

Misure di disuguaglianza

CONCENTRAZIONE

LA VARIABILITA' è l'attitudine delle unità di un collettivo ad assumere differenti modalità di un carattere.

Misurare la dispersione significa calcolare la variabilità rispetto a valori medi...

Misurare la disuguaglianza significa, invece, quantificare la variabilità tra le diverse modalità

CONCENTRAZIONE

Studia la variabilità esistente tra N unità statistiche rispetto ad un **carattere quantitativo TRASFERIBILE** (reddito, spesa, numero di dipendenti, quote di mercato...)

Obiettivo \longrightarrow verificare se esiste:

- ✓ **Equidistribuzione** (concentrazione nulla): tutte le unità statistiche possiedono la stessa quantità dell'ammontare totale del carattere (T)
- ✓ **Massima concentrazione**: l'ammontare totale è posseduto da un'unica unità statistica
- ✓ **Situazione intermedia** (concentrazione) rispetto alle prime due

FREQUENZE RELATIVE E CUMULATE

INDICANDO CON

$x_1 x_2 x_3 \dots x_n$ n osservazioni di un carattere X
quantitativo trasferibile (p.e. reddito) ordinate

$$x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$$

La concentrazione si misura a partire da due
grandezze

F_i : frequenze relative cumulate

Q_i : intensità relative cumulate

$$F_i = \frac{i}{n} \quad i = 1, 2, \dots, n$$

INTENSITA' RELATIVE CUMULATE

Ogni frazione di percettori (indicata con F_i) detiene una frazione (indicata con Q_i) dell'ammontare totale (A) del carattere

Definendo le intensità assolute cumulate come

$$A_1 = x_1$$

$$A_2 = x_1 + x_2 \quad (\text{il reddito delle prime due unità})$$

.....

$$A_i = x_1 + x_2 + \dots + x_i \quad (\text{il reddito delle prime } i \text{ unità}),$$

la generica intensità relativa cumulata Q_i è data da

$$Q_i = \frac{A_i}{A_n}, \quad i = 1, 2, \dots, n$$



$$Q_1 = \frac{A_1}{A_n}, \quad Q_2 = \frac{A_2}{A_n}, \quad \dots, \quad Q_n = \frac{A_n}{A_n} = 1$$

Concentrazione dei ricavi dell'azienda per punto vendita

Ricavi (valori ordinati)	Freq. rel. cum. F_i	Intens. ass. cum. A_i	Intens. rel. cum. Q_i
$X_{(1)}=180$	$1/9=0,11$	180	$180/2925=0,06$
$X_{(2)}=200$	$2/9=0,22$	380	$380/2925=0,13$
$X_{(3)}=205$	$3/9=0,33$	585	$585/2925=0,20$
$X_{(4)}=270$	$4/9=0,44$	855	$855/2925=0,29$
$X_{(5)}=280$	$5/9=0,56$	1135	$1135/2925=0,39$
$X_{(6)}=340$	$6/9=0,67$	1475	$1475/2925=0,50$
$X_{(7)}=350$	$7/9=0,78$	1825	$1825/2925=0,62$
$X_{(8)}=500$	$8/9=0,89$	2325	$2325/2925=0,79$
$X_{(9)}=600$	$9/9=1$	2925	$2925/2925=1$

Il primo terzo dei punti vendita (33%) realizza il 20% dei ricavi totali

I primi due terzi dei punti vendita (67%) realizzano il 50% dei ricavi totali

$$\Sigma = 2925$$

Concentrazione

Come si misura il grado di concentrazione?

PARTENDO DALLA SOMMA DELLE
DIFFERENZE $\sum (F_i - Q_i)$

SI HA:

$$\sum_{i=1}^{n-1} (F_i - Q_i) = \begin{cases} 0 & \text{equidistribuzione} \longrightarrow \text{Ogni unità (1/n) detiene esattamente una frazione del carattere pari a 1/n.} \\ \sum_{i=1}^{n-1} F_i & \text{massima concentrazione} \\ & Q_1 = Q_2 = \dots = Q_{n-1} = 0 \end{cases}$$

Rapporto di concentrazione

Un *indice relativo di concentrazione* è definito da:

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

$$0 \leq R \leq 1$$

Rapporto di concentrazione di Gini

$R=0$ se c'è equidistribuzione

$R=1$ se c'è massima concentrazione

Concentrazione dei ricavi dell'azienda per punto vendita

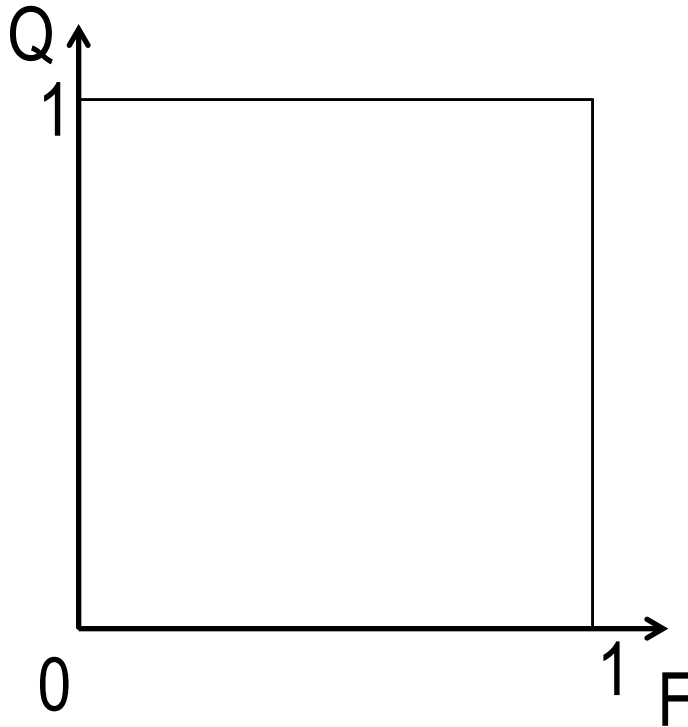
Ricavi (valori ordinati)	Freq. rel. cum. F_i	Intens. rel. cum. Q_i	$F_i - Q_i$
$X_{(1)}=180$	0,11	0,06	0,05
$X_{(2)}=200$	0,22	0,13	0,09
$X_{(3)}=205$	0,33	0,20	0,13
$X_{(4)}=270$	0,44	0,29	0,15
$X_{(5)}=280$	0,56	0,39	0,17
$X_{(6)}=340$	0,67	0,50	0,17
$X_{(7)}=350$	0,78	0,62	0,16
$X_{(8)}=500$	0,89	0,79	0,10
$X_{(9)}=600$	1	1	0

