

**RAGIONANDO  
DI SVILUPPO LOCALE:  
UNA LETTURA "NUOVA"  
DI TEMATICHE "ANTICHE"**

**a cura di  
Francesco Contò  
Mariantonietta Fiore**

**Università degli Studi di Foggia  
Dipartimento di Economia**

**FrancoAngeli**

OPEN  ACCESS



Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (<http://bit.ly/francoangeli-oa>).

**FrancoAngeli Open Access** è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più:

[http://www.francoangeli.it/come\\_publicare/publicare\\_19.asp](http://www.francoangeli.it/come_publicare/publicare_19.asp)

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: [www.francoangeli.it](http://www.francoangeli.it) e iscriversi nella home page al servizio "Informatemi" per ricevere via e-mail le segnalazioni delle novità.

**RAGIONANDO  
DI SVILUPPO LOCALE:  
UNA LETTURA "NUOVA"  
DI TEMATICHE "ANTICHE"**

**a cura di  
Francesco Contò  
Mariantonietta Fiore**

**FrancoAngeli**  
OPEN  ACCESS

Il lavoro si colloca nell'ambito del progetto SKIN – Short supply chain Knowledge and Innovation Network ([www.shortfoodchain.eu](http://www.shortfoodchain.eu)) finanziato dall'Unione Europea con il programma Horizon 2020, bando H2020-RUR-2016-2017 (Rural Renaissance – Fostering innovation and business opportunities), Grant Agreement n. 728055. Capofila Università degli Studi di Foggia – Dipartimento di Economia.

La stampa è stata finanziata dall'Università degli Studi di Foggia – Dipartimento di Economia, Delibera Consiglio di Dipartimento del 20 maggio 2020 punto 5bis, nell'ambito del progetto SKIN per la parte riguardante la stampa cartacea dei volumi, mentre l'edizione Open Access è stata finanziata da Tinada s.r.l. – Spin off dell'Università di Foggia.

*Supervisione scientifica:* prof. Francesco Contò, prof.ssa Mariantonietta Fiore.

La pubblicazione è stata sottoposta a processo di revisione tra pari (peer review).

Hanno curato la collocazione, l'organicità e la revisione dei testi del volume: prof. Francesco Contò, prof.ssa Mariantonietta Fiore.

*Coordinamento editoriale, elaborazioni, segreteria:* Società Tinada s.r.l. – Spin off dell'Università di Foggia.

Copyright © 2020 by FrancoAngeli s.r.l., Milano, Italy.

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore ed è pubblicata in versione digitale con licenza *Creative Commons Attribuzione-Non Commerciale-Non opere derivate 4.0 Internazionale* (CC-BY-NC-ND 4.0)

*L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito*  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it>

# INDICE

<b>Prefazione</b>	pag.	9
-------------------	------	---

## PARTE I AFFASCINANTI SFACCETTATURE E NUOVI RIVERBERI DELLO SVILUPPO

1. Sviluppo sostenibile. Un concetto trasversale, di <i>Leonardo Salvatore Alaimo</i>	»	29
2. Gli strumenti e gli incentivi per la salvaguardia ambientale e la tutela del paesaggio, di <i>Emilia Lamonaca</i>	»	36
3. Il sistema dell'innovazione e della conoscenza in agricoltura. Un'evoluzione in atto, di <i>Raffaele Dicecca</i>	»	47
4. Modelli e metodi di innovazione nel settore agroalimentare, di <i>Gianluigi De Pascale</i>	»	67
5. Il sistema agroalimentare e le perdite agroalimentari. Perdite o risorse? Questo è il problema, di <i>Mariantonietta Fiore</i>	»	80
6. Multifunzionalità e prospettive di sviluppo, di <i>Raffaele Dicecca</i>	»	103
7. La carbon footprint nella filiera agroalimentare, di <i>Roberto L. Rana</i>	»	115

### FOCUS SU... NUTRIZIONE E SANITÀ

1. Salute e alimentazione, di <i>Fiorella Pia Salvatore e Alberto Ametta</i>	»	143
2. Azioni di prevenzione e promozione della salute, di <i>Fiorella Pia Salvatore</i>	»	156
3. L'impatto economico della malnutrizione sul Sistema Sanitario Nazionale, di <i>Carmela Robustella, Giovanni Messina e Fiorenzo Moscatelli</i>	»	167
4. Economia, diritto ed etica nell'assistenza sanitaria, di <i>Prisco Piscitelli</i>	»	176

### FOCUS SU... SISTEMA IMPRESA

1. I sistemi logistici, di <i>Raffaele Silvestri, Savino Santovito e Piermichele La Sala</i>	»	209
2. La gestione del rischio: strumenti pubblici e privati, di <i>Raffaele Silvestri, Savino Santovito e Leonardo Di Gioia</i>	»	226

PARTE II  
PROGETTARE IDEE E FUTURO:  
LE OPPORTUNITÀ DELL'UE

1. La Strategia Europa 2020, di <i>Sara Djelveh e Fedele Colantuono</i>	pag. 245
2. Ricerca e innovazione nell'UE. Il funzionamento dei fondi comunitari e le recenti strategie europee, di <i>Francesco Fera e Fedele Colantuono</i>	» 272
3. Project design e metodologie di management, di <i>Sara Djelveh e Fedele Colantuono</i>	» 298

PARTE III  
VALUTIAMO LO SVILUPPO

1. La valutazione dei programmi di sviluppo rurale attraverso il modello I/O, di <i>Nicola Faccilongo e Leonardo Di Gioia</i>	» 317
2. Metodologie di valutazione, di <i>Nicola Faccilongo</i>	» 336
3. Valutazione dei PSR e modello I/O, di <i>Nicola Faccilongo</i>	» 348

