3.2. L'analisi dei risultati degli studenti

Nel campo delle scienze didattiche e del comportamento, la valutazione di un certo fenomeno è più generale della *misurazione* definita per le grandezze fisiche; in effetti quando non si dispone di una unità di misura è improprio parlare di misurazione; mentre ha senso **misurare** il peso di un individuo con l'unità di misura *chilogrammo* o l'altezza con l'unità di misura *metro*, non ha ad esempio senso parlare di misurazione quando abbiniamo a ciascun individuo il suo indirizzo di studio (parleremo in questo caso piuttosto di *classificazione*).
In generale si tratta di determinare il livello (inteso come valore o categoria, che quindi può essere qualitativo o quantitativo) di una proprietà posseduta da una determinata unità statistica.

3.2.1. Cosa è una misura

Nell'accezione più ampia del termine, misurare significa stabilire una corrispondenza tra i modi in cui una caratteristica variabile può manifestarsi e i valori (di solito numerici) di una scala di riferimento. Tali valori, detti anche modalità, possono far riferimento a manifestazioni qualitativamente differenti (ad esempio il colore dei capelli), dando così luogo ad una scala qualitativa, oppure a manifestazioni differenti sul piano quantitativo (ad esempio l'altezza espressa in metri), dando luogo ad una scala quantitativa.
Dunque le misure possono essere di diverso tipo e la loro scelta è determinata dal tipo di sperimentazione. Le misure possono essere divise in diverse categorie. Ad un estremo vi sono le misure "vere", rappresentate per esempio dall'altezza, lunghezza, temperatura. All'altro estremo abbiamo le osservazioni "qualitative": sesso, stato civile, colore dei capelli. Tra questi due estremi ci sono tipi intermedi di misure, per esempio i giudizi (Insufficiente, Scarso, Sufficiente, Buono, Ottimo), con i quali possiamo affermare che una "misura" è più grande o più piccola di un'altra, senza poter definire in modo preciso l'ampiezza di tale differenza.

3.2.2. Le scale di misura

La misurazione consiste nell'assegnare valori numerici alle cose secondo certe regole. Le scale di misura hanno il compito di stabilire queste regole: decidono così il modo in cui i numeri, assegnati nel processo di misurazione, corrispondono alle proprietà degli oggetti che vengono misurate (decidono, cioè, il significato di tali numeri in relazione alle proprietà misurate).

Possiamo effettuare misure con diversi tipi di scale, che godono di proprietà formali differenti; di conseguenza, esse ammettono operazioni differenti.

La distinzione fra le diverse scale di misura non è puramente teorica, è invece fondamentale, poiché per ogni tipo di scala (e quindi per ogni tipo di variabile ad esso corrispondente) esistono, come vedremo, indici sintetici di posizione dei soggetti, di tendenza centrale e di variabilità, rappresentazioni grafiche e procedure per il controllo delle ipotesi che dipendono dalla scala di misura utilizzata.

Le scale utilizzate nella ricerca didattica sono quattro e corrispondono a crescenti livelli di complessità. Sono la **scala nominale**, la **scala ordinale**, la **scala ad intervalli** e la **scala di rapporti**. Esiste una gerarchia cumulativa tra le scale di misura, per cui le tecniche statistiche usate sono tanto più precise quanto più elevata la scala di misurazione adottata; solo quando non è possibile adottare un tipo di scala superiore si ripiegherà verso un livello di scala inferiore.

La scala razionale e quella intervallare sono definite scale superiori rispetto alla scala ordinale, e questa è superiore rispetto alla nominale.

Ogni livello di misurazione superiore include le caratteristiche di quelli inferiori.

La scala nominale ha il minimo di informazione; la scala di rapporti ha il massimo di informazione.

Ciò che differenzia le scale tra di loro è semplicemente la regola di assegnazione dei numeri alle cose. Le singole scale possiedono caratteristiche proprie che consentono loro di darci specifici tipi di informazione; è importante ricordare che ogni scala conserva tutte le caratteristiche delle scale di livello precedente ed aggiunge ad esse nuove proprietà. (gerarchia cumulativa)

<table>
<tr>
<td rowspan="4">gerarchia cumulativa, informazione crescente

SCALE
DI
MISURA</td>
<td rowspan="2">scale di misura quantitative</td>
<td>scala
di rapporti
o razionale</td>
<td>il rapporto tra le misure coincide con quello tra le grandezze e lo zero è lo stesso su ciascuna scala (es: peso, statura, numero di risposte esatte ad un test)</td>
</tr>
<tr>
<td>scala
ad intervalli
o intervallare</td>
<td>E' possibile stabilire la distanza fra le classi; il rapporto tra due intervalli è indipendente dall'unità di misura e dallo zero, che sono arbitrari
(es.: temperatura)</td>
</tr>
<tr>
<td rowspan="2">scale di misura qualitative</td>
<td>scala
ordinale
o per ranghi
(o categoriale ordinata)</td>
<td>I dati vengono ordinati, ma non è possibile stabilire la distanza tra le classi. Ci si limita a stabilire se una è superiore o inferiore ad un'altra (es: Giudizi: insufficiente, scarso, sufficiente, buono, ottimo)</td>
</tr>
<tr>
<td>scala
nominale
o classificatoria
(o categoriale non ordinata)</td>
<td>I dati vengono etichettati e ognuno può essere solo uguale o diverso da un altro (es.: esito di un esame: promosso-respinto; stato civile: celibe, nubile, coniugato, vedovo, divorziato)</td>
</tr>
</table>

3.2.2.1 Scala nominale o classificatoria (o categoriale non ordinata)

Quando possiamo semplicemente controllare se i soggetti osservati hanno o no una determinata proprietà, ad esempio se sono o no maschi, oppure rilevare le modalità di una loro proprietà ed assegnarli ad una data classe, ad esempio le modalità "italiano" o "francese", sulla proprietà "nazionalità", possiamo dire che abbiamo operato una classificazione (ossia una ripartizione in classi diverse) dei soggetti, ossia che abbiamo effettuato una catalogazione di tipo nominale. Le variabili generate definendo operativamente queste proprietà vengono dette variabili nominali o categoriali non ordinate e il livello di scala corrispondente è la scala nominale.

E' la più elementare scala di misurazione, con cui non si fa altro che identificare le categorie nelle quali membri, oggetti o eventi possono essere classificati. A eventi o oggetti dello stesso tipo è assegnato il medesimo valore numerico, mentre a eventi o oggetti di tipo diverso è assegnato un diverso valore numerico.

