

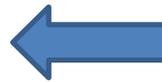
Le metodologie di campionamento e l'affidabilità dei dati delle rilevazioni nazionali INVALSI

Roma - 19 novembre 2020

Piero Demetrio Falorsi

Patrizia Falzetti

Roberto Ricci



Lavoro realizzato in base a una pluriennale e consolidata collaborazione tra l'Istat e l'INVALSI



Illustrare gli aspetti tecnico-metodologici e risultati del piano di campionamento delle prove INVALSI



Risultati presentati nel volume http://ojs.francoangeli.it/_omp/index.php/oa/catalog/book/445

La **metodologia rigorosa** permette di guadagnare la fiducia dell'opinione pubblica (*public trust*) nei risultati e si basa sui seguenti tre principi:

1. **Rispetto e dialogo** tra **Istituzioni** diverse
2. **Rispetto e dialogo** tra **competenze** diverse
3. **Buona pratica nella restituzione dell'informazione agli utenti** (accesso ai microdati e possibilità di calcolo autonomo dell'accuratezza)

Necessità di investire nel *capacity building*

- ✓ Non soluzioni chiuse **chiavi in mano**
- ✓ **Soluzioni rigorose**, ma **gestibili autonomamente** dall'ente che è responsabile di produrre la statistica
- ✓ Gestione trasparente **dei trade off** tra accuratezza e costi

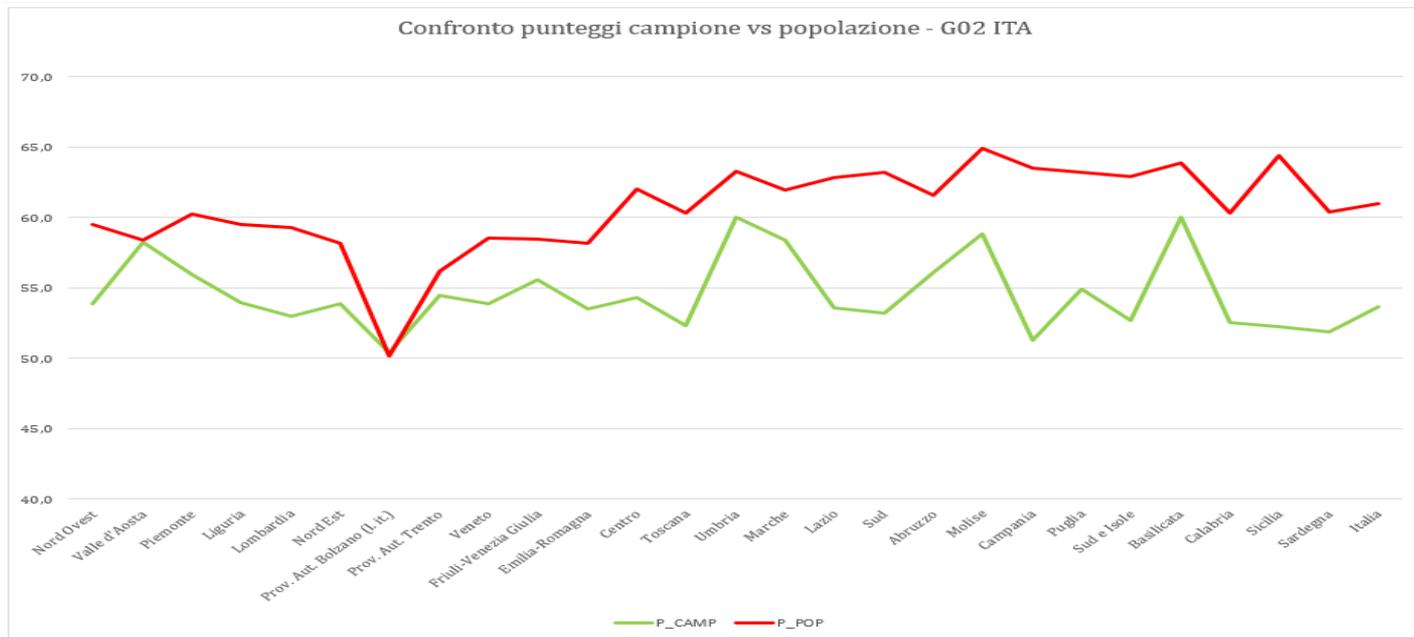
- Le rilevazioni INVALSI
- Il sistema delle prove INVALSI
- Obiettivi dell'indagine
- Campionamento: allocazione del campione
- Stime campionarie
- Stime della varianza
- I risultati: significatività, variabilità, affidabilità

- ✓ Analogamente a quanto è avvenuto in molti paesi industrializzati, anche l'Italia si è dotata abbastanza recentemente di un Sistema Nazionale di Valutazione (SNV).
- ✓ L'INVALSI è stato chiamato a effettuare precise scelte tecniche e strategiche per garantire la realizzazione di un SNV 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.

