STATISTICA, anno LXIV, n. 3, 2004

LA SCOMPOSIZIONE DELL'INDICE DI GINI NEL CASO DI DUE GRUPPI

Michele Costa

1. INTRODUZIONE

L'indice di concentrazione di Gini (Gini, 1914, 1939) rappresenta una delle misure di disuguaglianza maggiormente diffuse sia nel contesto degli studi teorici, sia nell'ambito delle analisi empiriche. Per un approfondimento delle numerose ragioni di questo successo si rimanda all'imponente letteratura sviluppatasi intorno a tale indicatore (Atkinson, 1970; Dagum, 1980, 1987; Frosini, 1985; Giorgi, 1999; Zenga, 1987); a solo titolo di esempio delle potenzialità dell'indice di Gini, si cita la possibilità di includere anche valori negativi della variabile oggetto di studio: ciò costituisce un innegabile vantaggio rispetto alle misure entropiche di disuguaglianza, come gli indici di Theil, di Hirschman-Herfindahl e di Bourguignon (Bourguignon, 1979; Theil, 1967).

Una delle caratteristiche di maggiore interesse delle misure di disuguaglianza è rappresentata dalla possibilità di scomporre la disuguaglianza complessiva nelle sue diverse componenti, così da poterne correttamente valutare la struttura e individuarne i principali fattori causali.

Nel caso di una popolazione suddivisa in k gruppi, una tradizione ormai consolidata vuole che la scomposizione degli indici di disuguaglianza sia ottenuta nell'ambito dell'analisi della varianza, misurando, pertanto, la disuguaglianza "tra" i k gruppi confrontando soltanto le rispettive k medie. A partire dal lavoro di Bhattacharya e Mahalanobis del 1967, anche la scomposizione dell'indice di Gini viene generalmente effettuata in questo contesto.

Nella storia della scomposizione delle misure di disuguaglianza, un ruolo di rilievo è svolto dalla proprietà di scomposizione additiva, introdotta da Shorrocks nel 1980 (Cowell, 1980; Foster *et al.*, 1984; Shorrocks, 1980). Poiché gli indici di entropia, a differenza dell'indice di Gini, godono di questa proprietà, numerosi Autori preferiscono far ricorso a misure di disuguaglianza entropiche.

È, tuttavia, necessario sottolineare come misurare la disuguaglianza "tra" i gruppi soltanto sulla base delle medie aritmetiche, porti a trascurare altre fondamentali caratteristiche differenzianti numerose variabili, tra le quali la ricchezza e il reddito. Le distribuzioni del reddito (e della ricchezza) in gruppi diversi della popolazione risultano, infatti, tipicamente differenti non solo per quanto riguarda le medie, ma anche, e in modo rilevante, per altri aspetti, come varianze, asimmetrie e curtosi. A questo proposito, è appena il caso di ricordare che anche l'analisi della varianza presuppone l'omoschedasticità. Risulta, così, sorprendente e inspiegabile osservare come la differenza tra le varianze dei gruppi non venga usualmente né considerata e, perfino, neppure citata.

È, inoltre, possibile dimostrare come le misure della disuguaglianza "tra" ottenute nell'ambito degli indici di entropia, ad esempio mediante l'indice di Theil, pur additivamente scomponibili, portino ad una sottostima dell'effettivo contributo alla disuguaglianza complessiva riconducibile alle differenze tra i gruppi (Dagum, 1997*b*; Costa e Morales, 1997).

Per quanto elegante e interessante dal punto di vista matematico, la proprietà di scomposizione additiva sembra essere adottata in modo spesso acritico, soprattutto se si tiene conto delle implicazioni logiche relative alla misura della distanza tra gruppi, e delle stesse caratteristiche delle distribuzioni delle variabili all'interno dei gruppi. In entrambi i casi, una misura della variabilità "tra" ottenuta soltanto in funzione delle medie dei gruppi può non essere ritenuta soddisfacente.

Una scomposizione dell'indice di Gini in grado di assicurare una misura della disuguaglianza "tra" corretta e completa è stata proposta da Dagum (1997*a*), prendendo in considerazione non solo le medie dei gruppi, ma tutte le differenze tra le distribuzioni dei gruppi.

