Società italiana di economia pubblica

XVI CONFERENZA

POLITICHE PUBBLICHE, SVILUPPO E CRESCITA

Pavia, Università, 7 - 8 ottobre 2004

LA DISUGUAGLIANZA MULTIDIMENIONALE: UNA RASSEGNA CRITICA

ERNESTO SAVAGLIO

pubblicazione internet realizzata con contributo della

COMPAGNIA di San Paolo

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

LA DISUGUAGLIANZA MULTIDIMENIONALE: UNA RASSEGNA CRITICA

ERNESTO SAVAGLIO

ABSTRACT. La recente letteratura economica sulla disuguaglianza ha esteso l'analisi dei confronti interpersonali in termini di disparità sociale ad altre caratteristiche individuali oltre al reddito. Le presenti note passano in rassegna i principali filoni di ricerca teorica sulla disuguaglianza multidimensionale.

JEL Classification: D31

AREA TEMATICA: 3.- *Distribuzione del reddito, misurazione dell'impatto delle politiche, analisi delle evidenze empiriche*

1. INTRODUZIONE

La letteratura economica relativa allo studio delle disuguaglianze si è prevalentemente concentrata sugli aspetti unidimensionali del problema. La distribuzione personale del reddito costituisce infatti la variabile rilevante su cui basare la misurazione della disparità economica di una data popolazione. Pare evidente, invece, che per analizzare lo status sociale delle persone le variabili da considerare come rilevanti siano più di una. Sen [30], Kolm [13], Maasoumi [19] e molti altri studiosi hanno in precedenza sottolineato come sia fondamentale analizzare le diverse caratteristiche degli individui, ai fini di misurare e valutare se una data popolazione (o nazione ad esempio), presenti un livello di disuguaglianza più elevata rispetto ad un'altra. Le persone sono diverse per reddito, istruzione, condizioni di salute, ecc. ed è quindi necessario estendere le nostre valutazioni ad un insieme di caratteristiche ulteriori rispetto al solo reddito, se vogliamo rispondere in maniera esaustiva alle domande poste da Sen (1981): "*Why inequality?*" e "*Inequality of What?*". Purtroppo, lo studio della disuguaglianza in un contesto multidimensionale (cioè con più variabili oltre al reddito), è difficile e la letteratura economica e matematica è piuttosto esigua, tecnicamente sofisticata, ma spesso con un basso valore euristico. Dal momento che il problema è complesso, è risultato difficile estendere i risultati ottenuti sul terreno della misurazione della disuguaglianza in termini della sola variabile esplicativa reddito al contesto multivariato. Una delle difficoltà principali risiede nell'ovvia interazione tra gli attributi individuali di tipo reddituale e quelli di tipo qualitativo quali l'istruzione o la salute.

Queste note passano in rassegna la recente letteratura economica riguardo la disuguaglianza multidimensionale e fanno parte di una serie di lavori collegati che analizzano i vantaggi dell'estendere l'analisi e la misurazione della disuguaglianza da un contesto univariato a uno multivariato. Il progetto di ricerca, di cui questo articolo fa parte, si articola in due parti:

Date: 6 Settembre 2004.

Key words and phrases. Ordinamenti, disuguaglianza multidimensionale, disparità sociale.

a: la misurazione teorica del benessere relativo a diverse distribuzioni di caratteristiche individuali;
b: lo studio empirico delle politiche pubbliche rivolte ad ottenere una diminuzione delle disuguaglianze misurate con lo strumento di cui al punto a;

La prima parte (punto a, si veda Savaglio [28]) del mio lavoro quindi consiste nei confronti tra diverse distribuzioni multivariate di individui-attributi. Una funzione di benessere sociale stabilirà la desiderabilità in termini di uguaglianza di una distribuzione rispetto ad un'altra. Successivamente, sarà possibile studiare come empiricamente una tassazione progressiva o l'introduzione di un insieme di beni pubblici migliorino l'uguaglianza individuale a livello di più dimensioni e quindi a livello complessivo.

Il presente lavoro contiene una prima sezione in cui discuto i risultati principali relativi alla misurazione della disuguaglianza unidimensionale. Nella terza sezione, spiego la notazione e introduco le principali definizioni dei criteri di confronto di popolazioni di individui che differiscono sulla base di diverse caratteristiche comuni. La sezione 4 rassegna i tre principali filoni di ricerca sulla disuguaglianza multidimensionale. Il primo riguarda l'approccio che misura la disuguaglianza degli individui che si differenziano sulla base di più attributi incluso il reddito. Questa misura viene effettuata a partire da una funzione di benessere sociale generalizzata di cui si studiano le caratteristiche su basi assiomatiche. Successivamente, mi soffermo sui pro e i contro della misurazione di distribuzioni multidimensionali attraverso degli indici. Infine, analizzo un filone di ricerca, analiticamente molto complesso, che studia la disuguaglianza multidimensionale facendo uso di strumenti di analisi convessa. La quinta sezione conclude con alcune osservazioni riguardanti i problemi ancora aperti e le sfide teoriche da affrontare in questo filone di ricerca ancora quasi del tutto inesplorato.

2. Maggiorazione univariata

All'inizio del secolo scorso, gli economisti hanno incominciato ad essere non solo interessati ai problemi legati alla disuguaglianza, ma anche a volerla misurare, cioè a dire hanno sentito l'esigenza di uno strumento che valutasse quando una determinata distribuzione di una caratteristica quantitativa di una popolazione fosse più o meno sperequata. Storicamente, il primo strumento che ha consentito di fare confronti in termini di disuguaglianza (del reddito) fu scoperto e studiato da Lorenz [18], il quale ha introdotto quella che oggi viene chiamata *curva di Lorenz.*

Si consideri una popolazione di n individui e sia x_i la ricchezza degli i individui, $i = 1, ..., n$. Si ordinino gli individui dal più povero al più ricco, ottenendo la distribuzione $x_{(1)}, ..., x_{(n)}$. Si traccino ora i punti di coordinate $(k/n, S_k/S_n)$, $k = 0, ..., n$, dove $S_0 = 0$ e $S_k = \sum_{i=1}^{k} x_{(i)}$ è la ricchezza totale del k-esimo gruppo di individui più poveri nella popolazione. Unendo questi punti con una spezzata, si ottiene una curva che unisce l'origine degli assi cartesiani col punto di coordinate $(1, 1)$. Se $x_1, ..., x_n$ è la distribuzione di una certa quota di ricchezza T fra n individui e $y_i, ..., y_n$ è una distribuzione alternativa di T. Allora l'idea di Lorenz è che:

Definition 1. $(x_1, ..., x_n)$ *rappresenta una distribuzione più egalitaria della ricchezza rispetto a* $(y_i, ..., y_n)$ *se e solo se*

$$\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)} \qquad k = 1, ..., n-1 \tag{2.1}$$

E' ovvio che,

$$\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)} = T \tag{2.2}$$

Le relazioni 2.1 e 2.2 sono una maniera di dire che la distribuzione x è meno disuguale, analiticamente "più liscia" (o dualmente è maggiorata da), della distribuzione y. Questo si indica con la scrittura $x \preceq y$.

