,A_{(h)}} tr\hat{\sigma}_{(h)}^2 [1 + tr\rho_{(h)} (tr\bar{a}_{(h)} - 1)]. \tag{7.9}$$

La varianza di una stima nazionale, può quindi essere determinata mediante semplice aggregazione delle stime regionali:

$$\hat{V}(tr\hat{y}) = \sum_{h=1}^{21} \hat{V}(tr\hat{y}_{(h)}). \tag{7.10}$$

Le precedenti espressioni (7.9) e (7.10) costituiscono la base per la determinazione di modelli per la presentazione sintetica degli errori per sottoclassi. La teoria per i sottoinsiemi distribuiti in modo approssimativamente uguale è riportata nel successivo paragrafo 7.3.2; mentre quella relativa a sottoinsiemi distribuiti in modo diseguale è illustrata nel paragrafo 7.3.3. Infine, nel paragrafo 7.3.4 viene descritto come un utente che disponga dei microdati dell’indagine possa calcolare autonomamente le varianze campionarie delle stime.

**7.3.2 Modelli per sottopopolazioni distribuite in modo approssimativamente uguale**

Sia  $b$  una particolare sottoclasse distribuita in modo approssimativamente uguale nel territorio, e sia  ${}_{tr}m_{b(h)}$  il numero di alunni campione rispondenti appartenenti alla sottoclasse nella regione  $h$  per l'anno  $r$  e la tipologia di scuola  $t$ . Sia ancora:

$${}_{tr}f_{b(h)} = \frac{{}_{tr}m_{b(h)}}{{}_{tr}a(h)}$$

la proporzione di alunni nella sottoclasse, con  ${}_{tr}f_{b(h)} < 1$ .

Sotto l'ipotesi che il coefficiente di variazione

$${}_{tr}c_{(h)}^2 = {}_{tr}\hat{\sigma}_{(h)}^2 / {}_{tr}\hat{y}_{(h)}^2, \tag{7.11}$$

nella sottopopolazione sia approssimativamente uguale a quello della popolazione nel suo complesso, la stima della varianza della stima  ${}_{tr}\hat{y}_{b(h)}$  riferita alla sottopopolazione  $b$  può essere ottenuta come:

$$\hat{V}({}_{tr}\hat{y}_{b(h)}) \cong \frac{1}{{}_{tr}f_{b(h)} \cdot {}_{tr}a(h)} \frac{({}_{tr}A(h) - {}_{tr}a(h))}{{}_{tr}A(h)} {}_{tr}c_{(h)}^2 \cdot {}_{tr}\hat{y}_{b(h)}^2 [1 + {}_{tr}\rho_{(h)} ({}_{tr}f_{b(h)} \cdot {}_{tr}\bar{a}_{(h)} - 1)]. \tag{7.12}$$

La varianza di una stima nazionale (o di aggregazione di regioni) viene quindi ottenuta come somma delle varianze regionali determinate secondo l'espressione (7.11).

**Box 7.2 - Formula per le regioni AR (Valle d'Aosta e Bolzano)**

La precedente relazione vale per le regioni NAR, nelle rimanenti regioni AR la formula da applicare é la seguente:

$$\hat{V}({}_{tr}\hat{y}_{b(h)}) \cong \frac{1}{{}_{tr}f_{b(h)} \cdot {}_{tr}a(h)} \frac{({}_{tr}A(h) - {}_{tr}a(h))}{{}_{tr}A(h)} {}_{tr}c_{(h)}^2 \cdot {}_{tr}\hat{y}_{b(h)}^2 [1 + {}_{tr}\rho_{(h)} ({}_{tr}f_{b(h)} \cdot {}_{tr}\bar{a}_{(h)[ij]} - 1)].$$

### 7.3.3 Modelli per sottoclassi distribuite in modo diseguale

Per tali sottoclassi si può partire dalla formula (8.3) di Cicchitelli *et al* (1992).

$$\hat{V}(\text{tr} \hat{y}_{b(h)}) \cong \frac{1}{\text{tr} f_b \text{tr} a(h)} \frac{(\text{tr} A(h) - \text{tr} a(h))}{\text{tr} A(h)} \text{tr} c(h)^2 \text{tr} \hat{y}_b^2 \times \\ \times \{[(1 + \text{tr} \rho(h) (\text{tr} f_{b(h)} \text{tr} \bar{a}(h) - 1))] + [(1 - \text{tr} f_{b(h)}) \text{tr} \hat{y}_{b(h)}^2]\}. \quad (7.13)$$

Per le regioni AR si adotta la formula precedente previa sostituzione di  $\text{tr} \bar{a}(h)$  con  $\text{tr} \bar{a}(h)_{[ij]}$ .

### 7.3.4 Calcolo delle varianze campionarie

Gli utenti e i ricercatori che dispongano dei microdati possono calcolare autonomamente le varianze campionarie delle proprie stime, adottando la procedura articolata nei seguenti passi:

- Calcolo della proporzione della sottoclasse  $\text{tr} f_b$ ;
- Calcolo della stima per la specifica sottoclasse  $\text{tr} \hat{y}_{bh}$ ;
- Calcolo della varianza  $\text{tr} \hat{\sigma}_{(h)}^2$  mediante un software standard statistico e calcolo del coefficiente di variazione mediante semplice rapporto  $\text{tr} c(h)^2 = \text{tr} \hat{\sigma}_{(h)}^2 / \text{tr} \hat{y}_{(h)}^2$ ;
- Calcolo della varianza campionaria mediante la (7.12) per le classi che si distribuiscono in modo approssimativamente uguale o la (7.13) per le altre sottoclassi.

**Tavola 7.2a – Statistiche per il calcolo delle varianze campionarie per stime di punteggi medi riferiti a specifiche sottopopolazioni – Italiano**

Ripartizione geografica	Livello 2			Livello 5			Livello 6			Livello 8		
	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)
Valle d'Aosta	1.169	542	0,3	1.121	450	0,2	1.241	594	0,2	1.228	274	1,6
Piemonte	38.169	1.367	0	37.593	1.313	0,1	40.028	1.555	0,2	39.164	1.560	0,5
Liguria	12.366	884	0,1	12.375	865	0,1	13.252	992	0,2	12.980	1.222	0,5
Lombardia	93.298	1.556	0,1	90.294	1.524	0,1	95.307	2.374	0,2	92.853	2.405	0,4
Prov. Aut. Bolzano (I. it.) <sup>13</sup>	1.241	1.067	0,5	1.282	1.056	0,7	1.313	1.060	0,8	1.315	1.124	3,8
Prov. Aut. Trento	5.376	886	0,2	5.442	888	0,1	5.612	1.025	0,2	5.608	777	0,1
Veneto	46.776	1.212	0,1	45.502	1.119	0,1	47.569	1.409	0,1	47.706	1.585	0,3
Friuli-Venezia Giulia	10.233	988	0,2	10.146	984	0,1	10.826	1.113	0,1	10.418	1.115	0,6
Emilia-Romagna	39.688	1.261	0,1	38.509	1.167	0,1	39.619	1.530	0,1	38.267	1.578	0,1
Toscana	32.131	1.259	0,1	31.510	1.179	0,1	32.860	1.340	0,1	31.670	1.470	0,6
Umbria	7.848	926	0,1	7.613	857	0,2	7.944	1.113	0,1	7.876	974	0,3
Marche	13.670	1.058	0,2	13.624	1.078	0,1	14.056	1.216	0,1	14.220	1.235	0,4
Lazio	52.849	1.598	0,1	51.741	1.537	0,1	54.424	1.573	0,2	53.255	1.856	0,9
Abruzzo	11.445	1.007	0,1	11.212	1.033	0,1	12.091	1.123	0,2	11.999	1.120	0,4
Molise	2.527	754	0,3	2.666	869	0,2	2.763	973	0,2	2.973	776	0,9
Campania	62.140	2.221	0,4	64.830	2.269	0,4	68.692	2.063	0,3	69.660	2.160	0,5
Puglia	39.049	1.982	0,2	41.237	1.939	0,2	44.029	1.884	0,2	44.066	1.975	0,5
Basilicata	4.797	909	0,3	5.374	890	0,2	5.692	1.059	0,3	5.759	883	0,9
Calabria	18.442	1.037	0,3	18.822	1.161	0,3	20.183	1.161	0,1	20.211	1.362	0,6
Sicilia	50.892	1.754	0,3	51.556	1.802	0,2	56.125	1.642	0,2	56.066	1.768	0,8
Sardegna	13.324	799	0,1	13.444	862	0,1	14.870	735	0,1	14.790	946	0,8

