

**QUADERNI DEL DIPARTIMENTO DI SCIENZE  
ECONOMICHE E SOCIALI**

**LA COSTRUZIONE DI UN INDICATORE DI  
SOSTENIBILITÀ DEI COMUNI ITALIANI  
CAPOLUOGO DI PROVINCIA:  
METODOLOGIA E APPLICAZIONI**

*Antonio Dallara*

*Serie Rossa: Economia – Quaderno N. 64 luglio 2010*



**UNIVERSITÀ CATTOLICA DEL SACRO CUORE  
PIACENZA**

# **La costruzione di un indicatore di sostenibilità dei comuni italiani capoluogo di provincia: metodologia e applicazioni**

**Antonio Dallara<sup>1</sup>**

---

<sup>1</sup> LEL - Laboratorio di Economia Locale, Facoltà di Economia, Università Cattolica del Sacro Cuore sede di Piacenza, Via Emilia Parmense n. 84, 29100 Piacenza, Italia  
Tel. 0523 599.311 Fax 0523 599.437 e-mail: antonio.dallara@unicatt.it

## Indice

<b>Sommario</b>	<b>3</b>
<b>Abstract</b>	<b>3</b>
<b>1. Lo stato dell'arte degli indicatori territoriali di sostenibilità dello sviluppo: applicazioni, punti di forza e punti di debolezza delle varie metodologie</b>	<b>4</b>
1.1 Gli indicatori di sostenibilità	4
1.2 I metodi utilizzati	6
1.3 Le tendenze dei nuovi indicatori di sostenibilità dello sviluppo	7
<b>2. Dal concetto di “sviluppo sostenibile” ad un indicatore di sintesi: la descrizione della metodologia e la sua applicazione ai comuni capoluogo</b>	<b>10</b>
2.1 Canoni e postulati di base	10
2.2 Il concetto di “sviluppo urbano sostenibile” e la sua rappresentazione quantitativa	10
2.3 L'analisi in componenti principali	12
2.4 La clusterizzazione dei 103 comuni capoluogo	17
2.5 L'analisi discriminante	19
2.6 L'indice di sviluppo urbano sostenibile	21
<b>3. I dati e le variabili</b>	<b>23</b>
3.1 Le variabili elementari per dimensione della sostenibilità	23
3.2 Il posizionamento delle singole città nelle dimensioni della sostenibilità	23
<b>4. L'indice di sviluppo urbano sostenibile nei 103 comuni capoluogo di provincia italiani</b>	<b>27</b>
4.1 La graduatoria dello sviluppo sostenibile nei comuni capoluogo di provincia	27
4.2 Le dimensioni della sostenibilità urbana: economia, società, ambiente	30
4.3 Le relazioni tra l'indice di sviluppo urbano e le tre dimensioni della sostenibilità	31
4.4 I confronti con le altre classifiche nazionali (Il Sole24 Ore, Italia Oggi, Legambiente)	33
<b>5. Conclusioni</b>	<b>34</b>
<b>6. Bibliografia</b>	<b>37</b>
Appendice 1	39
Appendice 2: introduzione teorica alle tecniche utilizzate nella costruzione dell'indicatore composito	43

## **Sommario**

Dopo la sintesi dello stato dell'arte degli indicatori di sostenibilità territoriale costruiti a livello internazionale e nazionale, in questo lavoro si richiamano alcune tecniche per l'analisi e la selezione di variabili territoriali e si propone un metodo per la costruzione di un indice di sviluppo urbano sostenibile.

L'indice composito proposto si ottiene combinando in sequenza tre tecniche di statistica multivariata, l'analisi in componenti principali, l'analisi cluster, l'analisi discriminante. Se ne formula la definizione formale, e si richiamano i risultati della sua applicazione ai 103 comuni capoluogo di provincia italiani, per gli anni 1999, 2000 e 2001.

## **Abstract**

This paper begins with the description of the state of the art of the sustainability indicators constructed at international and national level. Then are recalled some techniques for the analysis and selection of territorial quantitative variables and it is proposed a method for the definition of a composite index of urban sustainable development. This index is obtained arranging in sequence three techniques of multivariate statistical analysis: the principal component analysis, the cluster analysis, the discriminant analysis. It is defined in a formal way, and it is described by its Italian application to the 103 chief towns of Italian provinces, for the years 1999, 2000 and 2001.

**Keywords:** indicatori di sintesi; sviluppo territoriale sostenibile

**JEL Classification:** C4; P5; O; R1

## **1. Lo stato dell'arte degli indicatori territoriali di sostenibilità dello sviluppo: applicazioni, punti di forza e punti di debolezza delle varie metodologie**

Il paper comincia con una sintesi della letteratura empirica in cui si colloca l'indicatore composito (OECD, 2008), che nel seguito viene descritto nelle sue implicazioni metodologiche e nelle sue evidenze empiriche. In questo capitolo si descrivono alcuni indicatori di rilevanza internazionale e altri costruiti in Italia per l'analisi del nostro territorio a livelli diversi (in genere l'intera nazione e le regioni). Si analizzano anche i metodi con cui vengono elaborati e i principali temi in base ai quali si confrontano i territori (competitività, sostenibilità, qualità della vita).

### **1.1 Gli indicatori di sostenibilità**

Negli ultimi decenni l'analisi economica e sociale si è dotata in misura crescente di tecniche e metodologie statistiche di rilevazione per definire e misurare le performance e la sostenibilità dello sviluppo dei sistemi economici sia a livello nazionale che regionale. A partire dagli anni Sessanta, prima negli Usa e poi anche in Europa si sviluppa il cosiddetto "movimento degli indicatori sociali". In quegli anni agli indicatori sociali e economici si attribuivano, soprattutto da parte della Pubblica amministrazione e dei ricercatori, numerose funzioni: da un lato finalità di informazione relative ai fenomeni sociali e economici, per rendere possibile la percezione, l'individuazione e l'analisi dei mutamenti e dei cambiamenti socioeconomici; in secondo luogo obiettivi di costruzione di modelli esplicativi e di predizione (Sheldon-Moore, 1968). Dopo una fase di riflessione critica di carattere euristico e teorico, nel corso degli anni Ottanta, che ha contribuito comunque al consolidamento e alla maturazione delle tecniche (Andrews, 1990), gli indicatori socioeconomici nel corso degli ultimi anni hanno vissuto una fase di rivitalizzazione sia nella ricerca sia nell'utilizzo a fini politici e amministrativi (Zajczyk, 1997).

Agli indicatori socioeconomici nel corso degli anni si sono affiancati quelli di contesto ambientale-ecologico (livello delle pressioni, aree verdi disponibili, raccolta dei rifiuti, qualità di aria, acqua e suolo) e nel contempo si sono cominciati a ricercare metodi di sintesi per la definizione di indicatori aggregati di benessere sociale, qualità della vita, sostenibilità. Nell'ultimo decennio si sono diffusi anche indicatori più complessivi di competitività territoriale e di attrattività, parallelamente all'esplosione degli scambi commerciali e finanziari tra paesi e territori, che rappresenta il fenomeno principale del cosiddetto processo di globalizzazione (Sassen, 2002; Ohmae, 1998; Bauman, 2001; Stiglitz, 2002).

Nel 2008 l'OECD pubblica un manuale per la costruzione di indicatori compositi. In 10 punti riassume i principali passaggi che ritiene sia opportuno seguire per la costruzione di indicatori di analisi comparata delle prestazioni dei sistemi territoriali (OECD, 2008). Nel settembre 2009 la Commissione Stiglitz-Sen-Fitoussi pubblica un rapporto sulla "misurazione delle prestazioni economiche e del progresso sociale". In tre capitoli delinea i limiti del Pil come misura di performance economica, illustra la necessità di passare dalla misura della produzione alla misura del benessere umano e sociale, elenca raccomandazioni per individuare le dimensioni che caratterizzano la qualità della vita e lo sviluppo sostenibile a tutela della qualità ambientale (Stiglitz-Sen-Fitoussi Commission, 2009).

Gli indicatori sono infatti strumenti di analisi dei sistemi economici, di *benchmarking* tra ambiti territoriali che competono e cooperano, sono utili per individuare *swot* territoriali (punti di forza, punti di debolezza, opportunità e minacce) e quindi per definire strategie d'azione, obiettivi e più in generale per architettare e negoziare politiche pubbliche. Allo stesso modo, sono importanti per monitorare le politiche e i programmi di intervento, così come per valutare i risultati, gli effetti e gli impatti prodotti dalla loro implementazione.

Numerosi sono gli indicatori aggregati che vengono calcolati negli Stati Uniti, nei paesi europei, e in parte anche in Italia: in genere si tratta di indicatori stimati a livello nazionale e sovra-nazionale. La loro evoluzione rivela interessanti mutamenti teorici e possibili conseguenze in termini di policy e di

ranking comparato. Dopo la critica all'assolutizzazione del prodotto interno lordo (Stiglitz, 2002) quale indice principe per la definizione e la misurazione della competitività, perché non ritenuto in grado di misurare il benessere sociale e soprattutto non comprensivo di alcune dimensioni sempre più importanti nei sistemi socioeconomici (terzo settore e volontariato, costi ambientali, lavoro domestico, ...), uno dei primi passi compiuti dalla ricerca è stato quello di inserire nella contabilità nazionale valori e dimensioni del sistema ambientale (pil verde, impronta ecologica, indici di sostenibilità ecologica) (Tab. 1).

**Tab. 1 Indicatori di sostenibilità**

<b>Indicatore</b>	<b>Istituzione/ autore (anno di definizione)</b>	<b>Ambito territoriale di riferimento</b>	<b>Metodologia</b>
<b>PIL verde *</b>	ONU	nazionale	Manuale SEEA, 2003
	Eurostat	nazionale	Series 94; Ieeaf 2002
	Istat	nazionale	Namea 90
<b>HDI</b> (Human Development Index)	UNDP (1992)	nazionale	Composizione dei valori medi di tre indicatori: speranza di vita, livello di istruzione, pil pro-capite
<b>ISEW **</b> (Indicator of Sustainability Economic Welfare)	Daly e Cobb (1989)	nazionale	Somma della spesa privata per consumi e per investimenti netti dei benefici non direttamente monetizzabili, tratti da beni di consumo durevoli e dalla rete stradale. A questa somma si sottraggono spese private per sanità e istruzione, costo dell'inquinamento e deprezzamento del capitale naturale
<b>GPI ***</b> (Genuine Progress Indicator)	Redefining Progress (1994)	nazionale	Ai consumi privati si sottraggono i costi non considerati nel calcolo del pil, tra cui: valore netto dei prestiti, costo dei beni durevoli, costi sociali (criminalità, incidenti, pendolarismo, sottoccupazione), costi ambientali (inquinamento, perdita terre umide, perdita terreni agricoli, perdita foreste). Poi si aggiungono i benefici che non si considerano nel calcolo del pil, tra cui lavoro domestico, lavoro volontario, servizi offerti dalle strade.
<b>Spazio Ambientale</b>	Opschoor – Wuppertal Institute (1992)	nazionale	Stima la minima quantità del bene considerato (energia, materie prime non rinnovabili, legname, acqua) e, poi, il punto di equilibrio tra i tassi di rinnovo del bene e di prelievo/consumo da parte dell'uomo.
<b>TMR</b> (Total Material Requirement)	Wuppertal Institute	nazionale	Calcola l'ammontare totale in peso delle materie prime estratte e quelle effettivamente utilizzate dal sistema produttivo di un dato territorio. È in grado di distinguere i materiali effettivamente immessi nel processo di trasformazione e i materiali mobilitati per le estrazioni (zaino ecologico).
<b>Living planet index</b>	WWF	sovranazionale	Misura la qualità della vita delle popolazioni delle foreste, delle popolazioni delle acque dolci, e delle specie marine.
<b>Impronta ecologica</b>		nazionale, regionale, provinciale, comunale	Stima con metodi complessi e articolati quanti ettari di terra sono necessari per garantire la vita ad un essere umano.
<b>Indice ISSI</b>	Istituto per lo Sviluppo Sostenibile in Italia (2002)	nazionale	È dato dalla composizione di tre indicatori riferiti rispettivamente a: economia-società; ambiente; uso delle risorse.

\* Pil verde: integra la sfera economica con la sfera ambientale

\*\* Isew: integra le tre sfere della sostenibilità, l'economia, la società e l'ambiente

\*\*\* Gpi: integra la sfera economica con la sfera della società

Gli sviluppi che poi negli anni 90 si sono susseguiti sono stati un continuo tentativo di affinamento di questi metodi, sia in termini di contenuti misurati che di tecniche di aggregazione e ponderazione. Nella tabella seguente si propone una sintesi di alcuni degli indicatori più noti e diffusi a livello internazionale e nazionale, riportandone l'istituzione di riferimento, l'ambito territoriale di analisi e una breve descrizione dell'oggetto di rilevazione (Dallara, 2004, Ambrosanio-Bordignon-Dallara, 2004, Dallara-Rizzi, 2004, Dallara, 2006, Dallara, 2008, Ciciotti-Dallara-Rizzi., 2008).

In alcuni casi misure della qualità della vita e indicatori degli aspetti sociali e ambientali vengono proposti all'interno di rapporti periodici realizzati da istituzioni nazionali e sovranazionali. Spesso in questi rapporti non si costruiscono indici di sintesi, ma si propongono semplicemente serie di variabili elementari e sulla base di questi dati si pongono a confronto territori tra loro eterogenei. È il caso ad esempio del rapporto Ambiente Italia di Legambiente e dello Human Development Report dell'UNDP (Tab. 2).

**Tab. 2 Rapporti sull'ambiente, sullo sviluppo sociale e economico**

Titolo del rapporto	Istituzione	Ambito territoriale di riferimento
Ambiente Italia	Legambiente	nazione
Ecosistema Urbano	Legambiente	comuni capoluogo
La qualità della vita in Italia	Il Sole24 Ore	provincia
Rapporto sulla qualità della vita in Italia	Italia Oggi	provincia
Rapporto sullo Stato dell'Ambiente	Ministero Ambiente	nazione
Human Development Report	UNDP	nazione
World Development Indicator	World Bank	nazione

In Italia sono molto diffusi e dibattuti anche in ambito accademico i rapporti sulla qualità della vita redatti annualmente dai due maggiori quotidiani economici a livello nazionale (Il Sole 24Ore e Italia Oggi). Nell'uno e nell'altro caso si tratta di indagini che riaggregano dati provinciali e comunali, e portano ad indicatori settoriali e ad una misura di sintesi per singolo ambito territoriale. I dati utilizzati coprono tutti i principali aspetti economici, sociali e ambientali. Allo stesso tempo altri organismi hanno costruito altri indici e ranking sullo sviluppo in particolare delle province e delle città italiane (Censis, Unioncamere) che classificano i sistemi economici nazionali in base a livelli e dinamiche del reddito. Da ultimo sono stati elaborati alcuni indicatori di crisi industriali e contesto socioeconomico (Ministero Attività Produttive, IPI, con la collaborazione del LEL-Università Cattolica) e di attrattività territoriale (Fondazione Accenture). Diverso il rapporto realizzato da Legambiente, che costruisce misure della qualità ambientale per ogni comune capoluogo, senza indagare direttamente aspetti sociali e economici o quello di Crediop-Università Cattolica che cerca di formalizzare il concetto di sostenibilità in indici sintetici.

## 1.2 I metodi utilizzati

Gli indici divergono non solo per contenuto, natura e numero delle variabili analizzate e aggregate, ma soprattutto per approccio e metodi di operazionalizzazione: dati quantitativi e indicatori statistici; interviste-Delphi; indici misti. Risulta quindi problematico valutare l'eshaustività e la completezza di molti indicatori, soprattutto per comprendere i fattori esplicativi degli indici di posizionamento dei paesi o delle regioni. A titolo di esempio una lettura comparata tra il World Competitiveness dell'IMD e il Business Competitiveness del World Economic Forum, entrambi fondati su tecniche miste (statistiche e interviste), non consente di verificare se fattori endogeni possano sostenere la crescita futura o se la competitività sia legata ad altri fattori di natura internazionale o esogena. Ancora le nozioni di competitività e attrattività derivano e intercettano anche altre tipologie di variabili e quindi di indicatori, ad esempio trasparenza e corruzione, come nel caso degli indicatori Economic Freedom

del Fraser Institute e degli indici di trasparenza Corruption Perception Index della Transparency International.

Diventa pertanto difficile valutare le relazioni e le reciproche interrelazioni tra le variabili misurate, anche perché, nel caso, miglioramenti di efficienza e trasparenza pubblica non producono necessariamente risultati nel breve periodo. La natura ambigua delle rilevazioni da sondaggio o intervista produce ancora aggregati sintetici di difficile interpretazione, soprattutto se non comparati con gli stessi indicatori desunti ex-post da misure effettive.

In sintesi è possibile definire il posizionamento dei diversi indicatori internazionali e nazionali studiati in base a due prospettive analitiche: le metodologie di analisi/misurazione (statistiche, interviste, miste) e il contenuto euristico delle variabili (competitività, sostenibilità, qualità della vita).

**Tab. 3 Metodologie e temi trattati nei principali rapporti internazionali**

<b>Metodologia</b>	Statistiche	Interviste	Miste
<b>Contenuto variabili</b>			
Competitività	- World Development Indicator; - Censis Municipium - Sistema Italia-Unioncamere - Ministero Attività Produttive-IPI	- Corruption Perception Index-Trasparenza International - FDI Confidence Index A.T.Kerney	- Business Competitiveness-World Economic Forum - World Competitiveness-IMD - FDI Potential Index-Unctad - Economic Freedom-Fraser Institute
Sostenibilità	- Total Material Requirement - Pil Verde - Indice di Sviluppo Umano-UNDP - Living Placet Index - Impronta ecologica - Genuine Progress Indicator - Ecosistema Urbano - Indice ISSI-Istituto Italiano Sviluppo Sostenibile		
Qualità della vita	- Human Development Report-UNDP - Environment Regulatory Regime-World Economic Forum - Italia Oggi	- Quality of Life-Mercer	- Il Sole24 Ore

### 1.3 Le tendenze dei nuovi indicatori di sostenibilità dello sviluppo

Si può anche cercare una lettura integrata dell'evoluzione dei diversi indicatori di competitività e sostenibilità richiamati nella tavola precedente in termini di vantaggi e criticità dell'attuale processo di misurazione e classificazione/ranking dei paesi e territori.

Emerge come tensione positiva lo sforzo crescente di integrare le tradizionali misurazioni economiche basate su variabili di contabilità nazionale (pil, reddito, export, investimenti, consumi) con nuovi indicatori che misurino i costi sociali e ambientali dello sviluppo, come gli aspetti della qualità della vita (benessere, contenuti sociali, culturali e relazionali).



**Tab. 4 Punti di forza e di debolezza dei principali indicatori di sostenibilità**

<b>Punti di forza</b>	<b>Punti di debolezza</b>
<ul style="list-style-type: none"> <li>- inserimento variabili ambientali</li> <li>- inserimento variabili sociali</li> <li>- comparabilità internazionale (ranking)</li> <li>- oggettività delle rilevazioni statistiche</li> <li>- possibili benchmarking</li> <li>- enfasi su struttura e prestazioni</li> </ul>	<ul style="list-style-type: none"> <li>- comparazioni soprattutto nazionali, raramente regionali o locali</li> <li>- soggettività delle analisi a sondaggio</li> <li>- carente lettura delle strategie e delle politiche</li> <li>- debolezza delle tecniche statistiche (medie ponderate)</li> <li>- scarsa verifica della robustezza dei dati</li> <li>- confusione semantica (tra competitività, sostenibilità e attrattività)</li> </ul>

Un focus particolare può essere fatto per il caso italiano, dove emergono i primi tentativi di analisi regionale e provinciale, ma solo in alcuni casi su scala sub-provinciale (si ricorda tra gli altri Ministero Attività Produttive-IPI con il contributo del LEL-Università Cattolica, Ecosistema Urbano, Censis Municipium, Crediop-Università Cattolica), soprattutto per i comuni capoluogo di provincia.

In particolare emergono gli stessi limiti euristici che si osservano nei ranking internazionali, ovvero l'enfasi posta su indicatori che intercettano le variabili di prestazione territoriale (dinamica di pil, reddito, occupazione, investimenti, consumi, redditività aziendale) e di struttura socioeconomica (aspetti demografici, dati sul mercato del lavoro, struttura settoriale della produzione) o della dotazione di economie esterne (infrastrutture economiche e sociali, qualità della vita, capitale umano).

