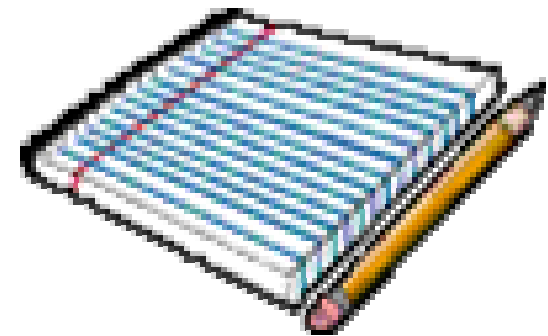
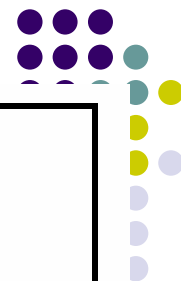


Lezione 5

La Concentrazione





- LA CONCENTRAZIONE:
 - Indice di concentrazione di GINI
 - Concentrazione con il metodo dei TRAPEZI

La concentrazione

Solo per caratteri che godono della “**cumulabilità**” e “**trasferibilità**”. (*ad es. il reddito, la quantità di libri posseduta, le percentuali di voti, ecc...*)

carattere equidistribuito

j	x_h
1°	20
2°	20
3°	20
4°	20
5°	20
totale	100

Concentrazione NULLA

una sola unità possiede
l'intero ammontare

j	x_h
1°	0
2°	0
3°	0
4°	0
5°	100
totale	100

Concentrazione MASSIMA



Ordino le unità statistiche secondo modalità crescente del carattere portato.

$$p_j > q_j \quad (j = 1, \dots, n-1)$$

j	x_h	p_j	q_j
1°	5	1/5	$5/100 = 1/20$
2°	15	2/5	$20/100 = 1/5$
3°	22	3/5	$42/100 = 21/50$
4°	28	4/5	$70/100 = 7/10$
5°	30	$5/5 = 1$	$100/100 = 1$
totale	100		

c'è concentrazione

$$p_j = \frac{j}{n}$$

$$q_j = \frac{Q_j}{Q_n} = \frac{x_1 + \dots + x_j}{x_1 + \dots + x_j + \dots + x_n} = \frac{\sum_{h=1}^j x_h}{\sum_{h=1}^n x_h}$$

Comportamento p_j e q_j

j	X_h	p_j	q_j
1°	20	1/5	$20/100 = 1/5$
2°	20	2/5	$40/100 = 2/5$
3°	20	3/5	$60/100 = 3/5$
4°	20	4/5	$80/100 = 4/5$
5°	20	5/5	$100/100 = 5/5$
totale	100		

Equidistribuzione

$$tutti \ p_j = q_j$$

$$p_n = q_n = 1$$

j	X_h	p_j	q_j
1°	0	1/5	0
2°	0	2/5	0
3°	0	3/5	0
4°	0	4/5	0
5°	100	5/5	$100/100 = 5/5$
totale	100		

Max concentrazione

$$tutti \ q_j = 0$$

tranne ultimo

Indice di concentrazione di GINI

$$R = \frac{\sum_{j=1}^{n-1} (p_j - q_j)}{\sum_{j=1}^{n-1} p_j} \quad 0 \leq R \leq 1$$

