

L'interpolazione della spezzata di Lorenz (*)

Agostino Tarsitano
Università degli studi della Calabria
Dipartimento di Economia e Statistica
87030 Arcavacata di Rende (Cs)
agotar@unical.it

Riassunto

La curva di Lorenz è la rappresentazione grafica della relazione che si instaura fra quote cumulate di redditi t_0 e frazioni cumulate di redditi t_1 una volta che i redditi siano stati ordinati in senso crescente. In questa nota la curva viene studiata con un approccio diretto, senza la mediazione del modello di ripartizione del reddito.

Dopo aver presentato i requisiti formali di un modello generale di curva di Lorenz si è studiata una forma funzionale ottenuta generalizzando la curva di Lorenz associata alla distribuzione di Pareto. In funzione di tale curva sono state ottenute le formule di vari indici di concentrazione e di asimmetria.

Si è discussa infine la stima dei parametri della curva con applicazioni nell'ambito della relazione tra sviluppo economico e distribuzione del reddito.

keywords: indici di concentrazione, ottimizzazione vincolata

(*) *Statistica anno XLVII, n.3, pp. 437-451*

1. Introduzione⁽¹⁾

L'interesse per l'adattamento di modelli statistici alla distribuzione empirica dei redditi si è ridestato anche per l'interpolazione della spezzata di Lorenz⁽²⁾. Il legame fra la curva di Lorenz ed il modello di ripartizione del reddito è molto stretto (Arnold, 1983, pp. 105-108); lo studio delle curve di concentrazione non può cioè essere considerato alternativo allo studio dei modelli di ripartizione⁽³⁾. Tuttavia, come osservava Gini (1932): "... *Il loro esame può essere più fruttuoso e la loro rappresentazione analitica più facile a darsi di quanto non avvenga per le rispettive curve di frequenza...*".

Lo scopo di questa nota è di discutere un insieme di condizioni di regolarità per definire una famiglia di funzioni che goda delle proprietà tipiche della curva di Lorenz. All'interno di questa famiglia verrà scelto un modello particolare, la curva di Pareto generalizzata, da adattare a dati storici e sezionali della distribuzione del reddito.

L'intento finale sarà di calcolare, a mezzo del modello interpolante, alcuni indici connessi alla curva di Lorenz in grado di caratterizzare la ripartizione di reddito osservate.

2. Un modello per la curva di Lorenz

la curva di Lorenz è la rappresentazione grafica della relazione che si instaura fra quote cumulate di reddito (indicate con q) e frazioni cumulate di redditi (indicate con p) una volta che i redditi stessi siano ordinati in senso non decrescente. In questo paragrafo verranno presentate delle condizioni di regolarità che trasmettono le peculiarità della curva di Lorenz empirica⁽⁴⁾ al modello analitico

$$q = L(p; d) \quad 0 \leq p \leq 1 \quad (2.1)$$

dove $d = (d_1, d_2, \dots, d_m)' \in D \subseteq \mathcal{R}^m$ è il vettore di parametri che individua l'elemento particolare nella famiglia di funzioni (2.1).

Lo spazio parametrico D , sottoinsieme dello spazio euclideo m -dimensionale, \mathcal{R}^m sarà al centro delle condizioni 4 e 5 laddove le condizioni 1, 2 e 3 riguarderanno la struttura generale del modello⁽⁵⁾.

Condizione_1: Positività

La curva di Lorenz $L(p;d)$ è una funzione reale definita sull'intervallo chiuso $[0,1]$ e che ha qui anche il suo codominio. In particolare

$$\lim_{p \downarrow 0} L(p;d) = 0 \quad e \quad \lim_{p \uparrow 1} L(p;d) = 1 \quad \text{per ogni } d \in D \quad (2.2)$$

Condizione_2: Nondecrescenza

Siano p_1 e p_2 due punti in $[0,1]$. Allora

$$\text{Se } p_1 \leq p_2 \Rightarrow L(p_1;d) \leq L(p_2;d) \text{ per ogni } d \in D \quad (2.3)$$

Condizione_3: Convessita ⁽⁶⁾.

Si abbiano p_1, p_2 ed α tutti in $[0,1]$ e si supponga che $p_1 < p_2$. Allora

$$L[\alpha p_1 + (1-\alpha)p_2; d] \leq \alpha L(p_1; d) + (1-\alpha)L(p_2; d) \text{ per ogni } d \in D \quad (2.4)$$

Condizione_4: Unicit  per l'equidistribuzione

Esiste uno ed un solo punto $d^+ \in D$ tale che

$$L(p; d^+) = p \text{ per ogni } p \in [0,1] \quad (2.5)$$

Tale condizione richiede in pratica che il modello adottato generi la curva di concentrazione minima per una unica combinazione di parametri. Un requisito essenziale per poter poi interpretare la situazione di distribuzione rappresentata dalla curva.