Due tipi di rilevazione:

- ✓ **Campione Nazionale Probabilistico** (con presenza di **osservatore esterno**) per controllare il *bias*
- ✓ **Rilevazione Censuaria** (**in presenza degli insegnanti**)

Ragione di questa scelta



In ogni classe campione della prova è prevista la presenza di un osservatore esterno. Il compito dell'osservatore esterno 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 da INVALSI (solo per la scuola primaria). Mentre per i gradi 8, 10 e 13 l'acquisizione delle risposte, grazie alla modalità di somministrazione CBT (Computer Based Testing) è automatizzata. Pertanto l'intervento dell'osservatore è a salvaguardia del protocollo di somministrazione.

La scelta di prevedere un'indagine campionaria con osservatori esterni è connessa al fatto che questi ultimi sono più efficaci nel garantire il pieno rispetto del protocollo di somministrazione delle prove (eliminazione o forte riduzione del fenomeno noto con il termine *cheating*) e, di conseguenza, le valutazioni che si ottengono, 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.

MATERIA	GRADO 2	GRADO 5	GRADO 8	GRADO 10	GRADO 13
ITALIANO	Carta e penna	Carta e penna	CBT	CBT	CBT
MATEMATICA	Carta e penna	Carta e penna	CBT	CBT	CBT
INGLESE	-	Carta e penna	CBT	-	CBT

Valori medi dei punteggi riportati alle prove (Lettura, Italiano, Matematica, Inglese):

$$\bar{y} = \frac{Y}{A}$$

Totale dei punteggi

Totale alunni

$$Y = \sum_h \sum_i \sum_j \sum_k y_{hijk}$$

$$A = \sum_h \sum_i \sum_j a_{hij}$$

k = alunno

j = classe

i = istituto

h = regione

Componenti della devianza tra i punteggi, distinguendo la variabilità

- all'interno della classe **D3**,
- tra classi **D2** dello stesso Istituto e
- tra Istituti differenti **D1**.

$$D = \sum_h \sum_i \sum_j \sum_k (y_{hij} - \bar{y})^2$$

$$D = D_1 + D_2 + D_3$$

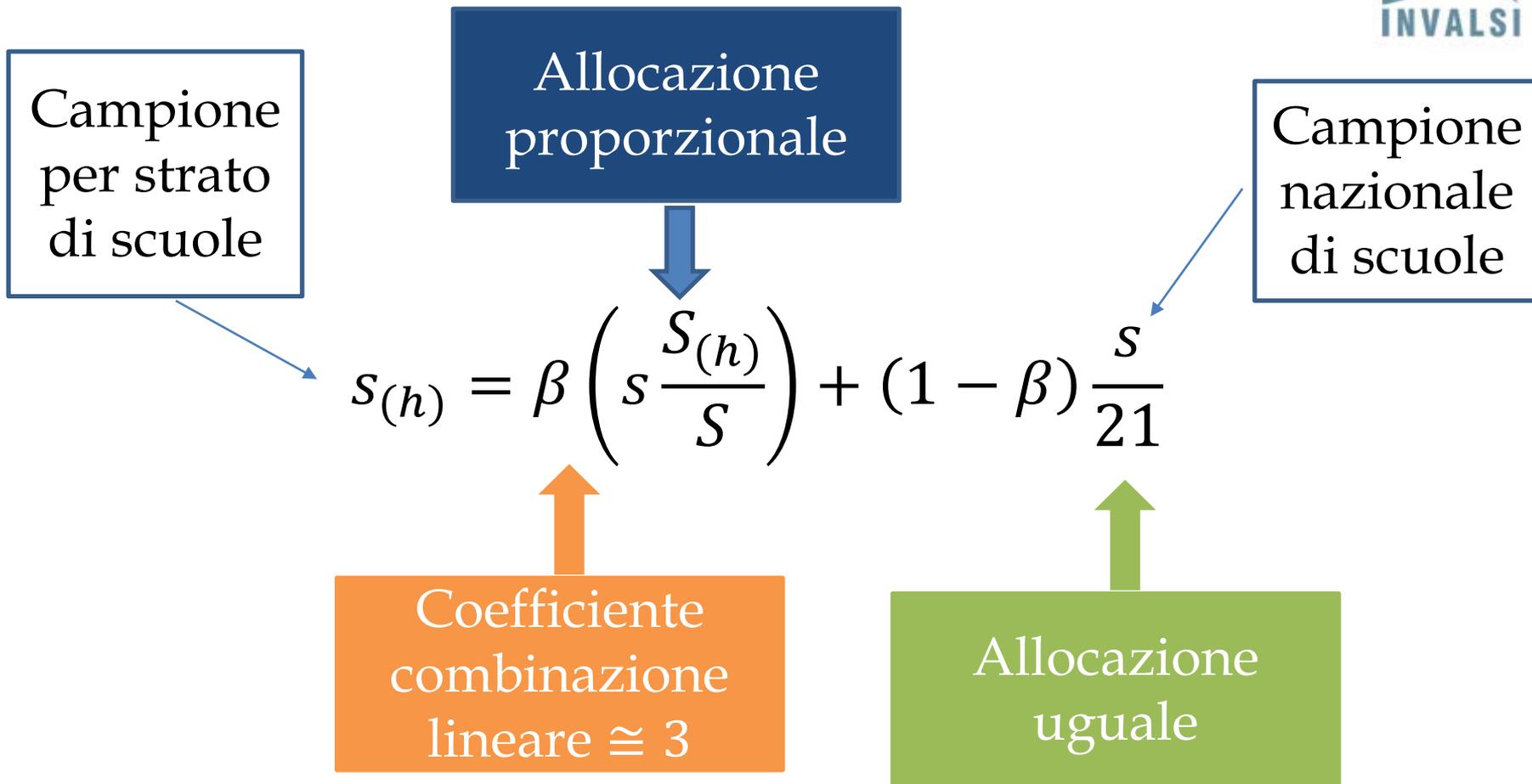
Domini di stima

$$286 = \begin{cases} 22 \text{ aree (Italia, Regioni + Prov aut.)} \\ \times \\ 13 \text{ tipologie di scuola} \end{cases}$$

