

Analisi Fattoriale

La soluzione fattoriale ed il modello ACP
Scelta del numero dei fattori

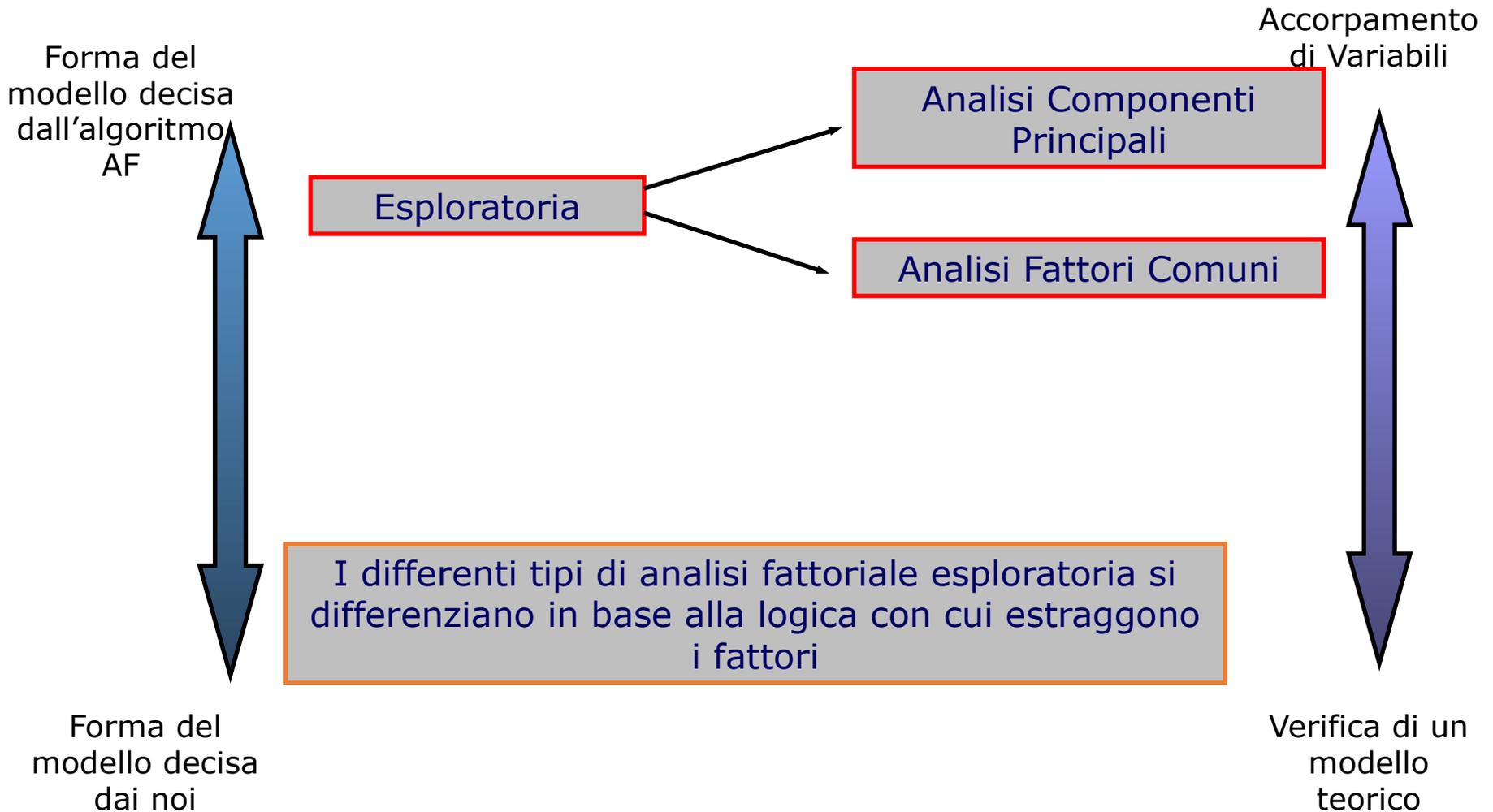
Marcello Gallucci

Milano-Bicocca

I passaggi fondamentali

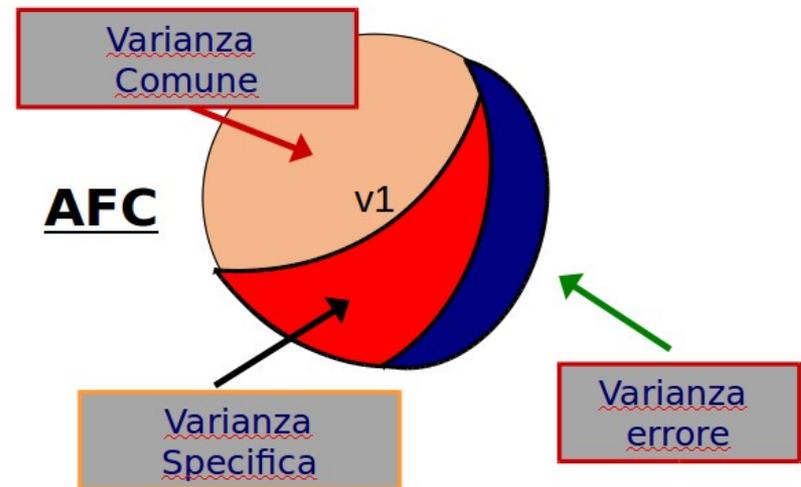
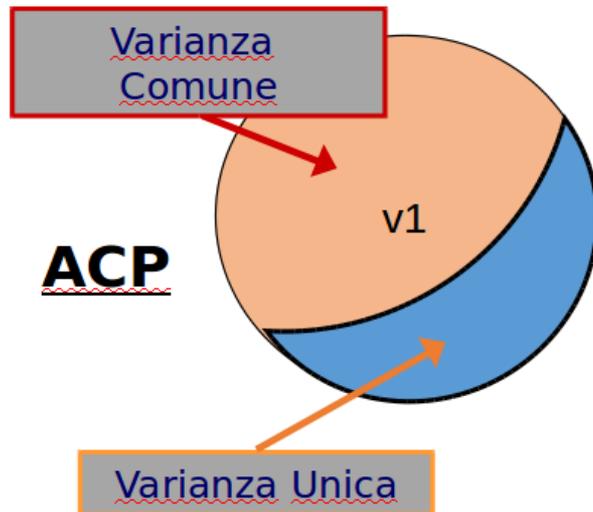
- 1) **Scelta del modello fattoriale**
- 2) Scelta del numero di fattori
- 3) Rotazione della struttura fattoriale
- 4) Selezione delle variabili
- 5) Interpretazione dei fattori
- 6) Punteggi fattoriali

Modelli di estrazione dei fattori



ACP vs. AFC

- La ACP considera tutta la varianza delle variabili e la divide in varianza comune (spiegata dai fattori considerati) and varianza unica (errore o residua)
- La AFC prima stima la varianza condivisa e distingue tra varianza comune tra le variabili (covarianza) e varianza di errore. Analizza solo la prima e la divide in varianza comune (spiegata dai fattori considerati) e varianza specifica.



