

FONDAMENTI E METODI PER L'ANALISI EMPIRICA NELLE SCIENZE SOCIALI

Università degli Studi di Macerata

Prof. Mattia Tassinari

Indicatori composti

Il percorso in sintesi...

- La sociologia come scienza empirica
- Domanda di ricerca e revisione della letteratura
- Strategie/tipi di ricerca e fasi della ricerca
- Raccolta dei dati: campionamento e raccolta dati primari (questionario, intervista, focus group)
- Elaborazione ed analisi statistica dei dati
- **Indicatori composti**

INDICATORE COMPOSTO (o INDICE)

- Indicatore sintetico che descrive un fenomeno **complesso** (non misurabile attraverso un solo indicatore)
- Fornisce una **classifica** (ranking) che indica, in termini quantitativi, come il fenomeno di interesse si presenta in ciascuna unità statistica relativamente alle altre
- Esempio: ***Classifica dei migliori atenei d'Italia*** (esempio di variabili rilevanti: *n. di iscritti, n. di pubblicazioni scientifiche, media voto di laurea studenti, % di occupati a 1 anno dalla laurea, ...*)

Costruire un Indice: 4 passaggi

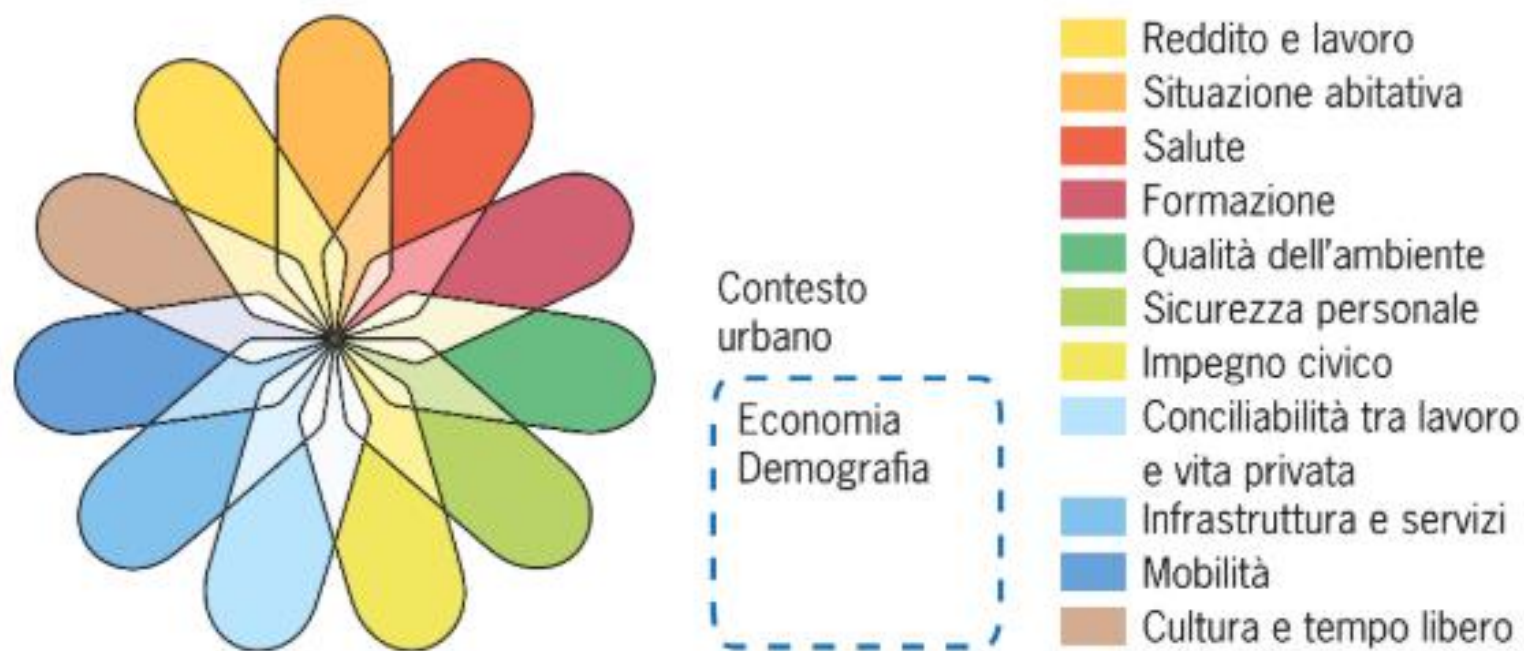
1. Definizione del fenomeno da studiare
2. Individuazione dei singoli indicatori rilevanti e selezione tramite analisi della correlazione
3. Normalizzazione degli indicatori per renderli confrontabili
4. Attribuzione del peso ai singoli indicatori e aggregazione in un unico valore

