Concentrazione

- Nello studio della distribuzione della ricchezza, è di fondamentale importanza l'aspetto della ***concentrazione***. Intuitivamente, la concentrazione è elevata quando poche unità della popolazione possiedono gran parte della ricchezza.La concentrazione è minima (equidistribuzione) quando tutte le unità hanno la stessa ricchezza.

- Un carattere *quantitativo trasferibile*, le cui modalità ordinate sono $x_1,\ldots,x_n$, si dice ***equidistribuito*** se ognuna delle n unità possiede una quota dell'ammontare del carattere pari a

$$\frac{1}{n}A \quad \text{dove} \quad A=\sum_{i=1}^{n}x_n$$

che coincide con la media aritmetica $\bar{x}$.

- Se non c'è *equidistribuzione* allora si ha ***concentrazione***.

- Si ha ***massima concentrazione*** quando una sola unità del collettivo possiede tutto l'ammontare del carattere e tutte le altre nulla, cioè

$$x_1=\ldots=x_{n-1}=0 \quad e \quad x_n=A$$

- **Esempio**: si hanno 100 soggetti e l'ammontare complessivo del reddito mensile è A = 50.000€. Se c'è equidistribuzione ogni soggetto ha reddito pari a 500€ mentre nel caso di massima concentrazione un solo soggetto ha reddito pari a 50.000€ e gli altri soggetti non hanno reddito.

Misurazione della concentrazione (distribuzione unitaria)

- **Ammontare del carattere** posseduto dalle i unità "più povere" dopo aver ordinato i termini della distribuzione ($x_1 \le x_2 \le \ldots \le x_n$)

$$A_i=x_1+x_2+\ldots+x_i=\sum_{j=1}^{i}x_j$$

- **Ammontare relativo del carattere** posseduto dalle i unità "più povere"

$$Q_i=\frac{A_i}{A}=\sum_{j=1}^{i}x_j \Big/ \sum_{j=1}^{n}x_j$$

- **Ammontare relativo del carattere** posseduto dalla i unità "più povere" nel caso (*ipotetico*) di *equidistribuzione*:

$$F_i=i/n$$

- Per *qualsiasi distribuzione* si ha: $F_i \ge Q_i, \quad \forall i \quad e \quad F_n=Q_n=1$

- All'aumentare della concentrazione aumentano le differenze: F_i-Q_i

- Nel caso di *massima concentrazione* si ha: $Q_1=Q_2=\ldots=Q_{n-1}=0$

- Come indice sintetico si usa il ***rapporto di concentrazione di Gini*** che si ottiene come rapporto tra $\sum_{i=1}^{n-1}(F_i-Q_i)$ e il suo valore massimo:

$$R=\frac{\sum_{i=1}^{n-1}(F_i-Q_i)}{\sum_{i=1}^{n-1}F_i}=1-\frac{\sum_{i=1}^{n-1}Q_i}{\sum_{i=1}^{n-1}F_i}$$

- L'indice di Gini cresce al crescere del livello di concentrazione ed è sempre compreso tra 0 (nel caso di *equidistribuzione*) e 1 (nel caso di *massima concentrazione*).

- Un altro strumento che permette di valutare il grado di concentrazione è la ***curva di Lorenz***. Si tratta di un grafico ottenuto unendo con dei segmenti i punti di coordinate (F_i, Q_i), per $i = 1,\ldots,n$. Maggiore è l'area tra la curva di Lorenz e la bisettrice, maggiore è la concentrazione.

Esempio

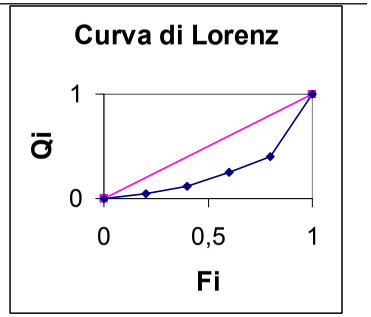
- Per un gruppo di 5 soggetti si ha la seguente distribuzione del reddito mensile

Unità (i)	Reddito (x_i)
1	250
2	650
3	3000
4	750
5	350
Totale	5000

- Non c'è equidistribuzione e quindi c'è concentrazione. Per quantificarne il livello si ordinano prima le modalità ottenendo

Reddito (x_i)	A_i	Q_i	F_i	F_i-Q_i
250	250	0,05	0,2	0,15
350	600	0,12	0,4	0,28
650	1250	0,25	0,6	0,35
750	2000	0,40	0,8	0,40
3000	5000	1,00	1	0,00
5000	--	--	2	1,18

da cui l'indice di Gini è pari a R =1,18/ 2 = 0,59



Misurazione della concentrazione (distribuzione di frequenze)

- **Ammontare del carattere** posseduto dalle unità "più povere" (ossia quelle che possiedono un ammontare $\leq x_h$)

$$A_h = x_1 n_1 + x_2 n_2 + \ldots + x_j n_j = \sum_{j=1}^{h} x_j n_j$$

- **Ammontare relativo del carattere** posseduto dalle unità con ammontare $\leq x_h$

$$Q_h = \frac{A_h}{A} = \sum_{j=1}^{h} x_j n_j \Big/ \sum_{j=1}^{K} x_j n_j$$

$$F_h = N_h / n$$

Misurazione della concentrazione (distribuzione in classi)

- **Ammontare del carattere** posseduto dalle j classi "più povere"

$$A_j = \bar{x}_1 n_1 + \bar{x}_2 n_2 + \ldots + \bar{x}_j n_j = \sum_{i=1}^{j} \bar{x}_i n_i$$

dove abbiamo fatto l'ipotesi di uniforme distribuzione del carattere all'interno della classe e usato il valore centrale della classe