Al contrario si registrano scarsi tentativi di misurare e monitorare gli aspetti volontaristici della competizione territoriale (LEL 1999, Ciciotti 2002), sia dal lato delle strategie pubbliche (indicatori di spesa pubblica, indici di managerialità, come nel caso di Valdani e Vicari sulla managerialità pubblica, input e output delle politiche locali) sia dal lato delle strategie private (livelli di internazionalizzazione e strategie di investimento aziendale) e soprattutto delle strategie di comunità (peso e ruolo del terzo settore, partecipazione civile, capitale sociale).

**Tab. 5 I contenuti dei principali indicatori di sostenibilità e di competitività**

<b>Contenuto euristico</b>	<b>Struttura e economie esterne</b>	<b>Prestazioni</b>	<b>Strategie pubbliche e private</b>
<b>Contenuto tematico</b>			
Competitività	<ul style="list-style-type: none"> <li>- Censis Municipium</li> <li>- Sistema Italia-Unioncamere</li> <li>- Ministero Attività Produttive-IPI</li> </ul>	<ul style="list-style-type: none"> <li>- Censis Municipium</li> <li>- Sistema Italia-Unioncamere</li> <li>- Ministero Attività Produttive-IPI</li> </ul>	<ul style="list-style-type: none"> <li>- Valdani e Vicari-II Sole24 Ore</li> </ul>
Sostenibilità	<ul style="list-style-type: none"> <li>- Ecosistema Urbano</li> <li>- Crediop-Università Cattolica</li> <li>- Indice ISSI-Istituto Italiano Sviluppo Sostenibile</li> </ul>	<ul style="list-style-type: none"> <li>- Ecosistema Urbano</li> <li>- Crediop-Università Cattolica</li> <li>- Indice ISSI-Istituto Italiano Sviluppo Sostenibile</li> </ul>	<ul style="list-style-type: none"> <li>- Ecosistema Urbano</li> <li>- Crediop-Università Cattolica</li> </ul>
Qualità della vita	<ul style="list-style-type: none"> <li>- Il Sole24 Ore</li> <li>- Italia Oggi</li> </ul>	<ul style="list-style-type: none"> <li>- Il Sole24 Ore</li> <li>- Italia Oggi</li> </ul>	

Gli stessi limiti metodologici riscontrati a livello internazionale per quanto riguarda la debolezza statistica delle analisi, quali l'eccessivo uso di medie ponderate senza esplicitazioni e giustificazioni dei pesi adottati, l'aggregazione di variabili senza adeguato studio preliminare sulla loro robustezza statistica e sulla loro multicollinearità, la superficialità dell'adozione di tecniche di clusterizzazione e

analisi in componenti principali (analisi fattoriale) senza una opportuna verifica dei test statistici suggeriti dalla letteratura scientifica (Dallara-Rizzi, 2004), la commistione ingiustificata di tecniche miste di rilevazione (soggettiva-oggettiva, statistiche-interviste, ...) ma soprattutto l'assenza di chiari riferimenti a modelli interpretativi che coniughino una solida teoria di sviluppo locale endogeno da un lato e un chiaro riferimento alle finalità di politiche territoriali dall'altro.

Ne nasce l'esigenza di nuove sperimentazioni di rating territoriale che contemplino i seguenti criteri ad oggi non del tutto soddisfatti a livello di ricerca applicata nazionale:

- l'ambito territoriale sub-provinciale
- l'utilizzo di tecniche statistiche di analisi in componenti principali (o in alternativa l'analisi fattoriale) e clusterizzazione e elaborazioni econometriche più solide
- verifica della multicollinearità e ridondanza statistica delle variabili utilizzate
- l'integrazione di variabili relative a prestazioni e struttura con indicatori inerenti le strategie territoriali (pubbliche, private, comunitarie)
- la ripetibilità della rilevazione e dei ranking
- adozione di un chiaro modello teorico di riferimento (approccio di sviluppo locale).

## **2. Dal concetto di “sviluppo sostenibile” ad un indicatore di sintesi: la descrizione della metodologia e la sua applicazione ai comuni capoluogo**

In questo capitolo si presenta una metodologia nuova per la costruzione di un indicatore di sintesi su base territoriale, ottenuto dalla combinazione sequenziale di tre tecniche di statistica multivariata (l'analisi in componenti principali, l'analisi cluster, l'analisi discriminante). Si presentano anche i risultati dell'applicazione della metodologia ai comuni capoluogo italiani, in particolare gli scores prodotti dalle singole tecniche e i test di significatività statistica. Mentre si rimanda al quarto capitolo il commento ai risultati da un punto di vista economico-territoriale.

### **2.1 Canoni e postulati di base**

Per costruire un indicatore di sintesi che misuri in termini quantitativi il livello di sviluppo sostenibile dei sistemi territoriali locali da un punto di vista economico-sociale-ambientale occorre fare un breve cenno, di carattere solo introduttivo, sia ai canoni fondamentali sociologici per la valutazione di fenomeni complessi proposti nel 1969 da Lazarsfeld sia ai principi di Bellagio del 1996.

I canoni di Lazarsfeld sono la descrizione delle fasi da seguire per costruire variabili di sintesi in grado di misurare l'attività umana:

- descrizione e specificazione di un “concetto”. Nel caso in esame il concetto è rappresentato dall'espressione “sviluppo economico-sociale-ambientale sostenibile” e può essere articolato in due categorie ordinali “basso sviluppo sostenibile” e “alto sviluppo sostenibile”;
- scelta delle variabili (quantitative) che descrivono il “concetto”;
- sintesi delle variabili in un unico indice che misuri in termini ordinali o cardinali il “concetto” indagato.

I “principi di Bellagio”<sup>2</sup> sono molto più recenti. Vogliono essere di guida in tutti i processi di valutazione dello sviluppo sostenibile, soprattutto nella scelta e nella costruzione di indicatori, nella loro interpretazione e nella comunicazione dei risultati delle analisi di sostenibilità.

Si nota che in genere i sistemi di indicatori per l'analisi della sostenibilità dello sviluppo soddisfano almeno in parte questi principi. I 10 principi di Bellagio sono di seguito elencati. La valutazione del progresso verso lo sviluppo sostenibile dovrebbe: 1. avere una chiara *vision* dello sviluppo sostenibile e chiari obiettivi da realizzare; 2. avere una visione olistica, ossia d'insieme e contemporaneamente delle singole parti componenti; 3. tener conto dei rapporti intergenerazionali; 4. avere una portata adeguata; 5. avere un orientamento e funzioni pratiche; 6. adottare metodi di valutazione “aperti” e accessibili a tutti; 7. avere una comunicazione efficace; 8. garantire una partecipazione ampia e allargata; 9. consentire analisi ripetute nel tempo; 10. capacità di guidare i policy maker nel prendere decisioni.

### **2.2 Il concetto di “sviluppo urbano sostenibile” e la sua rappresentazione quantitativa**

La costruzione di un indice di sintesi dello sviluppo sostenibile richiede la definizione preliminare di alcuni elementi teorici essenziali (Merlini-Vitali, 1999), tra cui il termine di “concetto”, il “verso logico” del concetto, l'associazione tra concetto e variabili elementari quantitative. Quindi per definire l'indice di sintesi sono necessari:

1. la definizione del concetto  $C$  “sviluppo urbano sostenibile in termini economici-sociali-ambientali” (come già per Lazarsfeld, si veda il paragrafo 2.1) e la definizione delle categorie ordinali in cui il concetto si declina,  $(c_1, \dots, c_h)$  tali per cui valga la relazione  $c_i \pi c_j$  (la categoria  $c_i$  è non preferita alla categoria  $c_j$ ). Ad esempio le categorie potrebbero essere “alto sviluppo sostenibile”, “medio”, “basso”).

---

<sup>2</sup> I 10 Principi di Bellagio vennero messi a punto nel novembre 1996 in occasione di un incontro internazionale organizzato a Bellagio (Italia) dall'International Institute of Sustainable Development (IISD), a cui parteciparono esperti provenienti dai cinque continenti. L'esperienza si realizzò quattro anni dopo il Summit della Terra tenuto a Rio de Janeiro nel 1992, voleva essere un'occasione a sostegno della strategia globale di Agenda 21.

Il concetto attribuibile ad una variabile di sintesi può essere positivo o negativo. In questo paper si sceglie di assegnare all'indicatore di sintesi il concetto positivo di "sviluppo sostenibile". Nel caso in cui si fosse scelto di attribuirle un concetto negativo si sarebbe fatto riferimento ad esempio al "degrado socio-economico-ambientale" o alla "decrescita dei sistemi locali";

2. la costruzione di una matrice  $X_{(n,p)}$  costituita da  $p$  variabili quantitative continue rilevate per  $n$  unità statistiche;
3. la definizione della associazione (positiva o negativa) tra ciascuna  $p$ -esima variabile della matrice  $X_{(n,p)}$  e il "concetto" di "sviluppo sostenibile". L'associazione tra "variabile" e "concetto" può essere positiva o negativa;
4. l'ordinamento delle  $n$  unità statistiche in base alle categorie ordinali  $(c_1, \dots, c_h)$  in cui si articola il concetto  $C$ , ordinamento possibile in base all'associazione di cui al punto 3 precedente. Ad esempio la criminalità nell'ambito sociale, l'inquinamento per l'ambiente, la disoccupazione per l'economia sono le variabili associate negativamente con il concetto di "buon sviluppo sostenibile". Nel senso che all'aumentare di almeno una di queste tre variabili il livello di sviluppo sostenibile di un sistema si riduce. Assegnare questa associazione tra il significato della variabile elementare e il concetto di sviluppo sostenibile significa attribuire un "verso logico" alla variabile di sintesi e al "concetto" ad essa attribuito.

#### *La funzione di sintesi*

Poste queste definizioni è possibile introdurre la nozione di funzione di sintesi, che rappresenta l'algoritmo statistico-matematico con il quale trasformare singole variabili elementari in un indicatore di sintesi. Quindi data la matrice  $X_{(n,p)}$ , con  $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$  il problema di ordinare le  $n$  osservazioni può porsi stabilendo che sia valida la relazione ordinale  $\{x_{(r_i)} \pi x_{(r_j)}\}, r_i \pi r_j$  in modo tale che

$$\{x_{(r_i)} \pi x_{(r_j)}\} \Rightarrow \{c_i \pi c_j\}$$

cioè " $x_{(i)}$  non preferito a  $x_{(j)}$ " implica che " $c_i$  è non preferita a  $c_j$ ".

Se le categorie ordinali del concetto  $C$  sono  $h$  e si stabilisce che ogni osservazione caratterizzi una e una sola classe del concetto  $C$ , allora si perviene al seguente ordinamento delle  $n$  unità statistiche

$$\{x_{(1)} \pi x_{(2)} \pi \dots \pi x_{(n)}\}$$

e è anche possibile costruire un vettore  $(X, R_X) = \{(x_1, r_1), \dots, (x_n, r_n)\}$  che alla misura  $x_i$  della  $i$ -esima unità statistica associa il rango

$$r_i = \sum_{j=1}^n k(x_i - x_j)$$

dove  $k(x_i - x_j) = 1$  se  $(x_i - x_j) \geq 0$ , e  $k(x_i - x_j) = 0$  se  $(x_i - x_j) < 0$ .

Una volta stabilita l'associazione tra le categorie ordinali del concetto  $C$  e le componenti della matrice  $X_{(n,p)}$ ,  $X = (X_1, \dots, X_p)$ ,  $X \in A$ ,  $A \subseteq \mathfrak{R}^p$ , per costruire l'ordinamento  $\{x_{(1)} \pi x_{(2)} \pi \dots \pi x_{(n)}\}$  è necessario definire una funzione di sintesi delle  $p$  variabili quantitative che assuma valori in un sottoinsieme  $B$  dei reali ( $B \subseteq \mathfrak{R}$ )

$$\varphi(X) : A \rightarrow B, B \subseteq \mathfrak{R}, A \subseteq \mathfrak{R}^p$$

funzione monotona crescente nelle componenti di  $X_{(n,p)}$  associate positivamente con il concetto  $C$  e monotona decrescente in quelle componenti di  $X_{(n,p)}$  associate negativamente con  $C$ .

In questo modo la funzione di sintesi  $\varphi(X)$  determina un ordinamento nello spazio unidimensionale dei reali o in un suo sottoinsieme  $B$  ( $B \subseteq \mathfrak{R}$ ), cosicché l'ordinamento  $\{x_{(1)} \pi x_{(2)} \pi \dots \pi x_{(n)}\}$  può essere scritto

$$(\varphi(X), R_X) = \{(\varphi(X_1), r_1), \dots, (\varphi(X_n), r_n)\}$$

dove  $r$  è il rango di  $\varphi(X)$  ottenuto dall'ordinamento numerico  $\varphi_{(1)} \leq \dots \leq \varphi_{(n)}$ .

Questo in termini formali è il risultato a cui si giunge al termine del presente lavoro, ordinando i comuni capoluogo di provincia italiani in base al concetto di “sviluppo urbano sostenibile”.

La funzione di sintesi  $\varphi(X)$  delle  $p$  variabili quantitative rilevate sulle  $n$  unità statistiche si può definire in vari modi e può assumere varie forme algebriche.

Si ricordano al riguardo tutti i lavori citati nel capitolo precedente (tra cui l'indice di sviluppo umano HDI dell'UNDP), e in cui si ricorda che i metodi utilizzati per definire funzioni di sintesi sono i più vari, dalla media aritmetica dell'HDI fino alla rielaborazione di valori della contabilità nazionale per il Pil verde.

Il metodo che si propone in queste pagine, e con il quale sono stati prodotti i risultati proposti nel caso studio dei comuni italiani si articola nei seguenti passaggi:

1. selezione delle variabili e loro sintesi nelle tre macro-variabili dello sviluppo sostenibile (mediante l'analisi in componenti principali)
2. suddivisione delle  $n$  unità statistiche in 2 cluster, in base a due categorie ordinali del concetto di “sviluppo sostenibile”, basso e alto (mediante l'Analisi Cluster)
3. costruzione della funzione di sintesi (mediante l'Analisi Discriminante lineare canonica).

### 2.3 L'analisi in componenti principali

Il primo passaggio per costruire l'indicatore di sintesi (dato dalla funzione di sintesi  $\varphi(X)$  di cui al paragrafo precedente) consiste nella selezione delle variabili elementari costitutive e nella loro sintesi in indicatori intermedi facendo ricorso all'analisi in componenti principali (ACP).

*La costruzione delle tre componenti principali della sostenibilità nei 103 comuni capoluogo*

Sia data una matrice  $X_{(103,p)}$  dei 103 comuni capoluogo italiani e  $p$  variabili quantitative in essi rilevate in un dato periodo temporale. All'interno della matrice  $X_{(103,p)}$  si stabilisce a priori che alcune delle  $p$  variabili sono espressione della dimensione economica della sostenibilità, altre diverse da quelle economiche sono espressione della dimensione sociale, altre ancora sono espressione della dimensione ambientale.

Quindi si applica l'ACP in modo che si costruiscano solo tre componenti principali: una componente principale (CP) che contenga solo variabili economiche, una seconda componente principale che contenga solo variabili sociali e una terza componente principale che contenga solo variabili ambientali. Questo si ottiene associando alla componente principale “economia” (CP-economia) le variabili economiche con il coefficiente di correlazione più alto in modulo con la CP-economia; associando alla componente principale “società” (CP-società) le variabili sociali con il coefficiente di correlazione più alto in modulo con la CP-società; associando alla componente principale “ambiente” (CP-ambiente) le variabili ambientali con il coefficiente di correlazione più alto in modulo con la CP-ambiente.

Si compie questa procedura in modo iterativo con tutte le variabili a disposizione nella matrice  $X_{(103,p)}$ , e si prosegue in modo da generare una CP-economia composta dalle variabili elementari economiche con alto coefficiente di correlazione con la CP-economia, mentre tutte le altre variabili economiche che non

presentano un alto coefficiente di correlazione con la CP-economia sono scartate dall'analisi e non rientrano nel calcolo della funzione di sintesi  $\varphi(X)$  (indicatore di sintesi).

Allo stesso modo si procede per la selezione delle variabili elementari di contesto sociale, e quindi con quelle ambientali. Così procedendo per tentativi in modo iterativo si costruiscono le tre componenti principali, ciascuna delle quali chiaramente riferita ad una delle tre dimensioni della sostenibilità: economia, società, ambiente. Nella letteratura è possibile trovare indicazioni a sostegno della procedura appena esposta. In particolare si richiamano Jolliffe (2002 e 1986) e Zani (2000). Ancora in Zani (2000, p.121) si afferma che è possibile stabilire il numero delle componenti principali a priori.

Quanto appena esposto si può rappresentare anche in termini formali. Si faccia l'ipotesi che le tre componenti principali abbiano il seguente ordine<sup>3</sup>: economia, società, ambiente.

#### *La componente principale "economia"*

Per trovare la componente principale "economia", si risolve il seguente problema di massimo vincolato

$$\begin{aligned} \max_{a_1} \text{Var}(\tilde{X}a_1) &= a_1'Ra_1 \\ \text{s.v. } a_1'a_1 &= 1 \end{aligned}$$

dove l'obiettivo consiste nel determinare combinazioni lineari delle  $p$  variabili elementari con varianza massima. E dove  $\tilde{X}$  rappresenta l'insieme delle variabili elementari trasformate (ad esempio centrate rispetto alla media, oppure trasformate in logaritmi o con l'algoritmo Box-Cox<sup>4</sup>),  $a_1$  è il primo autovettore della matrice  $R$ ,  $R$  è la matrice di correlazione di  $\tilde{X}$ .

La soluzione si trova ricorrendo al moltiplicatore di Lagrange ( $\lambda$ ) e individuando il massimo del Lagrangiano che corrisponde all'annullamento del gradiente della funzione

$$\phi(a_1, \lambda) = a_1'Ra_1 - \lambda(a_1'a_1 - 1).$$

Le soluzioni sono  $p$ , quante sono le variabili inserite nella matrice  $\begin{matrix} X \\ (103, p) \end{matrix}$ , e sono le radici dell'equazione caratteristica  $|R - \lambda I| = 0$ . Delle  $p$  soluzioni si sceglie quella di valore maggiore, perché obiettivo del problema è la massimizzazione della funzione obiettivo. Questa soluzione si indica con  $\lambda_1$ , si chiama primo autovalore della matrice di correlazione, ad esso corrisponde l'autovettore  $a_1$  dei coefficienti della prima componente principale detta CP-economia. Quindi la macro-variabile di sintesi della dimensione "economia" è:

$$Z_i^{\text{economia}} = a_1X_{1i} + a_2X_{2i} + \dots + a_kX_{ki}.$$

dove  $i$  rappresenta l' $i$ -esimo comune capoluogo e  $k$  è inferiore a  $p$  (il numero di variabili).

#### *La componente principale "società"*

La componente principale che descrive la dimensione sociale è data dalla soluzione al seguente problema di massimo vincolato

$$\begin{aligned} \max_{a_2} \text{Var}(\tilde{X}a_2) &= a_2'Ra_2 \\ \text{s.v.} \end{aligned}$$

<sup>3</sup> Ma non è detto che dall'analisi dei dati relativi ai comuni capoluogo delle province italiane relativi ad un dato anno l'ordine con cui vengono generate le componenti principali sia proprio questo.