**Tavola 7.2b – Statistiche per il calcolo delle varianze campionarie per stime di punteggi medi riferiti a specifiche sottopopolazioni – Italiano**

Ripartizione geografica	Licei			Tecnici			Professionali		
	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)
Valle d'Aosta	535	158	0,7	322	149	0,8	228	114	0,6
Piemonte	16.988	965	0,3	10.969	638	0,4	6.439	502	0,4
Liguria	6.512	585	0,5	3.193	481	0,4	2.519	345	0,2
Lombardia	37.470	1.531	0,5	26.107	1.161	0,3	12.843	792	0,2
Prov. Aut. Bolzano (I. it.) <sup>13</sup>	559	498	1,9	320	250	0,3	229	155	-0,1
Prov. Aut. Trento	2.441	495	0,8	1.773	451	0,6	126	49	-0,1
Veneto	18.273	1.005	0,5	15.001	960	0,2	8.005	724	0,1
Friuli-Venezia Giulia	4.459	691	0,7	3.379	549	0,3	1.576	384	0,3
Emilia-Romagna	15.725	947	0,5	12.294	742	0,4	8.271	498	0,3
Toscana	15.846	772	0,6	9.504	519	0,3	6.296	502	0,5
Umbria	4.017	637	0,6	2.248	502	0,5	1.490	343	0,5
Marche	6.519	867	0,4	4.865	633	0,2	3.102	408	0,3
Lazio	30.084	686	0,2	12.818	714	0,3	7.835	372	0,4
Abruzzo	6.641	706	0,5	3.826	541	0,8	1.805	398	0,5
Molise	1.656	590	1,2	1.108	382	1	492	157	0,9
Campania	33.072	1.370	0,3	19.512	1.060	0,5	13.759	920	0,6
Puglia	20.540	1.091	0,5	13.355	929	0,4	9.594	704	0,1
Basilicata	2.907	679	0,3	1.900	539	0,4	1.444	444	0,3
Calabria	9.663	766	0,6	6.782	707	0,5	3.920	429	0,7
Sicilia	25.795	963	0,5	14.749	721	0,2	10.840	600	0,4
Sardegna	7.487	338	0,1	4.728	222	0,1	3.021	172	0

**Tavola 7.3a – Statistiche per il calcolo delle varianze campionarie per stime di punteggi medi riferiti a specifiche sottopopolazioni -Matematica**

Ripartizione geografica	Livello 2			Livello 5			Livello 6			Livello 8		
	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)
Valle d'Aosta	1169	539	0,4	1121	452	0,2	1241	595	0,2	1228	274	0,1
Piemonte	38169	1357	0,1	37593	1316	0,3	40028	1557	0,1	39164	1560	0,3
Liguria	12366	892	0,2	12375	873	0,1	13252	1029	0,2	12980	1222	0,1
Lombardia	93298	1544	0,1	90294	1499	0,1	95307	2373	0,2	92853	2405	0,1
Prov. Aut. Bolzano (l. it.) <sup>13</sup>	1241	1050	1,1	1282	1035	1,1	1313	1063	0,9	1315	1124	1,5
Prov. Aut. Trento	5376	884	0,2	5442	883	0,2	5612	1025	0,2	5608	777	0,1
Veneto	46776	1216	0,1	45502	1150	0,1	47569	1410	0,1	47706	1585	0,2
Friuli-Venezia Giulia	10233	1019	0,3	10146	1003	0,2	10826	1112	0,1	10418	1115	0,1
Emilia-Romagna	39688	1259	0,1	38509	1191	0,2	39619	1528	0,1	38267	1578	0,3
Toscana	32131	1272	0,1	31510	1208	0,1	32860	1337	0,1	31670	1470	0,4
Umbria	7848	927	0,3	7613	852	0,3	7944	1110	0,1	7876	974	0,3
Marche	13670	1046	0,1	13624	1066	0,1	14056	1216	0	14220	1235	0,3
Lazio	52849	1613	0,3	51741	1528	0,2	54424	1570	0,2	53255	1856	0,6
Abruzzo	11445	1006	0,2	11212	1019	0,2	12091	1122	0,2	11999	1120	0,5
Molise	2527	745	0,3	2666	845	0,4	2763	971	0,3	2973	776	1,8
Campania	62140	2184	0,6	64830	2225	0,6	68692	2060	0,3	69660	2160	0,8
Puglia	39049	1961	0,5	41237	1935	0,3	44029	1868	0,2	44066	1975	0,6
Basilicata	4797	896	0,3	5374	887	0,3	5692	1058	0,3	5759	883	0,8
Calabria	18442	1024	0,4	18822	1116	0,6	20183	1161	0,2	20211	1362	0,7
Sicilia	50892	1753	0,5	51556	1762	0,4	56125	1640	0,3	56066	1768	0,8
Sardegna	13324	883	0,3	13444	928	0,3	14870	699	0,1	14790	946	0,3

**Tavola 7.3b – Statistiche per il calcolo delle varianze campionarie per stime di punteggi medi riferiti a specifiche sottopopolazioni -Matematica**

Ripartizione geografica	Licei			Tecnici			Professionali		
	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)	tr A(h)	tr a(h)	tr ρ(h)
Valle d'Aosta	535	158	0,8	322	148	0,8	228	113	0,4
Piemonte	16988	963	0,6	10969	613	0,6	6439	501	0,3
Liguria	6512	587	0,6	3193	497	0,3	2519	359	0,1
Lombardia	37470	1507	0,8	26107	1164	0,4	12843	788	0,2
Prov. Aut. Bolzano (l. it.) <sup>13</sup>	559	498	4,2	320	248	1,8	229	154	-0,1
Prov. Aut. Trento	2441	495	1,3	1773	451	1,3	126	48	0,5
Veneto	18273	1022	0,7	15001	956	0,4	8005	721	0,2
Friuli-Venezia Giulia	4459	681	0,9	3379	547	1,1	1576	384	0,2
Emilia-Romagna	15725	945	0,9	12294	741	0,3	8271	497	0,1
Toscana	15846	769	1	9504	539	0,4	6296	497	0,2
Umbria	4017	636	1,2	2248	501	0,7	1490	342	0,4
Marche	6519	868	0,7	4865	633	0,3	3102	414	0,3
Lazio	30084	749	0,5	12818	714	0,5	7835	367	0,4
Abruzzo	6641	706	0,9	3826	536	0,5	1805	394	0,6
Molise	1656	590	2,1	1108	384	0,1	492	157	0,2
Campania	33072	1376	0,5	19512	1061	0,5	13759	925	0,4
Puglia	20540	1088	0,6	13355	928	0,3	9594	703	0,3
Basilicata	2907	678	0,9	1900	539	1,3	1444	443	0,4
Calabria	9663	779	0,9	6782	707	0,2	3920	428	1
Sicilia	25795	963	0,5	14749	721	0,2	10840	600	0,4
Sardegna	7487	350	0,5	4728	220	0,2	3021	182	0

<sup>13</sup> Data la specificità della scuola in lingua italiana della Provincia autonoma di Bolzano e la limitata consistenza numerica della popolazione scolastica in essa presente, si è preferito prendere in considerazione l'intera popolazione scolastica e non un campione da essa estratto.

