Centro Servizi Amministrativi di Alessandria - Ufficio Studi e Programmazione

Elementi di statistica

per le prove INVALSI

- Media aritmetica
- Deviazione standard
- Istogrammi
- Quartili
- Risultati delle prove INVALSI

2004

pubblicato su sito web http://provstudi.volta.alessandria.it

Introduzione

Questa breve dispensa è basata sui lucidi presentati all'incontro del 17/02/04 con i Dirigenti Scolastici e i Coordinatori del Progetto Pilota 3 delle scuole della provincia di Alessandria.

Lo scopo è quello di offrire, senza alcuna pretesa né di completezza né di assoluto rigore, qualche elementare nozione di carattere statistico, utile soprattutto a quanti siano privi di conoscenze in materia.

Vengono focalizzati, in particolare, alcuni aspetti non sufficientemente evidenziati nella documentazione fornita dall'INVALSI, ma che appaiono essenziali per una corretta interpretazione dei risultati delle prove.

Pier Luigi Orsi
C.S.A. Alessandria
Ufficio Studi e Programmazione

Sommario

1 – Media aritmetica

La **media aritmetica** di un insieme di punteggi, molto familiare nella scuola, è anche il parametro più semplice utilizzato per confronti fra i risultati delle prove INVALSI.
In effetti la media riassume una serie di dati (ad esempio i risultati di un'intera classe) in un solo numero, ma occorre una certa prudenza per valutarne la significatività.

Prendiamo ad esempio due classi A e B di 20 alunni ciascuna ed esaminiamo i voti (nella classica scala da 1 a 10) ottenuti in una determinata prova. Nella tabella di Figura 1.1.a tali voti sono ordinati dal migliore al peggiore. La media è **5** per entrambe: si può ritenere allora che le due classi abbiano presentato lo stesso rendimento?

Per chiarire meglio la situazione costruiamo una tabella di frequenza (Figura 1.1.b): per ogni punteggio (da 1 a 10) contiamo quanti sono gli alunni che hanno ottenuto tale punteggio (**frequenza**). Rappresentiamo poi gli stessi risultati con i relativi grafici, denominati appunto diagrammi di frequenza o istogrammi (Figura 1.2).

Esaminando appunto i due istogrammi si nota facilmente la grande differenza fra le due classi:
- nella classe A la distribuzione dei voti è pressoché simmetrica, molti alunni hanno ottenuto proprio il voto 5 (media) e la maggior parte ha ottenuto voti vicini alla media;
- nella classe B invece vi è un gruppo di alunni con risultati buoni e un altro gruppo con risultati scadenti; nessuno ha avuto il voto 5 e pochissimi voti vicini al 5.

Riassumendo: per la classe A la media 5 è un dato indicativo dell'andamento globale, mentre per la classe B la media è solo il risultato della compensazione di risultati molto validi o molto scadenti.

Per discriminare fra le due situazioni bisogna considerare, oltre alla media, anche un altro parametro: di solito si utilizza la **deviazione standard** o **scarto quadratico medio**.

	A	B
1	9	10
2	8	9
3	7	9
4	7	8
5	6	8
6	6	8
7	6	8
8	5	7
9	5	7
10	5	6
11	5	3
12	5	3
13	4	3
14	4	2
15	4	2
16	4	2
17	3	2
18	3	1
19	2	1
20	2	1
media	5	5

(a)

Punt.	freq.A	freq.B
10	0	1
9	1	2
8	1	4
7	2	2
6	3	1
5	5	0
4	4	0
3	2	3
2	2	4
1	0	3
totale	20	20
media	5	5

(b)

