Capitolo 2. La distribuzione del reddito

2.1 Reddito individuale, reddito familiare e reddito equivalente

Nell'analisi della distribuzione dei redditi e delle ricchezze, oltre alla necessità di analizzare le varie componenti del reddito, è necessario anche individuare l'unità di riferimento. Ovviamente a seconda dei dati rilevati può essere più opportuno avere informazioni relative ai singoli individui (che non sono esposte alle differenze derivanti dalle dimensioni e dalla struttura delle famiglie) oppure alle famiglie (avendo cura di attribuire a ciascun individuo il reddito equivalente della famiglia di cui è membro, mediante l'impiego di "scale di equivalenza", che verranno analizzate nell'Appendice 2). Spesso, la scelta delle unità di rilevazione è condizionata anche dal tipo di fonte disponibile.

L'impostazione individualistica, tipica dell'economia del benessere, indica l'individuo come naturale punto di riferimento per le teorie normative. Senza dubbio, però, per ragioni demografiche ed economiche, il benessere individuale ha come importante punto di riferimento la famiglia. Infatti, la famiglia è essenziale alla sopravvivenza stessa nelle fasi in cui un individuo non è autosufficiente economicamente ed inoltre l'organizzazione della vita all'interno di una famiglia consente di realizzare numerose economie di scala.

Alla luce di queste affermazioni la famiglia appare come l'unità di riferimento più appropriata per una valutazione del reddito. La definizione dell'aggregato famiglia, però, non è un compito semplice; possiamo distinguere tra famiglia anagrafica (marito, moglie e figli a carico: *family* nella letteratura inglese) e famiglia in senso esteso, che alla precedente aggiunge anche altri individui residenti nella stessa abitazione (nonni, figli maggiorenni a carico o no, altri individui legati da vincoli non di sangue: *household* nella letteratura inglese).

Di seguito quando parleremo di famiglia faremo sempre riferimento a quella definita in senso esteso, ovvero individui legati tra di loro da un qualche vincolo, di sangue, matrimonio, affetto, o altro, che convivono nella stessa abitazione e mettono in comune la totalità o parte dei loro redditi.

Esiste però un problema sostanziale, ovvero è lecito ritenere che tutte le persone all'interno di una data famiglia godano di un uguale livello di benessere? Infatti un individuo che percepisce un reddito potrebbe disporne di una parte solo per sé senza metterlo in comune con gli altri. Esiste però un ostacolo per poter fare analisi basandosi su queste assunzioni, ovvero per tener conto della distribuzione intra-familiare spesso non si hanno le informazioni sufficienti.

Infatti, anche se per quasi tutte le passività e le attività di una famiglia è identificabile il titolare, come ad esempio il proprietario della casa, l'intestatario del conto corrente, o il percettore di una busta paga, è ragionevole ipotizzare che tali intestazioni seguano criteri convenzionali all'interno delle famiglie, legati ad aspetti pratici, che però non riflettono necessariamente l'effettiva disponibilità dei beni posseduti o l'utilizzo dei redditi percepiti.

Nelle analisi di seguito riportate non sarà fatta quindi un'analisi di tipo intra-familiare, ed a ciascun individuo sarà attribuito il reddito della famiglia di appartenenza.

Per confrontare i redditi di famiglie dalla diversa struttura, per numerosità e composizione, si fa uso delle scale di equivalenza, ovvero di un vettore di coefficienti che standardizza le eterogeneità demografiche associando a ciascuna tipologia familiare un numero di componenti equivalenti.

La divisione del reddito familiare (derivante dalla somma di tutti i redditi percepiti da ogni componente sommati ai trasferimenti che avvengono a livello familiare) per il coefficiente della scala, determina il reddito equivalente; una misura che non dipende dalle caratteristiche demografiche della famiglia. Questa sarà la variabile utilizzata nelle analisi esposte nei capitoli seguenti.

2.2 Modelli per la distribuzione del reddito

Quando si analizzano le distribuzioni di reddito, si può cercare di ricondurle a modelli teorici, sia per scopi puramente descrittivi, sia per formulare schematicamente una determinata teoria sui meccanismi distributivi.

Nel caso in cui si utilizzino i modelli per scopi descrittivi, devono essere trovate delle funzioni di densità o di ripartizione per interpolare i dati nel miglior modo possibile. Una volta individuati, questi modelli possono essere impiegati per stimare dati che non sono disponibili, per correggere errori non sistematici e per utilizzare le stime dei parametri in modelli più ampi.

Nel caso in cui si voglia formulare una teoria sui meccanismi distributivi, i modelli vengono proposti non solo per rappresentare mediante parametri le distribuzioni osservate, ma vengono utilizzati come schemi in grado di descrivere in forma essenziale la realtà economico-sociale che presumibilmente ha generato le distribuzioni stesse (Statistica Economica, a cura di Giorgio Marbach).

Questi modelli, spesso, vengono mutuati da altre discipline come la demografia, ma anche da discipline molto diverse dall'economia, come per esempio la meteorologia; altri modelli hanno invece un'origine puramente stocastica e derivano dall'applicazione di tecniche e risultati che provengono dal calcolo delle probabilità.

La valutazione di questi modelli è però abbastanza difficile, in quanto ad oggi mancano ancora termini di riferimento all'interno delle discipline economiche stesse.

Quello che si tende a fare, più semplicemente, è trovare funzioni matematiche che generino distribuzioni di frequenza che si adattino bene alle distribuzioni di reddito osservate, e che abbiano anche un fondamento teorico. Esistono numerosi modelli tra i quali scegliere la forma analitica che meglio si adatta ad interpolare la distribuzione empirica di nostro interesse. Di seguito verranno illustrati alcuni modelli parametrici.

Il primo modello è stato proposto da Pareto nel 1895, è il più antico ed anche il più noto, introdotto con l'obiettivo di interpolare la parte superiore delle code della distribuzione; lascia però a desiderare per quanto riguarda la descrizione delle code inferiori. Le analisi condotte da Pareto generarono numerosi dibattiti relativamente agli effetti della crescita economica sulla disuguaglianza del reddito. Gini non approvava le opinioni di Pareto, secondo le quali la crescita economica conduceva ad una minore disuguaglianza nella distribuzione dei redditi tra gli individui, e quindi decise di proporre una misura della disuguaglianza, che tuttora è di uso corrente.