$$\sum_{i=1}^{n-1} (F_i - Q_i) = 1,02$$

$$\sum_{i=1}^{n-1} F_i = 4,00$$

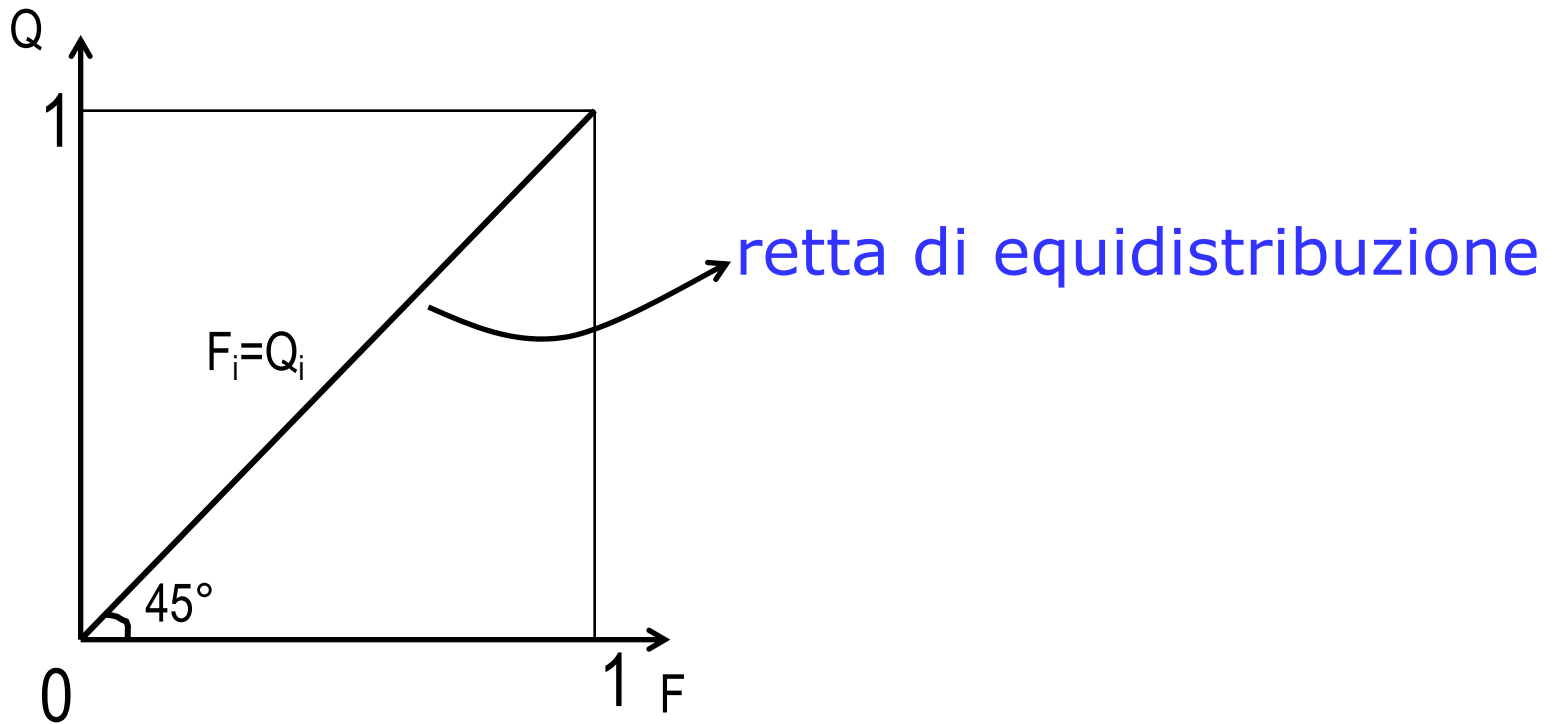
$$R = \frac{1,02}{4,00} = 0,255$$

Rappresentazione grafica

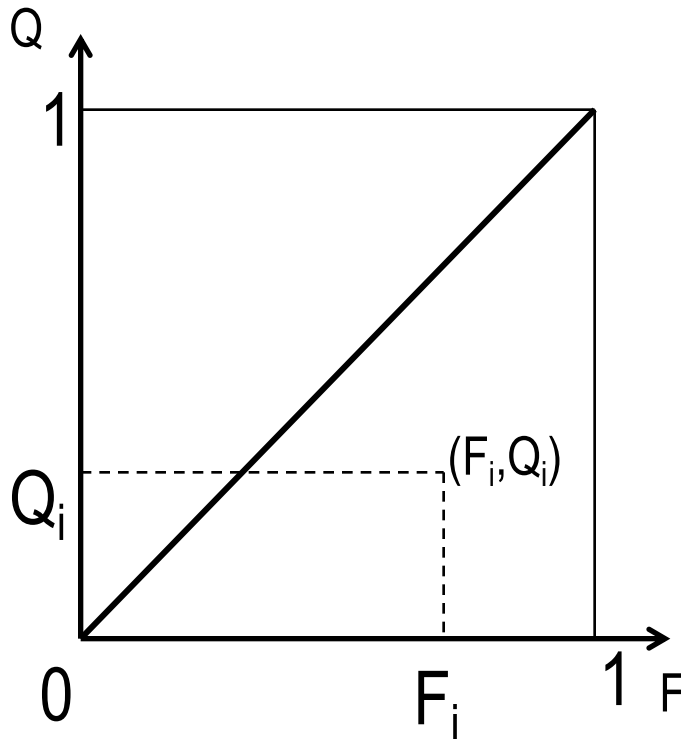


Tutti i punti di coordinate (F_i, Q_i) si trovano all'interno del quadrato di lato 1 (perché $0 \leq F_i \leq 1$ e $0 \leq Q_i \leq 1$)

