Giancarlo Bettuzzi

Considerazioni attorno al rapporto quadratico di concentrazione

Serie Ricerche 1994, n.1

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università degli studi di Bologna

Finito di stampare nel mese di Marzo 1994
presso le Officine Grafiche Tecnoprint
Via del Legatore 3, Bologna

1. Premesse

E' ben noto che gli studi sulla concentrazione risalgono a Vilfredo Pareto, e questo accenno alle remote origini storiche del problema è del tutto sufficiente per comprendere le difficoltà che si incontrerebbero qualora si tentasse di sviluppare, in questa sede, una disamina storica dei contributi apportati e, ancor più, un'analisi critica delle posizioni concettuali assunte nel corso del tempo in merito alla nozione stessa di concentrazione. Su questi temi, ed altri ancora, ci proponiamo di soffermarci adeguatamente in un lavoro di carattere monografico ove non si è costretti a contenere il discorso in un numero esiguo di pagine. E' certo, comunque, che qualunque siano le scelte metodologiche in proposito non si può prescindere dall'impostazione giniana ed, in particolare, dalla definizione del rapporto di concentrazione.

La costruzione paretiana, per addivenire alla misura della concentrazione, prevede la definizione di un modello teorico in cui vengono considerati modalità di reddito e numero di redditieri aventi reddito superiore o uguale a tali modalità. Fermo restando il valore storico dell'indicazione di Pareto, la critica del Gini appare fondata sull'argomentazione secondo cui ad entità reali, quali sono i numeri di redditieri con reddito superiore a un certo livello, si pongono a raffronto valori astratti quali appunto i livelli di reddito, e non invece entità concrete definite dall'ammontare del reddito spettante a tale numero di redditieri. Dall'impostazione giniana appare chiaramente che un elemento essenziale del ragionamento è quello dell'accumulo dei redditi, oltreché dei redditieri, il che significa affermare in generale che i caratteri per i quali ha senso parlare di misura della concentrazione devono essere trasferibili.

Gli sviluppi susseguitisi nel tempo a tale impostazione sono notissimi, numerosi e, in genere, rivolti ad un'approfondita analisi dell'impostazione data da Corrado Gini; al riguardo è doveroso ricordare alcuni nomi emblematici quali Pietra, Castellano, Bonferroni, De Vergottini, Amato e Fortunati che conferisce un pregnante significato logico alle cosiddette medie parziali che intervengono nella struttura del rapporto di concentrazione.

In verità, la possibilità di introdurre medie parziali nella costruzione di un indice di concentrazione era già stata indicata dallo stesso Gini, ma a noi preme in questa sede ricordare come tale indicazione metodologica sia stata recepita anche dal Bonferroni (1940) quando, nell'esporre un procedimento generale atto a dare indici di concentrazione, parla espressamente di media parziale "delle prime i misure disposte in ordine crescente"; il che riflette una precisa coincidenza di vedute con la impostazione di Gini e di Pietra che vuole lo schema di graduazione alla base del fondamento logico della nozione di concentrazione. E' ben noto, peraltro, che vi sono altri indirizzi di pensiero che danno della concentrazione una diversa definizione e quindi viene sviluppato un differente approccio metodologico talvolta non fondato sulla accumulazione. Ma il lavoro che intendiamo sviluppare in questa sede non nasce da un'intenzione di rassegna, possibilmente sistematica, delle varie posizioni quanto invece da alcune connotazioni dello scritto del Bonferroni del 1940 che ne riflettono un significato veramente emblematico risultante, a nostro sommesso avviso, da un non uniforme sviluppo della trattazione, nelle sue singole parti, certamente dovuta all'esigenza di estremo contenimento dell'esposizione che è una espressione tipica delle regole sempre adottate al riguardo delle comunicazioni presentate in occasione dei Congressi scientifici. Specificatamente, l'articolo del Bonferroni contiene, come già osservato, un preciso e chiaro riferimento alla determinazione degli accumuli e, quindi, delle medie parziali che però resta soltanto una indicazione di metodo quando poi concretamente passa alla costruzione del rapporto quadratico di concentrazione. Ci sembra allora legittimo e doveroso, anche in considerazione del valore intrinseco del contributo del Bonferroni, compiere una disamina per accertare se il rapporto quadratico di concentrazione è tale, cioè se è un rapporto di concentrazione nel senso giniano, o se, invece, il rapporto del Bonferroni debba soltanto riguardarsi come una premessa ed uno stimolo intellettuale verso una più appropriata definizione di rapporto quadratico di concentrazione.

Nelle pagine che seguono, e nello spirito delle indicazioni chiare ed esplicite degli studiosi ricordati, ci proponiamo di dare un coerente approccio alla costruzione ed alla interpretazione metodologica di un rapporto quadratico di concentrazione, per poi riprendere nel suo ambito il contributo del Bonferroni più volte richiamato.