Se $\mathbb{R}^n$ è l'insieme di tutte le distribuzioni del reddito di una popolazione di n individui, un *criterio di disuguaglianza* $\preceq$, come quello di Lorenz, è generalmente una relazione binaria definita su un sottoinsieme $\mathcal{A}$ di $\mathbb{R}^n$, che soddisfa le seguenti proprietà:

(Riflessivita') $\qquad x \preceq x \quad \text{per ogni} \quad x \in \mathcal{A}$

(Transitivita') $\qquad x \preceq y \;\text{ e }\; y \preceq z \quad \text{implica} \quad x \preceq z \quad \text{quando} \quad x, y, z \in \mathcal{A}$

(Antisimmetria) $\qquad x \preceq y \;\text{ e }\; y \preceq x \quad \text{implica} \quad x = y$

Una relazione binaria $\preceq$ per la quale valgono queste tre proprietà viene chiamata *ordinamento parziale*.[1] Se due distribuzioni del reddito $x, y \in \mathbb{R}^n$ soddisfano $x \preceq y$, allora diciamo che y è più diseguale (sperequata) di x.

Successivamente Dalton [4] notò che: "*If there are only two income-receivers and a transfer of income takes place from the richer to the poorer, inequality is diminishing, respecting the limiting condition that the transfer must not be so large as to more than reverse the relative positions of the two income receivers*". Formalmente, questo significa che se $y_i < y_j$ con y_k il reddito dell'individuo k con $k = 1, ..., n$ e un ammontare δ di reddito viene trasferito dall'individuo j all'individuo i, allora la disuguaglianza del reddito risulterà essere diminuita se $\delta \leq y_j - y_i$.

Muirhead [24] aveva già analizzato la nozione di trasferimento di Dalton. Infatti, aveva dimostrato che se le componenti delle distribuzioni x and y sono numeri interi non negativi, allora le seguenti condizioni risultano essere equivalenti:

(i) possiamo ottenere la distribuzione x dalla distribuzione y attraverso una serie finita di trasferimenti ognuno dei quali deve soddisfare la restrizione imposta da Dalton e citata sopra;

(ii) la somma delle prime k componenti più grandi di x deve essere inferiore o al più uguale alla somma delle corrispondenti k più grandi componenti di y, $k = 1, ..., n$, con l'uguale quando $k = n$.[2]

[1] In realtà, il criterio di dominanza di Lorenz è un preordine perchè esso implica una condizione più debole dell'antisimmetria: $x \prec y$ e $y \prec x$ insieme implicano che x è una permutazione degli elementi di y e allora dev'essere $x = y$.

[2] Si noti che la condizione (ii) è equivalente alle condizioni 2.1 e 2.2 nella Definizione 1.

Muirhead [24] aveva inoltre osservato che se y_i e y_j sono rimpiazzate da $y_i + \delta$ and $y_j - \delta$ in $\delta \leq y_j - y_i$, allora ciò equivale a rimpiazzare y_i e y_j con delle medie. Se $0 \leq \alpha = \delta / (y_j - y_i) \leq 1$ e $\overline{\alpha} = 1 - \alpha$, allora

$$y_i + \delta = \overline{\alpha} y_i + \alpha y_j \quad \text{e} \quad y_j - \delta = \alpha y_i + \overline{\alpha} y_j . \tag{2.3}$$

Medie ripetute di due redditi alla volta saranno allora equivalenti alla distribuzione x che si ottiene da y attraverso un numero finito di trasferimenti. Hardy, Littlewood and Polya [10] (da qui in avanti semplicemente HLP), successivamente dimostrarono che l'operaione 2.3 produce lo stesso risultato della sostituzione di y_i con x_i attraverso una media arbitraria del tipo:

$$\left\{ \begin{array}{c} x_i = y_1 p_{1i} + ... + y_n p_{ni}, \qquad i = 1, ..., n \\ \text{dove } p_{ij} \geq 0, \text{ per tutti } i, j, \text{ è tale che:} \\ \sum_{i=1}^{n} p_{ij} = 1 \quad \text{per ogni } j, \quad \text{e} \quad \sum_{j=1}^{n} p_{ij} = 1 \quad \text{per tutti gli } i. \end{array} \right. \tag{2.4}$$

HLP hanno inoltre dimostrato che l'operazione 2.4 è equivalente a premoltiplicare un vettore y per una matrice bistocastica $P = (p_{ij})$, cioè una matrice quadrata le cui componenti sono tutte non negative e dove la somma di ciascuna riga e/o di ciascuna colonna è pari a uno. A parole, questo risultato, molto elegante sotto il profilo analitico, significa che se $x \preceq y$, allora x può essere derivato da y attraverso una sequenza finita di trasferimenti, che soddisfano la restrizione enunciata da Dalton, se e solo se:

$$x = yP \text{ dove } P \text{ è una matrice bistocastica } n \times n.$$

Remark 1. *Si noti che le matrici di permutazione sono matrici quadrate nelle quali in ogni riga e/o colonna c'è un uno e tutte le altre componenti sono zero. Esse sono delle matrici bistocastiche di particolare interesse. Birkhoff (si veda Marshall and Olkin* [22]*) ha dimostrato che le matrici di permutazioni sono i punti estremi dell'insieme convesso delle matrici bistocastiche e che l'insieme delle matrici bistocastiche rappresenta l'involucro convesso delle matrici di permutazione.*

HLP [10] hanno dimostrato che un tipo particolare di trasformazione lineare, chimata T-*transform*, la cui matrice ha la forma

$$T = \lambda I + (1 - \lambda) Q$$

dove $\lambda \in [0, 1]$ e Q è una matrice di permutazione che interscambia solo due coordinate, è equivalente ad un trasferimento di tipo Dalton-Muirhead. La trasformazione yT ha la forma:

$$yT = (y_1, ..., y_{j-1}, \lambda y_j + (1 - \lambda) y_k, y_{j+1}, ..., y_{k-1}, \lambda y_k + (1 - \lambda) y_j, y_{k+1}, ..., y_n) .$$

HLP dimostrano che se, in base al criterio di Lorenz, la distribuzione x è meno disuguale della distribuzione y, allora x può essere derivato da y attraverso un numero finito di applicazioni successive di T-transform. Dal momento che una T-transform è una matrice bistocastica e che il prodotto (finito) di matrici bistocastiche è una matrice bistocastica, HLP mostrano che la seguente equivalenza è vera:

$$x = y \prod_{i=1}^{k} T = yP.$$

A parole, una sequenza di trasferimenti di Muirhead-Dalton (cioè P) può essere scomposta in un numero finito di trasferimenti elementari, cioè trasferimenti che coinvolgono solo due coordinate alla volta (cioè T). Questo risultato è cruciale nello studio della disuguaglianza sotto il profilo teorico. Infatti, esso suggerisce che è sufficiente analizzare una popolazione di due individui, cioè $n = 2$, per dimostrare i risultati principali passati in rassegna in questa sezione, e ciò senza perdere in generalità.