Il modello di Pareto si basa sulla constatazione che in numerose distribuzioni di reddito osservate, l'elasticità della funzione $G(y)=1-F(y)$ (frequenze retrocumulate delle unità statistiche) rispetto a y (limiti inferiori delle classi di reddito) risulta all'incirca costante. Da questo deriva la seguente formulazione matematica del modello di Pareto:

$$G(y) = 1 - F(y) = Ky^{-\alpha} \quad (2.1)$$

dove K e α sono costanti determinabili sulla base dei dati osservati. Pareto propose anche due forme più generali della (2.1):

$$G(y) = K(y+a)^{-\alpha} \tag{2.2}$$

$$G(y) = K(y+a)^{-\alpha} e^{\beta y} \tag{2.3}$$

La (2.1) è l'equazione di un'iperbole, che in forma logaritmica diventa:

$$\log G(y) = \log K - \alpha \log y \tag{2.4}$$

Questa è l'equazione di una retta con parametri *logK* e α, i quali possono essere stimati anche con semplici metodi di stima quali il metodo dei minimi quadrati ordinari (OLS).

Il parametro α è l'elasticità di *G(y)* rispetto a y e può essere interpretato come un indice di ineguaglianza (si veda il paragrafo 2.3).

Applicando questo modello ai dati in suo possesso, Pareto osservò che il parametro α variava poco nel tempo e nello spazio e questo; a suo parere, questo fatto avrebbe dimostrato che la distribuzione del reddito dipende da fattori non modificabili. Queste sue affermazioni furono oggetto di numerose critiche.

Il modello di Pareto si basa su funzioni di densità zero-modali, mentre le distribuzioni del reddito, nella realtà sono unimodali. Il modello si adatta solo alle distribuzioni di redditi elevati, non inferiori alla moda della distribuzione completa; Pareto era a conoscenza di questo limite, ma difendeva il suo modello affermando che la sua valenza era limitata all'intervallo di redditi assegnato. Relativamente alla poca variabilità del parametro α, successivi studi hanno dimostrato che si tratta di una caratteristica tecnica intrinseca al modello.

Un altro importante modello è il Lognormale. In generale, una variabile ha distribuzione Lognormale, se il suo logaritmo ha distribuzione normale:

$$F(y) = N\left[\left(\frac{\log y - \mu}{\sigma}\right);0,1\right] \tag{2.5}$$

dove μ è il logaritmo della media geometrica dei redditi e σ^2 la varianza dei redditi.

Questo modello richiama la legge dell'effetto proporzionale. Per tale legge, data una variabile y (nel nostro caso il reddito), e ammettendo che:

i) essa si modifichi per l'azione di innumerevoli cause indipendenti, ciascuna delle quali produce un effetto molto piccolo rispetto all'effetto totale;

ii) la variabile *dy* risulti proporzionale a y;

si ricava che *dy/y=d(logy)* ha distribuzione normale (Statistica Economica, a cura di Giorgio Marbach).

Il modello Lognormale è molto diffuso, grazie anche al contributo di Gibrat (1931), che fornì le basi teoriche per considerare questi modelli in due parametri per misurare l'entità delle distribuzioni del reddito. Esso è stato successivamente riesaminato anche da Aitchinson e Brown (1969), ma non può essere considerato uno dei migliori, soprattutto quando lo scopo consiste nell'interpolare le code della distribuzione.

In letteratura sono stati introdotti altri modelli distributivi in due parametri: la distribuzione Gamma, proposta da Ammon (1895) e successivamente reintrodotta da Salem e Mount (1974) per analizzare i redditi degli Stati Uniti, e la distribuzione di Weibull, presentata da Bartels e Van Metile (1975).

Altri modelli fanno riferimento alle distribuzioni generalizzate Gamma e Beta, oppure appartengono alla famiglia dei modelli di distribuzione di Burr: si tratta del modello Singh-Maddala (1976), noto anche come Burr12, e il modello di Dagum (1977), noto come Burr3.

Il modello di Singh-Maddala, si basa sull'ipotesi che l'elasticità della funzione *G(y)* (frequenze retro cumulate delle famiglie) rispetto a *y* (limiti inferiori delle classi di reddito), cresca all'aumentare dei livelli di reddito, per poi stabilizzarsi. Inoltre viene ipotizzato che la difficoltà a raggiungere livelli di reddito più elevati, vari al crescere di tali livelli di reddito con un andamento di tipo logistico.

Il modello ha la seguente forma:

$$F(y) = 1 - G(y) = 1 - (1 + ay^b)^{-c} \qquad (2.6)$$

I parametri (*a,b,c*) possono essere stimati con la regressione non lineare, minimizzando la somma dei quadrati degli scarti tra i valori osservati *G(y)* e i corrispondenti valori stimati.

Il modello di Dagum, si basa su altre ipotesi diverse da quelle viste per il precedente. Infatti Dagum affermava che l'elasticità della frazione cumulata delle famiglie *F(y)* rispetto a *y* è decrescente al crescere di tale frazione, e che il tasso di decremento è funzione della frazione stessa.

Questo modello ha tre parametri: *b* che è una costante di dimensione, che dipende dall'unità di misura della variabile *y*; *p* e *a* sono parametri che dipendono dal grado di ineguaglianza della distribuzione. Questo modello può essere generalizzato mediante l'introduzione di un ulteriore parametro α, ed assume la seguente forma:

$$F(y) = \alpha + (1-\alpha)\left(1 + y / b^{-a}\right)^{-p} \tag{2.7}$$

Possiamo avere tre tipi di distribuzione di Dagum: 1, 2 e 3, che corrispondono rispettivamente a α=0 (modello con 3 parametri), 0<α<1, e α<0. Il modello Dagum 2, ammette che possano esserci anche individui con reddito negativo o nullo con $F(0)= \alpha$. Il modello Dagum 3 richiede che tutti gli individui abbiano un reddito maggiore di un reddito minimo assegnato, y_0>0. Anche in questo caso, la stima dei parametri del modello può essere effettuata con vari metodi econometrici.

Vediamo adesso altri modelli come il generalizzato Beta di primo e secondo tipo (GB1 e GB2); si tratta di distribuzioni con quattro parametri, molto importati perché non solo interpolano bene i dati, ma anche perché includono tutti i modelli menzionati sopra, come casi limite o casi particolari.

Il successo empirico dei modelli GB2 è stato confermato da quelli teorici di generazione del reddito elaborati di Parker, che hanno mostrato come i guadagni seguano una distribuzione del tipo GB2.

Questo modello è stato applicato in numerosi studi, tra i quali quello condotto da Bandourian *et al.* (2003) applicando il Beta generalizzato ai dati di 23 paesi per vari anni. Gli studiosi hanno potuto osservare come i modelli di Weibull, Dagum e il Beta generalizzato del secondo tipo, sono solitamente quelli che si adattano meglio ai dati relativi al reddito. Per fare delle valutazioni comparative tra i risultati ottenuti utilizzando i vari modelli, non sempre possono essere adoperate come termine di paragone le devianze dei residui. Questo perché i vari modelli hanno un numero di parametri diversi, e possono essere quindi applicati l'analisi della varianza e il test F di Fisher, che consentono di accertare, con una certa probabilità, quale di questi fornisce un'interpolazione significativamente migliore degli altri.