Il problema della comunalità iniziale

- Per calcolare la varianza spiegata e quella di errore, dobbiamo prima sapere quanto è la varianza spiegabile
- Ma per sapere quanto è la varianza spiegabile, dobbiamo sapere quanto è la varianza spiegata e quella di errore
- Ciò crea un circolo vizioso che va risolto stimando precedentemente una quantità plausibile di varianza spiegabile ed iterando il procedimento di calcolo dei fattori finché tale quantità soddisfa alcuni criteri
- Ogni algoritmo di calcolo (*minimi quadrati*, *massima verosimiglianza*, ecc.) usa un criterio diverso
- Noi vediamo la logica sottostante

Comunalità iniziale: ACP

- Ma se tutta la varianza di un item non è inclusa nell'analisi, la matrice di correlazione iniziale non potrà avere 1 sulla diagonale

Varianza totale item= comunalità iniziale nella ACP

Matrice di correlazione

Correlazioni

		a1	a2	a3	a4
a1	Correlazione di Pearson	1	.084	.154	.242*
	Sig. (2-code)		.409	.126	.015
	N	100	100	100	100
a2	Correlazione di Pearson	.084	1	.514**	.231*
	Sig. (2-code)	.409		.000	.021
	N	100	100	100	100
a3	Correlazione di Pearson	.154	.514**	1	.588**
	Sig. (2-code)	.126	.000		.000
	N	100	100	100	100
a4	Correlazione di Pearson	.242*	.231*	.588**	1
	Sig. (2-code)	.015	.021	.000	
	N	100	100	100	100

*. La correlazione è significativa al livello 0,05 (2-code).

** . La correlazione è significativa al livello 0,01 (2-code).

Analisi delle componenti principali

- L'ACP estrae inizialmente tanti fattori quante sono le variabili osservate
- L'estrazione iniziale spiega il 100% della varianza degli items e di ogni item (diversamente dalla AFC)
- Tale soluzione non è soddisfacente, in quanto non rappresenta una soluzione efficiente e parsimoniosa
- Useremo dei metodi per decidere quanti fattori tenere che bastano a rappresentare in maniera efficiente le relazioni fra variabili

Estrazione iniziale

- Estrahendo tanti fattori quanti items (variabili osservate), spieghiamo tutta la varianza

Matrice di componenti^a

	Componente		
	1	2	3
v1	.717	-.022	-.697
v2	.609	-.659	.441
v3	.592	.705	.390

Metodo estrazione: analisi componenti principali.

a. 3 componenti estratti

La comunalità è 1, cioè i tre fattori spiegano il 100% di ogni item

Somma dei quadrati per riga

Somma dei quadrati per colonna

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	1.235	41.177	41.177	1.235	41.177	41.177
2	.932	31.072	72.249	.932	31.072	72.249
3	.833	27.751	100.000	.833	27.751	100.000

Metodo di estrazione: Analisi componenti principali.

La varianza spiegata è il 100%

Esempio con 2 fattori

- In questo esempio consideriamo due fattori

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2.520	25.197	25.197	2.520	25.197	25.197
2	1.486	14.859	40.057	1.486	14.859	40.057
3	.959	9.591	49.648			
4	.894	8.938	58.586			
5	.804	8.043	66.629			
6	.768	7.675	74.304			
7	.741	7.411	81.715			
8	.696	6.956	88.670			
9	.575	5.748	94.418			
10	.558	5.582	100.000			

Spieghiamo il 40% con due fattori, uno che spiega il 25 e l'altro il 14

Metodo di estrazione: Analisi componenti principali.

Dunque il 60% di varianza è unico agli items

Varianza e comunalità

- Nell'ACP la comunalità iniziale è 1 (100%)
- Dunque si assume che tutta la varianza di ogni item può essere spiegata
- Anche se poi ne spiegheremo solo una parte

Varianza spiegabile

Comunalità

	Iniziale	Estrazione
v1 Scherzoso	1.000	.431
v2 Estroverso	1.000	.282
v3 Espansivo	1.000	.269
v4 Divertente	1.000	.676
v5 Colorito	1.000	.432
v13 Diligente	1.000	.430
v14 Prevedente	1.000	.469
v15 Sereno	1.000	.549
v17 Stabile	1.000	.735

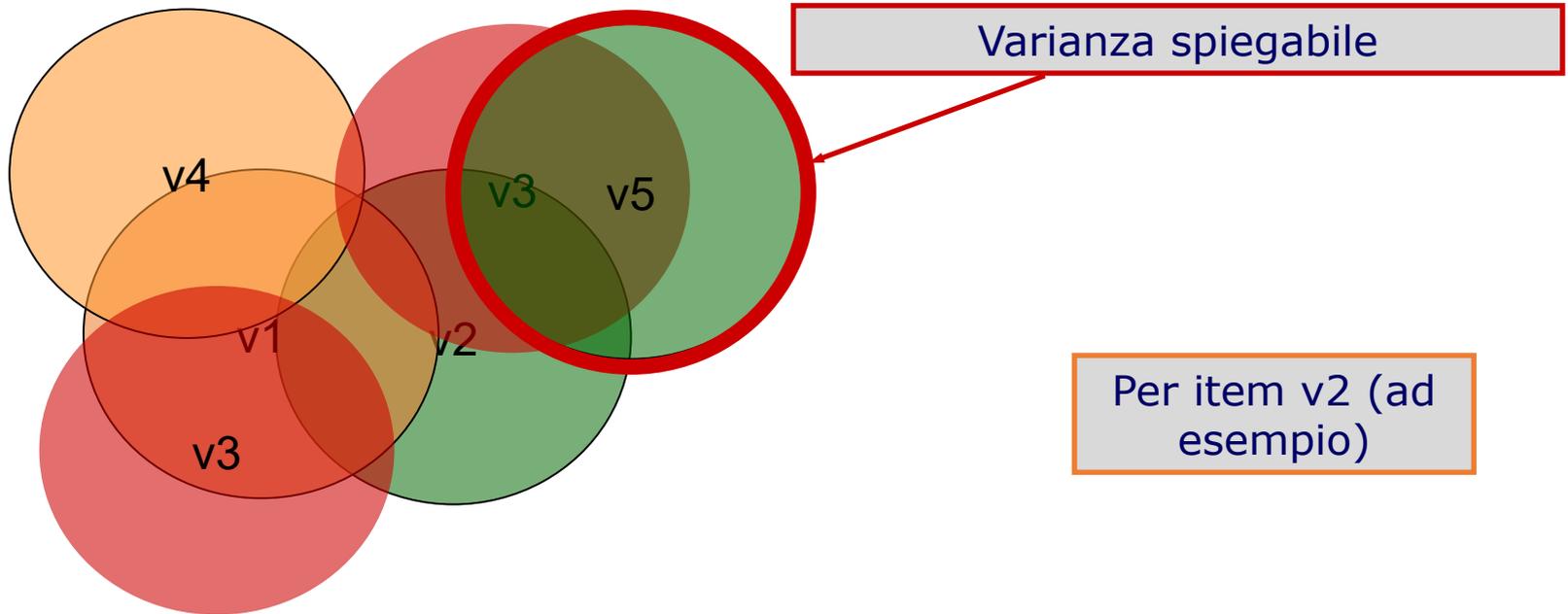
Varianza spiegata

Per ogni item

Metodo di estrazione: Analisi componenti principali.

