

**SESSIONE IV**  
METODOLOGIE DI SINTESI E ANALISI DEL  
TERRITORIO

---

Indici compositi per il BES

Pierpaolo Massoli, Matteo Mazziotta, Adriano Pareto e Claudia Rinaldelli

## Indici compositi per il BES

Pierpaolo Massoli, Matteo Mazziotta, Adriano Pareto, Claudia Rinaldelli

Istat

*pimassol@istat.it, mazziott@istat.it, pareto@istat.it, rinaldel@istat.it*

### Sommario

*In questo lavoro si descrive il percorso di studio che è stato seguito per la costruzione degli indici compositi dei domini o dimensioni del BES (Benessere Equo e Sostenibile). In particolare, si presenta un confronto tra sei metodologie di sintesi alternative, basato sui requisiti richiesti per l'indice composito, e si illustra il software che è stato prodotto per condurre le sperimentazioni.*

**Parole chiave:** indice sintetico, aggregazione, normalizzazione, software

### Abstract

*The paper describes the analyses followed for the construction of the composite indicators representing each domain (dimension) of the so-called Benessere Equo e Solidale – BES (Fair and Sustainable Well-Being). The focus is about a comparison among six different methodologies for aggregating individual indicators, based on the desirable requirements for the composite index. We also illustrate a specific software developed to carry out the experiments.*

**Key words:** composite indicator, aggregation, normalization, software

### 1. Il percorso di studio e lo sviluppo degli indici compositi

Nel Marzo del 2013 e a Giugno 2014 sono stati presentati, al cospetto delle più alte cariche dello Stato, rispettivamente il primo e il secondo rapporto sul Benessere Equo e Sostenibile (BES). Come ormai è noto, il progetto è nato da un'iniziativa del Cnel e dell'Istat e si inquadra nel dibattito internazionale denominato "Beyond GDP". L'idea di base si fonda sulla convinzione che i parametri sui quali valutare il progresso di una società non debbano essere solo di carattere economico, ma anche sociale e ambientale e, pertanto, debbano essere presenti anche misure di disuguaglianza e sostenibilità. È coscienza comune che numerosi fenomeni socio-economici non possano più essere misurati da un solo indicatore descrittivo e che, invece, essi debbano essere rappresentati su più dimensioni. Fenomeni quali lo sviluppo, il progresso, la povertà, la disuguaglianza sociale, il benessere, la qualità della vita, la dotazione infrastrutturale, ecc. richiedono, per essere misurati, la "combinazione" di dimensioni diverse, da considerarsi congiuntamente come delle proxy del fenomeno; tale combinazione può essere ottenuta applicando metodologie di sintesi note come indici sintetici.

La misura del BES è attualmente articolata in 12 dimensioni e 134 indicatori elementari. La sintesi degli indicatori elementari del BES è stato uno dei temi di discussione della

Commissione Istat per la misurazione del benessere; il percorso di studio e sperimentazione delle sintesi si è svolto a partire dalla fine dell'anno 2010, prendendo in particolare come riferimento il volume "Handbook on Constructing Composite Indicators. Methodology and user guide" (OECD, 2008) e una serie di metodi di sintesi (sperimentali e non) applicati a livello internazionale. Varie tappe si sono susseguite nel percorso di studio e di sperimentazione, sia all'interno dell'Istat che in ambito scientifico-accademico; in particolare, si ricordano gli incontri tra i ricercatori Istat e i ricercatori del JRC (Joint Research Centre of European Commission, ossia l'organizzazione che, nel corso degli ultimi anni, è divenuta un indiscusso punto di riferimento internazionale sul tema), da cui sono scaturiti una serie di esperimenti di sintesi su alcuni domini pilota del BES. Tuttavia dalle discussioni svolte in Commissione, è emerso che le principali metodologie di sintesi proposte in letteratura non godono di alcune proprietà importanti di cui Istat, facendo statistica ufficiale, deve tener conto.

