

Unità Didattica N° 19

La variabilità

- 01) Brevi considerazioni generali**
- 02) Campo di variazione**
- 03) Lo scostamento semplice medio**
- 04) Varianza e scostamento quadratico medio**
- 05) Differenza semplice media**
- 06) Il coefficiente di variazione**

La variabilità

INTRODUZIONE

Le medie consentono di esprimere sinteticamente un fenomeno collettivo , ma esse vanno integrate sempre con **parametri** atti a descrivere sinteticamente la loro **variabilità** cioè il **grado di dispersione** o di disuguaglianza dei valori individuali . La metodologia statistica distingue due aspetti della variabilità : la **dispersione** e la **disuguaglianza** . La prima caratterizza il maggiore o minore addensamento delle **modalità** rilevate intorno alla media , mentre la seconda mette in evidenza la diversità delle varie **modalità** fra di loro , senza riferimento ad alcuna media . Le misure della **variabilità** riducono gli svantaggi determinati dalla perdita d'informazione che si produce sintetizzando la distribuzione con una media ed è opportuno accompagnarla alla media per una loro migliore interpretazione . Naturalmente , come possono elaborarsi infinite medie , così possono costruirsi infiniti misure o **indici di variabilità** , anche se soltanto alcuni di essi presentano un interesse teorico o trovano concrete applicazioni nella pratica delle analisi statistiche .

Campo di variazione o campo di variabilità

Il **campo di variazione** o campo di variabilità di una distribuzione di n dati x_i è la differenza tra il dato più grande $x_{\max}$ e quello più piccolo $x_{\min}$:

$$\omega = x_{\max} - x_{\min}$$

Si tratta di un *indice posizionale* assoluto . Esso è di facile determinazione , ma è assai grossolano e imperfetto per i seguenti motivi :

- a) dipende soltanto dai due estremi , ignorando la struttura della distribuzione nel suo interno
- b) è influenzata da eventuali valori estremi aberranti o eccezionali , che possono avere scarsa importanza essendo in genere poco numerosi .

Se si vuole evitare di considerare solo i due termini estremi , spesso troppo distanti dalla massa principale delle osservazioni , si può trascurare il primo 25 % e l'ultimo 25 % dei termini , limitandosi a considerare solo il 50 % degli elementi che si trovano nella parte intermedia della distribuzione .

Lo scostamento semplice medio

Agli **scostamenti medi** si perviene attraverso le fasi seguenti : 1) si sceglie una media **M** (di solito la media aritmetica o la mediana) rispetto alla quale vengono determinati gli **scostamenti** delle singole intensità : $x_i - M$ 2) si procede ad una sintesi di detti **scostamenti** , calcolandone una opportuna media (di solito la media aritmetica o la media quadratica) : il risultato viene assunto come misura (assoluta) della variabilità .

Giova osservare che gli **scostamenti medi** , che si ricavano dal procedimento descritto , sono indici di variabilità desunti da tutte le informazioni disponibili : essi costituiscono , pertanto , misure molto più valide di quelle fornite dagli indici posizionali che definiremo in seguito .

Definiamo lo **scostamento semplice medio dalla media aritmetica** . Se gli scarti $x_i - M$ vengono determinati per $M = \bar{x} = m = A$, cioè rispetto alla media aritmetica , una loro sintesi mediante la media aritmetica non potrebbe fornire una misura di variabilità poiché , come è noto , la somma (e quindi la media aritmetica) degli scarti è nulla . D'altra parte , ai fini della misura della dispersione , non interessa sapere se gli scarti sono positivi o negativi , quanto conoscere l'entità della << distanza >> che intercorre tra le singole **modalità** x_i e la media **m** , e viceversa .

Se consideriamo gli scarti da $\bar{x} = A$ e ne calcoliamo la media aritmetica dei loro valori assoluti otteniamo lo **scostamento semplice medio dalla media aritmetica** che ci viene fornito da

$$S_A = \frac{\sum_{i=1}^n |x_i - A|}{n} \quad \text{per dati non raggruppati}$$

$$S_A = \frac{\sum_{i=1}^n |x_i - A| \cdot n_i}{N = \sum_{i=1}^n n_i} \quad \text{per dati raggruppati}$$

Nel caso di distribuzioni di frequenza secondo classi di intensità del carattere , si suole fare ricorso al modello che concentra le frequenze nel valore centrale di ogni singola classe .