Condizione_5: Asimmetria

Lo spazio parametrico D pu  essere suddiviso nei tre sottoinsiemi D_1, D_2, D_3 tali che

$$\bigcup_{i=1}^3 D_i = D, \quad D_i \cap D_j = \emptyset \text{ se } i \neq j \quad (2.6)$$

In tali insiemi succede che per ogni $p \in [0,1]$ si ha

$$L(1-p; d) = (1-p) \text{ per ogni } d \in D_1 \quad (2.7)$$

$$L(1-p; d) < (1-p) \text{ per ogni } d \in D_2 \quad (2.8)$$

$$L(1-p; d) > (1-p) \text{ per ogni } d \in D_3 \quad (2.9)$$

La condizione 5 impone che la curva di Lorenz $L(p; d)$ sia simmetrica (2.7) oppure presenti asimmetria destrorsa (2.8) ed asimmetria sinistrorsa (2.9) per combinazioni distinte di parametri. Lo scopo della condizione 5   anche in questo caso di facilitare la comprensione del fenomeno distributivo espresso dal modello.

Il concetto di asimmetria adottato per la curva di Lorenz   analogo a quello applicato alle distribuzioni di frequenza⁽⁷⁾. Le formule (2.7)-(2.8) diventano pi  chiare se il diagramma di Lorenz viene traslato nel sistema di coordinate⁽⁸⁷⁾

$$x = (p+q)/\sqrt{2} \quad ; \quad y = (p-q)/\sqrt{2} \quad (2.10)$$

che equivale in pratica ad osservare la curva di Lorenz in uno specchio posto sulla retta di equidistribuzione. In tale contesto la curva di Lorenz avr  asimmetria destrorsa se il massimo per y viene ottenuto per una ascissa maggiore di quella mediana ($x=1/\sqrt{2}$). L'asimmetria sar  invece sinistrorsa se il massimo   ottenuto per un'ascissa minore di quella mediana. Condizione necessaria perch  la curva di Lorenz sia simmetrica   che il massimo si abbia per $x=1/\sqrt{2}$.

3. Calcolo degli indici

In funzione della curva di Lorenz si possono calcolare le più comuni misure di concentrazione e di asimmetria. In questo paragrafo verranno date le formule degli indici più direttamente connessi alla curva. Un breve commento seguirà quelli meno usuali.

Indice di Gini⁽⁹⁾

$$G=1 - 2 A(1;d) \quad (3.1)$$

dove $A(x;d)$ indica l'area sottesa alla curva $L(p;d)$ nell'intervallo $[0,x]$. Come è ben noto l'indice (3.1) equivale al rapporto tra la differenza media ed il doppio della media aritmetica.

Indice del Pietra

$$E = \text{Massimo } \{p \in [0,1] \mid (p - L(p;d))\} \quad (3.2)$$

L'indice (3.2) è pari al rapporto tra lo scostamento semplice medio dalla media aritmetica ed il suo massimo (il doppio della media aritmetica).

Indici di Elteto e Frigyes

Sia p_E il frattile di popolazione che determina E e sia q_E il corrispondente frattile di reddito. Le due quantità consentono il computo degli indici proposti da Elteto e Frigyes (1968)

$$u = (p_E - q_E)/q_E; \quad w = (p_E - q_E)/(1 - q_E); \quad v = u + w - u*w \quad (3.3)$$

Tutti e tre gli indici variano tra zero (assenza di concentrazione) ed uno (concentrazione massima). Poichè $u*w/v = E$ gli indici (3.3) non sembrano apportare più informazione di quanto già non faccia E (Kakwani, 1980, p. 82). Tuttavia, Arnold (1983, p. 115) osserva che mentre E è una funzione dei (3.3) il converso non è vero e dunque il loro contenuto informativo supera quello del E .

Lunghezza della curva

$$S = \int_0^1 \sqrt{\{1 + L'(p;d)^2\}} dp \quad (3.4)$$

L'indice (3.4) varia⁽¹⁰⁾ tra $\sqrt{2}$ (assenza di concentrazione) e 2 (concentrazione massima).

Indici basati su punti specifici del diagramma di Lorenz

Fra i metodi per misurare la concentrazione Gini (1932) incluse

$$\gamma_r = L(1/2; d) \quad (3.5)$$

L'indice (3.5) è dato dal frattile di reddito spettante alla metà più povera dei redditeri. Ne consegue che $0 \leq \gamma_r \leq 1/2$. Pietra (1914-1915) dimostrò che (3.5) è pari al rapporto tra

lo scostamento semplice medio dalla mediana ed il doppio della media aritmetica. Un altro indice, analogo al precedente, è

$$\gamma_p = \text{Minimo}\{ p \in [0,1] \mid L(p;d)=1/2\} \quad (3.6)$$

L'indice (3.6) è dato dal frattile più piccolo di popolazione cui spetta il 50% del reddito totale⁽¹¹⁾. Ne consegue che $1/2 \leq \gamma_p \leq 1$. Gini (1932) dimostrò che (3.6) ha significato analogo a (3.5) ma per quella che egli chiamò "antiserie"⁽¹²⁾.