A partire da giugno 2012, sono stati quindi esplorati altri metodi di sintesi che godono di una serie di requisiti, sia teorici, che pratici; in particolare, è stata individuata una metodologia che possa ben 'sintetizzare' tutti i domini del BES, consentendo confronti temporali e spaziali tra le regioni italiane; una metodologia che possa, da un lato, rispettare gran parte delle proprietà desiderabili di un indice sintetico e dall'altro mantenere un livello di semplicità tale da poterne agevolare la comunicazione e la diffusione tra addetti ai lavori e non solo. La diffusione di un indice sintetico, infatti, deve il suo successo alla combinazione tra rigore statistico ed elevato livello di comunicabilità; non è un caso che tra gli indici sintetici più noti ci siano lo Human Development Index (HDI) e lo Human Poverty Index (HPI) che si basano su un numero esiguo di indicatori, non sostituibili, aggregati tra loro attraverso medie di potenze.

Per questo motivo, sono state studiate e sperimentate differenti metodologie di sintesi (sia di natura compensativa che non-compensativa) e i risultati sono stati accompagnati dall'analisi di influenza che, in questi casi, risulta assolutamente indispensabile al fine di valutare la 'performance' del metodo testato. I requisiti richiesti e i metodi utilizzati sono descritti nel par. 2, mentre la formulazione matematica degli indici compositi è riportata nel par. 3.

Al fine di realizzare la sperimentazione, è stato sviluppato un software che consente di applicare agevolmente questo ventaglio di tecniche ai domini del BES. Il software, denominato COMIC (COMposite Indices Creator) è descritto nel par. 4.

I risultati delle sperimentazioni sono stati presentati in diversi contesti scientifici (Massoli et al., 2013a; 2013b).

## 2. I requisiti degli indici compositi: metodi a confronto

Come introdotto nel par. 1, i possibili metodi di sintesi per il BES (in quanto statistica ufficiale) devono possedere una serie di requisiti, di seguito riportati: a) la comparabilità spaziale, ossia la possibilità di confrontare valori di sintesi tra unità territoriali; b) la comparabilità temporale, ossia la possibilità di confrontare valori di sintesi nel tempo; c) la non-sostituibilità degli indicatori elementari, ossia l'attribuzione dello stesso peso agli indicatori elementari e l'impossibilità di compensare il valore di uno con quello di un altro; d) la semplicità e trasparenza di calcolo; e) l'immediata fruizione e interpretazione dei risultati di output; f) la robustezza dei risultati ottenuti.

Sono stati quindi studiati e messi a confronto i seguenti metodi di sintesi:

- 1) *Media indici 0-1*, che consiste nella standardizzazione con metodo min-max degli indicatori elementari e aggregazione con la media aritmetica;
- 2) *Media z-scores*, che consiste nella trasformazione in scarti standardizzati degli indica-

- tori elementari e aggregazione con la media aritmetica;
- 3) *Mazziotta-Pareto Index (MPI)*, che consiste nella trasformazione degli indicatori elementari in variabili con media 100 e scostamento quadratico medio 10 e aggregazione con la media aritmetica penalizzata dalla variabilità «orizzontale» degli indicatori medesimi;
  - 4) *Indice di Jevons*, che consiste nella trasformazione degli indicatori elementari in numeri indici con base 'Italia' (indice statico) o numeri indici con base 'anno precedente' (indice dinamico) e aggregazione con la media geometrica;
  - 5) *Indice Media Geometrica (IMG)*, che consiste nella standardizzazione *min-max* degli indicatori elementari e aggregazione con la media geometrica;
  - 6) *Indice MPI corretto*, che consiste nella standardizzazione *min-max* degli indicatori elementari e aggregazione con la media aritmetica penalizzata dalla variabilità «orizzontale» degli indicatori medesimi.

Naturalmente, non sono stati considerati nelle sperimentazioni altri approcci, molto noti in letteratura, come l'Analisi in Componenti Principali, il metodo di Wroclaw o l'Analisi Multicriteria, poiché disattendono uno o più dei requisiti richiesti.

Si noti, tuttavia, che anche alcuni dei sei metodi testati hanno delle caratteristiche che non consentono il rispetto di tutti i requisiti. Per un'analisi dettagliata, si veda la tavola 1.