E' fondamentale notare che quando a tali categorie sono assegnati dei numeri, essi non hanno un significato numerico, ma rappresentano modi convenienti di etichettare le informazioni.
Per esempio, nel codificare un progetto sperimentale, si potrebbe assegnare le etichette in questo modo:

❶	ai maschi adulti
❷	alle femmine adulte
❸	ai bambini
❹	alle bambine

Tali etichette: ❶, ❷, ❸, ❹ non hanno significato quantitativo. Non possono essere sommate, sottratte, moltiplicate o divise, la loro funzione è semplicemente quella di identificare categorie differenti. Infatti il numero ❶ assegnato ai maschi adulti non indica che maschi adulti sono più grandi, più importanti o più belli delle femmine adulte, dei bambini e delle bambine: indica semplicemente che essi sono diversi dagli altri soggetti e che costituiscono una classe con caratteristiche particolari. L'attribuzione di numeri per identificare le varie categorie nominali, come avviene per individuare i giocatori nel gioco del calcio, è solamente un artificio che non può certamente autorizzare ad elaborare quei numeri come se fossero reali, ad esempio calcolandone la media.

Al posto dei numeri avremmo potuto utilizzare qualunque altro gruppo di simboli sufficientemente diversi tra di loro, ad esempio:

	ai maschi adulti
	alle femmine adulte
	ai bambini
	alle bambine

O ancora, in una serie di dati i maschi potrebbero essere indicati con 0, le femmine con 1; la condizione sociale maritale di un individuo potrebbe essere codificato con K se sposato, R se single, H se divorziato.
La scala nominale, attraverso l'assegnazione dei valori numerici, ci dà dunque soltanto informazioni sull'uguaglianza o diversità di due eventi o oggetti: se due oggetti hanno lo stesso numero allora sono uguali; se non hanno lo stesso numero, sono diversi. I dati nominali si possono *classificare* ma non *ordinare* o *misurare*.
Usando la scala nominale di misurazione, noi possiamo dunque assegnare individui ed oggetti a varie categorie e, naturalmente, contare i numeri all'interno di ogni categoria.
La scala nominale, in base al diverso numero e alle caratteristiche delle categorie alla quale può appartenere il nostro dato può essere distinto in: scala dicotomica o in scala politomica.

3.2.2.1.1 Scala dicotomica

La scala dicotomica è una scala nominale con due soli stati: alcuni esempi sono rappresentati dal sesso: maschio, femmina; dall'abitudine al fumo: fumatore, non fumatore; dal risultato di un test di laboratorio qualitativo: positivo, negativo; dall'effetto di un trattamento: miglioramento, non miglioramento; dall'esito di un esame: superato o non superato.

Le variabili misurabili a livello di scala dicotomica, pur appartenendo alla classe delle variabili nominali, possono essere trattate statisticamente con strumenti normalmente non applicabili alle variabili nominali, ma utilizzati per variabili collocate ad un livello superiore di scala.

3.2.2.1.2 Scala politomica

La scala politomica è una scala nominale con più di due stati: alcuni esempi sono rappresentati dallo stato civile: celibe, nubile, coniugato, vedovo, divorziato o dal gruppo sanguigno di appartenenza: "A", "B", "AB" e "0".

3.2.2.2 Scala ordinale o per ranghi (o categoriale ordinata)

Se possiamo ordinare i soggetti in base alla maggiore o minore presenza di una certa caratteristica (ad esempio l'aggressività oppure il profitto in storia) e stabilire in quale ordine si possono disporre in base all'intensità della caratteristica stessa, ossia dal più al meno aggressivo o dal soggetto con profitto in storia maggiore al soggetto con profitto in storia minore, le variabili generate definendo operativamente tale proprietà ordinabile vengono dette variabili ordinali o categoriali ordinate e il livello di scala corrispondente è la scala ordinale.

La scala ordinale di misurazione incorpora le funzioni di classificazione ed etichettatura della scala nominale, quindi ci dà informazioni sulla diversità degli eventi o oggetti rispetto alle varie manifestazioni (modalità) del carattere, ma stabilisce anche un ordine tra queste manifestazioni. I numeri ordinali sono usati solo per indicare l'ordine in una "classifica". La scala ordinale viene usata per inserire membri od oggetti in una sequenza ordinata in modo discendente secondo particolari caratteristiche misurate. I numeri ordinali assegnati a tale serie di oggetti specificano un ordine in relazione ai diversi gradi di possesso di un attributo, ma non specificano le differenze tra tali gradi. Per esempio, in un gruppo di bambini classificati e ordinati in modo discendente da un osservatore in base al loro grado di aggressività, non si può dedurre che la differenza del grado di aggressività fra i soggetti uno e due sia la stessa che si ottiene tra i soggetti nove e dieci, né che il soggetto uno possiede dieci volte la aggressività del soggetto dieci.

Vediamo un altro esempio:

Immaginiamo di avere una popolazione di 7 sassi e il carattere di essi che ci interessa è "grandezza", la scala ordinale attribuirà il numero 1 al sasso più piccolo, il 2 al sasso più piccolo dopo aver tolto il primo, e così via fino ad assegnare il numero 7 al sasso più grande. Osservando il numero assegnato ad un certo sasso, ad esempio 4, sapremo che quest'ultimo è più grande dei sassi cui sono stati attribuiti i numeri inferiori al 4 e più piccolo di quelli contrassegnati dai numeri 5, 6, 7. La scala ordinale non dice però quanto sia più piccolo o più grande degli altri sassi. Pur essendo la distanza numerica del 7 dal 6 identica a quella del 6 dal 5, il sasso numero 7 potrebbe essere il Monte Bianco, il numero 6 un sassolino da spiaggia e il numero 5 un sasso di poco più piccolo del numero 6: la scala ordinale tiene conto delle differenze dei gradi di un attributo (ad esempio, la "grandezza") e ordina tali gradi, ma non tiene conto della grandezza di queste differenze; in altre

parole, la scala ordinale afferma l'esistenza di una differenza, ma non ne specifica l'entità.
Va fatta una precisazione: nell'esempio precedente abbiamo assegnato il numero più piccolo all'oggetto più piccolo e quello più grande all'oggetto più grande; naturalmente avremmo potuto anche assegnare il numero più piccolo all'oggetto più grande e far funzionare l'attribuzione al contrario: ciò che conta è che le modalità dell'assegnazione siano funzionali alle proprietà della scala e vengano rese esplicite; inoltre invece di utilizzare numeri, avremmo potuto utilizzare le lettere dell'alfabeto o un qualunque altro insieme di segni ordinabili.
Il numero d'ordine che compete ad una certa modalità, cioè il posto che occupa in una certa graduatoria, viene detto *rango*.
Esempio classico di variabile misurabile a livello di scala ordinale o per ranghi è quello dei giudizi (es.: profitto insufficiente, sufficiente, buono, ottimo). I giudizi o punteggi nominali possono e sono spesso sostituiti da valori all'interno di scale convenzionali (es.: scala dei voti scolastici). Le variabili nominali ordinabili possono essere in tal modo valutate quantitativamente con valori discreti, convenzionali, rappresentati in genere da piccoli interi (es.: scarso: 0, mediocre: 1, buono: 2).
Sottintendono una scala ordinale anche misure che sono rappresentate con simboli, come:

☆	☆☆	☆☆☆	☆☆☆☆	☆☆☆☆☆

o anche:

☹☹	☹	😐	☺	☺☺

o ancora:

- -	-	=	+	+ +

Anche in questo ultimo caso vi è l'impossibilità di valutare quanto sia grande la distanza tra ”--" e “-“, e se sia inferiore o superiore alla distanza tra “+” e “++”.