Figura 1.1

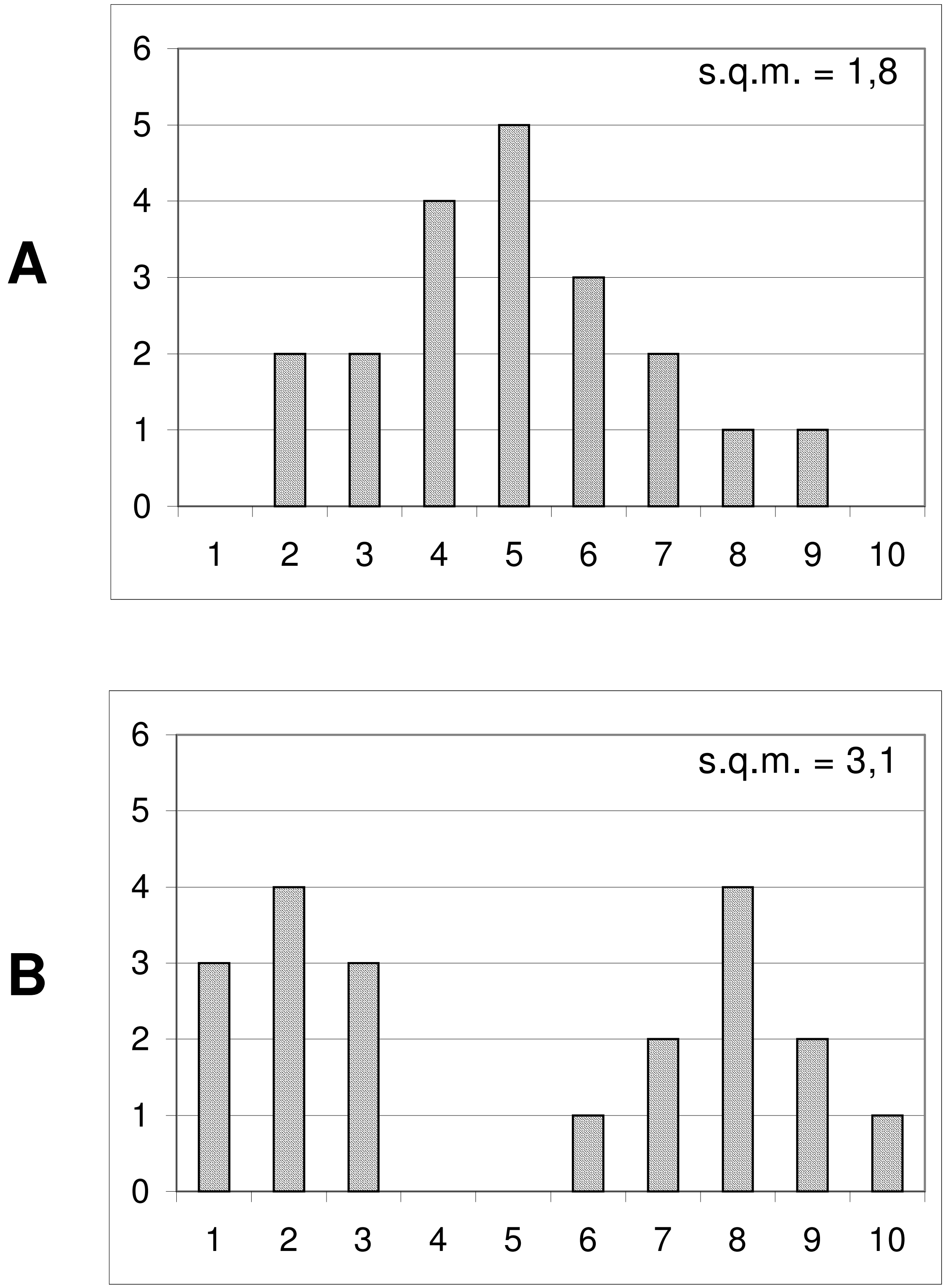


Figura 1.2

2 – Deviazione standard (**o scarto quadratico medio**).

Come abbiamo detto, per una lettura più significativa di un insieme di punteggi è opportuno affiancare alla media un cosiddetto **indice di dispersione**, cioè un numero che ci fa capire quanto i dati siano raggruppati attorno al valore centrale.

Vediamo come si può procedere.
Prendiamo la classe A (Figura 2.1) e calcoliamo, per ogni punteggio, la differenza dalla media (**scarto**); non ci interessa se la differenza sia positiva o negativa, ma solo quanto il punteggio sia "distante" dalla media (prendiamo perciò il valore assoluto).
Adesso facciamo la media di questi scarti: ciò che otteniamo è un numero che ci dice quanto i punteggi siano "distanti", in media, dalla media stessa (gioco di parole inevitabile).

Il risultato, come si vede, è **1,4** il che significa che i singoli punteggi differiscono in media di circa un voto e mezzo dalla media globale. Per la classe B (Figura 2.2) lo stesso procedimento porta ad un valore più che doppio (**3**).