<sup>4</sup> La trasformazione Box-Cox consiste nell'applicare alla variabile elementare con distribuzione asimmetrica o caratterizzata

da accentuata curtosi il seguente algoritmo: 
$$\tilde{X}(\lambda) = \begin{cases} \log X & \text{per } \lambda = 0 \\ \frac{X^\lambda - 1}{\lambda} & \text{per } \lambda \neq 0 \end{cases}.$$

$$a_2' a_2 = 1$$

$$a_1' a_2 = 0$$

dove il secondo vincolo garantisce che questa seconda componente principale sia ortogonale alla prima componente principale. Le soluzioni sono  $p$ , quante sono le variabili inserite nella matrice  $X_{(103,p)}$ , e sono le radici dell'equazione caratteristica  $|R - \lambda I| = 0$ . Delle  $p$  soluzioni si sceglie la seconda più grande, e la si indica con il simbolo  $\lambda_2$ . Si tratta del secondo autovalore, a cui corrisponde l'autovettore  $a_2$ . La seconda componente principale quindi è rappresentata dalla seguente equazione, che descrive la dimensione "società" dei sistemi territoriali considerati:

$$Z_i^{società} = b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki}.$$

#### *La componente principale "ambiente"*

La componente principale che descrive la dimensione ambientale è data dalla soluzione al seguente problema di massimo vincolato

$$\max_{a_3} \text{Var}(\tilde{X}a_3) = a_3' R a_3$$

s.v.

$$a_2' a_2 = 1$$

$$a_1' a_2 = 0$$

$$a_1' a_3 = 0$$

$$a_2' a_3 = 0$$

dove i vincoli successivi al primo garantiscono che le tre componenti principali estratte siano ortogonali tra di loro. La soluzione è il terzo autovalore in ordine di grandezza,  $\lambda_3$ , a cui corrisponde l'autovettore  $a_3$ . La terza componente principale CP-ambiente è quindi la seguente equazione

$$Z_i^{ambiente} = c_1 X_{1i} + c_2 X_{2i} + \dots + c_k X_{ki}.$$

#### *La selezione delle variabili*

L'ACP è utile per la costruzione di macro-variabili di sintesi e ancor prima nell'ordine operativo dell'applicazione della procedura è utile per compiere la selezione statistica delle variabili. Le variabili utilizzate per la costruzione delle tre componenti principali economia-società-ambiente sono state selezionate in base alla procedura iterativa descritta all'inizio del paragrafo<sup>5</sup> ed anche tenendo in considerazione:

1. il principio di fattorializzabilità della matrice di correlazione<sup>6</sup>
2. la multicollinearità delle variabili elementari
3. la necessità di contenere la ridondanza di informazioni, come da più parti ricordato, tra gli altri Jolliffe (2002 e 1986), Krzanowski (1988), Zani (2000), Ganugi (1999), e quindi la considerazione che delle  $p$  variabili a disposizione già le prime  $k$  (con  $k < p$ ) contengono la maggior parte delle

<sup>5</sup> Data la matrice  $X$ , si applica l'analisi in componenti principali e si estraggono solo le variabili elementari che si aggregano in una componente principale, che raccoglie solo variabili della medesima natura (ad esempio economica) e che hanno alta correlazione con la componente principale di sintesi "economia". Lo stesso si fa per le altre variabili elementari relative alla società e all'ambiente.

<sup>6</sup> La fattorializzabilità della matrice di correlazione viene in genere verificata con due test:

- test di sfericità di Bartlett: si pone come ipotesi zero  $R=I$  (l'uguaglianza tra la matrice di correlazione e la matrice identità); il test rifiuta l'ipotesi zero, quindi non accetta che le variabili siano incorrelate;
- il test KMO (Kaiser-Meyer-Olkin), occorre che sia superiore a 0,5 altrimenti la matrice non è fattorializzabile.

informazioni originarie, di conseguenza che le altre  $(p-k)$  variabili della matrice  $X_{(103,p)}$  sono ridondanti e la loro esclusione dall'analisi non comporta una perdita di informazioni significativa.

*L'applicazione dell'ACP ai dati (anno 2001)*

Tornando alla costruzione dell'indice di sviluppo urbano sostenibile proposto nel capitolo 2, l'ACP ha consentito di riaggregare 9 variabili in 3 macro-variabili, che complessivamente riescono a spiegare quasi il 69% della varianza totale del fenomeno.

Dalla tabella seguente (Tab. 6), emerge che, dopo la rotazione, la prima delle tre componenti principali (CP) da sola riesce a spiegare il 29% della varianza totale del fenomeno descritto dalle nove variabili di partenza, la seconda componente principale il 24%, la terza componente principale il 16%.

**Tab. 6 La varianza totale spiegata (anno 2001)**

Component	Initial		Cumulative	Extraction		Cumulative	Rotation		Cumulative
	Eigenvalues			Sums of Squared Loadings			Sums of Squared Loadings		
	Total	% of Variance	%	Total	% of Variance	%	Total	% of Variance	%
1	3,339	37,099	37,099	3,339	37,099	37,099	2,569	28,544	28,544
2	1,803	20,030	57,129	1,803	20,030	57,129	2,162	24,026	52,571
3	1,043	11,588	68,717	1,043	11,588	68,717	1,453	16,146	68,717
4	,802	8,915	77,632						
5	,685	7,608	85,240						
6	,531	5,903	91,143						
7	,390	4,336	95,479						
8	,266	2,957	98,436						
9	,141	1,564	100,000						

Extraction Method: Principal Component Analysis.

Inoltre, se si analizzano le comunaltà finali tra le variabili elementari e le variabili di sintesi (o CP), si nota che le tre CP estratte riescono a spiegare l'87% della varianza del tasso di disoccupazione, l'84% della varianza del valore aggiunto pro capite, l'81% della varianza dell'indice di dotazione di strutture sociali, e altrettanto alti sono i valori per altre variabili. Ciò a riprova del contenuto informativo rilevante che si è riusciti a trasferire nelle macro-variabili di sintesi costruite (le tre CP).

In più come si può notare dalla tabella seguente (Tab. 7), la prima componente principale (CP) è la macro-variabile sociale e cattura quattro variabili elementari che descrivono: la criminalità, la cultura, la dotazione di strutture sociali e il trasporto pubblico.

La dotazione di strutture a servizio della collettività locale racchiude in sé variabili che forniscono indicazioni sulla disponibilità e sulla qualità delle strutture per l'arricchimento culturale extra-scolastico, per lo svago e il tempo libero. Si tratta di musei, biblioteche, cinematografi, teatri, strutture per la pratica sportiva. Comprende anche dati sulla dotazione di strutture per la formazione scolastica, dalle scuole materne, alle elementari, le medie inferiori e superiori, fino alle università. Al suo interno si trovano anche dati relativi alle strutture sanitarie, tra cui soprattutto tutti gli ospedali, convenzionati e non convenzionati con il Servizio Sanitario Nazionale.

I valori positivi dei coefficienti di correlazione associati a queste variabili esprimono il fatto che la variabile di sintesi descrive il grado di offerta presente sul territorio considerato di ciascuna delle strutture sociali appena descritte.

La prima CP ha al proprio interno anche la dimensione negativa della criminalità.

La seconda componente principale è espressione dei caratteri prevalentemente economici dei comuni capoluogo di provincia, si può interpretare come misura del vigore economico di un territorio. È infatti



sintesi del valore aggiunto pro capite e del tasso di disoccupazione. Il legame tra le due variabili elementari economiche è inversamente proporzionale.

Un limite insito in questa macro-variabile è rappresentato dal fatto che le variabili elementari che la compongono non sono relative all'ambito territoriale comune capoluogo, ma al sistema locale del lavoro (si veda la letteratura prodotta da Sforzi-Istat e a partire dal seminario di Bologna del novembre 2003 anche da Banca d'Italia) a cui il comune capoluogo appartiene<sup>7</sup>.

**Tab. 7 La matrice delle componenti principali ruotate (anno 2001)**

Variabili elementari costitutive	Componente principale		
	1	2	3
	Sviluppo sociale	Sviluppo economico	Tutela ambientale
Tasso di disoccupazione	-0,147	<b>-0,896</b>	-0,204
Valore aggiunto pc	0,234	<b>0,878</b>	0,129
Concentrazione PM 10	-0,224	-0,098	<b>-0,742</b>
Politiche per l'ambiente	0,004	0,436	<b>0,453</b>
Depurazione dell'acqua	-0,129	0,108	<b>0,783</b>
Criminalità	<b>-0,790</b>	0,323	-0,073
Cultura	<b>0,757</b>	0,339	0,117
Strutture sociali	<b>0,876</b>	0,204	0,025
Trasporto pubblico	<b>0,680</b>	0,342	-0,081
Misura di adeguatezza campionaria KMO (Keiser Meyer Olkin)	0,706		
Test di sfericità di Bartlett	Chi-quadrato appross. gradi di libertà	355,353 36	
	p-value	0,000	

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.  
a Rotation converged in 6 iterations.

La terza componente principale è la macro-variabile ambientale ed esprime elementi tipici della qualità ambientale in aree urbane e in aree fortemente antropizzate: la concentrazione delle polveri, la capacità di depurazione delle acque reflue, le politiche per la tutela ambientale. Cattura quindi in sé le variabili che esprimono il peso dell'inquinamento atmosferico (causato dallo svolgimento di attività economiche e di attività sociali), cattura la dotazione di infrastrutture per la depurazione delle acque, racchiude in sé anche le politiche messe in atto a livello locale per la tutela e la difesa dell'ambiente.

Costruite le tre componenti principali, l'ACP produce gli *scores* per ciascuna di esse. È così possibile avere il valore della macro-variabile economica, della macro-variabile sociale e della macro-variabile ambientale per ciascuno dei comuni capoluogo. Gli *scores* prodotti si possono interpretare come la sintesi quantitativa delle variabili che hanno catturato. In questo modo, per ogni territorio non si ha più solo il dato del valore aggiunto e del tasso di disoccupazione, si ha anche un valore nuovo che si ritiene possa essere una buona sintesi di queste due variabili. Allo stesso modo per la società e per l'ambiente.

<sup>7</sup> Per superare questo limite si può procedere in due modi diversi:

- eliminazione della dimensione "economia", quindi conseguentemente calcolo dell'indice di sintesi di sostenibilità in base a due sole dimensioni (società e ambiente, perché per queste esistono variabili rilevate a livello di comune capoluogo)

- stima del valore aggiunto pro capite e di altre variabili economiche su scala di comune capoluogo, mediante le tecniche sviluppate da Rao N.K. (si veda un primo paper significativo pubblicato nel 1994 insieme a Ghosh M. su Statistical Science) e attualmente implementate da Istat per i sistemi locali del lavoro.

In più si potrebbe anche sostenere che ciascuna delle variabili sintetizzate dalle tre CP è tra le variabili maggiormente rappresentative rispettivamente della dimensione sociale, economica e ambientale. E partendo da queste variabili elementari gli attori locali e gli enti pubblici potrebbero compiere analisi di almeno una parte delle politiche locali di sviluppo.

#### *Test di verifica della significatività dell'ACP*

La scelta dell'uso dell'ACP è anche confortata dalla significatività dei test di controllo: sia il KMO (pari a 0,706) sia il test di sfericità di Barlett (pari a 355,353, con un *p-value* pari a 0,000)<sup>8</sup> sono entrambi significativi.

### **2.4 La clusterizzazione dei 103 comuni capoluogo**

Il secondo passaggio per la costruzione dell'indicatore di sintesi dello sviluppo sostenibile nei comuni capoluogo di provincia italiani consiste nel suddividere le 103 città in 2 cluster, in base a due categorie ordinali del concetto di "sviluppo sostenibile", basso e alto.

Nella costruzione dell'indice di sviluppo urbano sostenibile, dopo vari tentativi si è preferito tralasciare i metodi gerarchici dell'analisi dei gruppi, perché portano all'individuazione di un numero di gruppi troppo elevato, ma soprattutto perché l'individuazione del taglio quando le unità statistiche sono numerose risulta essere molto complesso. Nella costruzione dell'indice di sostenibilità urbana si è scelto di ricorrere all'approccio "quick cluster" noto anche come "clusterizzazione non gerarchica k-medie". Consiste nel porre pari a due il numero dei gruppi da costruire e si pone come condizione necessaria per procedere che i due centroidi finali dei due cluster

$\bar{X}_1 = [\bar{x}_{11}, \bar{x}_{21}, \dots, \bar{x}_{k1}]$  e  $\bar{X}_2 = [\bar{x}_{12}, \bar{x}_{22}, \dots, \bar{x}_{k2}]$  siano caratterizzati dai seguenti aspetti:

- per ciascuna delle *k* variabili il valore del centro finale nei due gruppi abbiano segno opposto
- un centroide deve essere espressione di "economia negativa", "società negativa", "ambiente negativo", e simmetricamente l'altro centroide deve essere espressione di "economia positiva", "società positiva", "ambiente positiva". Se queste due condizioni sono soddisfatte, e se il test di bontà<sup>9</sup>

$R^2 = 1 - \frac{W}{T}$  viene soddisfatto, allora si accettano le due partizioni ottenute.

#### *L'applicazione dell'analisi cluster ai dati (anno 2001)*

Si inseriscono nell'analisi cluster non gerarchizzata k-medie le 12 variabili selezionate con l'ACP. Si è stabilito di riaggregare le unità statistiche (i 103 comuni capoluogo di provincia) in due gruppi<sup>10</sup>. La procedura k-medie inizia con la costruzione dei centri dei cluster iniziali. Attraverso una serie di iterazioni viene modificata la composizione dei due gruppi, in modo da soddisfare gli obiettivi della tecnica ("coesione interna" e "separazione esterna" tra i due cluster).

Al termine delle iterazioni, quando il processo si stabilizza e raggiunge la convergenza, si ottengono alcuni risultati piuttosto interessanti. Si nota soprattutto che i centri dei cluster finali presentano valori con segni coincidenti con quelli attribuiti dall'ACP. E in più il cluster 1 ha centri finali di significato economico-sociale-ambientale negativo, mentre il cluster 2 ha centri finali con significato economico-

<sup>8</sup> Il Test KMO (Kaiser-Meyer-Olkin) esamina se le correlazioni parziali delle variabili sono basse; è necessario che il valore sia superiore a 0,5. Il Test di sfericità di Bartlett esamina se la matrice di correlazione è una matrice identità; è necessario che il valore del *p-value* sia nullo o molto prossimo allo zero. Anche i valori posti lungo la diagonale principale della "matrice anti-immagine" (misure di adeguatezza campionaria) sono tutti al di sopra di 0,5. Quindi anche questo test è superato. La media aritmetica di questi valori è esattamente pari al test KMO.

<sup>9</sup> Il test ricordato è preso da Zani (2000) pag. 219. Si basa sulla scomposizione della devianza totale (*T*) delle *p* variabili in devianza nei gruppi (within, *W*) e devianza tra i gruppi (between, *B*).

<sup>10</sup> In termini socioeconomici si rimanda al modello di sviluppo dualistico tipico dell'Italia degli anni Sessanta (si vedano al riguardo ad esempio Garofoli (1992), Ciciotti e Rizzi (2005))

sociale-ambientale positivo. E ancora tutti i centri nei cluster finali hanno segno coerente con il significato di sostenibilità ricercato.

**Tab. 8 I centri finali dei due cluster**

Variabili elementari utilizzate	Cluster	
	1	2
Valore aggiunto pc	-0,85016	0,68609
Disoccupazione	0,88945	-0,71780
Cultura	-0,58655	0,47336
Criminalità	0,09416	-0,07599
Strutture sociali	-0,59471	0,47994
Trasporto pubblico	-0,52023	0,41984
depurazione acque	-0,13491	0,10888
Concentrazione PM 10	0,33795	-0,27273
Politiche per l'ambiente	-0,32471	0,26204

Nota: le variabili elementari inserite nel calcolo dei cluster sono state standardizzate. La convergenza è stata raggiunta dopo 11 iterazioni.

Quindi nel primo cluster si trovano i comuni capoluogo di provincia con il livello più elevato di disoccupazione, di criminalità, e la maggior concentrazione di PM10.

Nel secondo cluster invece sono raggruppati i comuni capoluogo di provincia in cui si osservano le condizioni esattamente opposte: maggior livello di valore aggiunto, minor livello di disoccupazione, minori livelli di criminalità, maggior dotazione di strutture sociali, maggior attenzione all'ambiente, mediante l'adozione di politiche pubbliche per la tutela ambientale, buona dotazione di impianti per la depurazione delle acque.

**Tab. 9 Il numero di province per cluster e per ripartizione geografica**

Ripartizioni geografiche	Cluster		totali
	1	2	
Nord-Centro Italia	12	55	67
Sud Italia	34	2	36
Totale province	46	57	103

Nota: le province del Nord nel cluster 1 sono (4): Asti, Verbanò-Cusio-Ossola, Vercelli, Rovigo. Le province del Centro presenti nel cluster 1 sono (8): Ascoli Piceno, Frosinone, Grosseto, Latina, Massa Carrara, Rieti, Terni, Viterbo. Le province del Sud Italia presenti nel cluster 2 sono (2): L'Aquila e Pescara.

Si osservano valori quasi equamente distribuiti nei due cluster, con distribuzioni prossime alla normale (si veda in appendice Fig 6). Valori anomali si associano a realtà urbane effettivamente diverse dalla media di ciascun gruppo: nel Centro-Nord emergono positivamente rispetto alla media Trieste, Roma, Milano e Verona, mentre al Sud si nota che la dispersione è minore, i dati sono più raccolti attorno alla media e alla mediana, eccezion fatta per Napoli e Lecce, che hanno valori abbondantemente superiori alla media. In sintesi si nota che la variabilità all'interno dei due cluster è molto contenuta.

#### *La transvariazione*

L'analisi di transvariazione rientra nei test di selezione della singola variabile da inserire in una macrovariabile di sintesi, insieme ad altri test, tra cui l'analisi preliminare della variabilità e della forma della variabile, misure di correlazione, il test KMO relativo alla singola variabile, il dato della comunalità finale attribuita alla singola variabile. In termini operativi si ricorre alla transvariazione qualora, essendo il coefficiente di correlazione tra due variabili in valore assoluto superiore a 0,80,

allora si debba stabilire quale delle due variabili debba essere eliminata. In considerazione del fatto che una delle finalità della metodologia che si sta descrivendo, è l'analisi cluster, si può affermare che delle due si elimina quella con la transvariazione minore, congiuntamente a varianza between maggiore e varianza within minore (relative all'analisi congiunta della varianza tra gruppi e nei gruppi).

L'indice relativo di transvariazione assume valori nell'intervallo [0, 1] e precisamente risulta uguale a 0 quando le due distribuzioni non presentano alcuna sovrapposizione tra loro; risulta uguale a 1 quando le mediane dei due gruppi sono uguali, per cui è il medesimo il punto che discrimina il primo 50% ed il secondo 50% dei valori delle due distribuzioni.

Inoltre se  $Me_g > Me_h$ , estraendo a caso un elemento del primo gruppo ed un elemento del secondo, la probabilità che le due unità estratte presentino una disuguaglianza con verso opposto rispetto a quella delle mediane è nulla se  $f_{gh}^* = 0$  e tende ad 1/2 se  $f_{gh}^*$  tende a 1.

“Dire che la transvariazione tra due gruppi, rispetto ad un carattere, ha un certo valore significa che è nulla, piccola, grande l'area in comune delle due distribuzioni, le quali vengono così, in un certo senso, raffrontate reciprocamente (...). Alla domanda se le mediane del carattere, nei due gruppi, differiscono tra loro in modo da rappresentare quanto avviene per i casi singoli, l'indice di transvariazione fornisce una risposta perentoria” (Boldrini, 1968).

Nella tavola 10 sono riportati i valori del test di trans variazione per alcune variabili elementari escluse dall'indicatore di sintesi causa una bassa frequenza relativa di trans variazione, congiuntamente ad un'alta varianza between e una bassa varianza within, nei confronti a coppie con le variabili elementari che invece si è deciso di inserire nell'indice composito.

**Tab. 10 Transvariazione e scomposizione della varianza**

Variabili elementari selezionate con l'ACP	Frequenza relativa di transvariazione	Varianza totale	% Varianza nei gruppi (within)	% Varianza tra i gruppi (between)
Verde urbano fruibile mq/ab (2000)	0,43	2,63	79%	21%
Isole pedonali mq x ab (dato 2000)	0,63	1,03	95%	5%
Piste ciclabili metri lineari x ab (2000)	0,23	0,84	67%	33%
Omicidi (2000)	0,27	1,52	66%	34%
Furti di autoveicoli x 100.000 ab (2000)	0,21	0,66	55%	45%
Protesti di assegni per 1.000 ab (2000)	0,29	0,70	67%	33%

### 2.5 L'analisi discriminante (l'approccio lineare di Fisher)

Mediante l'analisi discriminante si vuole attribuire un ordinamento alle unità statistiche riportate nei due gruppi definiti con l'analisi cluster descritta nel paragrafo precedente.