3.2.2.2.1 Rango

Il rango è la posizione occupata all'interno di una graduatoria da un individuo o da un gruppo di individui in una lista ordinata quantitativamente.
E' il numero d'ordine che compete ad una certa modalità di una variabile ordinabile, (variabile misurabile a livello di scala ordinale) cioè il posto che occupa in una certa graduatoria.
Di solito si dà il rango 1 alla osservazione più piccola, il rango 2 a quella immediatamente più grande, ecc.; così per esempio se con un test a scelta multipla per la valutazione delle abilità matematiche per 6 studenti abbiamo ottenuto i seguenti punteggi: 37, 40, 21, 36, 32, 15, li possiamo ordinare nel modo seguente:

punteggio ottenuto	15	21	32	36	37	40
Rango	1	2	3	4	5	6

Se abbiamo 6 osservazioni, i ranghi vanno da 1 a 6; se abbiamo n osservazioni, i ranghi vanno da 1 a n. Se però nella nostra serie di dati abbiamo due o più osservazioni uguali, nasce il problema di quale rango assegnare ai numeri uguali. Consideriamo ad esempio la serie: 15, 18, 29, 29, 38, 40; ci sono sei osservazioni e così i ranghi devono andare da 1 a 6. Vi sono, però, due valori in terza e quarta posizione nella serie. In questo caso si attribuisce a ciascuna delle due osservazioni la media dei ranghi corrispondenti alle loro posizioni: quindi si assegnerà a ciascuno dei valori uguali il rango 3,5 (media fra 3 e 4). Si avrà quindi il seguente ordinamento:

punteggio ottenuto	15	18	29	29	38	40
Rango	1	2	3,5	3,5	5	6

In una serie di n osservazioni senza ranghi assegnati ex aequo la somma dei ranghi è 1 + 2 + ... + n = k. Questa regola si applica anche se vi sono osservazioni uguali. Sommando i ranghi assieme si deve ottenere lo stesso valore.
Per trasformare variabili misurabili da una scala quantitativa in variabili qualitative discontinue ordinabili è sufficente sostituire al valore dei dati il rango, ossia la posizione occupata nella disposizione in ordine crescente:

Esempio:

dati
1.3
7.2
3.8
10.4
2.1

disposti in ordine crescente	rango
1.3	1
2.1	2
3.8	3
7.2	4
10.4	5

La trasformazione in rango comporta una perdita di informazione (aspetto negativo) ma elimina anche eventuali condizionamenti della distribuzione che potrebbero rendere non applicabili alcune analisi statistiche basate sull'ipotesi che la distribuzione dei dati sia normale (aspetto positivo).

3.2.2.3 Scala ad intervalli

Si ha una scala ad intervalli quando i dati possono essere disposti in ordine di grandezza ed inoltre l'ampiezza tra le distanze di due dati è di dimensioni note; questa scala è più precisa della scala ordinale. Infatti la scala ad intervalli alle caratteristiche della scala ordinale aggiunge quella di misurare le distanze o differenze tra tutte le coppie di valori. La scala ad intervalli si fonda su una misura oggettiva e costante, pur se l'unità di misura e il punto di origine sono arbitrari (a differenza della scala di rapporti in cui vi è un punto di zero assoluto per la scala, ossia un punto in cui la proprietà in questione è oggettivamente assente). Tipici esempi di variabili misurate su scala ad intervalli sono la temperatura (misurata in gradi Celsius o Fahrenheit) ed il tempo (misurato secondo calendari differenti). I valori di temperatura infatti, oltre a poter essere facilmente ordinati secondo l'intensità del fenomeno, godono della proprietà che le differenze tra loro sono direttamente confrontabili e quantificabili; così pure le date in un calendario gregoriano, islamico, ebraico o cinese possono essere tra loro ordinate dalla più antica a quella più recente e le differenze temporali sono misurate con precisione oggettiva.
Altri esempi di variabili misurabili a livello di scale ad intervalli sono le altezze delle maree, e la misura della longitudine.
Nella scala ad intervalli: diversamente dalla scala ordinale (in cui "10-9" e "9-8" non danno lo stesso risultato), le differenze fra i valori numerici assegnati hanno un significato poiché espresse secondo la

stessa unità di misura; questo significa che gli intervalli possono essere addizionati o sottratti.
La scala ad intervalli permette certe procedure matematiche non possibili a livelli di misurazione nominali e ordinali; tuttavia solo per le differenze sono permesse tutte le operazioni: possono essere tra loro sommate, elevate a potenza oppure divise, determinando le quantità che stanno alla base della statistica parametrica.
Va evidenziato che lo zero della scala ad intervalli è arbitrario, non rappresenta cioè l'assenza della caratteristica misurata ma viene attribuito convenzionalmente; in altri termini non vi è un punto di zero assoluto (ossia un punto in cui la proprietà in questione è oggettivamente assente) ma solo un punto di zero arbitrario (ossia scelto soggettivamente da chi definisce la scala).
Per esempio, l'intervallo temporale tra l'inizio degli anni 2003 e 2002 è lo stesso di quello tra gli anni 1983 e 1984, cioè 365 giorni. Il punto zero, anno 1, è arbitrario; il tempo non incominciò in quel momento.
La scala ad intervalli ha alcuni limiti, non gode infatti di alcune proprietà di cui gode, ad esempio la scala di rapporti. Ad esempio, una temperatura di 80 gradi non è il doppio di una di 40 gradi, se la riferiamo alla temperatura corporea; ancora, se una persona ponesse la mano sinistra in una bacinella con acqua a 80 gradi e la mano destra in una con acqua a 10 gradi, non potrebbe affermare che la prima scotta 8 volte più della seconda, ma potrebbe dire solo che la prima è bollente e la seconda è fredda.
Da una scala d'intervalli è possibile scendere ad una scala di ranghi (utilizzando solo l'informazione dell'ordine dei valori) oppure ad una scala nominale (suddividendo in misure alte e basse, sopra o sotto un valore prefissato). Pertanto, la scala d'intervalli gode anche delle proprietà definite per le due scale precedenti.