**Tavola 1 – Metodi di sintesi e requisiti richiesti**

		Requisiti					
		a	b	c	d	e	f
Metodi sperimentati	1	SI	SI	NO	SI	NO	SI
	2	SI	NO	NO	SI	NO	SI
	3	SI	NO	SI	SI	SI	SI
	4	SI	SI	SI	SI	NO	NO
	5	SI	SI	SI	SI	SI	NO
	6	SI	SI	SI	SI	SI	SI

Le caratteristiche metodologiche dell'indice MPI corretto (Mazziotta e Pareto, 2013) consentono di rispettare tutte le proprietà sopra elencate, secondo quanto risulta dagli esperimenti eseguiti sui domini pilota del BES.

### 3. Formulazione degli indici compositi

#### 3.1. Media indici 0-1

Data la matrice  $\mathbf{X}=\{x_{ijt}\}$  di  $n$  righe (unità statistiche),  $m$  colonne (indicatori) e  $p$  strati (anni), si calcola la matrice trasformata  $\mathbf{R}=\{r_{ijt}\}$ , con:

$$r_{ijt} = \begin{cases} \frac{(x_{ijt} - \text{Min}_{x_j})}{(\text{Max}_{x_j} - \text{Min}_{x_j})} & \text{se l'indicatore } j \text{ ha polarità positiva;} \\ \frac{(\text{Max}_{x_j} - x_{ijt})}{(\text{Max}_{x_j} - \text{Min}_{x_j})} & \text{se l'indicatore } j \text{ ha polarità negativa;} \end{cases} \quad (1)$$

dove  $\text{Min}_{x_j}$  e  $\text{Max}_{x_j}$  sono, rispettivamente, il minimo e il massimo dell'indicatore  $j$ , tra le  $n$  unità statistiche, nei  $p$  anni considerati o dei valori forniti dall'esterno. L'indice sintetico, per l'unità  $i$  al tempo  $t$ , è dato dalla formula:

$$M_{it}^{0-1} = \frac{\sum_{j=1}^m r_{ijt}}{m}. \quad (2)$$

### 3.2. Media z-scores

Data la matrice  $\mathbf{X}=\{x_{ijt}\}$  di  $n$  righe (unità statistiche),  $m$  colonne (indicatori) e  $p$  strati (anni), si calcola la matrice trasformata  $\mathbf{Z}=\{z_{ijt}\}$ , con:

$$z_{ijt} = \begin{cases} \frac{(x_{ijt} - M_{x_{jt}})}{S_{x_{jt}}} & \text{se l'indicatore } j \text{ ha polarità positiva;} \\ -\frac{(x_{ijt} - M_{x_{jt}})}{S_{x_{jt}}} & \text{se l'indicatore } j \text{ ha polarità negativa;} \end{cases} \quad (3)$$

dove  $M_{x_{jt}}$  e  $S_{x_{jt}}$  sono, rispettivamente, la media e lo scostamento quadratico medio dell'indicatore  $j$  al tempo  $t$ .

L'indice sintetico, per l'unità  $i$  al tempo  $t$ , è dato dalla formula:

$$M_{it}^z = \frac{\sum_{j=1}^m z_{ijt}}{m}. \quad (4)$$

### 3.3. Mazziotta-Pareto index (MPI)

Data la matrice  $\mathbf{X}=\{x_{ijt}\}$  di  $n$  righe (unità statistiche),  $m$  colonne (indicatori) e  $p$  strati (anni), si calcola la matrice trasformata  $\mathbf{Z}=\{z_{ijt}\}$ , con:

$$z_{ijt} = \begin{cases} 100 + \frac{(x_{ijt} - M_{x_{jt}})}{S_{x_{jt}}} 10 & \text{se l'indicatore } j \text{ ha polarità positiva;} \\ 100 - \frac{(x_{ijt} - M_{x_{jt}})}{S_{x_{jt}}} 10 & \text{se l'indicatore } j \text{ ha polarità negativa;} \end{cases} \quad (5)$$

dove  $M_{x_{jt}}$  e  $S_{x_{jt}}$  sono, rispettivamente, la media e lo scostamento quadratico medio dell'indicatore  $j$  al tempo  $t$ .

