

## Lezione 10

L'analisi fattoriale: il modello di base

### Argomenti della lezione:

- ➔ **Concetti fondamentali**
- ➔ **Equazioni fondamentali**

### Scopo dell'Analisi Fattoriale

studiare le relazioni in un insieme di variabili per:

- ➔ ridurre l'informazione
- ➔ individuare "dimensioni latenti"

### Punto di partenza: matrice delle correlazioni tra le variabili (R)

	1	2	3	4	5	6
1.Determinato	1.00	.60	.55	.37	.36	.40
2.Dinamico	.60	1.00	.66	.40	.43	.44
3.Energico	.55	.66	1.00	.47	.50	.51
4.Affidabile	.37	.40	.47	1.00	.79	.67
5.Responsabile	.36	.43	.50	.79	1.00	.77
6.Scrupoloso	.40	.44	.51	.67	.77	1.00

### Punto di arrivo: matrice delle saturazioni

	F1	F2
Determinato	.63	.51
Dinamico	.74	.43
Energico	.78	.33
Affidabile	.30	-.41
Responsabile	.34	-.43
Scrupoloso	.32	-.33

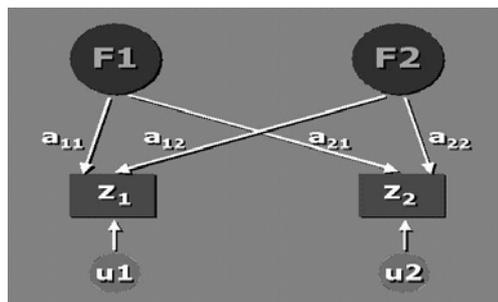
### Modello teorico dell'analisi fattoriale

Esame della varianza che le variabili hanno in comune

#### Ipotesi di base:

la correlazione tra le variabili è determinata da dimensioni non osservabili che "causano" le variabili osservate

## Modello dell'analisi fattoriale



### F: Fattori Comuni

- Rappresentano la variabilità condivisa tra le variabili in analisi
- Influenzano più di una variabile

### a: Saturazioni

- Relazioni tra variabili e fattori

### u: unicità della variabile

- Parte di varianza non condivisa
- Dovuta a cause sistematiche specifiche, o all'errore casuale di misurazione

### Punteggio (standardizzato) di un soggetto in una variabile:

somma "ponderata" del punteggio ottenuto dallo stesso soggetto:

- nei fattori comuni
- in una componente unica

### Equazione del modello teorico alla base dell'analisi fattoriale:

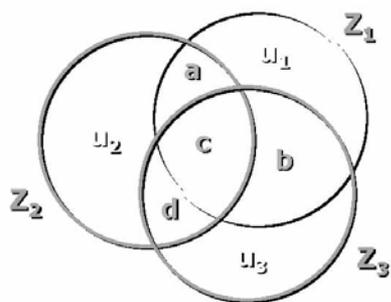
$$z_{ik} = F_{i1}a_{k1} + F_{i2}a_{k2} + \dots + F_{im}a_{km} + u_{ik}$$

### Espressione matriciale:

$$Z = FA' + U$$

- Z: punteggi nelle variabili
- F: punteggi nei fattori comuni
- A: matrice delle saturazioni
- U: matrice unicità

## Varianza unica e varianza comune



**Scomposizione della  
varianza di ogni variabile  
osservata:**

$$\text{Varianza totale} = 1 = h^2 + u^2$$

**Comunalità:  $h^2$**

Parte di varianza totale  
spiegata dai fattori comuni

**Unicità (varianza unica):**

$$u^2 = 1 - h^2$$

Parte di varianza totale  
non spiegata dai fattori comuni

**Assumiamo che:**

$$\text{Cov}(u_i, F_j) = 0$$

$$\text{Cov}(u_i, u_j) = 0$$

$$\text{Cov}(F_i, F_j) \neq 0$$

solo nell'analisi "obliqua"

In generale

$$\text{Cov}(F_i, F_j) = 0$$

**Applicando alcuni passaggi  
di calcolo matriciale si ha che:**

$$R = AA' + U^2$$

**A:** matrice delle saturazioni nei  
fattori comuni

**$U^2$ :** matrice delle varianze uniche  
di ogni variabile

→  **$AA'$ :** rende ragione degli elementi  
fuori della diagonale principale,  
e della comunalità di ogni variabile

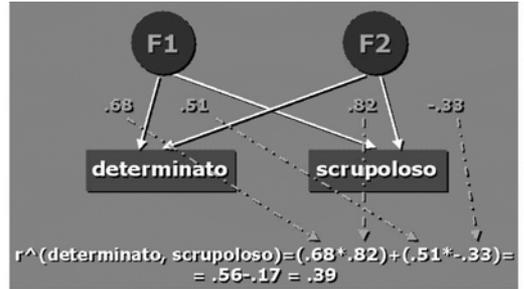
→  **$U^2$ :** contribuisce a rendere ragione  
degli elementi sulla diagonale  
principale di  $R$

**La correlazione tra due variabili (i e j) può essere riprodotta dalla somma dei prodotti delle loro saturazioni nei fattori comuni:**

$$r_{ij} = a_{i1}a_{j1} + \dots + a_{im}a_{jm} = \sum a_{ir}a_{jr} \quad (i \neq j)$$

$$r_{ii} = a_{i1}a_{i1} + \dots + a_{im}a_{im} + u_i^2 = \sum a_{ir}a_{ir} + u_i^2$$