## 8. IL CHEATING E L'AFFIDABILITA' DEI DATI

L'importanza crescente che a livello nazionale e internazionale rivestono le valutazioni standardizzate delle competenze degli studenti si accompagna a una crescente esigenza di disporre di dati qualitativamente affidabili e attendibili. In particolar modo, in Italia è molto forte l'attenzione dell'opinione pubblica verso le rilevazioni effettuate dall'INVALSI nell'ambito del Servizio Nazionale di Valutazione. Di conseguenza, la qualità dei dati raccolti è diventata un requisito indispensabile per renderli pienamente fruibili e "apprezzabili" da tutti gli *stakeholder* del sistema educativo: *policy maker*, insegnanti, ricercatori, famiglie e studenti. In questa prospettiva risulta necessario disporre di strumenti metodologici tecnicamente validi e condivisi in grado di massimizzare la qualità dei dati e di limitare l'effetto di componenti "distorsive", come quelle legate al fenomeno del *cheating*.

Il *cheating* in ambito educativo indica le pratiche messe in atto dagli studenti (*student cheating*) o degli insegnanti (*teacher cheating*) per alterare, sovrastimandoli, i risultati delle prove standardizzate di valutazione. La presenza del *cheating* tende a invalidare i risultati delle rilevazioni compromettendone fortemente l'accuratezza e di conseguenza l'attendibilità.

Il *cheating* è un argomento molto complesso e, dal punto di vista metodologico, difficile da identificare e correggere in quanto può essere attuato da molteplici attori, in diversi *step* del processo di rilevazione dei dati e con differenti modalità.

### 8.1 L'individuazione e correzione del *cheating* mediante l'approccio *fuzzy clustering*

L'INVALSI, sin dalle prime rilevazioni condotte nell'a.s. 2004/05 nell'ambito del SNV, ha messo in luce la presenza del fenomeno e ha cercato di correggerne e limitarne l'impatto adottando una procedura statistica di correzione costruita *ad hoc* per i dati del SNV (Quintano *et. al.*, 2009).

I dati rilevati con l'Indagine Campionaria di Controllo (ICC) risultano statisticamente accurati in virtù del ricorso a osservatori esterni in grado di garantire un corretto svolgimento dell'intero processo di somministrazione. La presenza degli osservatori si traduce nel produrre una base dati di elevata affidabilità caratterizzata dalla pressoché assenza di anomalie in termini di *cheating*.

La metodologia adottata, quindi, si applica principalmente sui dati provenienti dall'Indagine Standard, utilizzando tecniche statistiche di riduzione dimensionale e di raggruppamento (fattorizzazione e *fuzzy clustering*) e permette di stimare la probabilità di *cheating* di ogni classe di studenti mediante l'analisi di un insieme di indicatori: punteggio medio, variabilità interna, livello di eterogeneità nelle risposte e tasso di presenza di dati mancanti.

Sulla base di questa probabilità si costruisce un fattore di correzione che "penalizza" i punteggi medi di classe al crescere del grado di sospetto di *cheating*.

Si parte dal presupposto che se in una classe vengono messi in atto dei comportamenti opportunistici, determinando *pattern* anomali di risposta, da parte degli insegnanti oppure degli studenti, i risultati tenderanno a perdere la variabilità che naturalmente dovrebbe contraddistinguerli e, di conseguenza, si osserverà un elevato punteggio medio della classe (che tenderà verso il valore massimo), una variabilità interna minima o nulla e, inoltre, valori minimi dei tassi di mancata risposta, in quanto tutti gli studenti forniranno risposte al test che tenderanno ad essere uguali per tutta la classe (eterogeneità minima o assente nelle risposte).

## 8.2 Come viene trattato nel campione: il livello 8

La rilevazione nazionale degli apprendimenti nella classe III della scuola secondaria di I grado si svolge durante l'esame di Stato, i risultati derivanti da tale rilevazione, rispetto agli altri livelli scolari indagati, risultano essere affetti dal *cheating* in maniera più sostanziale. Le motivazioni possono essere diverse, trattandosi di un esame la propensione a "copiare" è maggiore anche negli studenti, inoltre nelle classi campione non è presente l'osservatore esterno ma viene chiesto al Presidente di Commissione di vigilare.

Al fine di fornire dei risultati comunque robusti, in questo caso il campione viene "corretto" dal *cheating* utilizzando un fattore di correzione calcolato a livello di classe. Tale fattore di correzione altro non è che il complemento a uno dell'Indice di propensione al *cheating* calcolato secondo la procedura evidenziata al paragrafo precedente. In pratica il punteggio di ciascuno studente viene moltiplicato per questo fattore di correzione, uguale per tutti gli studenti della classe, considerando la classe come l'unità minima di aggregazione.

---

## RIFERIMENTI BIBLIOGRAFICI

Cicchitelli G., Herzl A., Montanari G.E. (1992), *Il Campionamento Statistico*, Il Mulino.

Kish, L. (1965), *Survey Sampling*, Wiley, New Yoork.

Särndal C.-E., and Lundström S., (2005), *Estimation in Surveys with Nonresponse*.

Quintano C., Castellano R. e Longobardi S. (2009), *A fuzzy clustering approach to improve the accuracy of Italian student data. An experimental procedure to correct the impact of outliers on assessment test scores*, *Statistica&Applicazioni*, Vol. VII, n.2, Vita e Pensiero, ISSN 1824-6672.

## APPENDICE - Approfondimenti metodologici

### A.1 Riformulazione dei punteggi medi

I valori medi regionali possono essere riformulati nel modo seguente

$$\begin{aligned}
 tr \bar{y}(h) &= \frac{1}{tr A(h)} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} tr Y(h)_{ij} \\
 &= \frac{1}{tr A(h)} \sum_{i=1}^{tr S(h)} tr Y(h)_i \\
 &= \frac{tr Y(h)}{tr A(h)}.
 \end{aligned}$$

Una formulazione alternativa è data da

$$\begin{aligned}
 tr \bar{y}(h) &= \frac{1}{tr A(h)} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} tr A(h)_{ij} tr \bar{y}(h)_{ij} \\
 &= \frac{1}{tr A(h)} \sum_{i=1}^{tr S(h)} tr A(h)_i tr \bar{y}(h)_i.
 \end{aligned}$$

Analogamente, il valore medio nazionale è dato da

$$\begin{aligned}
 tr \bar{y} &= \frac{1}{tr A} \sum_{h=1}^{21} tr Y(h) \\
 &= \frac{tr Y}{tr A}, \\
 &= \frac{1}{tr A} \sum_{h=1}^{21} tr A(h) tr \bar{y}(h)
 \end{aligned}$$

essendo  $tr Y = \sum_{h=1}^{21} tr Y(h)$ .