Per avviare l'analisi discriminante si sono recuperati gli scores prodotti dall'ACP descritta nelle pagine precedenti e i due cluster definiti con la procedura non gerarchica k-medie. La tecnica definisce una funzione di classificazione per ognuno dei due cluster<sup>11</sup>. Le funzioni di classificazione sono utilizzate

<sup>11</sup> Dato un set di variabili indipendenti riferite a  $n$  unità statistiche e dati i gruppi (cluster) a cui appartengono le  $n$  unità statistiche, l'analisi discriminante definisce una relazione lineare tra queste variabili in modo da differenziare al meglio i due gruppi, e in modo che si rendano massime le differenze tra i gruppi. Le relazioni lineari tra le variabili indipendenti vengono chiamate “funzioni discriminanti” e si indicano con la seguente equazione:  $y_i = a_1X_{1i} + a_2X_{2i} + a_3X_{3i} + \dots + a_pX_{pi}$  dove:

per assegnare le unità statistiche (in questo caso i 103 comuni capoluoghi) ai due gruppi. Per ogni comune capoluogo ognuna delle due funzioni di classificazione calcola lo score per il gruppo corrispondente. Il modello discriminante assegna il comune capoluogo al gruppo per il quale le due funzioni di classificazione hanno calcolato il valore maggiore.

Dall'analisi del valore e del segno dei "coefficienti delle funzioni di classificazione", si nota che il primo gruppo è costituito da comuni capoluogo di provincia in cui le tre dimensioni della sostenibilità (economia, società, ambiente) hanno tutte espressione negativa, e quindi vanno lette come: declino economico, declino sociale, degrado ambientale. Significa che nel primo gruppo sono stati collocati i comuni capoluogo peggiori in tutte e tre le dimensioni. Nel gruppo due invece si aggregano comuni capoluogo di provincia in cui le stesse tre dimensioni hanno manifestazione positiva. Significa che nel secondo gruppo sono stati inseriti i comuni migliori sia in economia, sia in società sia in ambiente. Sostanzialmente attraverso l'analisi discriminante si sono riusciti a ricostruire i due gruppi definiti nel corso dell'analisi cluster k-means, distinguendo con evidenza i comuni migliori da quelli peggiori in termini di economia, società, ambiente.

**Tab. 11 I coefficienti della funzione di classificazione (anno 2001)**

	Cluster	
	1	2
società	-1,561	1,259
economia	-2,928	2,363
ambiente	-,511	,412
(constant)	-2,274	-1,722

Nota: Funzioni discriminanti lineari di Fisher

Come il risultato offerto dall'evidenza empirica dei comuni italiani nei tre anni 1999-2001, che conferma alcuni "sentire" comuni in merito al dualismo dello sviluppo in Italia e alle due velocità della crescita, così anche i test dell'analisi discriminante confermano la bontà statistica dei risultati. Infatti una volta analizzati i risultati delle funzioni lineari di Fisher, occorre verificare una serie di test che in sintesi confermano la significatività statistica dell'analisi proposta.

I test a cui sottoporre i risultati sono presentati brevemente di seguito, e sono nell'ordine:

- l'assenza di collinearità tra le variabili indipendenti (dall'analisi delle correlazioni "all'interno dei gruppi" tra le macro-variabili economia-società-ambiente si nota che in ogni anno i valori sono molto contenuti, il valore maggiore si registra nel 2001 tra economia e società (-0,47). Interessante notare che le correlazioni sono sempre negative);
- la non correlazione tra le medie e le varianze dei gruppi (dall'analisi dei dati emerge che questa condizione è soddisfatta);
- l'omogeneità delle matrici di covarianza (il test nel caso in esame risulta essere significativo).

Oltre a questi test è necessario verificare:

- il contributo delle singole variabili indipendenti all'analisi discriminante, attraverso: a) il test di uguaglianza tra le medie dei gruppi (in cui il p-value è significativo per economia e società, mentre per

-  $Y_i$  è il punteggio dell'unità statistica i-esima assegnato dalla funzione discriminante

-  $X_{pi}$  è il valore della variabile indipendente p-esima per l'unità statistica i-esima

-  $a_p$  è il coefficiente della funzione discriminante relativo alla p-esima variabile indipendente.

I coefficienti  $a_p$  vengono stimati in modo che la differenza tra i gruppi definiti dalla variabile Y sia massima. Ciò significa che si individua una relazione lineare che renda massima la devianza tra i gruppi e minima la devianza entro i gruppi rispetto alle variabili indipendenti.

la variabile ambiente è pari a 0,19, superiore al massimo fissato a 0,10); b) i coefficienti standardizzati della funzione discriminante canonica (da cui si evince che le variabili con maggior potere discriminante sono economia e società, mentre la variabile ambiente appare più debole); c) la matrice di struttura (da cui emerge che la correlazione più forte si ha tra la variabile economia e la variabile di sintesi stimata dalla funzione discriminante (correlazione pari a 0,68; segue la variabile “società” con 0,26; ultima delle tre la variabile “ambiente” con solo 0,08);

- il grado di adattamento ai dati da parte del modello discriminante, mediante due indicatori distinti: a) gli autovalori (che sono funzionali al calcolo della correlazione tra gli scores della funzione discriminante. Il valore del coefficiente nel caso dei dati utilizzati è molto alto, pari a 0,85); b) il test lambda di Wilks (misura la capacità della funzione discriminante di separare i comuni capoluogo nei due cluster territoriali. A questo si associa un chi-quadro che nel caso in esame risulta essere particolarmente significativo).

## 2.6 L'indice di sviluppo urbano sostenibile

La combinazione delle tre analisi appena descritte (l'analisi in componenti principali, l'analisi cluster, l'analisi discriminante) consente di costruire un indice di sostenibilità urbana, nuovo e originale nella struttura statistico-matematica<sup>12</sup>.

Attraverso l'ACP si compie la selezione delle variabili e si costruiscono le tre componenti economia-società-ambiente espressione dello sviluppo sostenibile. L'analisi cluster applicata a condizione che si soddisfino le condizioni richieste riparte i comuni capoluogo in due gruppi omogenei. L'analisi discriminante costruisce la funzione di sintesi. In questo modo le tre metodologie di statistica multivariata utilizzate portano alla definizione di un indicatore di sintesi.

### *Selezione delle variabili e definizione delle tre componenti principali economia-società-ambiente*

Utilizzando l'ACP si selezionano le variabili di base e si costruiscono tre componenti principali, ciascuna descrive una delle dimensioni della sostenibilità (economia, società, ambiente). In termini formali le tre componenti principali sono così sintetizzabili:

$$\begin{cases} Z_i^{economia} = a_1 X_{1i} + a_2 X_{2i} + \dots + a_k X_{ki} \\ Z_i^{società} = b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki} \\ Z_i^{ambiente} = c_1 X_{1i} + c_2 X_{2i} + \dots + c_k X_{ki} \end{cases}$$

dove  $i$  indica l' $i$ -esimo comune capoluogo di provincia italiano,  $X_{ki}$  indica il valore rilevato nell' $i$ -esimo comune capoluogo per la  $k$ -esima variabile selezionata,  $Z$  rappresenta gli scores prodotti dall'ACP per l' $i$ -esimo comune capoluogo per ognuna delle tre dimensioni della sostenibilità (economia, società, ambiente).

### *Clusterizzazione dei territori*

Utilizzando l'analisi cluster si suddividono le unità statistiche in due gruppi, ammesso che siano soddisfatte date condizioni. Si pone infatti come condizione necessaria per procedere che i due centroidi finali  $\bar{X}_1 = [\bar{x}_{11}, \bar{x}_{21}, \dots, \bar{x}_{k1}]$  e  $\bar{X}_2 = [\bar{x}_{12}, \bar{x}_{22}, \dots, \bar{x}_{k2}]$  dei cluster estratti siano caratterizzati dai seguenti aspetti:

- per ciascuna delle  $k$  variabili il valore del centro finale nei due gruppi abbiano segno opposto

<sup>12</sup> Il manuale OECD del 2008, che descrive le metodologie possibili per costruire indicatori compositi, tratta le tre tecniche, soffermandosi principalmente sull'ACP e sull'analisi cluster k-medie, ma non arriva a proporre una loro combinazione sequenziale come invece si propone in questo paper.

- un centroide deve essere espressione di “economia negativa”, “società negativa”, “ambiente negativo”, e simmetricamente l’altro centroide deve essere espressione di “economia positiva”, “società positiva”, “ambiente positiva”. Se queste due condizioni sono soddisfatte, e se il test di bontà<sup>13</sup>

$R^2 = 1 - \frac{W}{T}$  viene soddisfatto, allora si accettano le due partizioni ottenute.

#### *Sintesi delle variabili di clusterizzazione*

Utilizzando le tre componenti principali economia-società-ambiente ottenute con l’ACP, tenendo in considerazione la suddivisione dei comuni capoluogo nei due gruppi ottenuti con l’analisi cluster, mediante l’analisi discriminante si costruisce un indicatore di sintesi della sostenibilità urbana, così rappresentabile

$$Y_i = \alpha Z_i^{economia} + \beta Z_i^{società} + \gamma Z_i^{ambiente}$$

dove  $i$  indica l’ $i$ -esimo comune capoluogo di provincia.

---

<sup>13</sup> Il test ricordato è preso da Zani (2000) pag. 219. Si basa sulla scomposizione della devianza totale (T) delle  $p$  variabili in devianza nei gruppi (within, W) e devianza tra i gruppi (between, B).

### 3. I dati e le variabili

In questo capitolo si descrivono brevemente le variabili elementari selezionate per la costruzione dell'indicatore di sostenibilità dei comuni italiani capoluogo di provincia

#### 3.1 Le variabili elementari per dimensione della sostenibilità

L'indice di sviluppo urbano sostenibile è dato dalla combinazione statistica di variabili elementari. Le variabili utilizzate per ciascuno dei tre anni sono riportate nella tavola seguente (Tab. 12). La macro-variabile "economia" è costituita in tutti e tre gli anni analizzati dalle medesime variabili elementari osservate: il "valore aggiunto pro capite" e "il tasso di disoccupazione" rilevati a livello di sistema locale del lavoro. La macro-variabile "società" è data dall'aggregazione di quattro variabili che rimangono immutate nel corso degli anni: "cultura", "criminalità", "strutture sociali" e "trasporto pubblico". L'"ambiente" viene descritto con variabili in parte diverse anno per anno, comunque sempre relative a determinate dimensioni tipicamente espressione del patrimonio naturale e dell'azione dell'uomo su di esso: la risorsa acqua, l'inquinamento, le aree fruibili. Infatti come si nota nella tabella seguente nell'ambiente compare costantemente per tutti e tre gli anni la variabile elementare "depurazione dell'acqua"; la "concentrazione di PM10" utilizzata per i primi due anni nel terzo anno viene sostituita dalla "concentrazione di nitrati".

**Tab. 12 Le dimensioni della sostenibilità e le variabili costitutive**

Macro-variabili	anno 2001	anno 2000	anno 1999
"Economia"	- Valore aggiunto pro capite - Tasso di disoccupazione	- Valore aggiunto pro capite - Tasso di disoccupazione	- Valore aggiunto pro capite - Tasso di disoccupazione
"Società"	- Cultura - Criminalità - Strutture sociali - Trasporto pubblico	- Cultura - Criminalità - Strutture sociali - Trasporto pubblico	- Cultura - Criminalità - Strutture sociali - Trasporto pubblico
"Ambiente"	- Depurazione acqua - PM 10 - Politiche per l'ambiente	- Depurazione acqua - PM 10 - Zone a traffico limitato	- Depurazione acqua - Concentrazione nitrati - Isole pedonali

Due delle variabili aggregate nella macro-variabile "società" sono state ottenute attraverso sintesi parziali intermedie. La variabile "cultura" è ottenuta come aggregazione intermedia delle variabili elementari: strutture per le attività ricreative, strutture per attività culturali, strutture per l'istruzione. La variabile "criminalità" è il risultato di un'aggregazione, condotta facendo ricorso alla tecnica dell'ACP. Raccoglie in sé le seguenti variabili elementari: furti di auto, rapine, scippi, minori denunciati, furti in casa.

#### 3.2 Il posizionamento delle singole città nelle dimensioni della sostenibilità

In genere quando si descrive il dataset delle variabili elementari utilizzate nell'analisi si propongono tavole contenenti gli indici statistici di posizione (media, mediana, ...), gli indici di variabilità (varianza, ...) e gli indici di forma (asimmetria e curtosi). In questo caso è parso più interessante proporre il posizionamento delle città italiane. Quindi nelle tavole seguenti sono riportate le prime tre città per ognuna delle variabili inserite nell'indice di sintesi, raggruppate per dimensione, "economia", "società" e "ambiente".

##### *Le graduatorie delle città nelle variabili costitutive della dimensione "economia"*

L'economia è misurata in tutti i tre anni esaminati dal valore aggiunto pro capite e dal tasso di disoccupazione. Presenta caratteristiche note e ormai consolidate nell'opinione pubblica: l'Italia si presenta divisa nettamente in due parti chiaramente distinte: il Nord e il Sud. Al Nord primeggiano



Bolzano e prima ancora Milano per l'alto valore aggiunto prodotto, così come per il basso livello di disoccupazione registrato. Nel Centro Roma e Firenze hanno il valore aggiunto pro capite maggiore insieme a Siena che con Pesaro registra la disoccupazione minore. Al Sud L'Aquila nel 2001 è il capoluogo economicamente più avvantaggiato, insieme a Matera realtà distrettuale emergente e i territori urbani nuoresi della Sardegna che testimoniano i buoni risultati prestazionali realizzati dall'isola e dal Mezzogiorno in questi ultimi periodi.

**Tab. 13 Le prime tre città in graduatoria per ripartizione geografica secondo le variabili elementari della dimensione "economia"**

Ranking	anno 2001		anno 2000		anno 1999	
	Valore aggiunto pc (max)	Tasso di disoccupazione (min)	Valore aggiunto pc (max)	Tasso di disoccupazione (min)	Valore aggiunto pc (max)	Tasso di disoccupazione (min)
<b>Classifica assoluta Italia</b>						
1°	Milano	Bolzano	Milano	Bolzano	Milano	Bolzano
2°	Bolzano	Lecco	Bolzano	Trento	Bolzano	Trento
3°	Imperia	Trento	Imperia	Lecco	Imperia	Lecco
<b>Classifica nord Italia</b>						
1°	Milano	Bolzano	Milano	Bolzano	Milano	Bolzano
2°	Bolzano	Lecco	Bolzano	Trento	Bolzano	Trento
3°	Imperia	Trento	Imperia	Lecco	Imperia	Lecco
<b>Classifica centro Italia</b>						
1°	Roma	Siena	Roma	Siena	Roma	Siena
2°	Firenze	Pesaro	Firenze	Ancona	Pesaro	Arezzo
3°	Siena	Arezzo	Siena	Pesaro	Firenze	Pesaro
<b>Classifica sud Italia</b>						
1°	L'Aquila	Pescara	L'Aquila	Pescara	Matera	Pescara
2°	Matera	Teramo	Matera	Teramo	L'Aquila	L'Aquila
3°	Nuoro	L'Aquila	Nuoro	L'Aquila	Nuoro	Teramo

Occorre notare che struttura e prestazione economica mantengono per il triennio in esame le medesime città nei primi posti della classifica in Italia e in ogni ripartizione territoriale. Bolzano, Milano, Trento e Lecco primeggiano nell'intera penisola e in Italia settentrionale. Roma con Firenze e altre città toscane (prime tra tutte Siena e Arezzo) sveltano nell'Italia centrale. L'Aquila, Matera, Teramo, Nuoro si distinguono tra le altre città del Mezzogiorno.

*Le graduatorie delle città nelle variabili costitutive della dimensione "società" nell'anno 2001*

La dimensione sociale è misurata nel 2001 da quattro variabili: cultura, criminalità, trasporto pubblico e strutture sociali. Se l'economia è caratterizzata dalla netta predominanza delle città del Nord, la società vede una buona presenza di realtà urbane del centro affiancare città del nord nelle prime posizioni delle classifiche delle singole variabili. In più, mentre nell'economia alle grandi città si uniscono città del nord di medie dimensioni, caratterizzate da spiccate capacità imprenditoriali e vocazioni distrettualistiche, nella società la classifica relativa all'intero territorio nazionale mette ai primi posti città di grandi dimensioni del nord e del centro. Così Trieste primeggia nella "cultura", nella dotazione di strutture sociali e nel trasporto pubblico insieme a Roma e a Firenze. Nel Sud Italia per la cultura si distinguono Catania e Palermo. Ancora Catania con Napoli primeggia per le strutture sociali e per il trasporto pubblico. Il dato della "criminalità" appare in controtendenza con il comune sentire. Tra le città con minor criminalità in Italia figurano Enna e altre realtà del sud per le quali si tenderebbe ad affermare che la delittuosità in genere violenta ed estorsiva soprattutto siano superiori che altrove. Già il Sole24 Ore e Italia Oggi si sono soffermati ampiamente negli anni passati su questa anomalia apparente dei dati statistici ufficiali.

**Tab. 14 Le prime tre città in graduatoria per ripartizione geografica secondo le variabili elementari della dimensione “società” (anno 2001)**

	<b>Cultura</b>	<b>Criminalità (dal minore)</b>	<b>Trasporto pubblico</b>	<b>Strutture sociali</b>	<b>macro-variabile “Società”</b>
<b>Classifica assoluta Italia</b>					
1°	Trieste	Campobasso	Roma	Trieste	Trieste
2°	Verona	Mantova	Milano	Firenze	Roma
3°	Roma	Enna	Trieste	Roma	Milano
<b>Classifica nord Italia</b>					
1°	Trieste	Mantova	Milano	Trieste	Trieste
2°	Verona	Trento	Trieste	Padova	Milano
3°	Milano	Rovigo	Bologna	Milano	Bologna
<b>Classifica centro Italia</b>					
1°	Roma	Ancona	Roma	Firenze	Roma
2°	Firenze	Chieti	Siena	Roma	Firenze
3°	Lucca	Grosseto	Firenze	Pisa	Pisa
<b>Classifica sud Italia</b>					
1°	Catania	Campobasso	Napoli	Napoli	Napoli
2°	Pescara	Enna	Cagliari	Bari	Catania
3°	Palermo	Agrigento	Catania	Catania	Palermo

**Tab. 15 Le prime tre città in graduatoria per ripartizione geografica secondo le variabili elementari della dimensione “ambiente”**

	<b>Concentrazione PM10 (dal minore)</b>	<b>Politiche per l’ambiente</b>	<b>Depurazione acque</b>	<b>macro-variabile “Ambiente”</b>
<b>Classifica assoluta Italia</b>				
1°	Lecce	Genova	Lecce	Lecce
2°	Arezzo	Perugia	Sondrio	Vercelli
3°	Biella	Lecce	Bolzano	Caltanissetta
<b>Classifica nord Italia</b>				
1°	Biella	Reggio Emilia	Sondrio	Vercelli
2°	Gorizia	Genova	Bolzano	Sondrio
3°	Varese	Bolzano	Aosta	Gorizia
<b>Classifica centro Italia</b>				
1°	Arezzo	Perugia	Perugia	Latina
2°	Viterbo	Arezzo	Livorno	Perugia
3°	Grosseto	Viterbo	Massa	Grosseto
<b>Classifica sud Italia</b>				
1°	Lecce	Lecce	Lecce	Lecce
2°	Potenza	Napoli	Salerno	Caltanissetta
3°	Caltanissetta	Foggia	Sassari	Caserta

*Le graduatorie delle città nelle variabili costitutive della dimensione “ambiente” nell’anno 2001*

Se già con la macro-variabile “società” il nord segnava il passo e doveva cedere posizioni a favore di città del centro Italia, per la dimensione ambientale la tendenza si accentua ulteriormente, tanto che in alcune variabili le città del nord non compaiono neppure tra le prime cinque classificate.

Nel caso dell’inquinamento atmosferico Biella è la città del Nord con le minori concentrazioni di polveri fini (PM10). Le altre città nei primi 5 posti sono del sud (Lecce) e toscano-laziali (Arezzo, Grosseto, Viterbo). Invece nelle politiche attuate a tutela dell’ambiente il nord riconquista qualche posizione, ancora con Bolzano, Sondrio, Reggio Emilia.

Da segnalare le buone posizioni occupate dalle città di Lecce in tutte le variabili ambientali utilizzate. A differenza del nord in cui politiche ambientali vengono promosse e attuate in centri urbani di medio-piccole dimensioni (Sondrio, Reggio Emilia, Bolzano), nel meridione le politiche a tutela del patrimonio naturale vengono attuate in ambiti urbani tendenzialmente di medio-grandi dimensioni. La sensibilità ambientale al Sud si diffonde meno rapidamente nelle piccole realtà urbane.