Indice di Zanardi

Si indichi con y_c il livello del reddito tale che la frazione p_c di redditi possiede la frazione di reddito⁽¹³⁾ $q_c=(1-p_c)$. In pratica (p_c, q_c) è il punto di intersezione tra la curva $L(p;d)$ e la retta di equidistribuzione. Zanardi (1965) ha proposto di misurare l'asimmetria della curva di Lorenz con l'indice

$$Z = 2p_c q_c [G_s - G_i] = (4/G) \{ (1-p_c)^2 + (G-1)/2 + 2A[p_c;d] \} \quad (3.7)$$

Nella (3.7) G_s è l'indice di Gini della seriazione dei redditi superiori ad y_c e G_i è lo stesso indice per i redditi non superiori ad y_c . L'indice Z si annulla nel caso di perfetta simmetria. È positivo se la asimmetria è destrorsa ovvero quando la concentrazione è maggiore fra i gruppi più ricchi. È negativo se l'asimmetria è sinistrorsa cioè se la disparità è maggiore fra i gruppi più poveri. Il campo di variazione di Z è l'intervallo $[-1,1]$. La tendenza all'estremo inferiore si realizza quando ridottasi la curva di Lorenz alla estremante minore⁽¹⁴⁾

$$q = \begin{cases} 0 & \text{se } p \leq G \\ \frac{p-G}{1-G} & \text{se } p \geq G \end{cases} \quad (3.8)$$

la concentrazione tende a zero. La tendenza all'estremo superiore si ha quando la curva di Lorenz si riduce alla estremante maggiore⁽¹⁵⁾

$$q = \begin{cases} (1-G)p & \text{se } p < 1 \\ 0 & \text{se } p \geq 1 \end{cases} \quad (3.9)$$

e l'indice G tende a zero. Poiché Z è basato su di una differenza tra aree non si può escludere un suo valore nullo per curve non esattamente simmetriche. Si tratterà comunque di casi in cui l'effettiva asimmetria è trascurabile.

Indice di Taguchi

Taguchi (1968) ha proposto di misurare l'asimmetria della curva di Lorenz con il seguente indice

$$T = 2\cos^2(\theta) - 1 \quad (3.10)$$

dove θ è l'angolo formato con l'asse delle p dalla tangente alla curva di Lorenz nel punto (p_T, q_T) in cui la curva stessa ha la massima curvatura. In pratica

$$p_T = \text{Massimo}\{ p \in [0,1] \mid L''(p;d) / [1 + L'(p;d)^2]^{(3/2)} \} \quad (3.11)$$

Anche per l'indice di Taguchi il campo di variazione è l'intervallo $[-1,1]$, ma a differenza dell'indice di Zanardi associa valori negativi a curve asimmetri che destrorse e valori positivi a curve asimmetriche sinistrorse. La tendenza all'estremo inferiore si realizza quando ridotta la curva alla estremante maggiore, la concentrazione tende a quella massima. L'indice T tenderà ad uno quando la curva di Lorenz si riduce alla estremante minore ed il G tende pure ad uno. Un difetto dell'indice T è la sua incapacità di distinguere curve solo culminanti⁽¹⁶⁾ dalle curve simmetriche.

Indice di Hagerbaumer

Hagerbaumer (1977) ha suggerito di valutare la posizione dei relativamente poveri con la seguente misura⁽¹⁷⁾

$$H = \{(2-G)/[G(1-G)]\} \{P_h q_h - 2A[p_h; d]\} \quad (3.12)$$

dove (P_h, q_h) è il punto di intersezione della curva di Lorenz con la curva (3.9). L'indice H varia nell'intervallo $[0,1]$ e risulta crescente se, per un fisato G , cresce la quota posseduta dal 50% più povero dei redditeri in rapporto al massimo che questa può raggiungere. L'estremo inferiore si ottiene se la curva di Lorenz coincide con la (3.8) e l'estremo superiore se la curva di Lorenz si sovrappone alla (3.9).

4. La curva di Pareto generalizzata

I requisiti relativamente modesti voluti per il modello $L(p;d)$ danno spazio ad una ampia gamma di forme analitiche. L'attenzione, in questo lavoro si è limitata ad una curva biparametrica che ben si accorda alle condizioni discusse nel secondo paragrafo. La curva è

$$L(p;a,b) = 1 - (1-p^{1/a})^b \quad 0 < a, b \leq 1 \quad (4.1)$$

La presenza di almeno due parametri è essenziale se si vogliono cogliere i due aspetti più importanti della curva di Lorenz: l'area ad essa sottesa e la asimmetria.