Partendo da questa scomposizione, e riprendendo una semplificazione della disuguaglianza "tra" per il caso di due gruppi (Costa, 2004), nel presente scritto viene introdotta una misura della disuguaglianza "tra" in grado di valutare anche gli effetti dovuti alla presenza di transvariazione.

2. IL QUADRO TEORICO

Data una popolazione di n unità, disaggregata in k gruppi, ed essendo n_j la numerosità del gruppo j-esimo, con $n = \sum_{j=1}^{k} n_j$, è possibile rappresentare l'indice di Gini come

$$G = \frac{1}{2n^2 \overline{y}} \sum_{j=1}^{k} \sum_{h=1}^{k} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} | y_{ji} - y_{hr} |$$

dove $\{y_1, y_2, ..., y_n\}$ è il vettore delle osservazioni sulla variabile di interesse (reddito, ricchezza, ecc.), di media $\overline{y}$, k è il numero dei gruppi, n_j e n_h sono rispettivamente le numerosità dei gruppi j e h.

Indicando con $p_j = n_j / n$ la quota della popolazione j-esima e con $s_j = p_j \overline{y}_j / \overline{y}$ la rispettiva quota di reddito, l'indice di Gini tra i gruppi j e h può essere espresso come

$$G_{jh} = \frac{1}{n_j n_h (\overline{y}_j + \overline{y}_h)} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} | y_{ji} - y_{hr} |$$

Usando la misura G_{jh} è possibile ottenere l'indice di Gini per l'intera popolazione grazie ad una media ponderata, con pesi $p_j s_h$

$$G = \sum_{j=1}^{k} \sum_{h=1}^{k} G_{jh} p_j s_h$$

In questo contesto si può facilmente ricavare la componente di disuguaglianza "entro" i gruppi come

$$G_w = \sum_{j=1}^{k} G_{jj} p_j s_j$$

che è la media degli indici di Gini dei *k* gruppi ponderata con $p_j s_j$.

Questa misura della disuguaglianza "entro" i gruppi è ampiamente diffusa in letteratura, mentre la disuguaglianza "tra" i gruppi è oggetto di un ampio dibattito, relativo al ruolo, alla natura e agli obiettivi della scomposizione delle misure di disuguaglianza (Dagum, 1980; Deutsch e Silber, 1999; Frosini, 1989).

La scomposizione dell'indice di Gini è un tema che suscita da tempo l'interesse e l'attenzione del mondo scientifico e sono numerosi gli Autori che hanno proposto, a partire dal contributo di Soltow (1960), specifiche misure per valutare la componente di disuguaglianza "tra" i gruppi: Bhatthacharya e Mahalanobis (1967), Civardi (1987), Ferrari e Rigo (1987), Mehran (1975), Podder (1993*a,b*, 1997), Pyatt (1976), Rao (1968), Schechtman e Yitzhaki (1987, 1999), Yao (1997) solo per citarne alcuni.

Le diverse proposte si differenziano tra loro in funzione sia delle finalità rispetto alle quali sono state sviluppate, sia delle modalità di analisi dell'interazione tra i gruppi. Generalmente, la componente dell'indice di Gini relativa alla disuguaglianza "tra" i gruppi, G_b, viene ottenuta attraverso varianti della proposta di Bhatthacharya e Mahalanobis (1967)

$$G_b = \sum_{j=1}^{k} \sum_{\substack{h=1 \\ j \neq h}}^{k} p_j p_h | \overline{y}_j - \overline{y}_h |$$

Una proposta alternativa (Dagum, 1997*a,b,c*; Mehran, 1975) per misurare il contributo alla disuguaglianza complessiva attribuibile alle differenze tra i *k* gruppi è data da

$$G_b = \sum_{j=1}^{k} \sum_{\substack{h=1 \\ j \neq h}}^{k} G_{jh} p_j s_h$$

Si tratta di una quantità che non considera solo le differenze tra le medie dei k gruppi, ma anche tutti gli altri possibili elementi di diversità.

Dagum suggerisce una ulteriore scomposizione della disuguaglianza "tra" distinguendo disuguaglianza netta "tra" e contributo dovuto alla presenta di transvariazione; una discussione più analitica è sviluppata nei lavori citati in bibliografia.