Tipologia di scuola	Grado scolastico	Sottopopolazione
Scuola primaria	2	1
	5	2
Scuola secondaria di primo grado	8	3
Licei scientifici	10 - 13	4 - 5
Licei classici	10 - 13	6 - 7
Altri licei	10 - 13	8 - 9
Istituti tecnici	10 - 13	10 - 11
Istituti professionali	10 - 13	12 - 13

Il numero di scuole campione in una determinata regione, $S(h)$, viene determinato in due passi:

1. Viene definito a priori, attraverso un'analisi costi-benefici, il numero totale di scuole da campionare per un particolare grado scolastico, S
2. Il numero finale di scuole viene definito mediante allocazione di compromesso tra:
 - ✓ un **allocazione proporzionale** (privilegia accuratezza stima nazionale e penalizza piccole regioni)
 - ✓ un **allocazione uguale** (penalizza stima nazionale e grandi regioni e privilegia piccole regioni)



S → totale scuole

$S_{(h)}$ → totale scuole per strato

Il valore di β 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

Codice regione	campione ottimo nazionale	cv % campione ottimo nazionale
Valle D'Aosta	1	6,05
Piemonte	27	1,18
Liguria	9	2,07
Lombardia	61	0,73
Veneto	25	0,95
Friuli-Venezia Giulia	6	2,05
Emilia-Romagna	24	1,08
Toscana	22	1,26
Umbria	5	2,42
Marche	7	1,72
Lazio	39	1,04
Abruzzo	7	2,03
Molise	2	4,61
Campania	59	1,14
Puglia	30	1,27
Basilicata	3	3,40
Calabria	14	1,98
Sicilia	47	1,29
Sardegna	10	2,11
Prov. Aut. Bolzano (l. it.)	1	6,40
Prov. Aut. Trento	2	2,72
Totale	400	0,32



Come detto, viene definito a priori, attraverso un'analisi costi-benefici, il numero totale di scuole da campionare per un particolare grado scolastico. Nello specifico esempio che vedremo, Grado 8 (III secondaria di I grado), la numerosità del campione è pari a **400 scuole** e la prima allocazione che vedremo è l'**allocazione ottima**.

Codice regione	campione ottimo nazionale	cv % campione ottimo nazionale	campione equiriparto	cv % campione equiriparto
Valle D'Aosta	1	6,05	19	0,22
Piemonte	27	1,18	19	1,41
Liguria	9	2,07	19	1,35
Lombardia	61	0,73	19	1,33
Veneto	25	0,95	19	1,10
Friuli-Venezia Giulia	6	2,05	19	1,07
Emilia-Romagna	24	1,08	19	1,23
Toscana	22	1,26	19	1,37
Umbria	5	2,42	19	1,06
Marche	7	1,72	19	1,02
Lazio	39	1,04	19	1,50
Abruzzo	7	2,03	19	1,15
Molise	2	4,61	19	0,91
Campania	59	1,14	19	2,08
Puglia	30	1,27	19	1,60
Basilicata	3	3,40	19	1,28
Calabria	14	1,98	19	1,67
Sicilia	47	1,29	19	2,08
Sardegna	10	2,11	19	1,47
Prov. Aut. Bolzano (l. it.)	1	6,40	19	-
Prov. Aut. Trento	3	2,72	19	0,94
Totale	400	0,32	400	0,45

La seconda allocazione è **l'allocazione equiripartita.**



La terza allocazione è l'allocazione di compromesso.

Codice regione	campione ottimo nazionale	cv % campione ottimo nazionale	campione equiriparto	cv % campione equiripartito	campione di compromesso 0,3	cv campione di compromesso 0,3
Valle D'Aosta	1	6,05	19	0,22	14	0,74
Piemonte	27	1,18	19	1,41	21	1,33
Liguria	9	2,07	19	1,35	16	1,49
Lombardia	61	0,73	19	1,33	32	1,03
Veneto	25	0,95	19	1,10	21	1,05
Friuli-Venezia Giulia	6	2,05	19	1,07	15	1,22
Emilia-Romagna	24	1,08	19	1,23	21	1,18
Toscana	22	1,26	19	1,37	20	1,33
Umbria	5	2,42	19	1,06	15	1,25
Marche	7	1,72	19	1,02	15	1,14
Lazio	39	1,04	19	1,50	25	1,31
Abruzzo	7	2,03	19	1,15	15	1,29
Molise	2	4,61	19	0,91	14	1,25
Campania	59	1,14	19	2,08	31	1,61
Puglia	30	1,27	19	1,60	22	1,48
Basilicata	3	3,40	19	1,28	14	1,52
Calabria	14	1,98	19	1,67	18	1,75
Sicilia	47	1,29	19	2,08	28	1,72
Sardegna	10	2,11	19	1,47	16	1,61
Prov. Aut. Bolzano (l. it.)	1	6,40	19	-	14	0,83
Prov. Aut. Trento	3	2,72	19	0,94	14	1,13
Totale	400	0,32	400	0,45	400	0,38



