

SAPIENZA - UNIVERSITÀ DI ROMA

ANNALI DEL DIPARTIMENTO DI METODI
E MODELLI PER L'ECONOMIA,
IL TERRITORIO E LA FINANZA

2012 - 2013

PÀTRON EDITORE
Bologna 2014

Direttore Responsabile - Director

Alessandra De Rose

Direttore Scientifico - Editor in Chief

Roberta Gemmiti

Comitato Scientifico - Co-editors

Maria Giuseppina Bruno, Francesca Gargiulo, Roberta Gemmiti, Cristina Giudici, Ersilia Incelli, Antonella Leoncini Bartoli, Isabella Santini, Rosa Vaccaro.

Comitato di Redazione - Editorial Staff

Elena Ambrosetti, Maria Caterina Bramati, Filippo Celata, Augusto Frascatani, Maria Rita Sebastiani, Marco Teodori, Judith Turnbull, Mauro Rota.

Consulenti Scientifici - Advisory Board

Catherine Wihtol de Wenden (CERI-Sciences Po-CNRS Paris), Raimondo Cagiano de Azevedo (Sapienza - Università di Roma), Maria Antonietta Clerici (Politecnico di Milano), Marina Fuschi (Università di Chieti-Pescara), Alessandra Faggian (The Ohio State University), Alison Brown (Cardiff University), Luciano Pieraccini (Università degli Studi Roma Tre), Silvia Terzi (Università degli Studi Roma Tre), Gennaro Olivieri (Luiss Guido Carli), Giulio Fenicia (Università degli Studi di Bari), Angelo Moioli (Università Cattolica del Sacro Cuore), Filomena Racioppi (Sapienza - Università di Roma); Pablo Koch-Medina (Centro di Finanza e Assicurazioni, Università di Zurigo).

External Reviewers - Blind review

Copyright © 2014 by Pàtron editore - Quarto Inferiore - Bologna

I diritti di traduzione e di adattamento, totale o parziale, con qualsiasi mezzo sono riservati per tutti i Paesi. È vietata la riproduzione parziale, compresa la fotocopia, anche ad uso interno o didattico, non autorizzata.

PÀTRON Editore - Via Badini, 12
Quarto Inferiore, 40057 Granarolo dell'Emilia (BO)
Tel. 051.767003
Fax 051.768252
E-mail: info@patroneditore.com
<http://www.patroneditore.com>

Il catalogo generale è visibile nel sito web. Sono possibili ricerche per autore, titolo, materia e collana. Per ogni volume è presente il sommario, per le novità la copertina dell'opera e una breve descrizione del contenuto.

Stampa: Rabbi s.r.l., Bologna per conto di Pàtron editore.

ISBN: 978-88-555-3290-7

ISSN: 2385-0825

IL PROBLEMA DELLA VALUTAZIONE MEDIANTE INDICATORI COMPOSITI IN PRESENZA DI CORRELAZIONE TRA GLI INDICATORI ELEMENTARI

Riassunto: La valutazione ed il confronto tra le performance di unità tra loro antagoniste (ad esempio, paesi, università o altre istituzioni economiche, sociali o politiche) riguarda spesso un aspetto multidimensionale ed ascrivibile ad una variabile latente non osservabile nella sua interezza; diventa quindi necessario far ricorso ad una molteplicità di indicatori e ad un criterio di sintesi delle informazioni in essi contenute, che spesso risultano correlate tra loro. Il presente lavoro si prefigge di valutare gli effetti della presenza di correlazione tra gli indici elementari, sulla graduatoria costruita mediante un indicatore composito. La valutazione consiste nel verificare, sulla base di alcune simulazioni, se e come diversi livelli di correlazione tra le variabili elementari possano indurre variazioni nella graduatoria finale.