$$R = 0$$

equidistribuzione

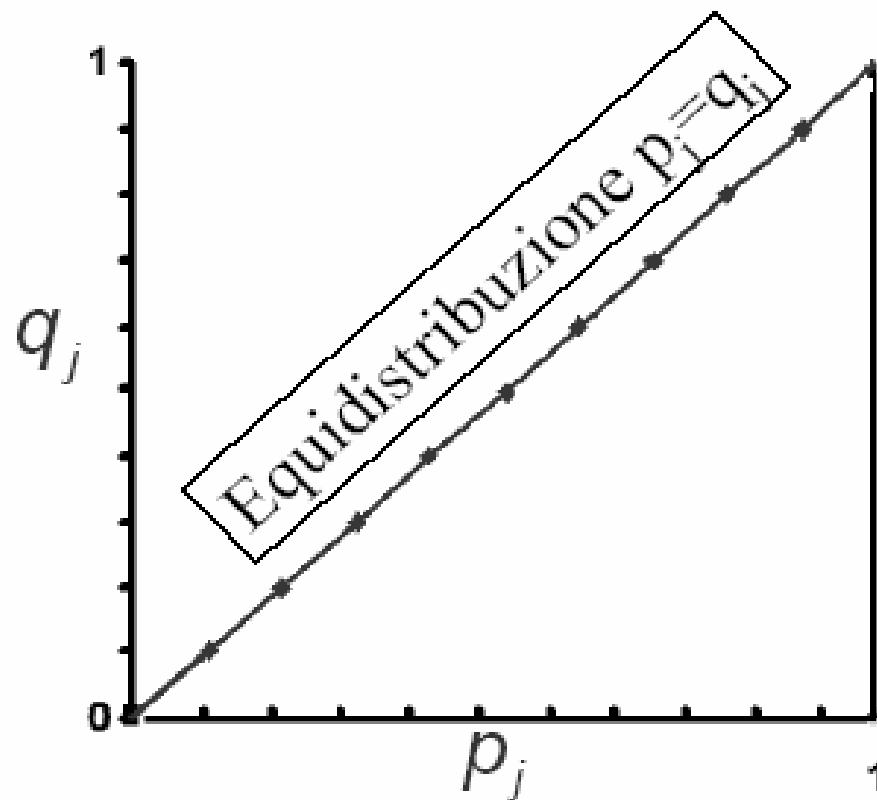
$$R = 1$$

concentrazione massima

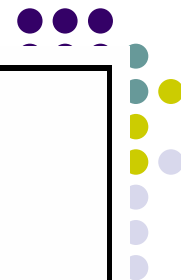
Metodo dei trapezi

Equidistribuzione:

$$p_j = q_j \quad (j = 1, \dots, n)$$



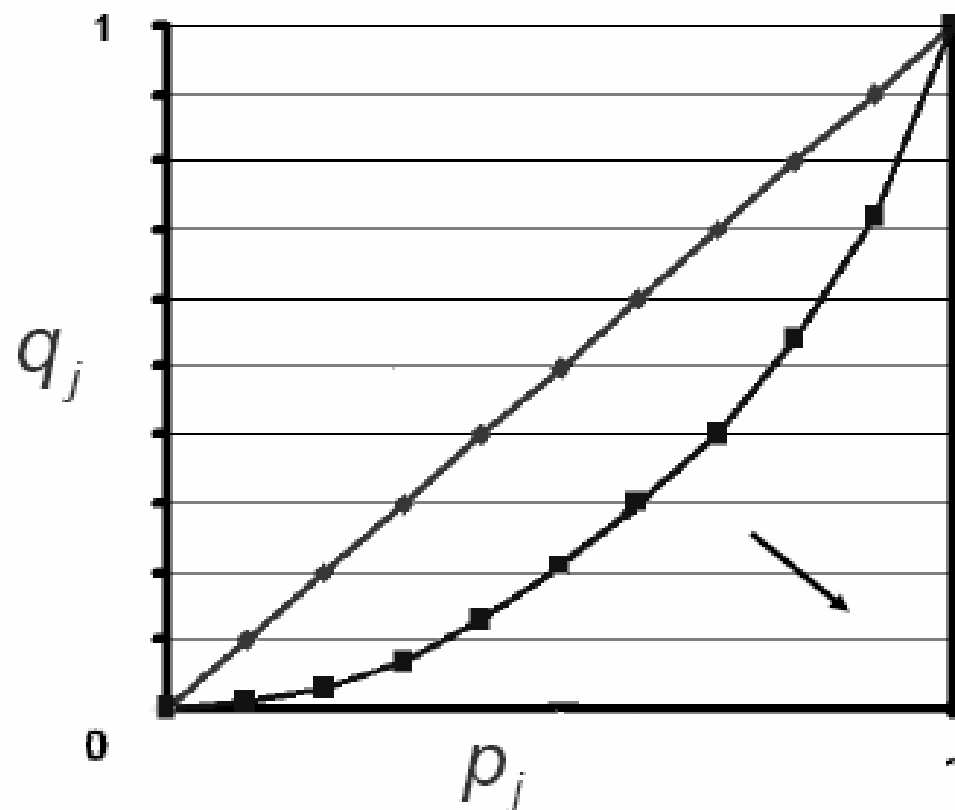
**Bisettrice del
1° quadrante**



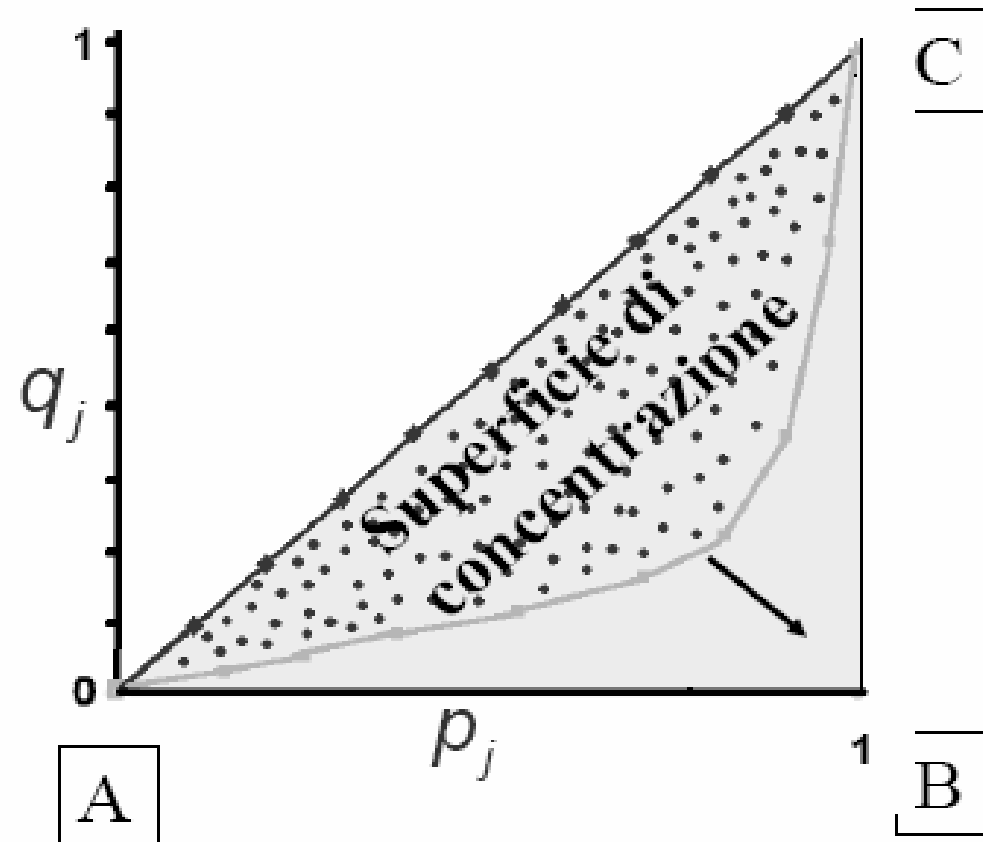
Concentrazione: spezzata congiungente i punti