1 - Definizione del fenomeno "*complesso*" da studiare

Esempio:

Come può essere misurata la qualità della vita?

Dimensioni della qualità della vita



Fonte: OCSE (2014) «How's Life in Your Region? Measuring Regional and Local Well-being for Policy Making», pubblicazione OCSE, Parigi;
Grafico adattato dall'UST

© UST 2016

2 – Individuazione dei singoli indicatori rilevanti e selezione tramite analisi della correlazione

⇒ Costruzione degli indicatori rilevanti

Esempi:

- Variazioni temporali
- Una parte sul tutto



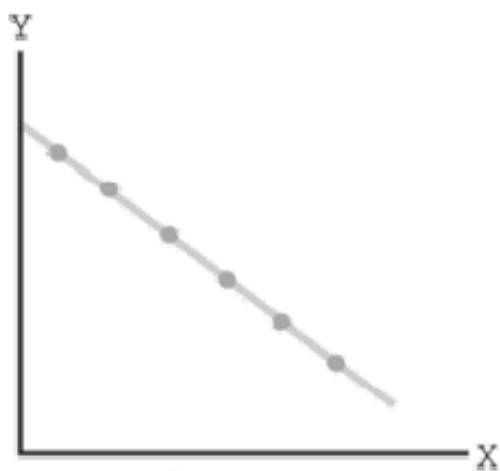
CORRELAZIONE (Indice di Pearson, r)

- Indicatore statistico che esprime la relazione tra due variabili
- Valori assunti: $[-1;1]$
- -1 : perfetta correlazione negativa
- $+1$: perfetta correlazione positiva
- 0 : nessuna relazione lineare tra le due variabili
- 0.5 punto di riferimento per considerare la correlazione rilevante o meno

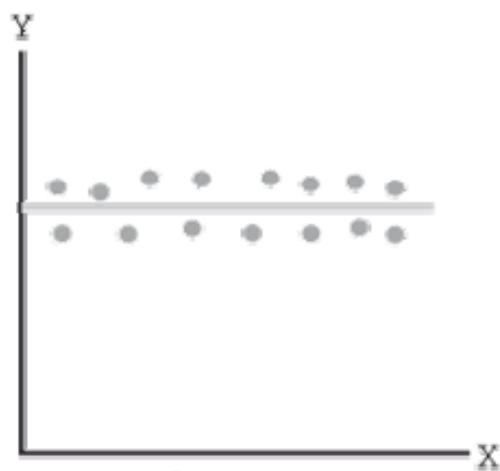
CORRELAZIONE (Indice di Pearson, r)

Indicatore statistico che esprime la relazione tra due variabili: X e Y

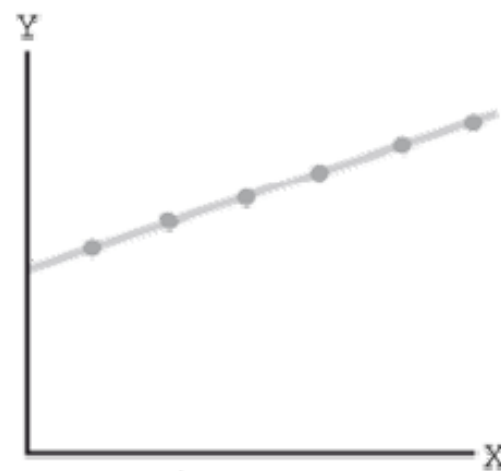
Esempio: correlazione positiva => all'aumentare di x aumenta anche y