1. Introduzione

Il problema della valutazione e del confronto tra le performance di unità tra loro antagoniste (ad esempio, paesi, università o altre istituzioni economiche, sociali o politiche) è sempre più ricorrente e riguarda i più svariati ambiti. Tuttavia, la realtà osservata è estremamente complessa e l'aspetto da valutare è spesso di tipo

* Dipartimento MEMOTEF - Sapienza - Università di Roma.

multidimensionale ed ascrivibile ad una variabile latente non osservabile nella sua interezza; diventa quindi necessario far ricorso ad una molteplicità di indicatori e ad un criterio di sintesi delle informazioni in essi contenute.

A questa finalità rispondono i cosiddetti indicatori composti, cioè degli indicatori aggregati che cercano di fornire una sintesi dei valori presentati dagli indici elementari in corrispondenza di una stessa unità di analisi.

Ovviamente, la descrizione della realtà ottenibile mediante questo processo di aggregazione non potrà essere esaustiva, ma potrà fornire una indicazione di massima, più o meno precisa ed interpretabile a seconda del procedimento di costruzione adottato.

Un procedimento di questo tipo, come diffusamente riconosciuto in letteratura¹, deve prevedere una serie di passi che riguardano, tra le altre cose, l'analisi dimensionale degli indicatori elementari selezionati. In particolare, scopo dell'analisi dimensionale è comprendere la natura sottostante dei dati utilizzati e controllare la validità degli indicatori elementari scelti per la rappresentazione del fenomeno latente oggetto di interesse.

In questa fase, particolare rilevanza assume lo studio delle eventuali relazioni esistenti tra gli indici elementari componenti, al fine di sintetizzare indicatori elementari che non siano ridondanti e sovrapposti. Infatti, se da una parte è fuori dubbio che vada del tutto evitata la presenza di perfetta collinearità tra questi, cioè la ridondanza e la sovrapposizione delle informazioni, dall'altra è quasi inevitabile che tra i diversi indicatori volti a misurare lo stesso fenomeno complesso vi sia una qualche relazione. Si tratta allora di stabilire in che misura l'esistenza di una correlazione comporti il problema del *double counting*, che potrebbe a sua volta condurre a risultati fuorvianti.

Per ovviare a tale problema, è possibile ricorrere ai tipici strumenti dell'analisi multivariata, come l'Analisi Fattoriale o l'Analisi in Componenti Principali, ed ottenere così nuove variabili componenti totalmente incorrelate tra loro. Tuttavia, questa soluzione non sempre è condivisa, a causa della perdita del significato originale delle variabili componenti e, conseguentemente, della loro interpretabilità. In alternativa, si potrebbe allora ricorrere ad una diversa ponderazione delle variabili correlate, che dovrebbero es-

¹ Si vedano, tra gli altri, Nardo et al. (2005) e Maggino (2009).

sere pesate meno per ridurre l'effetto della correlazione. C'è poi chi ritiene che in un contesto multicriterio l'esistenza di correlazione tra gli indicatori componenti sia una caratteristica saliente del fenomeno latente che si vuole misurare e che quindi non vada compensata in alcun modo (Saltelli et al., 2005). In molte altre situazioni, infine, gli indici componenti sono selezionati in modo arbitrario sulla base della concreta disponibilità dei dati, piuttosto che sulla base della loro rilevanza e della validità analitica, prestando quindi scarsa attenzione alle interrelazioni tra essi. Quanto finora osservato testimonia come in letteratura la questione sia ancora del tutto aperta e non esista una soluzione univoca alla gestione della eventuale presenza di correlazione tra gli indicatori elementari².

In tale contesto si inserisce la presente nota, il cui obiettivo è valutare gli effetti della presenza di correlazione tra gli indici elementari sulla graduatoria costruita mediante un indicatore composito.

2. L'effetto della presenza di correlazione tra gli indici elementari componenti sui valori di un indice composito: risultati di uno studio simulato

L'analisi che qui si presenta prende le mosse da una precedente analisi critica della graduatoria degli atenei italiani proposta dal Censis-Servizi per conto del quotidiano La Repubblica (Arezzo et al., 2006); pertanto il contesto di riferimento è quello della valutazione delle facoltà universitarie (unità di analisi) mediante un indice aggregato, ottenuto come sintesi di indicatori appartenenti a diverse aree, in ognuna delle quali è possibile individuare un numero variabile di indicatori elementari.