La curva (4.1) è una generalizzazione della curva di Lorenz che si ottiene quando il modello di ripartizione è quello di Pareto⁽¹⁸⁾. La curva è peraltro la funzione inversa di quella proposta da Raasche et al. (1984). Nella figura 1 vengono tracciate alcune curve del modello (4.1).

Dato che la (4.1) è continua in $[0,1]$ La condizione 1 del paragrafo 2 è di riscontro immediato. All'interno di $[0,1]$ si hanno

$$L'(p;a,b) = (b/a)(1-p^{1/a})^{b-1} p^{1/a-1} \quad (4.2)$$

$$L''(p;a,b) = [L'(p;a,b)/a] \{ (1-a)/p + (1-b)/(1-p^{1/a}) \} \quad (4.3)$$

La (4.2) è positiva per $p \in [0,1]$ e così pure la (4.3) (dati i vincoli sui parametri). Quindi

anche le condizioni 2 e 3 sono rispettate.

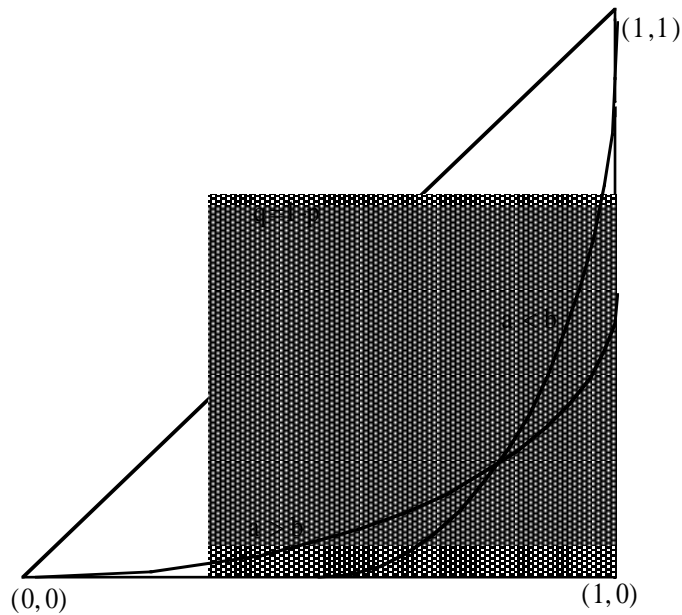


Figura 1: Esempi della curva di Pareto generalizzata.

La valutazione delle (2.7)-(2.9) permette di stabilire che la curva di Pareto generalizzata è simmetrica per $a=b$. In particolare da luogo alla retta di equidistribuzione se e solo se $a=b=1$. L'asimmetria è sinistrorsa per $a < b$ e destrorsa se $a > b$.

Le formule per il calcolo⁽¹⁹⁾ degli indici visti nel paragrafo 3 sono le seguenti

$$G(a,b) = [2ab/(a+b)] B(a,b) - 1 \quad (4.4)$$

dove $B(r,s)$ è la funzione matematica Beta. Da notare che $G(a,b) = G(b,a)$.

$$E(a,b) = \{p \in [0,1] \mid (b/a)(1-p^{1/a})^{b-1} p^{1/a-1} = 1\} \quad (4.5)$$

Dopo il calcolo di $E(a,b)$ quello degli indici di Elteto e Frigyes segue immediatamente.

$$S(a,b) = \int_0^1 \left[1 + \left[\left(\frac{b}{a} \right) \left(1 - p^{\frac{1}{a}} \right)^{b-1} p^{\frac{1}{a}-1} \right]^2 dp \quad (4.6)$$

$$\gamma_r(a,b) = 1 - [1 - (1/2)^{1/a}]^b \quad (4.7)$$

$$\gamma_p(a,b) = 1 - [1 - (1/2)^{1/b}]^a \quad (4.8)$$

Si vede subito che per la (4.1) si ha $\gamma_r(a,b) = \gamma_p(b,a)$.

$$Z(a,b) = (4/G)(p_c^2 + [(G+1)/2][1 - 2IB(p_c^{1/a}; a, b+1)]) \quad (4.9)$$

dove $p_c = \{p \in [0,1] \mid p_c^{1/a} + p_c^{1/b} = 1\}$ e $IB(x;r,s) = [B(r,s)]^{-1} \int_0^x t^{r-1} (1-t)^{s-1} dt$. Per l'indice di Zanardi, nella (4.1), vale⁽²⁰⁾ la relazione $Z(a,b) = -Z(b,a)$. L'indice di Taguchi non ha, nella (4.1), una formula semplice⁽²¹⁾ a presentarsi. Per l'indice di Hagerbaumer si ha

$$H(a,b) = \{(2-G)/[G(2-G)]\} \{(q_h - 2)p_h + (G+1)IB(p_h^{1/a}; a, b+1)\} \quad (4.10)$$

dove $p_h = \{p \in [0,1] \mid (1-G)p_h = 1 - (1-p_h^{1/a})^b\}$.