Rappresentazione grafica

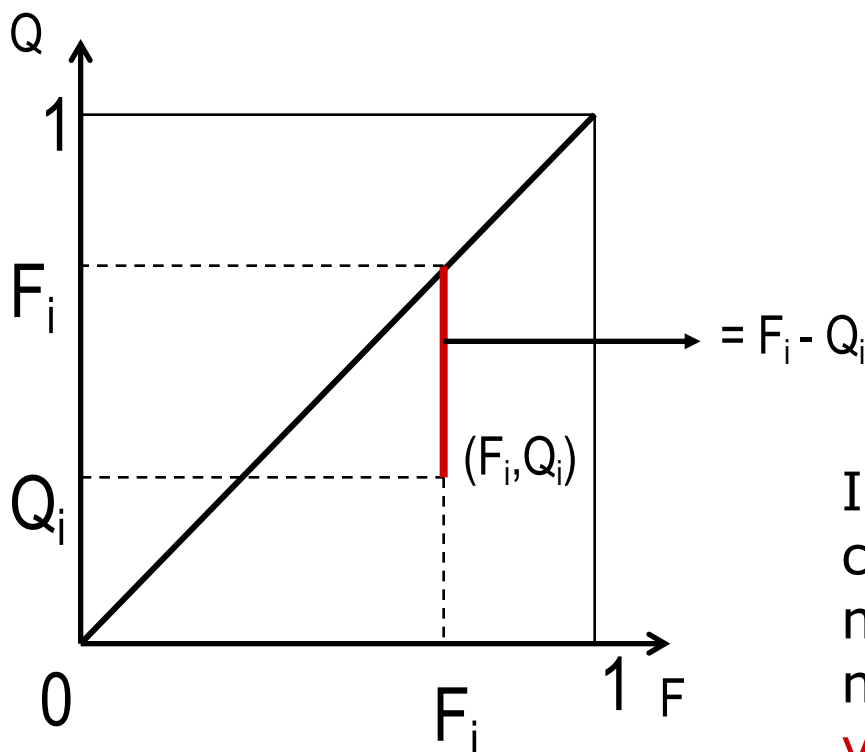


Rappresentazione grafica



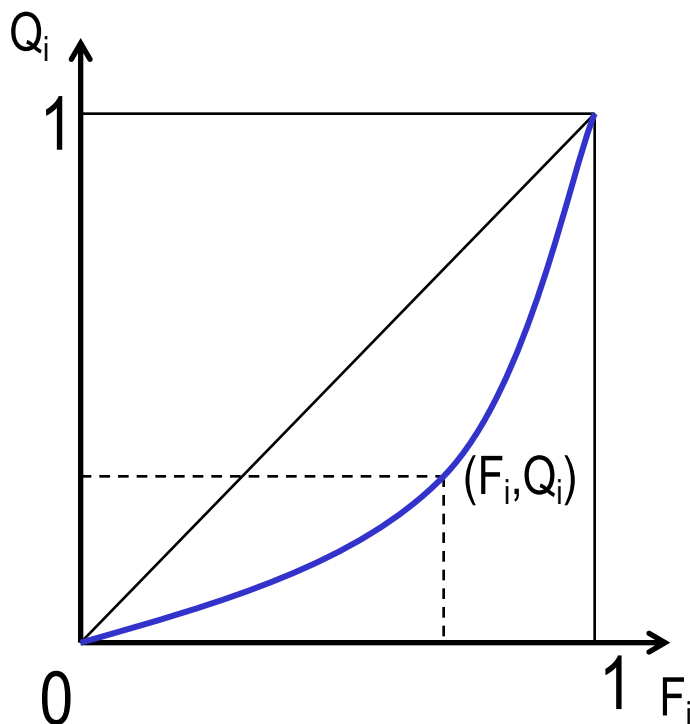
Tutti i punti di coordinate (F_i, Q_i) si trovano **al di sotto della retta a 45°** (perché $F_i \geq Q_i$)

Rappresentazione grafica



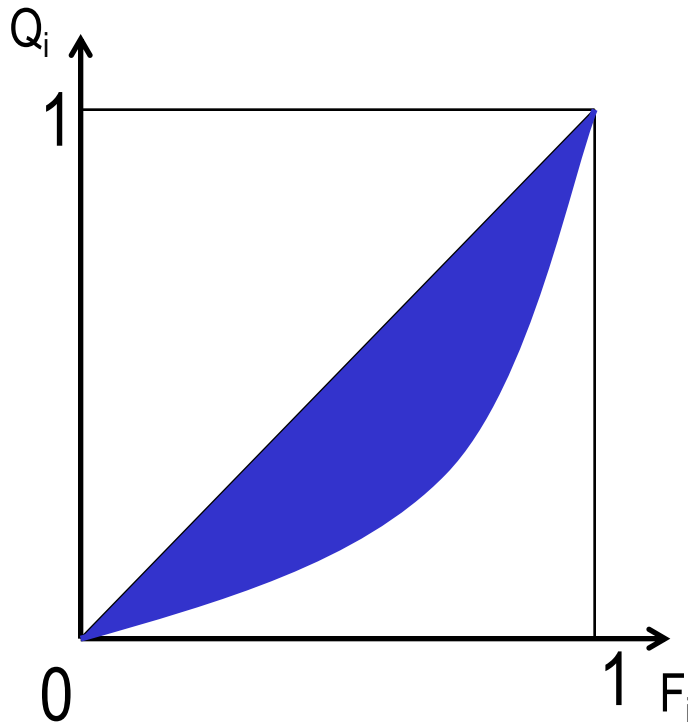
Il grado di concentrazione è tanto maggiore quanto maggiore è la **distanza verticale** di ogni punto dalla **retta di equidistribuzione**

Curva di Lorenz



Unendo tutti i punti (F_i, Q_i) si ottiene la **curva di Lorenz**

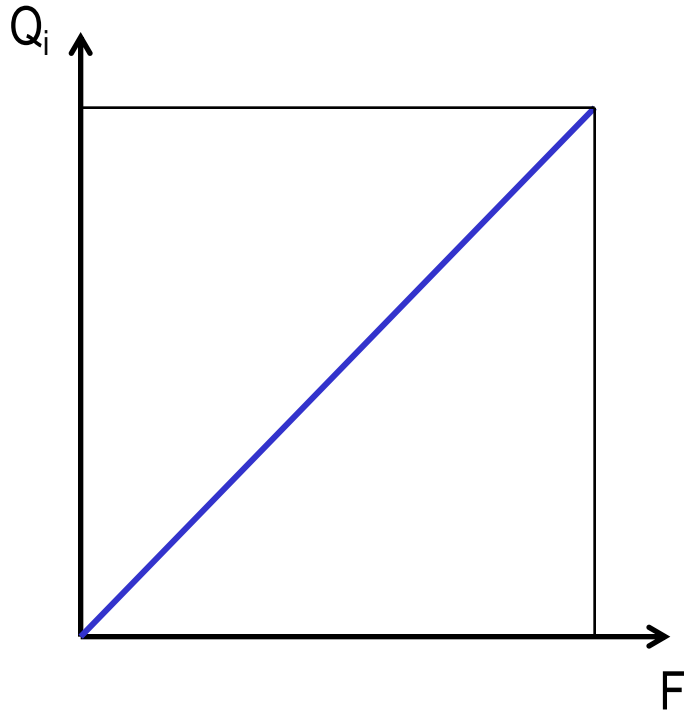
Area di concentrazione



L'area compresa tra la retta di equidistribuzione e la curva di Lorenz si chiama **area di concentrazione**