Invece se prendiamo in esame gli scarti dalla mediana M_e e ne calcoliamo la media aritmetica , otteniamo lo **scostamento semplice dalla mediana** , che risulta essere uguale a :

$$S_{M_e} = \frac{\sum_{i=1}^n |x_i - M_e|}{n} \quad \text{per dati non raggruppati}$$

$$S_{M_e} = \frac{\sum_{i=1}^n |x_i - M_e| \cdot n_i}{N = \sum_{i=1}^n n_i}$$

per **dati raggruppati**

Dei due **scostamenti semplici** medi il minore è quello calcolato rispetto ad M_e per la nota proprietà di cui gode la mediana . Nel caso in cui $m = M_e$ i due indici coincidono , cioè si ha : $S_{M_e} \leq S_m$

Anche per la determinazione di S_{M_e} nel caso di distribuzioni di frequenza secondo classi di intensità del carattere , si ricorre di solito alla sostituzione di x_i con il valore centrale delle varie classi .

Lo scostamento semplice medio dalla mediana è sempre minore dell'analogo scostamento semplice medio dalla media aritmetica .

La disuguaglianza è sempre valida , purché almeno due **modalità** siano diverse e deriva direttamente dalla già indicata proprietà della mediana secondo cui la somma dei valori assoluti degli scarti dalla mediana è minima .

Lo scostamento semplice medio assoluto dalla mediana M_e è **minimo** rispetto allo scostamento assoluto calcolato rispetto a qualsiasi altro valore medio .

Data la tabella

classi	16—21	21—26	26—31	31—36	36—41	41—46	N
frequenze assolute n_i	6.125	4.500	9.290	10.000	31.200	22.000	83.115

calcolare lo **scostamento semplice medio assoluto dalla mediana** M_e .

classi	frequenze assolute n_i	valori centrali x_i	frequenze cumulate C_i	
16—21	6125	18,5	6125	
21—26	4500	23,5	10.625	
26—31	9290	28,5	19.915	
31—36 x_{r-1}	10.000	33,5	29.915 C_{r-1}	$\bar{C} = \frac{N+1}{2}$
36—41 x_r	31.200	38,5	61.115 C_r	$\bar{C} = 41.580$

41—46	22.000	43,5	N = 83.115
-------	--------	------	------------

Individuo la **classe mediana**

$$\bar{C} = \frac{N+1}{2} = \frac{83115+1}{2} = 41.558 \quad \left\{ \begin{array}{l} \text{frequenza cumulata coincidente col punto} \\ \text{centrale della classe mediana } 36\text{—}41 \end{array} \right.$$

$$M_e = x_{r-1} + \frac{x_r - x_{r-1}}{C_r - C_{r-1}} (\bar{C} - C_{r-1}) = 36 + \frac{41 - 36}{61115 - 29915} (41558 - 29915) =$$

$$= 36 + \frac{5 \cdot 11643}{31200} = 36 + 1,86586 = 37,86586 \quad \boxed{M_e = 37,9}$$

classi	frequenze assolute n_i	valori centrali x_i	$ x_i - M_e $	$ x_i - M_e \cdot n_i$
16—21	6125	18,5	19,4	118.825
21—26	4500	23,5	14,4	64.800
26—31	9290	28,5	9,4	87.326
31—36 x_{r-1}	10.000	33,5	4	44.000
36—41 x_r	31.200	38,5	0,6	18.720
41—46	22.000	43,5	5,6	123.200
	$N = 83.115$		$\sum_{i=1}^6 x_i - M_e \cdot n_i$	456.871

$$S_{M_e} = \frac{\sum_{i=1}^6 |x_i - M_e| \cdot n_i}{\sum_{i=1}^6 n_i} = \frac{118825 + 64800 + 87326 + 44000 + 18720 + 123200}{83115} =$$

$$= \frac{456871}{83115} = 5,49685$$

$$|x_1 - M_e| \cdot n_1 = |18,5 - 37,9| \cdot 6125 = 118825 \quad , \quad |x_2 - M_e| \cdot n_2 = |23,5 - 37,9| \cdot 4500 = 64800$$

$$|x_3 - M_e| \cdot n_3 = |28,5 - 37,9| \cdot 9290 = 87326 \quad , \quad |x_4 - M_e| \cdot n_4 = |33,5 - 37,9| \cdot 10000 = 44000$$

$$|x_5 - M_e| \cdot n_5 = |38,5 - 37,9| \cdot 31200 = 18720 \quad , \quad |x_6 - M_e| \cdot n_6 = |43,5 - 37,9| \cdot 22000 = 123200$$