3. IL CASO DI DUE GRUPPI

L'analisi e la misura della disuguaglianza in una popolazione suddivisa in due soli gruppi risulta centrale in una vasta antologia di importanti situazioni che interessano settori anche estremamente diversi tra loro.

Facendo riferimento al quadro delineato nel paragrafo precedente, la scomposizione dell'indice di Gini, per il caso di due gruppi, in disuguaglianza "entro", G_w, e "tra", G_b, risulta

$$G = G_w + G_b = (G_1 p_1 s_1 + G_2 p_2 s_2) + (G_{12} p_1 s_2 + G_{21} p_2 s_1) \tag{1}$$

È possibile ottenere (Costa, 2004) una forte semplificazione della componente $(G_{12} p_1 s_{12} + G_{21} p_2 s_1)$, relativa alla misura del contributo alla disuguaglianza complessiva attribuibile alle differenze tra i due gruppi.

Quando la popolazione è divisa in due gruppi si ha

$$p_2 = 1 - p_1; \quad s_2 = 1 - s_1; \quad G_{12} = G_{21}$$

e, ipotizzando che i poveri siano nel primo gruppo e i ricchi nel secondo, cioè assumendo $| y_{1i} - y_{2r} | = (y_{2r} - y_{1i}) \geq 0 \quad \forall i, r$, la misura della componente di disuguglianza "tra" può essere semplificata attraverso i seguenti elementari passaggi

$$
\begin{aligned}
G_b &= G_{12} p_1 s_2 + G_{21} p_2 s_1 = G_{12}(p_1 s_2 + p_2 s_1) = \\
&= \frac{p_1 s_2 + p_2 s_1}{n_1 n_2 (\bar{y}_1 + \bar{y}_2)} \sum_{i=1}^{n_1} \sum_{r=1}^{n_2} | y_{1i} - y_{2r} | = \frac{p_1 s_2 + p_2 s_1}{n_1 n_2 (\bar{y}_1 + \bar{y}_2)} \sum_{i=1}^{n_1} \sum_{r=1}^{n_2} (y_{2r} - y_{1i}) = \\
&= \frac{p_1 s_2 + p_2 s_1}{n_1 n_2 (\bar{y}_1 + \bar{y}_2)} n_1 n_2 (\bar{y}_2 - \bar{y}_1) = \frac{p_1 s_2 + p_2 s_1}{(\bar{y}_1 + \bar{y}_2)} (\bar{y}_2 - \bar{y}_1) = \\
&= p_1 (s_2 + p_2 \bar{y}_1 / \bar{y}) \frac{(\bar{y}_2 - \bar{y}_1)}{(\bar{y}_1 + \bar{y}_2)} = p_1 p_2 \frac{(\bar{y}_2 + \bar{y}_1)}{\bar{y}} \frac{(\bar{y}_2 - \bar{y}_1)}{(\bar{y}_1 + \bar{y}_2)} = \\
&= p_1 p_2 \bar{y}_2 / \bar{y} - p_1 p_2 \bar{y}_1 / \bar{y} = p_1 (1 - s_1) - (1 - p_1) s_1 = \\
&= p_1 - p_1 s_1 - s_1 + p_1 s_1 = p_1 - s_1
\end{aligned}
$$

La disuguaglianza "tra" viene così misurata come differenza tra p_1, la quota di popolazione del gruppo più povero, e s_1, la rispettiva quota di reddito.

Un risultato analogo è ottenuto da Yitzhaki (2002), scomponendo l'indice di Gini nell'indice di Sen e in una serie di altre componenti.

Questa semplificazione della disuguaglianza "tra", pur risultando corretta nel rispetto delle premesse, richiede un aggiustamento in presenza di transvariazione tra i due gruppi, come si vedrà nel prossimo paragrafo.

Peraltro, in assenza di transvariazione e per una popolazione suddivisa in due gruppi, l'indice di Gini risulta scomponibile come

$$G = G_w + G_b = (G_1 p_1 s_1 + G_2 p_2 s_2) + (p_1 - s_1) \tag{2}$$

e il contributo alla disuguaglianza complessiva riconducibile alle differenze tra i due gruppi risulta misurato in modo più esaustivo rispetto alla tradizionale semplice differenza tra le medie dei due gruppi.

