FACTOR ANALYSIS

La **factor analysis** viene applicata per analizzare le correlazioni tra variabili quantitative al fine di identificarne una struttura più parsimoniosa

In particolare, partendo da *j* variabili osservate, la factor analysis mira a identificare *w* **fattori** (o componenti) latenti, con $j > w$

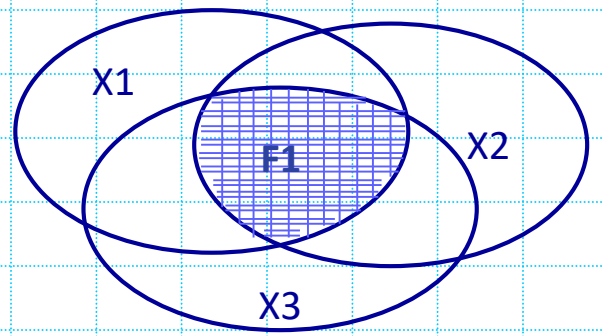
I fattori sono **variabili latenti** (nota: i componenti invece sono **combinazioni lineari**) non osservabili direttamente, che rappresentano la **parte comune a più variabili osservate** (indicatori o item)

Due tipologie: analisi fattoriale **esplorativa** (nessuna teoria, dai dati si inferisce la struttura latente) o **confermativa** (teoria da sottoporre a test confermativo, applicazione dei modelli di equazioni strutturali)

Applicazioni di marketing: riduzione della complessità per successive analisi (regressione, cluster analysis), analisi di posizionamento (ma la discriminant analysis è più efficace)

FACTOR ANALYSIS

Graficamente, i **fattori** rappresentano la parte di covarianza condivisa da più indicatori



X1, X2 e X3: indicatori

F1: fattore

A livello formale, i **fattori** vengono definiti analizzando **la matrice delle correlazioni tra gli indicatori**: gli indicatori che mostrano correlazioni alte tra di loro e basse rispetto ad altri indicatori genereranno un fattore

Indicatori	1	2	3	4	5	6	7	8
1. attrattività del luogo	1,00	0,54	0,76	0,05	0,24	0,11	0,07	0,11
2. Struttura	0,54	1,00	0,69	0,14	0,10	0,22	0,00	0,11
3. Intrattenimento	0,76	0,69	1,00	0,20	0,19	0,09	0,10	0,16
4. tempi di attesa	0,05	0,14	0,20	1,00	0,63	0,58	0,06	0,11
5. varietà dei menu	0,24	0,10	0,19	0,63	1,00	0,69	0,12	0,06
6. qualità cibo	0,11	0,22	0,09	0,58	0,69	1,00	-0,09	0,04
7. costi logistici	0,07	0,00	0,10	0,06	0,12	-0,09	1,00	0,52
8. costi di contatto	0,11	0,11	0,16	0,11	0,06	0,04	0,52	1,00

FACTOR ANALYSIS

Considerando tutti gli indicatori si spiega il 100% della varianza del fenomeno indagato

La factor analysis consente di identificare *w* **fattori**, capaci di sintetizzare in modo efficiente gli indicatori e di ridurne la complessità; **i fattori, però, spiegano una quantità di varianza inferiore al 100%**

I fattori, cioè, descrivono la **varianza comune** tra gli indicatori, ma non la parte di varianza "unica" dei singoli indicatori

Quindi, con la factor analysis, si decide di "sacrificare" una parte della varianza spiegata a favore di una maggiore **semplicità**

Per ogni fattore viene calcolato un **eigenvalue (autovalore)**: si tratta di un numero che esprime la varianza spiegata da quel fattore e può essere espresso anche in percentuale

È auspicabile che i fattori **spieghino**, in termini cumulati, almeno il **60-70% della varianza totale**

FACTOR ANALYSIS

Il funzionamento della factor analysis

Attraverso l'analisi della matrice delle correlazioni tra gli indicatori viene generato un set di fattori

Per ogni indicatore, rispetto a ogni fattore, sarà possibile calcolare/stimare un **coefficiente fattoriale (o factor loading)**; più è alto (in valore assoluto) un coefficiente fattoriale, più l'indicatore sarà associato a quel fattore

Uno degli output principali della factor analysis è, infatti, **la matrice dei coefficienti fattoriali**, che presenta gli indicatori sulle righe e i fattori sulle colonne

Gli indicatori più correlati tra loro mostreranno coefficienti alti sullo stesso fattore

ASPETTI RILEVANTI

Test di bontà del modello fattoriale

Indice KMO (preferibili valori > .70)

Test di sfericità di Bartlett (preferito $p < .01$)

Algoritmo di calcolo/stima

Analisi delle componenti principali, massima verosimiglianza, fattorizzazione dell'asse principale.