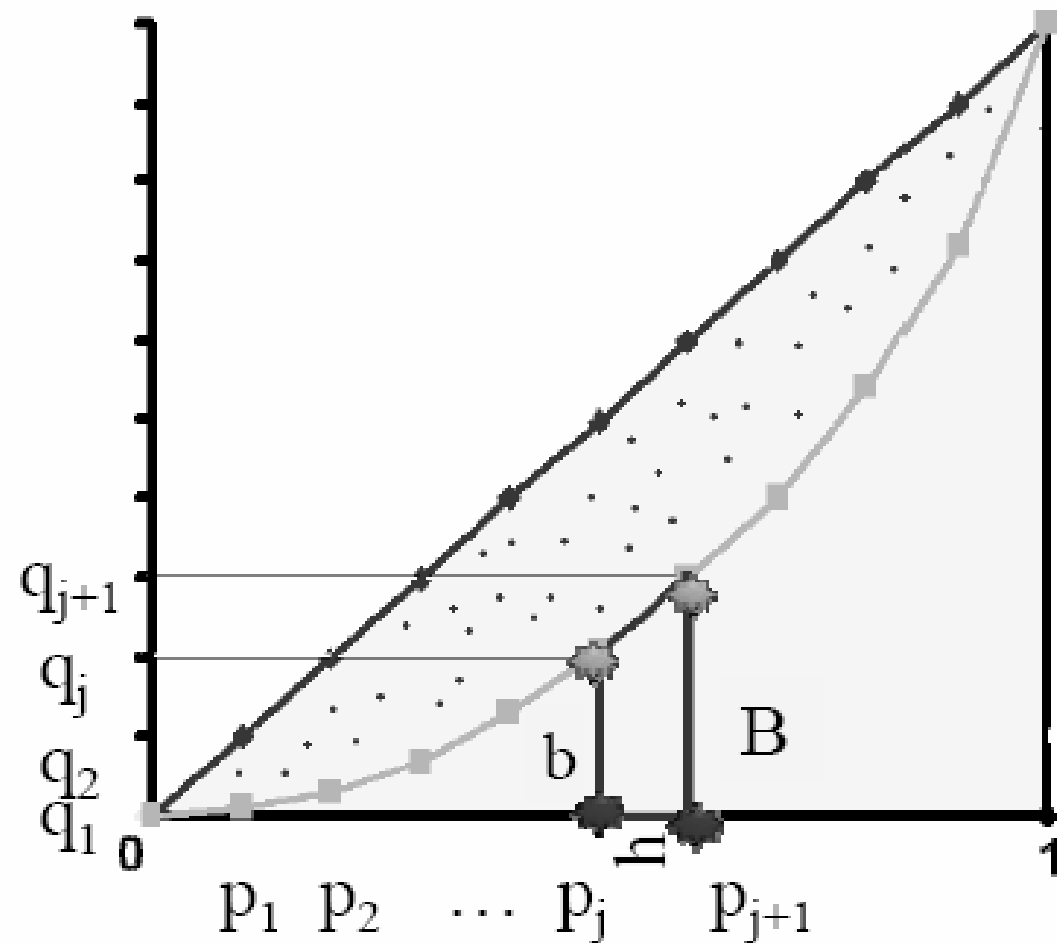
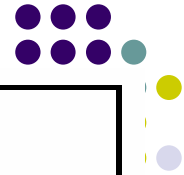
$$p_j > q_j \quad (j = 1, \dots, n)$$

$$(p_j, q_j)$$

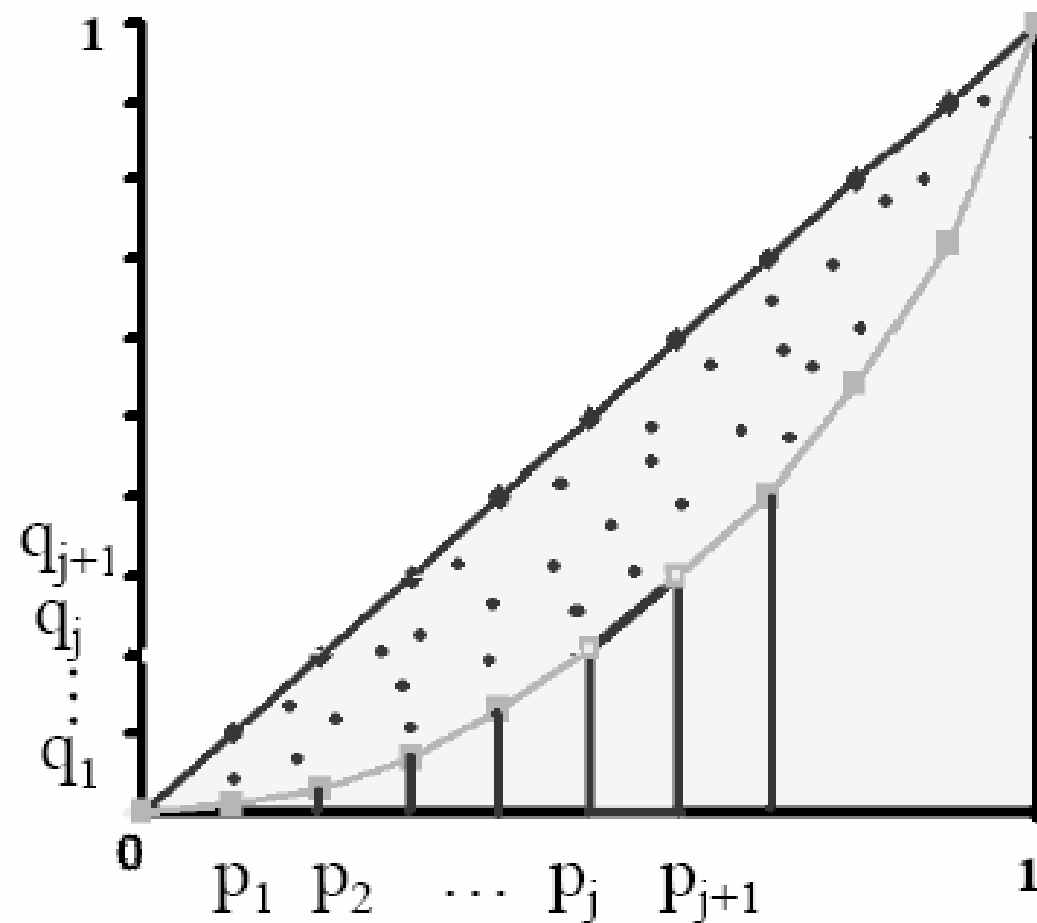
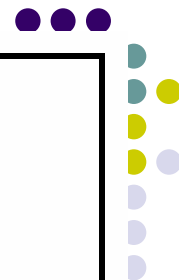