Per formalizzare il problema, indichiamo con I_{iat} l' i -esimo indice elementare, $i=1, 2, \dots, n_a$, rilevato per l'area a , $a=1, 2, \dots, A$, nella circostanza t , $t=1, 2, \dots, T$, e con I_{iatu} il valore da esso assunto in corrispondenza dell'unità u , $u=1, \dots, U$. Ad esempio, nella classifica Censis-Servizi degli atenei italiani, le aree di interesse sono cinque (produttività; didattica; ricerca; profilo docenti; rapporti

² Si veda, al riguardo, anche quanto osservato da Saltelli et al. (2005), che individuano proprio nella presenza di correlazione uno degli aspetti da approfondire nell'applicazione degli indici composti.

e cooperazione internazionale); gli indici elementari appartenenti all'area produttività sono il tasso di sopravvivenuti tra il primo ed il secondo anno, il tasso di studenti attivi, il tasso di iscritti in corso e così via; la circostanza t di interesse è tipicamente rappresentata dal tempo e le unità sono le diverse sedi universitarie considerate per una data facoltà.

Nel nostro studio, per semplicità verranno considerate due sole aree ($a=1,2$), per ciascuna delle quali si considerano due indici elementari ($i=1, 2$), in due diverse circostanze temporali ($t=1, 2$), rilevati su 44 unità ($u=1, \dots, 44$), corrispondenti alle diverse facoltà di Economia di tutti gli atenei pubblici italiani.

Nella prima serie di simulazioni, i valori degli indici I_{2a1} sono stati generati assumendo un basso livello di correlazione (imputabile soltanto al caso) tra le coppie di indici appartenenti alla stessa area; in particolare si è ipotizzato:

$$\begin{aligned} r(I_{111}, I_{211}) &= 0.05 \\ r(I_{121}, I_{221}) &= 0.05 \end{aligned}$$

Nella seconda serie di simulazioni, invece, sono stati ipotizzati i seguenti coefficienti di correlazione:

$$\begin{aligned} r(I_{111}, I_{211}) &= 0.9 \\ r(I_{121}, I_{221}) &= 0.2 \end{aligned}$$

In altre parole, le due situazioni differiscono tra loro solo per il grado di correlazione presente tra gli indici appartenenti alla stessa area; è importante sottolineare che i valori degli indici I_{111} e I_{121} non variano tra le due serie di simulazioni.

Al tempo 2 è stata poi selezionata una facoltà alla volta e si è ipotizzato che per questa facoltà si registrasse un incremento del 30%, tra il tempo 1 e il tempo 2, per l'indice I_{111} . Tale incremento riguarda esclusivamente l'unità statistica scelta (supponiamo, in generale, la f -esima); in altre parole, il vettore I_{111} ed il suo corrispondente al tempo 2, cioè I_{112} , sono identici tranne che per i valori corrispondenti alla facoltà selezionata, cioè I_{111f} e I_{112f} . Sono invece del tutto identiche le coppie di vettori (I_{121}, I_{122}) e (I_{221}, I_{222}) .

Nell'ipotesi di un legame lineare più o meno forte tra i due indici elementari dell'area 1, la variazione nel solo elemento I_{111f} del vettore I_{111} dovrebbe comportare una corrispondente variazione in

I_{211P} coerente con il grado di correlazione ipotizzato. Si è quindi stimato il valore che ci si dovrebbe attendere per I_{211F} in seguito all'aumento di I_{111P} a meno di una variazione accidentale.

I dati degli indici elementari sono stati poi trasformati secondo i criteri più frequentemente utilizzati in letteratura³ (precisamente: normalizzazione, standardizzazione, ranghi, trasformata MAD⁴ e trasformata IQR⁵), al fine di valutare un eventuale effetto differenziale sui punteggi finali dell'indicatore composito. Infine, coerentemente con la metodologia seguita dal Censis-Servizi, secondo la quale i principali indici elementari di area sono aggregati secondo una funzione lineare priva di struttura di ponderazione, l'indice aggregato per ciascuna delle due aree è stato ottenuto come media semplice dei due indici elementari; analogamente, l'indicatore finale è stato ottenuto come media semplice degli indicatori di area.