PARTE IV  
ATTREZZI PER L'ANALISI DELLA REALTÀ

1. Analisi di correlazione, di <i>Leonardo Salvatore Alaimo</i>	» 387
2. Indici di concordanza fra valutatori, di <i>Alessia Spada</i>	» 396
3. Analisi delle componenti principali, di <i>Leonardo Salvatore Alaimo e Carlotta Pacifici</i>	» 404
4. L'analisi fattoriale, di <i>Leonardo Salvatore Alaimo e Maria Barbato</i>	» 418
5. Analisi discriminante, di <i>Leonardo Salvatore Alaimo e Federica Nobile</i>	» 433

PARTE V  
“NUOVE” METODOLOGIE E CHIAVI DI LETTURA  
PER LA VALORIZZAZIONE DEL “VECCHIO” STRUMENTO DELLA FILIERA  
CORTA: UNA CARRELLATA DI EVIDENZE EMPIRICHE

1. L'istituzionalizzazione della filiera corta agroalimentare: tra processi di aggregazione strategica e governance territoriale partecipata, di <i>Claudio Nigro e Enrica Iannuzzi</i>	» 451
2. Comunità di pratica: uno strumento per l'agricoltura sostenibile. Il caso SKIN, di <i>Claudia Delicato e Nino Adamashvili</i>	» 478
3. Il progetto EnertMob per una maggiore sostenibilità dei trasporti nella filiera corta, di <i>Antonino Galati, Maria Crescimanno, Marcella Giacomarra, Alessandro Carollo e Antonio Tulone</i>	» 508

4. Prospettive delle filiere corte in Europa attraverso il progetto Smartchain, di <i>Vilma Xhakollari, Marco Medici, Maurizio Canavari, Alessandra Castellini</i>	pag. 523
5. Puglia Km 0 obiettivi e azioni della recente legge regionale pugliese, di <i>Vincenzo Colonna</i>	» 532

## PARTE VI

### FOCUS DI APPROFONDIMENTO SU ASPETTI EMERSI DAL PROGETTO SKIN

Introduzione, di <i>Francesco Contò, Mariantonietta Fiore e Fedele Colantuono</i>	» 543
1. La sostenibilità economica nella filiera corta agroalimentare: il progetto SKIN, di <i>Gianluigi De Pascale, Fedele Colantuono, Sara Djelveh e Francesco Contò</i>	» 545
Appendice – Best practices dal progetto SKIN: diversi approcci nella filiera corta	» 557
2. Scambio di conoscenze universitarie e il progetto SKIN, di <i>Sara Djelveh e Francesco Contò</i>	» 571
Appendice – Best practices dal progetto SKIN: networking e approccio multi-attore	» 588
3. La vendita diretta nel settore vitivinicolo, lezioni dalle cantine pugliesi, di <i>Mariantonietta Fiore</i>	» 591
Appendice – Best practices dal progetto SKIN: la filiera corta nei percorsi enogastronomici	» 610
4. Gli effetti dell’approccio “cloud” nell’amministrazione a filiera corta, di <i>Francesco Contò, Nicola Faccilongo, Massimo Carella e Piermichele La Sala</i>	» 620
Appendice – Best practices dal progetto SKIN: e-commerce e i servizi cloud nella filiera corta	» 638
5. Adozione di strumenti ICT da parte delle imprese agricole in Basilicata, di <i>Gianluigi De Pascale, Piermichele La Sala, Nicola Faccilongo e Claudio Zaza</i>	» 650
Appendice – Best practices dal progetto SKIN: innovazione e tecnologie nella filiera corta	» 665
6. Dalle parole ai fatti – La legge regionale 30 aprile 2018, n. 16 “Norme per la valorizzazione e la promozione dei prodotti agricoli e agroalimentaria km zero e in materia di vendita diretta dei prodotti agricoli”	» 670
7. Dalle parole ai fatti – L’app Orto+, di <i>Federico Angelo Franzese</i>	» 684

## 4. L'ANALISI FATTORIALE

di *Leonardo Salvatore Alaimo*, Istituto Nazionale di Statistica – Istat  
e *Maria Barbato*, Sapienza Università di Roma

Nel seguente capitolo verrà trattata l'analisi fattoriale, tecnica utile per riassumere e semplificare le relazioni di un insieme di variabili. È spesso presentata insieme con l'Analisi in Componenti Principali, da cui però si distanzia per vari aspetti, a partire dallo scopo. In effetti, l'analisi fattoriale si basa su un'ipotesi relativa alla struttura della covarianza o correlazione delle variabili: si assume l'esistenza di un set di variabili latenti che siano idonee a spiegare le interrelazioni tra le variabili originarie. Di contro, l'analisi in componenti principali è una mera trasformazione dei dati che non si basa su nessun tipo di assunzione relativa alla forma della matrice delle covarianze o delle correlazioni da cui origina.

L'analisi fattoriale è, quindi, una metodologia che permette di ridurre la dimensionalità delle variabili trattate, raggruppandole in fattori. Può essere condotta sia a fini esplorativi (EFA, Exploratory Factor Analysis) sia a fini confermativi (CFA, Confirmatory Factor Analysis), che verranno affrontate in seguito.