Il risultato che abbiamo ottenuto, detto **scarto medio** dalla media, non è per la verità molto utilizzato in statistica. Per ragioni teoriche è più utilizzato un altro parametro, detto, come avevamo anticipato, **deviazione standard** o **scarto quadratico medio**.

Il calcolo è solo leggermente più complicato: una volta calcolati gli scarti dalla media dei singoli punteggi, eleviamoli al **quadrato** e poi determiniamo la media di questi quadrati.
Per completare il calcolo non resta che estrarre la **radice quadrata** del risultato.

Lo scarto quadratico medio così ottenuto differisce di poco, almeno nei nostri casi, dallo scarto medio calcolato prima ed ha un significato strettamente simile.

A

	punteggio	|scarto|	scarto²
1	9	4	16
2	8	3	9
3	7	2	4
4	7	2	4
5	6	1	1
6	6	1	1
7	6	1	1
8	5	0	0
9	5	0	0
10	5	0	0
11	5	0	0
12	5	0	0
13	4	1	1
14	4	1	1
15	4	1	1
16	4	1	1
17	3	2	4
18	3	2	4
19	2	3	9
20	2	3	9
media	5	1,4	$\boxed{3,3}$

$\sqrt{\ }$

s.q.m. : $\boxed{1,8}$

Figura 2.1

B

	punteggio	$\lvert\text{scarto}\rvert$	scarto^2
1	10	5	25
2	9	4	16
3	9	4	16
4	8	3	9
5	8	3	9
6	8	3	9
7	8	3	9
8	7	2	4
9	7	2	4
10	6	1	1
11	3	2	4
12	3	2	4
13	3	2	4
14	2	3	9
15	2	3	9
16	2	3	9
17	2	3	9
18	1	4	16
19	1	4	16
20	1	4	16
media	5	3	$\boxed{9{,}9}$

$\sqrt{}$

s.q.m. : $\boxed{3{,}1}$

Figura 2.2

3 – Istogrammi

Per affrontare in modo più completo il problema della rappresentazione di grosse quantità di dati attraverso istogrammi, prendiamo in considerazione un caso più complesso.

La tabella di Figura 3.1 rappresenta gli ipotetici punteggi ottenuti da una classe di 25 alunni (classe X) in una prova INVALSI. In questo caso la rappresentazione attraverso istogramma delle frequenze dei singoli punteggi, come nei casi precedenti, porta ad un risultato del tutto illeggibile.

1	2	3	4	5	6	7	8	9	10	11	12	13
99,5	84,8	81,3	77,2	75,7	74,2	71,6	68,3	56,4	55,1	52,5	48,7	48,3

14	15	16	17	18	19	20	21	22	23	24	25	media
46,2	44,9	43,7	38,4	37,3	33,6	31,8	28,6	26,9	18,1	14,7	5,8	50,5

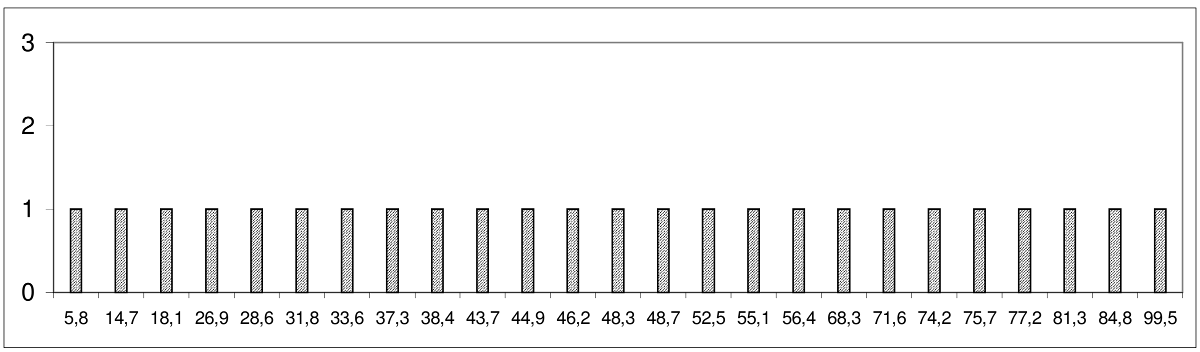


Figura 3.1

Una rappresentazione migliore può essere ottenuta considerando non singoli punteggi, ma intervalli di punteggio:
- ad esempio si può contare il numero di alunni che hanno ottenuto punteggi fra 0 e 5, 5 e 10, 10 e 15, ecc. (Figura 3.2.a);
- oppure si possono scegliere intervalli più ampi, come 0 – 10, 10 – 20, ecc. (Figura 3.2.b);
- o ancora 0 – 20, 20 – 40, ecc. (Figura 3.2.c).