#### 4. L'indice di sviluppo urbano sostenibile nei 103 comuni italiani capoluoghi di provincia

In questo quarto capitolo si presenta l'elaborazione dell'indice di sviluppo urbano sostenibile dei comuni capoluogo di provincia. Si tratta dei risultati dell'implementazione della metodologia descritta nel secondo capitolo.

##### 4.1 La graduatoria dello sviluppo sostenibile nei comuni capoluogo di provincia

Bolzano, Milano, Trieste sono le città che nel 2001 occupano le prime tre posizioni nella classifica dello sviluppo sostenibile urbano in Italia. Sono le città in cui meglio si è riuscito a combinare i tre fattori della sostenibilità: crescita economica, equità e dinamismo sociale, tutela e valorizzazione ambientale. L'azione degli attori locali, le interazioni di cooperazione-competizione con i sistemi esterni, i condizionamenti dei macro-trend, gli influssi e gli impatti prodotti da interventi dei governi nazionale e europeo sono stati tali da consentire una crescita più equilibrata e più armoniosa in termini economici, sociali e ambientali rispetto alle altre città italiane. Equilibrio e armonia sono intesi ovviamente in termini relativi: la graduatoria dello sviluppo urbano sostenibile è costruita ponendo a confronto le città tra di loro, non in termini assoluti.

Dopo le prime tre, scorrendo la classifica dell'anno 2001 nelle posizioni di testa si trovano città di medio-piccole dimensioni del nord e città medio-grandi del centro Italia. Città del nord Italia e grandi città del centro (Firenze e Siena, prime tra le altre) sono i contesti urbani in cui lo sviluppo nei primi anni del 2000 è avvenuto in modo maggiormente sostenibile. Tra le prime si trovano anche città alpine (Trento e Sondrio), città della Val Padana (Reggio Emilia, Parma, Cremona, Mantova, Modena), città metropolitane (Bologna, Genova, Firenze). Si tratta di città che son riuscite a ottenere buone prestazioni economiche offrendo al contesto sociale buoni livelli di sicurezza, di servizi e strutture per la cultura, per la formazione, strutture di incontro e di interazione, tutelando meglio (rispetto alle altre città italiane) il proprio patrimonio naturale. Nelle città di medie dimensioni del nord e del centro le traiettorie di sviluppo e le progressioni temporali si svolgono in modo da combinare simultaneamente crescita economica, equità sociale, tutela ambientale, in modo migliore rispetto ad ambiti urbani di maggiori dimensioni e rispetto al Mezzogiorno nel suo insieme.

Come già evidenziato in tanti lavori, nell'anno 2001 (così come nei due anni precedenti) il dualismo tipicamente economico tra nord e sud Italia si ritrova anche nella variabile "società" e nella variabile "ambiente". "...Non soltanto l'economia italiana mostra un carattere dualistico, ma tale doppio binario connota tutti gli aspetti della vita quotidiana. La lettura dei risultati della presente indagine non fa che confermare quelli dell'anno passato: è come se si analizzassero i dati di due nazioni differenti" così si esprimeva Ornello Vitali nell'introdurre l'indagine sulla "Qualità della vita in Italia" del quotidiano Italia Oggi pubblicato nel mese di gennaio del 2001.

**Tab. 16 Le prime 10 città dello sviluppo sostenibile**

<b>Graduatoria</b>	<b>anno 2001</b>	<b>anno 2000</b>	<b>anno 1999</b>
1°	Bolzano	Bolzano	Milano
2°	Milano	Trieste	Trieste
3°	Trieste	Bologna	Bologna
4°	Firenze	Milano	Firenze
5°	Verona	Verona	Bolzano
6°	Bologna	Parma	Verona
7°	Trento	Firenze	Parma
8°	Padova	Siena	Bergamo
9°	Parma	Trento	Roma
10°	Bergamo	Brescia	Trento

In centro classifica si posizionano città medio-piccole del nord e del centro Italia molte con spiccata tradizione economica di stampo distrettuale, con i distretti industriali localizzati nel cuore urbano o in prossimità del sistema cittadino in cui le ricadute socioeconomiche si ripercuotono visibilmente sul contesto ambientale (ad esempio: Lecco (nel 2001 42a), Como (43a); Biella (49a); Prato (53a)).

In fondo alla graduatoria si posizionano città del sud grandi e piccole senza un ordinamento gerarchico chiaro come invece si nota nelle città del nord. Al sud le grandi città per il buon livello di valore aggiunto prodotto avrebbero potuto posizionarsi prima delle piccole città, ma questo non avviene perché il dato economico viene corretto con la dimensione sociale e ambientale. Criminalità e scarsa qualità ambientale a Palermo, Napoli, Reggio Calabria, fanno sì che queste siano precedute da realtà urbane con dinamiche socioeconomiche più equilibrate come Nuoro, Ragusa, Sassari.

Nel 2000 si conferma una graduatoria analoga, con leggere variazioni nel posizionamento relativo e semplici sostituzioni, per effetto delle quali le città del centro-nord ancora in buone posizioni e il meridione ancora penalizzato da attività culturali, dotazione di strutture sociali, attenzione all'ambiente inferiori rispetto alle città del centro-nord Italia.

Nel 1999 Milano occupa la prima posizione seguita da Trieste e Bologna, rispettivamente seconda e terza come nell'anno 2000. Roma è nona, Firenze quarta, entrambe prime tra le città dell'Italia centrale e ben posizionate nella graduatoria complessiva assieme alle migliori città del nord.

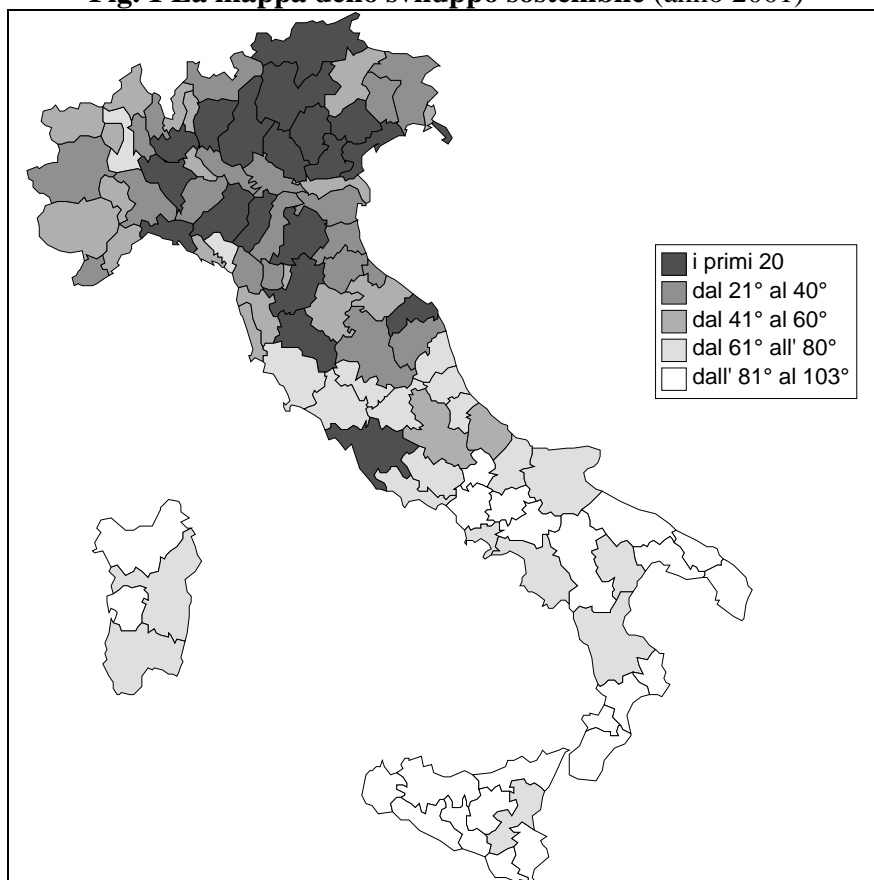
Nei tre anni città di medie dimensioni come Bolzano, Trento, Padova guadagnano posizioni in graduatoria, sia per i risultati conseguiti in ambito economico sia per le buone prestazioni di natura "sociale". Nel medesimo periodo la città di Bologna, come Bergamo e Parma perdono alcune posizioni, poche in realtà, si tratta sempre di variazioni minime. Comunque è come se la loro sostenibilità si contraesse un poco nel tempo. Ma l'aspetto più evidente è sicuramente dato dalla netta ripartizione tra nord e sud Italia. Nel corso dei tre anni analizzati la situazione sostanzialmente non cambia minimamente: il nord mantiene le sue posizioni di vetta e il sud occupa tutte le posizioni in fondo alla graduatoria.

**Tab. 17 Le prime tre città per ogni ripartizione geografica**

	anno 2001	anno 2000	anno 1999
<b>Le prime tre in Italia del Nord</b>			
1°	Bolzano	Bolzano	Milano
2°	Milano	Trieste	Trieste
3°	Trieste	Bologna	Bologna
<b>Le prime tre in Italia del Centro</b>			
1°	Firenze	Firenze	Firenze
2°	Roma	Siena	Roma
3°	Siena	Roma	Ancona
<b>Le prime tre in Italia del Sud</b>			
1°	L'Aquila	L'Aquila	L'Aquila
2°	Pescara	Pescara	Pescara
3°	Foggia	Cagliari	Salerno

Graduatorie stabili soprattutto per l'Italia centrale e meridionale sono determinate dall'indice di sviluppo urbano. Più soggetto a variazioni, anche se circoscritte, appare nell'Italia del nord. Probabilmente perché sensibile al dinamismo socioeconomico di queste aree territoriali.

**Fig. 1 La mappa dello sviluppo sostenibile (anno 2001)**



*Le grandi città e la sostenibilità*

Roma, Milano e Napoli sono le tre città italiane che all'inizio del 2000 hanno più di un milione di abitanti. Milano nella classifica dello sviluppo urbano sostenibile nell'anno 2001 risulta in seconda posizione sia rispetto a tutte le città italiane sia rispetto alle città del nord, Napoli è 79-esima in Italia e 11-esima nel Mezzogiorno, Roma 12-esima in Italia e seconda tra le città del centro.

**Tab. 18 La graduatoria delle città metropolitane**

(ex art. 22 del Testo Unico Enti Locali)

	<b>anno 2001</b>	<b>anno 2000</b>	<b>anno 1999</b>
Torino	25°	27°	16°
Milano	2°	4°	1°
Venezia	13°	14°	27°
Genova	14°	31°	31°
Bologna	6°	3°	3°
Firenze	4°	7°	4°
Roma	12°	13°	9°
Bari	87°	76°	75°
Napoli	79°	87°	76°

La dimensione urbana non determina il livello di sostenibilità dello sviluppo. Città di grandi dimensioni, riconosciute poli di aree metropolitane anche dalla legislazione nazionale, occupano posizioni di secondo piano, e indipendentemente anche dalla localizzazione geografica. Torino occupa

la 25-esima posizione nel 2001, piuttosto arretrata rispetto a Firenze e a Roma. Bari e Napoli sono lontane dalle prime in classifica, attardate da contesti socioeconomici importanti per lo sviluppo del Mezzogiorno, ma non confrontabili per gli aspetti strutturali e prestazionali con le realtà settentrionali. È quasi superfluo notare che in queste città la criminalità raggiunge dimensioni consistenti, tanto da porle in vetta alle graduatorie nazionali. Napoli, Roma, Bologna, Milano, Torino sono nell'ordine le peggiori in Italia per crimini e delitti commessi. Ma questo aspetto è solo uno degli elementi che costituiscono l'indice di sintesi, e il buono posizionamento relativo in altre dimensioni della sostenibilità dello sviluppo fa sì che le città metropolitane del nord Italia primeggino nella graduatoria nazionale.

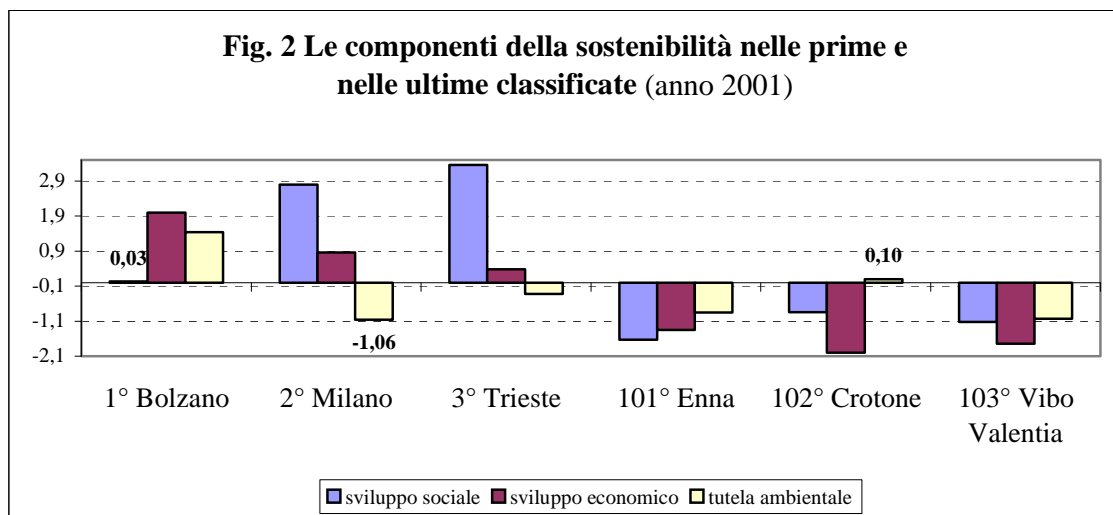
#### 4.2 Le dimensioni della sostenibilità urbana: economia, società, ambiente

Quali sono i fattori che determinano la graduatoria delle città italiane?

Il posizionamento competitivo di ciascuna città italiana rispetto alle altre nella graduatoria dell'indice di sviluppo urbano è determinato dalla combinazione di tre fattori che contemporaneamente caratterizzano ogni sistema territoriale: la dimensione economica, la dimensione sociale, e quella ambientale. Ognuno di questi fattori a sua volta è sintesi di variabili elementari rilevate dalle statistiche ufficiali in genere a livello di comune capoluogo di provincia (si veda il capitolo 3).

E così Bolzano prima in Italia si presenta superiore a Milano e a Trieste, rispettivamente seconda e terza, per crescita economica e per tutela ambientale, mentre in ambito sociale è inferiore a entrambe e pure a Firenze, prima delle città del centro Italia.

Il risultato economico di Bolzano è dato dall'ottimo posizionamento riportato contemporaneamente nelle due variabili elementari esplicative del fattore "economia": seconda dopo Milano per valore aggiunto pro capite e prima in Italia per minor tasso di disoccupazione.

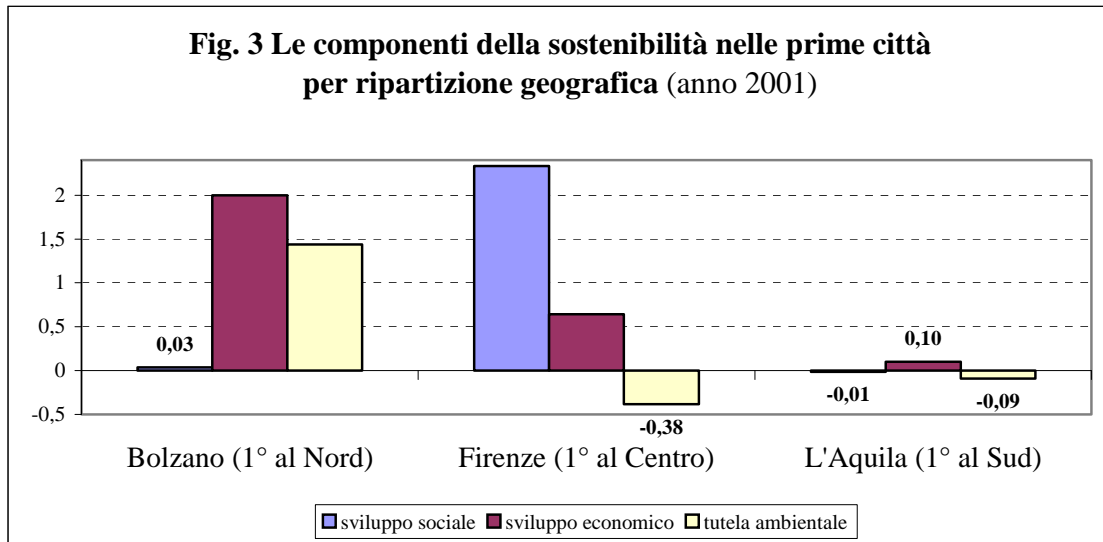


Milano e Trieste hanno una dimensione economica più compressa di Bolzano, addirittura la variabile ambientale negativa, ma in compenso la "società" molto forte. La combinazione di questi tre fattori con risultati così differenti le porta ad occupare le prime posizioni in graduatoria.

Milano nel 2001 per l'economia è seconda dopo Bolzano perché, pur essendo al primo posto per ammontare del valore aggiunto pro capite, occupa una posizione più arretrata in Italia per tasso di disoccupazione.

Le città ultime in classifica sono tutte del Mezzogiorno, in prevalenza appartengono alle regioni di Calabria e di Sicilia. In questo caso le tre componenti della sostenibilità registrano valori ampiamente negativi, senza eccezioni (salvo Crotona per l'ambiente).

L'Aquila prima delle città del Mezzogiorno presenta valori molto compressi se confrontati con quelli delle prime tre della graduatoria assoluta e con Firenze prima delle città dell'Italia centrale (Fig 3).



#### 4.3 Le relazioni tra l'indice di sviluppo urbano e le tre dimensioni della sostenibilità

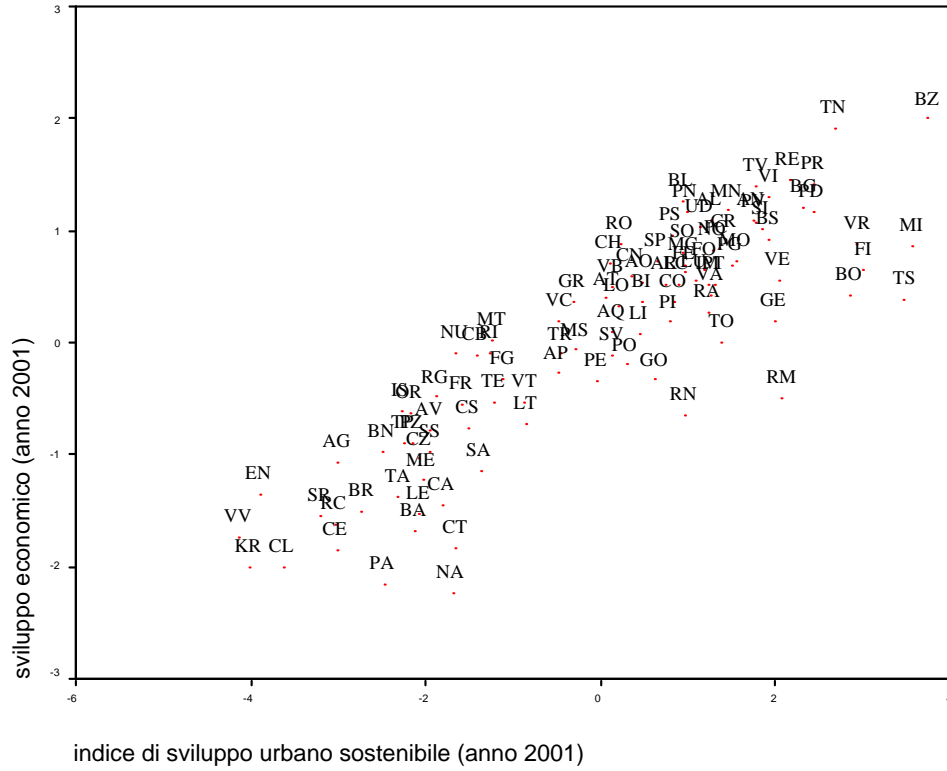
L'indice di sviluppo urbano sostenibile costruito per l'anno 2001 presenta una relazione diretta molto forte con la macro-variabile "economia" e con la macro-variabile "società", anche se in misura minore. L'indice di sviluppo urbano sostenibile assume valori via via maggiori con il crescere della variabile "sviluppo economico". Rappresentando in un grafico di dispersione le due macro-variabili si nota la posizione arretrata occupata dalle città del centro-sud Italia e la posizione avanzata delle città del nord Italia.