Sulla base dell'indicatore finale è stata costruita la graduatoria delle 44 unità statistiche, sia al tempo 1 sia al tempo 2.

La procedura appena descritta è stata seguita per entrambe le simulazioni, condizionatamente all'ipotesi correlativa tra I_{1a1} e I_{2a1} . È chiaro che la differenza di comportamento nelle graduatorie finali è imputabile interamente alla più elevata correlazione ipotizzata nella seconda situazione.

Ad esempio, con riferimento all'unità 19, si è riscontrato che il numero di posizioni cambiate tra le due circostanze temporali nel caso di bassa correlazione è pari a: 5, se si ricorre alla normalizzazione degli indici elementari; 3, con la standardizzazione; 5, con la trasformazione in ranghi; 3, con la trasformata MAD; 4, con la trasformata IQR. Nel caso di correlazione elevata, invece, il numero di posizioni cambiate per ciascuna delle cinque trasformate considerate risulta rispettivamente pari a: 13; 15; 14; 15; 15.

Indipendentemente dal tipo di trasformazione utilizzata, la presenza di forte correlazione tra indicatori comporta quindi uno spostamento consistente in graduatoria dell'unità statistica di in-

³ Per esigenze espositive non ci si è soffermati sulle diverse fasi della costruzione di un indicatore composto (delle quali la scelta del criterio di trasformazione dei dati originari è una delle più importanti), nonché sul modo in cui le diverse scelte adottate in ciascuna fase possano condizionare il punteggio finale dell'indicatore; per una trattazione più approfondita di questi aspetti, si veda tra gli altri Jacob et al. (2004).

⁴ Media degli scarti assoluti.

⁵ Differenza interquartile.

teresse. Infatti nel primo blocco di simulazioni, e quindi in una situazione di quasi assenza di correlazione tra indicatori, l'unità statistica si sposta mediamente di 4 posizioni per ciascuna trasformata; nel secondo blocco di simulazioni, cioè allorché gli indicatori sono molto correlati, lo spostamento medio è di 14.4 posizioni. Essendo le condizioni iniziali identiche tra i due esperimenti, tale differenza è imputabile interamente alla diversa struttura correlativa dei dati.

Ciò conferma quanto già osservato in precedenza e cioè che la presenza di una forte correlazione tra gli indici elementari introduce una distorsione nella valutazione fatta mediante l'indicatore composito e, ovviamente, la distorsione sarà tanto più elevata quanto maggiore è la correlazione.

Si può inoltre osservare che nella prima ipotesi correlativa, l'effetto maggiore si riscontra con la normalizzazione e la trasformazione in ranghi, mentre nella seconda ipotesi, alla normalizzazione corrisponde la più piccola variazione della posizione in graduatoria.

Si è inoltre riscontrato che il numero di posizioni cambiate tra i due tempi in generale dipende dalla posizione iniziale dell'unità statistica sulla quale si applica la variazione; per questo motivo l'analisi è stata ripetuta per tutte le unità statistiche facendole variare una alla volta e valutando le rispettive posizioni in graduatoria in entrambe le circostanze temporali. La tabella 1 mostra le statistiche di sintesi della distribuzione del numero di posizioni cambiate in graduatoria tra i due tempi.

Si nota che, indipendentemente dal tipo di trasformazione fatta, la presenza di forte correlazione tra indicatori ha come effetto uno spostamento medio superiore mostrando chiaramente la presenza del *double counting*. I risultati confermano quanto già osservato in merito all'unità 19: la normalizzazione si caratterizza come la trasformazione più sensibile alla variazione di I_{111} , mentre le meno sensibili sono i ranghi e l'IQR. Tuttavia si può osservare come l'aumento nel numero medio di posizioni cambiate, che si riscontra per ogni trasformazione tra le due ipotesi correlative considerate, sia più elevato per la standardizzazione (3,59 posizioni) piuttosto che per la normalizzazione (3,28) e molto più basso per la trasformazione in ranghi (solo 2 posizioni). Quest'ultima, quindi, è quella che risente meno del problema del *double counting*.