Per ridurre la complessità preservando la maggior parte della varianza → analisi delle componenti principali

Per analisi di concetti di interesse teorico (costrutti) → massima verosimiglianza

ASPETTI RILEVANTI

Numero di fattori da estrarre

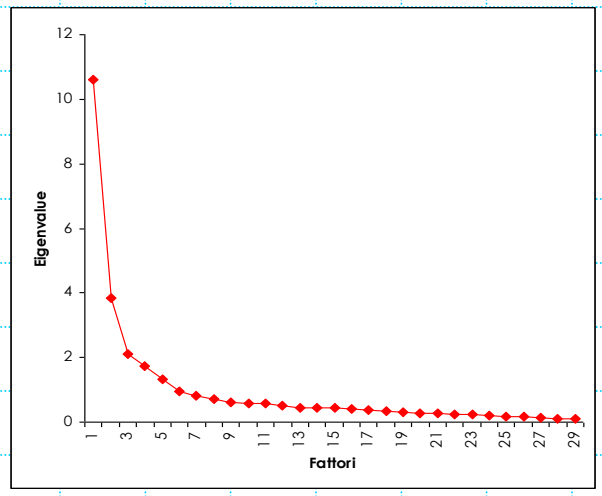
Regola di Kaiser: Eigenvalue (autovalori) maggiori di uno

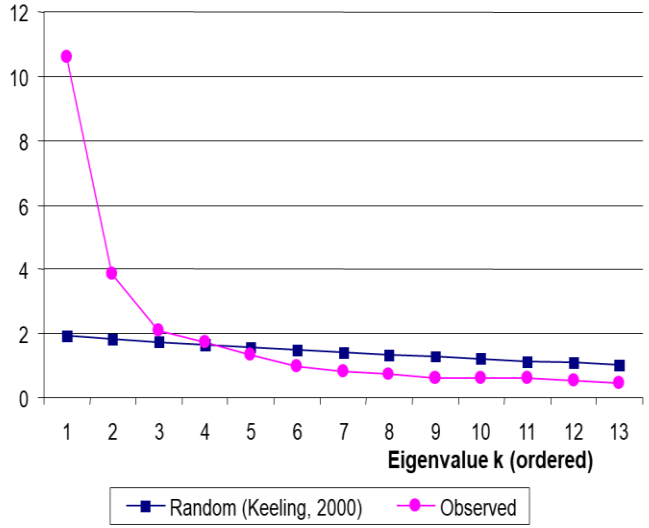
Teoria

Percentuale di varianza spiegata

Scree-plot

Parallel analysis





Più comune (anche se non sempre corretta…): eigenvalue maggiori di uno

ASPETTI RILEVANTI

Rotazione degli assi: i fattori vengono estratti in forma gerarchica (il primo fattore è quello che spiega la maggior parte della varianza comune, seguono il secondo, il terzo, ecc.)

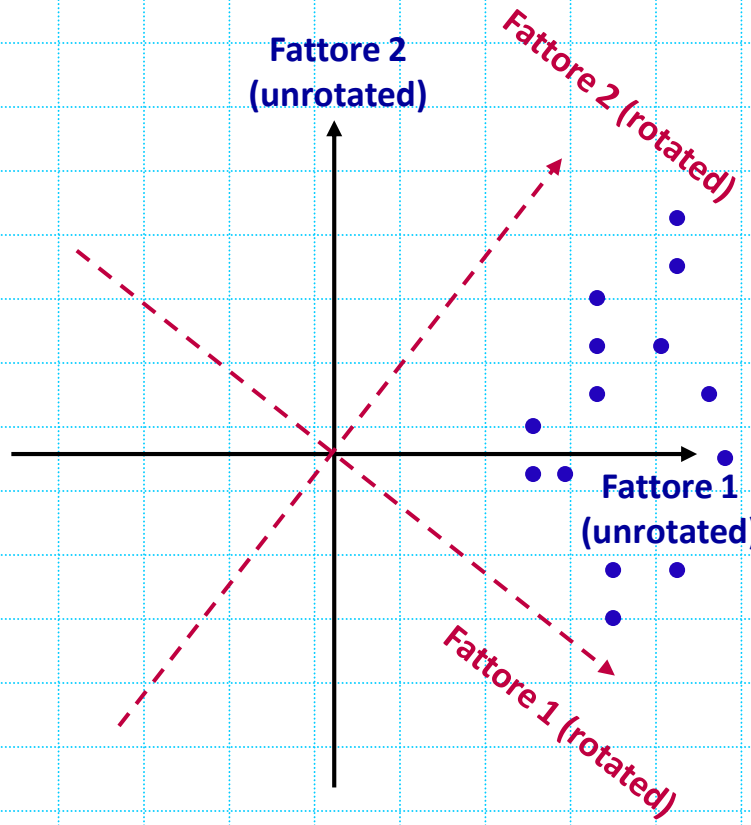
Di conseguenza, il primo fattore tende a mostrare coefficienti molto alti su tutti gli indicatori, rendendo difficile l'interpretazione dei risultati