Riquadro A
Perfetta correlazione
negativa ($= -1$)



Riquadro B
Assenza di correlazione
($= 0$)



Riquadro C
Perfetta correlazione
positiva ($= +1$)

=> Possibile eliminazione dall'indicatore composto
delle variabili non adeguatamente correlate (0.3/0.7)



3 - Normalizzazione degli indicatori per renderli confrontabili

Normalizzazione delle variabili:



$$Z_{ij} = \frac{x_{ij} - \min_i(x_{ij}) + \frac{1}{n}}{\max_i(x_{ij}) - \min_i(x_{ij}) + \frac{2}{n}} \in (0,1)$$

Dove:

- X_{ij} è il valore della i -esima osservazione ($i = 1, 2, \dots, n$) assunto dalla j -esima variabile ($j = 1, 2, \dots, k$);
- $\min_i (X_{ij})$ è il valore minimo assunto dalla variabile;
- $\max_i (X_{ij})$ è il valore massimo assunto dalla variabile;
- n è il numero di unità statistiche;
- Z_{ij} è il valore della variabile normalizzata.

4 - Attribuzione del peso ai singoli indicatori e aggregazione in un unico valore

Alcune possibili funzioni combinanti:

1. FISHER: $y_i = -\sum_{j=1}^k \omega_j \log(1 - z_{ij})$

2. LIPTAK: $y_i = \sum_{j=1}^k \omega_j \varphi^{-1}(z_{ij})$ con $\varphi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$

3. LOGISTICA: $y_i = \sum_{j=1}^k \omega_j \log\left(\frac{z_{ij}}{1-z_{ij}}\right)$

4. TIPPETT: $y_i = \max_j (\omega_j z_{ij})$

5. ADDITIVA: $y_i = \sum_{j=1}^k \omega_j z_{ij}$



Dove:

- k è il numero di variabili;
- Z_{ij} è il valore della variabile normalizzata della i -esima osservazione ($i = 1, 2, \dots, n$) assunto dalla j -esima variabile ($j = 1, 2, \dots, k$);
- ω_j è il peso della variabile X_j .

4 - Attribuzione del peso ai singoli indicatori e aggregazione in un unico valore

Trasformazione a **RANGO** delle osservazioni, cioè al loro **numero d'ordine** (al posto del valore delle osservazioni).



E' possibile calcolare lo stesso indice per **anni diversi** e **confrontare la variazione delle classifiche nel tempo** (come ciascuna unità statistica modifica il proprio rango nel tempo).



Esempi di applicazione di indicatori composti

Esempio 1:

CONFRONTO TRA INDICI DIVERSI

Obiettivo dell'analisi

**Costruire un indicatore composto
che fornisca un'informazione preliminare sui settori strategici
– l'Indice di Settore Strategico (SSI) –**

Il metodo ha lo scopo di rendere il processo di scelta dei settori strategici più rigoroso e trasparente in quanto:

Si evidenziano i criteri di individuazione dei settori strategici e i pesi ad essi attribuiti.

L'Indice di Settore Strategico (SSI): il metodo utilizzato

1. Definizione del fenomeno di analisi => la rilevanza "strategica" dei diversi settori dell'economia valutata sulla loro capacità di produrre reddito
2. Scelta delle variabili che descrivono l'obiettivo: analisi preliminare delle correlazioni (es. coefficienti di correlazione di Pearson).
3. Normalizzazione delle variabili:

$$z_{ij} = \frac{x_{ij} - \min_i(x_{ij}) + \frac{1}{n}}{\max_i(x_{ij}) - \min_i(x_{ij}) + \frac{2}{n}} \quad \in (0,1)$$

Dove:

- X_{ij} è il valore della i -esima osservazione ($i = 1, 2, \dots, n$) assunto dalla j -esima variabile ($j = 1, 2, \dots, k$);
- $\min_i(X_{ij})$ è il valore minimo assunto dalla variabile nella distribuzione;
- $\max_i(X_{ij})$ è il valore massimo assunto dalla variabile nella distribuzione;
- n è il numero unità statistiche della distribuzione;
- Z_{ij} è il valore della variabile normalizzata.