4. IL RUOLO DELLA TRANSVARIAZIONE

Il concetto di transvariazione, introdotto e approfonditamente studiato da Corrado Gini (per una discussione dettagliata si veda Gini, 1959; Dagum, 1959), risulta cruciale nell'analisi della disuguaglianza, e proprio alla presenza di transvariazione possono essere ricondotte molte delle differenze che caratterizzano le diverse scomposizioni dell'indice di Gini proposte nel corso del tempo e alle quali si è accennato nel paragrafo 2.

Si ha transvariazione quando la differenza tra i valori di almeno due unità appartenenti a gruppi diversi presenta un segno diverso dalla differenza tra le medie dei due gruppi.

Al fine di valutare gli effetti sulla disuguaglianza complessiva, riconducibili alla presenza di transvariazione, si fa ricorso ad un semplice esempio, illustrato in tavola 1.

Un insieme di n=50 unità viene diviso in due gruppi, di numerosità, rispettivamente, n_1=35 e n_2=15, prima in assenza di transvariazione, e, successivamente, ammettendone la presenza.

Nel primo caso, la disuguaglianza "entro", G_w, rappresenta il 31.7% della disugualianza complessiva, valore che sale al 51.9% nel secondo caso. La presenza di transvariazione porta, pertanto, ad un aumento del peso della disuguaglianza "entro" sulla disuguaglianza complessiva.

Per quanto riguarda la parte di disuguaglianza non "spiegata" da G_w, $G - G_w$, nel primo caso risulta $G - G_w = G_b = p_1 - s_1$, mentre, nel secondo caso, si ha $G - G_w = G_b + G_t$, dove G_t è il contributo alla disuguaglianza complessiva riconducibile alla presenza di transvariazione.

TAVOLA 1

Effetti della transvariazione

	Popolazione complessiva		Senza transvariazione I gruppo		Senza transvariazione II gruppo		Con transvariazione I gruppo		Con transvariazione II gruppo	
	y_i	n_i	y_i	n_i	y_i	n_i	y_i	n_i	y_i	n_i
	1	5	1	5	1	0	1	5	1	0
	2	8	2	8	2	0	2	7	2	1
	3	9	3	9	3	0	3	8	3	1
	4	7	4	7	4	0	4	5	4	2
	5	6	5	6	5	0	5	3	5	3
	6	5	6	0	6	5	6	2	6	3
	7	4	7	0	7	4	7	2	7	2
	8	3	8	0	8	3	8	2	8	1
	9	2	9	0	9	2	9	1	9	1
	10	1	10	0	10	1	10	0	10	1
$\overline{y}_i$	4.32		3.03		7.33		3.68		5.80	
p_i	1		0.70		0.30		0.70		0.30	
s_i	1		0.49		0.51		0.60		0.40	
G	0.3065		0.2415		0.0933		0.3216		0.2038	
G_w			0.0972				0.1591			
G_b			0.2093				0.1251			
G_t			0				0.0223			

Dalla tavola 1 si ha, nel primo caso, $G_b = 0.2093 = p_1 - s_1 = 0.7 - 0.4907$.

Per l'esame del secondo caso, facendo riferimento alla scomposizione (1) dell'indice di Gini per una popolazione suddivisa in due gruppi, si ha:

$$G - G_w = (G_{12} p_1 s_2 + G_{21} p_2 s_1) = \frac{p_1 s_2 + p_2 s_1}{n_1 n_2 (\overline{y}_1 + \overline{y}_2)} \sum_{i=1}^{n_1} \sum_{r=1}^{n_2} | y_{1i} - y_{2r} |$$

La presenza di transvariazione porta a distinguere differenze $(y_{2r} - y_{1i})$ positive, relative alla disuguaglianza "tra" i gruppi, e differenze $(y_{2r} - y_{1i})$ negative, relative alla presenza di transvariazione.