La nozione di uguaglianza o il suo duale (disuguaglianza) è una nozione intuitiva, ma storicamente fu resa analiticamente precisa solo col lavoro di Schur [29] sui determinanti di Hadamar delle matrici Hermitiane semidefinite. I risultati di Schur rappresentano il punto di partenza di una vasta letteratura matematica sulla disuguaglianza. Schur [29] studia le funzioni cosiddette che preservano l'ordinamento e che in economia vengono più semplicemente chiamate funzioni di valutazione sociale. Analiticamente, se $\preceq$ è un preordine definito su un qualche insieme $\mathcal{A} \subseteq \mathbb{R}^n$,

Definition 2. *Una funzione $\varphi : \mathcal{A} \subseteq \mathbb{R}^n \to \mathbb{R}$ è detta che preserva l'ordinamento (o isotonica) se*

$$x \preceq y \quad \textit{implica} \quad \varphi(x) \leq \varphi(y), \quad x, y \in \mathcal{A}.$$

Dal momento che la ricerca di Schur [29] è stata l'origine di tutta una vasta letteratura sulle funzioni che preservano l'ordine, tali funzioni vengono oggi chiamate anche "convesse nel senso di Schur", in onore dello sfortunato matematico di origini ebree, o più frequentemente "S-convesse", per contrapposizione alle funzioni "convesse nel senso di Jensen", oggi più semplicemente chiamate funzioni "convesse". Definiamo una funzione φ come convessa nel senso di Schur se vale la seguente:

Definition 3. *Una funzione $\varphi : \mathcal{A} \subseteq \mathbb{R}^n \to \mathbb{R}$ è S-convessa se $f(Bx) \leq f(x)$ per tutti gli $x \in \mathcal{A}$ e tutte le matrici $n \times n$ bistocastiche B e strettamnete S-convessa se la disuguaaglianza vale in senso stretto e B non è una matrice di permutazione.*

Dal lavoro di Schur, noi sappiamo che se φ è differenziabile e se

$$\varphi_{(k)}(z) = \partial\varphi(z) / \partial z_k$$

è la derivata parziale di φ rispetto al suo k-esimo argomento, allora il seguente teorema caratterizza la classe di funzioni *S-convesse*:

Theorem 1 (Schur, 1923; Ostrowski,1952)**.** *Sia $\mathcal{A} \subset \mathbb{R}$ un intervallo aperto e sia $\varphi : \mathcal{A}^n \to \mathbb{R}$ una funzione continua e differenziabile. Condizione necessaria e sufficiente affinchè φ sia Schur-convessa in $\mathcal{A}^n$ è che*

$$\varphi \textit{ sia simmetrica su } \mathcal{A}^n$$

e

$$\varphi_{(i)}(z) \textit{ sia crescente in } i = 1, ..., n \quad \textit{per tutti } z \in \mathbb{R} \cap \mathcal{A}^n,$$

dove $\mathbb{R} = \{(x_1, ..., x_n) : x_1 \leq ... \leq x_n\}$ è un campo ordinato.

Alternativamente, φ è Schur-convessa in $\mathcal{A}^n$ se e solo se φ è simmetrica e per tutti gli $i \neq j$,

$$(z_i - z_j)\left[\varphi_{(i)}(z) - \varphi_{(j)}(z)\right] \geq 0 \quad \textit{per ogni} \quad z \in \mathcal{A}^n \tag{2.5}$$

Si noti che nel dimostrare che una funzione φ è Schur-convessa è sufficiente dimostrare che $\varphi(x) \leq \varphi(y)$ quando $x \preceq y$ e x e y differiscono solo in due componenti. Questa è una conseguenza del fatto che se $x \preceq y$, allora x può essere derivato da y attraverso un numero finito di T-transform.

Il risultato di Schur è utile per identificare la classe di funzioni consistente con un dato ordinamento. Infatti, il Teorema 1 significa che se la distribuzione x è "più liscia", cioè meno disuguale, della distribuzione y, sulla base di un qualche criterio di disuguaglianza (cioè un preordinamento come quello di Lorenz), allora la disuguaglianza associata a y risulterà essere maggiore di quella associata a x se la misuriamo ricorrendo alla classe di funzioni di valutazione sociale φ (cioè alla classe di funzioni S-convesse se l'ordinamento utilizzato è quello di Lorenz).

A seguito del lavoro di Schur [29], la ricerca matematica sulla disuguagliaza è letteralmente esplosa. Al contrario, in economia, dobbiamo attendere l'inizio degli anni Settanta affinchè riprenda il dibattito sui problemi legati alla misurazione della disuguaglianza economica. Tutto ciò grazie ai pioneristici lavori di Kolm [14] e Atkinson [1]. Atkinson ha giustificato l'utilizzo delle curve di Lorenz per misurare la disuguaglianza del reddito all'interno di un approccio di tipo utilitarista. Ha mostrato, in altri termini, che se una funzione di valutazione sociale è additiva separabile ed è la somma di funzioni di utilità individuali (la stessa funzione di utilità per tutti gli n individui), allora l'ordinamento parziale di Lorenz sulle distribuzioni del reddito equivale all'ordinamento dato da una funzione di valutazione sociale così fatta, e dove le funzioni di utilità degli individui sono concave.

Questo risultato si basa in realtà su un risultato più generale che qui riporto:

Theorem 2 (HLP, 1934)**.** *Siano $x, y \in \mathbb{R}^n$, le seguenti condizioni sono equivalenti:*
(i) $x \preceq y$;
(ii) $x = yP$ *per qualche matrice bistocastica* P;
(iii) $\sum_1^n g(x_i) \leq \sum_1^n g(y_i)$,*per ogni funzione strettamente concava* g.[3]

Il Teorema 2 rappresenta la "pietra miliare" di tutta la letteratura economica sulla teoria della disuguagliaza. In esso si mostra che l'ordinamento delle distribuzioni del reddito prescelto deve avere un contenuto normativo e positivo. Infatti, la condizione (i) che rappresenta l'ordinamento di Lorenz, una misura statistica, quindi a forte contenuto positivo; la condizione (ii) è il cosiddetto criterio di trasferimento di Muirhead-Dalton, dal forte valore normativo; infine la condizione (iii) mostra una funzione di valutazione a valore normativo, legato alla scelta del particolare funzionale, e a valore descrittivo, essendo il duale di una misura del grado di disuguaglianza di una distribuzione.

3. Maggiorazione multidimensionale

In questa sezione passiamo in rassegna il problema di come modellizzare in astratto la disuguaglianza multidimensionale. Supponiamo che la componenti di due distribuzioni x e y siano punti in $\mathbb{R}^m$, cioè vettori colonna. In tal caso, x, y diventano matrici che indichiamo come segue:

$$X = (x_1^n, ..., x_m^n),$$

dove x_i^n sono tutti vettori colonna di lunghezza n. Con riferimento alla nozione di *T-transform,* la definizione seguente esprime l'idea che la matrice X è meno disuguale, cioè a dire rappresenta una distribuzioni di caratteristiche meno sperequate di quelle di Y:

Definition 4. *Siano X e Y due matrici $n \times m$. Allora X è detta maggiorata a catena da Y, scritto $X \prec\prec Y$, se $X = PY$ dove P è il prodotto di un numero finito di $n \times n$ T-transform.*

[3]Dasgupta et alii [6] hanno poi esteso questo risultato alla classe di funzioni S-convesse.