Le situazioni distributive che la curva di Pareto generalizzata è in grado di formalizzare possono essere brevemente illustrate usando gli indici di Gini e di Zanardi. Le derivate parziali di $G(a,b)$ rispetto ai due parametri sono

$$G'_a(a,b) = (1+G)[\Psi(a+1) - \Psi(a+1+b)] \quad (4.11)$$

$$G'_b(a,b) = (1+G)[\Psi(b+1) - \Psi(a+1+b)] \quad (4.12)$$

dove $\Psi(x)$ è la funzione matematica Digamma: continua, monotona e crescente per $x > 0$. (Abramowitz e Stegun, 1964, pp. 258-259). Quindi le due derivate parziali sono negative. La conseguenza è che se uno dei parametri aumenta e l'altro rimane costante, la concentrazione dei redditi diminuisce. La diminuzione può però essere meglio qualificata secondo l'asimmetria della curva.

Si supponga che sia a ad aumentare e che $a < b$ sia prima che dopo la variazione. Se, in valore assoluto, Z rimane elevato, i maggiori beneficiari del trasferimento dai gruppi più ricchi sono soprattutto i redditeri intermedi. Nel tendere di Z a zero migliorerà anche la posizione dei gruppi più poveri mentre i redditeri della fascia medio-alta vedranno prima rallentare il loro miglioramento per poi subire un peggioramento. Tale andamento si accentua se l'aumento è stato tale che ora $a > b$. In questo caso i peggioramenti si dilatano ai gruppi intermedi e risultano favoriti i gruppi a reddito medio basso o basso.

Si supponga che ora sia b ad aumentare e che si abbia $a > b$ sia prima che dopo il cambiamento. L'effetto della minore concentrazione riguarda in essenza le fasce di reddito medio-alte e solo con la progressiva riduzione di Z il miglioramento si allarga ai redditi intermedi. I gruppi a reddito basso non risentono apprezzabilmente di queste modifiche.

Quando l'aumento è al punto che $a < b$ non sono più solo i redditi estremi superiori a cedere quote, ma anche quelli inferiori, il tutto a favore delle fasce di reddito centrali.

5. Applicazioni del modello.

Il completamento del modello $L(p;d)$ richiede il calcolo dei parametri (d_1, d_2, \dots, d_m) . La tecnica usuale è quella dei minimi quadrati, ovvero la determinazione del vettore $d \in D$ cui corrisponde il minimo della funzione

$$Q(d) = \sum_{i=1}^m w_i [q_i - L(p_i; d)]^2, \quad w_i \geq 0, \quad \sum_{i=1}^m w_i = 1 \quad (5.1)$$

dove le n coppie (p_i, q_i) sono i punti di svolta della spezzata di Lorenz (esclusi gli estremi). Va detto subito che la curva di Pareto generalizzata è intrinsecamente non lineare in entrambi i parametri⁽²²⁾. Ciò non costituisce tuttavia il problema principale. Le difficoltà maggiori provengono dalla presenza dei vincoli che riducono lo spazio parametrico al rettangolo unitario.

Il trattamento dei vincoli si apre a diverse strategie:

- A) Non tenerne conto nella convinzione che se il modello è quello giusto i vincoli saranno “naturalmente” soddisfatti. Anzi sfruttare tale presupposto per una eventuale discriminazione fra modelli alternativi.
- B) Inserire i vincoli in forma implicita: come pseudo-osservazioni addizionali (Faliva, 1973); con una funzione di penalizzazione (Fletcher, 1970); oppure a mezzo di una ridefinizione dei parametri (Box, 1970).
- C) Coinvolgere esplicitamente i vincoli applicando una qualche tecnica di programmazione non lineare, ad esempio Rosen (1960).

La strategia A) è quella che comporta minori complicazioni e forzature. Produce inoltre l'adattamento migliore⁽²³⁾.

Le capacità di adattamento della curva (4.1) sono state saggiate in due diversi contesti

- 1) L'interpolazione dei decili della distribuzione del reddito disponibile alle famiglie (o assimilabile) relativi a nazioni diverse⁽²⁴⁾.
- 2) L'interpolazione dei decili della distribuzione del reddito familiare in Italia dal 1967 al 1983⁽²⁵⁾.

Nelle tabelle 1, 2, e 3 vengono esposti i risultati delle elaborazioni relative alla stima dei parametri ed al calcolo dell'indice di Gini e di Zanardi.