2. Rapporto di concentrazione del quadrato delle intensità

Nello studio delle misure di variabilità il problema generalmente viene affrontato mediante il confronto per differenza fra coppie di intensità, di cui una eventualmente può essere costante, e nella conseguente e conveniente scelta di una media per sintetizzare tali confronti. Si hanno, così, come è noto, misure semplici di variabilità se la media prescelta è quella aritmetica e misure quadratiche se la media in questione è quella quadratica. Queste osservazioni sono ovvie, e forse superflue anche in un manuale, ma a noi sembra che non sia privo di utilità richiamarle in questa sede perché nella costruzione di misure quadratiche di concentrazione il problema va posto, o comunque può essere posto, in maniera differente. Una misura di concentrazione, come ogni misura, è un confronto tra grandezze. Ci si deve allora domandare se nella costruzione di un indice quadratico di concentrazione si debba partire da confronti per differenza tra grandezze per poi elevarli al quadrato, o sia invece opportuno costruire quadrati di grandezze per porli convenientemente a confronto, e la questione non trova certo un elemento di semplicità nel fatto che la costruzione del rapporto di concentrazione richiede la definizione degli accumuli delle grandezze interessate. La prima via esperibile, cioè quella di costruire accumuli per poi elevarne al quadrato i confronti, non è certamente a priori da rifiutare e, semmai, la scelta di non proseguire l'indagine in tal senso è soprattutto dettata dalle difficoltà che insorgono negli sviluppi formali. La impostazione che quindi noi adottiamo è quella di costruire i quadrati delle singole intensità per poi riproporre su questi le tappe fondamentali della elaborazione giniana per pervenire ad un rapporto di concentrazione. La scelta non è di comodo ed ha un significato concettuale che può avere implicazioni specifiche nei confronti e nelle misure riferite a situazioni concrete proprio perché la differenza tra quadrati caratterizza in generale una diversità sostanziale che invece viene attenuata dal quadrato del confronto per differenza. Questo tipo di scelta, peraltro, è in linea con uno spunto metodologico contenuto in un lavoro del Ferreri (1978) in cui l'Autore, allo scopo di favorire l'affinamento di strumenti di indagine sulla distribuzione dei redditi coerentemente ai principi della teoria economica del benessere, suggerisce una estensione della curva di

Lorenz connessa alla distribuzione del reddito equivalente inteso in termini di utilità e, con ciò, una estensione del rapporto di concentrazione al fine della misura della concentrazione del benessere economico. In particolare, Ferreri introduce un diagramma di concentrazione di ordine r, limitatamente ai valori $0<r<1$, e con tale posizione configura il corrispondente rapporto di concentrazione di ordine r ; tale scelta è logicamente connessa alla specificità della ricerca da lui condotta che richiede, in coerenza ai principi della teoria economica, l'assunzione di funzioni di utilità individuali crescenti, con utilità marginale decrescente, al crescere del reddito.

Per quanto attiene l'elaborazione logico-formale delle premesse che noi poniamo è vero che, perlomeno nella fase iniziale, questa ripropone lo schema logico costruito dal Gini, ma è anche vero che nella articolazione degli sviluppi si ottengono risultati che non sempre trovano un parallelo con la metodologia dell'indice semplice di concentrazione e che presentano, in sè e per sè, un preciso interesse e significato. In particolare, vedremo la relazione che sussiste tra il rapporto di concentrazione da noi proposto e la differenza quadratica media; relazione, questa, che può contribuire ad evitare ambiguità di concetti tra variabilità e concentrazione come, invece, storicamente è accaduto considerando misure fondate sulla media aritmetica, e ciò stante la notissima relazione tra rapporto di concentrazione e differenza semplice media.

Ciò premesso, gli strumenti logico-formali sui quali cercare di costruire un rapporto quadratico di concentrazione sono quelli che qui di seguito sinteticamente introduciamo.

Consideriamo un collettivo di N unità ed indichiamo genericamente con x_i l'intensità afferente all' *i-esima* unità, assumendo i tutti i valori da 1 a N. Euristicamente una ipotesi siffatta è in genere priva di significato perché equivale a dire che le intensità rilevate non sono state classificate. L'ipotesi però risulta notoriamente assai comoda dal punto di vista delle elaborazioni logico-formali che si possono, o si devono, sviluppare. Restano quindi da introdurre, una volta pervenuti a risultati concettualmente soddisfacenti, alcune altre elaborazioni per dare ai risultati stessi una forma congrua anche dal punto di vista operativo, e cioè facendo riferimento a distribuzioni statistiche. Per limiti di spazio di tale aspetto non ci occupiamo nel presente lavoro.

Ritornando allora alle intensità afferenti alle singole unità del collettivo, supponiamo che queste siano graduate in ordine non decrescente, cioè che risulti

$$x_1 \leq x_2 \leq \ldots \leq x_N, \qquad [1]$$

per impostare, come scelta iniziale, ogni sorta di argomentazione e di elaborazione sui quadrati di tali intensità. Si ha allora:

$$x_1^2 \leq x_2^2 \leq \ldots \leq x_N^2 \qquad [2]$$

Va da sè che molte delle annotazioni che dovremo introdurre ripetono lo schema giniano e per suffragare tale affermazione è sufficiente osservare che nel proseguo lo strumento di sintesi sarà la media aritmetica dei quadrati delle intensità. Ragioni di opportunità in riferimento ai dati statistici inizialmente considerati consigliano poi di annotare tale media aritmetica come quadrato della media quadratica. Allora, iniziamo con l'osservare che è

$$\frac{\sum_1^i {}_h\, x_h^2}{i} \leq \frac{\sum_1^N {}_h\, x_h^2}{N} \qquad [3]$$

e che, escludendo l'ipotesi di equidistribuzione, la diseguaglianza è stretta per ogni i minore di N. In verità, nell'attuale contesto occorre un chiarimento sul termine equidistribuzione. Allora, giusta l'osservazione concettuale secondo cui ogni ragionamento è impostato in termini di media aritmetica dei quadrati, ricorreremo ancora alla nozione di media aritmetica per caratterizzare la equidistribuzione: va da sè che l'equidistribuzione si realizza, in questo contesto, quando ad ogni unità del collettivo è assegnata una quantità corrispondente al quadrato della media quadratica della distribuzione originaria, alla quale, in ogni caso, non occorre fare alcun esplicito riferimento.
Alla diseguaglianza [3] conviene dare la seguente forma

$$\frac{\sum_1^i {}_h x_h^2}{\sum_1^N {}_h x_h^2} \le \frac{i}{N} \qquad [4]$$

per introdurre in maniera classica i seguenti simboli

$${}_{(2)}q_i = \frac{\sum_1^i {}_h x_h^2}{\sum_1^N {}_h x_h^2} = \frac{i\, M_{2,i}^2}{N M_2^2} \qquad [5]$$