Un CV troppo alto porterebbe, ex-post, a un errore standard della stima regionale troppo elevato. L'idea è quella di uniformare il più possibile i CV regionali mantenendo inalterato il CV nazionale.

Ripartizione geografica	campione di compromesso 0,3 (Originale)	cv % campione di compromesso 0,3 (Originale)
Valle D'Aosta	15	1,63
Piemonte	28	1,86
Liguria	21	2,00
Lombardia	40	1,40
Veneto	28	1,56
Friuli-Venezia Giulia	20	1,79
Emilia-Romagna	27	1,60
Toscana	26	1,85
Umbria	19	1,81
Marche	21	1,81
Lazio	32	1,80
Abruzzo	21	2,18
Molise	18	2,49
Campania	42	2,55
Puglia	30	2,16
Basilicata	19	2,59
Calabria	24	3,01
Sicilia	35	2,56
Sardegna	20	1,87
Prov. Aut. Bolzano (l. it.)	12	1,66
Prov. Aut. Trento	19	1,58
Italia	517	0,55

Si individuano le regioni sulle quali intervenire e il margine di aumento della numerosità campionaria, quindi nello specifico esempio:

- è stato possibile aumentare la numerosità campionaria da 517 a 525 scuole
- è stato possibile intervenire riducendo il numero di scuole nelle regioni nelle quali il CV era basso

Ripartizione geografica	campione di compromesso 0,3 (Modificato)	cv % campione di compromesso 0,3 (Modificato)	campione di compromesso 0,3 (Originale)	cv % campione di compromesso 0,3 (Originale)
Valle D'Aosta	15	1,63	15	1,63
Piemonte	28	1,86	28	1,86
Liguria	21	2,00	21	2,00
Lombardia	▼ 36	1,48	40	1,40
Veneto	28	1,56	28	1,56
Friuli-Venezia Giulia	20	1,79	20	1,79
Emilia-Romagna	▼ 25	1,66	27	1,60
Toscana	26	1,85	26	1,85
Umbria	19	1,81	19	1,81
Marche	21	1,81	21	1,81
Lazio	32	1,80	32	1,80
Abruzzo	21	2,18	21	2,18
Molise	18	2,49	18	2,49
Campania	▲ 43	2,51	42	2,55
Puglia	30	2,16	30	2,16
Basilicata	▲ 21	2,45	19	2,59
Calabria	▲ 34	2,50	24	3,01
Sicilia	▲ 37	2,50	35	2,56
Sardegna	20	1,87	20	1,87
Prov. Aut. Bolzano (l. it.)	12	1,66	12	1,66
Prov. Aut. Trento	19	1,58	19	1,58
Italia	525	0,55	517	0,55

Tab. 1 – Totale delle classi e degli studenti per grado scolastico coinvolti nella Rilevazione per l'anno scolastico 2018/19

<i>Grado scolastico</i>	<i>Totale classi</i>	<i>Totale classi campione</i>	<i>Totale studenti</i>	<i>Totale studenti campione indagine ICC</i>
2°: II primaria	28.546	1.362	522.266	26.143
5°: V primaria	29.481	1.362	556.875	26.959
8°: III secondaria primo grado	27.333	1486	563.726	31.393
10°: II secondaria secondo grado	26.406	1.922	536.782	40.893
13°: V secondaria secondo grado	26.285	2.086	487.282	40.717
Totale	138.051	8.218	2.666.931	166.105

Una volta determinata la dimensione campionaria in termini di scuole per regione (unità di I stadio), si passa all'estrazione delle:

- **scuole**
- **classi** all'interno delle scuole campione di norma non superiore alle due unità per istituto (una classe nel caso in cui la scuola estratta ne abbia solo una). Tutti gli studenti delle classi estratte entrano a far parte del campione

Le stime sono ottenute ponderando i dati campionari

1. **Peso scuola** (primo stadio)

- ✓ Riproduce il totale di scuole.
- ✓ Moltiplicato per il totale alunni (da registro) riproduce il totale noto degli alunni per regione e strato.
- ✓ Permette di ottenere stime considerando le scuole come unità statistiche.

2. **Peso classe** (secondo stadio)

- ✓ Riproduce il totale di alunni della scuola per l'anno scolastico

Peso finale alunno = **Peso scuola** x **Peso classe**

- ✓ Riproduce il totale noto degli alunni per regione e strato.