La procedura di stima non lineare⁽²⁶⁾ non ha mai presentato problemi di convergenza o di sconfinamento dei vincoli. Tali problemi non sono però da escludere, soprattutto per curve di Lorenz empiriche molto vicine alle estremanti (3.8) e (3.9). L'adattamento del modello, misurato come coefficiente di correlazione fra decili osservati e decili stimati (colonna R^2 nelle tabelle) è risultato sempre molto soddisfacente, al punto da giustificare il ricorso alla forma analitica per tutto quel che concerne la distribuzione dei redditi.

La tabella 1 mostra due tendenze abbastanza marcate: quando da paesi in via di sviluppo si passa a paesi più sviluppati

- I) La concentrazione generale diminuisce.

- II) La concentrazione fra i “poveri” diminuisce meno rapidamente che fra i “non poveri” sia nelle fasi iniziali che avanzate dello sviluppo. Nei momenti di crescita piu’ regolari la dinamica si inverte e la disparità decade piu’ lentamente fra i non poveri⁽²⁷⁾.

La tesi che accomuna queste due tendenze vede un ruolo via via piu’ dominante dei percettori di quote intermedie di reddito. In tal senso i paesi sudamericani, caratterizzati da valori elevati di G e valori negativi di Z , sarebbero più indietro nello sviluppo rispetto ai paesi africani che ad analoghi valori di G affiancano valori positivi di Z . Nei primi infatti risultano pressoché assenti le fasce intermedie. La progressiva crescita delle classi centrali porterà poi al ripristino della asimmetria negativa, ma a livelli piu’ bassi di concentrazione.

La tendenza risulta in gran parte confermata nei dati temporali dell’Italia. È qui netta la diminuzione della concentrazione mentre la curva re sta sostanzialmente simmetrica fino al (1980). Questo si spiegherebbe con il lento ma sicuro aumento dei percettori di redditi medi.

Tabella 2-Stime dei parametri e calcolo degli indici per i dati italiani. Fonte: Banca

Anno	R ²	a	b	Z	G
67	0.9951	0.6484	0.6457	0.00	0.392
68	0.9950	0.6279	0.6536	-0.02	0.400
69	0.9981	0.6654	0.6398	0.02	0.386
70	0.9977	0.6397	0.6385	0.00	0.402
71	0.9939	0.6315	0.6618	-0.02	0.393
72	0.9968	0.6197	0.6860	-0.04	0.387
73	0.9966	0.6449	0.5291	0.03	0.428
74	0.9983	0.6475	0.6073	0.03	0.417
75	0.9999	0.6686	0.6257	0.03	0.393
76	0.9994	0.6712	0.6838	-0.01	0.356
77	0.9997	0.6712	0.6587	0.01	0.371
78	0.9994	0.6856	0.6666	0.01	0.358
79	0.9997	0.6594	0.6625	0.00	0.376
80	0.9983	0.7112	0.6159	0.07	0.375
81	0.9995	0.7060	0.6913	0.01	0.331
82	0.9982	0.6888	0.7228	-0.03	0.323
83	0.9983	0.6814	0.7033	-0.02	0.339

Tabella 3. Stime dei parametri e calcolo degli indici per i dati italiani

Anno	R ²	a	b	Z	G
1980	0.9996	0.7094	0.7107	0.0000	0.3170
81	0.9997	10.6898	0.7413	-0.0400	0.3120
82	0.9996	0.6984	0.7343	-0.0430	0.3100
83	0.9983	0.6910	0.7888	-0.0800	0.2850

Dal 1980 in poi⁽²⁸⁾ la caduta della concentrazione si allinea ad una asimmetria crescente in senso negativo. Questo puo’ facilmente interpretarsi dicendo che laddove in precedenza l’aumento di importanza delle classi centrali era affiancato da un lieve, ma

significativo incremento delle fasce piu' basse, quest'ultime ora si deprimono perdendo sempre piu' il loro peso nella ripartizione del reddito.

La relazione tra sviluppo economico e ripartizione del reddito non ha ancora ricevuto una sistemazione definitiva. I risultati di questo articolo, pur molto vicini a quelli di Kakwani (1980), aggiungono un contributo autonomo, per certi aspetti integrativo, ma che deve essere ulteriormente rì scontrato in piani di indagine piu' estesi ed accurati.

Notes

(1) Ringrazio i proff. G. Gagliani, C. Vitale e M. Zenga per aver letto una versione precedente e per gli utili suggerimenti critici.

(2) Kakwani e Podder (1973) hanno il merito della riscoperta di questo argomento che già Gini, nel 1932, aveva studiato ed impostato.

(3) Maddala e Singh (1977) hanno espresso forti riserve sulla interpolazione della spezzata di Lorenz a causa dei troncamenti, non sempre plausibili, nel campo di variazione del reddito per i modelli di ripartizione ottenuti a partire dalla curva di concentrazione.