Posto $\overline{y}_2 > \overline{y}_1$, è possibile ottenere la componente di disuguaglianza "tra" i gruppi come

$$G_b = \frac{p_1 s_2 + p_2 s_1}{n_1 n_2 (\overline{y}_1 + \overline{y}_2)} \sum_{i=1}^{n_1} \sum_{r=1}^{n_2} (y_{2r} - y_{1i}) \qquad \forall (y_{2r} - y_{1i}) \geq 0 \tag{3}$$

e il contributo alla disuguaglianza complessiva riconducibile alla presenza di transvariazione come

$$G_t = -\frac{p_1 s_2 + p_2 s_1}{n_1 n_2 (\overline{y}_1 + \overline{y}_2)} \sum_{i=1}^{n_1} \sum_{r=1}^{n_2} (y_{2r} - y_{1i}) \qquad \forall (y_{2r} - y_{1i}) < 0 \tag{4}$$

Nel paragrafo precedente, il risultato $G_b = p_1 - s_1$ assume come premessa $(y_{2r} - y_{1i}) \geq 0 \quad \forall i, r$. Si tratta di una ipotesi semplificatrice, che esclude la presenza di transvariazione.

Nel caso di grandezze transvarianti, una o più differenze $(y_{2r} - y_{1i})$ risultano negative e, di conseguenza, la quantità $(p_1 - s_1)$, uguale alla somma di tutte le differenze $(y_{2r} - y_{1i})$, sia positive, sia negative, sottostima la disuguaglianza tra (3), dove le differenze $(y_{2r} - y_{1i})$ sono tutte positive. Dalla tavola 1 si ha $(p_1 - s_1) = (0.70 - 0.5972) < G_b = 0.1251$.

È, tuttavia, ancora possibile ottenere una misura della disuguaglianza "tra" più semplice ed immediata della (3): a tal fine è sufficiente aggiungere a $(p_1 - s_1)$ il valore delle differenze $(y_{2r} - y_{1i})$ negative, cioè l'effetto sulla disguaglianza attribuibile alla presenza di transvariazione.

La disuguaglianza "tra" (3) può essere così semplificata nella seguente espressione

$$G_b = p_1 - s_1 + G_t$$

Dall'esempio precedente risulta:

$$G_b = 0.1251 = p_1 - s_1 + G_t = 0.7 - 0.5972 + 0.0223 .$$

La scomposizione (2) dell'indice di Gini diventa pertanto

$$G = G_w + G_b + G_t = (G_1 p_1 s_1 + G_2 p_2 s_2) + (p_1 - s_1 + G_t) + G_t \tag{5}$$

È facile osservare come, dati G, G_1 e G_2, dalla (5) sia possibile ricavare G_t

$$G_t = (G - G_w - p_1 - s_1)/2 \tag{6}$$

Rispetto alla formula tradizionale (4) per la misura del contributo della transvariazione alla disuguaglianza complessiva, la relazione (6) rappresenta una semplificazione rilevante.

Questa espressione per G_t consente, inoltre, di ricavare la disuguaglianza "tra" (3) conoscendo soltanto G, G_1 e G_2 come

$$G_b = p_1 - s_1 + (G - G_w - p_1 - s_1)/2 \tag{7}$$

Anche per quanto riguarda la misura della disuguaglianza "tra", è immediato osservare come, rispetto alla tradizionale formula (3), la relazione (7) rappresenti una notevole semplificazione.

Nella misura della disuguaglianza "tra", la differenza $(p_1 - s_1)$ rimane l'elemento principale e preponderante, ma, accanto ad essa, viene introdotta una correzione che tiene conto dell'eventuale presenza di transvariazione.

Nell'esempio di tavola 1, $(p_1 - s_1)/G_b = 0.1028/0.1251 = 0.82$. Vale a dire che l'82% della disuguaglianza "tra" è dato dalla differenza tra la quota di popolazione del gruppo più povero e la rispettiva quota di reddito. Si tratta di un semplice esempio, che risulta tuttavia rappresentativo di numerose situazioni reali, nelle quali la quantità $(p_1 - s_1)$ rappresenta la componente essenziale della disuguaglianza "tra".

5. CONCLUSIONI

La misura della disuguaglianza "tra" ottenuta nell'ambito della scomposizione dell'indice di Gini permette di evitare le principali difficoltà legate alle scomposizioni tradizionali delle misure di disuguaglianza e, allo stesso tempo, consente di valutare correttamente e di confrontare tra loro gli effetti di differenti fattori sulla disuguaglianza complessiva.