Calcolare anche lo **scostamento semplice medio dalla media aritmetica**

Scostamento quadratico medio

Il più usato degli indici di variabilità è dato dalla **media quadratica degli scarti dalla media aritmetica**, ed è noto come scostamento quadratico medio. . La sua espressione risulta :

per **dati non raggruppati**

$$S = \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

per **dati raggruppati**

$$S = \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 n_i}{\sum_{i=1}^n n_i}}$$

dove n_i rappresentano le frequenze assolute associate alle **modalità** x_i (per esempio i valori centrali delle classi).

Abbiamo detto che la misura di variabilità più usata è lo **scarto quadratico medio**. . Lo *scarto quadratico medio S di una distribuzione di dati è la media quadratica degli scarti dei singoli dati dalla loro media aritmetica*. Tale misura gode di alcune proprietà che la rendono particolarmente utilizzabile nelle elaborazioni matematiche dei dati.

Lo **scostamento quadratico medio** può essere calcolato quando sono note le **modalità** di tutte le unità del collettivo. Se la distribuzione statistica è presentata per classi di valori del carattere si ricorre di solito ai valori centrali delle classi.

E' immediato riconoscere che in alternativa allo scarto quadratico medio S risultante dalle differenze rispetto alla media aritmetica, potrebbero calcolarsi indici analoghi determinando gli scarti singoli con riferimento ad altra media (ad esempio la mediana, la media armonica, ...).

In ogni caso si otterrebbe una misura della variabilità maggiore di quella fornita da S che, nella famiglia degli scarti quadratici medi, ha il valore minimo.

Nella metodologia statistica, lo **scarto quadratico medio** rappresenta un indice privilegiato di dispersione: ciò segue dal fatto che esso si presta meglio di altre misure a particolari elaborazioni matematiche. Inoltre, se le **modalità** x_i sono tra loro diverse, essenzialmente per cause accidentali, lo scostamento quadratico medio risente meno di altre misure di variabilità dell'influenza del numero di osservazioni. Questa proprietà fa di esso un parametro fondamentale nella teoria delle rilevazioni campionarie.

Lo **scostamento quadratico medio** è meno sensibile di altre misure alla presenza di valori particolarmente alti o particolarmente bassi . Esso , perciò , amplifica le fluttuazioni attorno alla media , ciò che lo rende un indice di variabilità particolarmente ricettivo .

Il quadrato dello scostamento quadratico medio prende il nome di **varianza** ^(*)

$$S^2 = \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$S^2 = \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{N = \sum_{i=1}^n n_i}$$

Applicando le proprietà della sommatoria otteniamo una formula che ci permette di snellire i calcoli

della varianza S^2 .

$$S^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2$$

$$S^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^n n_i \cdot x_i^2}{N} - \left(\frac{\sum_{i=1}^n n_i \cdot x_i}{N} \right)^2$$

<< **La varianza è uguale alla media aritmetica dei quadrati dei dati meno il quadrato della media aritmetica dei dati** >> .

Quando è necessario distinguere lo scarto quadratico medio di una intera popolazione statistica dallo scarto quadratico medio di un campione tratto dalla popolazione stessa , spesso si usa il simbolo **S** per il secondo ed il simbolo **σ** per il primo . In questo caso il simbolo S^2 rappresenta la **varianza campionaria** mentre il simbolo σ^2 rappresenta la **varianza dell'intera popolazione statistica** .

Esempio di calcolo della varianza e dello scostamento quadratico medio .