Ciascun intervallo individua quella che viene chiamata una **classe di frequenza**.

Nelle tabelle la frequenza, oltre che come numero assoluto, viene indicata anche come frequenza **percentuale**, il che equivale a rapportare i singoli dati ad una classe di 100 alunni.
In genere questa è l'opzione più utilizzata, perché consente di confrontare popolazioni di grandezza anche molto diversa (ad esempio una singola classe di poche decine di alunni, con un ambito nazionale di milioni di alunni).
Bisogna però tenere presente, per inciso, che in una grossa popolazione la variazione, ad esempio, di un alunno in più o in meno non provoca alcuna variazione significativa in una percentuale, mentre in una piccola classe la percentuale rischia di oscillare in modo considerevole.

	freq.	%
95 -100	1	4
90 - 95	0	0
85 - 90	0	0
80 - 85	2	8
75 - 80	2	8
70 - 75	2	8
65 - 70	1	4
60 - 65	0	0
55 - 60	2	8
50 - 55	1	4
45 - 50	3	12
40 - 45	2	8
35 - 40	2	8
30 - 35	2	8
25 - 30	2	8
20 - 25	0	0
15 - 20	0	0
10 - 15	2	8
5 - 10	1	4
0 - 5	0	0
totale	25	100

(a)

	freq.	%
90 -100	1	4
80 - 90	2	8
70 - 80	4	16
60 - 70	1	4
50 - 60	3	12
40 - 50	5	20
30 - 40	4	16
20 - 30	2	8
10 - 20	2	8
0 - 10	1	4
totale	25	100

(b)

	freq.	%
80 -100	3	12
60 - 80	5	20
40 - 60	8	32
20 - 40	6	24
0 - 20	3	12
totale	25	100

(c)

Figura 3.2

La Figura 3.3 mostra gli istogrammi relativi alle distribuzioni sopra individuate, dove le singole frequenze sono espresse in percentuale.

Naturalmente l’esemplificazione potrebbe estendersi ad intervalli dalle più svariate ampiezze: in generale, considerando intervalli più piccoli, si aumenta il numero di dati e quindi la precisione della rappresentazione, ma si rende più complessa la lettura; al contrario con intervalli più ampi si riduce l’informazione e si semplifica la leggibilità.
Al limite possiamo avere, da un lato l’elencazione completa di tutti i dati (massima informazione), dall’altro un solo dato riassuntivo, quale può essere la media (massima leggibilità).