La variabile "sviluppo sociale" messa in relazione con l'indice di sviluppo urbano mostra un andamento positivo e crescente. Emergono con evidenza le grandi città, anche quelle con uno sviluppo urbano complessivo contenuto (tra cui Palermo, Catania, Napoli) si mettono in luce per l'alto sviluppo sociale.

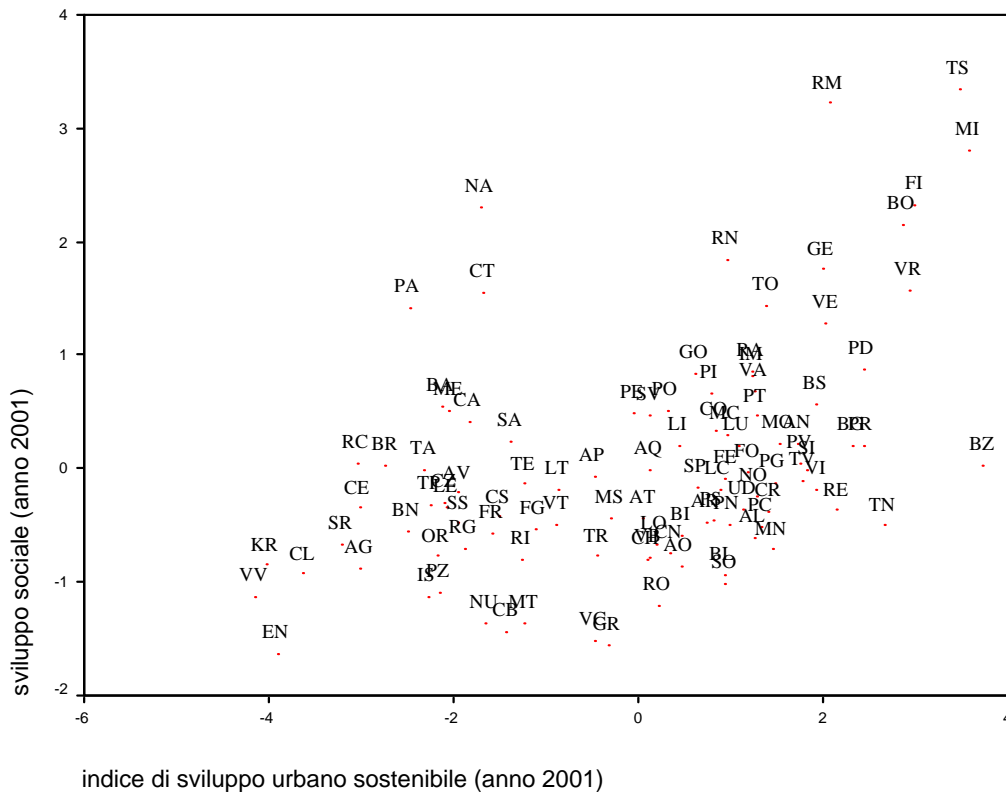
Il grafico a dispersione che mette in relazione l'indice di sviluppo urbano con la misura della tutela ambientale in apparenza sembra una nube di punti vaga, eccessivamente dispersa e priva di un qualche significato. A ben guardare però si può notare un andamento a "u rovesciata". Le città del nord con basso sviluppo urbano hanno anche bassa tutela ambientale. Salendo nel diagramma si nota che (come se si risalisse lungo la penisola, passando da Reggio Calabria a Napoli a Caserta a Viterbo ...) la tutela ambientale è maggiore e contemporaneamente anche lo sviluppo urbano appare maggiormente sostenibile nelle città medio-piccole del centro-nord Italia, per poi peggiorare nelle città medio-grandi (Parma, Padova, Trieste, Treviso, ..., Milano).



**Fig. 4 La relazione tra sviluppo economico e l'indice sintetico di sviluppo sostenibile**



**Fig. 5 La relazione tra sviluppo sociale e l'indice sintetico di sviluppo sostenibile**



#### 4.4 I confronti con le altre classifiche nazionali

In questo paragrafo si propone un confronto tra la graduatoria ottenuta con l'indice sintetico presentato nel paper con le graduatorie che nei medesimi periodi temporali sono state definite dai principali rapporti diffusi in Italia: "La qualità della vita del Sole24 Ore", quella di Italia Oggi, e l'"Ecosistema urbano" di Legambiente.

Innanzitutto si deve portare in evidenza che mentre il Sole24 Ore e Italia Oggi analizzano le province italiane utilizzando variabili che in qualche modo fanno riferimento ad aspetti economici-sociali ed ambientali dei territori assegnando a queste misure ruoli e rilevanza diverse, Legambiente analizza esclusivamente aspetti di carattere ambientale e solo nei comuni capoluogo di provincia.

Prendendo le mosse dalle graduatorie riprodotte nella tavola seguente (Tab. 19), si nota che Italia Oggi e Il Sole hanno un unico punto di contatto: Bolzano prima nel 2000 e seconda nel 2001 per entrambi i rapporti. Tutto il resto è diverso nelle loro rispettive graduatorie. Ed inevitabilmente è così se ci si limita a considerare le province una ad una, perché diversi sono i metodi statistici messi in campo e diverse sono le variabili utilizzate. Tanto che si può sostenere che la coincidenza di Bolzano sia puramente casuale.

Eppure indipendentemente dal metodo e dalle variabili, emerge nell'una e nell'altra classifica la prevalenza di alcuni territori, se considerati nel loro complesso come macro-aree: tra le prime province risultano quelle dell'Arco Alpino (Bolzano, Trento, Sondrio, Belluno), parte della Toscana (con Firenze e Siena), tratti della Valle del Po (con Mantova; Reggio Emilia, Parma (solo però per Italia Oggi). Un limite evidente nell'indagine del Sole è l'elevata variabilità delle posizioni occupate dalle città di anno in anno. Altro elemento caratterizzante, anche se non evidente dall'analisi della tavola sotto riprodotta (Tab. 19), è la variabilità altrettanto elevata sia nella numerosità del gruppo di variabili utilizzate che nella loro tipologia. Nel Sole24 Ore Sondrio, Bolzano, Trento, Siena, Belluno, Firenze sono le sole città che permangono nei due anni considerati. In Italia Oggi Trento, Bolzano, Reggio Emilia si scambiano le prime tre posizioni da un anno all'altro. Tra le prime 10 poi nei due anni permangono Mantova, Siena, Parma. Italia Oggi ha il pregio di utilizzare nel limite del possibile (a seconda della disponibilità dei dati forniti dalle fonti, ISTAT in primis) sempre le medesime variabili, ed utilizza tecniche statistiche solo in parte dichiarate, ma avanzate e sicuramente migliori rispetto a quelle del Sole (come già posto in evidenza dalla confutazione che Merlini e Vitali svilupparono in un ampio articolo del 1999).

**Tab. 19 Le prime 10 classificate nel 2001 e nel 2000 nelle principali graduatorie a diffusione nazionale**

Graduatoria	Il Sole24 Ore		Italia Oggi		Legambiente	
	anno 2001	anno 2000	anno 2001	anno 2000	anno 2001	anno 2000
1°	Sondrio	Bolzano	Trento	Bolzano	Cremona	Ferrara
2°	Bolzano	Sondrio	Bolzano	Reggio Emilia	Mantova	Mantova
3°	Trento	Trieste	Reggio Emilia	Trento	Bergamo	Massa
4°	Aosta	Bologna	Mantova	Modena	Sondrio	Pavia
5°	Siena	Belluno	Biella	Mantova	Pavia	Sondrio
6°	Mantova	Gorizia	Aosta	Belluno	Belluno	Prato
7°	Belluno	Rimini	Siena	Parma	Bolzano	Modena
8°	Firenze	Trento	Parma	Siena	Ferrara	Bolzano
9°	Milano	Firenze	Firenze	Cuneo	Arezzo	Livorno
10°	Forlì	Siena	Sondrio	Ancona	Biella	Bergamo

La graduatoria di Legambiente dal canto suo ha la caratteristica di premiare nel tempo per lo più città del nord Italia, e in particolare quelle di dimensioni medio-piccole. Le grandi aree metropolitane gravate da forti diseconomie di agglomerazione e costi sociali di natura ambientale non figurano

almeno per i due anni 2000 e 2001 nelle prime dieci classificate. Legambiente analizza solo territori comunali, quindi per questo si può ritenere più vicina all'indicatore proposto, si basa sempre sulle medesime variabili, raramente cambiano di anno in anno. Nel caso una variabile non venga più inserita, si tratta di una scelta strutturale, che porta gli autori a non riutilizzarla negli anni successivi (ad esempio la concentrazione di monossido di carbonio usata inizialmente negli anni più recenti non è più stato preso in considerazione).

Per completare il confronto si passa ad analizzare la graduatoria prodotta dall'indicatore di sintesi proposto nel paper. Colloca Bolzano in prima posizione, seguono Trieste, Milano, Firenze, Bologna, Verona, Parma, Trento che permangono nelle prime 10 posizioni per i due anni consecutivi 2001-2000 (gli stessi del confronto con Italia Oggi e Sole). Quindi a differenza delle classifiche nazionali le graduatorie definite con questo nuovo indicatore si rivelano maggiormente stabili nel posizionamento delle città al passare degli anni. In termini metodologici la stabilità è in parte garantita dalla procedura impiegata per aggregare i valori, in parte dalla qualità dei dati in fase di input (le variabili elementari prima di essere inserite nelle elaborazioni sono trasformate se hanno distribuzioni molto difformi dalla gaussiana), in parte dal ricorso costante nel tempo alle variabili utilizzate, per quanto possibile sempre le stesse.

In secondo luogo si conferma la permanenza nel tempo di alcune macro-regioni: parte dell'Arco Alpino (con Bolzano e Trento), parte della Toscana (con Firenze e Siena), tratti della Valle del Po (con Parma, Bologna e Verona). In questo nuovo indicatore prevalgono in graduatoria ambiti territoriali notoriamente caratterizzati da ricchezza economica, dinamicità imprenditoriale e apertura verso l'esterno (Firenze, Milano, Bologna, Parma, Padova, Bergamo, Brescia). La causa va ricercata nelle variabili utilizzate nella costruzione della sintesi, e va ricercata anche nel peso che la dimensione economica riveste nell'analisi discriminante (superiore solo in parte al peso della società e invece in modo rilevante rispetto all'ambiente).

## 5. Conclusioni

Il paper propone un approccio quantitativo all'analisi dello sviluppo urbano sostenibile. I dati presentati sono una misura della sostenibilità economica-sociale-ambientale nei comuni capoluogo di provincia italiani per il triennio 1999-2001.

Come già portato ad evidenza in numerosi lavori, nell'anno 2001 (e nei due anni precedenti) il dualismo tipicamente economico tra nord e sud Italia si ritrova anche nella variabile "società" e nella variabile "ambiente". Riprendendo le parole di Ornello Vitali in parte già citate, "...è come se si analizzassero i dati di due nazioni differenti".

Città metropolitane del centro-nord occupano le prime posizioni (Torino 25-esima è la più lontana dalla vetta della graduatoria, preceduta da Milano (2a), Firenze (4a), Bologna (6a), Roma (12a)). Città metropolitane del sud occupano posizioni anche molto arretrate in graduatoria: Palermo è 93-esima, Bari 87-esima, precedute da Catania (78a) e da Napoli (79a). Città in contesti rurali del nord e del centro occupano posizioni di media classifica e condividono le loro posizioni con città di medie dimensioni inserite in contesti industriali e produttivi differenziati. Mentre città in aree rurali del sud e città intermedie precedono spesso ambiti metropolitani.

L'immagine della sostenibilità economico-sociale-ambientale urbana si presenta stabile nel medio periodo. Le città mantengono nel periodo considerato le posizioni in graduatoria, subendo mediamente contenuti spostamenti. Le posizioni di testa si confermano nel tempo, mostrando solo lievi mutamenti con limitati scambi di posizione.

Se si focalizza l'attenzione sul sud Italia emerge la debolezza dei contesti urbani: Foggia (69a nel 2001) è la prima città del Mezzogiorno, se si escludono L'Aquila (57-esima) e le altre città abruzzesi di Chieti (59a) e Pescara (61a) (non a caso la regione Abruzzo è stata esclusa dall'obiettivo 1 della Programmazione Europea del periodo 2000-2006). Il sud, per quanto svantaggiato, mostra sue

peculiarità e specificità. Città sarde (Nuoro 77a, Cagliari 80a, Sassari 82a) si uniscono a realtà campane (Salerno 73a, Napoli 79a, Avellino 83a) e siciliane (Ragusa 81a, Catania 78a). Ma le ultime sei nella graduatoria del 2001 sono città della Sicilia e della Calabria. Reggio Calabria (98a) precede Siracusa (non aiutata dall'export di idrocarburi). Caltanissetta e Enna, seguite da Crotona e Vibo Valentia, staccate sensibilmente da Reggio Calabria nel valore dell'indice di sintesi.

Le grandi città meridionali non riescono ad emergere dal resto del sud, rimangono sommerse nelle difficoltà economiche-sociali-ambientali del territorio. Non così le grandi città del nord che in alcuni casi (Milano, Firenze, Bologna) trainano lo sviluppo integrato dei loro contesti.

Passando ad analizzare aspetti di metodo, l'indicatore proposto ha il pregio di essere fondato sull'applicazione di tecniche di statistica multivariata note e ampiamente usate. Originale è invece la sequenza con cui le tecniche sono organizzate e il modo in cui il risultato di ognuna viene utilizzato come input della tecnica posta immediatamente di seguito. Infatti l'indicatore è stato ideato in modo da integrare le tecniche in modo sequenziale: dapprima l'ACP che seleziona in termini statistici le variabili, quindi l'analisi cluster dei comuni usando le componenti principali prodotte nel passaggio precedente, quindi l'analisi discriminante che definisce una funzione di sintesi con cui si assegnano valori numerici agli elementi dei due cluster. Il nuovo indicatore però presenta allo stato attuale un limite rilevante che consiste nelle variabili economiche che sono state inserite negli algoritmi per costruire la sintesi. Il limite sarebbe facilmente superabile se si potessero individuare in ambito comunale variabili economiche per la misura di prestazioni e dinamicità imprenditoriali, aspetti che ha senso rilevare con frequenza annuale e per costruire indici aggiornabili anno per anno. In realtà dati economici a livello comunale disponibili in Italia sono prevalentemente di natura strutturale (il numero delle imprese, il numero degli addetti, per di più disponibili con cadenza annuale solo dall'anno 2005 grazie alla creazione dell'archivio ASIA). Per descrivere aspetti prestazionali e quindi di natura dinamica relativi alla dimensione economica, gli unici dati effettivamente utilizzabili con frequenza annuale sono il valore aggiunto e il tasso di disoccupazione. Stimati però per sistema locale del lavoro (non per comune). Stime prodotte con la tecnica delle "piccole aree" (si veda al riguardo il testo di riferimento di Rao, 2003; e il numero monografico della Rivista internazionale di scienze sociali dell'Università Cattolica del dicembre 2009).

A questo punto si può compiere un confronto metodologico tra l'indicatore proposto e gli altri tre indicatori nazionali già considerati in alcuni capitoli precedenti.

L'indicatore del Sole si ottiene dalla aggregazione di 36 variabili elementari suddivise in 6 sotto-indicatori di sintesi, relativi a specifici ambiti tematici (tenore di vita, affari e lavoro, servizi e ambiente, ordine pubblico, popolazione, tempo libero). Mentre i sei ambiti tematici permangono nel tempo, le 36 variabili sono estremamente mutabili, tanto che in una analisi temporale di una decina di rapporti consecutivi si è notato che solo il valore aggiunto pro capite e il tasso di disoccupazione sono le uniche due variabili delle 36 utilizzate che anno per anno figurano stabilmente nell'analisi (LEL, 2005). Il limite peggiore del Sole però risiede nel fatto che per sintetizzare le variabili, vengono tutte trasformate in un valore compreso tra zero e 1.000 e poi i valori così trasformati vengono sommati insieme. Ciò fa sì che tutte le variabili elementari entrino nella sintesi con il medesimo peso, il che significa che tutte hanno la medesima rilevanza in termini di qualità della vita. In più la somma delle variabili trasformate impedisce di tener conto delle correlazioni che possono esserci tra le variabili elementari utilizzate. Se due variabili sono legate da relazioni di linearità significa che nel sintetizzarle insieme le unità statistiche prime in classifica si consolidano nella loro posizione e unità statistiche ultime in classifica risultano essere sempre più penalizzate con il procedere dell'aggregazione. Per critiche dettagliate si rimanda a Merlini-Vitali (1999), già citato in precedenza, un articolo con il quale i due autori mostrando tutti i limiti metodologici degli indici di sintesi costruiti con la logica dell'indice proposto dal Sole24 Ore, spianavano la strada ad una loro proposta metodologica diversa che poco dopo si è concretizzata nell'indice di qualità della vita del quotidiano Italia Oggi. Il metodo adottato da

Italia Oggi non è stato completamente reso manifesto. Si articola in indicatori intermedi (popolazione, criminalità, disagio sociale, ambiente, affari e lavoro, servizi, tempo libero, tenore di vita), che di anno in anno rimangono invariati e rappresentano la chiave di lettura che gli autori propongono per l'analisi di ogni provincia italiana. Come a dire che ogni territorio, per quanto diverso dagli altri e pur presentando inevitabilmente proprie peculiarità e specificità può essere letto allo stesso modo dal nord al sud Italia.

Le variabili sono aggregate in indicatori intermedi con una tecnica di statistica multivariata che produce scores che a loro volta si addensano nell'indice di sintesi finale. Si esplicita quali sono le variabili elementari utilizzate legate negativamente con il concetto espresso dall'indicatore di sintesi, e vengono preliminarmente unite tra di loro per ciascuna area tematica, prima ancora di definire gli indici intermedi. Mentre Sole24 Ore e Italia Oggi utilizzano esclusivamente metodi statistici (solo negli ultimi anni il Sole ha aggiunto alle variabili oggettive anche una variabile soggettiva, costruita attraverso interviste a persone residenti in ciascuna provincia, espressione del cosiddetto "sentiment"), l'indice di sintesi di Legambiente è ottenuto pesando le variabili elementari con parametri definiti da esperti di settore. Questo fa sì che una data variabile entri in modo più o meno incisivo nella sintesi in base a quanto alcuni testimoni privilegiati pensano su quel dato tema. In base a quanto ritengano che un dato fenomeno possa essere influenzato da dati valori. Un limite di questo indice risiede nel fatto che non vengono condotte analisi di correlazione tra le variabili, quindi può accadere che nella sintesi si duplichi l'informazione, se si utilizzino contemporaneamente due variabili correlate tra di loro. In più si aggregano insieme variabili chiaramente contrapposte nella loro logica operativa. Ad esempio lo stato dell'ambiente misurato dalla quantità di verde fruibile e dalla concentrazione di inquinanti è determinato dalle attività di pressione causate dall'utilizzo di automobili, dal consumo di combustibili e in più lo stato può essere migliorato e la pressione può essere combattuta con politiche pubbliche locali (es. la realizzazione di piste ciclabili, l'apertura di zone a traffico limitato, ...). Legambiente nel proprio indice di sintesi tiene conto di tutto questo, ma somma i valori, nonostante possano essere spesso in relazione anche antitetica. Un altro limite di Legambiente consiste nell'assegnare i valori peggiori nelle variabili elementari a città per il semplice fatto che non forniscono la risposta su quel determinato dato. Questo sfavorisce pesantemente il sud Italia, che per Legambiente ha un ambiente di bassa qualità per lo più proprio a causa della mancata risposta ai questionari da parte degli operatori pubblici locali.

Un passo ulteriore da compiere in questo ambito di ricerca (relativo alla costruzione di indicatori di sintesi e più in generale relativo a metodi di descrizione dei sistemi locali territoriali) potrebbe consistere nella definizione di metodi per l'individuazione di relazioni tra le variabili e il ruolo degli enti locali nello sviluppo territoriale integrato (eventualmente seguendo anche quanto contenuto nel Manuale OECD del 2008 sulla costruzione degli indicatori compositi).

In questo caso si potrebbe anche andrebbe alla ricerca di effetti che politiche attuate dagli enti locali possono avere sullo sistema socioeconomico. Quindi si potrebbero studiare strategie per la definizione di politiche economiche locali che tengano in considerazione le variabili selezionate dalle tecniche statistiche, l'indice sintetico di sostenibilità, eventuali relazioni econometriche tra le variabili e al contempo riescano a coinvolgere non solo i policy maker locali ma anche gli attori pubblici, privati e la collettività locale.