- **Ammontare relativo del carattere** posseduto dalle j classi "più povere"

$$Q_j = \frac{A_j}{A} = \sum_{i=1}^{j} \bar{x}_i n_i \Big/ \sum_{i=1}^{K} \bar{x}_i n_i$$

- **Ammontare relativo del carattere** posseduto dalle j classi "più povere" nel caso (*ipotetico*) di *equidistribuzione*:

$$F_j = N_j / n$$

- Un'approssimazione del ***rapporto di concentrazione*** si ottiene in questo caso come

$$R = 1 - \sum_{j=0}^{K-1} \left(F_{j+1} - F_j\right)\left(Q_{j+1} + Q_j\right)$$

in cui F_0 e Q_0 sono state poste per convenzione uguali a zero.

- Se si ha a disposizione l'ammontare delle classi, questo viene utilizzato al posto dell'ammontare approssimativo $\bar{x}_j n_j$

Esempio

Fatturato	Freq. (n_j)	$\bar{x}_j$	$\bar{x}_j n_j$	A_j	Q_j	N_j	F_j
0-1	144	0,5	72	72	0,0113	144	0,158
1-5	457	3	1371	1443	0,2264	601	0,660
5-10	171	7,5	1282,5	2725,5	0,4277	772	0,847
10-25	112	17,5	1960	4685,5	0,7352	884	0,970
25-100	27	62,5	1687,5	6373	1	911	1

$$R = 1 - \sum_{j=0}^{K-1} \left(F_{j+1} - F_j\right)\left(Q_{j+1} + Q_j\right) =$$

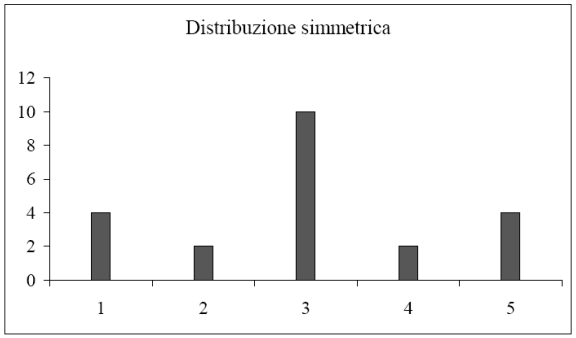
$$1 - 0{,}158 \cdot 0{,}0113 - (0{,}660 - 0{,}158)(0{,}0113 + 0{,}2264) - \ldots = 0{,}562$$

Asimmetria

- Una distribuzione di frequenza con *mediana* M_e è ***simmetrica*** se:

➢ $|x_1 - M_e| = |x_k - M_e|, \quad |x_2 - M_e| = |x_{k-1} - M_e|, \quad |x_3 - M_e| = |x_{k-2} - M_e|, \quad \ldots$

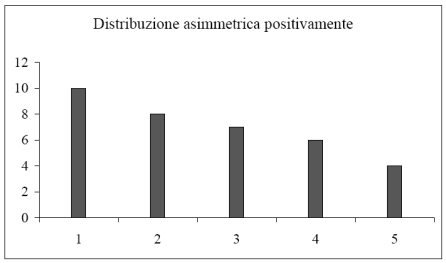
➢ $n_1 = n_k, \quad n_2 = n_{k-1}, \quad n_3 = n_{k-2}, \quad \ldots$

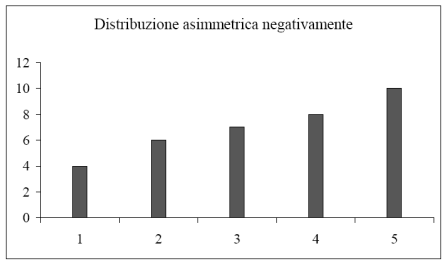


- Una distribuzione si dice ***asimmetrica*** se le condizioni precedenti non sono rispettate. In particolare, si può avere

➢ ***Asimmetria positiva***: sono più frequenti le modalità più piccole;

➢ ***Asimmetria negativa***: sono più frequenti le modalità più grandi.





Misurazione dell'asimmetria

- *L'indice di asimmetria* più utilizzato è quello di *Fisher*:

$$\beta_1 = \frac{1}{\sigma^3}\left[\frac{1}{n}\sum_{j=1}^{k}\left(x_j - \bar{x}\right)^3 n_j\right]$$

- Si noti che se la distribuzione di frequenza è simmetrica si ha:

$$M_e - Q_1 = Q_3 - M_e$$

  e quindi si può costruire un indice di asimmetria basato su statistiche d'ordine:

$$\beta_2 = \frac{\left(Q_3 - M_e\right) - \left(M_e - Q_1\right)}{Q_3 - Q_1}$$

- Usualmente, quando gli indici sono maggiori di 0 si ha asimmetria positiva e quando sono negativi si ha asimmetria negativa.

- Per la prima distribuzione del numero di figli per un collettivo di 25 famiglie, che ha media 2,04 e deviazione standard 0,871 si ha:

Popolazione 1

N. Figli (x_j)	Frequenze (n_j)	$\left(x_j - \bar{x}\right)^3 n_j$
0	1	-8,4897
1	4	-4,4995
2	15	-0,0010
3	3	2,6542
4	2	15,0591
Totale	25	4,7232

$$\beta_1 = \frac{1}{0{,}871^3}\left(\frac{4{,}7232}{25}\right) = 0{,}286$$