Esiste, ovviamente, un effetto "posizione iniziale"; infatti l'enti-

tà dello spostamento dipende anche dalla posizione occupata in graduatoria nella prima circostanza temporale. Per cogliere tale effetto, le 44 unità statistiche sono state suddivise in 4 gruppi; in particolare il gruppo 1 contiene le unità prime in graduatoria e il gruppo 4 le ultime. I risultati della simulazione sono riportati nella tabella 2.

Almeno due effetti sono degni di nota. Il primo è che, fatta eccezione per il gruppo 4, è confermato il maggior spostamento imputabile alla presenza di correlazione elevata; il secondo è che il gruppo 2 è quello con le fluttuazioni maggiori.

Tab. 1 - Statistiche di riepilogo del numero di posizioni cambiate in graduatoria da ciascuna unità di analisi tra le due circostanze temporali.

		Norma- lizza- zione	Stan- dardiz- zazione	Ranghi	MAD	IQR
Bassa correlazione	Media	9.45	8.82	8.05	8.09	8.02
	Mediana	8.50	7.50	7.00	7.50	7.00
	Q1	3.50	2.75	2.75	2.00	2.00
	Q3	15	13.5	13	14	13
	min	0	0	0	0	0
	Max	24	23	26	22	21
	Dev. Stand. Pop	7.01	6.69	6.37	5.96	6.04
	CV	0.74	0.76	0.79	0.74	0.75
Elevata correlazione	Media	12.73	12.41	10.05	11.18	11.07
	Mediana	12.50	13.50	10.00	12.50	12.00
	Q1	7.75	7.00	3.00	5.75	5.75
	Q3	18	17	14.25	16	16
	min	0	0	0	0	0
	Max	24	25	25	23	25
	Dev. Stand. Pop	6.54	6.30	6.17	6.19	6.23
	CV	0.51	0.51	0.61	0.55	0.56

Tab. 2 - Statistiche di riepilogo del numero di posizioni cambiate in graduatoria tra le due circostanze temporali, per gruppo di appartenenza nella prima.

		Media	Media- na	Q1	Q3	Min	Max	s.d.	CV	
ALTA CORRELAZIONE	Normaliz- zazione	Gruppo 1	15.00	14	12	18	7	24	4.71	0.31
		Gruppo 2	19.45	20	16.5	22.5	14	24	3.42	0.18
		Gruppo 3	12.36	12	10.5	14	9	18	2.64	0.21
		Gruppo 4	4.09	4	2	5.5	0	8	2.31	0.57
	Standar- dizzazione	Gruppo 1	15.18	16	12	18	7	24	4.47	0.29
		Gruppo 2	18.18	20	14	20.5	13	25	3.81	0.21
		Gruppo 3	12.36	14	10	14	8	16	2.57	0.21
		Gruppo 4	3.91	4	2	6	0	7	2.15	0.55
	Ranghi	Gruppo 1	13.36	15	11	17.5	2	21	5.96	0.45
		Gruppo 2	13.45	13	10	15.5	7	25	5.07	0.38
		Gruppo 3	9.64	11	8	13	0	15	4.46	0.46
		Gruppo 4	3.73	3	2	4.5	0	9	2.89	0.78
	MAD	Gruppo 1	13.82	13	11.5	17	2	23	5.49	0.40
		Gruppo 2	16.64	16	14.5	19	11	22	3.26	0.20
		Gruppo 3	10.82	11	8	13.5	5	16	3.49	0.32
		Gruppo 4	3.45	4	1.5	5	0	6	2.02	0.58
IQR	Gruppo 1	13.55	13	10.5	16.5	2	25	5.91	0.44	
	Gruppo 2	16.55	16	14.5	19	11	22	3.45	0.21	
	Gruppo 3	10.82	11	8	13.5	16	5	3.04	0.28	
	Gruppo 4	3.36	4	1.5	5	0	6	1.92	0.57	
BASSA CORRELAZIONE	Normaliz- zazione	Gruppo 1	11.27	10	6.5	17	0	21	6.63	0.59
		Gruppo 2	7.82	7	1.5	12	0	22	6.86	0.88
		Gruppo 3	9.09	9	4	11.5	0	24	6.58	0.72
		Gruppo 4	9.64	7	2.5	16	0	21	7.50	0.78
	Standar- dizzazione	Gruppo 1	10.18	11	4.5	14.5	1	21	6.39	0.63
		Gruppo 2	7.82	5	2.5	13	0	21	6.79	0.87
		Gruppo 3	8.82	8	4	10.5	2	23	6.34	0.72
		Gruppo 4	8.45	7	1.5	14.5	0	19	6.97	0.82
	Ranghi	Gruppo 1	8.82	7	4	14	0	17	5.65	0.64
		Gruppo 2	8.45	5	3.5	12	0	26	7.70	0.91
		Gruppo 3	8.00	7	4	12.5	0	18	5.75	0.72
		Gruppo 4	6.91	6	0	12	0	17	5.99	0.87
	MAD	Gruppo 1	8.82	9	4	14	0	17	5.56	0.63
		Gruppo 2	8.00	7	3	11.5	0	22	6.32	0.79
		Gruppo 3	7.73	8	3.5	10.5	1	16	5.12	0.66
		Gruppo 4	7.82	6	1.5	13.5	0	18	6.66	0.85
IQR	Gruppo 1	8.36	8	2.5	13.5	2	16	5.42	0.65	
	Gruppo 2	7.55	6	3	11	0	21	6.40	0.85	
	Gruppo 3	8.00	7	4.5	11	1	18	5.46	0.68	
	Gruppo 4	8.18	7	2.5	12.5	0	19	6.75	0.83	