### 4.1. Definizione, obiettivi e funzionalità

L'analisi fattoriale è una tecnica che riduce la dimensionalità ma non l'informazione utile derivante dall'analisi, permettendo di identificare uno o più “fattori” o “dimensioni latenti” – auspicabilmente tra loro indipendenti - che rendono conto delle similarità che accomunano una serie di variabili. Il risultato della sua applicazione è che pochi fattori “latenti” servono a rendere

ragione delle informazioni contenute in insiemi decisamente più ampi di variabili osservate.

L'analisi fattoriale si utilizza tutte le volte che si studiano fenomeni non direttamente osservabili (soddisfazione, capacità, abilità, etc.), quando si vuole indagare o verificare se la struttura teorica di un test corrisponde a quella empiricamente raggiunta. Non si tratta di un metodo di predizione delle variabili, piuttosto ha un duplice obiettivo: il primo è la riduzione dell'informazione di un insieme di dati; il secondo, è l'individuazione di "dimensioni latenti" che spiegano le relazioni tra le variabili che non emergono direttamente, per valutarne la forza esplicativa. Si parte dalla creazione di una matrice di correlazioni tra le variabili osservate per giungere alla *matrice delle saturazioni* che contiene una misura della relazione tra le variabili osservate ed i fattori "latenti".

## 4.2. La formalizzazione del modello

Data una variabile  $X_i = (X_{i1}, X_{i2}, \dots, X_{ij}, \dots, X_{ip})$  quale variabile di interesse, il modello fattoriale si definisce come (1):

$$\forall_{j=1, \dots, p} X_j = \mu_j + \lambda_{j1}f_1 + \dots + \lambda_{jm}f_m + u_j \quad (1)$$

dove:

- $\mu_j$  è l'intercetta,
- $\lambda_j$  determina l'influenza delle variabili esplicative,
- $f_m$  rappresenta le variabili latenti,
- $u_j$  rappresenta il fattore specifico, ovvero l'errore residuo che cambia al cambiare dell'equazione.

L'analisi fattoriale è differente dalla regressione lineare. Innanzitutto, si suppone che tutti gli elementi della variabile aleatoria di interesse  $X$  siano generati dagli stessi fattori comuni. Inoltre, nell'analisi fattoriale si osservano solo le  $X$  variabili di interesse e non le  $f$  variabili latenti al contrario di quanto avviene nel modello di regressione lineare.

Il modello formale in base al quale le variabili manifeste e latenti sono collegate è un modello di regressione multipla in cui ogni variabile osservata è il risultato della regressione sui fattori comuni. I coefficienti di regressione del modello sono definiti *factor loadings* e il termine di errore casuale è

definito *variazione specifica* e rappresenta la parte di una variabile osservata non spiegata dai fattori comuni. Si consideri il seguente sistema:

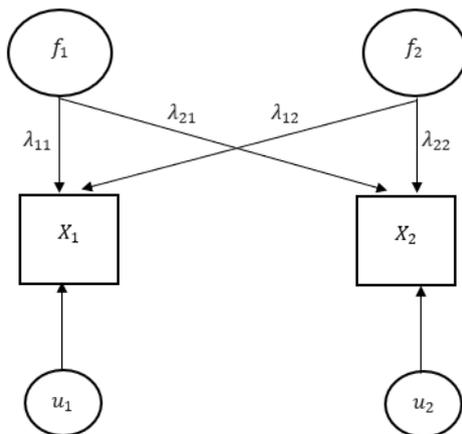
$$X_1 = \mu_1 + \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1m}f_m + u_1$$

$$X_2 = \mu_2 + \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2m}f_m + u_2$$

$$X_p = \mu_p + \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pm}f_m + u_p \quad (2)$$

Dal momento che i fattori sono latenti (non direttamente osservabili) si assume che si presentino nella forma di valori standardizzati con media nulla e varianza unitaria. Si assume anche che siano *incorrelati* e che non vi sia autocorrelazione tra i residui ed i fattori comuni. Ciò implica che le correlazioni tra le variabili osservate siano spiegabili in termini della loro relazione con i fattori comuni. Tali assunzioni verranno meglio specificate successivamente.

Fig. 1 – Rappresentazione grafica del modello di analisi fattoriale



Considerando solo le prime due variabili  $X_1$  e  $X_2$  e supponendo che vi siano solo due fattori  $f_1$  e  $f_2$ , si avrebbe la rappresentazione grafica del modello di analisi fattoriale mostrata in Figura 1.

I fattori latenti  $f_1$  e  $f_2$ , influenzano direttamente le variabili osservate  $X_1$  e  $X_2$  e questa influenza diretta è rappresentata dalle frecce unidirezionali che vanno dalle  $f$  alle  $X$  mentre l'intensità dell'influenza è rappresentata dai  $\lambda$ .

Le  $u$ , invece, rappresentano la variabilità residua. Al fine di ricavare informazioni su  $f_1$  si individua il legame tra le variabili  $X_1$  e  $X_2$ , tale legame è dato dalla loro covarianza (3):

$$\text{Cov}(X_1, X_2) = \lambda_{11}\lambda_{21} \text{Var}(f_1)$$

Tanto più la covarianza tra le due variabili è elevata (e di conseguenza il valore dei  $\lambda$  è elevato in valore assoluto), tanto più è probabile che le due variabili siano generate dal medesimo fattore. Il modello è di tipo *lineare* (4):

$$\forall_{j=1,\dots,p}; \forall_{m=1,\dots,M}; \text{con } M \ll p \quad (3)$$

$$X_j = \mu_j + \sum_{m=1}^M \lambda_{jm} f_m + u_j \quad (4)$$

e si basa sulle seguenti ipotesi:

- $E[X_j] = \mu_j$  e  $E[f_m] = E[u_j] = 0$
- $\text{Var}(f_m) = 1$  e  $\text{Var}(u_j) = \varphi_j$
- $\text{Cov}(f_m, u_j) = 0$  e  $\text{Cov}(u_j, u_l) = 0$

In forma matriciale, il modello è il seguente (5):

$$\begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} + \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_{p1} & \cdots & \lambda_{pm} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix} \quad (5)$$

mentre in forma contratta (6):

$$X = \mu + \Lambda F + U \quad (6)$$

dove le  $F$  sono i fattori comuni mentre le  $U$  sono i fattori specifici e tutte le quantità sono ignote. Pertanto, il modello non può essere stimato. Inoltre:

$$E(X) = \mu + \Lambda E(F) + E(U) \quad (7)$$

dove  $E(X) = \mu$  e di conseguenza anche  $E(F) = 0$  e  $E(U) = 0$ .

Alle ipotesi precedenti vanno, quindi, aggiunte le seguenti:

- $F$  è una variabile casuale multinormale con il vettore di medie 0 e matrice di varianze e covarianze  $I_m$ ;

- $U$  è una variabile casuale multinormale con vettore di medie 0 e matrice di varianze e covarianze  $\varphi$ ;
- le variabili casuali  $F$  e  $U$  sono indipendenti.
- $X$  è una variabile casuale multinormale. Il suo vettore di medie è pari a  $\mu$  e la sua matrice di varianze e covarianze si scrive in funzione di  $\Lambda$  e  $\varphi$ .

Poiché i fattori di nostro interesse sono latenti, i *factor loadings*, ovvero i coefficienti del modello, non possono essere stimati nello stesso modo in cui sono stimati i coefficienti di regressione nel modello di regressione multipla. Ci sono diversi metodi di stima e tra questi si annoverano l'analisi dei fattori principali e l'analisi di massima verosimiglianza dei fattori.

### 4.3. La matrice di varianze e covarianze

Dalle precedenti ipotesi, discendono diverse conseguenze. Tra queste: la matrice  $\Lambda$  non è identificata. Si conosce però, dalle prime tre ipotesi, che:

$$\text{Var}(X_j) = \sigma_j^2 = \lambda_{j_1}^2 \text{Var}(f_1) + \dots + \lambda_{j_m}^2 \text{Var}(f_m) + \text{Var}(u_j) = \quad (8)$$

e, valendo l'ipotesi (2):

$$= \sum_{m=1}^M \lambda_{j_m}^2 + \varphi_j^2 = c_j + \varphi_j^2 \quad (9)$$

dove:

- $c_j = \sum_{m=1}^M \lambda_{j_m}^2$  è la *comunalità*, rappresenta la parte di varianza “spiegata” dai fattori comuni.
- $\varphi_j^2$  è la *varianza specifica* ed è la parte residua della varianza di  $X_j$ , non spiegata in termini di variabilità dei fattori comuni.

In termini di covarianze tra le  $X_j$  avremo che, per  $j \neq s$ :

$$\begin{aligned}
\text{Cov}(X_j, X_s) &= \text{Cov}\left(\sum_{m=1}^M \lambda_{jm} f_m + u_j, \sum_{m=1}^M \lambda_{sm} f_m + u_s\right) \\
&= \sum_{m_1=1}^M \sum_{m_2=1}^M \lambda_{jm_1} \lambda_{sm_2} \text{Cov}(f_{m_1}, f_{m_2}) + \sum_{m=1}^M \lambda_{jm} \text{Cov}(f_m, u_s) \\
&\quad + \sum_{m=1}^M \lambda_{sm} \text{Cov}(f_m, u_j) + \text{Cov}(u_j, u_s) = \\
&= \sum_{m=1}^M \lambda_{jm_1} \lambda_{sm} \text{Var}(f_m) = \sum_{m=1}^M \lambda_{jm} \lambda_{sm} \tag{10}
\end{aligned}$$

Da questa relazione discende un risultato particolarmente interessante poiché si nota come in realtà la covarianza tra due generiche  $X_j$  e  $X_s$  sia interamente dovuta ai fattori comuni. Pertanto, è fondamentale cercare di spiegare la struttura di dipendenza lineare di queste variabili, ovvero la loro matrice di varianze e covarianze, in termini di presenza di fattori comuni.

In altre parole, le variabili di interesse sono tra loro interdipendenti perché dipendono dagli stessi fattori comuni, dalle stesse variabili inosservabili. In termini matriciali:

$$\Sigma = \Lambda \Lambda' + \varphi \tag{11}$$

dove  $\Sigma$  rappresenta la matrice di varianze e covarianze. Tuttavia, non conosciamo la matrice  $\Lambda$  che la determina, pertanto procederemo nel seguente modo. Si consideri una matrice ortogonale  $Q$  di ordine  $h$ , posto che  $\Lambda^* = \Lambda Q$  e fissate  $\Sigma$  e  $\varphi$ , si avrà:

$$\Lambda^* \Lambda^{*'} = \Lambda Q Q^1 \Lambda' = \Lambda \Lambda' \tag{12}$$

quindi:

$$\Sigma = \Lambda^* \Lambda^{*'} + \varphi \tag{13}$$

ciò implica che, alle stesse  $\Sigma$  e  $\varphi$  corrispondono diverse matrici  $\Lambda$ , tutte in grado di soddisfare  $\Sigma = \Lambda \Lambda' + \varphi$ . Si deduce quindi che, date  $\Sigma$  e  $\varphi$ ,  $\Lambda$  è determinata solo a meno di una moltiplicazione (a destra) per la matrice ortogonale.