L'area è tanto più ampia quanto maggiore è il grado di concentrazione

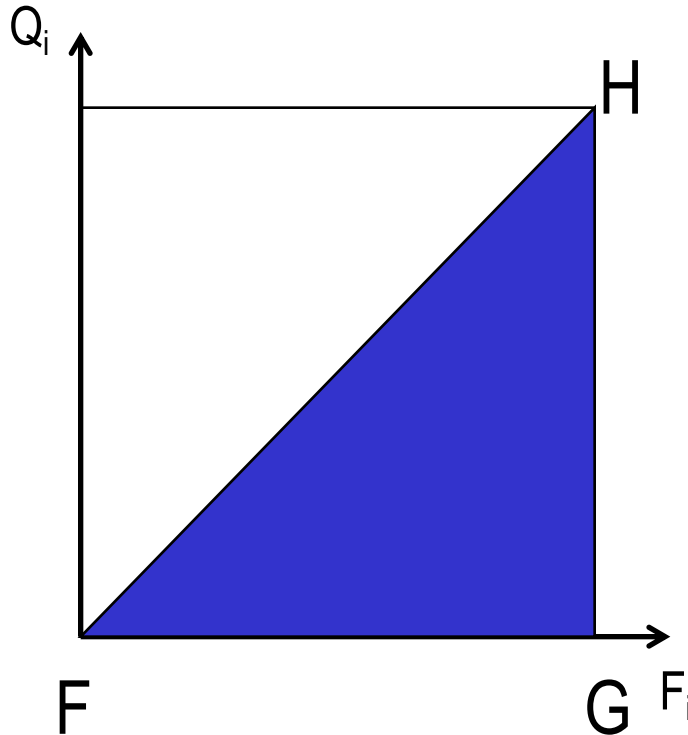
Concentrazione nulla



La **curva di Lorenz** coincide con la retta di equidistribuzione

L'**area di concentrazione** è nulla

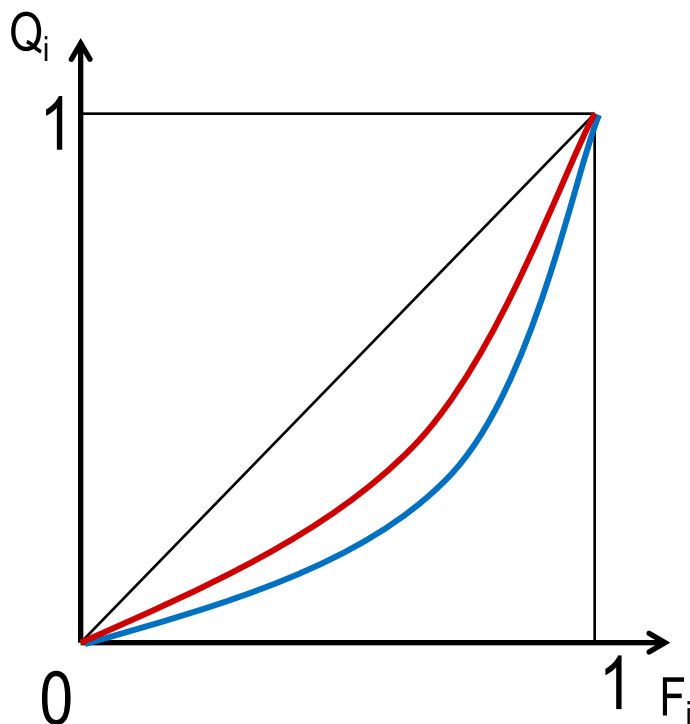
Concentrazione massima



La **curva di Lorenz** coincide con i cateti FG e GH del triangolo rettangolo

L'**area di concentrazione** è massima e pari a $\frac{1}{2}$ (è l'area del triangolo rettangolo)

Confronto tra due curve di Lorenz



La **curva rossa**, più vicina alla retta di equidistribuzione, segnala un grado di concentrazione inferiore a quello corrispondente alla **curva blu**

Concentrazione per una distrib. di freq. con classi di valori

Classi di valori	Frequenze	Intensità globali (conosciute)
$x_j - x_{j+1}$	n_j	X_j
$x_1 - x_2$	n_1	X_1
$x_2 - x_3$	n_2	X_2
...		
$x_j - x_{j+1}$	n_j	X_j
...		
$x_K - x_{K+1}$	n_K	X_K

L'ammontare del carattere detenuto **complessivamente** dalle n_j unità

Ognuna delle n_j unità detiene un ammontare del carattere compreso tra x_j e x_{j+1}

Esempio: Reddito delle famiglie (mgl.€)

Classi di reddito $x_j - x_{j+1}$	Numero di famiglie n_j	Intensità conosciute X_j
fino a 20	3	38
20 - 25	5	115
25 - 30	4	111
30 - 35	2	66
35 - 40	3	117
oltre 40	3	157
	20	604

Le tre famiglie della prima classe hanno complessivamente un reddito pari a 38 mgl. €

Ciascuna di loro ha un reddito non superiore a 20 mgl. €.

Calcolo di F_j

Per una distribuzione di frequenza,

$$F_j = \frac{N_j}{n} \quad j = 1, 2, \dots, K$$

N_j sono le frequenze cumulate

$$F_1 = \frac{N_1}{n}, \quad F_2 = \frac{N_2}{n}, \quad \dots, \quad F_K = \frac{N_K}{n} = 1$$

Calcolo di Q_i

Definendo le intensità assolute cumulate come

$$A_1 = X_1$$

.....

$$A_j = X_1 + X_2 + \dots + X_j$$

.....

$$A_K = X_1 + X_2 + \dots + X_K \quad (\text{ammontare complessivo del carattere})$$

$$Q_j = \frac{A_j}{A_K}, \quad j = 1, 2, \dots, K$$

cioè

$$Q_1 = \frac{A_1}{A_K}, \quad Q_2 = \frac{A_2}{A_K}, \quad \dots, \quad Q_K = \frac{A_K}{A_K} = 1$$

Esempio: Reddito delle famiglie

Classi di reddito $x_j - x_{j+1}$	Numero di famiglie n_j	Intensità conosciute X_j	Freq. cumul. N_j	$F_j = N_j / N$
fino a 20	3	38	3	0,15
20 - 25	5	115	8	0,40
25 - 30	4	111	12	0,60
30 - 35	2	66	14	0,70
35 - 40	3	117	17	0,85
oltre 40	3	157	20	1