3.2.2.4 Scala di rapporti o razionale

E' un livello di misurazione che non soltanto specifica la graduatoria, come nella misurazione ordinale, e la distanza tra le posizioni, come nella misurazione ad intervalli, ma che fissa un punto di zero assoluto (non arbitrario) per la variabile in questione.
È il più alto livello di misurazione, ed include i livelli d'intervallo, ordinali e nominali.
Se è possibile definire un punto di zero assoluto per la scala, ossia un punto in cui la proprietà in questione è oggettivamente assente, come nel caso del numero di risposte corrette ad un test di profitto, le

variabili generate definendo operativamente tale proprietà vengono dette variabili quantitative con scala di rapporti e il livello di scala corrispondente è la scala di rapporti.
La scala di rapporti ha tutte le caratteristiche di una scala ad intervalli, infatti l'unità di misura è ancora arbitraria, ma in più esiste un punto zero come vera origine. In essa il rapporto tra due punti qualsiasi è indipendente dall'unità di misura. Un esempio sono le scale di peso e lunghezza dove il rapporto tra due pesi o due lunghezze è lo stesso, sia che le misure vengano espresse in chili o in libbre o che siano espresse in pollici o in centimetri.
Sono tipiche scale di rapporti l'altezza, la distanza, l'età, il peso, il reddito, più in generale tutte quelle misure in cui 0 (zero) significa quantità nulla. Non solo le differenze, ma gli stessi valori possono essere moltiplicati o divisi per quantità costanti, senza che l'informazione di maggiore importanza, il rapporto tra essi, ne risulti alterata. Per il peso, ad esempio, si può affermare che 800 grammi è 200 grammi più pesante di 600 grammi e due volte più pesante di 400.
Alle variabili misurate con una scala di rapporti, il tipo di misurazione più sofisticato e completo, può essere applicato qualsiasi test statistico.

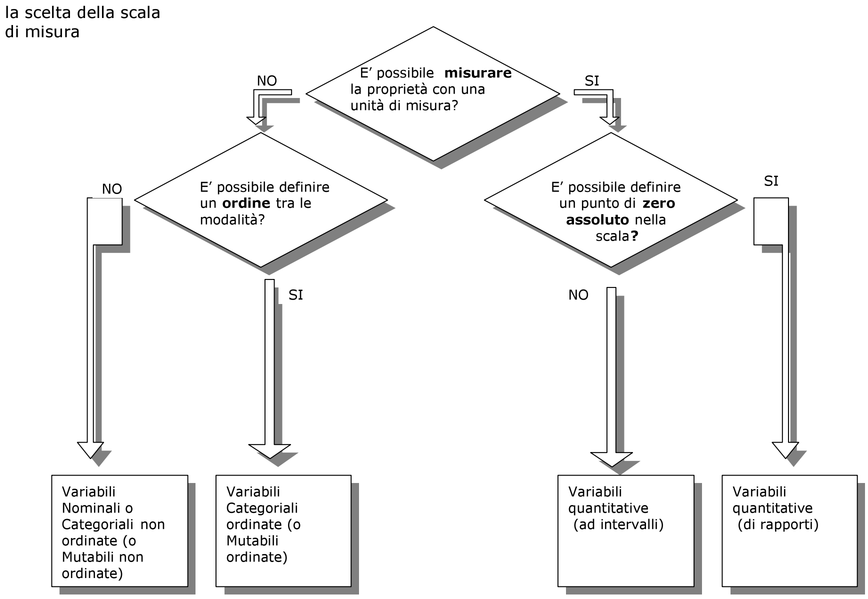
Con una scala di rapporti è possibile scendere nella scala di misurazione alla scala di intervalli, ordinale e nominale. Ovviamente, si ha una perdita ancor più rilevante della quantità d'informazione, rispetto alle scale precedenti; di conseguenza, rappresenta un'operazione che deve essere evitata, quando non imposta da altre condizioni dell'analisi statistica o dalle caratteristiche della distribuzione dei dati.
Le variabili misurate a livello di scala di intervalli ammettono operazioni di quantificazione (a differenza di quelle nominali ed ordinali) che ammettevano solo classificazione e ordinamento).

3.2.2.5.Scelta della corretta scala di misura

Per effettuare la scelta della corretta scala di misura tra:

- Scala nominale o classificatoria (o categoriale non ordinata)
- Scala ordinale o per ranghi (o categoriale ordinata)
- Scala a intervalli o intervallare
- Scala di rapporti o razionale

ci si può riferire allo schema qui riportato basato sulle caratteristiche tipiche delle scale (adattato da Trinchero R. , Pedagogia Sperimentale, Corso Online)



Solitamente è agevole individuare il livello di misura delle variabili, ma in alcuni casi va distinto un livello teorico di misurazione e un uso pratico.

Consideriamo per esempio i voti assegnati alla fine di un esame universitario: sono espressi in trentesimi ed espressi in una scala di rapporti (vanno da 0 a 30, e lo zero dovrebbe indicare preparazione nulla materia). Se però analizziamo dettagliatamente questa variabile, ci accorgiamo però che è, teoricamente, una scala misurata a livello ordinale. Infatti non possiamo affermare che la "quantità di conoscenza" che serve per passare da una votazione di 17 ad una di 18 è la stessa che serve per passare da 18 a 19 e da 29 a 30? Probabilmente no. Non abbiamo dunque certezza che l'intervallo esistente fra i 30 valori della votazione facciano riferimento ad una unità di misura comune. Nonostante ciò, questa scala ordinale viene comunemente utilizzata come se fosse una scala di rapporti continua di cui la votazione intera rappresenta solo un arrotondamento, vale a dire che una votazione di 27 corrisponderebbe ad un valore effettivo che oscilla fra 26,5 e 27,5.

3.2.2.6 Conversione fra scale di misura

La scala razionale e quella intervallare sono definite scale superiori rispetto alla scala ordinale, e questa è superiore rispetto alla nominale. Le osservazioni ottenute con una scala "superiore" possono essere trasferite in quelle di una scala "inferiore"; le misure intervallari possono essere trasformate in ordinali disponendo le osservazioni in ordine crescente o decrescente di grandezza, e le misure ordinali possono essere trasformate in nominali ad esempio scegliendo un punto arbitrario sulla scala e collocando le osservazioni in due classi, comprendenti rispettivamente quelle al di sopra e al di sotto del punto arbitrario. Non è possibile ovviamente eseguire il processo opposto, cioè esprimere una variabile con una scala di ordine superiore a quella che le è propria.

Ad esempio, consideriamo l'altezza, che è una variabile misurabile da una scala quantitativa. Se ho 10 individui, le loro stature possono essere misurate a livello di scala di rapporti; posso facilmente ridurmi ad una scala ordinale, considerando i 10 individui dal più basso al più alto. Ho però perso delle informazioni, poiché so che il 6° è più alto del 5°, ma non so di quanto. Posso ancora passare ad un livello di scala più basso, nominale, utilizzando le categorie: 1= soggetti più alti di 160 cm; 2= soggetti alti fino a 160 cm. Ho perso ancora informazioni.