In altre parole, l'idea di trasferimento introdotta da Muirhead [24] e Dalton [4] si estende anche la caso in cui le componenti di x e y sono vettori. Di fatto, se noi sostituiamo y_i e y_j con x_i e x_j otteniamo un nuovo vettore x dal vettore y, con i vincoli che:
i) x_i, x_j stanno nell'involucro convesso di y_i, y_j;
ii) $x_i + x_j = y_i + y_j$.
allora:

Definition 5. *Se X e Y sono due matrici $n \times m$, allora X viene detta maggiorata da Y, scritto $X \prec Y$, se $X = PY$, dove P è una matrice bistocastica $n \times n$.*

Dal momento che il prodotto di *T-transform* è una matrice bistocastica, la maggiorazione a catena implica la maggiorazione, $Y \prec\prec X \Rightarrow Y \prec X$ e nel caso $n = 1$, e quando $m = 2$, è vero anche l'inverso. In generale, per $n \geq 2$ e $m \geq 3$ la *maggiorazione* tra matrici non implica la nozione di *maggiorazione a catena*. Questa è una differenza fondamentale risptto al caso univariato dove, come abbiamo visto, esiste la perfetta equivalenza delle due nozioni, perchè una matrice bistocastica può sempre essere scomposta nel prodotto di un numero finito di T-transform. La Definizione 5 dice semplicemente che la media è un'operazione che liscia la distribuzione, cioè che rende le componenti della matrice X più "liscie" di quelle di Y.

Se definiamo l'involucro convesso di una matrice generica Y, indicata come:

$$H = co\left\{\left(y_i^1, ..., y_i^n\right), \quad i = 1, ..., m\right)\right\},$$

come la combinazione convessa dei vettori riga della matrice, una definizione equivalente di *maggiorazione* $\prec$ è la seguente:

Definition 6. *Siano $X, Y \in \mathbb{R}^{n \times m}$ due matrici, allora diciamo che X mostra un minore livello di disparità di Y, se X è contenuta nell'involucro convesso di tutte le permutazioni di Y.*

In un contesto multidimensionale, non esiste un Teorema equivalente al Teorema 2 e ben poco è noto riguardo le funzioni che preservano gli ordinamenti come quello di maggiorazione $\prec$. La difficoltà nel non riuscire a trovare un risultato equivalente all'elegante Teorema di HLP per l'ordinamento $\prec$ risiede nel fatto cher non esiste un risultato intermedio utile per la dimostrazione finale come quello dato dalla consizione (i) di Muirhead citata in precedenza. Rinott [27] ha caratterizzato la classe di funzioni che preservano l'ordinamento $\prec\prec$. Inoltre, è noto che se $Y \prec X$, allora $\sum_1^n \varphi\left(x_i^C\right) \leq \sum_1^n \varphi\left(y_i^C\right)$ per tutte le funzioni continue e convesse $\varphi : \mathbb{R}^m \to \mathbb{R}$, dove z_i^C è un vettore colonna di lunghezza m (si veda Marshall e Olkin [22] capitolo 15). Sulle funzioni multivariate che preservano l'ordinamento non è noto molto altro.

Un'altra nozione importate di maggiorazione tra matrici, nota agli economisti, è quella di *maggiorazione di prezzo*. Essa fu proposta da Marshall e Olkin [22] come un problema aperto. Formalmente:

Definition 7. *Siano X e Y due matrici, X è detta maggiorizzata direzionalmente da Y, e noi scriviamo $X \prec_d Y$, se $aX \prec aY$ per tutti gli $a \in \mathbb{R}^n$.*

Marshall e Olkin [22] dimostrano che $X \prec Y$ implica $X \prec_d Y$, in un contesto più generale, dove $X \prec_d Y$ significa $XA \prec YA$ per tutti gli $A \in \mathbb{R}^{m \times k}$ (e k

fissato). Inoltre, essi hanno formulato la questione se $X \prec_d Y$ implicasse a sua volta $X \prec X$ e in un lavoro importante, più tardi, Bhandari [3] ha fornito le condizioni sufficienti sotto cui è possibile affermare che la maggiorazione direzionale implica la maggiorazione multivariata ordinaria, cioè:

Theorem 3. *Si supponga che ogni vettore colonna di $Y_{(n\times m)}$ rappresenti un punto estremo dell'involucro convesso generato dalle colonne di X, e che esso abbia un volume positivo r-dimensionale, e che almeno $(m-r+2)$ di queste colonne siano coplanari.*[4] *Allora $X \prec_d Y$ implica $X \prec Y$ per tutte le X. Inoltre, $XP \prec YP$ per tutte le $P_{(m\times k)}$ implica $X \prec Y$.*

Cio che Bandhari [3] chiama maggiorazione direzionale, viene chiamata *maggiorazione attraverso combinazioni lineari* da Joe and Verducci [11] e *maggiorazione via i prezzi* dagli economisti come Kolm [13]. Infatti quando una distribuzione X è meno disuguale di una distribuzione Y, i.e. $X \prec Y$, ogni curva di Lorenz di X è in nessun punto al di sotto delle corrispondenti curve di Y per tutti i vettori di prezzo p (inclusi quelli negativi), e ciò se esse non sono permutazioni l'una dell'altra. Ciò implica che tutte le proprietà passate in rassegna in precedenza nel caso unidimensionale valgono anche in codesto contesto per tutte le distribuzioni del reddito derivate da Y e X, cioè per tutte le relazioni del tipo $pX \prec pY$, qualunque siano i prezzi utilizzati per l'aggregazione.

La nozione di maggiorazione via vetore dei prezzi è molto utile in economia per confrontare caratteristiche di tipo qualitativo tipo salute, istruzione ecc. In tal caso è sempre possibile valutare la disparità economica associata alle matrici le cui componenti sono di tipo qualitativo, assegnando un valore in termini di prezzo a variabili che a priori possono essere confrontate solo a livello ordinale. Ovviamente, così facendo il ricercatore riduce le caratteristiche qualitative a semplici variabili monetarie perdendo di vista l'informazione stessa che ci spinge ad estendere a più variabili la nostra analisi della disuguaglianza.

In un contesto meno astratto, immaginiamo ora che le righe di una matrice di una distribuzione multivariata rappresentino gli n individui di una popolazione e che le colonne della stessa matrice rappresentino le m caratteristiche rilevanti di quella stessa popolazione di individui. Vogliamo a questo punto dotarci di alcuni strumenti per misurare la disparità o dualmente il benessere associato ad una generica distribuzione X così costruita. I confronti tra matrici sulla base di diversi criteri di valutazione/misurazione della disuguaglianza è l'oggetto della sezione seguente, in cui ci soffermeremo sulle diverse soluzioni fornite dagli studiosi al problema della misura della disuguaglianza in un contesto multidimensionale.

