



Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Esercitazione 4

Statistica

Alfonso Iodice D'Enza
iodicede@gmail.com

Università degli studi di Cassino



Outline

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

- 1 La varianza
- 2 Altri indici di variabilità
- 3 Mutua variabilità
- 4 Esercizio sulla variabilità



Definizione di varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza un'indice che misura la variabilità di una variabile X rispetto alla media aritmetica. In particolare la varianza σ^2 data dalla media dei quadrati degli scarti (delle modalità dalla media)



Definizione di varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza un'indice che misura la variabilità di una variabile X rispetto alla media aritmetica. In particolare la varianza σ^2 data dalla media dei quadrati degli scarti (delle modalità dalla media)

$$\begin{aligned}\sigma^2 &= \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n} = \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2\end{aligned}$$



Definizione di varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza un'indice che misura la variabilità di una variabile X rispetto alla media aritmetica. In particolare la varianza σ^2 data dalla media dei quadrati degli scarti (delle modalità dalla media)

$$\begin{aligned}\sigma^2 &= \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n} = \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2\end{aligned}$$

per dati organizzati in frequenze (seriazione)

$$\begin{aligned}\sigma^2 &= \frac{(x_1 - \mu)^2 \times n_1 + (x_2 - \mu)^2 \times n_2 + \dots + (x_k - \mu)^2 \times n_k}{n_1 + n_2 + \dots + n_k} = \\ &= \frac{1}{n} \sum_{i=1}^k (x_i - \mu)^2 \times n_i\end{aligned}$$



Esempio di calcolo della varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Data la variabile X : *numero di esami sostenuti prima di quello di statistica* osservata su un collettivo di $n = 6$ studenti



Esempio di calcolo della varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Data la variabile X : *numero di esami sostenuti prima di quello di statistica* osservata su un collettivo di $n = 6$ studenti

<i>Esami sostenuti prima di 'statistica I'</i>	<i>Scarti dalla media $\mu=4.1667$</i>	<i>Scarti dalla media al quadrato</i>
3	$3-4.1667= -1.1667$	$(-1.1667)^2=1.3611$
2	$2-4.1667= -2.1667$	$(-2.1667)^2=4.6944$
4	$4-4.1667= -0.1667$	$(-0.1667)^2=0.0278$
10	$10-4.1667= 5.8333$	$(5.8333)^2=34.0278$
5	$5-4.1667= 0.8333$	$(0.8333)^2=0.6944$
1	$1-4.1667= -3.1667$	$(-3.1667)^2=10.0278$
<i>totale</i>	<i>0</i>	<i>50.8333</i>

Esempio di calcolo della varianza

Data la variabile X : *numero di esami sostenuti prima di quello di statistica* osservata su un collettivo di $n = 6$ studenti

<i>Esami sostenuti prima di 'statistica I'</i>	<i>Scarti dalla media $\mu=4.1667$</i>	<i>Scarti dalla media al quadrato</i>
3	$3-4.1667= -1.1667$	$(-1.1667)^2=1.3611$
2	$2-4.1667= -2.1667$	$(-2.1667)^2=4.6944$
4	$4-4.1667= -0.1667$	$(-0.1667)^2=0.0278$
10	$10-4.1667= 5.8333$	$(5.8333)^2=34.0278$
5	$5-4.1667= 0.8333$	$(0.8333)^2=0.6944$
1	$1-4.1667= -3.1667$	$(-3.1667)^2=10.0278$
<i>totale</i>	<i>0</i>	<i>50.8333</i>

La varianza sar dunque

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = \frac{50.8333}{6} = 8.4722$$





Massima variabilità

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità



Massima variabilità

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

- La varianza può crescere indefinitamente perchè gli scarti delle modalità dalla media possono essere illimitatamente grandi