## 4.4. I passi fondamentali dell'analisi fattoriale

L'applicazione dell'analisi fattoriale richiede alcuni passaggi fondamentali. Innanzitutto, vi è una fase di valutazione delle assunzioni che considera sia l'adeguatezza delle variabili, sia la valutazione della fattorializzabilità della matrice di correlazione. Successivamente, si sceglie il metodo di estrazione dei fattori, il numero dei fattori che si vogliono considerare ed il metodo di rotazione dei fattori. Infine, si procede con l'interpretazione e la validazione della soluzione fattoriale a cui si è giunti. Analizziamo le fasi più nel dettaglio.

- La *valutazione dell'adeguatezza delle variabili* riguarda il livello di misurazione e la forma della distribuzione (asimmetria e curtosi<sup>1</sup>). Si studia la scala di misurazione necessaria e valuta l'opportunità di ricorrere a trasformazioni di una o più variabili. Si tende a trasformare una variabile per migliorare la simmetria delle distribuzioni. La forma delle distribuzioni dei dati si approssima maggiormente ad una Normale (Gaussiana).

I casi outliers<sup>2</sup> che spesso si presentano in virtù dell'asimmetria della distribuzione possono essere rimossi. Si favorisce il confronto tra più distribuzioni multiple, rendendo la variabilità più simile. A completamento di questa fase preparatoria occorre verificare che le relazioni tra le variabili siano lineari. Se la distribuzione delle variabili è conforme alla normale multivariata, allora le relazioni tra le variabili sono sicuramente lineari. A tal fine, è necessario calcolare la *distanza di Mahalanobis*, che rappresenta il punto di partenza per il successivo calcolo del *coefficiente di Mardia*, per la rappresentazione dei quantili e per l'individuazione degli outliers multivariati.

- Per la *valutazione della fattorializzabilità della matrice di correlazione* è necessario calcolare, oltre ad alcune statistiche univariate, il *determinante della matrice di correlazione*, il *test di Keiser Meyer Olkin-KMO*, il *test di sfericità di Bartlett* e la matrice anti-immagine.

Il valore del determinante di una matrice deve essere diverso da zero affinché non ci siano variabili linearmente dipendenti. Se risulta pari a zero significa che c'è almeno una variabile che è combinazione lineare perfetta di

<sup>1</sup> Asimmetria e curtosi sono due aspetti della forma di una distribuzione statistica. In generale, per asimmetria si intende l'assenza di specularità di una distribuzione statistica rispetto a qualsiasi asse verticale; la curtosi fa riferimento alla maggiore o minore gibbosità di una curva rispetto al suo massimo.

<sup>2</sup> Il termine outlier si riferisce a un valore anomale e distante rispetto agli altri presenti in una distribuzione statistica.

altre variabili in analisi. In tal caso la matrice di correlazione non ammette inversa e, quindi, l'analisi fattoriale non può essere effettuata.

Il test di adeguatezza campionaria KMO (definire acronimo – *Test* di adeguatezza campionaria di Kaiser-Meyer-Olkin (Kaiser, 1974) consente di confrontare la grandezza delle correlazioni osservate rispetto alle correlazioni parziali. La seguente tabella.

Valori soglia	Interpretazione
>0,90	Eccellenti
Tra 0,80 e 0,90	Buoni
Tra 0,70 e 0,80	Accettabili
Tra 0,60 e 0,70	Mediocri
<0,60	Scarsi

Si può calcolare anche sulle singole variabili e i valori corrispondenti sono contenuti sulla diagonale principale della matrice di correlazione anti-immagine.

La matrice delle correlazioni anti-immagine contiene il complemento a 1 dei coefficienti di correlazione parziale. I coefficienti parziali sono calcolati eliminando da ogni coppia di variabili analizzate la varianza comune condivisa con tutte le altre variabili selezionate per l'analisi fattoriale. Quindi, se le correlazioni parziali sono basse, le correlazioni anti-immagine saranno alte e le variabili saranno apprezzabilmente correlate tra loro. Dunque, il modello fattoriale è buono.

Il test di sfericità di Bartlett verifica l'ipotesi nulla che la matrice di correlazione sia una matrice identità (valori di significatività inferiori a 0,05 indicano che la matrice di correlazione è significativamente diversa da una matrice di identità e quindi si rifiuta l'ipotesi nulla).

In questa fase, inoltre, si osserva come le variabili sono correlate tra loro, evidenziando alcuni gruppi interni di variabili più correlate tra loro rispetto ad altre. Si controlla la significatività statistica e, in ultimo, si ipotizza il numero di fattori da estrarre.

1. Esistono diversi metodi per la scelta del *metodo di estrazione dei fattori* tra cui: la fattorizzazione dell'asse principale, il metodo di massima verosimiglianza ed i minimi quadrati generalizzati, i quali richiedono che le variabili si conformino alla distribuzione normale multivariata; l'analisi in componenti principali e i minimi quadrati ordinari non richiedono invece nessuna assunzione.