L'indice sintetico, per l'unità  $i$  al tempo  $t$ , è dato dalla formula<sup>1</sup>:

<sup>1</sup> E' una forma generalizzata, poiché include due indici in uno. Se l'indicatore è di tipo crescente o positivo, ossia se a variazioni crescenti dell'indicatore corrispondono variazioni positive del fenomeno (per esempio, il benessere), si utilizza  $MPI^-$ . Viceversa, se l'indicatore è di tipo decrescente o negativo, ossia se a variazioni crescenti dell'indicatore corrispondono variazioni negative del fenomeno (per esempio, il malessere), si utilizza  $MPI^+$ .

$$\text{MPI}_{it}^{+/-} = M_{z_{it}} \pm S_{z_{it}} \text{cv}_{it} \quad (6)$$

$$\text{dove } M_{z_{it}} = \frac{\sum_{j=1}^m z_{ijt}}{m}; S_{z_{it}} = \sqrt{\frac{\sum_{j=1}^m (z_{ijt} - M_{z_{it}})^2}{m}}; \text{cv}_{it} = \frac{S_{z_{it}}}{M_{z_{it}}}. \quad (7)$$

### 3.4. Indice di Jevons

Data la matrice  $\mathbf{X}=\{x_{ijt}\}$  di  $n$  righe (unità statistiche),  $m$  colonne (indicatori) e  $p$  strati (anni), l'indice sintetico 'statico', per l'unità  $i$  al tempo  $t$ , è dato dalla formula:

$$\text{JS}_{it} = \prod_{j=1}^m \left( \frac{x_{ijt}}{x_{bjt}} 100 \right)^{\frac{1}{m}} \quad (8)$$

dove  $x_{bjt}$  è il valore base dell'indicatore  $j$  al tempo  $t$  (nel caso delle regioni italiane, è la media nazionale nell'anno di riferimento).

L'indice sintetico 'dinamico', per l'unità  $i$  al tempo  $t$ , è dato dalla formula:

$$\text{JD}_{it} = \prod_{j=1}^m \left( \frac{x_{ijt}}{x_{ij(t-1)}} 100 \right)^{\frac{1}{m}} \quad (9)$$

dove  $x_{ij(t-1)}$  è il valore dell'indicatore  $j$ , per l'unità  $i$  al tempo  $t-1$ .

### 3.5. Indice Media Geometrica (IMG)

Data la matrice  $\mathbf{X}=\{x_{ijt}\}$  di  $n$  righe (unità statistiche),  $m$  colonne (indicatori) e  $p$  strati (anni), si calcola la matrice trasformata  $\mathbf{R}=\{r_{ijt}\}$ , con:

$$r_{ijt} = \begin{cases} \frac{(x_{ijt} - \text{Min}_{x_{jt}})}{(\text{Max}_{x_{jt}} - \text{Min}_{x_{jt}})} 198 + 1 & \text{se l'indicatore } j \text{ ha polarità positiva;} \\ \frac{(\text{Max}_{x_{jt}} - x_{ijt})}{(\text{Max}_{x_{jt}} - \text{Min}_{x_{jt}})} 198 + 1 & \text{se l'indicatore } j \text{ ha polarità negativa;} \end{cases} \quad (10)$$

dove  $\text{Min}_{x_{jt}}$  e  $\text{Max}_{x_{jt}}$  sono, rispettivamente, il minimo e il massimo dell'indicatore  $j$ , tra le  $n$  unità statistiche, nei  $p$  anni considerati o dei valori forniti dall'esterno. Tali valori possono essere calcolati in modo da porre uguale a 100 un valore di riferimento (nel caso delle regioni italiane, la media nazionale nell'anno base).