3. Conclusioni e sviluppi futuri

Nella costruzione di indicatori compositi, il problema della correlazione esistente tra indici elementari è ben noto in letteratura.

Esso conduce a premiare/penalizzare una stessa unità statistica in maniera più o meno forte a seconda dell'intensità della correlazione esistente tra gli indici stessi (*double counting*). In questo lavoro, attraverso uno studio simulato, si è voluto vedere cosa accade alle graduatorie allorché si considerino coppie di indicatori con livelli di correlazione molto diversi tra loro. I risultati dell'analisi confermano la presenza di un effetto da *double counting*; tale effetto, inoltre, risulta più o meno forte, a seconda del tipo di trasformazione utilizzata (la standardizzazione dà lo spostamento massimo e i ranghi quello minimo), ma anche dalla posizione occupata dall'unità statistica nella situazione iniziale.

I risultati ottenuti suggeriscono l'opportunità di approfondire l'analisi, sia in termini di un maggior numero di indicatori per area, tra loro correlati, sia in termini di gestione del problema; in particolare, è interessante analizzare come opportune strutture di ponderazione, che tengano conto delle correlazioni tra gli indici elementari, possano limitare la ridondanza di questi.

Riferimenti bibliografici

- AREZZO M. F., GUAGNANO G., PETRELLA L. (2006), *Il rating delle università del Censis: un'analisi critica*, in *La valutazione della ricerca - Libro bianco*, Consiglio italiano per le Scienze sociali, ed. Marsilio.
- JACOB. R., SMITH P., GODDARD M. (2004), *Measuring Performance: An Examination of Composite Performance Indicators*, CHE Technical Paper Series 29, University of York.
- MAGGINO F. (2009), *La misurazione dei fenomeni sociali attraverso indicatori statistici*, Archivio e-print Università degli studi di Firenze.
- NARDO M., SAISANA M., SALTELLI A., TARANTOLA S., HOFFMAN A., GIOVANNINI E. (2005), *Handbook on Constructing Composite Indicators: Methodology and User guide*, OECD Statistics Working Paper.
- SALTELLI A., NARDO M., SAISANA M., TARANTOLA S. (2005), *Composite Indicators - The Controversy and the Way Forward*, in *Statistics, Knowledge and Policy - Key indicators to inform decision making*, OECD Publishing.