$$p_i = \frac{i}{N} \qquad [6]$$

avendo indicato con $M_{2,i}^{2}$ la media aritmetica delle prime i quantità della graduatoria [2]. Tenuto conto delle notazioni [5] e [6] è possibile riscrivere la [4] nella forma

$${}_{(2)}q_i \le p_i \qquad [7]$$

da cui segue

$$p_i - {}_{(2)}q_i \ge 0. \qquad [8]$$

Il valore nullo della [8] si ottiene soltanto nell'ipotesi di equidistribuzione della quantità $\sum_1^N {}_h x_h^2$ tra le N unità, cioè nella ipotesi in cui ad ogni unità sia assegnata come intensità la media aritmetica dei quadrati delle intensità. Le diseguaglianze indicate nella [8] devono essere considerate in numero di *N-1* poiché per *i=N* la [8] è identicamente nulla. Nella ipotesi massimante alle prime *N-1* unità della graduatoria compete intensità *0* ed alla *N-esima* il valore definito da N volte la media aritmetica dei quadrati delle intensità; quindi, in tale ipotesi, il valore della [8] è p_i.

Come già nell'impostazione giniana anche per pervenire al rapporto quadratico di concentrazione conviene prendere le mosse dal quoziente tra la [8] ed il valore massimo che questa può assumere, ponendo

$${}_{(2)}R_i = \frac{p_i - {}_{(2)}q_i}{p_i} \qquad [9]$$

Il valore che la [9] può realizzare è certamente compreso nell'intervallo [*0,1*] e per le proprietà possedute è atto a misurare la concentrazione rispetto al subcollettivo formato dalle prime i unità statistiche che possiamo denominare "più povere". Ponendo mente agli *N-1* modi di suddividere in due subcollettivi complementari il collettivo costituito da N unità statistiche e, quindi, alle corrispondenti espressioni del rapporto ${}_{(2)}R_i$ (con *i=1,2,...,N-1*), è del tutto immediato scrivere l'espressione della media aritmetica degli ${}_{(2)}R_i$ stessi, ponderati con le frequenze relative, e risulta:

$${}_{(2)}R = \frac{\sum_1^{N-1} {}_i\; {}_{(2)}R_i\, p_i}{\sum_1^{N-1} {}_i\; p_i} \qquad [10]$$

da cui

$$_{(2)}R = \frac{\sum_{1}^{N-1}{}_i \left(p_i - {}_{(2)}q_i\right)}{\sum_{1}^{N-1}{}_i \; p_i} \qquad [11]$$

Mediante sviluppi elementari, e recependo le indicazioni del Fortunati volte ad interpretare il rapporto di concentrazione come confronto tra valori medi, è agevole pervenire alla connotazione seguente:

$$_{(2)}R = \frac{M_2^2 - {}_2M^2}{M_2^2} \qquad [12]$$

in cui, appunto, $_{(2)}R$ risulta espresso in funzione di due medie: M_2^2, che è la media aritmetica dei quadrati delle intensità, e $_2M^2$, che è la media aritmetica delle medie aritmetiche parziali $M_{2,i}^2$ di tali quadrati. Per ottenere la [*12*] è sufficiente prendere l'avvio dalla [*11*] e, tenuto conto del terzo membro della [*5*] e del secondo membro della [*6*], è agevole pervenire all'espressione:

$$_{(2)}R = 1 - \frac{\dfrac{\sum_{1}^{N-1}{}_i \; i M_{2,i}^2}{\sum_{1}^{N-1}{}_i \; i}}{M_2^2} \qquad [13]$$

nella quale ponendo

$$_2M^2 = \frac{\sum_{1}^{N-1}{}_i \; i M_{2,i}^2}{\sum_{1}^{N-1}{}_i \; i} \qquad [14]$$

si giunge, appunto, alla [*12*].

3. Relazioni tra $_{(2)}R$ e le misure di variabilità dei quadrati delle intensità

Va da sè che rispetto alla graduatoria dei quadrati delle intensità si può sviluppare una metodologia delle misure di variabilità tutta equivalente a quella ben nota a proposito della graduatoria delle intensità stesse e l'equivalenza è tale da rendere per molti aspetti superfluo ogni indugio sull'argomento. Tuttavia, per un motivo che subito vedremo, conviene qui richiamare le tappe essenziali di una possibile trattazione al riguardo. Così, allora, la espressione generale dello scostamento semplice medio dei quadrati da una intensità x_k prescelta nella [*2*] è così definibile:

$$_{(2)}S_{x_k} \equiv S_{x_k^2}' = \frac{\sum_{1}^{N}{}_h \left|x_h^2 - x_k^2\right|}{N}, \qquad [15]$$

e ciò consente di introdurre la nozione di scostamento semplice medio senza ripetizione dei quadrati delle intensità (Gili 1985) ponendo

$$ {}_{(2)}S'_{x_k} = \frac{\sum_{1}^{N}{}_{h\neq k} \left| x_h^2 - x_k^2 \right|}{N-1} \qquad [16] $$

Analogamente, la differenza semplice media resta definita da

$$ {}_{(2)}\Delta = \frac{\sum_{1}^{N}{}_{h} \sum_{1}^{N}{}_{k\neq h} \left| x_k^2 - x_h^2 \right|}{N(N-1)} \qquad [17] $$

e sussiste, ovviamente, la relazione:

$$ {}_{(2)}\Delta = \frac{\sum_{1}^{N}{}_{k} \; {}_{(2)}S'_{x_k}}{N} \qquad [18] $$

Le argomentazioni sviluppate nel paragrafo *2* e le annotazioni qui esposte, consentono di affermare che sussiste la seguente eguaglianza

$$ {}_{(2)}R = \frac{{}_{(2)}\Delta}{2M_2^2} \qquad [19] $$

la quale anche appare scontata stanti le premesse poste. Ma sulla [*19*] subito ritorneremo; qui preme soltanto osservare che il denominatore della [*19*] rappresenta il valore che assume il numeratore nella ipotesi massimante, così come è stata configurata nel presente lavoro.