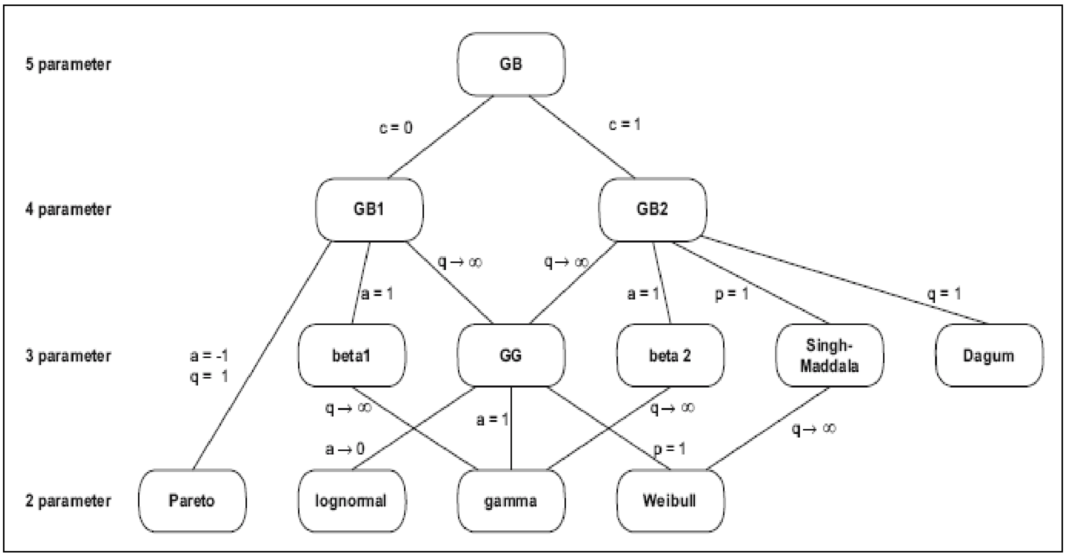
Analizziamo adesso nello specifico la famiglie delle distribuzioni Beta generalizzate.

Queste sono definite dalla loro funzione di densità delle probabilità come segue:

$$GB(y;a,b,c,p,q) = \frac{|a| y^{ap-1}\left(1-(1-c)(y/b)^a\right)^{q-1}}{b^{ap} B(p,q)\left(1+c(y/b)^a\right)^{p+q}} \text{ per } 0 < y^a < \frac{b^a}{1-c} \tag{2.8}$$

e zero altrimenti, dove 0≤c≤1, b,p,q>0, e $B(p,q)$ indica la funzione Beta.

Figura 2.1: Albero delle distribuzioni



Fonte: Dastrup *et al.* (2007).

Le distribuzioni GB1 e GB2 corrispondono alla distribuzione GB quando il parametro c è uguale rispettivamente a 0 ed a 1:

$$GB1(y;a,b,p,q)=\frac{|a|y^{ap-1}\left(1-(y/b)^a\right)^{q-1}}{b^{ap}B(p,q)}=GB(y;a,b,c=0,p,q) \tag{2.9}$$

$$GB2(y;a,b,p,q)=\frac{|a|y^{ap-1}}{b^{ap}B(p,q)\left(1+(y/b)^a\right)^{p+q}}=GB(y;a,b,c=1,p,q) \tag{2.10}$$

Vediamo adesso come i modelli di Dagum 1, modello con tre parametri, e i modello Singh-Maddala, siano casi particolari del modello GB:

DAGUM(y; a, b, p)=GB2(y; a, b, p, q=1) (2.11)

SM(y; a, b, q)=GB2(y; a, b, p=1, q) (2.12)

La distribuzione Gamma generalizzata, è un caso limite della distribuzione GB2:

$$GG(y;a,\beta,p)=\lim_{q\to\infty}GB\left(y;a,b=q^{1/\alpha}\beta,c=1,p,q\right)=\frac{y^{ap-1}e^{-(y/\beta)^a}}{\beta^a\Gamma(p)} \tag{2.13}$$

dove $\Gamma(\ )$ indica la funzione matematica Gamma. Il modello Gamma generalizzato include anche la distribuzione di Weibull e la distribuzione Gamma come casi speciali, cioè corrispondenti rispettivamente a p=1 e a=1.

Anche la Lognormale può essere espressa come un caso limite della Gamma generalizzata:

$$LN(y;\mu,\sigma)=\lim_{a\to 0} GG\left(y;a,\beta=\left(\sigma^2 a^2\right)^{1/a}, p=(a\mu+1)/\sigma^2 a^2\right) \tag{2.14}$$

Uno strumento molto utile per visualizzare le relazioni considerate sopra ed altri casi particolari menzionati, è l'utilizzo di un albero della distribuzione come quello riportato nella Figura 2.1. Tutte le funzioni di densità citate, sono riportate nella Tabella 2.1.

Tabella 2.1. Funzioni di densità della probabilità e coefficiente di Gini per i diversi modelli di reddito

Generalized Beta 1	$GB1(y;a,b,p,q)=\frac{\lvert a\rvert y^{ap-1}(1-(y/b)^a)^{q-1}}{b^{ap}B(p,q)}$	G_{GB1}
Generalized Beta 2	$GB2(y;a,b,p,q)=\frac{\lvert a\rvert y^{ap-1}}{b^{ap}B(p,q)(1+(y/b)^a)^{p+q}}$	G_{GB2}
Beta 1	$B1(y;b,p,q)=\frac{y^{p-1}(1-y/b)^{q-1}}{b^p B(p,q)}$	$\frac{B(p+q,1/2)B(p+1/2,1/2)}{B(q,1/2)\pi}$
Generalized Gamma	$GG(y;a,b,p)=\frac{\lvert a\rvert y^{ap-1}}{b^{ap}\Gamma(p)}\exp\left(-(y/b)^a\right)$	G_{GG}
Beta 2	$B2(y;b,p,q)=\frac{y^{p-1}}{b^p B(p,q)(1+(y/b)^a)^{p+q}}$	$\frac{2B(2p,2q-1)}{pB^2(p,q)}$
Singh–Maddala	$SM(y;a,b,q)=\frac{\lvert a\rvert q y^{a-1}}{b^a(1+y/b)^{q+1}}$	$1-\frac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(q-1/a)\Gamma(2q)}$
Dagum	$DAGUM(y;a,b,p)=\frac{\lvert a\rvert p y^{ap-1}}{b^{ap}(1+(y/b)^a)^{p+1}}$	$\frac{\Gamma(p)\Gamma(2p+1/a)}{\Gamma(p+1/a)\Gamma(2p)}-1$
Pareto	$PARETO(y;b,p)=\frac{pb^p}{y^{p+1}},$	$\frac{1}{2p-1}$
Lognormal	$LN(y;\mu,\sigma)=\frac{1}{y\sigma\sqrt{2\pi}}\exp\left(-\left(\frac{\ln y-\mu}{\sigma\sqrt{2}}\right)^2\right)$	$2N\left[\sigma/\sqrt{2};0,1\right]-1$
Gamma	$GAM(y;b,p)=\frac{y^{ap-1}}{b^p\Gamma(p)}\exp(-y/b)$	$\frac{\Gamma(p+1/2)}{\Gamma(p+1)\sqrt{\pi}}$
Weibull	$W(y;a,b)=\frac{\lvert a\rvert y^{a-1}}{b^a}\exp\left(-(y/b)^a\right)$	$1-\left(1/2^{1/a}\right)$