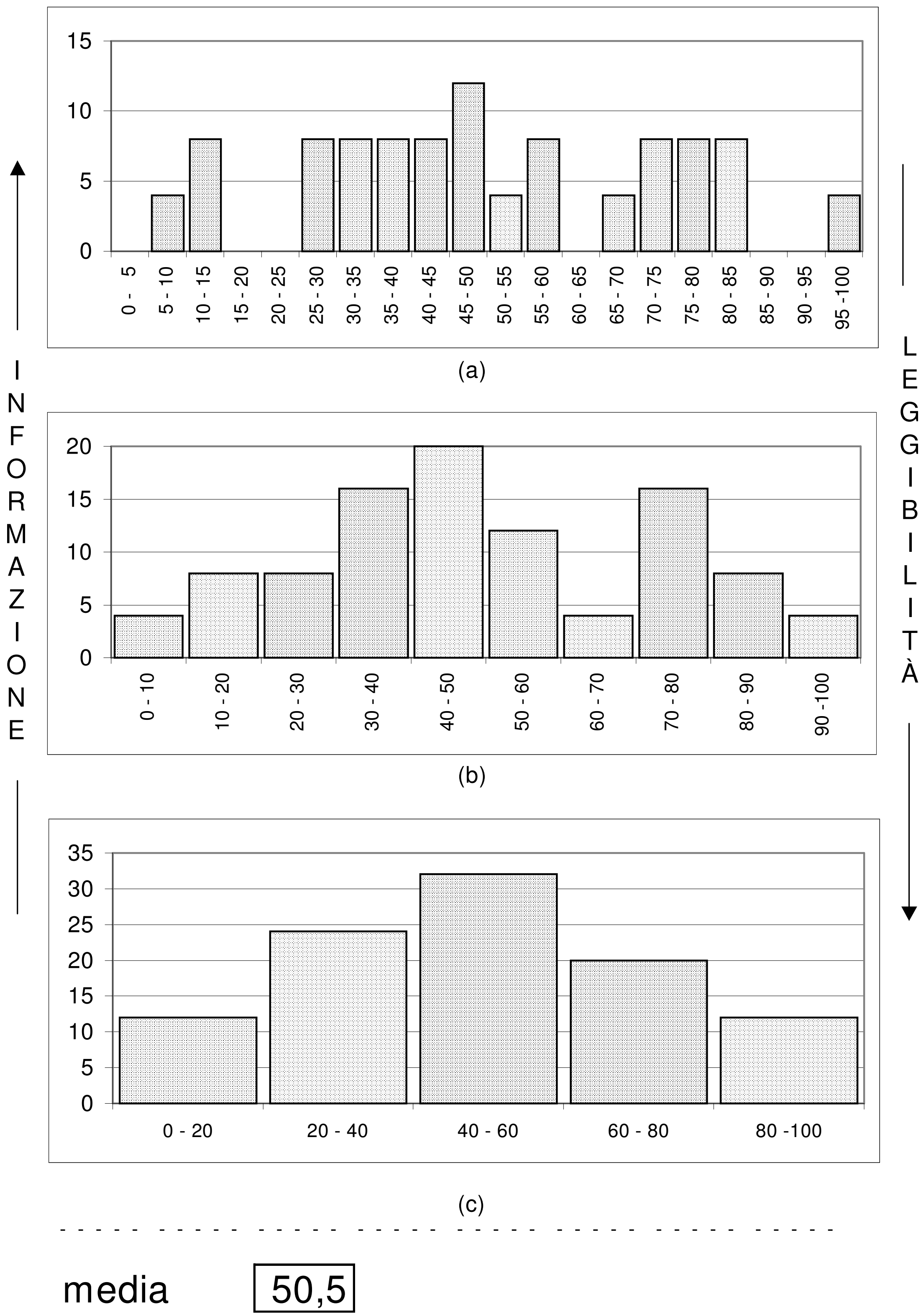
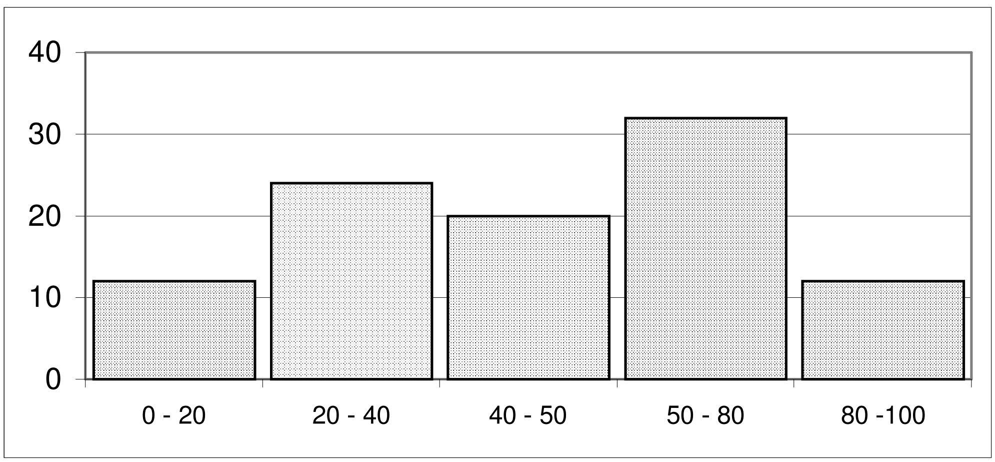

15
10
5
0
0 - 5
5 - 10
10 - 15
15 - 20
20 - 25
25 - 30
30 - 35
35 - 40
40 - 45
45 - 50
50 - 55
55 - 60
60 - 65
65 - 70
70 - 75
75 - 80
80 - 85
85 - 90
90 - 95
95 -100
(a)
20
15
10
5
0
0 - 10
10 - 20
20 - 30
30 - 40
40 - 50
50 - 60
60 - 70
70 - 80
80 - 90
90 -100
(b)
35
30
25
20
15
10
5
0
0 - 20
20 - 40
40 - 60
60 - 80
80 -100
(c)
INFORMAZIONE
LEGGIBILITÀ
media
50,5