2. Per quel che concerne la scelta del numero di fattori si può ricorrere a diversi metodi:
  - L'ispezione dello scree-plot (il grafico degli autovalori)
  - L'esame dello scree-test degli autovalori nella matrice delle correlazioni residue (Matrice riprodotta)
  - L'esame degli autovalori della matrice di correlazione, la percentuale di varianza spiegata da ognuna delle componenti principali calcolabili, la percentuale cumulata di varianza spiegata dalle componenti
  - L'esame degli autovalori della matrice ridotta limitatamente ai fattori comuni richiesti
3. La scelta del *metodo di rotazione dei fattori* ha come obiettivo quello di rendere la soluzione fattoriale più agevolmente interpretabile. I fattori vengono traslati nello spazio fattoriale rispetto all'originaria posizione. Ciò che cambia sono le saturazioni delle variabili originarie su fattori ruotati, quindi la percentuale di varianza spiegata dal singolo fattore non risulterà più omogenea. Esistono 5 diversi tipi di rotazione:
  - *Varimax*, *Quartimax* e *Equamax* sono soluzioni ortogonali;
  - *Promax* e *Oblimin* sono soluzioni oblique.
4. Il *livello interpretativo* richiede lo studio della percentuale di varianza spiegata e delle comunaltà delle variabili così da aiutare il ricercatore a comprendere la capacità della soluzione cui è pervenuto. Le saturazioni fattoriali della soluzione ruotata aiutano il ricercatore a dare un nome ai fattori che sono emersi dall'analisi.
5. L'ultimo step è quello della *validazione della soluzione*. A tal fine, si possono considerare elementi interni o esterni alla soluzione oppure si può far ricorso ai punteggi fattoriali ma questi sono disponibili direttamente solo per l'ACP, per l'AF occorre stimarli tramite regressione. L'interpretazione del coefficiente di determinazione fattoriale che si ricava sulla diagonale principale della matrice delle covarianze del punteggio fattoriale, varia tra 0 e 1. Valori superiori a 0,70 indicano che è elevata la coerenza interna del fattore e la qualità della misurazione.

## 4.5. Classificazione

L'analisi fattoriale si può condurre sia con scopi esplorativi (EFA, Explanatory Factor Analysis) sia con scopi confermativi (CFA, Confirmatory Factor Analysis).

Nel primo caso, i fattori vengono estrapolati a partire dai dati. Si tratta di un metodo non-condizionale che genera una nuova struttura fattoriale non formulata a priori, implica *induzione esplorativa* e considera un certo grado di affidabilità dei dati nell'indurre ipotesi su una struttura plausibile.

Di contro, nell'analisi fattoriale confermativa, è il ricercatore che pone dei vincoli sul proprio modello e verifica se tale modello sia coerente con i dati osservati. Si tratta, quindi, di un metodo condizionale che conferma le strutture fattoriali formulate a priori, che implica *logica deduttiva* e considera un certo grado di affidabilità dei dati nel riprodurre strutture plausibili.

#### 4.5.1. Analisi fattoriale esplorativa

L'analisi fattoriale esplorativa ha l'obiettivo di estrarre i fattori a partire dai dati, come detto in precedenza. Tra le tecniche più note di estrazione dei dati vi sono: il metodo dei fattori principali, la fattorizzazione per componenti principali, la stima di massima verosimiglianza.

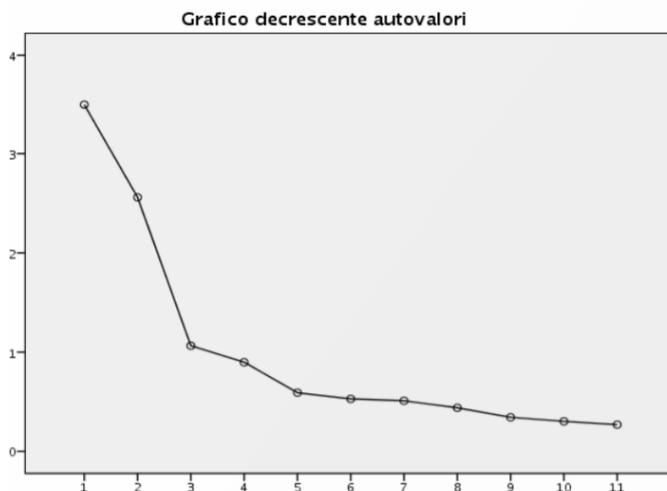
Il numero di fattori da considerare ai fini dell'analisi dipende dal problema che si intende affrontare e può essere scelto secondo diversi criteri:

- Il *criterio di Kaiser*: si considerano tutti i fattori il cui autovalore (la quota di variabilità "spiegata" dal fattore, che assume valori discendenti man mano che dal primo fattore ci si sposta verso l'ultimo) sia superiore o uguale a 1.
- L'analisi dello *scree-plot*: stabilisce il numero di fattori per via grafica andando a cercare il punto di flesso del grafico. Lo *scree plot* rappresenta gli autovalori inordinata ed il numero delle componenti in ascissa.

Al fine di determinare il numero di componenti ottimali per l'analisi si ricerca il punto di flesso del grafico. Tuttavia, vi sono due criteri che permettono di scegliere tale punto:

- secondo il *criterio di Harmann*, si considera l'ultimo punto prima del quale la curva dello scree plot inizia ad appiattirsi. Nell'esempio di scree plot sopra riportato, secondo il criterio di Harmann, il numero di componenti ottimali sarebbe pari a due perché a partire dalla terza componente la curva si appiattisce.
- secondo il *criterio di Cattell*, si considerano le componenti fino al punto in cui la curva si appiattisce, dunque si include, nel caso sopra riportato, anche la terza componente.

Fig. 2 – Esempio scree-plot



La *parallel analysis*: prevede un “doppio scree plot” sul quale sono rappresentati gli autovalori dei fattori determinati e uno stesso numero di fattori casuali. Il numero di dimensioni viene stabilito ricercando quale sia la dimensione dove gli autovalori “reali” diventino più piccoli di quelli “casuali”.