Massima variabilità

Esercitazione

4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

- La varianza può crescere indefinitamente perchè gli scarti delle modalità dalla media possono essere illimitatamente grandi
- La situazione di **massima variabilità** per un collettivo con media μ , si ha quando su n modalità, $n - 1$ sono nulle ed una sola modalità $x_i^{\neq 0} = n\mu$

$$\begin{aligned}\sigma^2 &\leq \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = \frac{1}{n} ((n-1)(0 - \mu)^2 + (n\mu - \mu)^2) = \\ &= \frac{1}{n} ((n-1)\mu^2 + \mu^2(n-1)^2) = \\ &= \frac{1}{n} ((n-1)\mu^2 + \mu^2(n^2 + 1 - 2n)) = \\ &= \frac{1}{n} (\mu^2 n(n-1)) = \mu^2(n-1)\end{aligned}$$



Le proprietà della varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza gode di alcune importanti proprietà di seguito riportate:



Le proprietà della varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza gode di alcune importanti proprietà di seguito riportate:

- 1 La varianza di X sempre un numero non negativo (≥ 0)
- 2 La varianza di X pari a 0 se e solo se X una costante
- 3 Se alla variabile X si aggiunge una costante, σ_x non cambia
- 4 Se si moltiplica la variabile X per una costante b , si avrà
$$\sigma_{x^*} = b^2 \sigma_x^2$$



Le proprietà della varianza

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Le proprietà 3 e 4 dipendono dalla proprietà di linearità della media aritmetica: si consideri $Y = a + bX$, con a e b costanti. Dalla proprietà risulta che $\mu_y = a + b\mu_x$. Calcolando la varianza di Y si avrà:

$$\begin{aligned}\sigma_y^2 &= \frac{1}{n} \sum_{i=1}^n (y_i - \mu_y)^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (y_i - (a + b\mu_x))^2 = \frac{1}{n} \sum_{i=1}^n (a + bx_i - a - b\mu_x)^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (bx_i - b\mu_x)^2 = \frac{1}{n} \sum_{i=1}^n b^2 (x_i - \mu_x)^2 = \\ &= b^2 \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2 = b^2 \sigma_x^2\end{aligned}$$



Lo scarto quadratico medio (standard deviation)

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Una difficoltà di interpretazione della varianza dipende dal fatto che tale indice espresso nell'unità di misura al quadrato della variabile cui si riferisce. Per ovviare a questo problema si utilizza lo *scarto quadratico medio* σ , dato da



Lo scarto quadratico medio (standard deviation)

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Una difficoltà di interpretazione della varianza dipende dal fatto che tale indice espresso nell'unità di misura al quadrato della variabile cui si riferisce. Per ovviare a questo problema si utilizza lo *scarto quadratico medio* σ , dato da

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



Lo scarto quadratico medio (standard deviation)

Esercitazione

4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Una difficoltà di interpretazione della varianza dipende dal fatto che tale indice espresso nell'unità di misura al quadrato della variabile cui si riferisce. Per ovviare a questo problema si utilizza lo *scarto quadratico medio* σ , dato da

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

dall'esempio precedente risulta dunque

$$\sigma = \sqrt{\frac{50.8333}{6}} = 2.9107$$



Il coefficiente di variazione (CV)

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza è un indice assoluto, dipende quindi dall'unità di misura della variabile. Un indice **relativo** di variabilità è il coefficiente di variazione CV . E' dato da

$$CV = \frac{\sigma}{\mu}$$

essendo un numero puro consente il confronto fra fenomeni rilevati in momenti diversi o espressi in unità di misura diverse



Il coefficiente di variazione (CV)

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

La varianza è un indice assoluto, dipende quindi dall'unità di misura della variabile. Un indice **relativo** di variabilità è il coefficiente di variazione CV . E' dato da

$$CV = \frac{\sigma}{\mu}$$

essendo un numero puro consente il confronto fra fenomeni rilevati in momenti diversi o espressi in unità di misura diverse