(4) Il riferimento piu' diretto è Taguchi (1967).

(5) Le prime tre condizioni sono ben note (cfr. Iritani e Kuga, 1983). Le condizioni 4 e 5 non sono ancora apparse nella letteratura sulle forme funzionali della curva di Lorenz. Si vedano ad esempio Aggarwal (1984), Gupta (1984) ed Arnold (1986).

(6) Una funzione convessa definita su $[0,1]$ è qui anche continua (Rockafeller, 1970, p. 82)

(7) Il riferimento piu' diretto è Taguchi (1967).

(8) Tale sistema è stato originariamente proposto da Gini (1932) e ripreso, su suggerimento di Arnold Zellner, da Kakwani e Podder (1976).

(9) Nelle formule che seguono è sottointesa la dipendenza degli indici dal vettore di parametri d .

(10) Kakwani (1980) presenta la trasformata lineare $S'=(S-\sqrt{2})/(2-\sqrt{2})$ come *new inequality measure*. L'indice è invece noto da diverso tempo. Cfr. E. Lombardo (1969).

(11) L'indice g_p è noto nella letteratura anglosassone come *Minimal Majority*. Si veda Alker e Russel (1964).

(12) Se le $\{X_i, i=1,2, \dots, n\}$ sono le modalità (supposte non nulle) della serie, ognuna presentatasi con frequenza f_i , l'antiserie avrà modalità $X_i^+=1/X_{n-i+1}$ con frequenza $f_i^+=X_{n-i+1}f_{n-i+1}$. Le curve di Lorenz associate alla serie ed all'antiserie sono mutualmente simmetriche secondo la definizione data da Taguchi (1968)

(13) Gini (1932) presentò la differenza (p_c-q_c) come una ulteriore misura di concentrazione. Zanardi (1965) ha dimostrato come (p_c-q_c) corrisponda al rapporto tra lo scostamento semplice medio da y_c e la somma di y_c alla media aritmetica.

(14) Tale curva, per G fissato, è quella in cui la frazione di redditi con reddito nullo è massima.

(15) Tale curva, per G fissato, è quella in cui la quota di reddito di pertinenza della prima metà di redditi è massima.

(16) La curva di Lorenz viene detta culminante quando, con riferimento al sistema di coordinate (2.10), l'altezza massima si ha in corrispondenza dell'ascissa mediana (La definizione è di Gini). Le curve simmetriche sono sempre culminanti, ma il converso non è vero.

(17) In realtà Hagerbaumer (1977) ha proposto $M=1-\{(2-G)/[G(1-G)]\}^{2A[p_k;d] - p_k q_k}$ con (p_k, q_k) punto di intersezione della curva di Lorenz con la (3.8). La ragione della formula (3.12) è di associare valori crescenti di H a posizioni man mano migliori dei relativamente poveri.

(18) Il modello di ripartizione de reddito che genera la (4.1) è ovviamente quello di Pareto per $a=1$. Se invece si ha $b=1$ il modello associato è $F(y)=[(a/\mu)y]^{(1-a)/a}$ con μ media aritmetica dei redditi. Se poi $a=b$ la (4.1) discende da un caso particolare del modello di Burr: $F(y)=[1+(y/\mu)^{-1/(1-b)}]^{-b}$. Negli altri casi non è disponibile una formula esplicita per la $F(y)$.

(19) La valutazione degli indici per la (4.1) richiede un uso esteso di tecniche di analisi numerica L'autore è disponibile per fornire copia del programma (in BASIC) che effettua il calcolo degli indici.

(20) Lo si puo' dimostrare usando le formule ricorsive della Beta incompleta discusse ad esempio da Majumder e Bhattacharjee (1973).

(21) Il ruolo dei parametri a e b suggerirebbero un pratico indice di asimmetria specifico della (4.1) e cioè $(a-b)$. Un tale indice avrebbe le stesse caratteristiche degli indici di Taguchi e di Zanardi, mancherebbe però della stessa naturale interpretazione.

(22) A causa della non linearità, gli stimatori ottenuti con i minimi quadrati non godono di proprietà ottimali nei piccoli campioni. La loro teoria è soprattutto asintotica. (Wu, 1981).

(23) Cfr. Tarsitano (1986).

(24) Sono stati usati i dati di van Ginneken e Park (1984, p. 108 e p. 176).

(25) Sono state usate due diverse fonti: I bollettini statistici della Banca d'Italia di cui al riferimento D. Gressani (1984) e le indagini ISTAT.

(26) Il programma adoperato è il P3R del package BMDP, versione 1981, con pesi uguali per ogni osservazione e con valori iniziali 0.5 per entrambi i parametri.