4. La misurazione delle disuguaglianza multidimensionale: tre approcci controversi

In questa sezione, consideriamo che una popolazione di individui sia caratterizzata da diversi attributi quali la salute, il reddito, l'istruzione, i talenti, le "capacità". La letteratura economica tradizionale considera il reddito come una valida proxy per stabilire il benessere di un gruppo di persone. Tuttavia diversi studiosi, quali Kolm [13], Atkinson and Bourguignon [2], hanno sottolineato il fatto ovvio che questo tipo di approccio all'analisi della disuguaglianza è insoddisfacente, perchè le

[4] L'ipotesi che almeno $(m-r+2)$ colonne di X_i siano co-planari equivale a dire che esse devono appartenere allo spazio affine bidimensionale di $\mathbb{R}^m$.

persone differiscono per tutta una serie di attributi. Storicamente, in economia, si sono seguiti due diversi e complementari approcci all'analisi della disparità. Il primo ordina le diverse distribuzioni multivariate sulla base di una funzione del benessere sociale (esempi significativi si trovano nei lavori di Atkinson and Bourguignon [2] e Kolm [13]). Un secondo approccio utilizza i numeri indici, cioè statistiche sintetiche di valori disaggregati (si vedano, fra gli altri, i lavori di Maasoumi [19], Tsui [33] e Gajdos and Weymark [8]). In questi lavori, gli attributi individuali vengono misurati ricorrendo ad una funzione di utilità, in modo tale da ottenere una distribuzione vettoriale univariata di utilità che vengono poi sintetizzate (cioè misurate), ricorrendo ad un tradizionale indice di disuguaglianza. Entrambi gli approcci non sono scevri da problemi, come ha sottolineato in un articolo molto arguto Dardanoni [5]. Più di recente, Koshevoy and Mosler [14], [15], [17], [16] hanno sviluppato un terzo approccio all'analisi della disuguaglianza multidimensionale, basato sull'ampio utilizzo di strumenti di analisi convessa. Di seguito, dirò sommariamente dei principali risultati raggiunti da questi tre filoni di ricerca.

4.1. **Ordinamento di matrici via funzioni del benessere sociale.** I confronti di benessere tra popolazioni con una diversa distribuzione di redditi, bisogni, istruzione, età, abilità ecc. devono essere effettuati sulla base di funzioni di valutazione che tengano conto delle diverse dotazioni degli individui. Iniziamo col considerare il lavoro di Kolm [13], il quale, per primo, ha posto il problema: "*When is one multivariate distribution more spread out than another one?*". Kolm misura la disuguaglianza multidimensionale facendo uso di una funzione di benessere sociale (SWF)

$$W : \mathbb{R}_+^{n\times m} \to \mathbb{R}$$

definita sull'insieme di tutte le matrici rettangolari semidefinite. Secondo la terminologia adottata in precedenza, una SWF non è nient'altro che una funzione che preserva un ordinamento, funzione con determinate proprietà. I meriti di Kolm però non finiscono qui. Egli ha introdotto per primo in economia la nozione di maggiorazione $\prec$ tra matrici e proposto un'altra nozione di maggiorazione matriciale.

Si consideri una matrice generica X con righe $x_1^R, ..., x_m^R$:

Definition 8. *Per ogni $X, Y \in \mathbb{R}_+^{n\times m}$, si dice che X è maggiorizzata per righe da Y, scritto $X \prec_{row} Y$ se e solo se esiste una matrice bistocastica P tale che $x_i^R = y_i^R P$ con $i = 1, ..., n$.*

$X \prec_{row} Y$ può essere scritto anche come:

$$(x_1^R, ..., x_n^R) = (y_1^R, ..., y_n^R) \begin{bmatrix} P & 0 & ... & 0 \\ 0 & P & ... & 0 \\ . & . & . & . \\ 0 & 0 & ... & P \end{bmatrix} \tag{4.1}$$

Ciò implica che:

$$x_i^R \preceq y_i^R \text{ per ogni } i = 1, ..., n \tag{4.2}$$

dove il simbolo $\preceq$ è da intendersi come nella Definizione 1. Dal momento che P potrebbe essere il prodotto di un numero finito di *T-transform*, la $\prec$*-maggiorazione* presumibilmente implica 4.1 and 4.2. Tuttavia, essa implica semplicemente che esistono $P_1, ..., P_m$ matrici bistocastiche tali che $x_i^R = y_i^R P_i, i = 1, ..., m$, ma assolutamente nulla garantisce che $P_i = P_j \forall i \neq j$, (si veda Marshall and Olkin [22] capitolo 15).

La critica principale a questo filone di ricerche che si basa su una valutazione della disuguaglianza multidimensionale via le SWF, si riferisce al fatto che i risultati principali che si ottengono impiegando tale metodologia partono dal presupposto forte che non esiste alcuna interrelazione tra le diverse caratteristiche individuali, cioè a dire, si assume che ad esempio, il reddito e la salute siano delle variabili non correlate o più semplicemente che la loro correlazione sia irrilevante ai fini valutativi della disuguaglianza. In realtà, tali correlazioni sono molto importanti ai fini dell'analisi e lo testimoniano i lavori di Atkinson e Bourguignon [2] e Rietveld [26].

Atkinson e Bourguignon [2] si sono soffermati sui confronti in termini di disuguaglianza che utilizzano la dominanza stocastica come strumento analitico. Per la precisione, essi si soffermano sul caso bidimensionale, utilizzando alcuni risultati di dominanza stocastica multivariata per selezione di portafoglio.[5] Generalizzando i risultati sull'ordinamento di Lorenz unidimensionale, Atkinson e Bourguignon analizzano come diverse forme di deprivazione (ad esempio, basso reddito, bassa istruzione, bassa qualità della vita ecc.) tendono ad essere associate insieme. Essi hanno usato una SWF[6] per valutare il benessere associato a diversi *vettori etichettati* x^i,[7] cioè vettori che rappresentano la percentuale della quantità totale della merce i-esima allocata all'individuo j-esimo, e hanno quindi analizzato le implicazioni di diverse ipotesi riguardo la SWF e i differenti livelli di interdipendenza tra gli elementi di x^i.

