Varianza

Esistono due tipi di varianza : varianza campionaria e varianza empirica.

Si definisce VARIANZA CAMPIONARIA

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

(strumento corretto per analizzare un campione)

Si considera la somma di distanze al quadrato dei dati dalla media , la si valuta e la si divide per n-1 e non per la numerosità dei dati.

Nella VARIANZA EMPIRICA invece si divide per la numerosità

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

(ha a che fare con l'intera popolazione)

La varianza indica quanto i dati discostano dalla media.

Presi gli Esempi di prima si vede che nonostante i tre esempi hanno la solita media i diagrammi sono diversi e la varianza cambia da caso a caso.

Varianze piccole mi dicono che i dati sono tutti vicino alla media mentre varianze grandi mi dicono che i dati sono distribuiti più lontani dalla media.

Esempi:

Esempio 1 :

$x = (1,1,2,2,2,3,3,3,3,4,4,4,5,5), \qquad n = 14, \qquad \bar{x} = 3$

$$var(x) = \frac{1}{13}(4+4+1+1+1+0+0+0+0+1+1+1+4+4) = \frac{1}{13} * 22 = \frac{22}{13}$$

$$var_e(x) = \frac{1}{14}(4+4+1+1+1+0+0+0+0+1+1+1+4+4) = \frac{1}{14} * 22 = \frac{22}{14}$$

Esempio 2 :

$x = (1,1,1,2,2,2,2,3,3,3,5,5,6,6), \qquad n = 14, \qquad \bar{x} = 3$

$$var(x) = \frac{1}{13}(4+4+4+1+1+1+1+0+0+0+4+4+9+9) = \frac{1}{13} * 42 = \frac{42}{13}$$

$$var_e(x) = \frac{1}{14}(4+4+4+1+1+1+1+0+0+0+4+4+9+9) = \frac{1}{14} * 42 = 3$$

Esempio 3 :

$x = (1,1,1,1,1,1,1,5,5,5,5,5,5,5), \qquad n = 14, \qquad \bar{x} = 3$

$$\mathrm{var(x)} = \frac{1}{13}(4+4+4+4+4+4+4+4+4+4+4+4+4+4) = \frac{1}{13} * 56 = \frac{56}{13}$$
$$\mathrm{var_e(x)} = \frac{1}{13}(4+4+4+4+4+4+4+4+4+4+4+4+4+4) = \frac{1}{14} * 56 = 4$$

OSSERVAZIONE

Se provo a sviluppare

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

Ottengo che

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

E lo posso riscrivere come

$$\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x}$$

L'ultimo termine $2\bar{x} = n\bar{x}$ *e allora*

$$\sum_{i=1}^{n} x_i^2 - n\bar{x}^2$$

OSSERVAZIONE

Cosa succede se la varianza è uguale a zero?

$$var(x) = 0 \Leftrightarrow \sum_{i=1}^{n}(x_i - \bar{x})^2 = 0 \Leftrightarrow (x_i - \bar{x})^2 = 0 \ \forall\, i$$

Questo è possibile solo quando tutti i valori coincidono con la media e questo è possibile solo quando sono tutti uguali quindi se

$x_i = \bar{x} \ \forall\, i$

- SCARTO QUDRATICO
  Lo scarto quadratico medio (o deviazione standard) campionaria è : $\sigma(x) = \sqrt{var(x)}$

Lo scarto quadratico medio (o deviazione standard) empirica è : $\sigma_e(x) = \sqrt{var_e(x)}$

Si introduce in quanto in molte situazioni sarà più utile parlare della radice quadrata della varianza

Continua su Scarto quadratico medio