4. Scelta delle funzioni combinanti:

1. FISHER: $y_i = -\sum_{j=1}^k \omega_j \log(1 - z_{ij})$
2. LIPTAK: $y_i = \sum_{j=1}^k \omega_j \varphi^{-1}(z_{ij})$ con $\varphi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$
3. LOGISTICA: $y_i = \sum_{j=1}^k \omega_j \log\left(\frac{z_{ij}}{1-z_{ij}}\right)$
4. TIPPETT: $y_i = \max_j (\omega_j z_{ij})$
5. ADDITIVA: $y_i = \sum_{j=1}^k \omega_j z_{ij}$

Dove:

- k è il numero di variabili;
- Z_{ij} indica il valore *normalizzato* i -esima osservazione ($i = 1, 2, \dots, n$) assunto dalla j -esima variabile ($j = 1, 2, \dots, k$);
- ω_i è il peso attribuito alla variabile X_j .

I valori ottenuti dalle funzioni combinanti vengono a loro volta normalizzati, in modo che l'indice assuma un valore compreso tra 0 e 1.

Applicazione al caso dell'industria USA

Obiettivo: costruire un indice che ordini i settori dell'economia americana in base alla **capacità del settore di produrre reddito nazionale**. Anno di riferimento 2012, settori **MANIFATTURIERI**, classificazione NAICS.

Le variabili considerate:

- 1) Valore aggiunto di settore/GDP totale.
- 2) Tasso di crescita del Valore aggiunto di settore (2012/2007)
- 3) Tasso di crescita dell'occupazione nel settore (2012/2007)
- 4) Valore aggiunto per addetto (produttività)
- 5) Investimento privato per addetto
- 6) Esportazioni nette per addetto
- 7) Tasso di crescita del Valore aggiunto per addetto (2012/2007)
- 8) Tasso di crescita del salario per addetto (2012/2007)

Pearson coefficient between variables (-1:1)

Pearson coefficient between variables	VA as perc of GDP (2012)	Growth rate VA (2012- 2007)	Growth rate full-time equivalent employees (2012- 2007)	VA/Full- time equivalent employees (2012)	Growth rate compensati on of employees/ Full-time equivalent employees (2012- 2007)	Net export/Full- time equivalent employees (2012)	Investment in private fixed assets/Full- time equivalent employees (2012)
VA as perc. of GDP (2012)	1,0000						
Growth rate VA (2012- 2007)	0,8130	1,0000					
Growth rate full-time equivalent employees (2012-2007)	0,6698	0,8698	1,0000				
VA/Full-time equivalent employees (2012)	0,3972	0,4008	0,4403	1,0000			
Growth rate compensation of employees/Full-time equivalent employees (2012-2007)	0,3552	0,5538	0,5483	0,6808	1,0000		
Net export/Full-time equivalent employees (2012)	0,2226	0,1602	0,3383	0,1839	0,2239	1,0000	
Investment in private fixed assets/ Full-time equivalent employees (2012)	0,7732	0,6400	0,4917	0,7876	0,5606	0,1506	1,0000

Un primo ranking:

- 1 Petroleum and coal products
- 2 Chemical products
- 3 Other transportation equipment
- 4 Food and beverage and tobacco products
- 5 Computer and electronic products
- 6 Machinery
- 7 Miscellaneous manufacturing
- 8 Plastics and rubber products
- 9 Fabricated metal products
- 10 Paper products
- 11 Electrical equipment, appliances, and components
- 12 Primary metals
- 13 Motor vehicles, bodies and trailers, and parts
- 14 Wood products
- 15 Printing and related support activities
- 16 Nonmetallic mineral products
- 17 Textile mills and textile product mills
- 18 Furniture and related products
- 19 Apparel and leather and allied products

L'Indice di Settore Strategico "GREEN"

Possibilità di focalizzarsi su altri obiettivi di policy. Esempio: riduzione impatto ambientale, efficienza energetica dei settori, ... per premiare i virtuosi (o viceversa).