La tabella, tratta da Dastrup *et al.* (2007), riporta anche un indice di concentrazione teorica, l'indice di Gini, che verrà introdotto nel seguente paragrafo 2.3. Da notare che l'indice di Gini per le distribuzioni GB1, GB2 e GG può essere calcolato come:

$$G_{GB1}=\frac{B(2p+1/a,q)}{B(p,q)B(p+1/a,q)p(ap+1)}\,{}_4F_3\left[\begin{matrix}2p+1/a,p,p+1/a,1-q;1\\ 2p+q+1/a,p+1,p+1/a+1;\end{matrix}\right]$$

$$G_{GB2}=\frac{B(2q-1/a,2p+1/a)}{B(p,q)B(p+1/a,q-1/a)}\left\{\begin{matrix}\left(\frac{1}{p}\right){}_3F_2\left[\begin{matrix}1,p+q,2p+1/a;1\\ p+1,2(p+q);\end{matrix}\right]\\ -\left(\frac{1}{p+1/a}\right){}_3F_2\left[\begin{matrix}1,p+q,2p+1/a;1\\ p+1/a+1,2(p+q);\end{matrix}\right]\end{matrix}\right\}$$

$$G_{GG}=\frac{1}{2^{2p+1/a}B(p,p+1/a)}\left\{\left(\frac{1}{p}\right){}_2F_1\left[\begin{matrix}1,2p+1/a;\frac{1}{2}\\ p+1;\end{matrix}\right]-\left(\frac{1}{p+1/a}\right){}_2F_1\left[\begin{matrix}1,2p+1/a;\frac{1}{2}\\ p+1/a+1;\end{matrix}\right]\right\}$$

I parametri dei modelli visti sopra possono essere stimati in vari modi, tra cui la massima verosimiglianza, che si adatta sia per le osservazioni riportate a livello di singoli individui, sia per i valori raggruppati in classi. Rispettivamente i parametri della distribuzione sono selezionati per massimizzare le seguenti funzioni:

$$l(\theta) = \sum_{i=1}^{N} \ln\left(f_d\left(y_i : \theta\right)\right) \tag{2.15}$$

$$l(\theta) = \ln(N!) + \sum_{i=1}^{g} \left\{ n_i \ln\left[p_i(\theta)\right] - \ln(n_i!) \right\} \tag{2.16}$$

dove $p_i(\theta)=F_d(Y_i: \theta)-F_d(Y_{i-1}: \theta)$, $f_d(\,)$ e $f_d(\,)$ denotano la funzione di densità e la funzione cumulata di tipo d; θ è un vettore contenente la distribuzione dei parametri; Y_i e Y_{i-1} sono l'estremo superiore e inferiore dell'*i-esimo* dei g gruppi di dati; n_i è il numero di osservazioni dell'*i-esimo* gruppo, e N è il numero totale di osservazioni.

Per stimare i parametri non noti possono essere utilizzati dei metodi numerici di ottimizzazione.

Il test basato sul rapporto delle verosimiglianze, definito come segue:

$$LR = 2\left[\hat{l} - \hat{l}^*\right] \stackrel{a}{\sim} \chi^2(r) \tag{2.17}$$

può essere utilizzato per confrontare distribuzioni che si trovano sullo stesso ramo (si veda la Figura 2.1). Nel test LR $\hat{l}$ e $\hat{l}^*$ rappresentano rispettivamente i valori di verosimiglianza corrispondenti ad un modello senza vincoli e ad uno vincolato, invece r è il numero di gradi di libertà del Chi-quadro calcolato come differenza tra il numero di parametri stimati nei due modelli specificati. Il miglior adattamento, ai dati rilevati, del modello GB2 rispetto al modello di Dagum, può quindi essere verificato con una distribuzione Chi-quadro con un grado di libertà.

Invece la statistica LR non può essere utilizzata per confrontare modelli che non si trovano sullo stesso ramo, ad esempio la distribuzione Gamma generalizzata e la distribuzione Beta del secondo tipo.

In questi casi è necessario utilizzare i valori della somma dei quadrati degli errori (SSE), la somma degli errori assoluti (SAE)[1], e la bontà di adattamento delle misure, mediante il Chi-quadro; questi sono definiti come segue:

---

[1] L'acronimo SAE utilizzato nel presente Capitolo 2 non deve essere confuso con l'acronimo di *Small Area Estimation* utilizzato nel Capitolo 5.

$$SSE = \sum_{i=1}^{g} \left(\frac{n_i}{N} - p_i\left(\hat{\theta}\right) \right)^2 \qquad (2.18)$$

$$SAE = \sum_{i=1}^{2} \left| \frac{n_i}{N} - p_i\left(\hat{\theta}\right) \right| \qquad (2.19)$$

$$\chi^2 = N \sum_{i=1}^{g} \left[\left(\frac{n_i}{N} - p_i\left(\hat{\theta}\right) \right)^2 \Big/ p_i\left(\hat{\theta}\right) \right] \qquad (2.20)$$

dove $\hat{\theta}$ è il vettore delle stime dei parametri. Il χ^2 è distribuito in modo asintotico come un χ^2 con un numero di gradi di libertà pari a uno meno la differenza tra il numero dei gruppi e il numero di parametri stimati (Cox e Hinckley, 1974).

Solitamente questi tre criteri, SSE, SAE, e χ^2 sono in accordo relativamente alla bontà di adattamento del modello ai dati, ma non sempre il risultato è univoco.

Le relazioni tra le distribuzioni, illustrate nella Figura 2.1, garantiscono che un modello che si trova all'apice del ramo, interpola bene i dati almeno come quelli che si trovano lungo ciascuno dei rami da esso originati. Questo comunque non afferma la superiorità in modo statisticamente significativo di questi modelli rispetto agli altri.