20

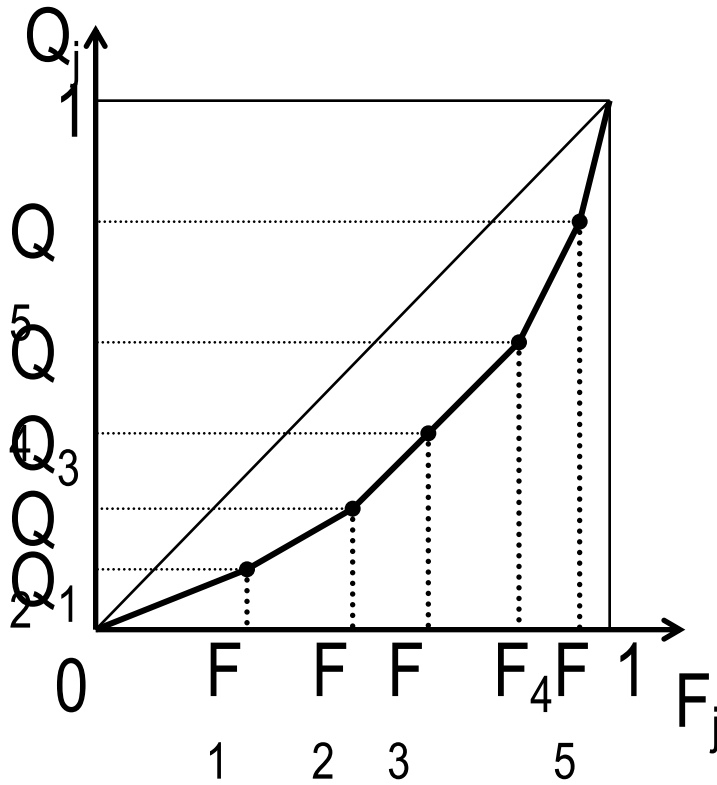
604

Non è possibile visualizzare l'immagine.

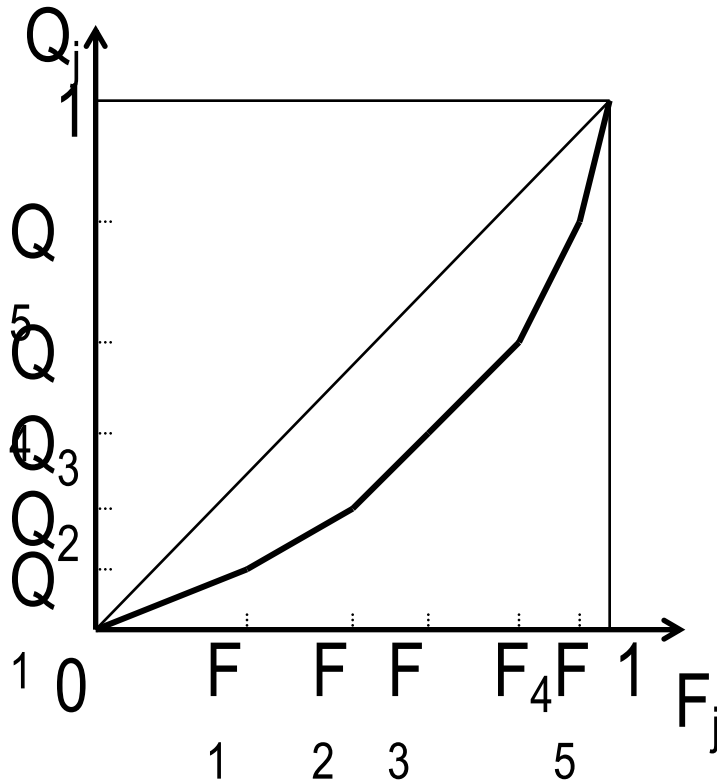
Esempio: Reddito delle famiglie

Classi di reddito $x_j - x_{j+1}$	Numero di famiglie n_j	Intensità conosciute X_j	Freq. cumul. N_j	$F_j =$ N_j/N	Intens. cumul. A_j	$Q_j =$ A_j/A_K
fino a 20	3	38	3	0,15	38	0,06
20 - 25	5	115	8	0,40	153	0,25
25 - 30	4	111	12	0,60	264	0,44
30 - 35	2	66	14	0,70	330	0,55
35 - 40	3	117	17	0,85	447	0,74
oltre 40	3	157	20	1	604	1
	20	604				