Vediamo un altro esempio: se volessi trasformare variabili misurabili con una scala quantitativa (ad esempio punteggio ottenuto in un test a scelta multipla da un gruppo di alunni) in variabili qualitative discontinue ordinabili basta sostituire al valore dei dati il rango, cioè la posizione occupata nella disposizione in ordine crescente:

dati
23
16
28
33
32
24
20
35
40
39
36

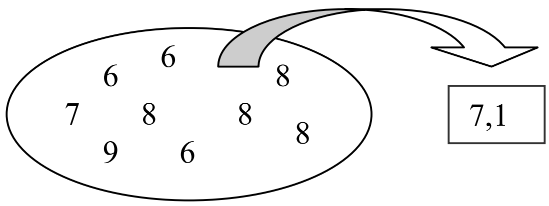
disposti in ordine crescente	rango
16	1
20	2
23	3
24	4
28	5
32	6
33	7
35	8
36	9
39	10
40	11

Si può notare che la trasformazione in rango comporta una perdita di informazione (aspetto negativo) ma elimina anche eventuali condizionamenti della distribuzione che potrebbero rendere non applicabili alcune analisi statistiche basate sull'ipotesi che la distribuzione dei dati sia normale (aspetto positivo). Ad esempio i cosiddetti metodi non-parametrici di verifica delle ipotesi, sono svincolati dal tipo di distribuzione ed applicabili ai ranghi.

Va evidenziato ai fini delle analisi statistiche, che ciò che conta non è la scala di misura che sarebbe possibile applicare a una variabile ma quella che è stata effettivamente applicata. Se dunque trovassimo conveniente per la nostra ricerca esprimere l'altezza in modo ordinale (come nell'esempio visto), dovremmo tenere ben presente che le operazioni possibili su tale variabile sono quelle tipiche di questa scala.

3.2.3. Le misure di tendenza centrale

E' spesso necessario "condensare" le informazioni derivanti da molti numeri in un unico numero significativo, che sintetizzi una certa serie di dati. Ad esempio, è necessario conoscere la media dei voti finali di ciascuno studente per assegnare il credito scolastico o una borsa di studio. Se lo studente ha ottenuto i seguenti voti in pagella: 6, 7, 6, 8, 8, 9, 8, 7, 8, basta sommare i numeri e dividere per il numero delle materie per ottenere il risultato (la media) che è un numero



sintetico.

Le *misure di tendenza centrale* servono per individuare il valore intorno al quale i dati sono raggruppati; la tendenza centrale è la misura più appropriata per sintetizzare l'insieme delle osservazioni, se una distribuzione di dati dovesse essere descritta con un solo valore; è la prima indicazione della dimensione del fenomeno.

Le misure di tendenza centrale più utilizzate sono tre: **la media, la moda e la mediana.**

La scelta della misura di tendenza centrale di una serie di dati dipende dalle caratteristiche della distribuzione e dal tipo di scala.
Tutte e tre i numeri sintetici, media, moda, e mediana misurano una *tendenza centrale*, ossia una caratteristica dei dati a disporsi a grappolo attorno a qualche valore.

Ad esempio, se consideriamo le risposte corrette date da un gruppo classe ad un test a scelta multipla sulla geometria analitica, potrei descrivere in tre modi diversi i punteggi ottenuti: potrei dire:

- in media hanno dato 26 risposto corrette; oppure
- per metà la classe è sopra 28 risposte corrette; oppure, ancora
- il numero di risposte corrette più frequente è 24.

In statistica si dà un nome specifico ai numeri sintetici appena espressi: la media, la mediana e la moda.

Media: totale di tutti i valori, diviso per il numero dei casi. La si chiama usualmente anche valore medio.
E' possibile utilizzarla con dati misurati con scala ad intervalli o di rapporti (ossia solo per variabili quantitative)

Mediana: valore che *occupa la posizione centrale in un insieme ordinato di dati*.
In sostanza è quel valore tale che il 50% dei casi la uguagliano o stanno al di sotto di esso, e il 50% restante la uguagliano o stanno al di sopra di esso.
E' una misura di tendenza centrale poco influenzata dalla presenza di dati anomali ed è possibile utilizzarla con dati misurati con scala ad intervalli, di rapporti o ordinale.
La sua utilizzazione è indispensabile nel caso di scale ordinali o di ranghi, non potendosi calcolare la media.

Moda: valore più frequente all'interno di un insieme di dati.
Il significato della moda è esattamente quello che noi attribuiamo generalmente alla parola "moda". Un abito di moda, una scarpa di moda... sono tali perché si presume che siano usati e preferiti dalla maggior parte delle persone, ovvero che siano scelte con maggior frequenza.

Se all'interno di un gruppo di dati vi è una sola moda, la distribuzione è chiamata "unimodale"; se ve ne sono due, sarà chiamata "bimodale". In teoria potrebbe esistere anche una distribuzione "trimodale" o "multimodale", ma quando ci sono troppe mode il potere descrittivo statistico della moda perde di informazione.
E' possibile utilizzarla con dati misurati con scala ad intervalli, di rapporti, ordinale o nominale.
E' l'unico indicatore di tendenza centrale utilizzabile con dati nominali.
La tabella che segue mostra gli indici di tendenza centrale utilizzabili con dati misurabili con le diverse scale.

Indici di tendenza centrale utilizzabili con dati misurabili con le diverse scale				
Livello di Scala	**Nominale**	**Ordinale**	**Ad intervalli**	**Di rapporti**
Esempi	Giudizi: Approvato-non approvato	Giudizi: Insufficiente, Scarso, Sufficiente, Buono, Ottimo	Voto di diploma	numero di risposte corrette ad un test di profitto
Indici di tendenza centrale utilizzabili	Moda	Moda e Mediana	Moda, Mediana, Media	Moda, Mediana, Media

Esempio
Immaginiamo di avere i seguenti dati, che indicano le risposte corrette ad una prova strutturata di inglese in una classe (sezione B):

37	31	26	22	25	25	28	30	32	40	18	39	19	34	29	36	27	28	25

E’ possibile calcolare sia moda che mediana che media.

Calcoliamo la **media**:
Sommiamo tutti i valori: totale = 551
Contiamo il numero di studenti =19
Dividiamo il totale per il numero di studenti
Media=29

Calcoliamo la **mediana**
Per calcolare la mediana di un gruppo di dati, occorre disporre i valori in una fila ordinata in modo crescente oppure decrescente;

18	19	22	25	25	25	26	27	28	28	29	30	31	32	34	36	37	39	40

Se il numero di dati è dispari come in questo caso, la mediana corrisponde al valore numerico del dato centrale:

18	19	22	25	25	25	26	27	28	**28**	29	30	31	32	34	36	37	39	40

Mediana = 28

Se il numero di dati fosse stato pari, la mediana si sarebbe trovata calcolando la media dei due valori centrali.