Così, per esempio, si può considerare che la distribuzione Gamma generalizzata possa fornire miglioramenti statisticamente significativi rispetto alle distribuzioni di Weibull e Gamma. Le differenze tra la distribuzione GB2 e le distribuzioni della Dagum e SM, sono statisticamente significative, dal momento in cui le differenze tra la GB e la GB2 non utilizzano valori critici convenzionali.

La scelta della funzione di distribuzione da utilizzare, può influenzare la stima del livello di disuguaglianza misurato dall'indice di Gini, mediante l'utilizzo delle formule riportate nella Tabella 2.1.

Nel tempo sono stati effettuati numerosi studi e applicazioni dei modelli sopra riportati. Un esempio è costituito dallo studio effettuato da Dastrup, Hartshorn e McDonald (2007), riportato in un articolo pubblicato sul *Journal of Economic Inequality*, in una *special issue* edita in memoria del compianto Prof. Dagum (Betti e Lemmi, 2007).

In tale studio vengono confrontate le capacità di 11 diverse funzioni di distribuzione di adattarsi ai dati rilevati in 11 paesi diversi, per un certo periodo di tempo. La variabile oggetto di studio è il reddito, e questo viene definito in tre modi diversi utilizzando i dati del LIS (*Luxembourg Income Study*): si parla di guadagni, di reddito totale e di reddito disponibile personale.

In ogni caso è stata valutata la capacità dei modelli di interpolare i dati, ed è stato calcolato l'indice di Gini. Relativamente alla forma funzionale in due parametri, quella che si adatta meglio ai dati, risulta la distribuzione di Weibull per i guadagni e la Gamma per il reddito totale e per il reddito disponibile. Relativamente alle distribuzioni in tre parametri, è la distribuzione di Dagum quella che si adatta meglio ai guadagni, mentre per le altre due connotazioni di reddito, è la distribuzione Gamma generalizzata. Si potrebbe così continuare per altre distribuzioni, ma questo ci fa capire come non sia possibile fornire una valutazione univoca della capacità del modello di adattarsi ai dati, e di come questa vari in base alle diverse definizioni di reddito utilizzate.

Sono stati inoltre calcolati i coefficienti di Gini, utilizzando le diverse definizioni riportate nella Tabella 2.1. È emerso che gli indici stimati con il modello Lognormale, presentano un valore più alto rispetto a quelli calcolati mediante tutti gli altri modelli.

Studi simili sono stati effettuati anche su dati italiani. Sono stati interpolati tre importanti modelli, quello Lognormale, quello di Singh-Maddala e quello di Dagum, ai dati ISTAT della distribuzione del reddito del 1985. È emerso che il modello di Dagum migliora significativamente l'interpolazione rispetto al modello Lognormale e anche, al livello di significatività del 10%, rispetto a quello di Singh-Maddala (Statistica Economica, a cura di Giorgio Marbach). Ciò conferma quanto detto sopra, ovvero che non è possibile esprimere a priori un giudizio relativamente alla bontà di un modello per interpolare i dati, in quanto questo giudizio può variare a seconda delle situazioni e dei dati disponibili; ciò suggerisce nelle varie situazioni di applicare più modelli e effettuare successivamente le valutazioni del caso.

2.3 Indici di Concentrazione e Disuguaglianza

Lo scopo di questo paragrafo consiste nel prendere in considerazione i più importanti indici di concentrazione e disuguaglianza[2]. Il più noto tra questi è sicuramente l'indice di Gini. Prima di analizzarlo nello specifico, consideriamo un altro strumento molto

[2] I termini disuguaglianza, diseguaglianza e ineguaglianza sono sinonimi dello stesso concetto.

utile per l'analisi della disuguaglianza, la curva di Lorenz, la quale consente poi di calcolare l'indice appena citato.

La curva di Lorenz (Max O. Lorenz, 1905) è uno strumento grafico proposto per l'analisi della disparità nella distribuzione di un certo attributo. Solitamente questo attributo è il reddito, ma non sono escluse altre numerose applicazioni di questo strumento. Di seguito ci concentreremo proprio sul reddito.

La curva di Lorenz è una misura relativa della disuguaglianza che consente di rappresentare graficamente la quota di reddito totale percepita da una porzione (frazione cumulata) di popolazione ordinata per livelli non decrescenti di reddito.

La curva di Lorenz è la relazione che lega ciascuna quota cumulata della popolazione con la corrispondente quota del reddito totale posseduta da queste persone.

Esistono due casi estremi di distribuzione tra gli individui: l'equidistribuzione, dove ciascun individuo ha la stessa quantità della media cioè il k% della popolazione ha il k% del reddito totale, e la massima concentrazione, nella quale un solo individuo ha tutto il reddito e gli altri hanno un reddito pari a zero.

Poiché per analizzare la disuguaglianza gli individui vengono ordinati per redditi crescenti, avremo una curva che si trova sempre al di sotto della bisettrice, in quanto il k% più povero della popolazione possiede sicuramente meno del k% del reddito complessivo, ed avrà una inclinazione positiva e crescente.

Da un punto di vista di analisi delle funzioni, la curva è sempre convessa, e questo deriva dal fatto che – come appena accennato sopra - per costruirla tutte le unità di analisi sono ordinate per livelli di reddito non decrescenti.

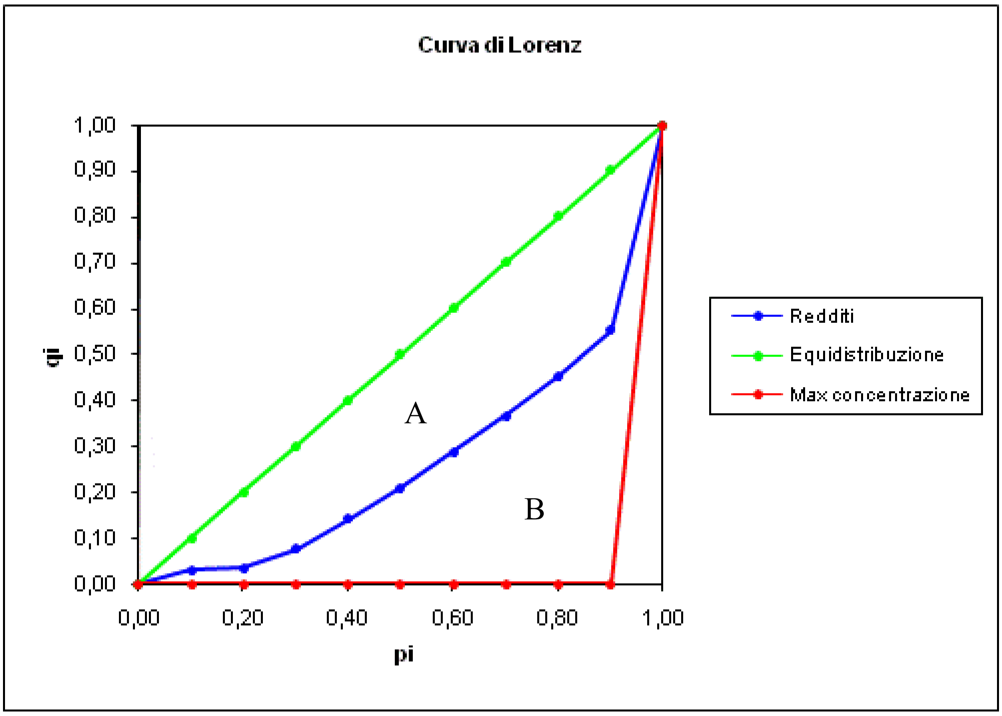
Come già affermato la curva di Lorenz informa sul grado di disuguaglianza della distribuzione, e quanto più questa curva si trova vicina alla bisettrice e tanto più la distribuzione è egualitaria.