Limiti di utilizzo del CV

- è definito solo se $\mu > 0$
- se $\mu \rightarrow 0$ il CV tende a diventare molto grande



Variabilità e modalità ordinate

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

In caso di variabili con modalità ordinabili è possibile definire degli indici di variabilità derivati dalla funzione di ripartizione empirica. Data la distribuzione unitaria ordinata di modalità

$$\{1, 5, 7, 13, 14, 15, 18, 18, 22, 23, 24, 25, 27, 28, 29\}$$



Variabilità e modalità ordinate

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

In caso di variabili con modalità ordinabili è possibile definire degli indici di variabilità derivati dalla funzione di ripartizione empirica. Data la distribuzione unitaria ordinata di modalità

$\{1, 5, 7, 13, 14, 15, 18, 18, 22, 23, 24, 25, 27, 28, 29\}$

- il **range** (o campo di variazione) è dato da
 $R(X) = \max(x_i) - \min(x_i) = 29 - 1 = 27$
- il **range inter-quartile** (o campo di variazione interquartile) è dato da $IQR(X) = Q_3 - Q_1 = 25 - 13 = 12$



Variabilità e modalità ordinate

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

In caso di variabili con modalità ordinabili è possibile definire degli indici di variabilità derivati dalla funzione di ripartizione empirica. Data la distribuzione unitaria ordinata di modalità

$\{1, 5, 7, 13, 14, 15, 18, 18, 22, 23, 24, 25, 27, 28, 29\}$

- il **range** (o campo di variazione) è dato da
 $R(X) = \max(x_i) - \min(x_i) = 29 - 1 = 27$
- il **range inter-quartile** (o campo di variazione interquartile) è dato da $IQR(X) = Q_3 - Q_1 = 25 - 13 = 12$



Mutua variabilità

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

In presenza di **caratteri trasferibili** (reddito, risorse energetiche, consumo di beni) è di maggior interesse lo studio della variabilità tra le singole unità statistiche piuttosto che la variabilità rispetto a un centro.

Mutua variabilità

Esercitazione
4

A. Indice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

In presenza di **caratteri trasferibili** (reddito, risorse energetiche, consumo di beni) è di maggior interesse lo studio della variabilità tra le singole unità statistiche piuttosto che la variabilità rispetto a un centro.

Differenza media semplice

tale indice rappresenta la media dei valori assoluti delle differenze calcolate rispetto a tutte le possibili coppie di modalità. Esso corrisponde a

$$\Delta = \frac{\sum_{i \neq j=1}^n |x_i - x_j|}{n(n-1)}$$

la quantità al denominatore ($n(n-1)$) rappresenta il numero di possibili coppie di n osservazioni.



Mutua variabilità

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

**Mutua
variabilità**

Esercizio sulla
variabilità

Dato un carattere X osservato su $n = 4$ osservazioni

$\{7, 14, 18, 24\}$



Mutua variabilità

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Dato un carattere X osservato su $n = 4$ osservazioni

$$\{7, 14, 18, 24\}$$

Differenza media semplice

Il valore di Δ sarà in questo caso

$$\begin{aligned}\Delta &= \frac{|7 - 14| + |7 - 18| + |7 - 24| + |14 - 7| +}{12} \\ &\quad + \frac{|14 - 18| + |14 - 24| + |18 - 7| + |18 - 14| + |18 - 24| +}{12} \\ &\quad + \frac{|24 - 7| + |24 - 14| + |24 - 18|}{12} = \\ &= \frac{110}{12} = 9.1667\end{aligned}$$



mutua variabilità: minimo e massimo

Esercitazione
4

A. Indice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

- l'indice Δ assume il valore minimo ($\Delta = 0$) quando tutte le modalità coincidono: in questo caso le differenze semplici sono nulle
- l'indice Δ assume il valore massimo quando tutte le modalità tranne una sono nulle: in tal caso si ha che $\Delta = 2\mu$

Dunque Δ assume valore nell'intervallo $[0, 2\mu]$: è possibile ottenere una versione **normalizzata**:

$$R = \frac{\Delta}{2\mu}$$

tale indice viene denominato **rapporto di concentrazione di Gini**

Esercizio

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Si consideri un campione di $n = 100$ sportelli bancari e sia X : numero di operazioni effettuate presso uno sportello nell'ultima settimana.