### Correlazioni riprodotte



### Correlazione residua

**Differenza tra la correlazione osservata e la correlazione riprodotta tramite le saturazioni**

$$r(\text{determinato, scrupoloso}) = .40$$

$$r^{\wedge}(\text{determinato, scrupoloso}) = .39$$

$$r_{\text{residua}}(e) = .40 - .39 = .01$$

$$R = AA' + U^2$$

**“equazione fondamentale della analisi fattoriale” (Thurston)**

**Mette in relazione il punto di partenza con il punto di arrivo**

→ **Per riprodurre le correlazioni tra le variabili, e le comunaltà, sono necessari solo i fattori comuni**

→ **Per riprodurre la varianza totale delle variabili sono necessarie anche le unicità**

### Equazione fondamentale:

$$R^* = AA' \quad (1)$$

$$R = AA' + U^2 \quad (2)$$

**$R^*$  è la matrice delle correlazioni che contiene le comunaltà sulla diagonale principale**

**Come ricavare A, la matrice delle saturazioni nei fattori comuni, in maniera tale che il numero di fattori comuni sia strettamente minore del numero di variabili osservate**

### Componenti principali

**Calcolo di alcuni elementi che caratterizzano R:**

→ **radici caratteristiche (autovalori, L)**

→ **vettori ad essi associati (autovettori, V)**

**Elementi che sintetizzano l'informazione relativa alla varianza delle variabili, e alla correlazione tra le variabili**

**Passo preliminare per il calcolo delle soluzioni di analisi fattoriale**

### Autovalori e autovettori (1)

$$\begin{pmatrix} 1.84 & 0 & 0 \\ 0 & .70 & 0 \\ 0 & 0 & .46 \end{pmatrix} = \begin{pmatrix} -.54 & -.63 & -.57 \\ .76 & -.08 & -.64 \\ .36 & -.78 & .52 \end{pmatrix} * \begin{pmatrix} 1 & .45 & .30 \\ .45 & 1 & .50 \\ .30 & .50 & 1 \end{pmatrix} * \begin{pmatrix} -.54 & .76 & .36 \\ -.63 & -.08 & -.78 \\ -.57 & -.64 & .52 \end{pmatrix}$$

$$L = V' * R * V$$

$$L = V' * R * V$$

**Questa operazione non distorce R, semplicemente la "riasmonta"**

**La varianza di R è condensata nei suoi autovalori**

### Autovalori e autovettori (2)

$$\begin{pmatrix} 1 & .45 & .30 \\ .45 & 1 & .50 \\ .30 & .50 & 1 \end{pmatrix} = \begin{pmatrix} -.54 & .76 & .36 \\ -.63 & -.08 & -.78 \\ -.57 & -.64 & .52 \end{pmatrix} * \begin{pmatrix} 1.84 & 0 & 0 \\ 0 & .70 & 0 \\ 0 & 0 & .46 \end{pmatrix} * \begin{pmatrix} -.54 & -.63 & -.57 \\ .76 & -.08 & -.64 \\ .36 & -.78 & .52 \end{pmatrix}$$

$$R = V * L * V'$$

**I autovalore: più elevato di tutti, associato alla prima componente che spiega una proporzione di varianza maggiore delle altre componenti**

**II autovalore: è il più elevato dopo il primo, associato alla seconda componente**

**Progressione decrescente degli autovalori, che corrisponde alla progressione della varianza spiegata dalle componenti**

**Le componenti non sono correlate (sono ortogonali)**

**Da V e da L è possibile ricavare la matrice A delle correlazioni tra variabili e fattori**

**In particolare,  $A = VL^{\cdot 5}$**

**Nell'esempio precedente la matrice A è data da:**

$$\begin{pmatrix} -.73 & .64 & .24 \\ -.85 & -.07 & -.53 \\ -.77 & -.54 & .35 \end{pmatrix} = \begin{pmatrix} -.54 & .76 & .36 \\ -.63 & -.08 & -.78 \\ -.57 & -.64 & .52 \end{pmatrix} * \begin{pmatrix} 1.36 & 0 & 0 \\ 0 & .84 & 0 \\ 0 & 0 & .68 \end{pmatrix}$$

$$A = V * L^{\cdot 5}$$

**In questo modo si può ricavare la matrice delle saturazioni fattoriali A che consente di riprodurre R**

**Infatti:**

$$R = VL^{\cdot 5}(VL^{\cdot 5})' = AA'$$

**dove  $A = VL^{\cdot 5}$**

**L'analisi sostituisce le variabili iniziali con nuove variabili (le componenti principali) che:**

- le sintetizzano
- non sono correlate tra loro
- hanno varianza massima
- hanno importanza decrescente

### Autovalori e varianza spiegata

- somma delle saturazioni elevate al quadrato per ogni fattore (colonna) = autovalore associato al fattore
- autovalore diviso per il numero di variabili osservate = proporzione di varianza totale spiegata dal fattore

### Autovalori e varianza spiegata

	F1	F2
Determinato	.68	.51
Dinamico	.74	.48
Energico	.73	.33
Affidabile	.30	-.41
Responsabile	.34	-.43
Scrupoloso	.32	-.33
<b>Autovalori</b>	<b>3.55</b>	<b>1.08</b>
<b>Varianza Spiegata</b>	<b>.61</b>	<b>.18</b>

### CONCLUSIONE

- Scopi e Modello teorico
- Equazioni Fondamentali
- Componenti Principali