4. L'indice quadratico di concentrazione del Bonferroni e sue relazioni col rapporto di concentrazione del quadrato delle intensità

Nello stesso lavoro in cui, come già accennato, introduce nozioni e concetti di carattere generale intorno alla teoria della concentrazione Bonferroni definisce formalmente il rapporto quadratico di concentrazione come quoziente tra il quadrato della differenza quadratica media ed il valore che questa assume in riferimento ad uno schema massimante teorico. Pone, cioè, per definizione

$$ R_2 = \frac{{}^{2}\Delta^2}{2M_2^2} \qquad [20] $$

Tra l'espressione [*20*] e quella che stabilisce il nesso tra il rapporto di concentrazione costruito da Gini e la differenza semplice media esiste una precisa diversità di carattere concettuale. E' noto infatti, che l'eguaglianza $R = \Delta / 2M_1$ è il risultato di una dimostrazione che sconta la conoscenza sia dell'espressione della differenza media, sia di quella del rapporto di concentrazione. Nella [*20*], invece, il rapporto quadratico di concentrazione introdotto da Bonferroni, che per ammissione esplicita dell'Autore non implica necessariamente la graduazione delle intensità, è, per definizione, determinato dal secondo membro della [*20*] stessa. Se si pensa alla impostazione giniana nel cui ambito la misura della concentrazione si fonda sulla graduazione e sulla condizione di trasferibilità, che è la premessa per conferire un senso agli accumuli delle intensità, si deve ragionevolmente concludere che la denominazione di "rapporto quadratico di concentrazione" riservata dal Bonferroni al suo indice sembra inadeguata; tuttavia, non è privo di interesse lo studio della relazione che intercorre tra l'indice quadratico del Bonferroni ed il rapporto di concentrazione del quadrato delle intensità da noi proposto. A tal fine, è comodo ricordare che dalla espressione del quadrato della differenza quadratica media senza ripetizione si perviene assai agevolmente al terzo membro della [*21*] che segue:

$$^{2}\Delta^2 = \frac{\sum_{1}^{N}{}_i \sum_{1}^{N}{}_{j\neq i} \left(x_i - x_j\right)^2}{N(N-1)} =$$

$$= \frac{2\left[(N-1)\sum_{1}^{N}{}_i\ x_i^2 - \sum_{1}^{N}{}_i \sum_{1}^{N}{}_{j\neq i}\ x_i x_j\right]}{N(N-1)} \qquad [21]$$

dove il coefficiente *2* segna il passaggio, come è ben noto, dalla struttura dispositoria a quella combinatoria. Seguendo Bonferroni, poniamo:

$$P_2^2 = \frac{\sum_{1}^{N}{}_i \sum_{1}^{N}{}_{j\neq i}\ x_i x_j}{N(N-1)} \qquad [22]$$

per cui, in definitiva, la [*21*] si riduce a:

$$^{2}\Delta^2 = 2\left(M_2^2 - P_2^2\right) \qquad [23]$$

In definitiva, quindi, il rapporto quadratico di concentrazione introdotto da Bonferroni, di cui alla [*20*], come d'altra parte è noto, può essere tradotto nella forma

$$R_2 = \frac{M_2^2 - P_2^2}{M_2^2} \qquad [24]$$

E' immediato riconoscere che il rapporto quadratico di concentrazione del Bonferroni, espresso come nella [*24*], assume i

valori estremi nelle stesse ipotesi delineate per il rapporto quadratico di concentrazione da noi introdotto. Ciò legittima il confronto per differenza tra i due indici, ed allora dalla [*12*] e dalla [*24*] si ha:

$$_{(2)}R - R_2 = \frac{P_2^2 - {}_2M^2}{M_2^2} \qquad [25]$$

Ricordando la [*22*] e la [*14*], si trova assai agevolmente

$$P_2^2 - {}_2M^2 = \frac{\sum_{1}^{N}{}_i \sum_{1}^{N}{}_{j\neq i}\ x_i x_j}{N(N-1)} - \frac{\sum_{1}^{N-1}{}_i \sum_{1}^{i}{}_h\ x_h^2}{\frac{1}{2}N(N-1)} =$$

$$= \frac{2\left[\sum_{1}^{N}{}_i \sum_{1}^{N}{}_{j>i}\ x_i x_j - \sum_{1}^{N-1}{}_i \sum_{1}^{i}{}_h\ x_h^2\right]}{N(N-1)} \qquad [26]$$

Essendo

$$\sum_{1}^{N}{}_i \sum_{1}^{N}{}_{j>i}\ x_i x_j = \sum_{1}^{N}{}_i\ x_i \sum_{i+1}^{N}{}_j\ x_j \qquad [27]$$

e così pure

$$\sum_{1}^{N-1}{}_i \sum_{1}^{i}{}_h\ x_h^2 = \sum_{1}^{N-1}{}_i\ (N-i)x_i^2 = \sum_{1}^{N}{}_i\ (N-i)x_i^2 \qquad [28]$$

dalla [*26*], in definitiva, si ha:

$$P_2^2 - {}_2M^2 = \frac{2\sum_1^N{}_i\; x_i\left[\sum_{i+1}^N{}_j\; x_j - (N-i)x_i\right]}{N(N-1)} \qquad [29]$$