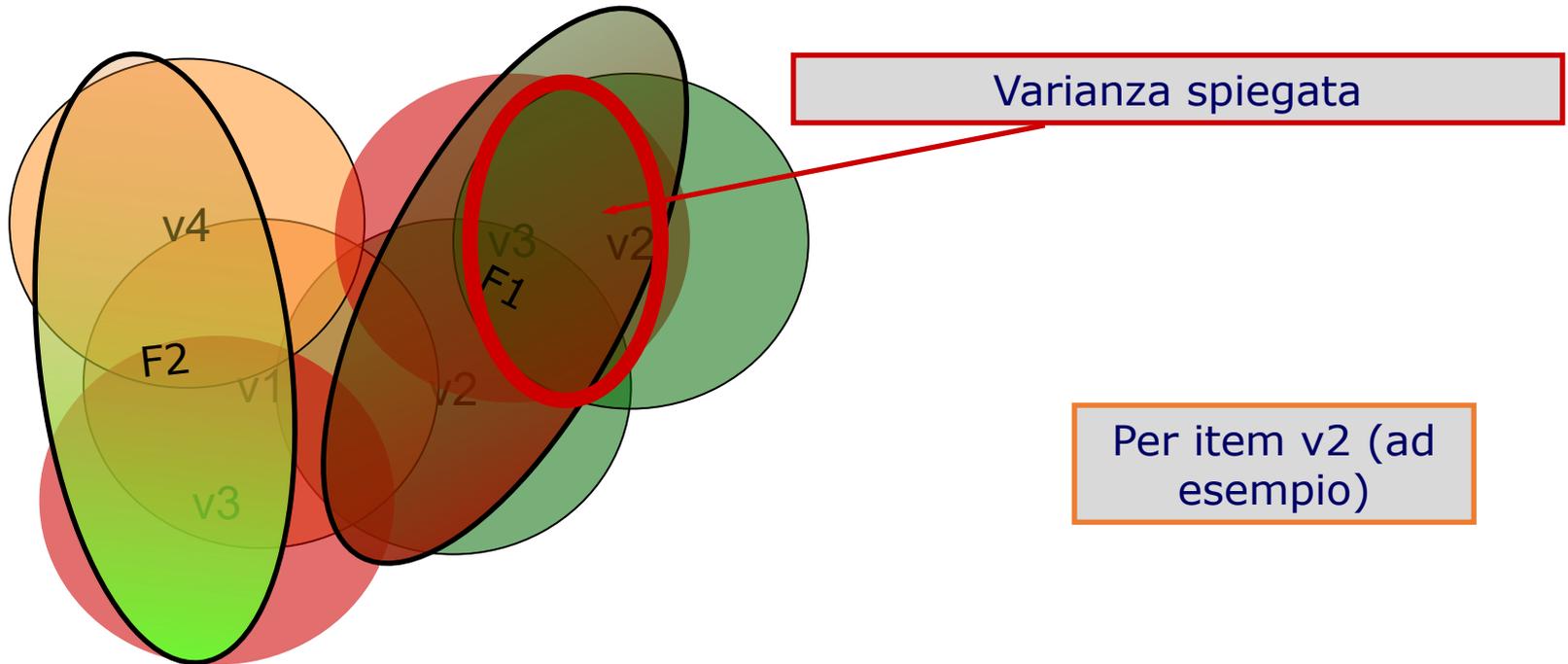
Varianza e comunalità

- Nell'ACP la comunalità iniziale è 1 (100%)
- Dunque si assume che tutta la varianza di ogni item può essere spiegata
- Anche se poi ne spiegheremo solo una parte



Varianza e comunalità

- Nell'ACP la comunalità iniziale è 1 (100%)
- Dunque si assume che tutta la varianza di ogni item può essere spiegata
- Anche se poi ne spiegheremo solo una parte



Comunalità iniziale: AFC

- Dovremmo stimare la quantità di varianza spiegabile (che sarà poi divisa in spiegata ed errore)

Varianza spiegabile dell'item=
comunalità iniziale nella AFC

Matrice di correlazione

Correlazioni

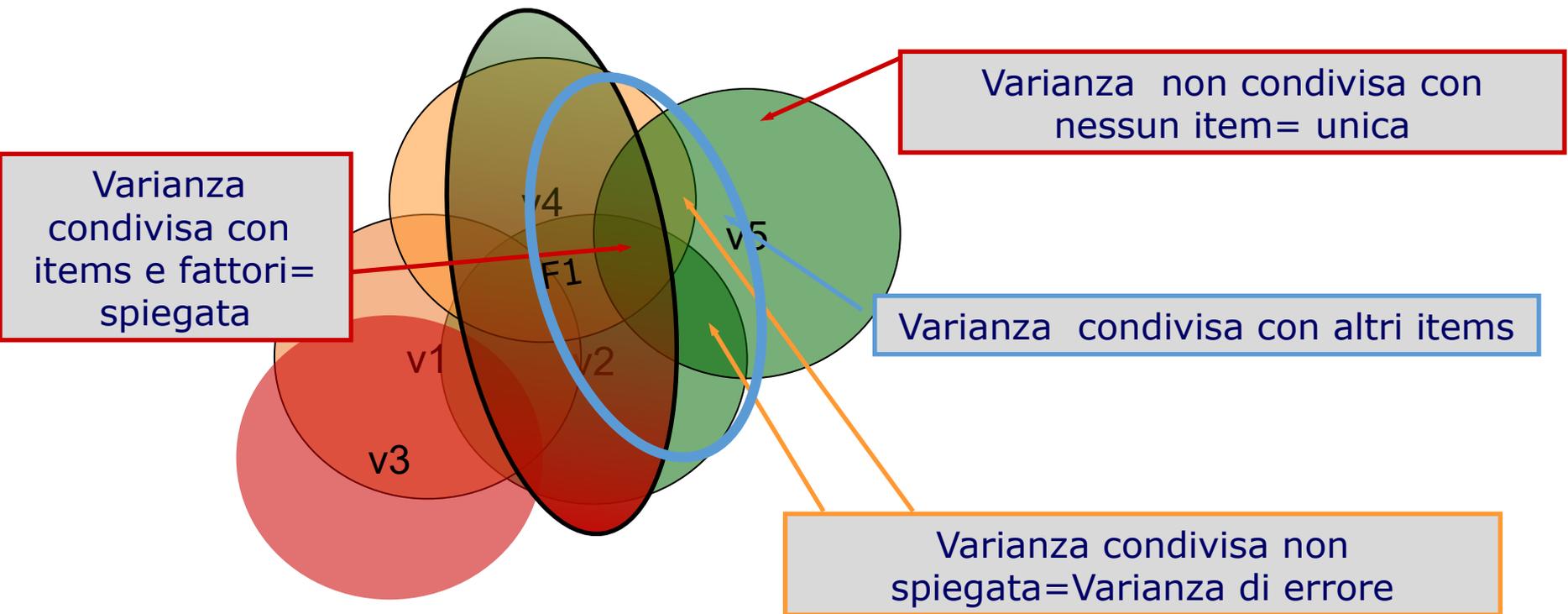
		a1	a2	a3	a4
a1	Correlazione di Pearson	?	.084	.154	.242*
	Sig. (2-code)		.409	.126	.015
	N	100	100	100	100
a2	Correlazione di Pearson	.084	?	.514**	.231*
	Sig. (2-code)	.409		.000	.021
	N	100	100	100	100
a3	Correlazione di Pearson	.154	.514**	?	.588**
	Sig. (2-code)	.126	.000		.000
	N	100	100	100	100
a4	Correlazione di Pearson	.242*	.231*	.588**	?
	Sig. (2-code)	.015	.021	.000	
	N	100	100	100	100

*. La correlazione è significativa al livello 0,05 (2-code).