## A.2 Differenti componenti della devianza

Essendo i doppi prodotti nulli:

$$tr D = \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A(h)ij} (tr y_{(h)ijk} - tr \bar{y} + tr \bar{y}_{(h)i} - tr \bar{y}_{(h)i} + tr \bar{y}_{(h)ij} - tr \bar{y}_{(h)ij})^2 =$$

$$tr D = \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A(h)ij} (tr \bar{y}_{(h)i} - tr \bar{y})^2 +$$

$$+ \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A(h)ij} (tr \bar{y}_{(h)ij} - tr \bar{y}_{(h)i})^2 +$$

$$+ \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A(h)ij} (tr y_{(h)ijk} - tr \bar{y}_{(h)ij})^2.$$

Si ottiene il risultato, riportato nel paragrafo 4.3, considerando che

- $\sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A(h)ij} (tr \bar{y}_{(h)i} - tr \bar{y})^2 = \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} tr A(h)i (tr \bar{y}_{(h)i} - tr \bar{y})^2$ ,
  - $\sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} \sum_{k=1}^{tr A(h)ij} (tr \bar{y}_{(h)ij} - tr \bar{y}_{(h)i})^2$
- $$= \sum_{h=1}^{21} \sum_{i=1}^{tr S(h)} \sum_{j=1}^{tr S(h)i} tr A(h)ij (tr \bar{y}_{(h)ij} - tr \bar{y}_{(h)i})^2.$$

### A.3 Condizione di autoponderazione

Nell'ipotesi che la dimensione delle classi in termini di alunni sia pressappoco costante nello strato e che ogni classe sia approssimativamente costituita da  ${}_t\bar{A}_{(h)}$  alunni, valgono le seguenti condizioni

$$\begin{aligned} \text{a)} \quad & {}_{tr}A_{(h)ij} \cong {}_t\bar{A}_{(h)} \\ \text{b)} \quad & {}_{tr}A_{(h)i} \cong {}_tS_{(h)i} {}_t\bar{A}_{(h)}. \end{aligned} \tag{a}$$

Valendo la (a), si ha

$$\begin{aligned} \bullet \quad & {}_t\ddot{A}_{(h)i} \cong {}_t\bar{A}_{(h)} {}_tS_{(h)i} \frac{2}{{}_tS_{(h)i}} \\ \bullet \quad & \sum_{i=1}^{{}_tS_{(h)}} {}_t\ddot{A}_{(h)i} \cong {}_t\bar{A}_{(h)} {}_t\bar{S}_{(h)} \end{aligned} \tag{b}$$

Sostituendo le precedenti espressioni (a) e (b) nella (5.4) si ha:

$${}_{tr}\pi_{(h)ijk} = {}_tS_{(h)} \frac{{}_t\ddot{A}_{(h)i}}{\sum_{i=1}^{{}_tS_{(h)}} {}_t\ddot{A}_{(h)i}} \frac{{}_tS_{(h)i}}{{}_tS_{(h)i}} \cong {}_tS_{(h)} \frac{{}_t\bar{A}_{(h)} {}_tS_{(h)i}}{{}_t\bar{A}_{(h)} {}_t\bar{S}_{(h)}} \frac{2}{{}_tS_{(h)i}} \frac{{}_tS_{(h)i}}{{}_tS_{(h)i}} = \frac{2 {}_tS_{(h)}}{{}_t\bar{S}_{(h)}} \tag{c}$$

## A.4 Algoritmo per il peso calibrato

### 1. Prima Iterazione

$${}^1_{tr}W_{(h)i} = \frac{1}{{}_t\pi_{(h)i}} \frac{{}_tS_{(h)}}{\sum_{i=1} {}_tS_{(h)} \frac{1}{{}_t\pi_{(h)i}}}$$

### 2. Iterazione sul numero di alunni

$${}^2_{tr}W_{(h)i} = \frac{{}_1_{tr}W_{(h)i}}{\sum_{i=1} {}_tS_{(h)} \frac{{}_1_{tr}W_{(h)i}}{{}_t\pi_{(h)i}}} \frac{{}_tA_{(h)}}{\sum_{i=1} {}_tS_{(h)} \frac{{}_1_{tr}W_{(h)i}}{{}_t\pi_{(h)i}} \frac{{}_tA_{(h)}}{{}_t\pi_{(h)i}}}$$

### 3. Iterazione sul numero di scuole

$${}^3_{tr}W_{(h)i} = \frac{{}_2_{tr}W_{(h)i}}{\sum_{i=1} {}_tS_{(h)} \frac{{}_2_{tr}W_{(h)i}}{{}_t\pi_{(h)i}}} \frac{{}_tS_{(h)}}{\sum_{i=1} {}_tS_{(h)} \frac{{}_2_{tr}W_{(h)i}}{{}_t\pi_{(h)i}}}$$

### 4. Generica iterazione (v) sul numero di alunni

$${}^v_{tr}W_{(h)i} = \frac{{}^{v-1}_{tr}W_{(h)i}}{\sum_{i=1} {}_tS_{(h)} \frac{{}^{v-1}_{tr}W_{(h)i}}{{}_t\pi_{(h)i}}} \frac{{}_tA_{(h)}}{\sum_{i=1} {}_tS_{(h)} \frac{{}^{v-1}_{tr}W_{(h)i}}{{}_t\pi_{(h)i}} \frac{{}_tA_{(h)}}{{}_t\pi_{(h)i}}}$$

### 5. Generica iterazione sul numero di scuole

$${}^v_{tr}W_{(h)i} = \frac{{}^{v-1}_{tr}W_{(h)i}}{\sum_{i=1} {}_tS_{(h)} \frac{{}^{v-1}_{tr}W_{(h)i}}{{}_t\pi_{(h)i}}} \frac{{}_tS_{(h)}}{\sum_{i=1} {}_tS_{(h)} \frac{{}^{v-1}_{tr}W_{(h)i}}{{}_t\pi_{(h)i}}}$$

### 6. Interruzione della procedura

La procedura si interrompe dopo 4/5 iterazioni allo stabilizzarsi, definendo il peso di primo stadio  ${}_{tr}W_{(h)i}$ .

## A.5 Stima delle componenti della devianza

### Stima della devianza totale D

Per la dimostrazione della (6.5), si considerino i seguenti passaggi in cui, per non appesantire eccessivamente la dimostrazione si sono omessi gli indici  $t, r$  e  $(h)$ :

$$\begin{aligned} \hat{D} &= \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} (y_{ijk} - \hat{y})^2 = \\ &= \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} y_{ijk}^2 + \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} \hat{y}^2 + \\ &\quad - 2 \hat{y} \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} y_{ijk} \\ &= \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} y_{ijk}^2 + \hat{A} \hat{y}^2 - 2 \hat{y} \hat{A} \hat{y} \\ &= \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} y_{ijk}^2 - \hat{A} \hat{y}^2. \end{aligned}$$

Si considerino i seguenti risultati

$$E_p \left( \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} y_{ijk}^2 \right) = \sum_{i=1}^S \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{i,j,k}^2$$

$$E_p(\hat{A} \hat{y}^2) \cong E_p(\hat{A}) E_p(\hat{y}^2) = A(V(\hat{y}) + \bar{y}^2).$$

Di conseguenza, una stima approssimativamente corretta di  $D$  è data da:

$$\hat{D}_{cor} = \sum_{i=1}^S w_i \sum_{j=1}^{S_i} \sum_{k=1}^{A_{ij}} w_{j|i} y_{ijk}^2 - \hat{A}(\hat{y}^2 - \hat{V}(\hat{y})).$$

### Stima della componente D1

Per la dimostrazione della (6.6), si considerino i seguenti passaggi in cui, per non appesantire eccessivamente la dimostrazione si sono omessi gli indici  $t, r$  e  $(h)$

$$\hat{D}1 = \sum_{i=1}^S w_i \hat{A}_i (\hat{y}_i - \hat{y})^2 =$$

$$\begin{aligned} \hat{D}1 &= \sum_{i=1}^S w_i \hat{A}_i \hat{y}_i^2 + \hat{y}^2 \sum_{i=1}^S w_i \hat{A}_i - 2\hat{y} \sum_{i=1}^S w_i \hat{A}_i \hat{y}_i \\ &= \sum_{i=1}^S w_i \hat{A}_i \hat{y}_i^2 + \hat{A} \hat{y}^2 - 2\hat{y} \hat{A} \hat{y} \\ &= \sum_{i=1}^S w_i \hat{A}_i \hat{y}_i^2 - \hat{A} \hat{y}^2 . \end{aligned}$$