(*) La varianza coincide col **secondo momento centrale** $\overline{m_2}$ (cioè col secondo m momento rispetto alla media) . Esiste tuttavia un'altra espressione della varianza in cui al denominatore compare $n - 1$ invece di n . In tal caso si

parla di **varianza corretta** . Tra le due **varianze** esiste la relazione $S_c^2 = S^2 \frac{n}{n - 1}$, dove S_c^2 è la varianza corretta

Classi di valori	Valori Centrali x_i	frequenze assolute n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot n_i$
125—130	127,5	1	-38,74	1500,7876	1500,7876
130—135	132,5	3	-33,74	1138,3876	3415,1628
135—140	137,5	11	-28,74	825,9876	9085,8636
140—145	142,5	38	-23,74	563,5876	21416,3288
145—150	147,5	106	-18,74	351,1876	37225,8856
150—155	152,5	152	-13,74	188,7876	28695,7152
155—160	157,5	238	-8,74	76,3876	18180,2488
160—165	162,5	305	-3,74	13,9876	4266,2180
165—170	167,5	315	1,26	1,5876	500,0940
170—175	172,5	302	6,26	39,1876	11834,6552
175—180	177,5	234	11,26	126,7876	29668,2984
180—185	182,5	128	16,26	268,3876	33841,6128
185—190	187,5	50	21,26	451,9876	22599,3800
190—195	192,5	19	26,26	689,5876	13102,1644
195—200	197,5	3	31,26	977,1876	2931,5628
					238.263,9780

$A = m = \bar{x} \cong 166,24 = \text{media aritmetica}$

$$S^2 = \frac{238263,978}{1905} = 125,0729 = \text{varianza}$$

$$S = \sqrt{125,0729} = 11,18 = \text{scostamento quadratico medio}$$

. Quest'ultima è preferibile a quella non corretta quando si tratta di stimare per mezzo di un campione, la varianza dell'universo da cui il campione è tratto. La correzione è insignificante per n elevato.

Differenza semplice media

La **differenza media** è un indice di disuguaglianza fondato sui confronti diretti delle osservazioni . La maniera più opportuna per realizzare il confronto tra le osservazioni x_i ed x_j è quella di considerare le differenze $x_i - x_j$. Gli indici di variabilità studiati fin qui , fatta eccezione per il campo di variazione , sono stati calcolati utilizzando le differenze dei dati da un valore medio , in genere la media aritmetica . Un altro modo per misurare la variabilità di una distribuzione di dati è quello che tiene conto di tutte le differenze di ciascun dato da tutti gli altri . Si parla in questo caso di **differenze medie** e si indicano con Δ . Le differenze medie sono utilizzate soprattutto per le distribuzioni in cui la scelta di una media piuttosto che un'altra risulta arbitraria . In questi casi si risolve il problema di definire la variabilità dei dati confrontandoli direttamente tra loro . Dati n valori $x_1, x_2, \dots, x_n$ calcoliamo tutte le differenze fra ogni termine e ciascuno degli altri n termini (compreso se stesso) . Si ottiene un quadro come è indicato nella seguente tabella :

x_1	x_2	x_3		x_n	
$x_1 - x_1$	$x_1 - x_2$	$x_1 - x_3$		$x_1 - x_n$	x_1
$x_2 - x_1$	$x_2 - x_2$	$x_2 - x_3$		$x_2 - x_n$	x_2
.....					.
$x_n - x_1$	$x_n - x_2$	$x_n - x_3$		$x_n - x_n$	x_n

Si tratta di un **quadrato numerico** di n righe ed n colonne , i cui termini posti sulla **diagonale principale** (quelli cioè del tipo $x_i - x_i$) sono tutti nulli .

Riferendoci al caso di n dati semplici e distinti , il numero totale delle differenze è n^2 , di cui n sono nulle e quindi le rimanenti $n^2 - n = n(n - 1)$ sono diverse da zero .

Dalla **tabella delle differenze** si deducono alcuni indici di variabilità , chiamati **differenze medie** così definite . Si chiama **differenza media assoluta** la media aritmetica dei valori assoluti delle differenze . Tale differenza è con ripetizione se si considerano tutte le n^2 differenze , **senza ripetizione** se si considerano le $n(n - 1)$ differenze ottenute escludendo i termini tutti nulli della diagonale principale .