Nel caso di una popolazione suddivisa in due gruppi, la scomposizione dell'indice di Gini consente di ottenere una misura semplice ed immediata dei contributi alla disuguaglianza complessiva riconducibili sia alle differenze tra i gruppi, sia alla presenza di transvariazione.

In particolare, è possibile ricavare le tre componenti dell'indice di Gini, ossia disuguaglianza "entro", G_w, disuguaglianza "tra", G_b, e contributo della transvariazione, G_t, conoscendo soltanto l'indice di Gini per l'intera popolazione, G, per il gruppo più povero, G_1, e per il gruppo più ricco, G_2.

Le misure che si ottengono per G_b e per G_t rappresentano, infine, un rilevante miglioramento rispetto alle espressioni finora disponibili in letteratura, in quanto sono caratterizzate da una estrema semplicità e da una notevole agilità calcolatoria.

Dipartimento di Scienze Statistiche
Alma Mater Studiorum - Università di Bologna

MICHELE COSTA

RIFERIMENTI BIBLIOGRAFICI

A.B. ATKINSON, (1970), *On the measurement of inequality*, "Journal of Economic Theory", 2, pp. 224-263.

N. BHATTACHARYA, B. MAHALANOBIS, (1967), *Regional disparities in household consumption in India*, "Journal of the American Statistical Association", 62, pp. 143-161.

F. BOURGUIGNON, (1979), *Decomposable income inequality measures*, "Econometrica", 47, pp. 901-921.

M. CIVARDI, (1987), *Proposte di scomposizione dell'indice di Gini,* in M. ZENGA, *La distribuzione personale del reddito,* Vita e Pensiero, Milano.

M. COSTA, (2004), *Notes on the Gini index decomposition*, Atti della XLII Riunione Scientifica SIS, Bari, pp. 187-190.

M. COSTA, A. FERNANDEZ, (1997), *Livello di istruzione e disuguaglianza nella distribuzione dei redditi: la scomposizione degli indici di Gini e di entropia generalizzata*, "Statistica", 57, pp. 309-323.

F.A. COWELL, (1980), *On the structure of additive inequality measures*, "Review of Economic Studies", 47, pp. 521-531.

C. DAGUM, (1959), *Transvariazione fra più di due distribuzioni,* in C. GINI, *Memorie di metodologia statistica,* vol. II, Libreria goliardica, Roma.

C. DAGUM, (1980), *Inequality measures between income distributions with applications*, "Econometrica", 48, pp. 1791-1803.

C. DAGUM, (1987), *Gini ratio*, in *The New Palgrave Dictionary of Economics*, MacMillan Press, London, 2, pp. 529-532.

C. DAGUM, (1997*a*), *A new decomposition of the Gini income inequality ratio*, "Empirical Economics", 22, pp. 515-531.

C. DAGUM, (1997*b*), *Scomposizione ed interpretazione delle misure di disuguaglianza di Gini e di entropia generalizzata*, "Statistica", 57, pp. 295-308.

C. DAGUM, (1997*c*), *Decomposition and interpretation of Gini and Theil entropy inequality measures*, Proceedings of the Business and Economic Statistics Section, American Statistical Association, pp. 200-205.

J. DEUTSCH, J. SILBER, (1999), *Inequality decomposition by population subgroups and the analysis of interdistributional inequality,* in J. SILBER, *Handbook on income inequality measurement,* Kluwer, Boston, pp. 363-397.

G. FERRARI, P. RIGO, (1987), *Sulla scomposizione del rapporto di concentrazione di Gini,* in M. ZENGA, *La distribuzione personale del reddito,* Vita e Pensiero, Milano.

J.E. FOSTER, J. GREER, E. THORBECKE, (1984), *Notes and comments. A class of decomposable poverty measures*, "Econometrica", 59, pp. 687-709.

B.V. FROSINI, (1985), *Comparing inequality measures*, "Statistica", 45, pp. 299-317.

B.V. FROSINI, (1989), *Aggregate units, within-group inequality and the decomposition of inequality measures,* "Statistica", 44, pp. 349-369.

C. GINI, (1914), *Sulla misura della concentrazione e della variabilità dei caratteri*, Atti del R. Istituto Veneto di Scienze, Lettere e Arti, LXXIII, parte II, pp. 1203-1248.