Nella [*29*] il generico termine che compare entro parentesi quadra è definito dal confronto tra due somme ciascuna di *N-i* addendi: nella somma a minuendo gli addendi, in genere diversi tra loro, sono tutti non inferiori all'addendo costante nella somma a sottraendo per cui, in genere, il valore della espressione entro tale parentesi è positivo. Poiché il ragionamento sviluppato si deve ripetere per ogni valore dello indice i , si può concludere che la somma di tali differenze, ciascuna ponderata con la intensità x_i , è nulla soltanto nell' ipotesi di equidistribuzione ed in quella di massima concentrazione, risultando positiva in ogni altro caso. In altri termini, quindi, il rapporto quadratico di concentrazione da noi proposto assume un valore, tranne nei casi estremi, che è superiore a quello dell'indice quadratico di concentrazione stabilito dal Bonferroni, vale a dire è ${}_{(2)}R \geq R_2$. Il risultato conseguito si poteva considerare atteso se si pensa che il rapporto di concentrazione da noi costruito è fondato sulla differenza tra quadrati delle intensità mentre quello del Bonferroni si basa sui quadrati delle differenze; e su questo punto già ci siamo soffermati allo inizio del lavoro. Allora, proprio dalla [*25*] è immediato desumere l'espressione

$${}_{(2)}R = R_2 + \frac{P_2^2 - {}_2M^2}{M_2^2} \qquad [30]$$

che esprime, appunto, la relazione tra il rapporto di concentrazione da noi proposto e l'indice quadratico di concentrazione del Bonferroni.

La relazione testè illustrata ci consente anche di dare una risposta all'interrogativo, che il Bonferroni si pone proprio al termine del paragrafo *7* del suo lavoro del 1940, volto ad appurare se il suo indice possa essere dedotto dalla formula generale degli indici di concentrazione che l'Autore dà nel paragrafo *5* dello stesso scritto. Ci sembra, sulla scorta della [*30*], di poter affermare che l'indice quadratico R_2 non possa essere ricondotto alla espressione generale degli indici di concentrazione del Bonferroni, mentre ciò è invece possibile per l'indice ${}_{(2)}R$, da noi proposto, quando si scelga come media parziale delle prime i intensità della graduatoria la media quadratica (al quadrato), vale a dire, la media aritmetica dei quadrati delle intensità

5. Relazione tra ${}_{(2)}R$ e il rapporto di concentrazione del Gini

E' del tutto spontaneo ritenere che il rapporto di concentrazione del quadrato delle intensità presenti valori maggiori di quelli assunti dal rapporto di concentrazione di Gini, se si escludono i valori estremi che sono comuni ai due indici con riferimento alle rispettive ipotesi di equidistribuzione e di massima concentrazione. D'altra parte, se si pone mente alla espressione dell'indice del Gini

$$R = \frac{\sum_1^{N-1}{}_i\;(p_i - q_i)}{\sum_1^{N-1}{}_i\; p_i} \qquad [31]$$

in cui, com'è noto, è posto $q_i = \sum_1^i{}_h\; x_h / \sum_1^N{}_h\; x_h$, e la si raffronta con la formula [*11*] del presente scritto, è immediato riconoscere che i differenti valori dei due indici dipendono unicamente dalle quantità q_i e ${}_{(2)}q_i$ che in essi figurano. Dal confronto di q_i , che compare in R, con la quantità ${}_{(2)}q_i$ del nostro indice, come risulta espressa al secondo membro della [*5*], consegue la relazione

$${}_{(2)}q_i \leq q_i, \quad \forall i \qquad [32]$$

ossia

$$\frac{\sum_{1}^{i} {}_h x_h^2}{\sum_{1}^{N} {}_h x_h^2} \leq \frac{\sum_{1}^{i} {}_h x_h}{\sum_{1}^{N} {}_h x_h} \qquad [33]$$

proprio perché la [*33*] stessa è suscettibile della trasformazione

$$\frac{\sum_{1}^{i} {}_h x_h^2}{\sum_{1}^{i} {}_h x_h} \leq \frac{\sum_{1}^{N} {}_h x_h^2}{\sum_{1}^{N} {}_h x_h} \qquad [34]$$

che esprime la relazione in cui si trova la media antiarmonica calcolata sulle prime i intensità graduate in ordine non decrescente, con la media antiarmonica afferente all'intero collettivo. La [*32*] consente, pertanto, di stabilire la relazione

$${}_{(2)}R \geq R \qquad [35]$$

tra rapporto di concentrazione del quadrato delle intensità e rapporto di concentrazione del Gini.

La struttura logico formale di ${}_{(2)}R$ e le relazioni con l'indice R_2 del Bonferroni e l'indice R del Gini, ripropongono la questione generale, per qualsiasi modello statistico, dell'approfondimento critico della portata euristica in riferimento ai contesti fenomenici in cui lo strumento di conoscenza può convenientemente calarsi. Ben si comprende come non risulti spontaneo prefigurare situazioni rispetto alle quali l'indice ${}_{(2)}R$ possa proporsi quale soluzione razionale ed opportuna specie se si rammenta che il nostro contributo, come si è

avuto modo di specificare nei paragrafi introduttivi, è stato sollecitato, in via del tutto immediata, dalla lettura di uno scritto metodologico del Bonferroni. E' corretto ritenere che siano le esigenze che si prospettano nello sviluppo della ricerca sostanziale a suggerire la scelta dello strumento statistico e ciò, naturalmente, dovrà valere anche per l'indice proposto nel presente lavoro, tenuto conto delle sue specificità.

6. Calcolo grafico del rapporto di concentrazione del quadrato delle intensità

Del rapporto di concentrazione è possibile dare una rappresentazione grafica; il procedimento che ora illustreremo, proprio in quanto è coerente ad un'ipotesi di continuità posta nei confronti delle grandezze oggetto della rappresentazione, che ovviamente è assente nell'impostazione algebrica in precedenza descritta, non è privo di una propria ed autonoma forza interpretativa che, se da un lato contribuisce a chiarire alcuni aspetti della precedente formulazione algebrica, dall'altro consente di evidenziare elementi distintivi ed innovativi. A tal fine, iniziamo col rappresentare nel piano cartesiano gli N punti di coordinate

$(i;\sum_{1}^{i} {}_h x_h^2)$, per $i=1,2,\ldots,N$ che, congiunti con segmenti di retta, consentono di definire la spezzata di concentrazione *OABC..DE..FGS* Nella impostazione attuale il rapporto di concentrazione è concepito come rapporto tra l'area di concentrazione ed il valore massimo che tale area può presentare.