## 6. Bibliografia

- Ambrosanio M., Bordignon M., Dallara A. (2003), *Enti locali e sviluppo sostenibile*, Dexia-Crediop, Roma
- Anderson R. J. (1970), *A note on economic base studies and regional econometric forecasting models*, in *Journal of Regional Science*, vol.10, pp.325-333
- Andrews F. (1990), *Evolution of a movement*, *Journal of Public Policy*, n.9, pp. 401-405
- Barbaranelli C. (2003), *Analisi dei dati*, Edizioni Led, Milano
- Bertuglia C.S., La Bella A. (1991), *I sistemi urbani*, vol. I, Franco Angeli, Milano
- Bohrstedt G. W., Knoke D. (1994), *Statistica per le scienze sociali*, Il Mulino, Bologna
- Bossel H. (1998), *Earth at a crossroads: paths to a sustainable future*, Cambridge University Press, Cambridge
- Camagni R., Capello R., Nijkamp P. (2001), *Managing sustainable urban environments*, in Paddison, R. (a cura di), *Handbook of urban studies*, Sage, Londra
- Camagni R. (1996), *Economia e pianificazione della città sostenibile*, Il Mulino, Bologna
- Cattell R. B. (1966), *The scree test for the number of factors*, in *Multivariate Behavioral Research*, vol.1, pp.245-276
- Ciciotti E., Dallara A., Rizzi P. (2008), *Una geografia della sostenibilità dei sistemi locali italiani*, in Bellini N., Calafati A. (a cura di), "Internazionalizzazione e sviluppo regionale", Franco Angeli, Milano
- Ciciotti E., Rizzi P. (a cura di) (2005), *Politiche per lo sviluppo territoriale*, Carocci, Roma
- Ciciotti E. (1997), *Sviluppo locale e nuovi approcci alle politiche regionali*, in AA.VV., *Mercato, Stato e Giustizia Sociale*, Giuffrè, Milano
- Comrey A. L., Lee H. B. (1992), *A first course in factor analysis*, Lawrence Erlbaum Associates, Hillsdale (NJ)
- Crouch C., Le Gales P., Trigilia C., Voelzkow H. (2001), *Local production systems in Europe*, Oxford University Press, Oxford
- Dallara A. (2008), *Un metodo per la descrizione quantitativa dei sistemi locali*, in Bellini N., Calafati A. (a cura di), "Internazionalizzazione e sviluppo regionale", Franco Angeli, Milano
- Dallara A. (2006), *La descrizione quantitativa dei sistemi locali italiani*, in *Rivista Italiana di Economia, Demografia e Statistica*, vol.LX n.3-4, luglio-dicembre, p.255-263
- Dallara A., Rizzi P. (2005), *Le politiche di sviluppo e i sistemi territoriali*, in Ciciotti E., Rizzi P. (a cura di), *Politiche per lo sviluppo territoriale*, Carocci, Roma
- Dallara A. (2004), *Gli indicatori di sostenibilità e sviluppo*, LEL, Università Cattolica di Piacenza, Quaderni LEL n.70
- Dallara A., Rizzi P. (2004), *Il Patto per Piacenza: un processo per la promozione concentrata dello sviluppo locale*, in V. Fedeli e F. Gastaldi (a cura di), *Pratiche strategiche di pianificazione*, Franco Angeli, Milano
- Dallara A., Rizzi P. (2004), *Metodologie e Indicatori per la rilevazione della aree di crisi in Lombardia*, IRER, Università Cattolica di Piacenza, Quaderni LEL n.66
- Dasgupta P. (2004), *Benessere umano e ambiente naturale*, Vita & Pensiero, Milano
- Fusco Girard L., Nijkamp P. (a cura di) (2005), *Energia, bellezza, partecipazione: la sfida della sostenibilità. Valutazioni integrate tra conservazione e sviluppo*, Franco Angeli, Milano
- Fusco Girard L., Nijkamp P. (a cura di) (1997), *Le valutazioni per lo sviluppo sostenibile della città e del territorio*, Franco Angeli, Milano
- Ganugi P. (a cura di) (1999), *Ricerche quantitative per la politica economica nei distretti industriali*, FrancoAngeli, Milano
- Garofoli G. (1992), *Economia del territorio*, Etas, Milano
- Hardi P., Zdan T. (1997), *Assessing sustainable development: principles in practice*, IISD, Winnipeg
- Hardle W., Simar L. (2003), *Applied multivariate statistical analysis*, Springer-Verlag, Berlin Heidelberg
- Hatcher L. (1994), *A step by step approach to using Sas<sup>®</sup> for factor analysis and structural equation modelling*, Sas Institute Inc. ,Cary (NC)

- Johnson J. E. (2003), *A user's guide to principal components*, Wiley & Sons, Inc., Hoboken (NJ)
- Jolliffe I.T. (2002), *Principal Component Analysis*, second edition, Springer, New York
- Jolliffe I.T. (1986), *Principal Component Analysis*, Springer, New York
- Kaiser H. F. (1974), *An index of factorial simplicity*, Psychometrika, vol. 39, pp. 31-36
- Kaiser H. F. (1970), *A second generation Little Jiffy*, Psychometrika, vol. 35, pp. 401-415
- Kaplan D. E. (2003), Measuring our competitiveness - a critical examination of the IMD and WEF competitiveness indicators for South Africa, in *Development Southern Africa*, vol. 20
- Kim J.O., Mueller C. W. (1978), *Factor analysis: statistical methods and practical issues*, Sage, Beverly Hills (CA)
- Krzanowski W.J. (1988), *Principles of Multivariate Analysis: A User's Perspective*, Oxford University Press, Oxford
- Krzanowski W.J. (1987), Cross-validation in principal component analysis, *Biometrics*, vol. 43, pp.575-584
- Lazarsfeld P. F. (1969), Dai concetti agli indici empirici, in Boudon R., Lazarsfeld P.F., *L'analisi empirica nelle scienze sociali*, vol.1, Il Mulino, Bologna
- LEL (2005), Il posizionamento competitivo di Piacenza, Laboratorio di Economia Locale, Università Cattolica di Piacenza, Quaderni LEL n.91
- LEL (1999), Analisi delle capacità competitività dei sistemi economici locali, Laboratorio di Economia Locale, Università Cattolica di Piacenza, Quaderni LEL n.19
- Lowry I.S. (1964), *Modello di una metropoli*, Guida editori, Napoli
- Meadows D. (1998), *Indicators and information systems for sustainable development*, The Sustainability Institute, Hartland Four Corners
- Merlini A., Vitali O. (1999), La qualità della vita: metodi e verifiche, *Rivista Italiana di Economia Demografia e Statistica*, vol. LIII, n.2, pp.5-93
- OECD, *Handbook on Constructing Composite Indicators: Methodology and User Guide*, European Commission, Paris, 2008
- Rey G. (2004), Politica economica e intervento pubblico a livello locale, in *Scienze Regionali*, vol. 3, n.3
- Sadocchi S. (1987), *Manuale di analisi statistica multivariata*, Franco Angeli, Milano
- Sforzi F. (1999), I confini territoriali della regolazione istituzionale, in Arrighetti-Serravalli 1999
- Sheldon E.B., Moore W.E. (1968), *Indicators of social change: concepts and measurements*, Russel Sage Foundation, New York
- Signorini L.F. (2001) (a cura di), *Lo sviluppo locale. Un'indagine della Banca d'Italia sui distretti industriali*, Donzelli, Roma
- Stevens J. (1986), *Applied multivariate statistics for the social sciences*, Lawrence Erlbaum Associates, Hillsdale (NJ)
- Stiglitz J.E. (2002), *La globalizzazione e i suoi oppositori*, Einaudi, Torino
- Stiglitz-Sen-Fitoussi Commission, *Report by the Commission on the Measurement of Economic Performance and Social Progress*, Paris, 2009
- Tabachnick B. G., Fidell L. S. (2001), *Using multivariate statistics*, Pearson Allyn and Bacon, Needham Heights (MA)
- Triglia C. (2005), *Sviluppo locale. Un manifesto per l'Italia*, Editori Laterza, Bari
- Viesti G. (2000), *Politiche economiche e sviluppo locale: alcune riflessioni*, in *Sviluppo locale*, vol. 7, n.14
- Zajczyk F. (1997) *Il mondo degli indicatori sociali*, Carocci, Roma
- Zani S. (2000), *Analisi dei dati statistici. Osservazioni multidimensionali*, Giuffrè, Milano
- Zani S., (a cura di) (1996), *Misure della qualità della vita*, FrancoAngeli, Milano
- Zani S. (1994), *Analisi dei dati statistici. Osservazioni in una e due dimensioni*, Giuffrè, Milano

## **Appendice 1 - Tabelle**



**Tab. 20 Matrice di correlazione delle variabili**

		Conc. PM10	biossido di azoto	Conc. nitrati	politiche per l'ambiente (eco management)	Industrie certificate Iso 14001	Trasporto pubblico	Capacità depurazione acqua	Verde urbano fruibile	Piste ciclabili	Zone a traffico limitato	Isole pedonali	Raccolta differenziata	Produzione rifiuti urbani
Concentrazione PM 10	Pearson	1,00	0,23	-0,02	<b>-0,22</b>	0,11	-0,07	-0,28	-0,23	-0,20	-0,24	-0,04	-0,29	-0,31
	Sig.		0,02	0,87	0,03	0,28	0,47	0,00	0,02	0,04	0,01	0,68	0,00	0,00
Biossido di azoto	Pearson	0,23	1,00	0,11	-0,16	-0,04	-0,12	-0,10	-0,24	-0,12	-0,12	-0,03	-0,18	-0,18
	Sig.	0,02		0,26	0,12	0,67	0,22	0,34	0,01	0,24	0,23	0,78	0,07	0,06
Concentrazione nitrati	Pearson	-0,02	0,11	1,00	0,00	-0,20	0,03	-0,19	0,01	0,08	-0,05	0,08	0,07	0,21
	Sig. (2-tailed)	0,87	0,26		0,97	0,05	0,73	0,06	0,94	0,45	0,64	0,41	0,50	0,04
politiche per l'ambiente (eco management)	Pearson	-0,22	-0,16	0,00	1,00	0,01	0,17	<b>0,25</b>	0,41	0,41	0,30	0,16	0,35	0,28
	Sig. (2-tailed)	0,03	0,12	0,97		0,92	0,08	0,01	0,00	0,00	0,00	0,12	0,00	0,00
Industrie certificate Iso 14001	Pearson	0,11	-0,04	-0,20	0,01	1,00	0,07	0,22	-0,10	0,04	-0,06	-0,03	0,06	-0,09
	Sig. (2-tailed)	0,28	0,67	0,05	0,92		0,48	0,02	0,30	0,65	0,55	0,75	0,57	0,34
Trasporto pubblico	Pearson	-0,07	-0,12	0,03	0,17	0,07	1,00	0,01	0,22	0,06	0,30	0,12	0,22	0,28
	Sig. (2-tailed)	0,47	0,22	0,73	0,08	0,48		0,91	0,03	0,56	0,00	0,23	0,02	0,00
Capacità depuraz. acqua	Pearson	-0,28	-0,10	-0,19	0,25	0,22	0,01	1,00	0,19	0,21	0,16	0,13	0,22	0,09
	Sig. (2-tailed)	0,00	0,34	0,06	0,01	0,02	0,91		0,05	0,03	0,11	0,20	0,03	0,39
Verde urbano fruibile	Pearson	-0,23	-0,24	0,01	<b>0,41</b>	-0,10	0,22	0,19	1,00	<b>0,60</b>	<b>0,43</b>	0,24	<b>0,55</b>	<b>0,39</b>
	Sig. (2-tailed)	0,02	0,01	0,94	0,00	0,30	0,03	0,05		0,00	0,00	0,02	0,00	0,00
Piste ciclabili	Pearson	-0,20	-0,12	0,08	<b>0,41</b>	0,04	0,06	0,21	0,60	1,00	0,32	0,34	<b>0,56</b>	<b>0,44</b>
	Sig. (2-tailed)	0,04	0,24	0,45	0,00	0,65	0,56	0,03	0,00		0,00	0,00	0,00	0,00
Zone a traffico limitato	Pearson	-0,24	-0,12	-0,05	<b>0,30</b>	-0,06	0,30	0,16	0,43	0,32	1,00	0,34	0,41	0,37
	Sig. (2-tailed)	0,01	0,23	0,64	0,00	0,55	0,00	0,11	0,00	0,00		0,00	0,00	0,00
Isole pedonali	Pearson	-0,04	-0,03	0,08	0,16	-0,03	0,12	0,13	0,24	0,34	0,34	1,00	0,34	0,33
	Sig. (2-tailed)	0,68	0,78	0,41	0,12	0,75	0,23	0,20	0,02	0,00	0,00		0,00	0,00
Raccolta differenziata	Pearson	-0,29	-0,18	0,07	<b>0,35</b>	0,06	0,22	0,22	0,55	0,56	0,41	0,34	1,00	0,39
	Sig. (2-tailed)	0,00	0,07	0,50	0,00	0,57	0,02	0,03	0,00	0,00	0,00	0,00		0,00
Produzione rifiuti urbani	Pearson	-0,31	-0,18	0,21	0,28	-0,09	0,28	0,09	<b>0,39</b>	<b>0,44</b>	<b>0,37</b>	0,33	0,39	1,00
	Sig. (2-tailed)	0,00	0,06	0,04	0,00	0,34	0,00	0,39	0,00	0,00	0,00	0,00	0,00	
	N	103	103	101	103	103	103	103	103	103	103	103	103	103

**Tab. 21 Elenco delle variabili disponibili a livello di comune capoluogo di provincia**  
(anni 2001, 2000, 1999)

<b>Gruppi di variabili</b>	<b>Variabile controllabile dal comune capoluogo</b>
<b>Variabili della dimensione “ambiente”</b> (Fonte: Legambiente) (dati disponibili a livello di comune capoluogo di provincia)	
- Indice di sintesi di Legambiente	
- Concentrazione PM 10	
- Concentrazione biossido di azoto	
- Concentrazione di nitrati	
- Politiche per l'ambiente (eco management)	sì
- Consumo idrico pro capite	
- Verde urbano fruibile (mq/ab.)	sì
- Superficie aree verdi (mq/ettari)	sì
- Piste ciclabili	sì
- Zone a traffico limitato	sì
- Isole pedonali	sì
- Capacità di depurazione acque reflue	sì
- Auto in circolazione	
- Produzione rifiuti urbani	
- Raccolta differenziata	sì
- Trasporto pubblico	sì
<b>Demografia</b> (fonte: ISTAT) (dati disponibili a livello di comune capoluogo di provincia)	
- Popolazione residente	
- Densità abitativa (nel SLL)	
- Indice di invecchiamento	
- Rapporto di mascolinità	
- Indice di vecchiaia	
- Indice di dipendenza dei giovani	
- Indice di dipendenza degli anziani	
- Indice di dipendenza totale	
- Indice di ricambio	
- Indice di ricambio della popolazione in età attiva	
- Indice di lavoro potenziale *	
- Donne in età feconda per 100 donne	
- Indice del carico di figli per donna feconda	
- Tasso di crescita totale	
- Tasso di crescita naturale	
- Tasso di natalità	
- Tasso di fecondità	
- Rapporto di mascolinità alla nascita	
- Tasso di mortalità - Maschi + Femmine	

\* \* (Popolazione residente in età 15-64 anni / Popolazione residente) \* 100

<b>Gruppi di variabili</b>	<b>Variabile controllabile dal comune capoluogo</b>
<b>Salute</b> (fonte: ISTAT) (dati disponibili a livello di comune capoluogo di provincia)	
- Tasso di mortalità per tumore	
- Tasso di mortalità per malattie del metabolismo	
- Tasso di mortalità per malattie cardiocircolatorie	
- Tasso di mortalità per malattie respiratorie	
- Tasso di mortalità per malattie dell'apparato digerente	
- Tasso di mortalità per cause esterne	
<b>Cultura</b> (fonte: ISTAT) (dati disponibili a livello di comune capoluogo di provincia)	
- Rappresentazioni teatrali e musicali per 100.000 abitanti	
- Biglietti venduti per rappresentazioni teatrali e musicali per 1.000 abitanti	
- Numero medio di biglietti venduti per rappresentazione teatrale o musicale	
- Spesa media per abitante per rappresentazioni teatrali e musicali	
<b>Criminalità</b> (fonte: ISTAT) (dati disponibili a livello di comune capoluogo di provincia)	
- Indice di delittuosità - Totale	
- Indice di delittuosità - Criminalità violenta	
- Indice di delittuosità - Furti	
- Indice di delittuosità - Altri delitti	
- Omicidi dolosi per 100.000 abitanti	
- Lesioni dolose per 100.000 abitanti	
- Estorsioni per 100.000 abitanti	
- Borseggi per 100.000 abitanti	
- Scippi per 100.000 abitanti	
- Furti in negozi e appartamenti per 100.000 abitanti	
- Furti in auto in sosta per 100.000 abitanti	
- Furti di autoveicoli per 100.000 abitanti	
- Truffe per 100.000 abitanti	
- Reati di contrabbando per 100.000 abitanti	
- Reati di produzione e spaccio di stupefacenti per 100.000 abitanti	
- Reati di sfruttamento e favoreggiamento della prostituzione per 100.000 abitanti	
- Protesti di assegni per 1.000 abitanti	
- Protesti di tratte accettate per 1.000 abitanti	
- Protesti di assegni emessi da individui per 1.000 abitanti	
- Ammontare medio delle tratte accettate protestate	
- Ammontare medio degli assegni protestati	
<b>Economie esterne</b> (dato a livello provinciale, fonte Tagliacarne) (dato determinato dalla dinamicità del comune capoluogo)	
- Indice di dotazione di strutture culturali e ricreative	
- Indice di dotazione di strutture per l'istruzione	
- Indice di dotazione di infrastrutture sociali	

## Appendice 2: introduzione teorica alle tecniche utilizzate nella costruzione dell'indicatore composito

### L'analisi in componenti principali: introduzione teorica

In termini generalizzati data una matrice  $\begin{matrix} X \\ (n,p) \end{matrix}$  di  $n$  unità statistiche e  $p$  variabili quantitative, definita  $S$  la matrice di covarianza corrispondente, l'analisi in componenti principali (ACP) consiste nella soluzione del problema di massimizzazione vincolata così espresso:

$$\max_a \text{Var}(\tilde{X}a) = a'Sa$$

$$\text{s.v. } a'a = 1$$

dove l'obiettivo consiste nel determinare combinazioni lineari delle  $p$  variabili con varianza massima.  $\tilde{X}$  rappresenta la matrice delle  $p$  variabili trasformate,  $a$  è un vettore di coefficienti definiti nell'insieme dei reali. Risolvere il problema di ottimo vincolato equivale a individuare il massimo del lagrangiano

$$\phi(a, \lambda) = a'Sa - \lambda(a'a - 1)$$

le condizioni necessarie di primo ordine sono

$$\begin{cases} \frac{\partial \phi}{\partial a} = 2Sa - 2\lambda a = 0 \\ \frac{\partial \phi}{\partial \lambda} = a'a - 1 = 0 \end{cases}$$

che danno luogo ad un sistema omogeneo di  $p$  equazioni in  $p$  incognite, che in forma matriciale è così rappresentabile  $(S - \lambda I)a = 0$ , dove  $a$  soddisfa il vincolo di normalizzazione  $a'a = 1$ .

Il sistema ammette soluzioni non tutte nulle se il suo determinante è uguale a zero, ossia se l'equazione caratteristica della matrice  $S$

$$|S - \lambda I| = 0$$

ammette soluzioni. L'equazione caratteristica è un polinomio di grado  $p$ , con  $p$  soluzioni dette autovalori. Per costruzione tutti gli autovalori sono non negativi. Ad ogni autovalore corrisponde un autovettore. Quindi il problema di ottimizzazione vincolata presentato coincide con quello di determinare gli autovalori e gli autovettori della matrice di covarianza  $S$ .

### L'analisi cluster: introduzione teorica

Obiettivo dell'analisi cluster è la creazione di una partizione di unità statistiche  $P(O)$  a partire dall'insieme originario delle unità statistiche  $O$ . Classifica le unità statistiche in gruppi con due caratteristiche:

- *coesione interna* (le unità assegnate ad un medesimo gruppo devono essere tra loro simili)

- *separazione esterna* (i gruppi devono essere il più possibile distinti).