Stima rapporto

Stima totale dei punteggi

$$\hat{y} = \frac{\hat{Y}}{\hat{A}}$$

Stima totale alunni

dove

$$\hat{Y} = \sum_h \sum_i \sum_j \sum_k y_{hijk} w_{hijk}$$

$$\hat{A} = \sum_h \sum_i \sum_j \sum_k w_{hijk}$$

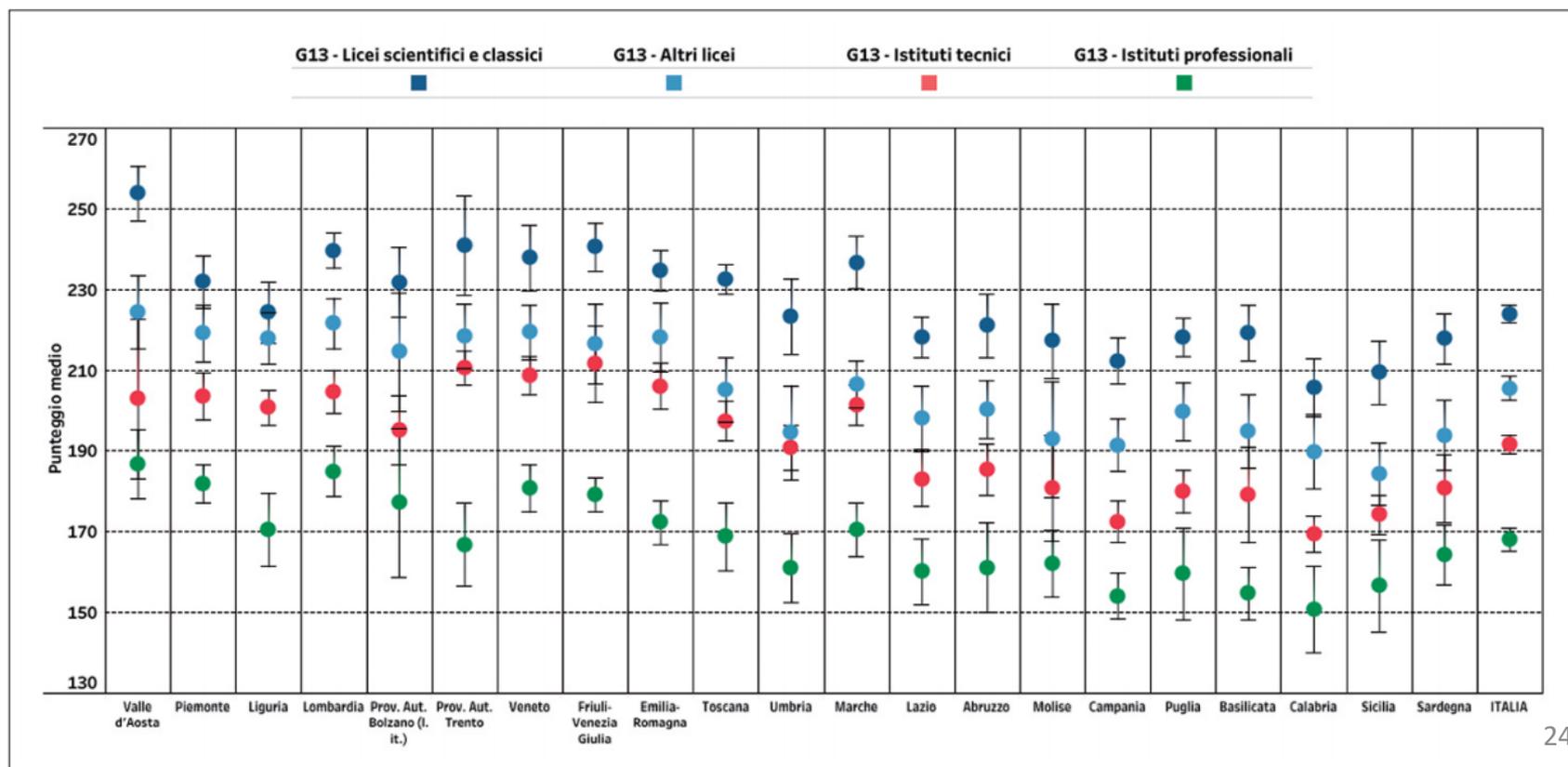
Peso campionario

Somme sul campione

k = alunno
j = classe
i = istituto
h = regione

La varianza delle stime \hat{y} e $\hat{y}_{(h)}$ viene stimata con metodi standard. In tal modo è possibile calcolare gli intervalli di confidenza