Al fine di giungere alla soluzione fattoriale, ovvero al set di fattori e di relazioni tra variabili e fattori che rappresenta la soluzione al problema fattoriale, è necessario considerare i seguenti parametri:

- la quantità di *variabilità (varianza) spiegata* dal complesso dei fattori considerati e da ciascun singolo fattore;
- la  *saturazione* (factor loading), che descrive la forza della relazione tra il fattore e la variabile misurata (al di sotto della soglia di 0,30/0,40 si esclude l’esistenza della relazione);
- la *comunalità* (communality): esprime la parte di varianza riprodotta da un numero ridotto di fattori. È la somma dei coefficienti di determinazione di una variabile riprodotta dai fattori. La comunalità iniziale di ogni componente è pari a 1 poiché ogni componente spiega sé stessa e le variabili sono standardizzate. Sarebbe inferiore ad uno solo nel caso di *fattori comuni* perché vi sarebbe un fattore di unicità non spiegabile dai fattori;
- l’*unicità* (uniqueness): complemento a 1 della comunalità.

### 4.5.2. Analisi fattoriale confermativa

L'analisi confermativa utilizza modelli di equazioni strutturali. È il ricercatore che identifica la struttura fattoriale, quindi definisce degli indici che analizzano quanto il modello che stima le saturazioni si adatti ai dati.

Tali indici si suddividono in:

- Indici assoluti
  - *Root Mean Square Residual* (RMR), ovvero la radice quadrata della media dei residui al quadrato ed è ricavabile dalla seguente formula:

$$RMR = \sqrt{2 \sum_i \sum_j \frac{(r_{ij} - \hat{r}_{ij})^2}{(q(q+1))}} \quad (14)$$

dove  $q$  è il numero totale delle variabili,  $r_{ij}$  è la correlazione osservata e  $\hat{r}_{ij}$  è la correlazione riprodotta tra le variabili  $i$  e  $j$ . L'RMR indica la media della correlazione residua, cioè non spiegata dal modello. Indica un buon adattamento per valori  $<0,5$ .

- *Root Mean Square Error of Approximation* (RMSEA): è la radice quadrata dell'errore di approssimazione. Valori  $<0,05$  suggeriscono che l'errore di approssimazione è minimo, valori compresi tra  $0,05$  e  $0,08$  implicano un errore accettabili, valori  $>0,08$  indicano che il modello fattoriale non regge. La formula relativa è la seguente:

$$RMSEA = \sqrt{\frac{\chi^2 - df}{N} \frac{1}{df}} \quad (15)$$

dove  $df$  sono i gradi di libertà mentre  $N$  è il numero dei soggetti. Corrisponde alla bontà di adattamento del modello nella popolazione, ponderata per i gradi di libertà del modello.

- *Indici relativi*: sono indici che valutano l'adeguatezza del modello fattoriale del ricercatore rispetto ad un modello nullo, che dà un valore di *baseline*. Nel modello nullo si ipotizza che non vi siano relazioni tra le variabili.
  - *Tucker and Lewis Index* (TLI):

$$TLI = \frac{\left(\frac{\chi^2_{null0}}{df_{null0}}\right) - \left(\frac{\chi^2_{target}}{df_{target}}\right)}{\left(\frac{\chi^2_{null0}}{df_{null0}}\right) - 1} \quad (16)$$

dove:

- $\chi^2_{null0}$  all'interno di un modello in cui non si definisce una struttura dei dati, rappresenta il limite inferiore del *fit*.
- $\chi^2_{target}$  definisce la bontà di adattamento nel modello ideato dal ricercatore.

I valori *df* sono rispettivamente i gradi di libertà del modello nullo e del modello target.

Un coefficiente del  $TLI > 0,9$  indica un buon adattamento del modello ai dati empirici.

- *Comparative Fit Index* (CFI):

$$CFI = 1 - \frac{\max(\chi^2_{target} - df_{target}, 0)}{\max(\chi^2_{null0} - df_{null0}, \chi^2_{target} - df_{target}, 0)} \quad (17)$$

dove i termini sono uguali a quelli definiti nel TLI. Rispetto al TLI, il CFI stima la bontà di adattamento *nella popolazione* e non su una sua partizione. Un coefficiente del  $CFI > 0,9$  indica un buon adattamento ai dati.

- *Indici comparativi*

- *Akaike's Information Criterion* (AIC): è un criterio che confronta e valuta diversi modelli. La formula che lo caratterizza è la seguente:

$$AIC = 2k - 2\ln(L) \quad (18)$$

dove *k* è il numero di parametri nel modello ed *L* è il valore massimizzato della funzione di verosimiglianza del modello stimato.

- *Bayesian Information Criterion* (BIC): è un metodo per selezionare un modello tra una classe di modelli parametrici che hanno diversi parametri. La formula che si ottiene è la seguente:

$$BIC = -2\ln(L) + k \ln(n) \quad (19)$$

dove  $k$  e  $L$  valgono esattamente come per l'AIC mentre  $n$  è il numero di osservazioni.

## 4.6 Confronto AF e ACP

Come già fatto presente ad inizio di questo capitolo, l'Analisi Fattoriale è spesso trattata di pari passo con l'Analisi in Componenti Principali (trattata al capitolo 3).