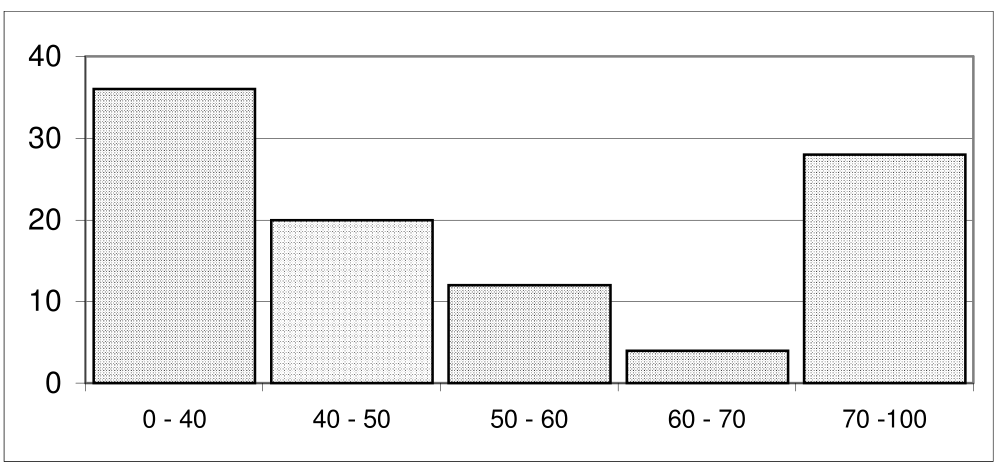
Figura 3.3

N.B. Per completezza occorre precisare come devono essere trattati i limiti degli intervalli: in pratica bisogna decidere a priori se il punteggio coincidente con un limite (ad esempio 10) debba essere conteggiato nell' intervallo inferiore (5 – 10) o superiore (10 – 15).

Le classi di frequenza di un istogramma devono avere tutte la stessa ampiezza. Quando ciò non si verifica l'interpretazione diventa più difficile e può essere fonte di equivoci.
I diagrammi di Figura 3.4.a e 3.4.b riprendono i dati precedenti in modo abbastanza simile alla 3.3.c, ma le ampiezze di alcune classi sono state modificate.
Come si vede, l'aspetto complessivo del grafico risulta fortemente alterato e può indurre facilmente a considerazioni erronee.



(a)



(b)

Figura 3.4

4 – Quartili

Un altro tipo di raggruppamento usato dall'INV ALSI, e alternativo alle classi di frequenza, è costituito dai **quartili**: sostanzialmente consiste nel suddividere l'insieme dei punteggi in 4 parti uguali, identificando i valori che delimitano ogni parte.

Vediamo come si procede.
Per prima cosa, ordinati come al solito i punteggi della classe X dal migliore al peggiore, occorre individuare il punteggio "centrale", che divide la distribuzione in due parti uguali: tale punteggio (in questo caso **48,3**) viene detto anche **mediana** della distribuzione.
Adesso procediamo allo stesso modo sulle due metà, inferiore e superiore, dividendole ancora in due parti. In questo caso però, il numero di elementi è pari e non esiste un valore centrale: per ovviare, creiamo questo punteggio facendo la media fra i due punteggi più vicini (Figura 4.1).

Così si ottiene una ripartizione per cui al di sotto del 1° quartile (32,7) si trova un quarto (25%) degli alunni; al di sotto del 2° quartile (mediana) la metà (50%) e al di sotto del 3° i 3/4 (75%). Per questo motivo il 1° quartile viene chiamato anche 25° percentile, il 2° 50° percentile, e così via.

La differenza fondamentale con le classi di frequenza consiste nel fatto che, in questo caso, si ripartiscono gli alunni (con i loro punteggi) in quattro fasce ugualmente numerose, senza tenere conto della "distanza" fra i punteggi, cosa che è invece fondamentale nelle classi di frequenza.

Nella Figura 4.1 le 4 fasce sono indicate con la nomenclatura utilizzata dall'INVALSI (V. par. successivo).

È possibile anche, in modo analogo, procedere ad una ripartizione in 5 parti (**quintili**), 10 parti (**decili**), 100 parti (**percentili**), ecc..