X	Freq (n_i)
[60, 62]	5
[63, 65]	18
[66, 68]	42
[69, 71]	27
[72, 74]	8
	100

Quesiti

- Calcolare un indice di tendenza centrale (media)
- Misurare la variabilità rispetto ad un centro (scostamento medio semplice e scarto quadratico medio)
- Misurare la mutua variabilità

Esercizio: svolgimento

Quesito 1

- Calcolare un indice di tendenza centrale (media)

Calcolo della media aritmetica

Per calcolare la media aritmetica bisogna individuare i centri di ciascuna classe e tenere conto delle frequenze. Ricordando la formula della media aritmetica

$$\mu = \frac{\sum_{i=1}^k c_i \times n_i}{n}$$

X	Freq= n_i	centri= c_i	$c_i \times n_i$
[60, 62]	5	61	305
[63, 65]	18	64	1152
[66, 68]	42	67	2814
[69, 71]	27	70	1890
[72, 74]	8	73	584
	100		6745

$$\mu = \frac{\sum_{i=1}^k c_i \times n_i}{n} = \frac{305 + 1152 + 2814 + 1890 + 584}{100} = 67.45$$



Esercizio: svolgimento

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Quesito 2

- Misurare la variabilità rispetto ad un centro : **scostamento medio semplice**

Calcolo dello scostamento medio semplice

Per ottenere lo scostamento medio semplice bisogna calcolare il valore assoluto degli scarti dei centri delle classi dalla media e tenere conto delle frequenze. Ricordando la formula dello scostamento medio semplice

$$MD = \frac{\sum_{i=1}^k |c_i - \mu| \times n_i}{n}$$

X	Freq= n_i	centri= c_i	$ c_i - \mu $	$ c_i - \mu \times n_i$
[60, 62]	5	61	6.45	32.25
[63, 65]	18	64	3.45	62.10
[66, 68]	42	67	0.45	18.90
[69, 71]	27	70	2.55	68.85
[72, 74]	8	73	5.55	44.40
	100			226.5

$$MD = \frac{\sum_{i=1}^k |c_i - \mu| \times n_i}{n} = \frac{32.25 + 62.1 + 18.9 + 68.85 + 44.4}{100} = \frac{226.5}{100} = 2.265$$



Esercizio: svolgimento

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Quesito 2

- Misurare la variabilità rispetto ad un centro : **scarto quadratico medio**

Calcolo della varianza

Per ottenere lo scarto quadratico medio, si procede al calcolo della varianza che rappresenta il **valore medio dei quadrati degli scarti dalla media aritmetica**: bisogna calcolare i quadrati degli scarti dei centri delle classi dalla media e tenere conto delle frequenze. Ricordando la formula della varianza

$$s^2 = \frac{\sum_{i=1}^k (c_i - \mu)^2 \times n_i}{n}$$

X	Freq= n_i	centri= c_i	$(c_i - \mu)^2$	$(c_i - \mu)^2 \times n_i$
[60, 62]	5	61	41.60	208.0125
[63, 65]	18	64	11.90	214.245
[66, 68]	42	67	0.20	8.505
[69, 71]	27	70	6.50	175.5675
[72, 74]	8	73	30.80	246.42
		100		852.75

$$s^2 = \frac{208.0125 + 214.245 + 8.505 + 175.5675 + 246.42}{100} = \frac{852.75}{100} = 8.5275$$

da cui lo scarto quadratico medio $s = \sqrt{8.5275} = 2.92$

Esercizio: svolgimento

Quesito 2

- Misurare la variabilità rispetto ad un centro : **scarto quadratico medio**

Calcolo della varianza

Ripetendo il calcolo utilizzando la formula alternativa per il calcolo della varianza

$$s^2 = \frac{\sum_{i=1}^k c_i^2 \times n_i}{N} - \left(\frac{\sum_{i=1}^k c_i \times n_i}{N} \right)^2 = \frac{\sum_{i=1}^k c_i^2 \times n_i}{N} - (\mu)^2$$