Fig. 1 – Risultati per tipologia di istituto – Italiano grado 13

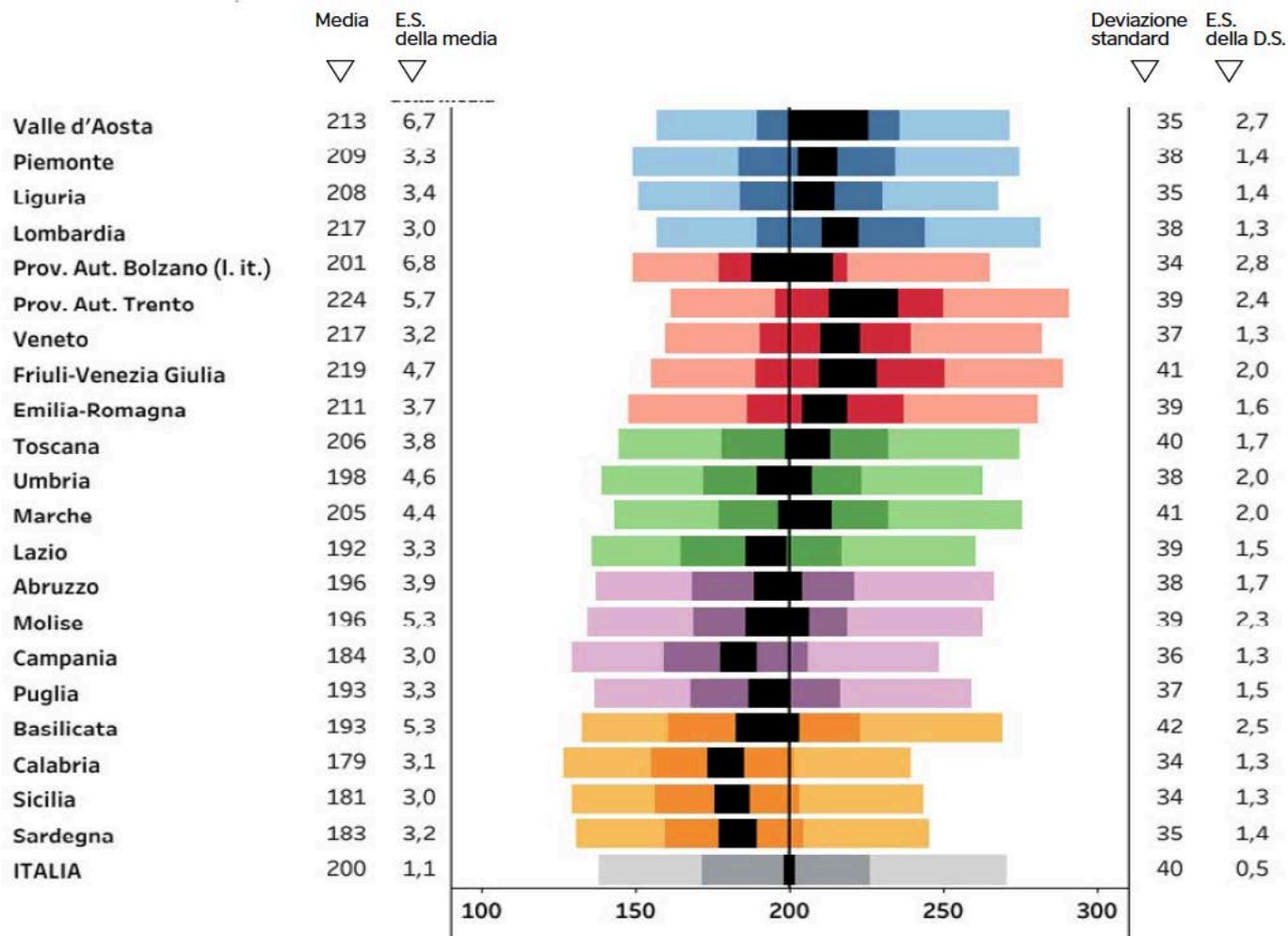


Gli utenti che hanno accesso ai dati elementari possono calcolare autonomamente l'accuratezza delle proprie stime