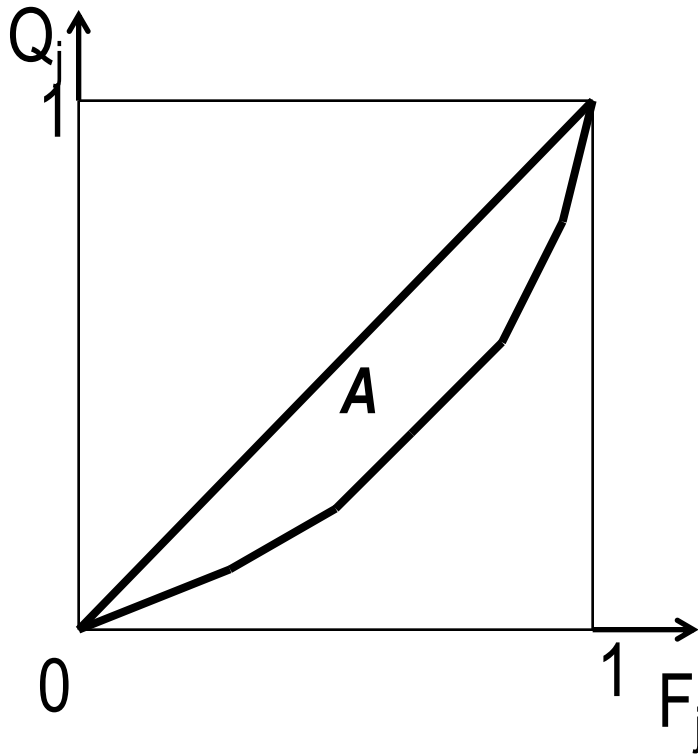
Rappresentazione grafica



Spezzata di concentrazione



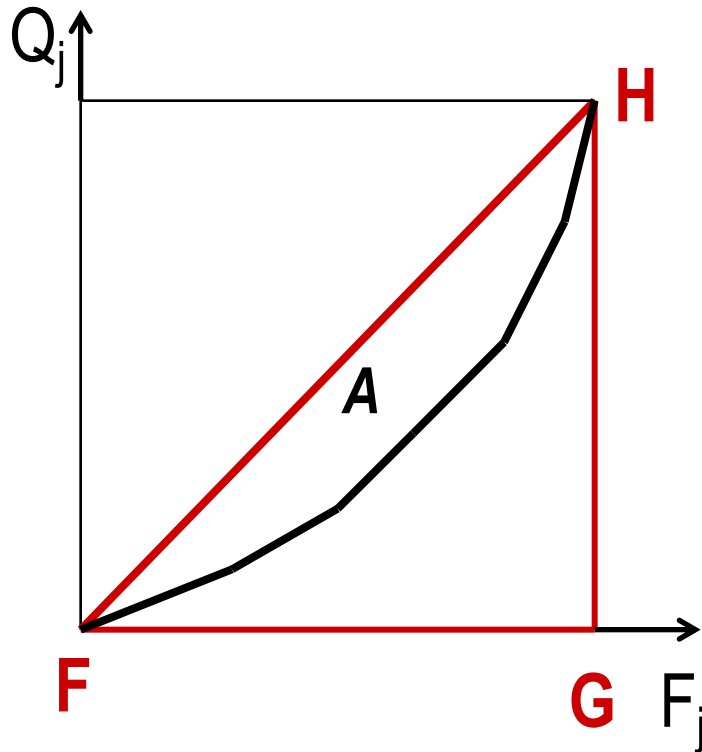
Area di concentrazione



A = Area di concentrazione

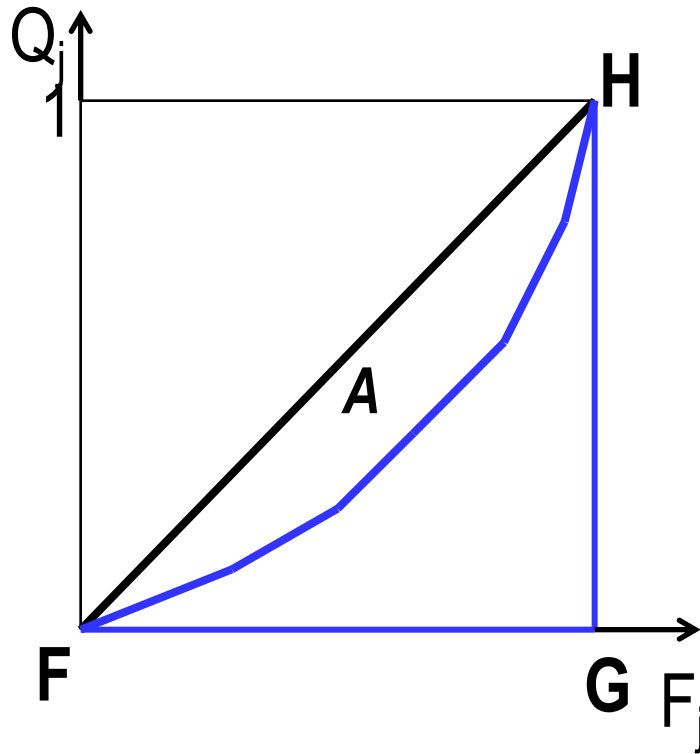
$$0 \leq A \leq 1/2$$

Calcolo dell'area di concentrazione



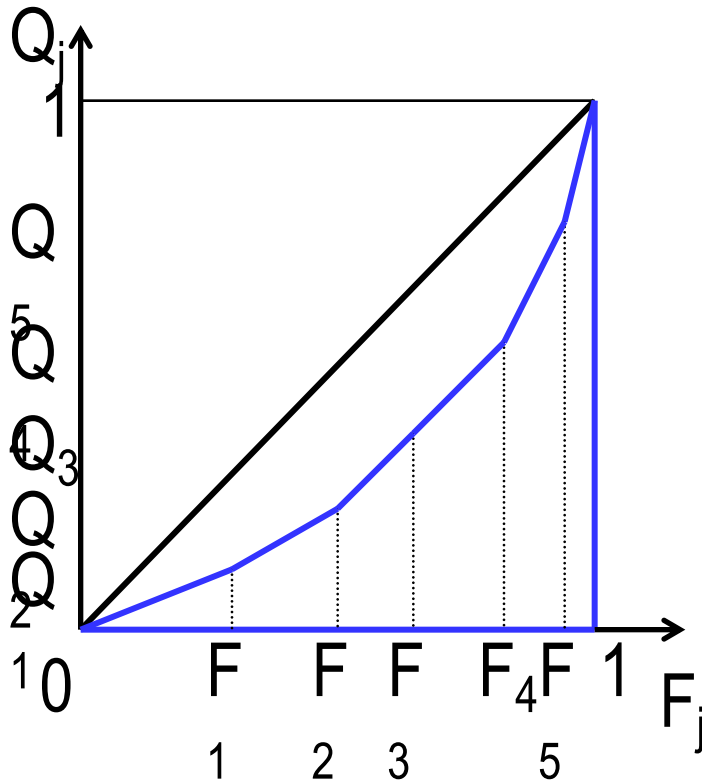
L'area di concentrazione **A** si può ottenere come differenza tra l'area dell'intero triangolo FGH (in rosso) uguale a $1/2 \dots$