1	99,5	
2	84,8	fascia alta
3	81,3	
4	77,2	
5	75,7	
6	74,2	72,9 (3° quartile)
7	71,6	
8	68,3	
9	56,4	fascia medio-alta
10	55,1	
11	52,5	
12	48,7	
13	$\boxed{48,3}$	48,3 (2° quartile - mediana)
14	46,2	
15	44,9	
16	43,7	fascia medio-bassa
17	38,4	
18	37,3	
19	33,6	32,7 (1° quartile)
20	31,8	
21	28,6	
22	26,9	
23	18,1	fascia bassa
24	14,7	
25	5,8	
media	50,5	

Figura 4.1

5 – Risultati delle prove INVALSI

L'INVALSI presenta i risultati delle prove classificati in 5 fasce: bassa, medio-bassa, medio-alta, alta, top. Le prime 4 fasce sono individuate semplicemente dai quartili del Campione Nazionale (il campione di scuole appositamente selezionato per essere rappresentativo dell'intera popolazione scolastica italiana).

Prendendo come esempio i risultati di Matematica – 1ª media del PP2, per conteggiare il 25% degli alunni del Campione Nazionale, bisogna arrivare, partendo dal basso, fino al punteggio 36 (1° quartile);

- per il 50% bisogna arrivare a 50 (2° quartile);
- per il 75% fino a 65 (3° quartile).

Allora la fascia **bassa** è costituita dagli alunni che hanno punteggi inferiori (o uguali) a 36,
la fascia **medio-bassa** fra 36 e 50 (per la precisione > 36 e <=50),
la fascia **medio-alta** fra 50 e 65 (>50 e <=65),
la fascia **alta** oltre 65.

La fascia "**top**" è costituita dai punteggi che superano il 90° percentile (90% degli studenti) e perciò non è una fascia a sé stante, ma una parte della fascia alta.

Tutto questo per il Campione Nazionale: vediamo ora i risultati delle singole scuole e classi. Una volta determinati i limiti di punteggio delle fasce, che altro non sono se non i quartili del Campione Nazionale, si determina per la singola scuola (o classe) la percentuale di alunni che ha riportato un punteggio compreso nella fascia stessa.
Qualora una scuola avesse l'identica distribuzione del Campione Nazionale, avrebbe il 25% di alunni nelle prime 4 fasce e il 10% nella fascia "top". Le differenze da tale distribuzione evidenziano la "distorsione" che si verifica rispetto al Campione Nazionale.

La Figura 5.1.a illustra quale sarebbe il risultato della solita classe X, in base al Campione Nazionale INVALSI per Matematica – 1ª media del PP2. Per confronto vengono anche rappresentate, in colore bianco, le fasce relative al Campione Nazionale che, naturalmente, hanno tutte il 25% di frequenza.

Come già sottolineato, la somma delle prime 4 fasce porta al 100%, mentre, se si considerano tutte le 5 fasce, si supera il 100% perché la fascia "top" è un'aggiunta (rappresenta una parte della fascia alta).

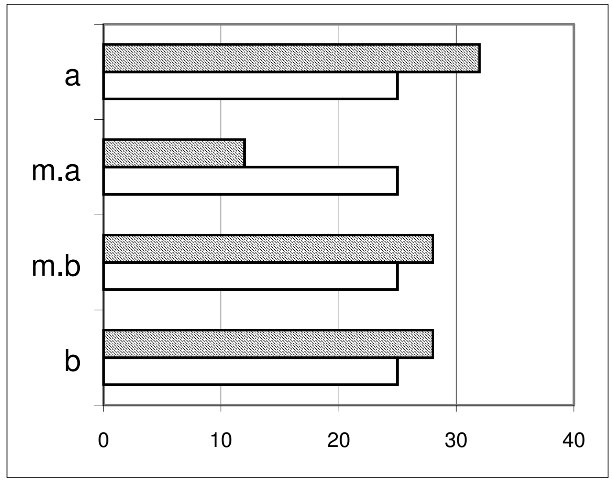
L'interpretazione dei risultati INVALSI impone una certa cautela.

La Figura 5.1.b mostra quale sarebbe stato il risultato della classe X se i limiti di punteggio del Campione Nazionale fossero stati leggermente diversi: qualche fascia molto affollata si restringe e viceversa.