X	Freq= n_i	centri= c_i	c_i^2	$c_i^2 \times n_i$
[60, 62]	5	61	3721	18605
[63, 65]	18	64	4096	73728
[66, 68]	42	67	4489	188538
[69, 71]	27	70	4900	132300
[72, 74]	8	73	5329	42632
	100			455803

$$s^2 = \frac{\sum_{i=1}^k c_i^2 \times n_i}{N} - \left(\frac{\sum_{i=1}^k c_i \times n_i}{N} \right)^2 = \frac{455803}{100} - \left(\frac{6745}{100} \right)^2 = 4558.03 - 4549.503 = 8.5275$$

da cui lo scarto quadratico medio $s = \sqrt{8.5275} = 2.92$



Esercizio: svolgimento

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Quesito 3

- Misurare la mutua variabilità : **differenza media semplice**

Calcolo della mutua variabilità

La differenza media semplice corrisponde al valore medio delle differenze tra le possibili coppie di modalità.

$$\Delta = \frac{\sum_{i \neq j=1}^N |x_i - x_j|}{N(N-1)}$$

la quantità al denominatore ($N(N-1)$) rappresenta il numero di possibili coppie di n osservazioni. Nel caso di intervalli di valori e di distribuzione di frequenze si fa riferimento ai centri delle classi c_i e alle frequenze. La formula da utilizzare è la seguente

$$\Delta = \frac{\sum_{i \neq j=1}^k |x_i - x_j| \times n_i \times n_j}{N(N-1)}$$



Esercizio: svolgimento

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Quesito 3

- Misurare la mutua variabilità : **differenza media semplice**

X	Freq= n_i	centri= c_i
[60, 62]	5	61
[63, 65]	18	64
[66, 68]	42	67
[69, 71]	27	70
[72, 74]	8	73
	100	

	$ c_i - c_j $	$n_i \times n_j$	$ c_i - c_j \times n_i \times n_j$
1	3	90	270
2	6	210	1260
3	9	135	1215
4	12	40	480
5	3	90	270
6	3	756	2268
7	6	486	2916
8	9	144	1296
9	6	210	1260
10	3	756	2268
11	3	1134	3402
12	6	336	2016
13	9	135	1215
14	6	486	2916
15	3	1134	3402
16	3	216	648
17	12	40	480
18	9	144	1296
19	6	336	2016
20	3	216	648
			31542



Esercizio: svolgimento

Esercitazione
4

A. Iodice

La varianza

Altri indici di
variabilità

Mutua
variabilità

Esercizio sulla
variabilità

Quesito 3

- Misurare la mutua variabilità : **differenza media semplice** metodo rapido.

c_i/c_j	61	64	67	70	73
61	25	90	210	135	40
64	90	324	756	486	144
67	210	756	1764	1134	336
70	135	486	1134	729	216
73	40	144	336	216	64

$c_i - c_j$	$n_i \times n_j$
3	90
6	210
9	135
12	40
3	756
6	486
9	144
3	1134
6	336
3	216

$$\begin{aligned}\Delta &= \frac{2 \times \sum_{i=1}^{k-1} \sum_{j=(i+1)}^k |x_i - x_j| \times n_i \times n_j}{N(N-1)} = \\ &= \frac{(3 \times 90) + (6 \times 210) + (9 \times 135) + \dots + (3 \times 216)}{100 \times 99} = 3.186\end{aligned}$$