Rietveld [26] analizza i problemi della scomposizione della disuguaglianza in componenti che congiuntamente concorrono alla disparità. La sua analisi si basa sul criterio di Lorenz. Se definiamo la curva di Lorenz per ogni componente di reddito, risulta che la disuguaglianza nel reddito totale non è maggiore della disuguaglianza fatta registrare dalla componente più sperequata. Da ciò segue che, in generale, la curva di Lorenz del reddito totale si trova al di sopra della media pesata delle curve di Lorenz delle componenti di reddito. Dunque, esisterebbe una sorta di *effetto aggravante* causato dalla correlazione tra le diverse componenti di reddito. Inoltre, misurando la disuguaglianza di una data distribuzione attraverso una funzione $f : \mathbb{R}^n \to \mathbb{R}$ omogenea e Schur-convessa, Rietvield dimostra che la considerazione congiunta delle componenti reddituali produce una sorta di *attenuazione* della disuguaglianza del reddito totale per un'ampia classe di misure della disuguaglianza, omogenee di grado zero, e rappresentabili come la somma di funzioni convesse. Tutto ciò, comunque non vale in generale. Se si prende, ad esempio, in considerazione il coefficiente di Gini, è possibile osservare che esso possiede la proprietà di "attenuazione" sopra descritta, ma è noto che esso non possa essere scritto come somma di funzioni convesse. Rietveld sostiene quindi che l'omogeneità e la convessità nel senso di Schur non sono consizioni sufficienti per ottenere la proprietà di attenuazione in un contesto multidimensionale, concludendo che le interazioni tra

[5] Analiticamente, il confronto tra distribuzioni bivariate in [2] si effettua sulla base delle differenze nelle utilità attese e delle proprietà che caratterizzano varie classi di SWF.

[6] Si assume che le SWF siano additive separsabili nelle utilità e simmetriche rispetto agli individui.

[7] Il temine bene etichettato è stato usato per la prima volta da Hahn [9] nell'analisi dell'equilibrio con costi di transazione.

le componenti che concorrono al benessere di una collettività sono rilevanti quando si fanno confronti in termini di disuguaglianza.

Di notevole interesse, in questo filone di letteratura economica, è anche il lavoro di Mosler [23]. Questi considera diversi attributi che descrivono gli stati sociali individuali e diversi criteri di valutazione delle consizioni di vita prodotte da detti stati sociali. I confronti di benessere sono fatti sulla base di applicazioni simultanee di un insieme di funzioni di valutazione sociale, che dipendono dalle diverse dotazioni degli individui. Mosler utilizza funzioni di valutazione sociale che possono essere espresse come somma di funzioni di valutazione dei diversi stati individuali, confrontando i livelli di benessere sociale in un modello di tipo ordinalista. L'approccio usato è di tipo assiomatico e diversi criteri di ordinamenti parziali multidimensionali vengono utilizzati, mentre alcune classi di funzioni di valutazione sociale vengono mostrate essere coerenti con detti criteri. L'originalità del lavoro di Mosler risiede nel suo approccio agli ordinamenti di benessere che coinvolge l'analisi dell'uguaglianza multidimensionale. Egli, in un setting puramente ordinalistico, introduce un valore soglia, una sortea di linea di povertà multidimensionale in termini di merci, sulla base del quale è possibile fare dei confronti delle diverse dotazioni di attributi posseduti dagli individui di una data popolazione.

4.2. **Indici di disuguaglianza multidimensionale[8].** In questa sezione analizziamo le proprietà degli indici di disuguaglianza nel caso multidimensionale. Questo approccio sintetizza gli attributi che gli individui posseggono attraverso una valutazione di tipo utilitarista. In tal modo, per ognuno degli n individui di una popolazione abbiamo un numero indice che lo rappresenta. Per tanto, si ottiene, dopo aver sintetizzato le caratteristiche individuali attraverso un indice di benessere, una distribuzione delle utilità individuali. A questo punto, un indice di disuguaglianza unidimensionale è applicato alla distribuzione delle utilità. Questo approccio detto "a due stadi", raccoglie l'informazione relativa alla disuguaglianza multidimensionale di una popolazione di individui attraverso una statistica.

Definition 9. *Un indice di disuguaglianza multidimensionale può essere scritto come una funzione a valori reali* $U_1(x_1), ..., U_n(x_n)$*, dove* $U_i(\cdot) : \mathbb{R}^m \to \mathbb{R}$ *indica una funzione di utilità individuale e* x_j *un vettore tiga di una matrice* X $n \times m$*.*

Questo tipo di esercizio teorico comporta due tipi di problemi. Innanzi tutto, occorre scegliere una funzione di utilità: scelta ovviamente arbitraria, anche quando è fatta su base assiomatica, perchè la scelta di una funzione al posto di un'altra significa sottolineare una determinata classe di preferenze e magari trascurare altri spazi valutativi che potrebbero essere molto importanti. Secondo, il processo di aggregazione del vettore delle utilità attraverso un indice di disuguaglianza è un'operazione in cui si perde molta dell'informazione per la quale e con la quale l'esercizio assume valore.

Sulla base della moderna teoria dell'informazione, Maasoumi [19] sostine che quando il benessere individuale è il *primum movens*, la classe di funzioni dell'entropia generalizzata (nel cui insieme stanno molte delle funzioni di utilità più utilizzate in economia) risulta essere la soluzione ottimale al primo dei due problemi citati più sopra. Seguendo l'approccio di Kolm [13], Maasoumi considera una matrice Y che rappresenta una ipotetica società composta da n individui, con come dotazione

[8] Per un'ampia rassegna di questo approccio si veda Weymark [35].

iniziale una certa quantità di m merci.[9] Moltiplicando una matrice Y per una matrice bistocastica P, Maasoumi ottiene una nuova matrice X che dovrebbe rappresentare una distribuzione multivariate meno sperequata. Tale trasformazione dovrebbe, inoltre essere identificata da qualsiasi indice di disuguaglianza. Questa tesi si basa su un argomento discusso in Kolm [13]. Quest'ultimo nota come una trasformazione via una matrice bistocastica sia condizione necessaria e sufficiente per un miglioramento evidente del benessere relativo ad una distribuzione multivariata. Tuttavai, Dardanoni [5] ha notato che, in un contesto multidimensionale, gli effetti di una moltiplicazione di una matrice per una bistocastica sono spesso ambigui. Di fatto, ha dimostrato che permutando le componenti di una matrice Y, 3×3, (cioè una matrice che rappresenta una popolazione di 3 individui distinti sulla base della distribuzione di 3 attributi):

$$Y = \begin{bmatrix} 100 & 1 & 1 \\ 1 & 90 & 100 \\ 90 & 100 & 90 \end{bmatrix},$$

con una matrice bistocastica, si ottiene una nuova matrice:

$$Y' = \begin{bmatrix} 1 & 1 & 1 \\ 90 & 90 & 90 \\ 100 & 100 & 100 \end{bmatrix}$$

le cui righe rappresentano una distribuzione iniqua delle dotazioni iniziali dei 3 individui. Sfortunatamente, questa evidente conclusione non si ottiene con l'applicazione di qualsiasi indice di disuguaglianza multidimensionale. Esistono ad esempio alcuneSWF, che appartengono alla classe di funzioni dell'entropia generalizzata usata da Maasoumi, che mostrano una diminuzione della disuguaglianza dopo una trasformazione come quella che ci fa ottenere la matrice Y'. Dardanoni [5] ha dimostrato che affinchè un indice di disuguaglianza multidimensionale registri una diminuzione dell'uguaglianza nel passaggio da Y a Y' occorre che la SWF sia additiva separabile nelle utilità individuali. Ciò, per quanto osservato da Fishburn [7], non è un'ottima rappresentazione delle preferenze degli individui, e in particolare contraddice le valutaioni del benessere individuale in tutti quei casi che prevedono correlazione tra gli attributi personali.