Si tratta di due tecniche che hanno come scopo la riduzione della complessità di un set di variabili tuttavia, il loro utilizzo si differenzia a seconda dell'obiettivo del ricercatore. Se il ricercatore è interessato ad un modello di sintesi dei dati che però prescinda da particolari assunzioni, allora la semplicità teorica dell'analisi delle componenti principali è preferibile. Proprio per questa sua semplicità teorica, l'ACP si adatta bene anche alla ricerca di strutture latenti, pur non ottenendo i medesimi risultati dell'analisi fattoriale. In effetti, se il ricercatore è interessato ad all'individuazione di variabili latenti che spieghino le relazioni implicite tra le variabili originarie, si predilige l'analisi fattoriale poiché ha un costrutto di ipotesi alla base che rende i risultati più robusti.

Una seconda differenza si riscontra al momento della formalizzazione dei due modelli. Se, da una parte l'ACP si presenta come una combinazione lineare delle variabili originarie al fine di individuare le componenti, l'analisi fattoriale si comporta diversamente: le variabili di interesse sono le variabili originarie e queste sono spiegate proprio dai fattori. Nel dettaglio, le componenti dell'ACP sono inizialmente note poiché espresse in funzione della matrice dei dati osservati, il processo di stima si riferisce invece alla massimizzazione della varianza di ciascuna componente. Nell'analisi fattoriale, invece, i termini noti sono le variabili originarie a sinistra dell'equazione; il resto dei termini risulta inizialmente ignoto e, quindi, da stimare.

Una volta formalizzato il modello, l'attenzione di entrambe le analisi si rivolge alla matrice di varianze e covarianze. Tuttavia, l'ACP si concentra prettamente sugli elementi diagonali di tale matrice, mentre, l'analisi fattoriale si focalizza sulle covarianze, quindi sui termini extra-diagonali.

Un'ultima differenza tra le due analisi discende dal comportamento che assumono in seguito all'aumento della dimensionalità. Se la dimensione aumenta, nell'ACP, le componenti che già facevano parte del modello non verranno influenzate dalle nuove. Nell'analisi fattoriale, invece, l'aumento del numero di fattori, porta alla completa ridefinizione di tutti i fattori.

In conclusione, quindi, dovrebbe essere chiaro da quest'ultimo paragrafo che l'ACP non è migliore dell'analisi fattoriale e viceversa perché l'applicazione di un modello non va a scapito dell'altro.

## Riferimenti bibliografici

- Barbaranelli, C. (2007). *Analisi dei dati. Tecniche multivariate per la ricerca psicologica e sociale*. 2<sup>a</sup> ed. Milano: LED, pp. 119-203.
- Cattell, R.B. (1978). *The scientific use of factor analysis*. New York, Plenum Press.
- Harman, H.H. (1976). *Modern factor analysis* (3<sup>rd</sup> ed.), Chicago, University of Chicago Press.
- Kaiser, H. (1974). An index of factor simplicity. *Psychometria*, 39, 31-36.
- Norman, G. R., & Streiner, D. L. (1998). *PDQ Statistics*. Pmph USA Ltd.
- Vitali, O. (1993). *Statistica per le scienze applicate*. Volume secondo. Cacucci, Bari, pp. 579-643.

Il presente manuale è una raccolta collettanea di contributi che presentano una nuova chiave di lettura dello sviluppo locale, declinando in prospettiva scientifico-divulgativa “nuove” teorie e tematiche classiche che hanno costituito da sempre lo schema di ciò che normalmente viene identificato sviluppo locale. Vi è, quindi, una multiformità di tematiche che potrebbero, a primo impatto, apparire eterogenee e distanti dagli obiettivi di sviluppo: come il tema della salute e della sanità. Il Covid-19 ha confermato, però, che immaginare oggi uno sviluppo locale a prescindere dalle tematiche sanitarie di un territorio rappresenta sicuramente un’assurdità. Ma nelle vecchie teorie dello sviluppo locale questa tematica non è riscontrabile. Questa declinazione dello sviluppo locale è inserita all’interno della nuova visione della cosiddetta “Economia di Francesco” che è il chiaro riferimento all’evoluzione dell’economia civile di A. Genovesi arricchita dalle suggestioni francescane e benedettine compendiate nella “Laudato Sii” di San Francesco e nella “Regola” di San Benedetto. L’occasione di questo manuale è data dalla presentazione degli atti del progetto SKIN (Short supply chain Knowledge and Innovation Network, finanziato nell’ambito del programma Horizon 2020) che ha avuto come Lead Partner l’Università di Foggia (Dipartimento di Economia) e si è appena concluso dando alla luce un’importante rete tematica europea sulla filiera corta.

**Francesco Contò** è professore ordinario presso l’Università di Foggia, Dipartimento di Economia. È stato Direttore del Dipartimento di Economia e coordinatore di corsi di dottorato. Dal 1977 gli sono state affidate importanti posizioni scientifico-accademiche e professionali. È autore di oltre 270 pubblicazioni nazionali e internazionali. Ha ricoperto incarichi di ricerca e insegnamento presso università e centri di ricerca qualificati e attualmente è direttore scientifico di alcuni importanti laboratori di ricerca regionali. È coordinatore di numerosi progetti di ricerca europei, nazionali e regionali.

**Mariantonietta Fiore** è professore associato presso l’Università di Foggia, Dipartimento di Economia. È membro della Scuola di dottorato internazionale della Warsaw University of Life Sciences e del Board of Directors dell’International Food and Agribusiness Management Association. È Fellow dell’EuroMed Academy of Business. È stata vicecoordinatrice scientifica del progetto SKIN ed esperto junior del Ministero dell’Ambiente: attualmente è responsabile o membro di progetti scientifici, gruppi di ricerca e comitati editoriali internazionali e nazionali. Ha ricevuto oltre dieci premi scientifici.