

Indici di mutabilità

Per mutabilità si intende l'attitudine di un carattere qualitativo ad assumere modalità differenti.

Indici di omogeneità e di eterogeneità

Una popolazione ripartita, rispetto ad un carattere, in k modalità è omogenea (si intende rispetto a quel carattere) se le N unità di cui si compone presentano tutte la stessa modalità. In altre parole una popolazione è omogenea, rispetto ad un carattere, se la totalità delle frequenze è concentrata in una sola modalità.

Se invece le frequenze sono ripartite in parti uguali fra le k modalità del carattere, allora si è in presenza della minima omogeneità.

Accanto al concetto di omogeneità si può considerare anche il concetto opposto, cioè la eterogeneità. Se l'omogeneità si può configurare come l'attitudine di un carattere a concentrare le frequenze su poche modalità, l'eterogeneità si può configurare come la tendenza di un carattere ad assumere tutte le modalità con uguale frequenza.

Deriva da quanto detto, che un indice di omogeneità deve assumere il valore massimo se le frequenze si concentrano in una sola modalità ed il valore minimo se le stesse si ripartiscono uniformemente fra tutte le modalità. L'opposto deve accadere per un indice di eterogeneità.

Indici di omogeneità e di eterogeneità di Gini

Come è noto date due modalità A_i e A_j di un carattere qualitativo sconnesso è possibile stabilire fra esse solamente la relazione di uguaglianza o di disuguaglianza. Volendo assegnare ad esse una distanza $d(i,j)$ si può porre

$$d(i, j) = \begin{cases} 0 & \text{se } i = j \\ 1 & \text{se } i \neq j \end{cases}$$

Gini ha ricavato un indice di eterogeneità calcolando la media aritmetica semplice di tutte le N^2 distanze che si ottengono confrontando ogni unità con se stessa e con tutte le rimanenti.

Esempio:

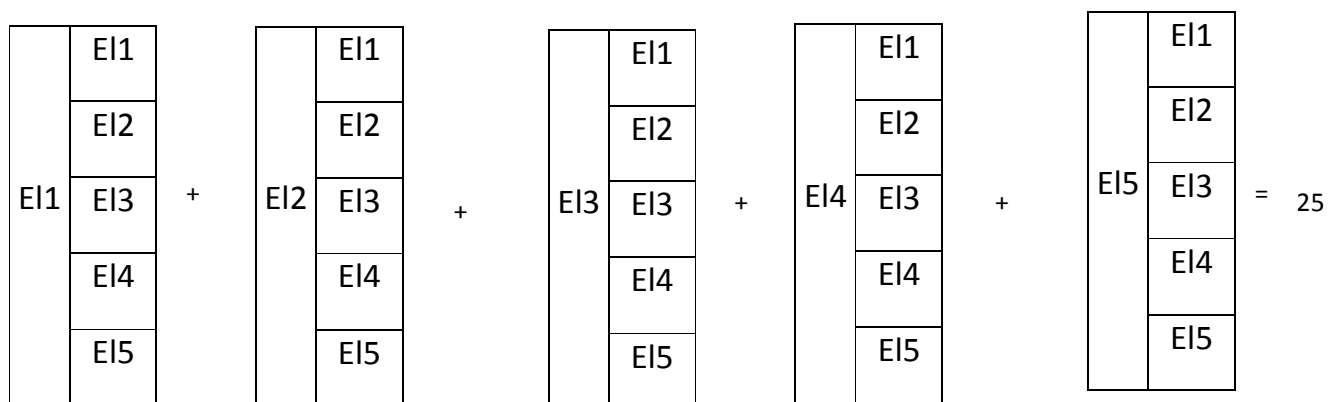
Data la seguente tabella del titolo di studi di una collettività di 16 studenti determinare il grado di eterogeneità e di omogeneità.