L'indice sintetico, per l'unità  $i$  al tempo  $t$ , è dato dalla formula:

$$\text{IMG}_{it} = \prod_{j=1}^m (r_{ijt})^{\frac{1}{m}}. \quad (11)$$

### 3.6. Indice MPI corretto

Data la matrice  $\mathbf{X}=\{x_{ijt}\}$  di  $n$  righe (unità statistiche),  $m$  colonne (indicatori) e  $p$  strati (anni), si calcola la matrice trasformata  $\mathbf{R}=\{r_{ijt}\}$ , con:

$$r_{ijt} = \begin{cases} \frac{(x_{ijt} - \text{Min}_{x_{jt}})}{(\text{Max}_{x_{jt}} - \text{Min}_{x_{jt}})} 60 + 70 & \text{se l'indicatore } j \text{ ha polarità positiva;} \\ \frac{(\text{Max}_{x_{jt}} - x_{ijt})}{(\text{Max}_{x_{jt}} - \text{Min}_{x_{jt}})} 60 + 70 & \text{se l'indicatore } j \text{ ha polarità negativa;} \end{cases} \quad (12)$$

dove  $\text{Min}_{x_{jt}}$  e  $\text{Max}_{x_{jt}}$  sono, rispettivamente, il minimo e il massimo dell'indicatore  $j$ , tra le  $n$  unità statistiche, nei  $p$  anni considerati o dei valori forniti dall'esterno. Tali valori possono essere calcolati in modo da porre uguale a 100 un valore di riferimento (nel caso delle regioni italiane, la media nazionale nell'anno base). L'indice sintetico, per l'unità  $i$  al tempo  $t$ , è dato dalla formula:

$$\text{MPI}_{it}^{+/-} = M_{r_{it}} \pm S_{r_{it}} \text{cv}_{it} \quad (13)$$

$$\text{dove } M_{r_{it}} = \frac{\sum_{j=1}^m r_{ijt}}{m}; S_{r_{it}} = \sqrt{\frac{\sum_{j=1}^m (r_{ijt} - M_{r_{it}})^2}{m}}; \text{cv}_{it} = \frac{S_{r_{it}}}{M_{r_{it}}}. \quad (14)$$

### 4. Comic: un software generalizzato per la sintesi degli indicatori

Il software COMIC nasce dall'esigenza di calcolare ed eventualmente confrontare gli indici composti prodotti mediante i metodi descritti nel par. 3, in modo agevole, senza dover ricorrere a programmi informatici scritti *ad hoc*.

COMIC è scritto in SAS ed è semplice da usare, poiché consente all'utente, attraverso una serie di menù *user-friendly*, di applicare uno o più metodi di sintesi agli indicatori elementari specificati dall'esterno.

Il software fornisce in output una reportistica efficace; in particolare, produce le sintesi secondo il metodo scelto, le relative graduatorie per unità geografica specificata, le mappe geografiche e le serie storiche degli indici composti, grafici, confronti tra metodi (correlazione di Spearman tra graduatorie e differenza media assoluta di rango), ecc.

Infine, COMIC consente di effettuare un'analisi di influenza delle sintesi calcolate. In particolare, è possibile valutare se, e con quale intensità, le graduatorie delle unità geografiche cambiano a seguito dell'eliminazione di un indicatore elementare dall'insieme di partenza.

**Bibliografia**

- Mazziotta, M., Pareto, A. (2013). *A Non-compensatory Composite Index for Measuring Well-being over Time*. Cogito. Multidisciplinary Research Journal, Vol. V, 4.
- Massoli, P., Mazziotta, M., Pareto, A., Rinaldelli, C. (2013). *La misura del BES una sperimentazione per l'aggregazione degli indicatori dell'istruzione e della formazione*. Primo Convegno Nazionale dell'AIQUAV, Luglio 2013.
- Massoli, P., Mazziotta, M., Pareto, A., Rinaldelli, C. (2013). *Metodologie di sintesi sperimentali per i domini del BES*. XXXIV Conferenza Italiana di Scienze Regionali (AISRE), Agosto 2013.
- OECD (2008). *Handbook on Constructing Composite Indicators. Methodology and user guide*. OECD Publications, Paris.