Più c'è concentrazione più la spezzata
tende ad identificarsi con il triangolo ABC





$$\text{Area trapezio} = \frac{(Base + base) \cdot h}{2}$$



$$R = \frac{\text{area triangolo} - \text{area trapezi}}{\text{area triangolo}}$$

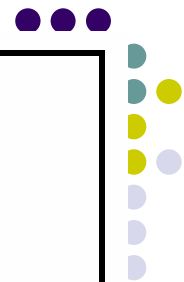


Rapporto di concentrazione:

$$R = \frac{\text{area di concentrazione}}{\text{area di max concentrazione}}$$

area di concentrazione: area della superficie compresa tra la spezzata di concentrazione e il segmento di equidistribuzione

area di massima concentrazione: area del triangolo definito da: bisettrice, ascissa massima, ordinata massima.



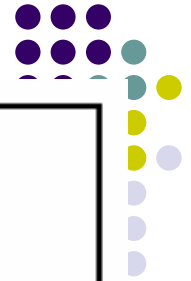
Calcolo di R: metodo dei trapezi

$$R = \frac{\text{area triangolo} - \text{area trapezi}}{\text{area triangolo}}$$

$$R^* = \frac{\frac{1}{2} - \frac{1}{2} \sum_{j=0}^{n-1} \overbrace{(p_{j+1} - p_j)}^h (q_{j+1} + q_j)}^B}{\frac{1}{2}}$$

Max
concentraz

$$R^* = 1 - \sum_{j=0}^{n-1} (p_{j+1} - p_j)(q_{j+1} + q_j)$$



L'indice di concentrazione del Gini "R" e l'indice di concentrazione calcolato con il metodo dei trapezi "R*", sono legati dalle seguenti relazioni:

$$R = \frac{n}{n-1} R^* \quad \Rightarrow \quad R^* = R \frac{n-1}{n}$$

Determinare la concentrazione della % di voti nei partiti in base a R e R*

E' data la seguente distribuzione di voti riportati dai diversi partiti nelle elezioni del Senato nel 1988:

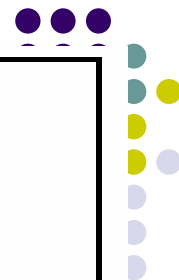
partiti	voti %
DC	33,6
PCI	28,3
PSI	10,9
MSI	6,5
PRI	3,9
PSDI	2,4
PLI	2,2
PR	1,8
VERDI	2
ALTRI	8,4



ordinare

n=10 ossia
il numero di
partiti

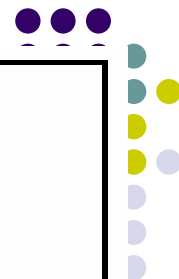
j	X_j
1	1,8
2	2
3	2,2
4	2,4
5	3,9
6	6,5
7	8,4
8	10,9
9	28,3
10	33,6



n=10

j	X _j	p _j =j/n	Q _j	q _j =Q _j /Q _n	(p _j -q _j)
1	1,8	0,1	1,8	0,018	0,082
2	2	0,2	3,8	0,038	0,162
3	2,2	0,3	6	0,060	0,24
4	2,4	0,4	8,4	0,084	0,316
5	3,9	0,5	12,3	0,123	0,377
6	6,5	0,6	18,8	0,188	0,412
7	8,4	0,7	27,2	0,272	0,428
8	10,9	0,8	38,1	0,381	0,419
9	28,3	0,9	66,4	0,664	0,236
10	33,6		100	1,000	
		4,5			2,672

$$R = \frac{\sum_{j=1}^9 (p_j - q_j)}{\sum_{j=1}^9 p_j} = \frac{2,672}{4,5} = 0,5938$$