Se indichiamo con Δ_R e Δ rispettivamente le differenze medie con ripetizione e senza , abbiamo :

$$\Delta_R = \frac{|x_1 - x_1| + |x_1 - x_2| + \dots + |x_1 - x_n| + |x_2 - x_1| + \dots + |x_n - x_n|}{n^2} = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{n^2}$$

$$\Delta = \frac{|x_1 - x_2| + \dots + |x_1 - x_n| + |x_2 - x_1| + \dots + |x_n - x_{n-1}|}{n(n-1)} = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{n(n-1)}$$

Nel caso di dati raggruppati è necessario tenere conto di tutte le frequenze e si ottiene la formula :

$$\Delta = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \cdot n_i \cdot n_j}{N(N-1)} \quad N = \sum_{i=1}^n n_i$$

$$\Delta_R = \frac{\sum_{i=1}^5 \sum_{j=1}^5 |x_i - x_j| \cdot n_i \cdot n_j}{N^2}$$

in quanto al valore x_i è associata la frequenza assoluta n_i , al valore x_j è associata la frequenza n_j .

Data la seguente distribuzione di frequenza calcolare la differenza media con ripetizione (Δ_R) e senza ripetizione (Δ) .

x_i	2	6	7	9	12	$N = n_1 + n_2 + n_3 + n_4 + n_5$
n_i	3	5	6	4	2	20

Per eseguire i calcoli si può elaborare la seguente tabella all'interno della quale disponiamo tutte le possibili differenze $|x_i - x_j| \cdot n_i \cdot n_j$.

		$n_1 = 3$	$n_2 = 5$	$n_3 = 6$	$n_4 = 4$	$n_5 = 2$	Totali
		$x_1 = 2$	$x_2 = 6$	$x_3 = 7$	$x_4 = 9$	$x_5 = 12$	Parziali
$n_1 = 3$	$x_1 = 2$	————	$ x_1 - x_2 n_1n_2 = 60$	90	84	60	294
$n_2 = 5$	$x_2 = 6$	60	————	30	60	60	150
$n_3 = 6$	$x_3 = 7$	90	84	————	48	60	282
$n_4 = 4$	$x_4 = 9$	84	60	48	————	24	216
$n_5 = 2$	$x_5 = 12$	60	60	60	24	————	204
						Totale generale	1146

Si noti che i valori riportati nella tabella sono simmetrici rispetto alla diagonale principale (diagonale discendente) e quindi in realtà dobbiamo calcolare soltanto $\frac{n(n-1)}{2}$ differenze .

$$\Delta_R = \frac{\sum_{i=1}^5 \sum_{j=1}^5 |x_i - x_j| \cdot n_i \cdot n_j}{N^2} = \frac{1146}{400} = 2,865 \quad , \quad \Delta = \frac{\sum_{i=1}^5 \sum_{j=1}^5 |x_i - x_j| \cdot n_i \cdot n_j}{N(N-1)} = \frac{1146}{380} = 3,016$$

Schema generale per 4 dati con relative frequenze assolute

		n_1	n_2	n_3	n_4	Totali
		x_1	x_2	x_3	x_4	Parziali
n_1	x_1	————	$ x_1 - x_2 n_1n_2$	$ x_1 - x_3 n_1n_3$	$ x_1 - x_4 n_1n_4$	
n_2	x_2	$ x_2 - x_1 n_2n_1$	————	$ x_2 - x_3 n_2n_3$	$ x_2 - x_4 n_2n_4$	
n_3	x_3	$ x_3 - x_1 n_3n_1$	$ x_3 - x_2 n_3n_2$	————	$ x_3 - x_4 n_3n_4$	
n_4	x_4	$ x_4 - x_1 n_4n_1$	$ x_4 - x_2 n_4n_2$	$ x_4 - x_3 n_4n_3$	————	
						Totale Generale

Il coefficiente di variazione

Le misure di variabilità viste finora (campo di variazione , scarti semplici medi assoluti e scarto quadratico medio , differenze medie) sono espresse nella stessa unità di misura dei dati della distribuzione e sono dette **misure assolute** della variabilità .

Ne risulta che i confronti tra indici di variabilità non sono possibili se i fenomeni ai quali si riferiscono non sono espressi nella stessa unità di misura .

Ad esempio non è possibile confrontare lo scarto semplice medio delle precipitazioni mensili con quello delle temperature medie mensili in una stazione meteorologica in quanto il primo è espresso in **mm** , il secondo in **gradi** .