L'"approccio a due stadi" di Maasoumi è stato utilizzato da Tsui [33], [34] per caratterizzare assiomaticamente la classe di indici di disugliaglianza a' la Atkinson-Kolm-Sen e quella di Shorrocks [32] per il caso multidimensionale. Gajdos e Weymark [8], che generalizzano la classe di indici di Gini al contesto multidimensionale, invertono la procedura di Maasoumi aggregando prima gli attributi individuali e poi i valori così ottenuti.

4.3. **Ordinamento di Lorenz multivariato e indice di Gini.** Il lavoro congiunto di Koshevoy e Mosler merita una menzione speciale. A partire da un lavoro di Rado [25], essi introducono l'*analisi convessa* nella teoria della disuguaglianza multidimensionale. Nel suo primo lavoro su questi temi, Koshevoy [14] considera una popolazione di n agenti economici tra i quali un insieme di beni vengono distribuiti. Considera quindi una matrice $X = (x_{ij}) \in \mathbb{R}_+^{n \times m}$ che assegna un vettore di

[9] Si noti che Maasoumi [20] estende tale approccio al caso in cui la distribuzione degli attributi è continua e che Maasoumi e Zandvakili [21] applicano tale modello teorico alla misurazione della mobilità sociale.

beni ad ogni i-esimo agente e si pone il seguente problema: "*Given two distribution matrices X and Y, which one contains the lower level of disparity?*". Nel dare una risposta a questo quesito, Koshevoy generalizza la nozione di curva di Lorenz facendo uso della nozione di corpo convesso, cioè un poliedro convesso in $\mathbb{R}_{+}^{m}$, che rappresenta la curva di Lorenz nello spazio $m+1$ dimensionale. La generalizzazione multivariata della curva di Lorenz viene battezzata col nome di *Lorenz zonotope*, e indicata come $LZ(X)$. Il criterio di Lorenz multivariato sarà allora il seguente:

Definition 10. *Siano X e Y due matrici, X è maggiorata nel senso di Lorenz da Y, scritto come $X \preceq_{LO} Y$, se $LZ(X) \subseteq LZ(Y)$.*

Koshevoy [14] dimostra che la nozione di maggiorazione nel senso di Lorenz è equivalente a quella di maggiorazione via vettore dei prezzi. Tuttavia, noi sappiamo dal lavoro di Bhandari [3], che la maggiorazione via i prezzi $\prec_d$ è equivalente alla nozione di maggiorazione ordinaria $\prec$, e ci aspetteremmo quindi che, per la proprietà transitiva, la maggiorazione $\prec$ sia analiticamente equivalente a la maggiorazione che utilizza il criterio di Lorenz $\prec_{LO}$. Purtroppo, ciò che si verificava nel caso unidimensionale, non si verifica nel multidimensionale e cioè la maggiorazione $\prec$ implica l'ordinamaneto di Lorenz, ma, nel caso multidimensionale, il contrario, in generale, non vale. Ciò perchè analiticamente il criterio di Lorenz nel caso multidimensionale non è compatibile con l'assunzione di separabilità tra gli attributi individuali. Utilizzando un argomento simile a quello usato da Dardanoni [5] per criticare il "processo a due stadi" di Maasoumi, Koshevoy fa osservare che la relazione d'ordine tra matrici $X \preceq_{LO} Y$ non si preserva se selezioniamo una qualunque submatrice di X e Y caratterizzata da un numero inferiore di attributi individuali. Analiticamente, $X \preceq_{LO} Y$ non implica $X^M \preceq_{LO} Y^M$, con Z^M qualunque sottomatrice fatta da n righe e s colonne con $s < m \in M = \{1, ..., m\}$, l'insieme di tutti gli attributi individuali presi in considerazione.

In [15], Koshevoy estende la nozione di Lorenz zonotope a quella di cono d'ordine[10] sviluppando un'approccio geoemtrico all'analisi della disuguaglianza multidimensionale. Di fatto, interpretando le coordinate di una direzione in un cono come i pesi (prezzi) da attribuire alle caratteristiche individuali, l'estenzione al cono d'ordine della maggiorazione multivariata di Lorenz risulta essere equivalente all'estenzione al cono d'ordine della maggiorazione via i prezzi, dove una distribuzione multivariata si dice sia maggiorizzata da un'altra nel cono d'ordine se le spese delle famiglie per qualunque vettore di prezzi in un cono sono meno disperse per la prima che per la seconda distribuzione. L'approccio geometrico allo studio della disuguaglianza multidimensionale ha il pregio che l'insieme delle matrici maggiorizzate da una matrice data può essere descritto, grazie alla nozione di cono d'ordine, attraverso un numero finito di disuguaglianze. Nel caso del cono d'ordine con un numero finito di raggi estremanti, il verificare che vale la relazione d'ordine data dalla maggiorazione via i prezzi estesa al cono equivale a verificare la disuguaglianza univariata per la spesa di una famiglia per un numero finito di prezzi (le direzioni).

Koshevoy e Mosler [17] hanno studiato l'estensione dell'intera classe degli indici di Gini alla misura della disuguaglianza quando le persone sono dotate di vari attributi. Essi hanno utilizzato la nozione di distanza e quella di volume di un corpo convesso di un insieme nello spazio $(m+1)$ dimensionale per definire l'indice di Gini multivariato. Si osserva in [17] come molte delle proprietà che valgono nel

[10]Si veda Marshall e Olkin [22] per una definizione formale del concettto di cono d'ordine.

caso univariato si estendono naturalmente alla classe di funzioni che definiscono la classe di indici di Gini. Ciò costituisce un importante risultato dal momento che l'indice di Gini in un contesto unidiemnsionale rappresenta l'indice di misura della disuguaglianza sicuramente più utilizzato nelle analisi empiriche.

Infine, Koshevoy e Mosler [16] hanno esteso le nozioni di curva di Lorenz e ordinamento di Lorenz al caso di distribuzioni multidimensionali *empiriche*. Essi generalizzano la nozione di curva di Lorenz del caso univariato utilizzando la nozione di zonoid, cioè l'insieme di tutti i punti tra il grafico del duale della curva di Lorenz multivariata e il grafico della curva di Lorenz. Il Lorenz zonoid è un sottoinsieme chiuso e convesso dell'ipercubo unitario in $\mathbb{R}^{m+1}$. E' un politopo convesso di una distribuzione discreta. Due distribuzioni multivariate empiriche vengono ordinate attraverso l'inclusione dei loro rispettivi Lorenz zonoid. Da notare che il criterio di inclusione degli zonoidi equivale alla nozione di maggiorazione via i prezzi delle due distribuzioni empiriche multivariate considerata dall'esercizio.