Calcolo dell'area di concentrazione



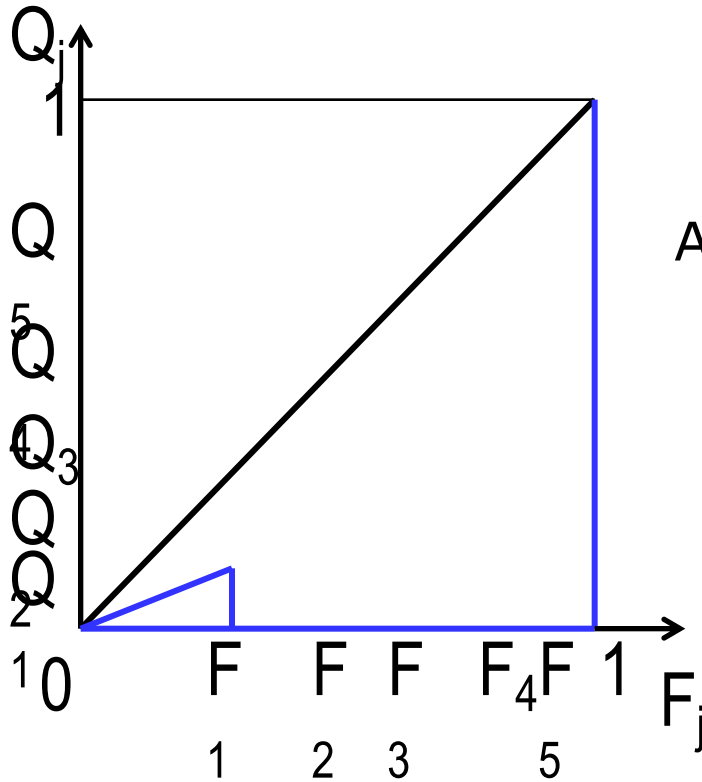
... e l'area della figura in blu

Calcolo dell'area di concentrazione



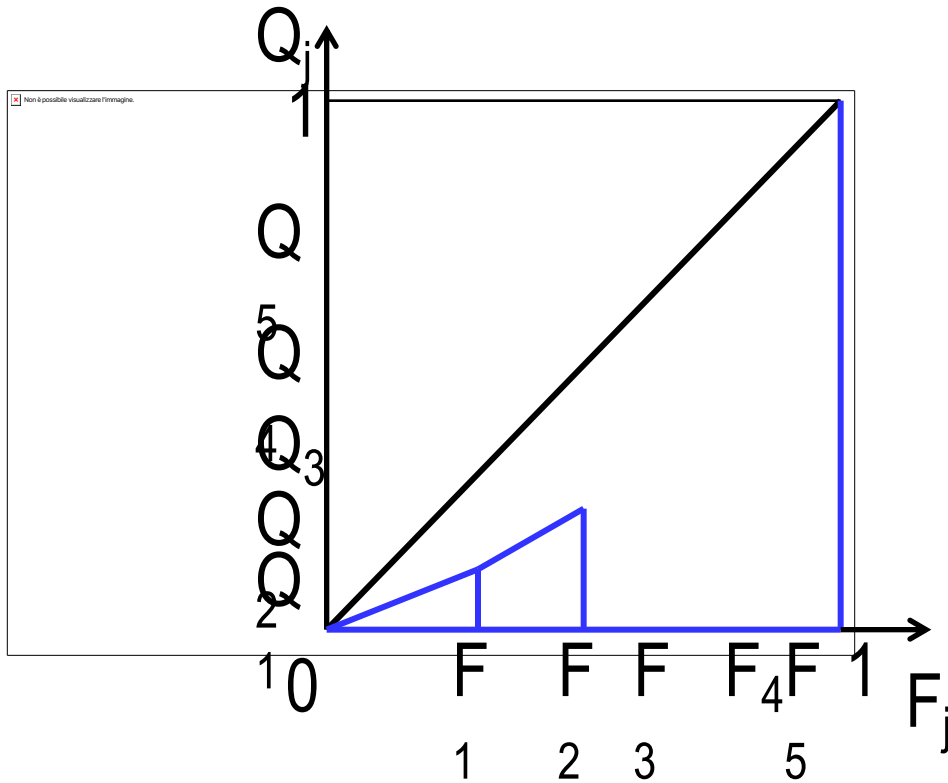
L'area in blu è formata da un triangolo e da trapezi