Un altro ostacolo è costituito dal fatto che , anche a parità di unità di misura , le entità dei fenomeni potrebbero essere diverse , si da rendere privo di significato il confronto tra i relativi indici di variabilità .

Ad esempio lo scarto quadratico medio del peso delle formiche (espresso in grammi) non può essere confrontato con lo scarto quadratico medio del peso degli elefanti (anche se espresso in grammi) . Una dispersione di 20g nella misurazione di una grandezza che ha peso 1Kg è molto più rilevante di una stessa dispersione riscontrata nella misurazione di una grandezza avente peso medio di 10 Kg .

Allo scopo di rimuovere i due suddetti ostacoli al confronto fra indici di variabilità si suole rapportare tali indici (detti **indici assoluti**) alla media aritmetica , pervenendo agli **indici relativi di variabilità** . Così operando si ottengono numeri puri , ossia privi di unità di misura e svincolati dalla grandezza media del fenomeno . In genere gli indici relativi di variabilità vengono moltiplicati per 100 in modo da esprimere la misura della variabilità come percentuale della media .

Si definisce **indice relativo di variabilità** il rapporto tra un indice assoluto ed un valore medio .

$$\text{indice relativo di variabilità} = \frac{\text{indice assoluto}}{\text{valore medio}}$$

Nel caso in cui l'indice assoluto è lo scarto quadratico medio ed il valore medio è la media aritmetica , l'indice relativo si chiama **coefficiente di variazione** e si indica con :

$$C.V. = \frac{S}{m} = \frac{S}{\bar{x}}$$

Il **coefficiente di variazione** di una distribuzione è il rapporto tra lo scarto quadratico medio e la media aritmetica della distribuzione .

In genere , per evitare troppe cifre decimali , il coefficiente di variazione viene espresso in forma percentuale :

$$C.V. = \frac{S}{\bar{x}} \cdot 100$$

Calcolare il coefficiente di variazione della seguente distribuzione :

x_i	6	7	8	9	10	
n_i	12	22	31	18	17	$N = 100$

La media aritmetica è : $\bar{x} = 8,06$

Lo scarto quadratico medio è : $S = 1,248$

$$V.C. = \frac{1,248}{8,06} = 0,155 = 15,5\%$$

Si consideri un gruppo di reclute baresi che ad una visita di leva ha presentato le seguenti caratteristiche fisiche :

Carattere	medie	Scarti quadratici medi
Statura	167,51 cm	6,48 cm
Torace	87,87 cm	5,16 cm
Peso	60,64 Kg	8,64 Kg

Gli scarti quadratici medi di statura e torace non sono confrontabili perché si riferiscono a grandezze diverse (la statura è circa il doppio del torace) . Il confronto con lo scarto quadratico medio del peso è poi

ostacolato anche dalla diversa unità di misura . I coefficienti di variabilità sono :

$$\frac{6,48 \text{ cm}}{167,51 \text{ cm}} 100 = 3,9\% \quad \text{per la statura} \quad , \quad \frac{5,16 \text{ cm}}{87,87 \text{ cm}} 100 = 5,9\% \quad \text{per il torace}$$

$$\frac{8,64 \text{ Kg}}{60,64 \text{ Kg}} \cdot 100 = 14,2\% \quad \text{per il peso}$$

Da essi si vede che la statura (che dipende quasi esclusivamente da fattori ereditari) è il carattere meno variabile , mentre il peso (che dipende prevalentemente da fattori esogeni) è il carattere più variabile . Intermedia la variabilità del torace .

Le distribuzioni dei pesi e delle stature di un gruppo di studenti hanno presentato media e scarto quadratico medio come dal seguente prospetto . Risulta maggiore la variabilità della distribuzione dei pesi o delle stature ?

	Media $\bar{x}$	Scarto Quadratico Medio
peso in Kg	59,40	7,5
statura in cm	173,2	8,2

Poiché non si possono confrontare direttamente i Kg con i cm , calcoliamo il coefficiente di variazione di ciascuna distribuzione :

$$C.V. \text{ peso} = \frac{7,5}{59,5} 100 = 12,6\% \quad , \quad C.V. \text{ statura} = \frac{8,2}{173,2} 100 = 4,7\%$$

Risulta maggiore la variabilità della distribuzione dei pesi .