Più formalmente (ipotizzando che la partizione sia in due soli gruppi) essa produce una raccolta di insiemi disgiunti e esaustivi (detti cluster) dell'insieme originario delle unità statistiche  $O$ :

$$P(O) = \{C_1, C_2\}$$

$$\text{ove } C_i \cap C_h = \emptyset \quad \forall i, h \in \{1, 2\}$$

tali che  $i \neq h$  e

$$\bigcup_{i=1}^2 C_i = O.$$

La partizione  $P(O)$  viene prodotta in modo tale che i profili delle unità statistiche che appartengono allo stesso cluster sono tra di loro molto simili, mentre i profili delle unità statistiche appartenenti a cluster differenti sono molto diversi tra di loro:

$r, s \in C \Leftrightarrow x^r \sim x^s \Leftrightarrow p_{rs}^d$  è minima

$r \in C_1, s \in C_2 \Leftrightarrow x^r \neq x^s \Leftrightarrow p_{rs}^d$  è massima

dove  $p_{rs}^d$  è una generica misura di “dissomiglianza”<sup>14</sup>.

### Clusterizzazione non gerarchica k-medie

Quando si conosce il numero di cluster in cui aggregare le unità statistiche la tecnica nota con il nome “Quick cluster” o “clusterizzazione non gerarchica k-medie” è efficiente nel riaggregare i casi in gruppi. È un algoritmo che produce una classificazione delle  $n$  unità statistiche in  $k$  gruppi distinti, con  $k$  fissato a priori. È un buon metodo quando le unità statistiche sono superiori a 100.

Si tratta di una tecnica basata su di una procedura iterativa, ciascuna iterazione è composta da tre passi, di seguito descritti.

#### Passo 1: selezione dei centri iniziali dei cluster

Si scelgono “i  $k$  centri iniziali”, che sono i  $k$  punti in uno spazio  $p$ -dimensionale (dove  $p$  sono le variabili usate) che rappresentano i centroidi dei  $k$  cluster iniziali:

a) se  $\min_i d(\mathbf{x}_k, \mathbf{M}_i) > \min_{i,j} d(\mathbf{M}_i, \mathbf{M}_j)$  e  $d(\mathbf{x}_k, \mathbf{M}_m) > d(\mathbf{x}_k, \mathbf{M}_n)$

allora la media del cluster  $\mathbf{M}_n$  viene sostituita dal vettore della  $k$ -esima osservazione  $\mathbf{x}_k$ .

se  $\min_i d(\mathbf{x}_k, \mathbf{M}_i) > \min_{i,j} d(\mathbf{M}_i, \mathbf{M}_j)$  e  $d(\mathbf{x}_k, \mathbf{M}_m) < d(\mathbf{x}_k, \mathbf{M}_n)$

allora la media del cluster  $\mathbf{M}_m$  viene sostituita dal vettore  $\mathbf{x}_k$ .

Ciò sta a significare che se la distanza tra il vettore della  $k$ -esima osservazione  $\mathbf{x}_k$  e la media del cluster a lui più vicino è più grande della distanza tra le due medie più vicine ( $\mathbf{M}_m$  e  $\mathbf{M}_n$ ), allora il vettore  $\mathbf{x}_k$  sostituisce o  $\mathbf{M}_m$  o  $\mathbf{M}_n$ , quella tra le due che è più vicina al vettore  $\mathbf{x}_k$ .

b) se  $\mathbf{x}_k$  non sostituisce la media di un cluster nel passaggio precedente a), allora si compie un secondo test:

- si ponga per ipotesi che  $\mathbf{M}_q$  sia la media del cluster più vicino a  $\mathbf{x}_k$

- si ponga per ipotesi che  $\mathbf{M}_p$  sia la media del secondo cluster più vicino a  $\mathbf{x}_k$

se  $d(\mathbf{x}_k, \mathbf{M}_p) > \min_i d(\mathbf{M}_q, \mathbf{M}_i)$

allora  $\mathbf{M}_q = \mathbf{x}_k$

questo sta a significare che se il vettore  $\mathbf{x}_k$  è più distante dal centro del secondo cluster più vicino piuttosto che dal centro del cluster più vicino, si mette  $\mathbf{x}_k$  al posto del centro del cluster più vicino.

#### Passo 2: aggiornamento dei centri iniziali dei cluster

Si calcola la distanza di ciascun elemento  $n$  da tutti i centroidi costruiti al passo precedente, e si assegna ciascun elemento al cluster con il centroide più vicino all'elemento. In altre parole, partendo dalla prima unità statistica, ogni unità statistica viene assegnata al cluster più vicino, e di conseguenza la media del cluster viene aggiornata. Si noti che il centro iniziale del cluster è inserito in questa media. Le medie aggiornate dei cluster sono i centri dei cluster di classificazione.

#### Passo 3: assegnazione delle unità statistiche al cluster più vicino

Se la distanza di una unità statistica dal centroide non è minima o se la distanza minima per una data unità non è in corrispondenza del centroide del gruppo a cui l'unità è stata assegnata, allora

---

<sup>14</sup> Nel caso della Quick Cluster di Mc Queen 1967 la misura di dissomiglianza utilizzata è la distanza euclidea.

questa unità si sposta dal cluster in cui era stata inserita inizialmente e viene inserita nel cluster il cui centroide le risulta essere il più vicino.

Quindi il terzo passo assegna ogni unità statistica al cluster più vicino, dove la distanza da un cluster è la distanza euclidea tra quel caso e i centri di classificazione aggiornati. Le medie dei cluster finali sono quindi calcolati come le medie delle variabili per i casi assegnati ad ogni cluster. Le medie dei cluster finali non contengono i centri di classificazione.

Si continuano le iterazioni finché il modello converge. La distanza calcolata è sempre in genere la distanza euclidea, perché garantisce la convergenza del modello. All'iterazione  $t$ -esima la distanza tra l'unità statistica  $i$  (con  $i = 1, 2, \dots, n$ ) e il centroide del gruppo  $l$  (con  $l = 1, 2, \dots, g$ ) è

$d(x_i, \overline{x}_l^{(t)}) = \sqrt{\sum_{s=1}^p (x_{is} - \overline{x}_{s,l}^{(t)})^2}$ , dove  $\overline{x}_l^{(t)}$  è il centroide del gruppo  $l$ . Quindi il metodo delle  $k$ -medie si

basa sul criterio di minimizzazione della devianza all'interno dei gruppi (Within,  $W$ ). Da qui un buon indice per verificare la bontà della soluzione ottenuta è  $R^2 = 1 - \frac{W}{T} = \frac{B}{T}$ . Il metodo  $k$ -medie produce cluster sferici e di dimensioni abbastanza simili tra di loro.

### L'analisi discriminante (l'approccio lineare di Fisher): introduzione teorica

Mediante l'analisi discriminante si vuole attribuire un ordinamento alle unità statistiche riportate nei due gruppi definiti con l'analisi cluster descritta nel paragrafo precedente.

Dato un set di variabili indipendenti riferite a  $n$  unità statistiche e dati i gruppi a cui appartengono le  $n$  unità statistiche, l'analisi discriminante definisce una relazione lineare tra queste variabili in modo da differenziare al meglio i due gruppi, e in modo che si rendano massime le differenze tra i gruppi. Le relazioni lineari tra le variabili indipendenti vengono chiamate "funzioni discriminanti" e si indicano con la seguente equazione:

$$y_i = a_1 X_{1i} + a_2 X_{2i} + a_3 X_{3i} + \dots + a_p X_{pi}$$

dove

- $y_i$  è il punteggio dell'unità statistica  $i$ -esima assegnato dalla funzione discriminante
- $X_{pi}$  è il valore della variabile indipendente  $p$ -esima per l'unità statistica  $i$ -esima
- $a_p$  è il coefficiente della funzione discriminante relativo alla  $p$ -esima variabile indipendente.

Il coefficienti  $a_p$  vengono stimati in modo che la differenza tra i gruppi definiti dalla variabile  $Y$  sia massima. Ciò significa che si individua una relazione lineare che renda massima la devianza tra i gruppi e minima la devianza entro i gruppi rispetto alle variabili indipendenti.

In termini formali significa che (una volta assunto che la devianza della  $Y$  si può scomporre nella somma della devianza all'interno dei gruppi (within) con la devianza tra i gruppi (between), formalizzato in questa espressione:  $a^T T a = a^T W a + a^T B a$ ) il problema dell'analisi discriminante

consiste nel trovare il vettore  $a$  che massimizza il rapporto seguente  $\frac{a^T B a}{a^T W a}$  (detto anche rapporto di

correlazione della variabile  $Y$ ), ossia:  $\max_a \frac{a^T B a}{a^T W a}$ , il che significa annullare il gradiente calcolato

rispetto ad  $a$ : 
$$\frac{\partial}{\partial a} \left( \frac{a^T B a}{a^T W a} \right) = 0$$

che con passaggi successivi porta a risolvere la seguente espressione  $Ba + \lambda Wa = 0$ , e poi  $W^{-1}B = \lambda a$ ,

dove  $a$  è l'autovettore della matrice  $W^{-1}B$  associato all'autovalore  $\lambda$  più elevato.

Per costruzione si ha che il rango della matrice  $W^{-1}B$  è pari al minore tra il numero delle variabili e il numero dei gruppi di ripartizione delle unità statistiche meno uno:  $\text{rango}(W^{-1}B) = \min(p, g - 1)$ .

Quindi il primo autovalore della matrice  $W^{-1}B$  fornisce la massima separazione tra i “g” gruppi, e gli elementi dell’autovettore associato ad esso consentono di ricavare i coefficienti della prima funzione discriminante, che quindi è quella che differenzia meglio i gruppi. Infatti i punteggi di questa combinazione lineare determinano la migliore separazione possibile tra i gruppi rispetto a qualsiasi altro punteggio ricavabile da altre combinazioni lineari. Come per l’ACP si assume che gli autovettori siano di lunghezza unitaria e che siano ortogonali.

Le ipotesi poste alla base dell’analisi discriminante sono le seguenti:

- le variabili indipendenti presentano coefficienti di correlazione bivariata le une con le altre non alti
- la media e la varianza di ogni variabile indipendente non sono correlate
- la correlazione tra due variabili indipendenti si mantiene costante nei gruppi in cui si ripartono le unità statistiche
- i valori delle variabili indipendenti sono normalmente distribuiti.

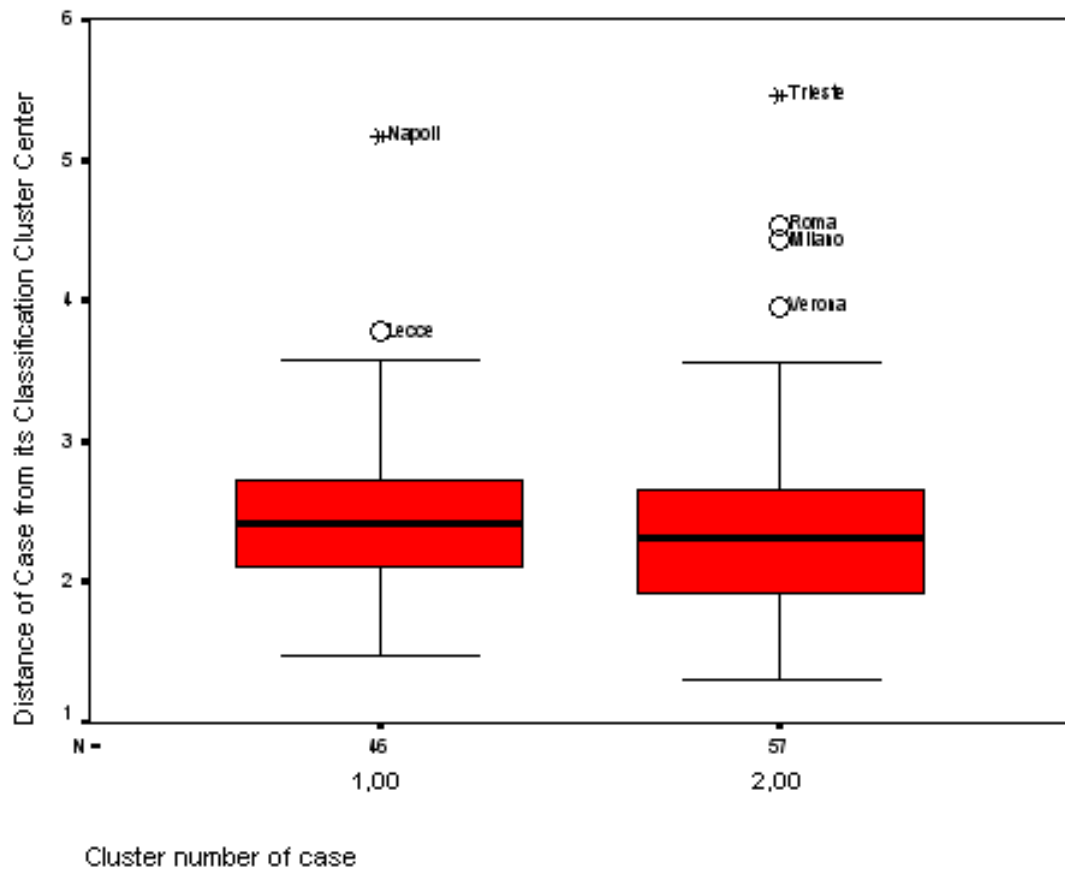
**Tab. 22 Output dell'analisi cluster non gerarchica k-medie**

Case Number	Comuni capoluogo di provincia	Cluster	Distance
1	Agrigento	1	2,596
2	Alessandria	2	2,152
3	Ancona	2	2,284
4	Aosta	2	2,726
5	Arezzo	2	2,603
6	Ascoli Piceno	1	2,170
7	Asti	1	2,379
8	Avellino	1	2,214
9	Bari	1	2,668
10	Belluno	2	2,634
11	Benevento	1	2,034
12	Bergamo	2	1,741
13	Biella	2	2,466
14	Bologna	2	3,557
15	Bolzano	2	3,287
16	Brescia	2	1,465
17	Brindisi	1	2,015
18	Cagliari	1	2,794
19	Caltanissetta	1	2,712
20	Campobasso	1	2,541
21	Caserta	1	2,003
22	Catania	1	3,488
23	Catanzaro	1	1,483
24	Chieti	1	2,191
25	Como	2	1,905
26	Cosenza	1	1,970
27	Cremona	2	1,863
28	Crotone	1	2,423
29	Cuneo	2	2,438
30	Enna	1	2,695
31	Ferrara	2	1,579
32	Firenze	2	3,525
33	Foggia	1	1,906
34	Forlì	2	1,305
35	Frosinone	1	1,800
36	Genova	2	3,141
37	Gorizia	2	2,865
38	Grosseto	1	3,583
39	Imperia	2	3,241
40	Isernia	1	1,826
41	La Spezia	2	2,600
42	L'Aquila	2	2,301
43	Latina	1	2,873
44	Lecce	1	3,778
45	Lecco	2	2,058
46	Livorno	2	2,531
47	Lodi	2	2,166
48	Lucca	2	2,119
49	Macerata	2	2,167
50	Mantova	2	2,079
51	Massa	1	2,717
52	Matera	1	2,523
53	Messina	1	2,414
54	Milano	2	4,426



55	Modena	2	1,706
56	Napoli	1	5,168
57	Novara	2	1,574
58	Nuoro	1	2,839
59	Oristano	1	2,192
60	Padova	2	2,460
61	Palermo	1	3,376
62	Parma	2	1,396
63	Pavia	2	2,325
64	Perugia	2	2,315
65	Pesaro	2	2,593
66	Pescara	2	2,017
67	Piacenza	2	2,662
68	Pisa	2	2,365
69	Pistoia	2	1,456
70	Pordenone	2	2,587
71	Potenza	1	2,274
72	Prato	2	1,726
73	Ragusa	1	2,332
74	Ravenna	2	1,414
75	Reggio Calabria	1	2,193
76	Reggio Emilia	2	2,017
77	Rieti	1	1,930
78	Rimini	2	3,493
79	Roma	2	4,536
80	Rovigo	1	2,640
81	Salerno	1	2,226
82	Sassari	1	2,061
83	Savona	2	2,054
84	Siena	2	2,022
85	Siracusa	1	2,420
86	Sondrio	2	2,958
87	Taranto	1	1,836
88	Teramo	1	2,114
89	Terni	1	2,589
90	Torino	2	2,704
91	Trapani	1	2,135
92	Trento	2	2,552
93	Treviso	2	2,492
94	Trieste	2	5,455
95	Udine	2	1,924
96	Varese	2	2,014
97	Venezia	2	1,821
98	Verbano Cusio Ossola	1	2,455
99	Vercelli	1	3,453
100	Verona	2	3,960
101	Vibo Valentia	1	3,237
102	Vicenza	2	1,831
103	Viterbo	1	2,751

Fig. 6 Il box-plot per il controllo della variabilità dei casi all'interno dei due cluster



**Tab. 23 Output dell'analisi discriminante: la classifica dei comuni capoluogo di provincia (anno 2001)**

<b>ranking</b>	<b>comuni capoluogo</b>	<b>scores AD</b>
1	Bolzano	3,73455
2	Milano	3,56886
3	Trieste	3,4722
4	Firenze	2,99554
5	Verona	2,92832
6	Bologna	2,85585
7	Trento	2,66535
8	Padova	2,44093
9	Parma	2,4347
10	Bergamo	2,32103
11	Reggio Emilia	2,16047
12	Roma	2,07204
13	Venezia	2,03403
14	Genova	1,98802
15	Vicenza	1,92013
16	Brescia	1,91834
17	Siena	1,83737
18	Treviso	1,77877
19	Pavia	1,76523
20	Ancona	1,73879
21	Modena	1,54108
22	Perugia	1,48977
23	Mantova	1,45269
24	Cremona	1,41799
25	Torino	1,38325
26	Piacenza	1,34533
27	Pistoia	1,29219
28	Novara	1,27987
29	Alessandria	1,26392
30	Varese	1,26352
31	Imperia	1,23849
32	Ravenna	1,23077
33	Forli	1,18627
34	Udine	1,14152
35	Lucca	1,09501
36	Pordenone	0,99184
37	Macerata	0,95712
38	Rimini	0,95213
39	Ferrara	0,95051
40	Sondrio	0,94151
41	Belluno	0,92927
42	Lecco	0,88188
43	Como	0,83942
44	Pesaro	0,81602
45	Pisa	0,78859
46	Arezzo	0,75162
47	La Spezia	0,63983
48	Gorizia	0,62717
49	Biella	0,47193
50	Aosta	0,46433

51	Livorno	0,44031
52	Cuneo	0,34814
53	Prato	0,30903
54	Rovigo	0,23387
55	Lodi	0,20585
56	Savona	0,13406
57	L'Aquila	0,12794
58	Verbano Cusio Ossola	0,11531
59	Chieti	0,10513
60	Asti	0,0576
61	Pescara	-0,03952
62	Massa	-0,28495
63	Grosseto	-0,32537
64	Terni	-0,44929
65	Vercelli	-0,47831
66	Ascoli Piceno	-0,47996
67	Latina	-0,86294
68	Viterbo	-0,88817
69	Foggia	-1,12182
70	Teramo	-1,23118
71	Matera	-1,24171
72	Rieti	-1,27009
73	Salerno	-1,37521
74	Campobasso	-1,42012
75	Cosenza	-1,51322
76	Frosinone	-1,58604
77	Nuoro	-1,65373
78	Catania	-1,66725
79	Napoli	-1,69324
80	Cagliari	-1,81843
81	Ragusa	-1,88867
82	Sassari	-1,9438
83	Avellino	-1,95276
84	Messina	-2,03882
85	Lecce	-2,07187
86	Catanzaro	-2,0875
87	Bari	-2,12005
88	Potenza	-2,15709
89	Oristano	-2,16985
90	Trapani	-2,25472
91	Isernia	-2,28164
92	Taranto	-2,31234
93	Palermo	-2,46278
94	Benevento	-2,49076
95	Brindisi	-2,7285
96	Agrigento	-3,00557
97	Caserta	-3,02075
98	Reggio Calabria	-3,04106
99	Siracusa	-3,21672
100	Caltanissetta	-3,62321
101	Enna	-3,89831
102	Crotone	-4,01627
103	Vibo Valentia	-4,14901