(27) Vale a dire che la curva di Lorenz cambia due volte asimmetria: da negativa a positiva e viceversa.

(28) Da notare la difformità delle stime, conseguenza di quella presente nelle fonti, per i dati dall'80 in poi.

Bibliografia

Abramowitz M., Stegun I.A. (1964). Handbook of mathematical functions. Dover, New York.

Aggarwal V. (1984). On optimal aggregation of income distribution data". *Sankhya, B*, 46, 343-355.

Arnold B.C. (1983). Pareto distributions. International Co-operative Publishing House, Fairland, Ma.

Arnold B.C. (1986). A class of hyperbolic Lorenz curves. *Sankhya, B*, 48

Box M.J. (1966). A comparison of several current optimization methods and the use of transformations in constrained problems. *Computre Journal*, 9, 67-77

Faliva M. (1973) Stimatori lineari ottimali dei parametri in un modello di regressione lineare con vincoli di disuguaglianza. *Statistica*, 33, 363-394

Elteto O., Frigyes F. (1968). New inequality measures as efficient tools for causal analysis and planning. *Econometrica*, 36, 383-396.

Fletcher R. (1970). A class of methods for non linear programming with termination and convergence properties. In Integer and non linear programming edito da J. Abadie. North-Holland, Amsterdam.

Gini C. (1932). Intorno alle curve di concentrazione. *Metron*, anno 9, 3-76

van Ginneken W., Park J.G. (1984). Generating internationally comparable income distribution estimates. International Labour Office, Ginevra.

Gressani D. (1984). I bilanci delle famiglie italiane nell'anno 1983
Bollettino Statistico della Banca D'Italia, anno 39

Gupta M.R. (1984). Functional form for estimating the Lorenz curve. *Econometrica*, 52, 1313-1314.

Hagerbaumer J.B. (1977). The Gini concentration ratio and the minor concentration ratio: A two parameter index of inequality. *The Review of Economics and Statistics*, 59, 377-379.

Iritani J, Kuga K. (1983). Duality between the Lorenz curve and the income distribution function. *Economic Studies Quarterly*, 34, 9-21.

ISTAT (1983). La distribuzione quantitativa del reddito in Italia nelle indagini sui bilanci di famiglia anni 1980-1981-1982. Supplemento al Bollettino Mensile di Statistica n.25.

ISTAT (1985). La distribuzione quantitativa del reddito in Italia nelle indagini sui bilanci di famiglia anno 1983.
Supplemento al Bollettino Mensile di Statistica n.2.

- Kakwani N.C., Podder N. (1973). On the estimation of Lorenz curves from grouped observations. *International Economic Review*, 14, pp. 278-292.
- Kakwani N.C., Podder N. (1976). Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations. *Econometrica*, 44, 137-148.
- Kakwani N.C. (1980). Income inequality and poverty: method of estimation and policy application. Oxford University Press, Oxford
- Lombardo E. (1969). Nota sulla concentrazione secondo Gini-Lorenz. Istituto di Statistica dell'Università di Roma, Quaderno n. 5.
- Maddala G.S., Singh S.K. (1977). Estimation problem in size distribution of incomes". *Economie Appliquée*, 30,461-480.
- Majumder K.L., Bhattacharjee G.P. (1973). Algorithm AS 63: The incomplete Beta integral". *Applied Statistics*, 22, 409-411.
- Pietra G. (1914-1915). Delle relazioni tra gli indici di variabilità. Nota I. *Atti del Reale Istituto Veneto di S.L.A.* tomo 74, parte II, 775-792,793-804.
- Raasche R.H., Gaffney J., Koo A.Y.C., Obst N. (1980) Functional forms for estimating the Lorenz curve. *Econometrica*, 48,1061-1062.
- Rockafellar R. T. (1970). *Convex Analysis*. Princeton University, Princeton, NJ.
- Rosen J.B. (1960). The gradient projection method for non linear programming. Part I: linear constraints. *SIAM Journal of applied Mathematics*, 8,181-217.
- Taguchi T. (1967). On some properties of concentration curve and its application. *Metron*, anno 17, 381-395.
- Taguchi T. (1968). Concentration-curve methods and structures of skew population. A methodology for the analysis of economic data. *Annals of institute of Statistical Mathematics*, 20,107-141.
- Tarsitano A. (1986). Regressioni lineari con vincoli: uno studio di simulazione sugli stimatori misti a confronto con gli stimatori vincolati. In *Studi in onore di Francesco Brambilla*, vol. 2, pp. 831-842, Edizioni Bocconi Comunicazioni, Milano
- Wu C. (1981). Asymptotic theory of non linear least squares estimation. *The Annals of Statistics*, 9,501-513.
- Zanardi G. (1965). L'asimmetria statistica delle curve di concentrazione. *Ricerche Economiche*,19, 355-396.