Per ovviare a tale problema, si applica la rotazione degli assi (dei fattori)

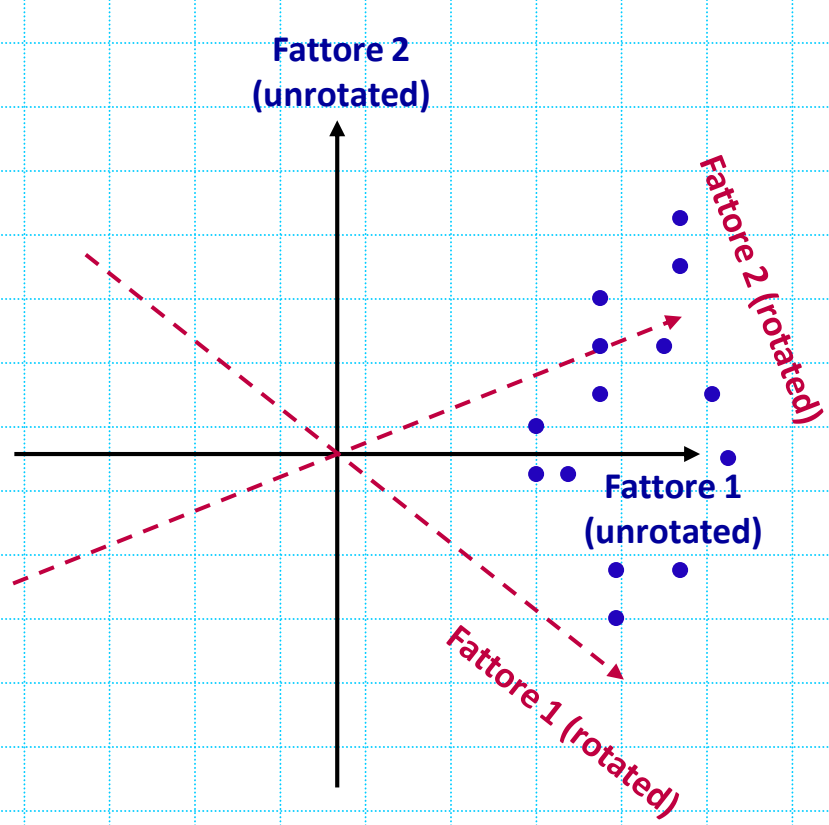
I risultati dell'analisi fattoriale che vengono interpretati concernono **la matrice dei coefficienti fattoriali ruotata**

Metodi di rotazione: ortogonali (e.g., Varimax) e obliqui (e.g., Promax)

ROTAZIONE DEI FATTORI



Rotazione ortogonale



Rotazione obliqua

ASPETTI RILEVANTI

Comunalità

Prima dell'applicazione dell'analisi fattoriale, gli indicatori mostrano comunalità (varianza spiegata) uguale a 1; le comunalità indicano, dopo l'applicazione dell'analisi fattoriale, la parte di varianza spiegata di ogni indicatore, considerando il modello fattoriale stimato

Sono preferibili comunalità maggiori di 0,50 per tutti gli indicatori

Punteggi fattoriali

Dopo aver identificato una struttura fattoriale soddisfacente è possibile salvare (tramite SPSS) **i punteggi fattoriali**

I punteggi fattoriali sono dei punteggi virtuali (standardizzati) calcolati per ogni soggetto rispetto a ogni fattore

Quindi, mentre a priori si analizzano i valori di ogni soggetto rispetto a ogni indicatore (variabile osservata), a posteriori è possibile ottenere i valori di ogni soggetto rispetto a ogni fattore (variabile latente)

I punteggi fattoriali possono essere utilizzati in successive analisi: regressione e cluster analysis