** . La correlazione è significativa al livello 0,01 (2-code).

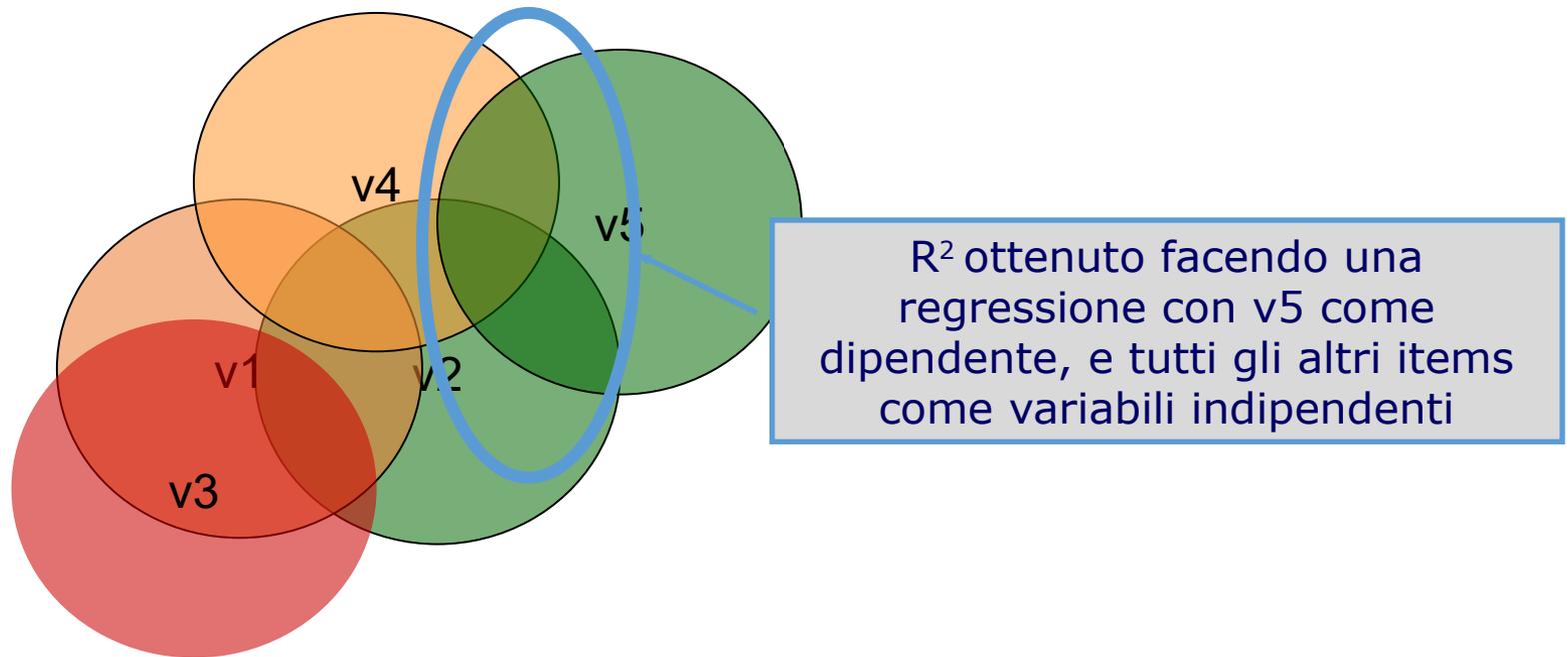
Stima iniziale della comunaltà

- Se la parte spiegabile di varianza di un item deve essere comune agli items e ai fattori, sicuramente non potrà essere più piccola della parte di varianza che l'item condivide con gli altri items



R² multiplo

- Quanto è tale varianza?
- Dallo studio della regressione sapremo che la varianza condivisa da un insieme di variabili indipendenti ed una dipendente è dato dall'R² della regressione



AFC: R^2 come comunaltà iniziale

- Ripetendo tale stima per tutti gli items, abbiamo una stima iniziale delle varianze spiegabili per ogni item

Varianza spiegabile dell'item = comunaltà iniziale nella AFC

Matrice di correlazione

Correlazioni

		a1	a2	a3	a4
a1	Correlazione di Pearson	R_{a1}^2	.084	.154	.242*
	Sig. (2-code)		.409	.126	.015
	N	100	100	100	100
a2	Correlazione di Pearson	.084	R_{a2}^2	.514**	.231*
	Sig. (2-code)	.409	.000	.021	
	N	100	100	100	100
a3	Correlazione di Pearson	.154	.514**	R_{a3}^2	.588**
	Sig. (2-code)	.126	.000	.000	
	N	100	100	100	100
a4	Correlazione di Pearson	.242*	.231*	.588**	R_{a4}^2
	Sig. (2-code)	.015	.021	.000	
	N	100	100	100	100

*. La correlazione è significativa al livello 0,05 (2-code).

** . La correlazione è significativa al livello 0,01 (2-code).

Esempio con 2 fattori (ACP vs. AFC)

ACP

Communalities

	Initial	Extraction
iffp1	1,000	,369
iffp6	1,000	,570
iffp11	1,000	,570
iffp16	1,000	,332
iffp7	1,000	,374
iffp12	1,000	,345
iffp17	1,000	,576
iffp26	1,000	,504
iffp32	1,000	,476
iffp37	1,000	,159

Extraction Method: Principal Component Analysis.

AFC

Communalities

	Initial	Extraction
iffp1	,203	,230
iffp6	,344	,472
iffp11	,325	,463
iffp16	,180	,197
iffp7	,157	,197
iffp12	,174	,183
iffp17	,263	,479
iffp26	,265	,358
iffp32	,215	,314
iffp37	,066	,068

Extraction Method: Principal Axis Factoring.

ACP vs AFC

- ACP è più semplice da eseguire in quanto possiamo farci suggerire dai dati le varie scelte da compiere (numero di fattori, aggregazione degli items)
- L'ACP si usa quando il nostro atteggiamento è più esplorativo, cioè non sappiamo quanti fattori estrarre e quali fattori
- AFC è più accurata ed i risultati più stabili, ma richiede di decidere a priori il numero di fattori
- AFC è più usata quando vogliamo confermare una struttura già nota (da altre ricerche o analisi precedenti)

I passaggi fondamentali

- 1) Scelta del modello fattoriale
- 2) **Scelta del numero di fattori**
- 3) Rotazione della struttura fattoriale
- 4) Selezione delle variabili
- 5) Interpretazione dei fattori
- 6) Punteggi fattoriali

Decidere il numero di fattori

- Esistono diversi metodi per decidere quanti fattori/componenti tenere
- Non esiste un criterio “oggettivo” (ad es., algoritmo) che determina in maniera inequivocabile ed assoluta il numero dei fattori da scegliere
- Perché?
- I fattori rappresentano un modello della realtà e diversi modelli possono essere plausibili
 - Alcuni sono più plausibili, altri sono meno plausibili, ed altri ancora sono non plausibili
 - C'è un elemento soggettivo nella decisione
- Soggettivo NON significa arbitrario

Metodi principali

- Autovalore maggiore di 1 (Kaiser-Guttman, mineigen)
- Scree-Test (Cattell)
- Analisi Parallela (Horn)
- Interpretabilità dei fattori

Criterio Mineigen (Autovalore > 1)

- Ricordiamo che **autovalore** = varianza spiegata dal fattore

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	1.235	41.177	41.177	1.235	41.177	41.177
2	.932	31.072	72.249	.932	31.072	72.249
3	.833	27.751	100.000	.833	27.751	100.000

Metodo di estrazione: Analisi componenti principali.

