

VARIANZA di Y

- se i dati relativi a Y sono "alla rinfusa" (ossia anche ripetuti), ossia se

$$Y = (\underbrace{y_1, \dots, y_N}_{\text{possono coincidere!}}) \text{ allora}$$

$$\bar{Y} = \frac{y_1 + y_2 + \dots + y_N}{N}$$

$$s_y^2 = \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \dots + (y_N - \bar{y})^2}{N}$$

$$\uparrow \sum_{i=1}^N (y_i - \bar{y})^2$$

Modo

COMPATTO

ESEMPIO (sulla dispersione ES. 7.3)

15 CAVIE

Resi 28 32 37 29 31 31
 " " " " " "
 y_1 y_2 y_3 y_4 y_4 y_{15}

$$\sigma_y^2 = \frac{(28 - 30.2)^2 + (32 - 30.2)^2 + \dots}{15}$$

$$\text{dove } 30.2 = \frac{28 + 32 + 37 + \dots + 31}{15}$$

- Se organizziamo i dati considerando solo quelli distinti.

$\{z_1, \dots, z_k\}$ ^{TUTTI} $\neq$
 e costruiamo la tabella delle frequenze assolute

Hod.	$z_1 z_2 \dots$	z_k
Freq ASS	$f_1 f_2$	f_k

allora $\bar{y} = \frac{f_1 \cdot z_1 + f_2 \cdot z_2 + \dots + f_k \cdot z_k}{\sum_{i=1}^k f_i}$

e $\sigma_y^2 = \frac{f_1 \cdot (z_1 - \bar{y})^2 + \dots + f_k \cdot (z_k - \bar{y})^2}{\sum_{i=1}^k f_i}$

che in modo compatto possiamo

scrivere come

$$s_y^2 = \frac{\sum_{i=1}^k f_i (z_i - \bar{y})^2}{\sum_{i=1}^k f_i}$$

} dimensione
del campione
 $N =$

Allora per calcolare la varianza
si può aggiungere alla nostra
tabella una riga con gli
scarti quadratici $(z_i - \bar{y})^2$
ESEMPIO (delle cavie)

COSTRUIAMO LA TABELLA

PESI	26	27	28	...	...	37
Freq. ASS	1	1	2	...	...	1
SCARTI QUADR	$(26-30.2)^2$	...	...	...	...	$(37-30.2)^2$

ossia la tabella

PESI	26	27	28	...	32	37
SCARTI QUAD.	19.64	10.24	...	...	3.24	46.24
Freq. Ass	1	1	...	...	3	1

↑ CAMBIO ORDINE per mettere le
frequenze sotto a tutto

$$\Rightarrow \sigma_y^2 = \frac{(19.64) \cdot 1 + (10.24) \cdot 1 + \dots + (3.24) \cdot 3 + (46.24) \cdot 1}{15}$$

$(Y_1, \dots, Y_N)$



	$z_1 \dots \dots$	z_k	
Freq. ASS	$p_1 \dots \dots p_k$		$\bar{y} = \frac{\sum_{i=1}^k z_i \cdot f_i}{N}$
Scarti quadr	$(z_1 - \bar{y})^2 \dots \dots (z_k - \bar{y})^2$		
	$\sigma_y^2 = \frac{\sum_{i=1}^N (z_i - \bar{y})^2}{N}$		

Se conosciamo invece le frequenze relative

Mod.	z_1	$\dots$	z_k
Freq. REL.	p_1	$\dots$	p_k

allora $\bar{Y} = \sum_{i=1}^N z_i \cdot p_i$

e $\sigma_Y^2 = \sum_{i=1}^K p_i (z_i - \bar{Y})^2$

(perché $p_i = f_i \cdot N$ e
e' N che è il denominatore
viene semplificato)

FORMULA FURBA

$$\sigma_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N} =$$

$$\uparrow \quad \frac{\sum_{i=1}^N y_i^2 + \bar{y}^2 - 2y_i \cdot \bar{y}}{N}$$

quadrato
di binomio

A FATTOR
COMUNE

$$\uparrow \quad \frac{\sum_{i=1}^N y_i^2 + \sum_{i=1}^N \bar{y}^2 - 2 \sum_{i=1}^N y_i \cdot \bar{y}}{N}$$

SPEZZO

LA SOMMA

(PROPR.

ASSOCIATIVA

DELLA SOMMA)

sono tutti $\bar{y}^2$
gli addendi = a $\bar{y}$
quindi tale somma
è $N \cdot \bar{y}^2$

$$= \frac{\sum_{i=1}^N y_i^2 + N \bar{y}^2 - 2 \left(\sum_{i=1}^N y_i \right) \cdot \bar{y}}{N}$$

↑
A FATTOR
COMUNE

$$= \frac{\sum_{i=1}^N y_i^2}{N} + \frac{N \bar{y}^2}{N} - 2 \frac{\sum_{i=1}^N y_i}{N} \cdot \bar{y}$$

$$= \frac{\sum_{i=1}^N y_i^2}{N} + \bar{y}^2 - 2 \bar{y} \cdot \bar{y}$$

$$= \frac{\sum y_i^2}{N} + \bar{y}^2 - 2 \bar{y}^2$$

$$= \boxed{\overline{y^2} - \bar{y}^2}$$

↙
MEDIA
DEI QUADRATI

↘
quadrato
della MEDIA

Per usare la formula furba

occorre costruire la tabella con

- le modalità $\{z_1, \dots, z_k\}$ distinte
- i loro quadrati $z_1^2, z_2^2, \dots, z_k^2$
- le frequenze assolute.

QUINDI

CALCOLARE

$$1) \bar{y} = \frac{\sum_{i=1}^k f_i \cdot z_i}{\sum_{i=1}^k f_i}$$

$$2) (\bar{y})^2$$

3) Calcolare la media dei quadrati

$$\overline{y^2} = \frac{\sum_{i=1}^k f_i \cdot z_i^2}{\sum_{i=1}^k f_i}$$

ESEMPIO (con la formula furba)

Riprendiamo le caviglie

PESO	26	27	28	...	32	37
PESI AL QUADRATO	676	729	784	...	...	1369
Freq. ASS	1	1	2		3	1

$$\overline{Y^2} = \frac{676 \cdot 1 + 729 \cdot 1 + \dots + 1369 \cdot 1}{15}$$

$$= 918,6 \text{ g}^2$$

$$\sigma_Y^2 = \overline{Y^2} - (\overline{Y})^2 = 918,6 - 912,04$$
$$= 6,56 \text{ g}^2$$

Infine definiamo deviazione standard o SCARTO QUADRATICO MEDIO

la radice quadrata della varianza, quindi se $Y = (Y_1, \dots, Y_N)$

$$\sigma_Y = \sqrt{\frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N}}$$

DEVIAZIONE
STANDARD

oppure se uso la freq. assolute

$$\sigma_Y = \sqrt{\frac{\sum_{i=1}^N f_i (z_i - \bar{Y})^2}{N}}$$

$$\sum_{i=1}^N \underbrace{P_i}_{\text{num. di CREDITI}} \underbrace{Y_i}_{\text{voto dell' } i\text{-mo esame}}$$

= numero di esami

VOTO = _____



di ammissione
alla laurea

$$\sum_{i=1}^N \underbrace{P_i}_{\text{SOMMA DEI CREDITI}}$$