Calcoliamo la **moda**
Per calcolare la moda di un gruppo di dati, occorre contare il numero di volte in cui ogni valore appare nella serie di dati e scegliere quello più frequente.

18	19	22	**25**	**25**	**25**	26	27	28	28	29	30	31	32	34	36	37	39	40

Moda = 25.

3.2.4. Le misure della dispersione

Nell'esempio precedente, avevamo a disposizione i dati relativi alle risposte corrette ad un test di inglese in una classe di 19 alunni (sezione B):

18	19	22	25	25	25	26	27	28	28	29	30	31	32	34	36	37	39	40

ed abbiamo calcolato la media delle risposte corrette, ottenendo il valore 28.
Immaginiamo che in un'altra classe di 19 alunni (sezione A) i punteggi siano stati:

29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29

Ed in un'altra classe ancora (sezione C):

19	19	19	19	19	19	19	19	19	20	40	40	40	40	40	40	40	40	40

Anche in questi altri casi la media è 29, ma la situazione è profondamente diversa dalla prima.
Si dice che nel terzo caso la dispersione è massima, nel secondo caso la dispersione è nulla dato che tutti i numeri coincidono con la media della classe e la situazione è completamente omogenea. Il primo è un caso intermedio.

Dunque si comprende come i valori delle misure di tendenza centrale danno informazioni solo sull'andamento complessivo della prova; se considerate da sole, infatti, possono condurre anche a fraintendimenti circa la distribuzione dei punteggi, ovvero delle prestazioni rilevate. Uno stesso valore di media può scaturire da situazioni ben diverse!

Per ovviare a tali inconvenienti vengono utilizzati gli indici di dispersione che rappresentano la variabilità o la omogeneità dei valori numerici rilevati.

Un primo indice di dispersione è rappresentato dal *range,* o *campo di variazione* (o anche *gamma*).

Il range si ottiene ordinando i dati e calcolando la differenza tra il risultato più alto ed il più basso.

Ad esempio se ho i dati 65, 73, 89, 56, 73, 52, 47, otterrò :
range= (valore massimo – valore minimo) = 89-47 = 42.

Vi è da dire che dal momento che solo i valori più grandi e più piccoli vengono considerati nel calcolo del range, ed i dati restanti sono ignorati, viene trascurata una grande quantità di informazioni; inoltre il range di è enormemente influenzato dalla presenza di un solo valore insolito molto grande o molto piccolo in una serie di dati.

Per misurare la dispersione dei valori attorno alla media l'indice più utilizzato è la deviazione standard, indicata con la lettera greca sigma, σ∍00 per i dati di una popolazione, e con s per un campione (un'altra maniera di chiamare la deviazione standard è scarto quadratico medio, s.q.m.).

Una deviazione standard più piccola indica che i valori sono in genere *più prossimi* alla media; una deviazione standard più ampia indica che i valori sono *più lontani* dalla media, più sparsi, più variabili. La deviazione standard è uguale a zero quando tutti i valori osservati sono *uguali* alla media; in questo caso non vi è alcuna dispersione intorno alla media.

La deviazione standard, oltre a dare indicazioni sulla reale distribuzioni di una serie di dati, come vedremo, rappresenta un parametro su cui si basano molti dei criteri di standardizzazione dei punteggi grezzi di un test.

Ma come si calcola la deviazione standard?

1) Si calcola per ciascun dato lo *scarto* rispetto alla media (lo *scarto* rispetto alla media è la differenza tra il singolo dato e la media)
2) Si eleva ciascuno *scarto* al quadrato.
3) Si sommano tutti gli scarti elevati al quadrato e si divide tale somma per il numero di casi
4) Si calcola la radice quadrata del valore così ottenuto.

Esempio

Riprendiamo i punteggi ottenuti al test di inglese nelle tre sezioni e organizziamo i dati per calcolare la deviazione standard nei tre casi:
Per ciascuna sezione nella prima colonna inseriamo il numero di risposte corrette di ciascun alunno; nella seconda inseriamo gli scarti dalla media, quindi gli scarti al quadrato. Sommiamo gli scarti al quadrato, dividiamo per il numero di alunni e facciamone la radice

quadrata, ed otteniamo la deviazione standard per le tre sezioni, rispettivamente pari a 0, 6,10 e 10,44.

Sez. A	scarti (x-media)	scarti al quadrato	**Sez.B**	scarti (x-media)	scarti al quadrato	**Sez. C**	scarti (x-media)	scarti al quadrato
29	0	0	18	-11	121	19	-10	100
29	0	0	19	-10	100	19	-10	100
29	0	0	22	-7	49	19	-10	100
29	0	0	25	-4	16	19	-10	100
29	0	0	25	-4	16	19	-10	100
29	0	0	25	-4	16	19	-10	100
29	0	0	26	-3	9	19	-10	100
29	0	0	27	-2	4	19	-10	100
29	0	0	28	-1	1	19	-10	100
29	0	0	28	-1	1	20	-9	81
29	0	0	29	0	0	40	11	121
29	0	0	30	1	1	40	11	121
29	0	0	31	2	4	40	11	121
29	0	0	32	3	9	40	11	121
29	0	0	34	5	25	40	11	121
29	0	0	36	7	49	40	11	121
29	0	0	37	8	64	40	11	121
29	0	0	39	10	100	40	11	121
29	0	0	40	11	121	40	11	121
	somma degli scarti al quadrato	**0**		somma degli scarti al quadrato	**706**		somma degli scarti al quadrato	**2070**
Dev. standard= $\sqrt{\frac{0}{19}} = 0$			Dev. standard= $\sqrt{\frac{706}{19}} = 6{,}10$			Dev.standard= $\sqrt{\frac{2070}{19}} = 10{,}44$		

Delle tre sezioni la prima (sez.A) è quella totalmente omogenea, la seconda (sez.B) meno dispersa e la terza (sez.C) la più dispersa.

L'operazione di calcolo della deviazione standard è assolutamente agevole e immediata con un foglio elettronico come Excel: per calcolare la deviazione standard, basta selezionare i dati ed utilizzare la funzione DEV.ST.POP.

Un altro indice di variabilità in grado di confrontare la dispersione di dispersione di distribuzioni con medie molto diverse come per esempio delle serie di punteggi ottenute con prove diverse e con studenti diversi è il coefficiente di variazione (CV), che si ottiene dividendo la deviazione standard per la media.

Il coefficiente di variazione è un numero puro, che annullando l'effetto della media è quindi utile per comparare la dispersione di distribuzioni con medie molto diverse. Secondo molti statistici, appunto perché un rapporto, avrebbe significato solamente se calcolato per variabili misurate con una scala di rapporti.
Va anche detto che la deviazione standard è utilizzabile con dati misurati a scala ad intervalli o di rapporti, mentre per le altre scale esistono indici diversi, elencati elencano nella tabella che segue.