Formuliamo adesso in termini matematici la definizione della curva di Lorenz. Per prima cosa ordiniamo gli individui per livelli non decrescenti di reddito ed otteniamo y(i), tale che y(1)≤ y(2)≤…≤ y(N). Successivamente determiniamo la quota di persone con redditi inferiori o uguali al reddito posseduto dall'n-esima persona, ovvero n/N, e la corrispondente frazione del reddito totale, $\sum_{i=1}^{n} y_{(i)} \Big/ \sum_{i=1}^{N} y_i$.

La curva di Lorenz sarà l'insieme dei punti di coordinate $\left(\frac{n}{N}, \frac{\sum_{i=1}^{n} y_{(i)}}{\sum_{i=1}^{N} y_i} \right)$, dove $1 \le n \le N$, e n/N coincide con F(y(n)).

Applicando questa definizione abbiamo che L(0)=0 e L(1)=1.

Figura 2.2. Curva di Lorenz



Nel grafico riportato nella Figura 2.2 sono state inserite, oltre ad una curva di Lorenz esemplificativa, anche la retta relativa all'equidistribuzione o equiripartizione (la retta a 45°), e quella relativa alla massima concentrazione (una L rovesciata, in rosso nel grafico).

Se avessimo due distribuzioni del reddito, per esempio X e Y, e la curva di X giacesse sempre al di sopra della curva di Y, si potrebbe dire che "X domina Y secondo Lorenz", e questo indicherebbe che il reddito complessivo di X è distribuito in modo meno diseguale di quello di Y. Se le due curve si incrociassero, allora non potremmo fare un confronto tra le due distribuzioni e non potremmo affermare niente sulla maggiore o minore uguaglianza di una distribuzione rispetto all'altra. Possiamo quindi affermare che l'ordinamento stabilito dal criterio di dominanza di Lorenz è incompleto.

La curva non fornisce informazioni sul reddito medio della popolazione.

A partire dall'area compresa tra la curva e la retta di equidistribuzione si possono definire gli indici di concentrazione, che valutano la tendenza della variabile oggetto di studio a concentrarsi su poche delle n unità statistiche oggetto d'indagine. Come già detto, una misura sintetica della disuguaglianza molto diffusa è l'indice di Gini (1914).

Questo indice fornisce un'immediata interpretazione geometrica della curva di Lorenz.

L'indice di Gini è calcolato come il rapporto tra l'area compresa tra la diagonale (la linea di uguaglianza perfetta, retta a 45°) e la curva di Lorenz, e l'area del triangolo sottesa alla diagonale. Utilizzando la simbologia riportata nella Figura 2.2:

$$G = \frac{A}{A+B} \tag{2.21}$$

Poiché l'area (A+B) e pari a ½, allora G=2A=2(1/2-B)=1-2B. Se il reddito fosse distribuito in modo perfettamente uguale la curva di Lorenz corrisponderebbe alla bisettrice, A=0 e G=0. Se invece tutto il reddito appartenesse ad un solo individuo, l'area B sarebbe uguale a 0 e l'area A sarebbe uguale a ½, da cui G=1.

Esistono molte altre formule alternative per calcolare il valore dell'indice di Gini (Xu, 2004). Una tra le più diffuse si può ottenere attraverso alcuni passaggi partendo dalla definizione (2.21):

$$G = 1 + \frac{1}{N} - \frac{2}{N^2 \bar{y}} \sum_{i=1}^{N} (N+1-i) y_{(i)} \tag{2.22}$$

L'indice in questa formulazione può essere interpretato come il gap atteso tra i redditi di due individui selezionati in modo casuale dalla popolazione:

$$G = \frac{2}{N^2 \bar{y}} \sum_{i=1}^{N} i \left(y_{(i)} - \bar{y} \right) \tag{2.23}$$

dove N è la numerosità della popolazione, $\bar{y}$ è il reddito medio, F(y) è la funzione di densità cumulata degli individui, ordinati per valori non decrescenti del reddito.

Per calcolare l'indice con la (2.23) è sufficiente conoscere la media del reddito e la covarianza tra il reddito e la sua funzione di densità cumulata.

Ovviamente, per tutte le formulazioni proposte in letteratura, l'indice assume valori compresi tra 0, dove tutti gli individui percepiscono lo stesso reddito (assenza di ineguaglianza) e 1, dove solo un individuo possiede tutto il reddito disponibile (perfetta ineguaglianza). Maggiore risulta essere il valore di questa area e maggiore risulta essere la disuguaglianza nella distribuzione del reddito.

Nel prossimo Capitolo 3 sono introdotte – attraverso un certo numero di assiomi – alcune proprietà alle quale è auspicabile i principali indici di disuguaglianza e povertà possano attenersi. Alcuni di questi assiomi sono stati formulati con il preciso scopo di proporre nuovi indici di povertà (ad esempio l'indice di Sen, 1976); per questa ragione si è deciso di presentare in maniera compatta tutti gli assiomi nel Capitolo 3, al quale rimandiamo per comprendere meglio il commento agli indici di disuguaglianza riportati qui di seguito.

L'indice di Gini soddisfa le proprietà di simmetria, di indipendenza dalla media e dalla popolazione e il principio del trasferimento. Facendo riferimento a questa ultima proprietà, e considerando la formula (2.22), possiamo notare come la sensibilità di G ad un trasferimento di reddito da un ricco ad un povero non dipende dai livelli di reddito dei due individui, bensì dalla loro differenza di rango nella scala dei redditi[3].

L'indice di Gini non soddisfa il principio di trasferimento decrescente, e nemmeno la proprietà di scomponibilità esatta tra gruppi della popolazione.