$$V(\hat{y}) = V_{(srs)}(\hat{y}) deff$$

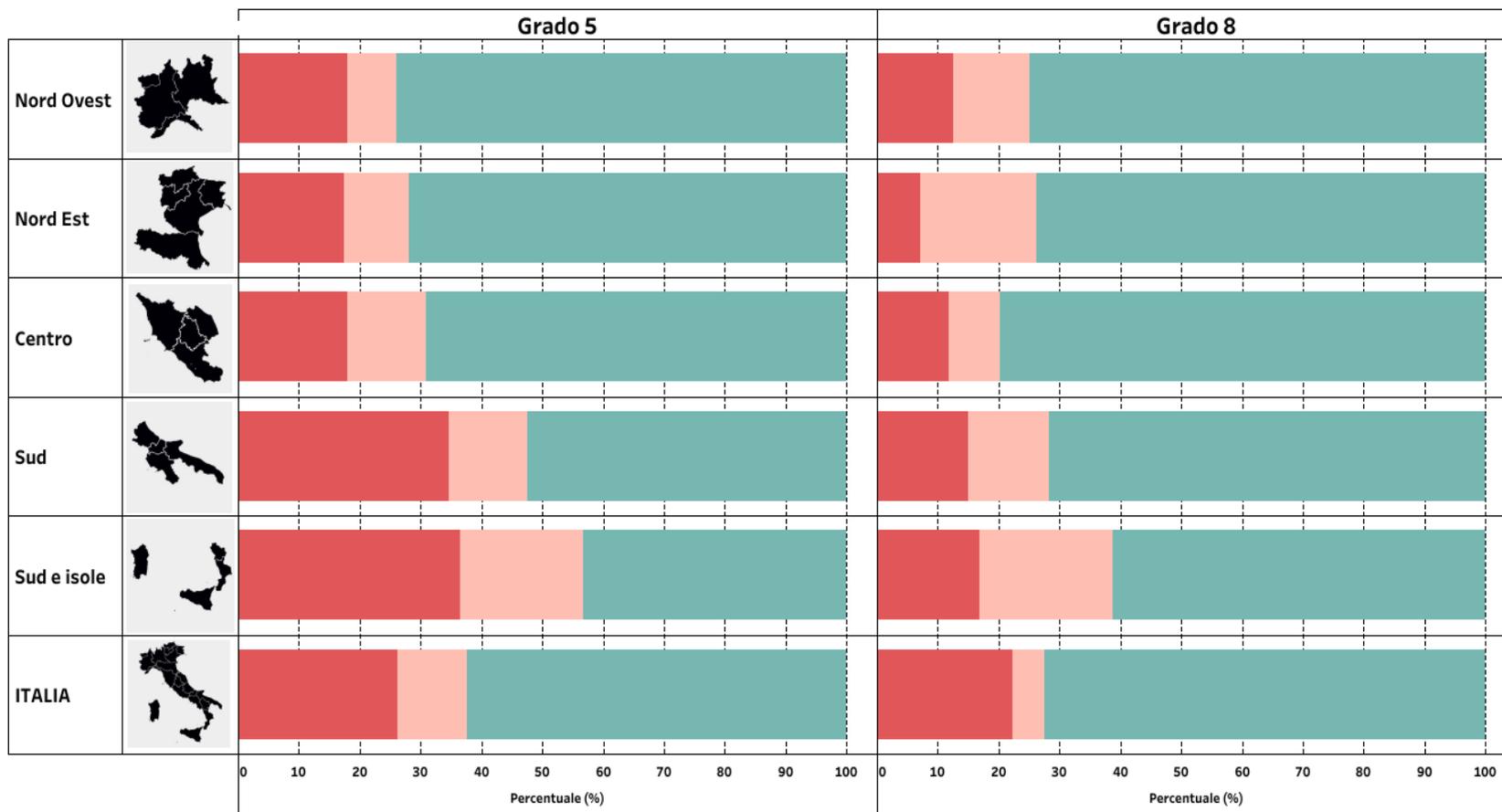
$$deff = 1 + \rho(\bar{a} - 1)$$

Nel **PAR 7.3.** sono forniti gli elementi per cui gli utenti possano calcolare autonomamente il *deff*



La variabilità *macro* dei risultati di Inglese-*listening*

■ Tra le scuole
 ■ Tra le classi entro le scuole
 ■ Tra gli alunni entro le classi



Istituzione scolastica nel suo complesso

Grafico 1a: [] - Scuola Secondaria di Secondo Grado - Classi seconde - Prova di Italiano: Incidenza della variabilità TRA le classi rispetto alla variabilità totale nel punteggio di ogni prova e nell'indice di background familiare (ESCS)³

Grafico 1a: [] Scuola Secondaria di Secondo Grado - Classi seconde - Prova di Italiano: Incidenza della variabilità TRA le classi rispetto alla variabilità totale nel punteggio di ogni prova e nell'indice di background familiare (ESCS)³