- Ogni variabile spiega se stessa, dunque spiega 1 (variabili standardizzate)
- Dunque un fattore che spiega meno di 1, spiega meno di una variabile
- Usare fattori che spiegano meno di 1 non è efficiente (sarebbe più efficiente usare una variabile osservata)
- Criterio detto *Mineigen* o di Kaiser-Guttman

Esempi

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2,614	21,780	21,780	2,614	21,780	21,780
2	2,553	21,278	43,058	2,553	21,278	43,058
3	1,306	10,881	53,939	1,306	10,881	53,939
4	1,226	10,220	64,159	1,226	10,220	64,159
5	,944	7,867	72,027	,944	7,867	72,027
6	,723	6,026	78,052	,723	6,026	78,052
7	,607	5,060	83,113	,607	5,060	83,113
8	,570	4,749	87,861	,570	4,749	87,861
9	,504	4,201	92,062	,504	4,201	92,062
10	,445	3,709	95,772	,445	3,709	95,772
11	,262	2,181	97,952	,262	2,181	97,952
12	,246	2,048	100,000	,246	2,048	100,000

Metodo di estrazione: Analisi componenti principali.

Esempi

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2.419	16.125	16.125	2.419	16.125	16.125
2	1.948	12.986	29.111	1.948	12.986	29.111
3	1.279	8.530	37.641	1.279	8.530	37.641
4	1.085	7.236	44.877	1.085	7.236	44.877
5	1.027	6.845	51.722	1.027	6.845	51.722
6	.958	6.385	58.107	.958	6.385	58.107
7	.886	5.905	64.012	.886	5.905	64.012
8	.849	5.663	69.675	.849	5.663	69.675
9	.820	5.467	75.143	.820	5.467	75.143
10	.724	4.824	79.967	.724	4.824	79.967
11	.670	4.464	84.431	.670	4.464	84.431
12	.646	4.306	88.737	.646	4.306	88.737
13	.593	3.955	92.692	.593	3.955	92.692
14	.579	3.859	96.551	.579	3.859	96.551
15	.517	3.449	100.000	.517	3.449	100.000

Metodo di estrazione: Analisi componenti principali.

Esempi

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	12.470	12.470	12.470	12.470	12.470	12.470
2	9.019	9.019	21.489	9.019	9.019	21.489
3	5.831	5.831	27.320	5.831	5.831	27.320
4	4.981	4.981	32.301	4.981	4.981	32.301
5	3.912	3.912	36.213	3.912	3.912	36.213
6	2.281	2.281	38.494	2.281	2.281	38.494
7	2.135	2.135	40.629	2.135	2.135	40.629
8	1.918	1.918	42.547	1.918	1.918	42.547
9	1.906	1.906	44.453	1.906	1.906	44.453
10	1.744	1.744	46.197	1.744	1.744	46.197
11	1.721	1.721	47.917	1.721	1.721	47.917
12	1.640	1.640	49.557	1.640	1.640	49.557
13	1.620	1.620	51.178	1.620	1.620	51.178
14	1.499	1.499	52.677	1.499	1.499	52.677
15	1.456	1.456	54.133	1.456	1.456	54.133
16	1.344	1.344	55.477	1.344	1.344	55.477
17	1.330	1.330	56.806	1.330	1.330	56.806
18	1.295	1.295	58.102	1.295	1.295	58.102
19	1.269	1.269	59.371	1.269	1.269	59.371
20	1.225	1.225	60.595	1.225	1.225	60.595
21	1.177	1.177	61.773	1.177	1.177	61.773
22	1.155	1.155	62.928	1.155	1.155	62.928
23	1.139	1.139	64.067	1.139	1.139	64.067
24	1.094	1.094	65.161	1.094	1.094	65.161
25	1.048	1.048	66.209	1.048	1.048	66.209
26	1.041	1.041	67.250	1.041	1.041	67.250
27	1.022	1.022	68.273	1.022	1.022	68.273
28	1.004	1.004	69.276	1.004	1.004	69.276
29	.973	.973	70.250			
30	.953	.953	71.202			
31	.921	.921	72.124			
32	.901	.901	73.025			
33	.876	.876	73.901			
34	.851	.851	74.751			
35	.826	.826	75.579			

Problemi

- Tende a sovrastimare il numero di fattori (tipicamente intorno ad $1/3$ delle variabili)
- E' più utile considerarlo come indicatore del **numero massimo** di fattori da considerare (con buon senso...)
- Qual è la differenza tra un autovalore di 1.01 ed uno di 0.99? (ma questo vale per qualsiasi soglia)
- E' l'opzione di default di SPSS
 - eppure è il criterio più debole tra quello che considereremo
- sconsigliato

Scree-test

- I fattori sono estratti in sequenza, secondo la loro capacità di spiegare varianza

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2.419	16.125	16.125	2.419	16.125	16.125
2	1.948	12.986	29.111	1.948	12.986	29.111
3	1.279	8.530	37.641	1.279	8.530	37.641
4	1.085	7.236	44.877	1.085	7.236	44.877
5	1.027	6.845	51.722	1.027	6.845	51.722
6	.958	6.385	58.107	.958	6.385	58.107
7	.886	5.905	64.012	.886	5.905	64.012
8	.849	5.663	69.675	.849	5.663	69.675
9	.820	5.467	75.143	.820	5.467	75.143
10	.724	4.824	79.967	.724	4.824	79.967
11	.670	4.464	84.431	.670	4.464	84.431
12	.646	4.306	88.737	.646	4.306	88.737
13	.593					
14	.579					
15	.517					