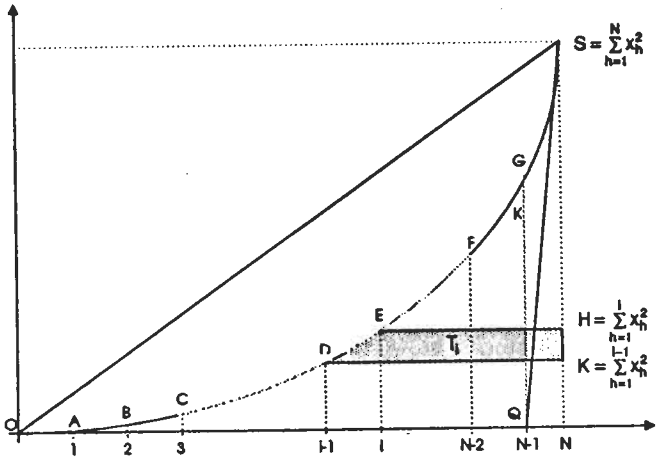

grafico 1

Com'è noto, l'area di concentrazione è la porzione di piano racchiusa tra il segmento di equidistribuzione e la spezzata di concentrazione; di contro, l'area di massima concentrazione si identifica con la superficie del triangolo rettangolo che ha per ipotenusa il segmento di equidistribuzione. Prima di procedere alla determinazione delle superfici ora menzionate ci sia consentito di indugiare brevemente sul significato che, nello specifico contesto in cui stiamo operando, assume il segmento di equidistribuzione *OS*; è ancora opportuno rammentare che la quantità oggetto della uguale ripartizione tra le unità del collettivo è costituita dalla somma dei quadrati delle intensità del carattere trasferibile e da ciò, ovviamente, consegue che l'entità costante da assegnare ad ogni unità, perché si possa realizzare l'equidistribuzione, è rappresentata dalla media aritmetica del quadrato delle intensità, ossia dal quadrato della media quadratica delle intensità.

Passiamo ora a determinare l'espressione dell'area di concentrazione quando la si intenda come differenza tra l'area di massima concentrazione e la superficie che è somma dei trapezi sottostanti alla spezzata di concentrazione. L'area di massima

concentrazione ${}_{\max}A_C$, è pari alla superficie del triangolo *ONS* e assume il valore

$$ {}_{\max}A_C = \frac{1}{2} N \sum_{1}^{N} {}_h \, x_h^2 = \frac{N^2}{2} M_2^2 \qquad [36] $$

Per quanto concerne la superficie del generico trapezio *DEHK*, che annotiamo con T_i, valgono i seguenti sviluppi formali:

$$ T_i = \frac{1}{2}\{[N-i]+[N-(i-1)]\}.\left\{\sum_{1}^{i} {}_h \, x_h^2 - \sum_{1}^{i-1} {}_h \, x_h^2\right\} \qquad [37] $$

ed anche

$$ T_i = \frac{1}{2}\{2(N-i)+1\}x_i^2 \qquad [38] $$

Va da sè che la superficie sottesa alla spezzata di concentrazione, che è la somma dei trapezi T_i (per $i=1,2,\ldots,N$), è pari a

$$ \sum_{1}^{N} {}_i \, T_i = \frac{1}{2}\sum_{1}^{N} {}_i \, \{2(N-i)+1\}x_i^2 \qquad [39] $$

Se si pone mente alla nota relazione $\sum_{1}^{N} {}_i \, [2(N-i)+1] = N^2$ è agevole ravvisare nell'espressione

$$ {}_2M_*^2 = \frac{\sum_{1}^{N}{}_i\left[2(N-i)+1\right]x_i^2}{N^2} \qquad [40] $$

la media aritmetica delle medie aritmetiche parziali dei quadrati delle intensità, e riconsiderare la [39] in funzione di tale media; allora risulta:

$$ \sum_{1}^{N}{}_i\ T_i = \frac{N^2}{2}\,{}_2M_*^2 \qquad [41] $$

Pertanto, siamo in grado di determinare il valore dell'area di concentrazione A_C compresa tra il segmento di equidistribuzione OS e la spezzata di concentrazione; si ha

$$ A_C = {}_{\max}A_C - \sum_{1}^{N}{}_i\ T_i = \frac{N^2}{2}\left(M_2^2 - {}_2M_*^2\right) \qquad [42] $$

E' poi immediato riconoscere nel rapporto

$$ {}_{(2)}R_* = \frac{A_C}{{}_{\max}A_C} = \frac{M_2^2 - {}_2M_*^2}{M_2^2} \qquad [43] $$

l'espressione del rapporto di concentrazione determinata in conformità ad una impostazione grafica e formalmente omogenea all'espressione [12] di ${}_{(2)}R$. Anche ${}_{(2)}R_*$, tramite la [43], risulta espresso in funzione della media aritmetica del quadrato delle intensità $M_2^{\;2}$ e della media aritmetica ${}_2M_*^{\;2}$ delle medie aritmetiche parziali di tali quadrati.