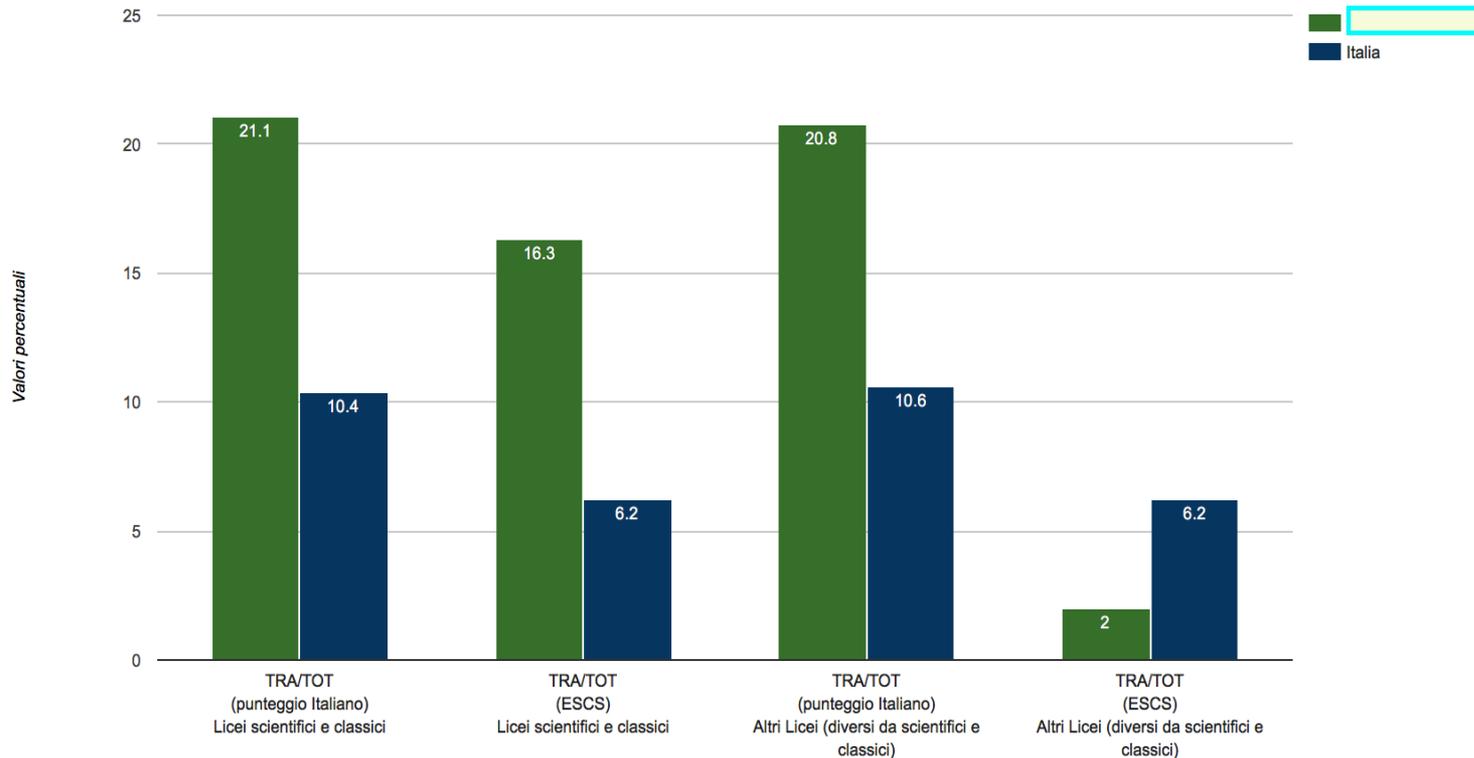
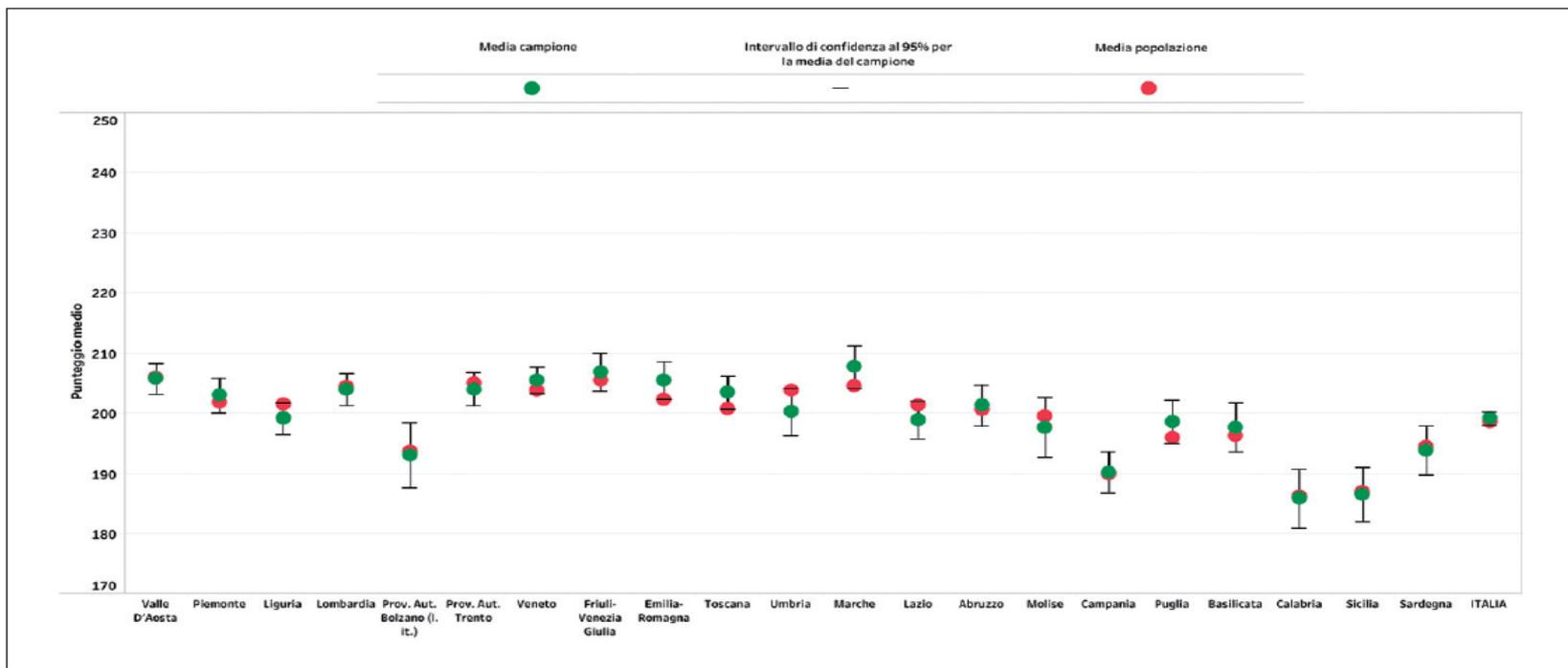


Fig. 11 – Andamento campione vs popolazione – G8 Italiano



Grazie per l'attenzione