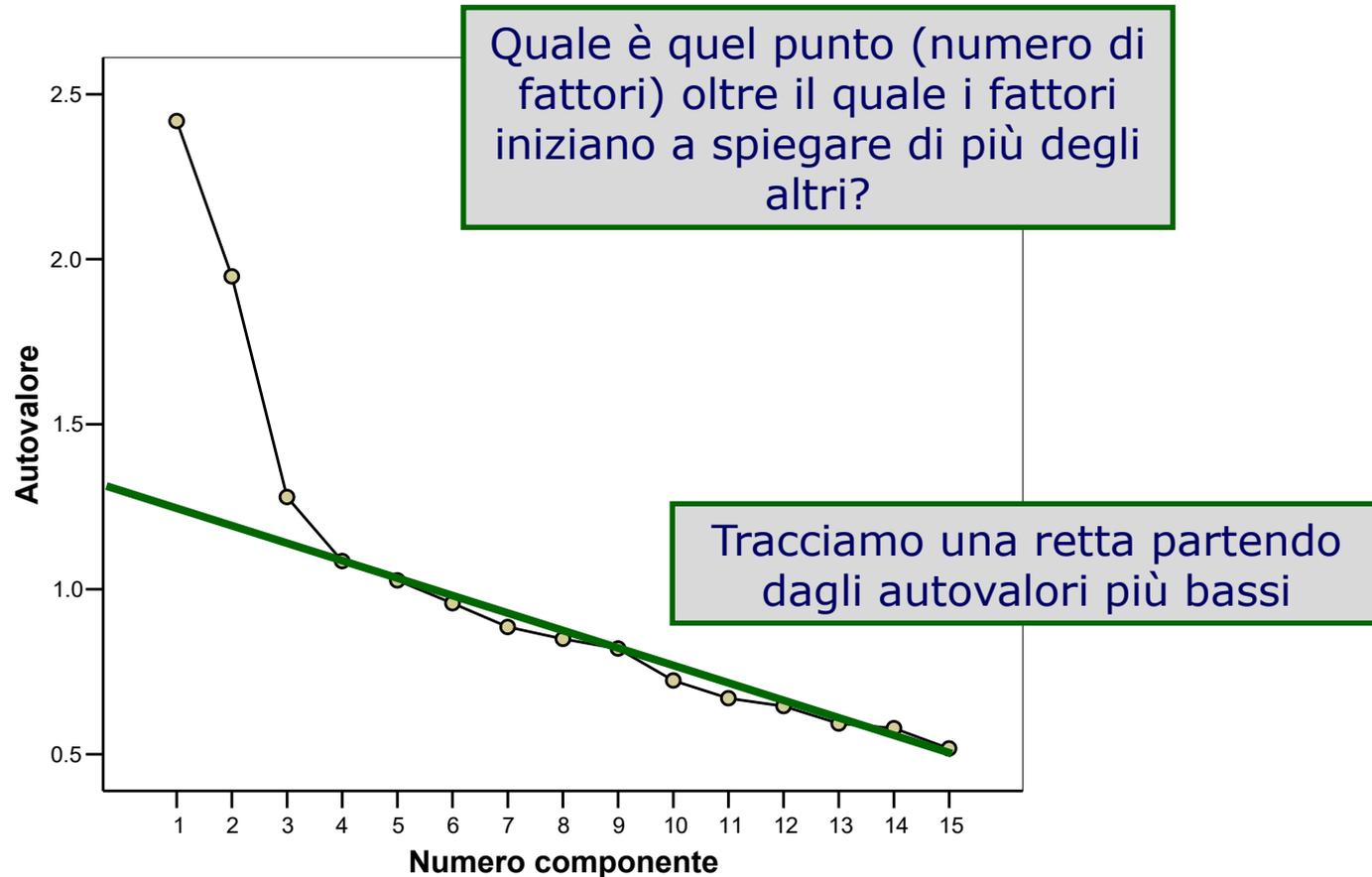
Metodo di estrazione: Analisi

Vi sarà un punto in questa sequenza dopo il quale aggiungere fattori non cambia molto la capacità di spiegare varianza

Scree-test

- Attraverso il **grafico degli autovalori**, è possibile stimare tale punto

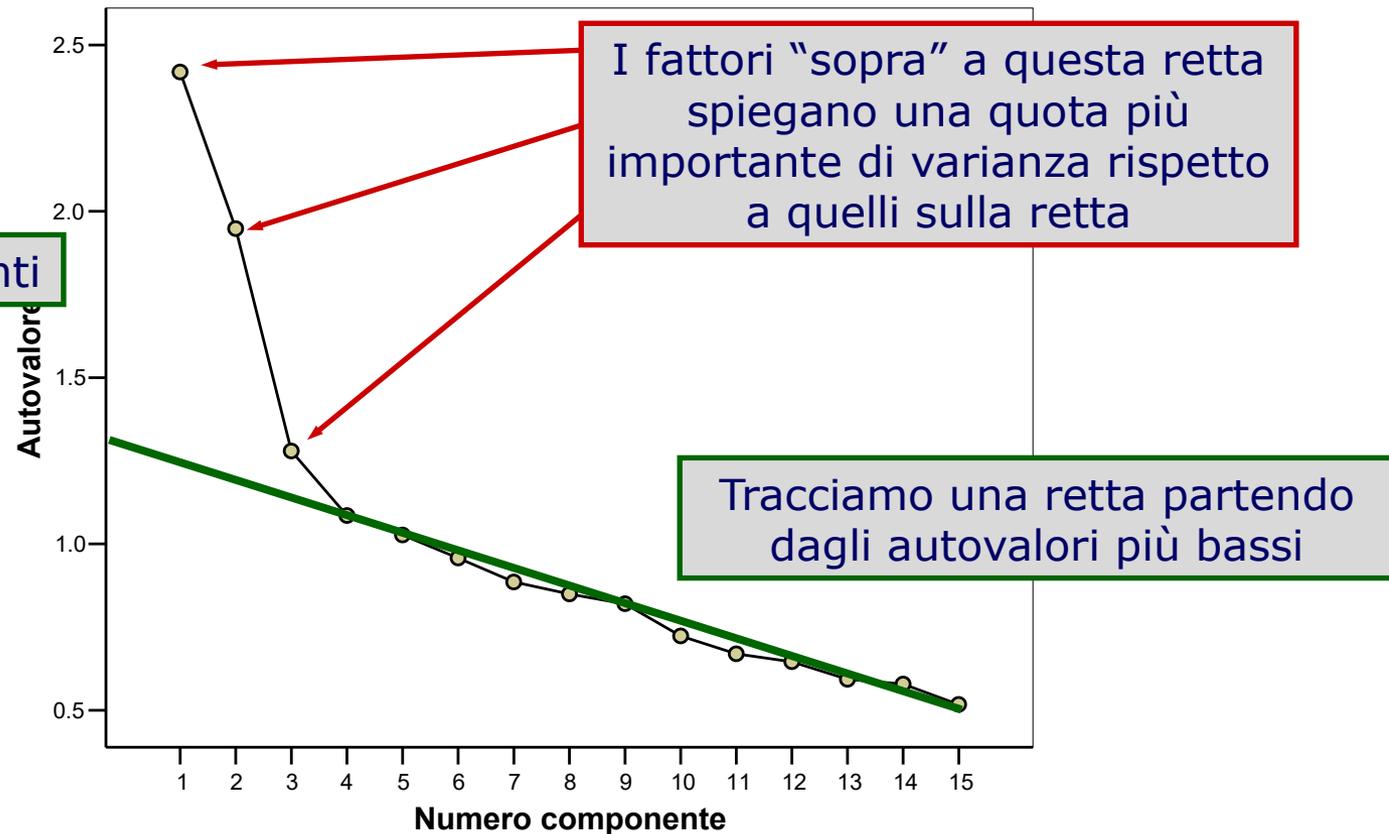
Grafico decrescente degli autovalori



Scree-test

- Attraverso il **grafico degli autovalori**, è possibile stimare tale punto

Grafico decrescente degli autovalori



Scree-test e Autovalore > 1

- Con gli stessi dati, il criterio Autovalore > 1 suggerisce la presenza di 5 componenti

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2.419	16.125	16.125	2.419	16.125	16.125
2	1.948	12.986	29.111	1.948	12.986	29.111
3	1.279	8.530	37.641	1.279	8.530	37.641
4	1.085	7.236	44.877	1.085	7.236	44.877
5	1.027	6.845	51.722	1.027	6.845	51.722
6	.958	6.385	58.107	.958	6.385	58.107
7	.886	5.905	64.012	.886	5.905	64.012
8	.849	5.663	69.675	.849	5.663	69.675
9	.820	5.467	75.143	.820	5.467	75.143
10	.724	4.824	79.967	.724	4.824	79.967
11	.670					
12	.646					
13	.593					
14	.579	3.859	96.551	.579	3.859	96.551
15	.517	3.449	100.000	.517	3.449	100.000

Il criterio scree-test è preferibile, purchè i fattori tenuti abbiano autovalori > 1

Metodo di estrazione: Analisi componenti principali.