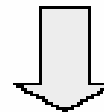


j	X _j	p _j =j/n	Q _j	q _j =Q _j /Q _n	p _{j+1} -p _j	q _{j+1} +q _j	prod
0					0,1	0,0180	0,0018
1	1,8	0,1	1,8	0,0180	0,1	0,0560	0,0056
2	2	0,2	3,8	0,0380	0,1	0,0980	0,0098
3	2,2	0,3	6	0,0600	0,1	0,1440	0,0144
4	2,4	0,4	8,4	0,0840	0,1	0,2070	0,0207
5	3,9	0,5	12,3	0,1230	0,1	0,3110	0,0311
6	6,5	0,6	18,8	0,1880	0,1	0,4600	0,0460
7	8,4	0,7	27,2	0,2720	0,1	0,6530	0,0653
8	10,9	0,8	38,1	0,3810	0,1	1,0450	0,1045
9	28,3	0,9	66,4	0,6640	0,1	1,6640	0,1664
10	33,6	1	100	1,0000			
		4,5					0,4656

$$R^* = 1 - \sum_{j=0}^9 (p_{j+1} - p_j)(q_{j+1} + q_j) = 1 - 0,4656 = 0,534$$

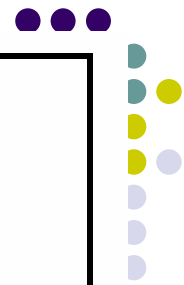
$$R = \frac{\sum_{j=1}^9 (p_j - q_j)}{\sum_{j=1}^9 p_j} = 0,5938$$

$$R^* = 1 - \sum_{j=0}^9 (p_{j+1} - p_j)(q_{j+1} + q_j) = 0,534$$



$$R = \frac{n}{n-1} R^*$$

$$R = \frac{10}{9} 0,534 = 0,5938$$

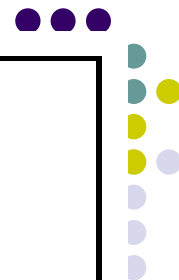


Distribuzione di frequenza per intervalli.

Si consideri la seguente classificazione di 1.000 individui secondo il reddito anno dichiarato (in milioni di lire):

Classi di reddito (milioni di lire)	Numero individui (n_i)
0 – 3	327
3 – 9	558
9 – 25	68
25 – 50	32
50 – 100	15
TOTALE	1.000

In quale proporzione la distribuzione dei redditi nell'insieme considerato realizza quella ipotetica di massima concentrazione?



$$N_i = \sum_{h=1}^i n_h$$

$$Q_i = \sum_{h=1}^i \hat{x}_h n_h$$

Classi di reddito	n_i	Valori centrali \hat{X}_i	$\hat{X}_i \cdot n_i$	N_i	Q_i	$p'_i = N_i/N$	$q'_i = Q_i/Q_N$	$p'_{i+1} - p'_i$	$q'_{i+1} + q'_i$	prod
	0	0	0	0	0	0,000	0,000	0,327	0,067	0,02191
0 - 3	327	1,5	491	327	491	0,327	0,067	0,558	0,591	0,32978
3 - 9	558	6	3348	885	3839	0,885	0,524	0,068	1,206	0,08201
9 - 25	68	17	1156	953	4995	0,953	0,682	0,032	1,528	0,04890
25 - 50	32	37,5	1200	985	6195	0,985	0,846	0,015	1,846	0,02769
50 - 100	15	75	1125	1000	7320	1,000	1,000	-----	-----	-----
TOTALE	1.000									0,51029

METODO DEI TRAPEZI $R^* = 1 - \sum_{i=0}^4 (p_{i+1} - p_i)(q_{i+1} + q_i) = 1 - 0,5103 = \mathbf{0,4897}$

pertanto in tale distribuzione si verifica il 48,9% della massima concentrazione possibile.