Anno di riferimento 2012, settori MANIFATTURIERI, classificazione NAICS.

Variabili considerate:

- 1) Energy/Full-time equivalent employees (2012)
- 2) Growth rate Energy/Full-time equivalent employees (2012-2007)
- 3) Energy-Related Carbon Dioxide Emissions/Full-time equivalent employees (2012)
- 4) Growth rate Energy-Related Carbon Dioxide Emissions/Full-time equivalent employees (2012-2011)
- 5) VA/Energy (2012)
- 6) Growth rate VA/Energy (2012-2007)

Il problema di possibili TRADE-OFF:

ISS

- 1 **Petroleum and coal products**
- 2 Chemical products
- 3 Other transportation equipment
- 4 Food and beverage and tobacco products
- 5 Computer and electronic products
- 6 Machinery
- 7 Miscellaneous manufacturing
- 8 Plastics and rubber products
- 9 Fabricated metal products
- 10 Paper products
- 11 Electrical equipment
- 12 Primary metals
- 13 Motor vehicles, bodies and trailers
- 14 Wood products
- 15 Printing and related support activities
- 16 Nonmetallic mineral products
- 17 Textile mills and textile product mills
- 18 Furniture and related products
- 19 Apparel and leather and allied products

ISS “Green”

- 1 Computer and electronic products
- 2 Apparel and leather and allied products
- 3 Miscellaneous manufacturing
- 4 Electrical equipment
- 5 Printing and related support activities
- 6 Furniture and related products
- 7 Textile mills and textile product mills
- 8 Other transportation equipment
- 9 Machinery
- 10 Motor vehicles, bodies and trailers
- 11 Fabricated metal products
- 12 Plastics and rubber products
- 13 Wood products
- 14 Food and beverage and tobacco products
- 15 Chemical products
- 16 Nonmetallic mineral products
- 17 Paper products
- 18 Primary metals
- 19 **Petroleum and coal products**

Esempio 2:

**CONFRONTO DELLO STESSO
INDICE TRA REGIONI DIVERSE**

Applicazione ad alcune regioni italiane.

Obiettivo: costruire un indice che ordini i settori regionali in base alle potenzialità del settore di produrre reddito – 9 settori manifatturieri – Fonte dei dati: ISTAT.

Variabili:

- 1) **Indice di specializzazione settoriale regionale - 2011:**

$$\frac{\text{Occupazione regionale di settore}}{\text{Occupazione regionale totale}} \bigg/ \frac{\text{Occupazione nazionale di settore}}{\text{Occupazione nazionale totale}}$$

- 2) **V.A. PER ADDETTO di settore (produttività di settore) – 2011.**

- 3) **Tasso di crescita V.A. di settore – 2011-2006.**

- 4) **Tasso di crescita V.A. PER ADDETTO di settore – 2011-2006.**

- 5) **Scostamento % della produttività di settore regionale dalla produttività di settore nazionale – 2011:**

$$\frac{\text{VA per addetto regionale di settore} - \text{VA per addetto nazionale di settore}}{\text{VA per addetto nazionale di settore}}$$

Ranking regionali

Lombardia	
Chimica	0,65
Prodotti alimentari	0,58
Elettronica	0,57
Industria pesante	0,48
Tessile e abb.to	0,45
Gomma e plastica	0,34
Arredamento	0,33
Legno e carta	0,30
Mezzi di trasporto	0,15

Piemonte	
Mezzi di trasporto	0,72
Prodotti alimentari	0,47
Chimica	0,40
Gomma e plastica	0,33
Arredamento	0,31
Legno e carta	0,30
Elettronica	0,23
Industria pesante	0,22
Tessile e abb.to	0,17

Liguria	
Mezzi di trasporto	0,78
Arredamento	0,53
Chimica	0,52
Prodotti alimentari	0,51
Gomma e plastica	0,50
Industria pesante	0,46
Elettronica	0,43
Legno e carta	0,25
Tessile e abb.to	0,23

Toscana	
Tessile e abb.to	0,81
Elettronica	0,57
Industria pesante	0,57
Chimica	0,53
Legno e carta	0,47
Arredamento	0,42
Prodotti alimentari	0,37
Gomma e plastica	0,35
Mezzi di trasporto	0,16