Un altro esempio

- PCA su 10 item

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2.520	25.197	25.197	2.520	25.197	25.197
2	1.486	14.859	40.057	1.486	14.859	40.057
3	.959	9.591	49.648	.959	9.591	49.648
4	.894	8.938	58.586	.894	8.938	58.586
5	.804	8.043	66.629	.804	8.043	66.629
6	.768	7.675	74.304	.768	7.675	74.304
7	.741	7.411	81.715	.741	7.411	81.715
8	.696	6.956	88.670	.696	6.956	88.670
9	.575					
10	.558					

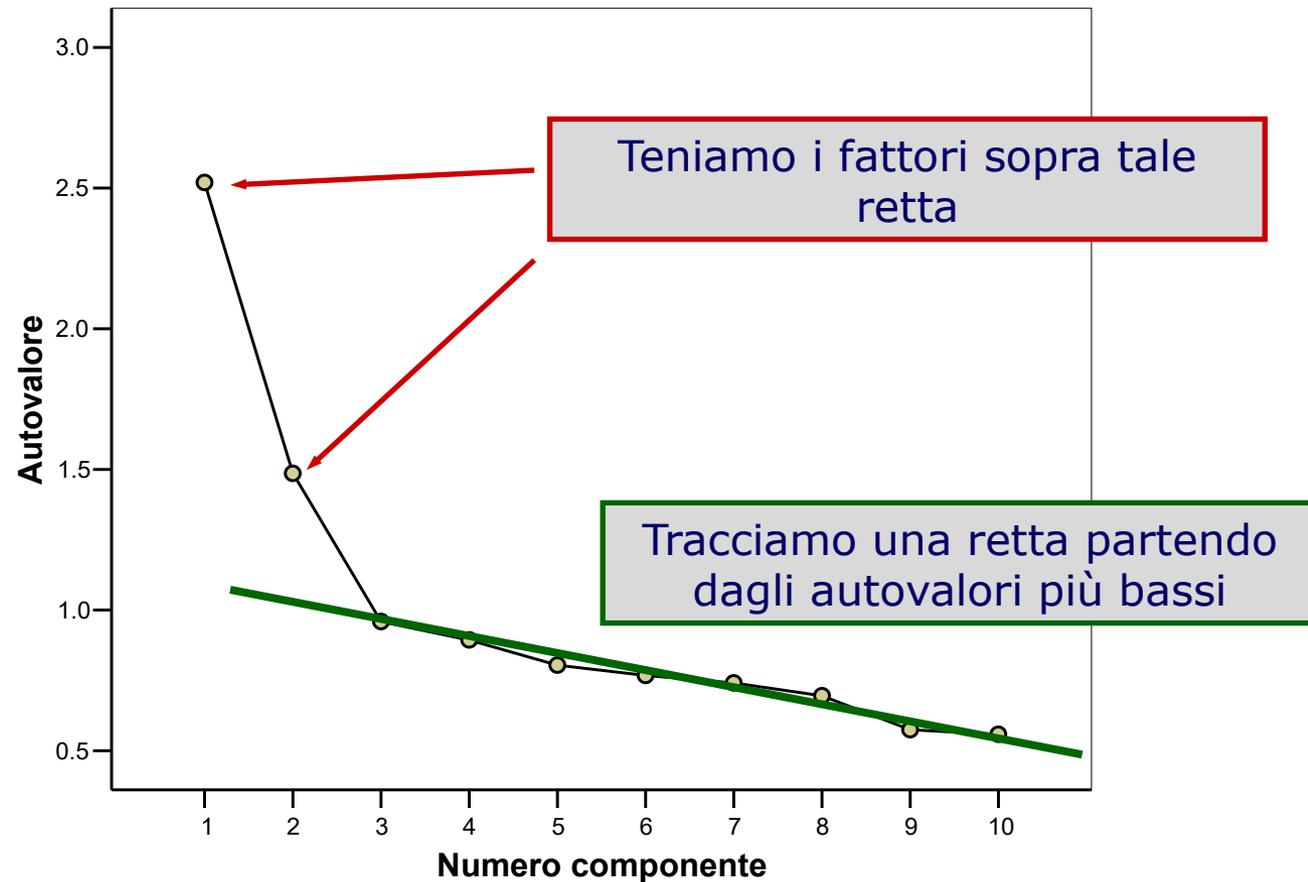
Il criterio Mineigen suggerisce la presenza di 2 componenti

Metodo di estrazione: Analisi componenti principali.

Scree-test

- anche scree test suggerisce 2 componenti.