C. GINI, (1939), *Variabilità e concentrazione*, Giuffrè, Milano.

C. GINI, (1959), *Transvariazione*, Libreria goliardica, Roma.

G.M. GIORGI, (1999), *Income inequality measurement: the statistical approach,* in J. SILBER, *Handbook on income inequality measurement,* Kluwer, Boston, pp. 245-267.

F. MEHRAN, (1975), *A statistical analysis of income inequality based on a decomposition of the Gini index*, Proceedings of the 40th ISI Session, Warsaw, pp. 580-585.

N. PODDER, (1993*a*), *A new method of disaggregating the Gini index by groups*, "Sankhya", Series B, 55.

N. PODDER, (1993*b*), *The disaggregation of the Gini coefficient by factor components and ist application to Australia*, "Review of Income and Wealth", pp. 51-61.

N. PODDER, (1997), *A method of analyzing subgroup poverty using a new decomposition of the Gini coefficient*, "Statistica", 57, pp. 21-30.

G. PYATT, (1976), *On the interpretation and disaggregation of Gini coefficients*, "Economic Journal", 86, pp. 243-255.

V.M. RAO, (1969), *Two decompositions of concentration ratio*, "Journal of the Royal Statistical Society", Series A, 132, pp. 418-425.

E. SCHECHTMAN, S. YITZHAKI, (1987), *A measure of association based on Gini's mean difference*, "Communications in Statistics", pp. 207-231.

E. SCHECHTMAN, S. YITZHAKI, (1999), *On the proper bounds of the Gini correlation*, "Economics Letters ", pp. 133-138.

A.F. SHORROCKS, (1980), *The class of additively decomposable inequality measures*, "Econometrica", 48, pp. 613-625.

L. SOLTOW (1960), *The distribution of income related to changes in the distribution of education, age, and occupation*, "Review of Economics and Statistics", 42, pp. 450-453.

H. THEIL, (1967), *Economics and information theory*, North Holland, Amsterdam.
S. YAO, (1997), *Decomposition of Gini coefficients by income factors: a new approach and application*, "Applied Economics Letters ", 4, pp. 27-31.
S. YITZHAKI, (2002), *Do we need a separate poverty measurement?*, "European Journal of Political Economy", 48, pp. 61-85.
M. ZENGA, (1987), *Concentration measures*, in A. NADDEO, *Italian contributions to the methodology of statistics,* CLEUP, Padova, pp. 42-51.

RIASSUNTO

La scomposizione dell'indice di Gini nel caso di due gruppi

Il rapporto di concentrazione di Gini è da tempo al centro di un intenso dibattito relativo alla scomposizione della disuguaglianza complessiva nelle due componenti di disuguaglianza "entro" e "tra". Tradizionalmente, la disuguaglianza "tra" è misurata seguendo un percorso analogo all'analisi della varianza, e quindi sulla base delle sole medie aritmetiche dei gruppi. Alternativamente, viene suggerito di considerare, nella misura della disuguaglianza "tra", tutti gli elementi di diversità tra le distribuzioni dei gruppi. In questo secondo contesto, il lavoro propone una semplificazione della misura della disuguaglianza "tra". Si dimostra che, in assenza di transvariazione, la disuguaglianza attribuibile alla diversità tra due gruppi è uguale alla differenza tra la quota di popolazione del gruppo più povero e la rispettiva quota nella variabile oggetto di studio. Vengono, inoltre, proposte, anche in presenza di transvariazione, espressioni semplici ed immediate per la misura sia della disuguaglianza "tra", sia del contributo della transvariazione.

SUMMARY

Gini index decomposition for the case of two subgroups

Gini inequality index decomposition into within and between inequality components arises since a long time a controversial, wide and interesting discussion. The measurement of inequality between is traditionally obtained in the framework of analysis of variance as a function of only the means of the subgroups. An alternative methodology to measure inequality between states the possibility to take into account all differences between subgroups distributions. By referring to this new approach, the paper introduces a relevant simplification of the measurement of inequality between and of the transvariation contribution for the case of a population divided in two subgroups. Without transvariation, inequality between can be obtained as the difference of the poor share in the population and the poor share in the variable of interest. Also for the case of two overlapping subgroups simple and straightforward expressions for the transvariation contribution and inequality between are derived.