Indici di dispersione utilizzabili con dati misurabili con le diverse scale				
Livello di scala	**Nominale**	**Ordinale**	**Ad intervalli**	**Di rapporti**
Esempi	Giudizi: Approvato-non approvato	Giudizi: Insufficiente, Scarso, Sufficiente, Buono, Ottimo	Voto di diploma	numero di risposte corrette ad un test di profitto
Indici di dispersione utilizzabili	Squilibrio, Entropia	Squilibrio, Entropia, Gamma (campo di variazione), Gamma interdecilica, Gamma interquartilica, Indice di Leti	Come scala precedente + Devianza, Varianza, deviazione standard (scarto quadratico medio)	Come scala precedente + Coefficiente di variazione

3.2.5. Le implicazioni valutative e didattiche delle misure di tendenza centrale e di dispersione

Dopo aver calcolato per una serie di punteggi ottenuti in una prova gli indicatori di tendenza centrale (media, moda e mediana) e di dispersione (deviazione standard e coefficiente di variazione) , è possibile fare alcune considerazioni di tipo valutativo-didattico sull'andamento della prova.

Il primo confronto possibile è tra il *massimo* punteggio teorico della prova e la *media* dei punteggi del gruppo; se infatti vi è grande differenza tra questi valori, possiamo dedurre che gli obiettivi didattici che si intendeva verificare con quella prova non sono stati raggiunti, o in alternativa, che la prova utilizzata non era valida per verificarli.

Se si va a confrontare il valore della media dei punteggi ottenuti con la mediana, potremo sapere se la maggior parte degli studenti hanno ottenuto punteggi superiori, inferiori o pari a quello medio Dunque se:

- il valore della media aritmetica è superiore a quello della mediana: vuol dire che la maggior parte degli studenti (almeno la metà più uno) ha ottenuto un punteggio *inferiore* a quello medio e quindi solo un residuo gruppo di studenti ha raggiunto gli obiettivi programmati. Si potrebbe in questo caso pensare ad un recupero curriculare, durante il normale orario di lezione.
- il valore della media aritmetica è inferiore a quello della mediana: vuol dire che almeno la metà più uno degli studenti ha ottenuto punteggi maggiori del valore medio e quindi solo un gruppo minoritario di studenti ha ottenuto punteggi molto bassi Si potrebbe in questo caso pensare ad un intervento di recupero mirato anche fuori dall'orario scolastico;
- il valore della media e quello della mediana sono uguali:
  vuol dire che la metà degli studenti ha ottenuto punteggi maggiori del valore medio, la metà minori; per prendere decisioni didattiche motivate occorre verificare la dispersione dei punteggi.

Un'altra analisi possibile è quella del valore della *moda*; ad esempio la presenza di due mode agli estremi dei punteggi ci indicherebbe nella classe la presenza di due tronconi, uno con livelli bassi di apprendimento, l'altro a livelli alti di prestazione, e ci dà evidentemente indicazioni chiare su ciò che è avvenuto in termini di apprendimento.

La *deviazione standard*, come visto, dà indicazioni sulla maggiore o minore variabilità nell'ambito di uno stesso gruppo: se piccola essa indica che i valori sono in genere *più prossimi* alla media; una deviazione standard più ampia indica che i valori sono *più lontani* dalla media, più sparsi, più variabili; se la deviazione standard è uguale a zero tutti i valori osservati sono *uguali*; in questo caso non vi è alcuna dispersione intorno alla media.

Vediamo che indicazioni didattiche possono scaturire dalla lettura del *coefficiente di variazione* calcolato su una prova oggettiva somministrata in una classe: quando il valore del coefficiente di variazione è inferiore a 10 allora le misure ottenute sono sostanzialmente omogenee, e quindi possiamo dedurre che anche la classe di studenti è omogenea; quando il valore supera il valore di 20 allora ciò indica una notevole eterogeneità nelle misure e di conseguenza anche nella classe. Va ricordato che a parità di valore del coefficiente di variazione possono essersi verificate due situazioni differenti:

- media dei punteggi alta, e con forte dispersione, ossia alta deviazione standard.
- media dei punteggi bassa e con dispersione ridotta ossia bassa deviazione standard.

Nel primo caso la classe indica una forte eterogeneità rispetto al percorso formativo verificato con la prova e quindi diviene indispensabile attuare percorsi didattici, che saranno ben diversi nella tipologia di quelli necessari in una classe con risultati mediamente bassi ma notevolmente omogenei.

3.2.6.Dai punteggi grezzi al confronto del singolo con il gruppo

Il difetto del punteggio grezzo, quello che otteniamo direttamente dalla correzione di una prova, è che non può essere confrontato con nessun altro punteggio; quando si vogliono confrontare dati diversi di uno stesso soggetto, o di due o più soggetti diversi, non è possibile farlo se i dati disponibili appartengono a distribuzioni differenti.

Come fare dunque a paragonare la prestazione di un soggetto a due prove differenti se il punteggio massimo del primo test è 30 e quello del secondo è 70? Oppure, come confrontare la prestazione ad uno stesso test di conoscenza lessicale inglese di due ragazzi, uno di terza superiore e l'altro di quinta, se la media dei ragazzi più grandi è in genere più alta a quel test rispetto a quella dei più piccoli?

Per risolvere questo tipo di problemi si effettuano sui punteggi alcune trasformazioni che rendano paragonabili tra loro distribuzioni diverse. I principali metodi utilizzati per confrontare in modo significativo i punteggi sono gli **indici di posizione** e **i punti standardizzati.**
Tali metodi sono utilizzabili con dati misurabili a livello almeno di scala ordinale i primi, e almeno a livello di scala ad intervalli, i secondi, come indicato in tabella.

Metodo di confronto utilizzabili con dati misurabili con le diverse scale				
Livello di scala	**Nominale**	**Ordinale**	**Ad intervalli**	**Di rapporti**
Esempi	Giudizi: Approvato-non approvato	Giudizi: Insufficiente, Scarso, Sufficiente, Buono, Ottimo	Voto di diploma	numero di risposte corrette ad un test di profitto
Metodo di confronto utilizzabile	Nessuno	Indici di posizione: Centili, Decili, Quartili	Indici di posizione Punti standardizzati, punti T, L, C di Guilford	Indici di posizione Punti standardizzati, punti T, L, C di Guilford

3.2.6.1 Gli indici di posizione

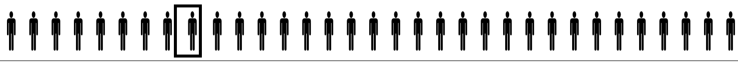
Gli indici posizionali servono a un'idea della posizione dei singoli soggetti all'interno della distribuzione di tutti i soggetti.
Già calcolando la mediana avevamo cercato il valore che si dispone al centro del gruppo. Seguendo lo stesso principio, è possibile trovare risultati che siano il valore più alto del 2°, del 3°, del 4° gruppo…. dopo aver diviso il gruppo totale in 4, 10 o 100 sottogruppi.
Gli indici posizionali più comuni sono:

- I centili, ossia i punti che dividono la distribuzione in cento parti uguali.
- I decili, ossia i punti che dividono la distribuzione in dieci parti uguali.
- I quartili, ossia i punti che dividono la distribuzione in quattro parti uguali.