Calcolo dell'area di concentrazione



Area del triangolo = $\frac{F_1 Q_1}{2}$

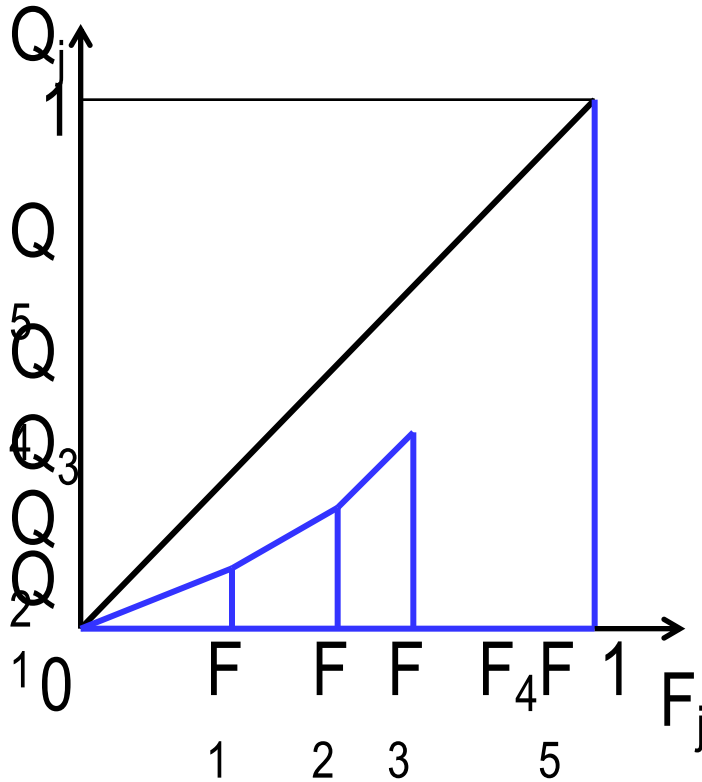
Calcolo dell'area di concentrazione



Area del primo trapezio =

$$= \frac{(Q_2 + Q_1)(F_2 - F_1)}{2}$$

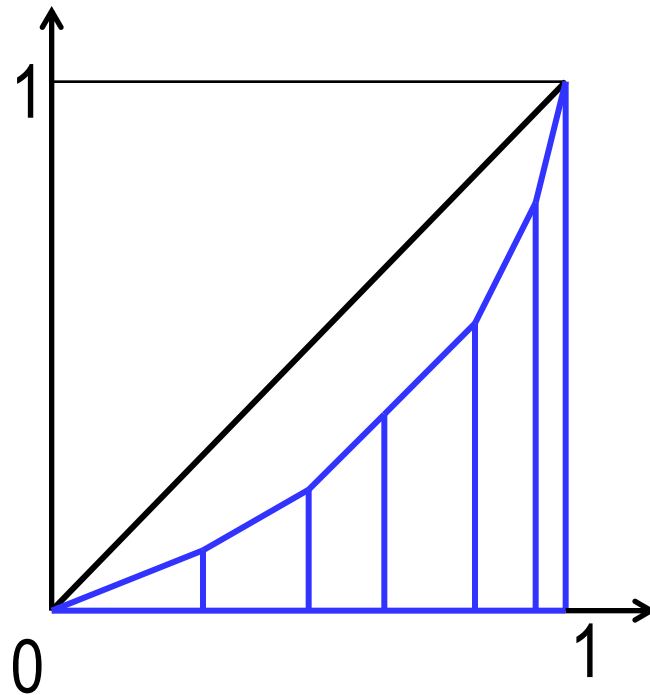
Calcolo dell'area di concentrazione



Area del secondo trapezio =

$$= \frac{(Q_3 + Q_2)(F_3 - F_2)}{2}$$

Calcolo dell'area di concentrazione

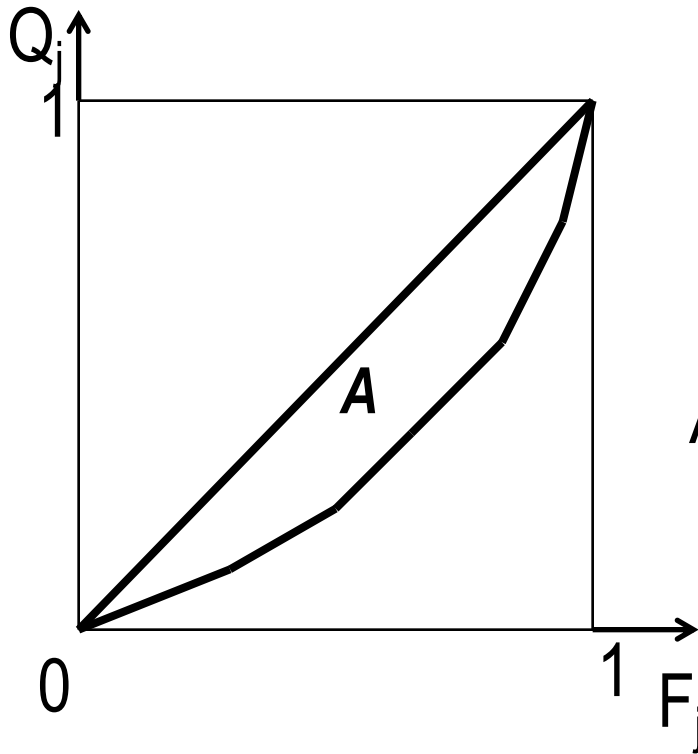


Ponendo $F_0=0$ e $Q_0=0$,
l'area della figura
in blu si può
scrivere come

$$\frac{\sum_{j=0}^{K-1} (F_{j+1} - F_j)(Q_{j+1} + Q_j)}{2}$$

(per $j=0$ si ottiene
l'area del triangolo,
per $j=1$ l'area del
primo trapezio,...)

Area di concentrazione



$$A = \frac{1}{2} - \frac{\sum_{j=0}^{K-1} (F_{j+1} - F_j)(Q_{j+1} + Q_j)}{2}$$

Rapporto di concentrazione

Per avere un indice che varia tra 0 e 1, si considera il doppio dell'area

$$R = 2A = 1 - \sum_{j=0}^{K-1} (F_{j+1} - F_j)(Q_{j+1} + Q_j)$$

$$0 \leq R \leq 1$$

R=0 → equidistribuzione

R=1 → massima concentrazione

Esempio: Reddito delle famiglie

F_j	Q_i	$F_{j+1}-F_j$	$Q_{j+1}+Q_j$	$(F_{j+1}-F_j)(Q_{j+1}+Q_j)$
0,15	0,06	0,15	0,06	0,009
0,40	0,25	0,25	0,31	0,078
0,60	0,44	0,20	0,69	0,138
0,70	0,55	0,10	0,99	0,099
0,85	0,74	0,15	1,29	0,194
1	1	0,15	1,74	0,261

$$\Sigma = 0,779$$

$$R = 1 - 0,779 = 0,221$$

Concentrazione per una distrib. di freq. con classi di valori

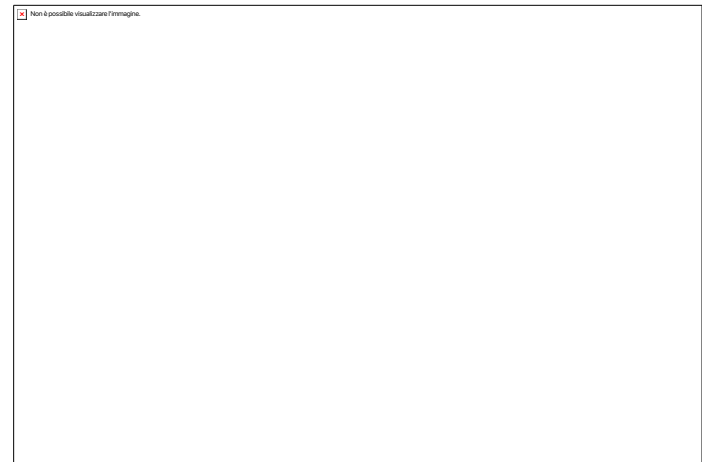
Classi di valori	Frequenze	Intensità globali (stimate)
$x_j - x_{j+1}$	n_j	$x'_j n_j$
$x_1 - x_2$	n_1	$X_1 = x'_1 n_1$
$x_2 - x_3$	n_2	$X_2 = x'_2 n_2$
...		
$x_j - x_{j+1}$	n_j	$X_j = x'_j n_j$
...		
$x_K - x_{K+1}$	n_K	$X_K = x'_K n_K$

$$x'_j = \frac{x_j + x_{j+1}}{2}$$

valore centrale
della classe

Esempio: Aziende per numero di addetti

Classi di addetti $x_j - x_{j+1}$	Numero di aziende n_j	Intensità h $x'_j n_j$
1 - 5	9	27
5 - 10	10	75
10 - 50	18	540
50 - 100	27	2025
100 - 500	2	600
	66	3267



Esempio:

Aziende per numero di addetti

Classi di addetti $x_j - x_{j+1}$	Numero di aziende n_j	Intensità h $x'_j n_j$	Freq. cumul. N_j	$F_j = N_j / n$
1 - 5	9	27	9	0,14
5 - 10	10	75	19	0,29
10 - 50	18	540	37	0,56
50 - 100	27	2025	64	0,97
100 - 500	2	600	66	1

66 3267

Esempio:

Aziende per numero di addetti

Classi di addetti $x_j - x_{j+1}$	Numero di aziende n_j	Intensità h $x'_j n_j$	Freq. cumul. N_j	$F_j =$ N_j/n	Intensità cumul. A_j	$Q_j =$ A_j/A_K
1 - 5	9	27	9	0,14	27	0,01
5 - 10	10	75	19	0,29	102	0,03
10 - 50	18	540	37	0,56	642	0,20
50 - 100	27	2025	64	0,97	2667	0,82
100 - 500	2	600	66	1	3267	1
	66	3267				

Esempio: Aziende per numero di addetti

F_j	Q_j	$F_{j+1}-F_j$	$Q_{j+1}+Q_j$	$(F_{j+1}-F_j)(Q_{j+1}+Q_j)$
0,14	0,01	0,14	0,01	0,001
0,29	0,03	0,15	0,04	0,006
0,56	0,20	0,27	0,23	0,062
0,97	0,82	0,41	1,02	0,418
1	1	0,03	1,82	0,055

$$\Sigma = 0,542$$

$$R = 1 - 0,542 = 0,458$$