Lazio	
Chimica	0,64
Elettronica	0,61
Arredamento	0,45
Tessile e abb.to	0,44
Industria pesante	0,35
Legno e carta	0,34
Gomma e plastica	0,31
Mezzi di trasporto	0,22
Prodotti alimentari	0,20

Emilia Romagna	
Mezzi di trasporto	0,75
Prodotti alimentari	0,73
Elettronica	0,60
Gomma e plastica	0,59
Tessile e abb.to	0,39
Arredamento	0,39
Industria pesante	0,32
Chimica	0,31
Legno e carta	0,23

Veneto	
Arredamento	0,69
Tessile e abb.to	0,67
Elettronica	0,65
Industria pesante	0,63
Gomma e plastica	0,53
Legno e carta	0,50
Prodotti alimentari	0,46
Chimica	0,43
Mezzi di trasporto	0,39

Marche	
Tessile e abb.to	0,75
Legno e carta	0,69
Arredamento	0,65
Industria pesante	0,56
Elettronica	0,41
Mezzi di trasporto	0,41
Chimica	0,38
Gomma e plastica	0,36
Prodotti alimentari	0,15

Campania	
Mezzi di trasporto	0,77
Prodotti alimentari	0,55
Elettronica	0,55
Industria pesante	0,40
Legno e carta	0,36
Tessile e abb.to	0,34
Chimica	0,32
Arredamento	0,30
Gomma e plastica	0,25

Piemonte		Lombardia		Veneto	
Mezzi di trasporto	0,72	Chimica	0,65	Arredamento	0,69
Prodotti alimentari	0,47	Prodotti alimentari	0,58	Tessile e abb.to	0,67
Chimica	0,40	Elettronica	0,57	Elettronica	0,65
Gomma e plastica	0,33	Industria pesante	0,48	Industria pesante	0,63
Arredamento	0,31	Tessile e abb.to	0,45	Gomma e plastica	0,53
Legno e carta	0,30	Gomma e plastica	0,34	Legno e carta	0,50
Elettronica	0,23	Arredamento	0,33	Prodotti alimentari	0,46
Industria pesante	0,22	Legno e carta	0,30	Chimica	0,43
Tessile e abb.to	0,17	Mezzi di trasporto	0,15	Mezzi di trasporto	0,39
Toscana		Emilia Romagna		Marche	
Tessile e abb.to	0,81	Mezzi di trasporto	0,75	Tessile e abb.to	0,75
Elettronica	0,57	Prodotti alimentari	0,73	Legno e carta	0,69
Industria pesante	0,57	Elettronica	0,60	Arredamento	0,65
Chimica	0,53	Gomma e plastica	0,59	Industria pesante	0,56
Legno e carta	0,47	Tessile e abb.to	0,39	Elettronica	0,41
Arredamento	0,42	Arredamento	0,39	Mezzi di trasporto	0,41
Prodotti alimentari	0,37	Industria pesante	0,32	Chimica	0,38
Gomma e plastica	0,35	Chimica	0,31	Gomma e plastica	0,36
Mezzi di trasporto	0,16	Legno e carta	0,23	Prodotti alimentari	0,15
Liguria		Lazio		Campania	
Mezzi di trasporto	0,78	Chimica	0,64	Mezzi di trasporto	0,77
Arredamento	0,53	Elettronica	0,61	Prodotti alimentari	0,55
Chimica	0,52	Arredamento	0,45	Elettronica	0,55
Prodotti alimentari	0,51	Tessile e abb.to	0,44	Industria pesante	0,40
Gomma e plastica	0,50	Industria pesante	0,35	Legno e carta	0,36
Industria pesante	0,46	Legno e carta	0,34	Tessile e abb.to	0,34
Elettronica	0,43	Gomma e plastica	0,31	Chimica	0,32
Legno e carta	0,25	Mezzi di trasporto	0,22	Arredamento	0,30
Tessile e abb.to	0,23	Prodotti alimentari	0,20	Gomma e plastica	0,25