7. Relazione tra ${}_{(2)}R_*$ e la differenza media del quadrato delle intensità

L'espressione di ${}_{(2)}R_*$, di cui alla [43], può essere riproposta nella forma

$$ {}_{(2)}R_* = \frac{N\sum_{1}^{N}{}_h\ x_h^2 - \sum_{1}^{N}{}_i\left[2(N-i)+1\right]x_i^2}{N^2M_2^2} \qquad [44] $$

quando si rammenti l'espressione di ${}_2M_*^{\;2}$ data tramite la [40]. Per quanto concerne il numeratore della [44] è del tutto agevole verificare che esso si identifica con la espressione

$$ \sum_{1}^{N}{}_h\sum_{1}^{N}{}_{k\le h}\left|x_k^2 - x_h^2\right| = \frac{1}{2}\sum_{1}^{N}{}_h\sum_{1}^{N}{}_k\left|x_k^2 - x_h^2\right| \qquad [45] $$

La [44] può, allora, essere riscritta nella seguente forma

$$ {}_{(2)}R_* = \frac{\sum_{1}^{N}{}_h\sum_{1}^{N}{}_k\left|x_k^2 - x_h^2\right|}{2N^2M_2^2} \qquad [46] $$

Se si annota la differenza semplice media (con ripetizione) del quadrato delle intensità con

$$ {}_{(2)}\Delta_R = \frac{\sum_{1}^{N}{}_h\sum_{1}^{N}{}_k\left|x_k^2 - x_h^2\right|}{N^2} \qquad [47] $$

è del tutto immediato individuare l'espressione

$$_{(2)}R_* = \frac{_{(2)}\Delta_*}{2M_2^2} \qquad [48]$$

che pone in relazione il rapporto di concentrazione $_{(2)}R_*$ e la differenza semplice media con ripetizione del quadrato delle intensità. Tra gli indici $_{(2)}R$ e $_{(2)}R_*$ sussiste una elementare relazione che può essere evidenziata rammentando quella che intercorre tra differenza media con ripetizione e differenza media senza ripetizione:

$$_{(2)}\Delta_R = {}_{(2)}\Delta \frac{N-1}{N} \qquad [49]$$

In base alla [49] possiamo riscrivere la [48] nella seguente forma

$$_{(2)}R_* = \frac{_{(2)}\Delta}{2M_2^2} \; \frac{N-1}{N} \qquad [50]$$

pervenendo, quindi, al risultato

$$_{(2)}R_* = \frac{N-1}{N} \; {}_{(2)}R \qquad [51]$$

che esprime la relazione cercata. L'espressione [48],se confrontata con la [19], consente inoltre di affermare che l'indice $_{(2)}R_*$, a differenza di $_{(2)}R$, soddisfa il principio di somiglianza, ossia non muta di valore se calcolato nei confronti di distribuzioni tra loro simili, poiché a tale principio si uniforma la differenza media con ripetizione e non quella senza ripetizione.
Infine, è possibile confrontare l'espressione [43] con la [48]; si ha

$$\frac{M_2^2 - {}_2M_*^2}{M_2^2} = \frac{_{(2)}\Delta_R}{2M_2^2} \qquad [52]$$

da cui segue immediatamente una ulteriore formulazione della differenza semplice media (con ripetizione) dei quadrati delle intensità:

$$_{(2)}\Delta_R = 2\left(M_2^2 - {}_2M_*^2\right) \qquad [53]$$

Il raffronto della [53] con la [42] consente poi di definire la relazione

$$A_C = \frac{N^2}{4} \; {}_{(2)}\Delta_R \qquad [54]$$

tra l'area di concentrazione e la differenza media in questione.

Già si è detto che nella costruzione dell'indice $_{(2)}R_*$ interviene un modello di massima concentrazione - costituito dai cateti del triangolo rettangolo che ha per ipotenusa il segmento *OS* di equidistribuzione - formalmente e concettualmente diverso da quello utilizzato per definire $_{(2)}R$ e proprio dal differente riferimento teorico discendono le peculiari caratteristiche illustrate per i due indici. E' possibile, anche, impostare in termini grafici il problema della rappresentazione del rapporto di concentrazione del quadrato delle intensità con riferimento ad un modello massimante di concentrazione coincidente con quello assunto per l'indice $_{(2)}R$, ossia con riferimento alla distribuzione che veda accentrata in una sola unità statistica la quantità totale $\sum_{1}^{N} {}_i \; x_i^2$.

Detta impostazione induce ad interpretare il rapporto di concentrazione come rapporto tra l'area di concentrazione *A* ed il suo massimo valore ora espresso dal triangolo *OQS*, del grafico *f*, in cui *Q* è il punto di coordinate (*N-1* , *0*); si ha

$$_{(2)}R = \frac{A_C}{O\overset{\Delta}{Q}S} \qquad [55]$$

Per quanto concerne il valore di A_c nulla cambia rispetto al procedimento grafico illustrato in precedenza e, quindi, vale il risultato acquisito con la [*42*]. Ciò che ora muta è la definizione di area di massima concentrazione, riconosciuta nel triangolo *OQS*, il cui valore è pari a

$$O\overset{\Delta}{Q}S = \frac{(N-1)N}{2}\,M_2^2 \qquad [56]$$

E che l'interpretazione espressa mediante la [*55*] sia congrua è avvalorata dal risultato a cui si perviene rapportando tra loro i valori delle aree che in essa figurano; risulta

$$\frac{A_C}{O\overset{\Delta}{Q}S} = \frac{N}{N-1}\;\frac{M_2^2 - {}_2M_*^2}{M_2^2} \qquad [57]$$

che è l'espressione dell'indice $_{(2)}R$ in funzione di quella di $_{(2)}R_*$, direttamente desumibile dalla relazione [*51*] tenuto conto del terzo membro della [*43*].

8. Interpretazione geometrica dell'indice quadratico di concentrazione del Bonferroni

Nel presente paragrafo cercheremo di attribuire una adeguata interpretazione geometrica all'indice quadratico di concentrazione del Bonferroni mediante la ricerca di un suo conveniente inquadramento nel grafico di rappresentazione del rapporto di concentrazione del quadrato delle intensità, e ciò anche allo scopo di evidenziare sotto il profilo geometrico le relazioni intercorrenti tra i due indici. A tal fine è opportuno considerare il grafico *2* che si distingue dal grafico *1* per la introduzione del segmento *OP* di coefficiente angolare pari a M_1^2; detto segmento illustra il processo di accumulazione, caratteristico dello schema di concentrazione, che si realizzerebbe qualora venisse attribuita ad ognuna delle N unità la medesima quantità M_1^2, corrispondente al quadrato della media aritmetica delle intensità.