Titolo studio	n_i
Lic. Elementare	5
Lic. Media	8
Lic. Superiore	2
Laurea	1
Totale	16 (N)

Il numero dei confronti possibili delle 5 unità della modalità Lic. Elementare sono così ottenute: $5 \cdot 16 (n_1 \cdot N) = 80$

Di queste : $5^2 = n_1^2$ confronti fra le unità della stessa modalità

$80 - 5^2 = n_1 \cdot N - n_1^2$ confronti con le altre modalità



I confronti tra le 5 unità della modalità Lic. Elementare e le 8 unità della Lic. Media sono $5 \cdot 8 = 40$

In tale modo si ottiene la tabella dei confronti , nella quale i numeri in rosso rappresentano i confronti tra le stesse modalità . Il loro totale è pari a 94

di	El	Med	Sup	Lau
El	0	1	1	1
Med	1	0	1	1
Sup	1	1	0	1
Lau	1	1	1	0

confronti	El	Med	Sup	Lau	
El	25	80	10	5	80
Med	40	64	16	8	128
Sup	10	16	4	2	32
Lau	5	8	2	1	16
	80	128	32	16	256

Si ottengono così gli indici:

$$\text{Eterogeneità} = E_2 = \frac{N^2 - \sum_i n_i^2}{N^2} = \frac{256 - 94}{256} = 0.63$$

$$\text{Omogeneità} = O_2 = \frac{\sum_i n_i^2}{N^2} = \frac{94}{256} = 0.37$$

Casi estremi

Titolo studio	n_i	f_i
Lic. Elementare	0	0
Lic. Media	16	1
Lic. Superiore	0	0
Laurea	0	0
Totale	16 (N)	1

Minima eterogeneità

Massimo omogeneità

$$n_i = N \dots \dots \dots f_i = 1$$

$$n_j = 0 \dots \dots \dots \forall j \neq i$$

Titolo studio	n_i	f_i
Lic. Elementare	4	$\frac{1}{4}$
Lic. Media	4	$\frac{1}{4}$
Lic. Superiore	4	$\frac{1}{4}$
Laurea	4	$\frac{1}{4}$
Totale	16 (N)	1

Massimo eterogeneità

Minimo omogeneità

$$n_1 = n_2 = \dots \dots \dots = n_k$$

$$f_1 = f_2 = \dots \dots \dots = f_k = \frac{1}{k}$$

Normalizzazione dell'indice del Gini

Per essere significativo un indice deve essere relativizzato onde fornire una misura che sia confrontabile con altre al fine di stabilire se il valore trovato è più o meno elevato.

Per avere informazioni sul grado di maggiore o minore elevatezza dell'eterogeneità è necessario quindi normalizzare l'indice.

Eterogeneità

$$I_N = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \dots\dots\dots > \dots 0 \leq I_N \leq 1$$

$$E_2 = 1 - \sum_{i=1}^k f_i^2$$

$$E_2^N = \frac{E_2 - E_{\min}}{E_{\max} - E_{\min}}$$

$$E_{2..min} = 1 - (0 + 1 + 0 + 0) = 0$$

$$E_{2..max} = 1 - \sum_{i=1}^k \left(\frac{1}{k}\right)^2 = 1 - k * \frac{1}{k^2} = \frac{k-1}{k}$$

$$E_2^N = \frac{\frac{E_2}{k-1}}{\frac{k}{4}} = \frac{0.63}{3} = 0.84 \text{ -----} > \text{grande..eterogeneità}$$

Omogeneità

$$O_2 = 1 - E_2 = \sum_{i=1}^k f_i^2$$

$$O_{2..min} = \sum_{i=1}^k \left(\frac{1}{k}\right)^2 = \frac{1}{k} = \frac{1}{4}$$

$$O_{2..max} = \sum_{i=1}^k f_i^2 = 1$$

$$O_2^N = \frac{O_2 - O_{2..min}}{O_{2..max} - O_{2..min}} = \frac{0.37 - \frac{1}{4}}{1 - \frac{1}{4}} = 0.16$$