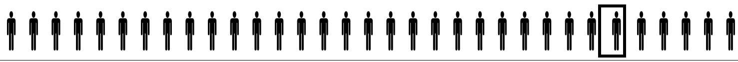
- La mediana, già vista, che divide la distribuzione in due parti uguali, e coincide con il 50° centile, il 5° decile, il 2° quartile.

Tali indici presuppongono che sia possibile misurare i dati almeno su scala ordinale.

Ad esempio il primo quartile è quel punteggio che lascia alla sua sinistra il 25% dei casi e alla sua destra il 75% dei casi ordinati in modo crescente.



L'ottantacinquesimo centile è quel punteggio che lascia alla sua sinistra l'85% dei casi e alla sua destra il 15% dei casi ordinati in modo crescente.



La procedura per trasformare di un punteggio in centili è simile a quella per il calcolo della mediana, (infatti il 50° centile corrisponde alla mediana).

Dopo aver ordinato i dati in modo crescente si verifica quale posizione occupa un determinato punteggio all'interno della serie, quindi si trasforma tale posizione in percentuale rispetto al totale dei punteggi tramite una proporzione.

Otteniamo che il punteggio viene rapportato ad una distribuzione di 100 punteggi, indipendentemente da quale fosse in origine la numerosità della classe o degli studenti.

Esempio

Se un punteggio occupa il 36° posto in una distribuzione di 45 punteggi, quale posizione occuperà in una distribuzione di 100 punteggi?

Dalla proporzione 36:45 = X:100 si ottiene $x=(100 \cdot 36)/45=80$.
Dunque il punteggio in questione corrisponde all'80° centile.

Se lo stesso studente ottenesse un punteggio corrispondente al 60° centile in un altro test, potremmo senz'altro affermare che la sua prestazione è stata migliore al primo test in quanto ha superato una maggiore percentuale di punteggi.

3.2.6.2 La standardizzazione di una variabile

Proviamo a chiederci:

- E' (relativamente) più alto un 28 all'esame di pedagogia sperimentale o un 26 all'esame di informatica?
- E' (relativamente) migliore una votazione di 102 per una laurea in medicina o un 105 per una laurea in scienze della formazione?
- E' (relativamente) più alto un 7 in matematica con il prof. Alfa o un 8 in matematica con il prof. Beta?
- E' (relativamente) più pesante un anguria di 28 Kg o una fragola di un etto?

In altri termini ci chiediamo: relativamente alla propria distribuzione, chi è più in alto/migliore?

Un modo corretto per effettuare confronti relativi è quello di standardizzare i termini del confronto. A tal fine basta sottrarre le medie delle popolazioni di provenienza e dividere quanto ottenuto per la relativa deviazione standard. Si ottengono così i valori che si vogliono confrontare espressi in termini di "*numero di deviazioni standard dalla (propria) media*"

$$\textbf{VALORE STANDARDIZZATO} = z = (x - \mu) / \sigma$$

Vediamo un esempio:
E' relativamente più alto un uomo alto 185 cm o una donna alta 176 cm?
Se le medie e le deviazioni standard delle altezze di uomini e donne italiane sono:
Uomini: media = 173 cm ; deviazione standard = 6 cm
Donne: media = 166 cm ; deviazione standard = 4 cm

Possiamo dunque calcolare:

185 standardizzato = (185 – 173)/6 = 2 deviazioni standard sopra la media
176 standardizzato = (176 – 166)/4 = 2,5 deviazioni standard sopra la media

e concludere che è <u>relativamente</u> più alta una donna di 176 cm che un uomo di 185cm.

3.2.6.2.1. I punti Z o punti standard

Mentre i punteggi centili, decili, quartili fanno riferimento alla *posizione occupata dal soggetto nella distribuzione* i **punti *z*** esprimono *distanze dalla media*.

I centili, decili, quartili possono essere calcolati anche per scale ordinali, mentre i **punti *z*** possono essere calcolati *solo* per scale ad intervalli o di rapporti, perché deve essere possibile calcolare media e deviazione standard.

Infatti i punti z rispondono sostanzialmente alla domanda: "Quante deviazioni standard posso misurare fra il valore *x* e la media di questa distribuzione?".

Se usiamo la deviazione standard come nuova unità di misura ed esprimiamo i valori della distribuzione secondo questo nuova unità di misura (quindi se dividiamo per la deviazione standard), quello che otteniamo è di avere una nuova scala che z mantiene tutte le relazioni esistenti nel sistema relazionale numerico della distribuzione di partenza ma ci fornisce molte informazioni in più. I punti z, tramite il segno, ci danno informazioni sul fatto che un certo valore sia inferiore (segno -) o superiore alla media (segno +), e su *quanto* il valore si discosti dalla media. Infatti se z è vicino allo zero il voto "grezzo" dello studente è vicino alla media della classe.

I nuovi punteggi standardizzati infatti assumono valori positivi quando il punteggio grezzo è superiore al valore medio, negativi (quando il punteggio grezzo è inferiore alla media del gruppo), nulli quando i punteggi grezzi coincidono col risultato medio.

Inoltre, con tale nuova unità di misura che ci possiamo confrontare fra loro i valori provenienti da distribuzioni diverse.

Il punteggio Zeta è definito dalla relazione:

$$Punt.Zeta = \frac{Punt.Grezzo - Media}{\text{Deviazione.\_standard}}$$

Esempio:

L'alunno Tizio è stato sottoposto a due differenti test di conoscenza, uno di inglese, l'altro di matematica. I due test avevano un numero differente di domande, ed egli ha ottenuto i due risultati:

TEST DI INGLESE=30

TEST DI MATEMATICA=70.

E' impossibile confrontare direttamente i due risultati. Se però trasformo i due valori in punti z, il confronto è assolutamente agevole:

	Test INGLESE	**Test MATEMATICA**
media	26	65
dev. standard	6	5
risultato ottenuto	**30**	**70**
punto z	**0,67**	**1**

Verifichiamo immediatamente che:

- Entrambi i punteggi sono buoni (poiché superiori alla media)
- Il punteggio ottenuto al test di matematica è migliore (1 è maggiore di 0,67).