Un altro limite di questo indice, che lo accomuna con tutte le misure statiche della dispersione, consiste nel fatto che non fornisce informazioni sul grado di asimmetria di una distribuzione. Dopo una dettagliata descrizione dell'indice di Gini, di seguito sono riportati altri indici di disuguaglianza.

La varianza, può essere considerata un indicatore di ineguaglianza, e misura la media degli scarti dalla media, elevati al quadrato[4]:

$$V = \frac{\sum_{i=1}^{N}(y_i - \bar{y})^2}{N} \tag{2.24}$$

Questo indice soddisfa le proprietà di anonimità, di indipendenza dalla popolazione, il principio del trasferimento ed è scomponibile per gruppi. La varianza non soddisfa l'indipendenza dalla media, perché se tutti i redditi vengono moltiplicati per una costante k, allora V risulta moltiplicata per la costante k^2.

---

[3] Per una dimostrazione si veda Baldini e Toso (2004).

[4] $\bar{y}$ indica la media della distribuzione: $\bar{y} = \sum_{i=1}^{N} \frac{y_i}{N}$

Il coefficiente di variazione è invece il rapporto tra lo scarto quadratico medio (ovvero la radice quadrata della varianza) e la media:

$$C = \frac{\sqrt{V}}{\bar{y}} = \frac{\sigma}{\bar{y}} \tag{2.25}$$

Questo indice è un numero puro, ovvero indipendente dall'unità di misura, e quindi rispetta il principio dell'indipendenza dalla media. Inoltre, rispetta anche gli assiomi di anonimità, di simmetria, di indipendenza dalla popolazione e dalla media, e l'assioma di trasferimento di Bonferroni, Pigou e Dalton.

Sia la varianza che il coefficiente di variazione non rispettano il principio del trasferimento decrescente. Per avere un indice che rispetti tale assioma si deve effettuare una trasformazione logaritmica dei redditi.

La varianza dei logaritmi è una misura molto diffusa soprattutto per studiare la dispersione dei redditi da lavoro[5].

$$VL = \frac{\sum_{i=1}^{N} \left(\ln y_i - \ln \mu_g \right)^2}{N} = \frac{\sum_{i=1}^{N} \left[\ln\left(\frac{y_i}{\mu_g} \right) \right]^2}{N} \tag{2.26}$$

Questa misura oltre al principio del trasferimento decrescente, soddisfa gli assiomi di simmetria, di anonimità, di indipendenza dalla media e dalla popolazione; non rispetta però in generale l'assioma di trasferimento di Bonferroni, Pigou e Dalton.

Esiste poi un indice chiamato deviazione media relativa dei redditi, definita nel seguente modo:

$$RMD = \frac{\sum_{i=1}^{n} \left| y_i - \bar{y} \right|}{2n\bar{y}} \tag{2.27}$$

Un alto valore di questo indice denota grande ineguaglianza. Può essere interpretato come la percentuale media di reddito da trasferire da coloro che sono sopra la media a quelli che sono sotto la media, per raggiungere l'uguaglianza perfetta.

[5] $\mu_g = \left(\prod_{i=1}^{N} y_i \right)^{\frac{1}{N}}$

Consideriamo adesso un importante indice di disuguaglianza definito “etico”, in quanto deriva da una funzione di benessere sociale, identificando la disuguaglianza con la perdita di benessere sociale causata da una distribuzione disuguale dei redditi.

Atkinson (1970) ha presentato questo indice rivoluzionando la ricerca nell’ambito delle misure di disuguaglianza, evidenziando che ogni indice incorpora un giudizio di valore.

Questo indice si basa su una struttura delle preferenze collettive corrispondenti ad una funzione di benessere sociale, separabile in senso additivo e simmetrica nei redditi individuali.

Si possono definire più indici di Atkinson in base al valore di un parametro interpretato come coefficiente di avversione alla disuguaglianza; quando il parametro aumenta l’indice aumenta, poiché viene data più importanza alle code basse della distribuzione A_1, A_2, A_3. Questo indice si può definire nel modo seguente:

$$A = 1 - \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i}{\bar{y}} \right)^{1-\varepsilon} \right]^{1/(1-\varepsilon)} \tag{2.28}$$

dove ε è l’avversione alla disuguaglianza.

Molto importante sono anche gli indici di entropia generalizzati, *general entropy* (GE); si tratta di una classe di indici basata sulla teoria che misura il valore informativo (entropia) di un sistema di eventi incerti.

L’assunto su cui si basa la teoria è che quanto minore è la probabilità che l’evento incerto si verifichi, tanto maggiore è il valore informativo dell’evento stesso (e viceversa). Avvalendosi di alcune analogie formali e reinterpretando opportunamente alcuni concetti di base della teoria dell’informazione, è possibile riformulare la misura dell’entropia in termini di disuguaglianza (Cowell, 1995). Questi indici sono espressi nella seguente formula generale:

$$\mathrm{GE}(\alpha) = \frac{1}{\alpha^2 - \alpha} \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{y_i}{\bar{y}} \right)^{\alpha} - 1 \right] \tag{2.29}$$

La misura della GE attribuisce pesi diversi a diverse parti della distribuzione del reddito, a seconda del valore assegnato al parametro α, che può assumere qualsiasi valore reale. Per bassi valori di α, GE è più sensibile alle variazioni lungo le code della distribuzione, e per alti valori di α, GE è più sensibile alle variazioni nella parte centrale della distribuzione. I valori più utilizzati di α sono 0, 1 e 2.

GE(0) = L, detto anche deviazione logaritmica media (*mean logarithmic deviation*):

$$\text{GE}(0) = \text{L} = \frac{1}{N}\sum_{i=1}^{N}\ln\frac{\bar{y}}{y_i} \tag{2.30}$$

GE(1) = T, la misura di entropia di Theil (1967):

$$\text{GE}(1) = \text{T} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{\bar{y}}\ln\frac{y_i}{\bar{y}} \tag{2.31}$$

GE(2) è la metà del quadrato del coefficiente di variazione:

$$\text{GE}(2) = \frac{1}{2}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i}{\bar{y}}\right)^2 - 1\right] \tag{2.32}$$

La *general entropy* (GE) è una misura che varia tra 0 e 1, dove zero rappresenta una distribuzione equa, e alti valori dell'indice rappresentano più alti livelli di ineguaglianza, 1 si ha in presenza di massima ineguaglianza.

Tutti gli indici di entropia generalizzata soddisfano gli assiomi di simmetria, indipendenza dalla media e dalla popolazione; l'indice L è inoltre coerente con l'assioma B di trasferimento di Bonferroni, Pigou e Dalton.