In base ai risultati precedenti, una stima approssimativamente corretta di  $D1$  è data da:

$$\hat{D}1_{cor} = [\sum_{i=1}^S w_i \hat{A}_i (\hat{y}_i^2 - \ddot{V}(\hat{y}_i))] - [A(\hat{y}^2 - \hat{V}(\hat{y}))]$$

$$\hat{D}1_{cor} = \hat{D}1 - \sum_{i=1}^S \ddot{V}(\hat{y}_i) + A\hat{V}(\hat{y})$$

essendo

$$\ddot{V}(\hat{y}_i) = \begin{cases} \hat{V}(\hat{y}_i) & \text{per } i \in ( >1 )_t \bar{S}(h) \\ \tilde{V}(\hat{y}_i) & \text{per } i \in ( =1 )_t \bar{S}(h) \end{cases}$$

Negli istituti con almeno 2 sezioni campione tale variabilità viene stimata come:

$$\hat{V}(tr \hat{y}_i) = \sum_{j=1}^{t^{S(h)i}} \frac{t^{S(h)i}}{t^{S(h)i} - 1} \left( tr \hat{R}_{ij} - tr \hat{R}(h)_i \right)^2 ,$$

essendo,

$$tr \hat{R}(h)_{ij} = \frac{1}{t \pi(h)_{j|i}} \frac{1}{tr \hat{A}(h)_i} (tr Y(h)_{ij} - \frac{tr \hat{Y}(h)_i}{tr \hat{A}(h)_i} tr A^*(h)_{ij}) \quad , \quad tr \hat{R}(h)_i = \sum_{j=1}^{t^{S(h)i}} \frac{1}{t^{S(h)i}} tr \hat{R}(h)_{ij} ,$$

$$tr \hat{Y}(h)_i = \sum_{j=1}^{t^{S(h)i}} \frac{1}{t \pi(h)_{j|i}} tr A^*(h)_{ij} \quad , \quad tr \hat{A}(h)_i = \sum_{j=1}^{t^{S(h)i}} \frac{1}{t \pi(h)_{j|i}} tr A^*(h)_{ij} .$$

Negli istituti con un'unica sezione campione tale variabilità viene stimata come

$$\hat{V}_{(tr \hat{y}_i)} = \sum_{k=1}^{tr A_{hij}} \frac{1}{tr A_{hij} - 1} \left( tr z_{(h)ijk} - tr \bar{z}_{(h)ij} \right)^2.$$

A.6 Espressione alternativa della stima della varianza regionale

Riformulando, la (7.1) al fine di mettere in evidenza il termine  ${}_{t}S_{(h)}$  si ottiene:

$$V({}_{tr}\hat{y}_{(h)}) = \frac{1}{{}_{tr}A_{(h)}^2} \left[ \frac{{}_{tr}K_{(h)}}{{}_{tr}S_{(h)}} - {}_{tr}G_{(h)} \right], \quad (6.3)$$

in cui

$${}_{tr}K_{(h)} = \sum_{i=1}^{{}_{tr}S_{(h)}} \frac{1}{{}_{tr}\alpha_{(h)i}} \left[ {}_{tr}Z_{(h)i}^2 + \frac{{}_{tr}S_{(h)i}({}_{tr}S_{(h)i} - {}_{tr}S_{(h)})}{{}_{tr}S_{(h)i}} {}_{tr}\sigma_{z(h)i}^2 \right], \quad (6.4)$$

$${}_{tr}G_{(h)} = \sum_{i=1}^{{}_{tr}S_{(h)}} {}_{tr}Z_{(h)i}^2. \quad (6.5)$$

## A.7 Un approccio *multistep* per l'individuazione e la correzione del *cheating* nei risultati del SNV

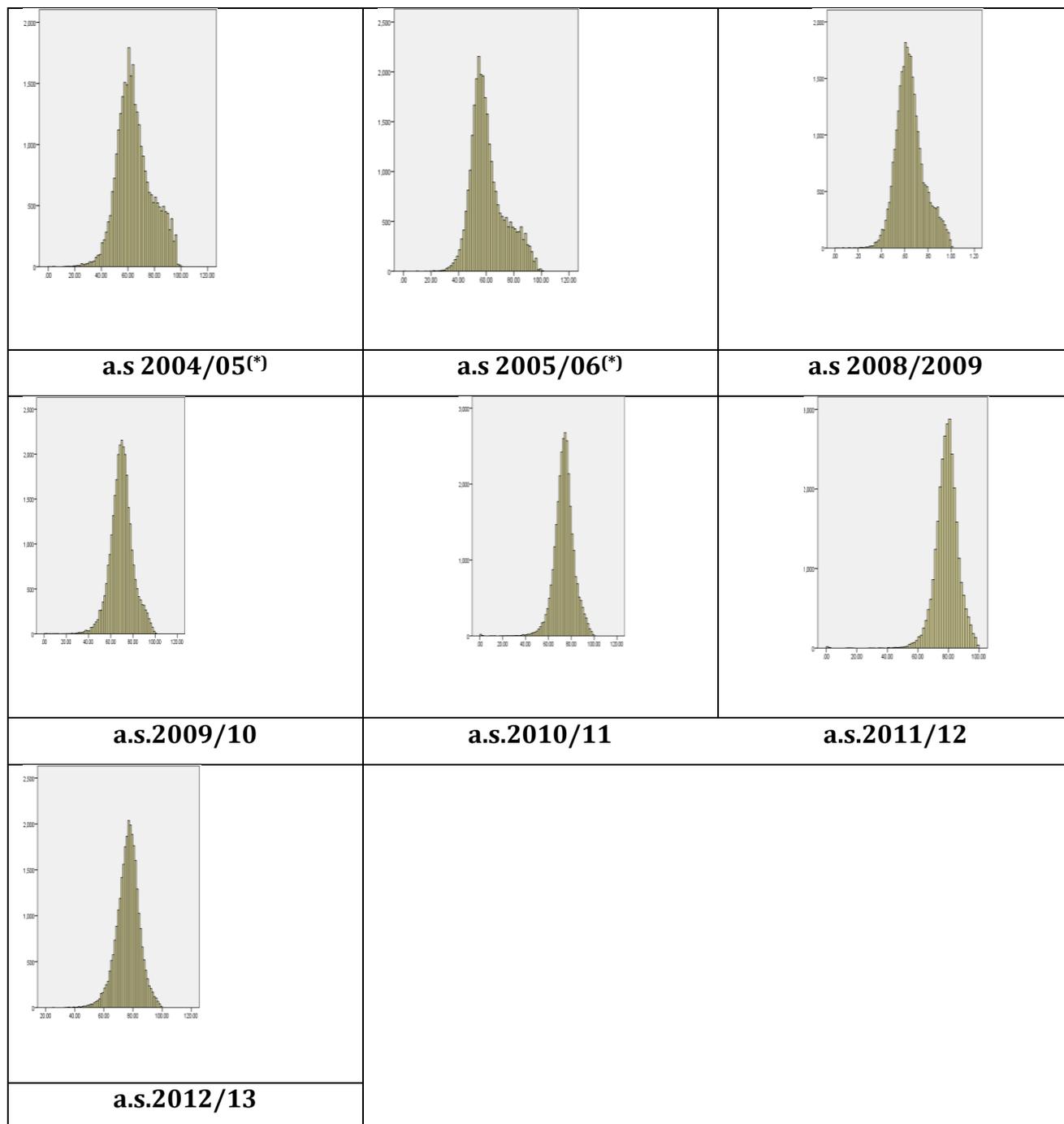
(in collaborazione con Sergio Longobardi)

Questo approccio, utilizzato fino all'a.s. 2011/12, è stato progettato per correggere il *cheating* in popolazioni caratterizzate da un elevato grado di anomalia, mentre si è rivelato poco idoneo a catturare il fenomeno quando la sua intensità è meno marcata. In particolare, si è osservato che la metodologia tende a sovrastimare il fenomeno in situazioni con “basso *cheating*”, attribuendo probabilità elevate di *cheating* anche a classi che “genuinamente” raggiungono punteggi alti, generando delle stime erronee del *cheating*, i cosiddetti “falsi positivi”. La presenza di falsi positivi è cresciuta negli ultimi anni in corrispondenza del minore ricorso al *cheating* da parte dei docenti e/o degli studenti. Questa evoluzione del fenomeno ha reso necessaria una revisione della procedura al fine di migliorarne la capacità identificativa e correttiva.