5. Conclusioni e possibili sviluppi

Abbiamo passato in rassegna i diversi modi di ordinare delle matrici che rappresentano la distribuzione di beni e merci tra persone usando una SWF, notando come questa operazione faccia perdere informazione o restringa la classe ammissibile di funzioni di valutazione. Poi abbiamo discusso alcuni risultati sugli indici di disuguaglianza multidimensionale, misure sintetiche del grado di disparità tra individui distinti sulla base di diverse caratteristiche. L'applicazione di un indice significa rifurre tutte le variabili che si vogliono analizzare a scalari, inficiando in questo modo lo scopo stesso dell'esercizio, che è quello di acquisire molta più informazione di quella fornitaci dalla sola variabile esplicativa reddito. Infine, si è riflettuto sui lavori di Koshevoy and Mosler, i quali hanno esteso la nozione di ordinamento di Lorenz ad un contesto multidimensionale attraverso l'utilizzo di strumenti di analisi convessa. I loro risultati sono analiticamente sofisticati, ma non si discostano molto da quelli ottenuti in teoria della maggiorazione. In definitiva, l'analisi della disuguaglianza multidimensionale costituisce un argomento di frontiera. Molto lavoro aspetta gli studiosi e sicuramente i principali argomenti di studio saranno il problema dei trasferimenti in un contesto tanto più sofisticato e la caratterizzazione delle funzioni che preservano l'ordinamento su matrici.

References

[1] Atkinson, A.B. (1970) - On the measurement of inequality - *Journal of Economic Theory*, 2, 244-263.

[2] Atkinson, A.B. and F. Bourguignon (1982) - The comparison of multidimensioned distributions of economic status - *Review of Economic Studies* 39, 183-201.

[3] Bhandari, S.K. (1988) - Multivariate majorization and directional majorization: positive results - *Sankhyā: The Indian Journal of Statistics*, vol.50 pp.199-204.

[4] Dalton, H. (1920) - The measurement of inequality - *Economic Journal*, 348-361.

[5] Dardanoni, V. (1992) - On multidimensional inequality measurement - In: Dagum, C., Lemmi, A. (Eds.), *Income Distribution, Social Welfare, Inequality,and Poverty*. Vol. 6 of Research on Economic Inequality. JAI Press, Stamford, CT, pp. 201–207.

[6] Dasgupta, P., A.K. Sen, and D. Starret (1973) - "Notes on the measurement of inequality" - *Journal of Economic Theory*, 6, 180-187.

[7] Fishburn, P.C., (1988) - *Nonlinear preferences and utility theory* - Baltimore, Johns Hopkins University Press.

[8] Gajdos, T. and J. A. Weymark (2003) - Multidimensional generalized Gini indices - Working Paper No. 03-W11, Department of Economics, Vanderbilt University.

[9] Hahn, F.H. (1971) - Equilibrium with transaction costs - *Econometrica* 39.

[10] Hardy, G.H., J.E. Littlewood, and G. Polya, (1934, 1952) - *Inequalities* - Cambridge University Press, London.

[11] Joe, H. and J. Verducci (1993) - Multivariate majorization by positive combination - In: *Stochastic Inequalities IMS Lecture Notes:* Monograph Series, vol.22, 159-181.

[12] Kolm, S.C. (1969) - The optimal production of social justice - In: J. Margolis and H. Guitton (eds), *Public Economics*, Macmillan, London and St. Martin's Press, New York, 145–200.

[13] Kolm, S.C. (1977) - Multidimensional egalitarianism - *Quarterly Journal of Economics* 91, 1-13.

[14] Koshevoy, G. (1995) - Multivariate Lorenz majorization - *Social Choice and Welfare* 12, 93-102.

[15] Koshevoy, G. (1998) - The Lorenz zonotope and multivariate majorizations - *Social Choice and Welfare* 15, 1-14.

[16] Koshevoy, G. and K. Mosler (1996) - The Lorenz zonoid of a multivariate distribution - *Journal of American Statistical Association* 91, 873-882.

[17] Koshevoy, G. and K. Mosler (1999) - Multivariate Gini indices - *Journal of Multivariate Analysis* 53, 112-126.

[18] Lorenz, M.O. (1905) - Methods for measuring concentration of wealth - *Journal of American Statistical Association* 9, 209-219.

[19] Maasoumi, E. (1986) - The measurement and decomposition of multidimensional inequality - *Econometrica* 54, 991-997.

[20] Maasoumi, E. (1989) - Continuously distributed attributes and measures of multivariate inequality - *Journal of Econometrics* 42, 131-144.

[21] Maasoumi, E. and S. Zandvakili (1990) - Generalized entropy measures of mobility for different sexes and income levels - *Journal of Econometrics* 43, 121-134.

[22] Marshall, A.W., and I. Olkin, (1979) - *Inequalities: Theory of Majorization and its Applications* - Academic Press: New York.

[23] Mosler, K. (1991) - Multidimensional Welfarism - In Eichhorn, W. (Ed.), *Models and Measurement of Welfare and Inequality* - Springer-Verlag, 808-820.

[24] Muirhead, R. F. (1903) - Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters - *Proceedings of Edinburgh Mathematical Society* 21, 144-157.

[25] Rado, R. (1952) - An inequality - *Journal of London Mathematical Society* 27, 1-6.

[26] Rietveld, P. (1990) - Multidimensional inequality comparisons - *Economics Letters* 32, 187-192.

[27] Rinott, Y. (1973) - Multivariate majorization and rearrangement inequalities with some applications to probability and statistics - *Israel Journal of Mathematics* 15, 60-77.

[28] Savaglio, E. (2004) - Multidimensional majorization with variable population size - *Economic Theory*, forthcoming.

[29] Schur, I. (1923) - Uber eine Klasse von Mittelbildungen mit Anwendungen die Determinanten - *Theorie Sitzungsber Berlin Matematische Gesellschaft* 22, 9-20 (*Issai Schur Collected Works* (A. Brauer and H. Rohrbach, eds.) vol.II. pp.416-427, Springer-Verlag, Berlin 1973).

[30] Sen, A.K. (1997) - *On Economic Inequality*" - (Second eds), extended edition with J. Foster, Oxford: Clarendon Press.

[31] Sen, A.K (1980) - "Equality of what?" - in McMurrin, S. ed., *Tunner Lectures on Human Values*; vol.1, Cambridge: Cambridge University Press.

[32] Shorrocks, A.F. (1984) - Inequality decomposition by population subgroups - *Econometrica* 52, 1369-1385.

[33] Tsui, K.Y. (1995) - Multidimensional generalizations of the relative and absolute inequality indices: The Atkinson-Kolm-Sen approach - *Journal of Economic Theory* 67, 251-265.

[34] Tsui, K.Y. (1999) - Multidimensional inequality and multidimensional generalized entropy measures: an axiomatic derivation - *Social Choice and Welfare* 16, 145-157.

[35] Weymark, J.A. (2004) - The normative approach to the measurement of multidimensional inequality - in this volume.

[36] Webster, R. (1994) - *Convexity* - Oxford Science Publication, Oxford University Press.

Dipartimento di Economia "R. M. Goodwin", Università di Siena - TEL. +39.338.300.25.26, FAX. +39.0577.232.661, E-MAIL: ERNESTO.SAVAGLIO@UNISI.IT