Campione
nazionale

100
65
50
36
0

	n°	%
a	8	32
m.a	3	12
m.b	7	28
b	7	28
totale	25	100

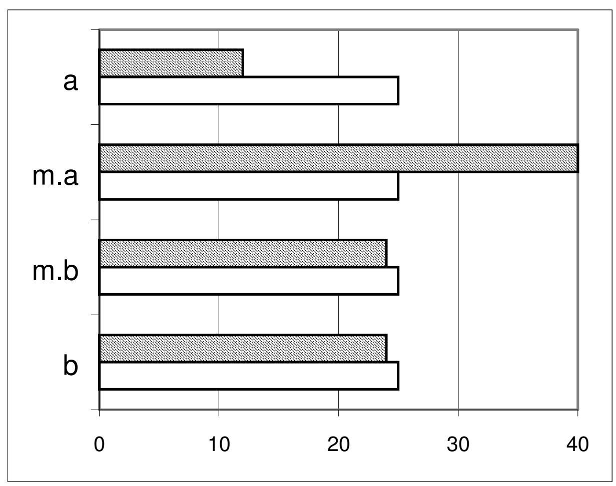


(a)

Campione
nazionale

100
80
48
32
0

	n°	%
a	3	12
m.a	10	40
m.b	6	24
b	6	24
totale	25	100



(b)

Figura 5.1

Risultati ancora più sorprendenti possono verificarsi simulando qualche cambiamento più consistente del Campione Nazionale (Figura 5.2): la classe sembra mostrare un rendimento molto scadente o molto positivo.

Sottolineiamo il fatto che si tratta sempre dello stesso risultato ottenuto dalla stessa classe: ciò che cambia è solo la fisionomia del Campione Nazionale.

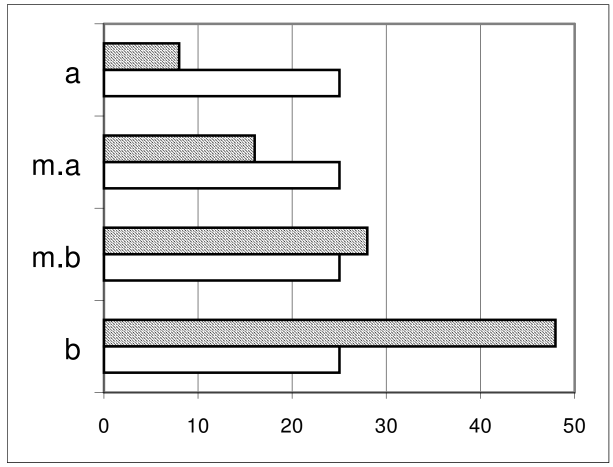
Tutto ciò è conseguenza di quanto detto prima a proposito degli istogrammi: se le ampiezze degli intervalli delle classi di frequenza vengono cambiate (cfr. sopra Figura 3.4), l'aspetto complessivo della distribuzione subisce forti alterazioni.

Questo non toglie nulla al valore comparativo dei risultati, qualora ci si limiti però a confronti fra dati omogenei (stessa tipologia di scuola e classe, e stesso anno). Non è invece per nulla immediato confrontare dati di prove diverse, perché riferiti, per forza di cose, a Campioni Nazionali diversi.

Campione nazionale

100
82
73
48
0

	n°	%
a	2	8
m.a	4	16
m.b	7	28
b	12	48
totale	25	100

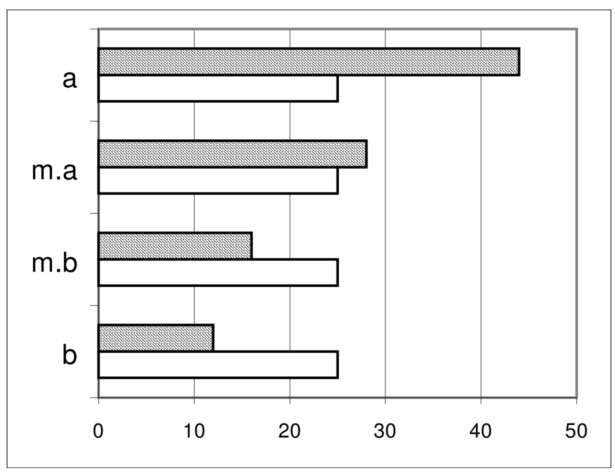


(a)

Campione nazionale

100
50
35
25
0

	n°	%
a	11	44
m.a	7	28
m.b	4	16
b	3	12
totale	25	100



(b)

Figura 5.2