### A.7.1 L'evoluzione del *cheating*

Nel corso degli anni si è assistito a un ricorso decrescente al *cheating* che da “problema” con elevata diffusione e forte connotazione territoriale e di livello scolastico, interessando in modo specifico le classi della scuola primaria del Centro-Sud, è diventato un fenomeno con minore intensità, caratterizzato da una concentrazione territoriale meno spiccata e interessando in misura maggiore le classi III della scuola secondaria di I grado, dove i risultati della prova contribuiscono a definire il voto finale dell'esame di Stato. Questa “involuzione” del fenomeno è evidente osservando le distribuzioni dei punteggi medi di classe nel corso delle rilevazioni che si sono succedute negli anni (Figura A.7.1).

**Figura A.7.1 -Distribuzione dei punteggi medi di classe in Italiano delle classi V della scuola primaria nelle rilevazioni 2004/05, 2005/06, 2008/09, 2009/10, 2010/11, 2011/12, 2012/13(\*)**



(\*) Negli a.s. 2004/05 e 2005/06 la rilevazione ha interessato le classi IV della scuola primaria. Nell'a.s. 2006/07 non sono state effettuate le rilevazioni mentre nell'a.s. 2007/08 è stata effettuata soltanto la Prova Nazionale per gli studenti delle classi III della scuola secondaria di I grado.

Da un semplice confronto grafico si nota che la distribuzione dei punteggi medi di classe nei primi anni delle rilevazioni del SNV risultava caratterizzata da valori modalì in corrispondenza sia del valore centrale che del valore massimo della distribuzione (bimodalità); nelle rilevazioni condotte negli ultimi anni scolastici queste anomalie si sono notevolmente attenuate in virtù di

una significativa riduzione del fenomeno del *cheating*. Il minore ricorso al *cheating*, probabilmente dovuto anche a una maggiore sensibilizzazione delle scuole da parte dell'INVALSI e da alcuni strumenti correttivi adoperati in fase di somministrazione della prova, ha reso più complicata la sua individuazione rendendo necessaria una revisione della procedura adottata fino all'a.s. 2011/12, in modo da adeguarla alla nuova connotazione del fenomeno.

#### *A.7.2 La revisione della procedura di cheating*

Nell'ottica di una revisione della procedura di controllo e correzione del *cheating*, al fine di affinare e rendere più preciso questo strumento, in modo da poter individuare correttamente il fenomeno anche laddove non sia fortemente connotato, è stata sviluppata una nuova metodologia statistica di correzione. Questa metodologia (procedura 2013) che si colloca a valle della procedura già utilizzata va a costituire un procedimento *multistep* che affianca a una logica "meccanicistica" (alto punteggio e bassa variabilità) un approccio "da modello", che permette di rimodulare i valori della probabilità di *cheating*.

La logica è quella di "premiare" (diminuire il valore dell'indice di propensione al *cheating*) le classi che presentano media e deviazione standard del punteggio (all'interno della classe) coerenti con quelle previste da un modello di regressione stimato sui dati del campione e, inoltre, una relazione concordante tra voti al I quadrimestre e punteggi INVALSI (indice di correlazione di Bravais-Pearson<sup>14</sup>).

Nel dettaglio, i dati grezzi di ciascuna classe vengono analizzati utilizzando la procedura precedente in modo da ottenere una prima misura della presenza di anomalie. Successivamente, restringendo il campo di analisi ai dati delle sole classi campione<sup>15</sup>, si stimano dei modelli di regressione lineare esplicativi della media e della variabilità interna (deviazione standard) dei risultati di classe. A tale scopo, mediante approccio *forward stepwise*, vengono impiegate diverse covariate relative a fattori, prevalentemente di contesto socio-economico, che influenzano le *performance* delle classi.

<sup>14</sup> Il coefficiente di correlazione di Bravais-Pearson è un indice che permette di misurare il grado di legame, lineare, tra due variabili statistiche.

<sup>15</sup> I dati rilevati sul campione hanno una minore incidenza di dati mancanti (*missing data*) per quanto riguarda le variabili di contesto. Di conseguenza i dati campionari costituiscono un valido campo di prova per la stima dei modelli lineari del punteggio medio e della variabilità.

I modelli, calcolati sulla base dei dati campionari vengono successivamente “estesi” all’intera popolazione (*fitting over sample*) in modo da stimare, per ogni classe, il punteggio medio e la sua deviazione standard. I valori stimati vengono utilizzati, insieme all’indice di correlazione lineare che esprime per ogni classe il grado di relazione tra punteggi INVALSI e voti al I quadrimestre, per definire un insieme di tre indicatori che modificano l’indice di *cheating* calcolato mediante l’approccio precedente.

L’insieme di indicatori è costituito da:

- a) indicatore di plausibilità della deviazione standard ( $IP^\sigma$ );
- b) indicatore di plausibilità della relazione tra punteggi SNV e votazione I quadrimestre ( $IP^{corr}$ );
- c) fattore correttivo ( $C$ ) costruito sulla base del confronto tra punteggi osservati e stimati.

L’indicatore di plausibilità della deviazione standard ( $IP^\sigma$ ) è calcolato considerando le differenze (residui) tra i valori della deviazione standard del punteggio di classe, stimati dal modello ( $\hat{\sigma}_i$ ), e quelli effettivamente osservati ( $\sigma_i$ ):

$$RES_i^\sigma = \hat{\sigma}_i - \sigma_i$$

Queste differenze, se negative poste uguali a zero<sup>16</sup>, vengono standardizzate su scala 0-1 in modo da ottenere l’indice  $IP^\sigma$ :

$$IP_i^\sigma = \frac{RES_i^\sigma - RES_{min}^\sigma}{RES_{max}^\sigma - RES_{min}^\sigma}.$$

L’indicatore  $IP^\sigma$  presenterà valori prossimi all’unità in corrispondenza di una variabilità anomala dei punteggi (eccessivamente bassa rispetto a quella che prevede il modello) e, quindi, una necessità di correzione dal *cheating*.

La logica è quella di confermare i valori dell’indice di *cheating*, calcolati con la procedura precedente, quando la variabilità empirica risulta inferiore a quella stimata, addebitando questa differenza a un “abbattimento” artificioso della variabilità a causa del *cheating*.

<sup>16</sup> Differenze negative tra valori teorici ed empirici della deviazione standard sono indicative di assenza di *cheating* in quanto la variabilità riscontrata nei punteggi risulta maggiore di quella prevista dal modello, di conseguenza il *cheating* è ritenuto poco plausibile ed il relativo indicatore di plausibilità viene posto uguale a zero.

L'indicatore di plausibilità della correlazione ( $IP^{corr}$ ) è costruito sulla base della correlazione (interna alla classe) tra voti al I quadrimestre e punteggi INVALSI. Per ogni classe  $i$ -esima si calcola il corrispondente valore del coefficiente di correlazione di Bravais-Pearson:

$$p_i = \frac{\sum_s (Q_{si} - \bar{Q}_i)(P_{si} - \bar{P}_i)}{\sum_s (Q_{si} - \bar{Q}_i)^2 \sum_s (P_{si} - \bar{P}_i)^2}$$

dove  $Q_{si}$  e  $P_{si}$  indicano, rispettivamente, il voto al I quadrimestre e il punteggio alla prova INVALSI di ogni studente  $s$ -esimo della classe  $i$ -esima.