FACTOR ANALYSIS: UN ESEMPIO

Ricerca sulla soddisfazione del cliente degli uffici postali

24 attributi dell'offerta

Applicazione dell'analisi fattoriale esplorativa

Algoritmo: componenti principali, criterio degli eigenvalue > 1, rotazione Varimax

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,890
Bartlett's Test of Sphericity	Approx. Chi-Square	2186,226
	df	276
	Sig.	,000

Il **KMO** è uguale a 0.89, mentre il test di **Bartlett** è significativo; quindi concludiamo che il modello fattoriale è adeguato

FACTOR ANALYSIS: UN ESEMPIO

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8,950	37,293	37,293	8,950	37,293	37,293	5,663	23,596	23,596
2	2,549	10,619	47,912	2,549	10,619	47,912	4,201	17,505	41,101
3	2,151	8,962	56,874	2,151	8,962	56,874	3,274	13,643	54,744
4	1,123	4,677	61,551	1,123	4,677	61,551	1,537	6,406	61,150
5	1,056	4,399	65,951	1,056	4,399	65,951	1,152	4,801	65,951
6	,915	3,811	69,762						
7	,845	3,522	73,284						
8	,759	3,164	76,448						
9	,639	2,662	79,110						
10	,585	2,436	81,546						
11	,525	2,188	83,734						
12	,450	1,877	85,610						
13	,447	1,864	87,475						
14	,425	1,770	89,245						
15	,387	1,611	90,855						
16	,352	1,465	92,321						
17	,311	1,294	93,615						
18	,300	1,249	94,864						
19	,261	1,087	95,951						
20	,246	1,027	96,978						
21	,213	,889	97,867						
22	,195	,811	98,677						
23	,184	,765	99,443						
24	,134	,557	100,000						

Extraction Method: Principal Component Analysis.

L'analisi estrae 5 fattori (quelli che hanno eigenvalue > 1); il modello a 5 fattori spiega il 65,95% della varianza

FACTOR ANALYSIS: UN ESEMPIO

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
Professionalità	,849				
Formazione del personale	,827				
Risoluzione tempestiva dei problemi	,819				
Accuratezza delle informazioni fornite	,814				
Disponibilità	,795				
Cortesia e competenza sportellisti	,746				
Cortesia e competenza direttori	,700				
Comprensione delle esigenze dei clienti	,631			,324	
Semplicità dei servizi offerti		,778			
Facilità nell'utilizzo dei pd		,765			
Praticità		,723			
Varietà della gamma dei pd		,710		,355	
Funzionalità	,441	,640			
Economicità		,611			
Capillarità		,609	,354		-,318
Velocità		,556			
Ambiente climatizzato			,835		
Pulizia locali			,774		
Adeguatezza locali			,731		
Assenza di barriere architettoniche		,308	,677		
Rispetto vietato fumare			,629		
Cortesia e competenza postini				,808	
Attenzione nella distribuzione della posta	,443			,608	
Presenza dei parcheggi					,880

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

È stato richiesto di non visualizzare i coefficienti < |0.30|

FACTOR ANALYSIS: UN ESEMPIO

Il **primo fattore** mostra coefficienti alti (quindi descrive) sugli indicatori relativi a professionalità, formazione, accuratezza, risoluzione problemi, disponibilità, cortesia e competenza di sportellisti e direttori, comprensione delle esigenze dei clienti – da tale evidenza è possibile definire questo fattore come **"rapporto personale"**

Il **secondo fattore** mostra coefficienti alti (quindi descrive) sugli indicatori relativi a semplicità, facilità, praticità, varietà gamma, funzionalità, economicità, capillarità e velocità – da tale evidenza è possibile definire questo fattore come **"accessibilità dell'offerta"**

Il **terzo fattore** mostra coefficienti alti (quindi descrive) sugli indicatori relativi a ambiente climatizzato, pulizia locali, adeguatezza locali, assenza barriere architettoniche e rispetto divieto di fumo – da tale evidenza è possibile definire questo fattore come **"struttura"**

Il **quarto fattore** mostra coefficienti alti (quindi descrive) sugli indicatori relativi a cortesia e competenza dei postini e attenzione nella distribuzione della posta – da tale evidenza è possibile definire questo fattore come **"aspetti distributivi"**

Il **quinto fattore** mostra coefficienti alti (quindi descrive) solo sull'indicatore relativo ai parcheggi – da tale evidenza è possibile definire questo fattore come **"parcheggi"**