Gli indici di entropia generalizzata godono della importante proprietà di scomposizione; possono essere scomposti per sottopopolazioni, individuando la quota di disuguaglianza presente in una popolazione dovuta alle diversità fra i gruppi (Regioni, Province, ecc…) e la quota di disuguaglianza dovuta alle diversità delle unità statistiche (famiglie o individui) all'interno dei gruppi. Per questa proprietà tali indici, ed in particolare l'indice T di Theil, sono molto utilizzati nell'analisi della disuguaglianza per piccole aree (si veda il Capitolo 5).

Infine, possiamo citare anche il rapporto interdecilico, ovvero una misura che esprime il rapporto tra i redditi dei soggetti che si trovano in diversi quantili della distribuzione. Questo rapporto si concentra su sezioni puntuali della distribuzione, ignorando del tutto ciò che accade in altre parti della stessa: non rispetta quindi il principio del trasferimento.

2.4 Un caso di studio: la scomposizione dell'indice di Theil in Albania[6]

… "Table 10 reports decomposition of one of the general entropy class inequality measures (GE(1), Theil Index) into its within area and between area components at various levels of aggregation. By definition, all of the inequality is within group when the group in question is the whole country or is the rural area or urban area, and all of it is between groups when each household is considered a separate group. GE(1) index is decomposable so that we are able to distinguish among the inequality due to differences between a certain level of disaggregated areas (prefectures, Districts, Communes, etc…) and the inequality due to the differences between households present in the disaggregated area. From Table 10 we can see that in the whole country and in both rural and urban areas, a large portion of the inequality is due to within-group inequality, even when the groups are relatively small, such as Communes or Municipalities.
Approximately, only 9% of the inequality in Albania is between Prefectures, 12,5% between Districts, and 17% between Communes/Municipalities." …

Table 10: Decomposition of the GE(1) inequality index (Theil).

Level of Decomposition	Number of Units	Within-Group Inequality	Between-Group Inequality	% Between-Group Inequality
Albania	**1**	**15.05**	**0**	**0**
Urban - rural	2	13.85	1.20	8.0
Strata	4	14.51	0.54	3.6
Strata – urban / rural	6	13.65	1.40	9.3
Prefectures	12	13.71	1.34	8.9
Districts	36	13.17	1.88	12.5
Communes/ Municipalities	374	12.50	2.55	16.9
Rural	**1**	**13.74**	**0**	**0**
Communes	309	11.40	2.34	17.0
Urban	**1**	**14.02**	**0**	**0**
Municipalities	65	13.27	0.75	5.3

Source: Neri, Ballini and Betti (2005).

[6] Questo paragrafo costituisce un approfondimento. Si tratta di una ricerca commissionata dalla *World Bank* i cui risultati sono stati pubblicati in Neri, Ballini e Betti (2005).

…"In this report we have estimated various measures of welfare for small administrative units in Albania, combining the 2001 Population and Housing Census with the 2002 Living Standards Measurement Study survey data.
Our estimates of mean expenditure, poverty, and inequality at the Stratum level, the level at which the household survey is representative, are comparable to those calculated using the sample survey; in fact, welfare ranking of the four strata are completely consistent. Welfare rankings of any administrative unit using various measures of poverty are consistent as well; poverty is more pronounced and less heterogeneous in rural areas than in urban areas. We have produced poverty rates very precise at Stratum, Prefecture and District level, and rates that are precise enough to be of value to researchers and policy-makers even at Communes level.
Some interesting findings of this report could be potentially very useful for policy-makers; they are reported as follows: in Albania there is considerable heterogeneity of poverty rates across administrative units; this heterogeneity is very pronounced between rural and urban areas.
When we compare different level of disaggregation we observe a large spatial heterogeneity among Prefectures and among Municipalities within the District to which they belong. This spatial heterogeneity is much less pronounced among Districts within the same Prefecture.
We therefore conclude this report recommending policy-makers to focus their attention on "Large scale" project at Prefecture level and more specific and "well oriented" project at Commune/ Municipality level rather than at District level." …

Riferimenti Bibliografici

Aitchinson, J. e Brown, J.A.C. (1969), *The Lognormal Distribution with Special Reference to Its Uses in Economics*, Cambridge University Press, Cambridge.

Ammon, O. (1895), *Die Gesellschaftsordnung Und Ihre Naturlichen Grundlagen*.

Atkinson, A.B. (1970), On the Measurement of Inequality. *Journal of Economic Theory*, 2, pp. 244-263

Baldini, M. e Toso, S. (2004), *Diseguaglianza, povertà e politiche pubbliche*, Il Mulino, Bologna.

Bartels, C.P.A. e Van Metile, H. (1975), Alternative Probability Density Functions of Income. Vrije University Amsterdam: Research memorandum 29.

Betti, G. e Lemmi, A. (2007), Guest Editors' Introduction, *Journal of Economic Inequality*, 5(3), pp. 259-262.

Bandourian, R., McDonald, J.B. e Turley R.S. (2003), A comparison of parametric models of income distribution across countries and over time. *Estadistica*, 55, pp. 135–152.

Cox, D.R. e Hinckley, D.V. (1974), *Theoretical Statistics*. Chapman and Hall, London.

Dagum, C. (1977), A new model for personal income distribution: specification and estimation, *Review of Applied Economics*, 30, pp. 413–437.

Dastrup, S.R., Hartshorn, R. e McDonald, J.B. (2007), The impact of taxes and transfer payments on the distribution of income: A parametric comparison. *Journal of Economic Inequality*, 5, pp. 353–369.

Gibrat, R. (1931), *Les Inegalites Economiques*. Sirey, Paris.

Gini, C. (1914), Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 73, pp. 1203-1240.

Lorenz, M.O. (1905), Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9, pp. 209–219.

Marbach, G. (a cura di) (1990), *Statistica Economica*, UTET Libreria, Torino.

Neri, L., Ballini, F. e Betti G. (2005), Poverty and inequality mapping in transition countries, *Statistics in Transition*, 7(1), pp. 135-157.

Pareto, V. (1895), *Teoria matematica del commercio internazionale*, in Pareto, V. (1952), pp. 306-328.

Salem, A.B. e Mount, T.D. (1974), A convenient descriptive model of income distribution: The gamma density, *Econometrica*, 42 (6), pp. 1115–1127.

Sen, A.K. (1976), Poverty: an ordinal approach to measurement. *Econometrica*, 44(2), pp. 219-231.

Singh, S.K. e Maddala, G.S. (1976), A function for the size distribution of incomes. *Econometrica*, 44, pp. 963–970.

Xu, K. (2004), How has the Literature on Gini's Index Evolved in the Past 80 Years? Halifax, Dalhousie University, Economics Department, Working Paper.