Si parte dal presupposto che, sebbene i punteggi INVALSI e le votazioni al I quadrimestre siano relativi a valutazioni tra di loro differenti, ci sia comunque una relazione concordante tra i due punteggi (questa relazione viene confermata anche dai dati del campione). Di conseguenza, calcolando il coefficiente di correlazione all'interno delle classi (correlazione *within*), si può supporre poco plausibile il *cheating* se questa relazione è molto forte, mentre si considerano valori bassi della correlazione come segnali di un'effettiva presenza dello stesso. Per costruire il corrispondente indicatore di plausibilità si standardizza su scala 0-1 il coefficiente di correlazione<sup>17</sup> uniformandone la direzionalità a quella dell'indicatore di plausibilità della deviazione standard:

$$IP_i^{corr} = \frac{p_{max} - p_i}{p_{max} - p_{min}}$$

In sintesi, l'indicatore  $IP^{corr}$  presenterà valori prossimi all'unità se, per la classe  $i$ -esima, si riscontra un'assenza di relazione tra voti al I quadrimestre e punteggi INVALSI (correlazione prossima allo zero); tale assenza induce a sospettare dell'attendibilità dei punteggi conseguiti alla prova e di conseguenza a confermare i valori dell'indice di *cheating*. Al contrario, l'indicatore  $IP^{corr}$  tenderà a zero quando la relazione tra i voti al I quadrimestre e i punteggi INVALSI è forte: in questo caso si potrà ritenere che eventuali punteggi elevati alla prova INVALSI siano coerenti con le *performance* scolastiche della classe e, quindi, che non vada effettuata una correzione per il *cheating* minimizzando l'indice di *cheating* originale.

<sup>17</sup> Valori negativi dell'indice di correlazione (relazione discordante) oppure una mancanza del dato relativo al voto del I quadrimestre vengono considerati come una spia di *cheating* e, di conseguenza, si pone lo stesso coefficiente pari a zero.

Un ulteriore indicatore utilizzato nella rimodulazione dell'indice di *cheating* è il fattore di correzione  $C_o$ , che è calcolato a seguito della correzione del punteggio medio di classe, sulla base dei punteggi medi stimati dal modello di regressione, effettuata propedeuticamente alla "revisione" dell'indice di *cheating*. Questa correzione consiste in una trasformazione dei punteggi medi di ogni classe attraverso la media tra il punteggio osservato nella classe e quello teorico previsto dal modello. La media viene calcolata, esclusivamente per le classi che non rientrano nel campione e per le quali il punteggio empirico risulta maggiore del teorico, utilizzando l'indice di determinazione lineare del modello di regressione come variabile di ponderazione:

$$P_i^{new} = \begin{cases} (P_i R_{adj}^2) + [\hat{P}_i(1 - R_{adj}^2)] & \text{se } P_i > \hat{P}_i \\ P_i & \text{se } P_i \leq \hat{P}_i \end{cases}$$

dove  $P_i^{new}$  indica il "nuovo" punteggio della classe, mentre  $\hat{P}_i$  indica il punteggio stimato dal modello e  $R_{adj}^2$  il coefficiente di determinazione lineare corretto.

Di conseguenza, il fattore di correzione  $C_o$  indica il grado di "aggiustamento" del punteggio originale della classe scaturito dal confronto con il punteggio del modello<sup>18</sup>:

$$C_o = \frac{P_i^{new}}{P_i}$$

Utilizzando i due indici di plausibilità  $IP^\sigma$  e  $IP^{corr}$  e il fattore di correzione  $C_o$  viene rimodulato l'indice di propensione al *cheating* utilizzando la seguente formulazione:

$$ICH_{new} = 1 - [C_o (1 - ICH_{old} M_g(IP))]$$

dove  $ICH_{old}$  è l'indice di *cheating* calcolato con la procedura precedente,  $C_o$  è il fattore di correzione,  $M_g(IP)$  è la media geometrica<sup>19</sup> dei due indicatori di plausibilità:

$$M_g(IP) = \sqrt{IP_i^\sigma \cdot IP_i^{corr}}$$

<sup>18</sup> Poiché la correzione del punteggio medio di classe è applicata solo se il punteggio osservato risulta maggiore di quello stimato si verificherà che  $P_i^{new} \leq P_i$  e di conseguenza  $0 \leq C_o \leq 1$ .

<sup>19</sup> Se uno dei due indicatori di plausibilità risulta nullo si ricorrerà alla media aritmetica.

### A.7.3 Valutazione dei risultati della revisione

La valutazione dell'impatto della revisione della procedura di *cheating* verrà effettuata confrontando i dati corretti con la nuova procedura sia con quelli originali (non corretti per il *cheating*), sia con quelli corretti con la procedura precedente. Risulta importante premettere che la revisione della metodologia è stata condotta con lo scopo principale di minimizzare il numero di falsi positivi, cioè quelle situazioni in cui vengono rilevate delle *performance* elevate a livello di classe "generate" da un livello effettivamente elevato delle competenze degli studenti ma che, erroneamente, la procedura di correzione attribuisce alla presenza di *cheating*. In questa ottica, un primo effetto positivo della revisione si può riscontrare nel significativo decremento delle segnalazioni inviate dalle scuole per rivendicare la bontà delle loro *performance*. Nell'a.s. 2011/12 le scuole che "lamentavano"<sup>20</sup> un'attribuzione, erronea, di un indice di *cheating* elevato ammontavano a oltre 150, mentre nell'a.s. 2012/13 sono arrivate a circa 45.

Per valutare analiticamente gli effetti delle modifiche apportate alla procedura di individuazione e correzione del *cheating* si prenderanno in considerazione i dati relativi alla prova di Italiano nella classe V della scuola primaria<sup>21</sup>, per facilitare il confronto la procedura adottata fino all'a.s. 2011/12 sarà definita "procedura 2012", mentre l'approccio *multistage*, adottato dall'a.s. 2012/13, verrà definito come "procedura 2013". Nella Tavola A.7.1 si riportano alcuni valori caratteristici delle distribuzioni dei punteggi di classe. I risultati del confronto evidenziano che sia la procedura 2012 che quella 2013 tendono a correggere la distribuzione intervenendo sui percentili più elevati della distribuzione.

<sup>20</sup> Le motivazioni che le scuole adducono a riprova di un'attribuzione errata di alti valori di *cheating* consistono, generalmente nell'aver svolto uno scrupoloso controllo durante le prove e/o un elevato livello medio di competenze della classe comprovato già dai voti degli anni passati e del I quadrimestre.

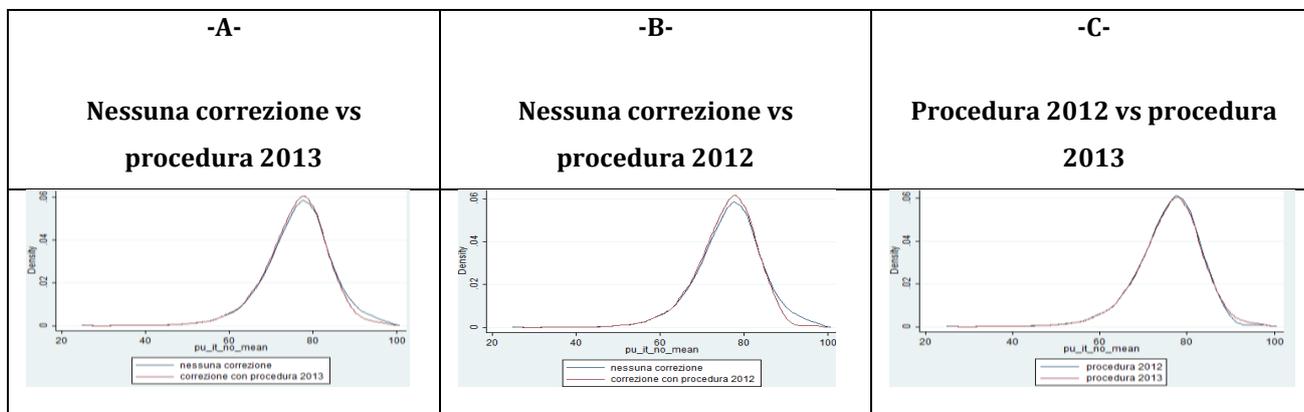
<sup>21</sup> La scelta della V primaria è dovuta alla possibilità di analizzare i dati di contesto (non rilevati per la II primaria e per la III secondaria di I grado) e ad un livello di *cheating* in genere più elevato rispetto a quello osservato nella I secondaria di I grado e nella II secondaria di II grado.