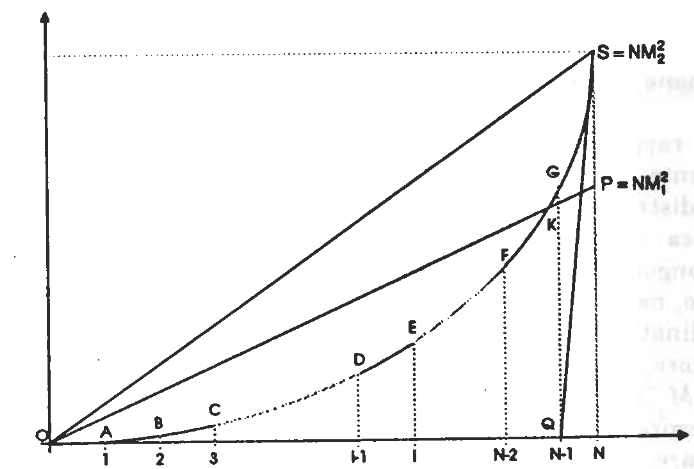

grafico 2

E' del tutto agevole accertare che l'indice R_2 del Bonferroni è suscettibile di una particolare interpretazione geometrica che si traduce nel rapporto tra le superfici di due triangoli:

$$R_2 = \frac{O\hat{P}S}{O\hat{Q}S} \qquad [58]$$

Infatti, per quanto concerne il numeratore, risulta:

$$O\hat{P}S = \frac{\left(NM_2^2 - NM_1^2\right)N}{2} = \frac{N^2}{2}\left(M_2^2 - M_1^2\right) \qquad [59]$$

e, quindi:

$$O\hat{P}S = \frac{N^2}{4}\,{}^2\Delta_R^2 \qquad [60]$$

mentre al riguardo del denominatore vale il risultato espresso nella *[56]*. Va da sè che il rapporto della *[60]* alla *[56]* riproduce proprio l'espressione *[20]* con ciò accreditando l'interpretazione data mediante la *[58]*.
Anche il rapporto di cui alla *[58]* assume valori nell'intervallo [*0,1*]. Ovviamente il valore *0* si presenta quando, per effetto dell'equidistribuzione dell'intensità totale media aritmetica e media quadratica coincidono; in tale evenienza i punti S e P si sovrappongono rendendo nulla l'area *OPS*.
Di contro, nell'ipotesi di massima concentrazione il punto S presenta una ordinata pari a $N^2M_1^2$ (giacché ora risulta $M_2^2=NM_1^2$) e numeratore e denominatore della *[58]* realizzano lo stesso valore $[N^2(N-1)M_1^2]:2$ riconducendo a *1* quello del rapporto.

L'interpretazione di R_2 data a mezzo della *[58]* consente di evidenziare, anche sul piano grafico, l'inidoneità dell'indice del Bonferroni di proporsi come rapporto di concentrazione nel senso giniano. Tale indice infatti, risulta definito unicamente in funzione di modelli teorici costituiti, rispettivamente, dal segmento *OS* di

equidistribuzione della quantità $\sum_1^N{}_i\, x_i^2$, dal segmento *OP* di

equidistribuzione della intensità totale $\sum_1^N{}_i\, x_i$ e dalla spezzata *OQS*

posta ad interpretazione dell'ipotesi massimante; non compare in alcun modo la spezzata di concentrazione che, come è noto, illustra la graduazione delle intensità osservate ed il conseguente processo di accumulazione tipico della impostazione del Gini.

Giancarlo Bettuzzi

Riferimenti bibliografici

V. Amato (1948), Sulla misura della concentrazione dei redditi, *Rivista Italiana di Demografia e Statistica.*

V. Amato (1949), Di un indice di concentrazione, *Statistica n. 1.*

C. Bonferroni (1933), *Elementi di statistica generale,* E. Gili, Torino.

C. Bonferroni (1940), Un indice quadratico di concentrazione, *Atti del secondo Congresso dell'Unione Matematica Italiana*, Bologna 1940.

V. Castellano (1933), Sull'interpretazione dinamica del rapporto di concentrazione, *Giornale dell'Istituto Italiano degli Attuari.*

S. De Simoni (1966), A proposito dei limiti di medie parziali caratteristiche nell'ambito della curva di concentrazione, *Statistica n.3.*

M. De Vergottini (1940), Sul significato di alcuni indici di concentrazione, *Giornale degli Economisti e Annali di Economia.*

C. Ferreri (1978), Significato e possibilità di taluni indicatori di concentrabilità economica dei redditi per una misura delle variazioni nel benessere economico, *Scritti in onore di Giuseppe de Meo*, Istituto di Statistica economica, Roma.

P. Fortunati (1952), Appunti sulle misure statistiche della variabilità, *Statistica n.3.*

P. Fortunati (1957), Rapporto di concentrazione valori medi e schemi teorici di distribuzione massimante e minimante di variabilità, *Statistica n.2.*

P. Fortunati (1961), Relazioni tra le misure di variabilità, *Statistica n.2.*

A. Gili (1985), *Introduzione alla tecnica statistica*, CLUEB, Bologna.

C. Gini (1914), Sulla misura della concentrazione e della variabilità dei caratteri, *Atti del R.Istituto Veneto di Scienze, Lettere ed Arti,* Tomo LXXIII, Parte II.

G. Pietra (1915), Delle relazioni tra gli indici di variabilità, *Atti del R.Istituto Veneto di Scienze, Lettere ed Arti,* Tomo LXXIV Parte II