**Tavola A.7.1 - Valori caratteristici della distribuzione dei punteggi medi di classe della V primaria per l'a.s. 2012/13 (ambito Italiano). Confronto tra distribuzioni originali (non corrette per il *cheating*), corrette con procedura 2012 e corrette con procedura 2013**

Valori caratteristici	Punteggi non corretti per il <i>cheating</i>	Punteggi corretti con procedura 2012	Punteggi corretti con procedura 2013	
Media	76,5267	75,8236	76,0204	
Deviazione std.	7,90178	7,35051	7,55954	
Asimmetria	-0,467	-0,735	-0,604	
Curtosi	1,307	1,650	1,492	
Percentili	90th	85,8537	84,3089	84,8036
	91th	86,3415	84,6669	85,1441
	92th	86,8564	85,0333	85,5691
	93th	87,4238	85,3659	86,0374
	94th	88,1098	85,8537	86,5854
	95th	88,8401	86,3415	87,1642
	96th	89,8084	86,9106	87,8691
	97th	90,9892	87,6098	88,7805
	98th	92,5206	88,4972	90,2439
	99th	94,6085	90,1720	92,5855

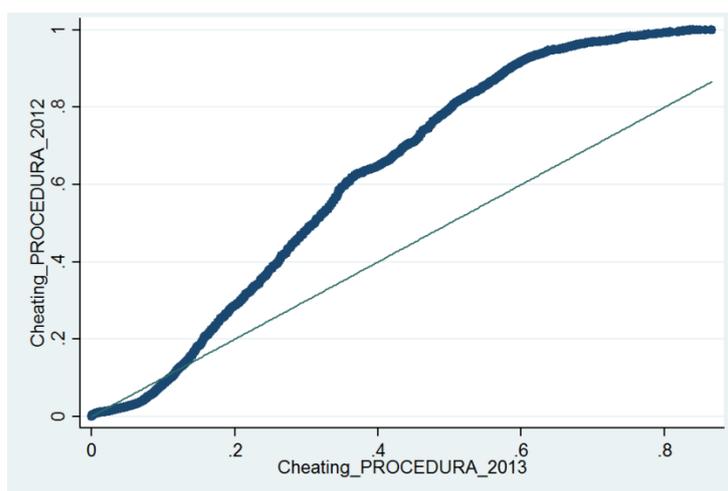
Di conseguenza sia i valori medi che la variabilità dei punteggi risultano inferiori rispetto a quelli non corretti. Si osserva che la correzione apportata dalla procedura 2013 risulta meno marcata salvaguardando maggiormente la variabilità della distribuzione, la simmetria e la curtosi. Inoltre, al crescere dei percentili della distribuzione, tra le due procedure si notano delle differenze più sostanziali, che si traducono in una minore propensione a segnalare falsi positivi da parte della procedura 2013. Una conferma di queste considerazioni si ottiene dal confronto grafico tra le distribuzioni dei punteggi (Figura A.7.2). La comparazione rileva come le differenze tra le due metodologie si concentrano nella parte “alta” delle distribuzioni, in particolare si nota (pannelli A e B) che entrambi gli approcci intervengono sulle classi con valori prossimi ai massimi della distribuzione, spingendo verso il basso la curva in corrispondenza dei percentili più elevati, anche se la correzione avviene in modo più marcato con la procedura 2012 rispetto a quella 2013. A conferma, dal confronto diretto (pannello C) tra correzione con il metodo del 2012 e quello del 2013, si nota che la nuova procedura preserva parte delle classi *top performers*, che vengono invece “annullate” dalla metodologia precedente.

**Figura A.7.2 - Rappresentazione grafica delle distribuzioni dei punteggi medi di classe della V primaria per l'a.s. 2012/13 (ambito Italiano). Confronto tra distribuzioni originali (non corrette per il *cheating*), corrette con la procedura 2012 e corrette con la procedura 2013**



Il minore impatto della procedura 2013 si può osservare anche focalizzandosi sulle distribuzioni dell'indice di *cheating* calcolato sulla base delle due procedure. Il confronto degli indici di *cheating* per ogni percentile delle loro distribuzioni, cfr. il grafico quantile-quantile in Figura A.7.3, riprova che la procedura revisionata interviene in modo più *soft* attribuendo, specialmente per i percentili più elevati, punteggi di *cheating* minori di quella precedente.

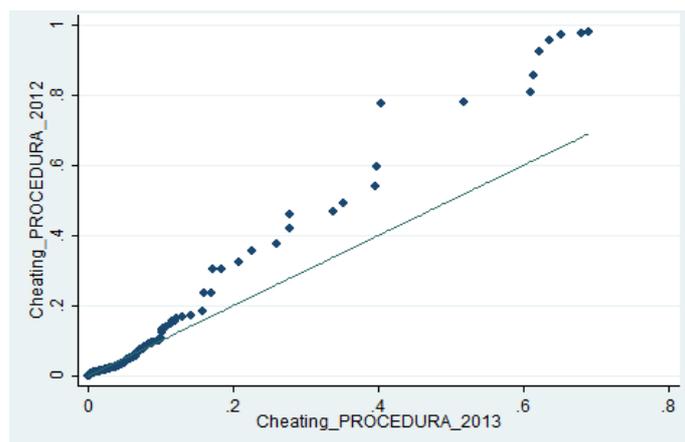
**Figura A.7.3 - Confronto tra percentili degli indici di *cheating* elaborati utilizzando la procedura 2012 e quella 2013 sulla base dei dati della V primaria per l'a.s. 2012/13 (Italiano)**



La nuova procedura di correzione presuppone che le classi campione non siano affette dal *cheating* e assegna un fattore correttivo (peso) pari a 1 (corrispondente a *cheating* nullo) a queste classi. Allo scopo di valutare i risultati della nuova procedura si è simulato il calcolo dell'indice di *cheating* anche per queste unità e si è proceduto al confronto tra la metodologia vecchia e la nuova, restringendo l'analisi alle sole classi campione (Figura A.7.4). In particolare,

ipotizzando che la presenza di osservatori esterni abbia reso nullo il *cheating* nelle classi campione, si valuterà positivamente la procedura che rende nullo o comunque minimizza la stima della propensione al *cheating* per queste unità.

**Figura A.7.4 - Confronto tra percentili degli indici di *cheating* elaborati utilizzando la procedura 2012 e quella 2013 sulla base dei dati campionari della V primaria per l'a.s. 2012/13 (Italiano)**



Il grafico *quantile-quantile* (Q-Q plot) permette di evidenziare come entrambe le procedure attribuiscono bassa propensione al *cheating* per la maggior parte delle 1.378 classi campione. In particolare, se si considera la procedura 2012 il 98° percentile è pari a 0,32 mentre risulta pari a 0,13 se si osservano i valori della procedura 2013. Questo dato attesta, ulteriormente, la maggiore correttezza della procedura 2013 nella stima dei valori di *cheating*.