Grafico decrescente degli autovalori



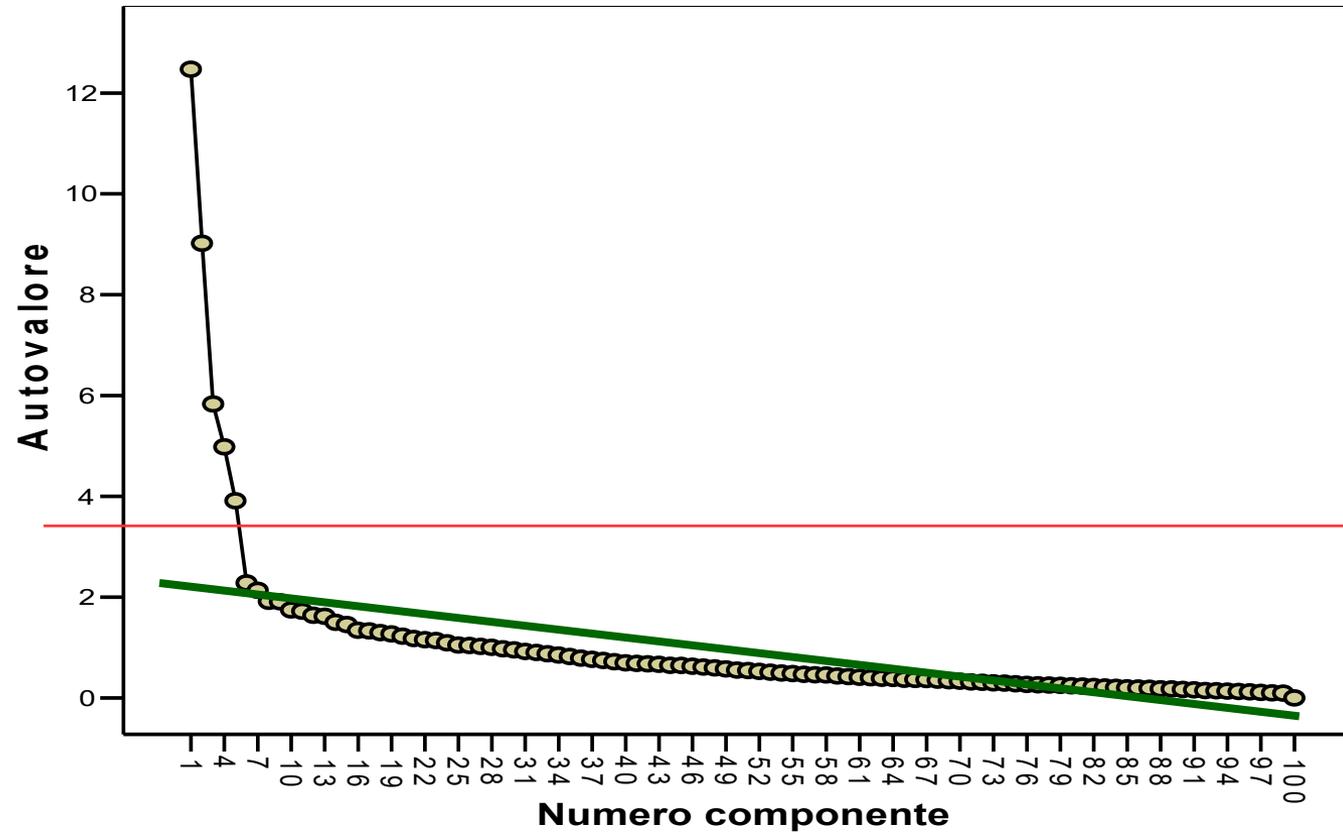
Esempi: 1

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	12.470	12.470	12.470	12.470	12.470	12.470
2	9.019	9.019	21.489	9.019	9.019	21.489
3	5.831	5.831	27.320	5.831	5.831	27.320
4	4.981	4.981	32.301	4.981	4.981	32.301
5	3.912	3.912	36.213	3.912	3.912	36.213
6	2.281	2.281	38.494	2.281	2.281	38.494
7	2.135	2.135	40.629	2.135	2.135	40.629
8	1.918	1.918	42.547	1.918	1.918	42.547
9	1.906	1.906	44.453	1.906	1.906	44.453
10	1.744	1.744	46.197	1.744	1.744	46.197
11	1.721	1.721	47.917	1.721	1.721	47.917
12	1.640	1.640	49.557	1.640	1.640	49.557
13	1.620	1.620	51.178	1.620	1.620	51.178
14	1.499	1.499	52.677	1.499	1.499	52.677
15	1.456	1.456	54.133	1.456	1.456	54.133
16	1.344	1.344	55.477	1.344	1.344	55.477
17	1.330	1.330	56.806	1.330	1.330	56.806
18	1.295	1.295	58.102	1.295	1.295	58.102
19	1.269	1.269	59.371	1.269	1.269	59.371
20	1.225	1.225	60.595	1.225	1.225	60.595
21	1.177	1.177	61.773	1.177	1.177	61.773
22	1.155	1.155	62.928	1.155	1.155	62.928
23	1.139	1.139	64.067	1.139	1.139	64.067
24	1.094	1.094	65.161	1.094	1.094	65.161
25	1.048	1.048	66.209	1.048	1.048	66.209
26	1.041	1.041	67.250	1.041	1.041	67.250
27	1.022	1.022	68.273	1.022	1.022	68.273
28	1.004	1.004	69.276	1.004	1.004	69.276
29	.973	.973	70.250			
30	.953	.953	71.202			
31	.921	.921	72.124			
32	.901	.901	73.025			
33	.876	.876	73.901			
34	.851	.851	74.751			

Esempi

Grafico decrescente degli autovalori



Esempi

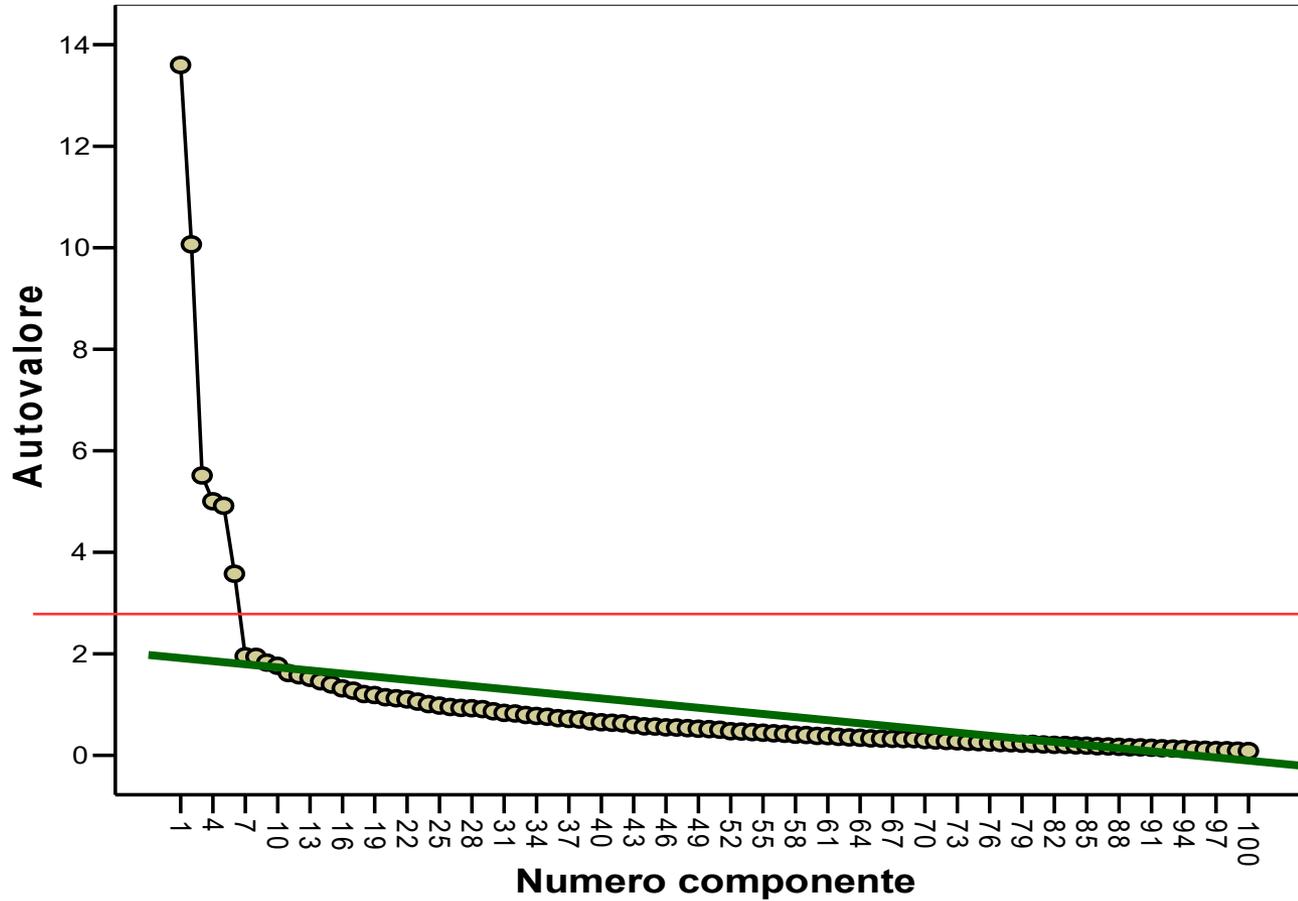
Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2,614	21,780	21,780	2,614	21,780	21,780
2	2,553	21,278	43,058	2,553	21,278	43,058
3	1,306	10,881	53,939	1,306	10,881	53,939
4	1,226	10,220	64,159	1,226	10,220	64,159
5	,944	7,867	72,027	,944	7,867	72,027
6	,723	6,026	78,052	,723	6,026	78,052
7	,607	5,060	83,113	,607	5,060	83,113
8	,570	4,749	87,861	,570	4,749	87,861
9	,504	4,201	92,062	,504	4,201	92,062
10	,445	3,709	95,772	,445	3,709	95,772
11	,262	2,181	97,952	,262	2,181	97,952
12	,246	2,048	100,000	,246	2,048	100,000

Metodo di estrazione: Analisi componenti principali.

Esempi

Grafico decrescente degli autovalori



Esempi - 3

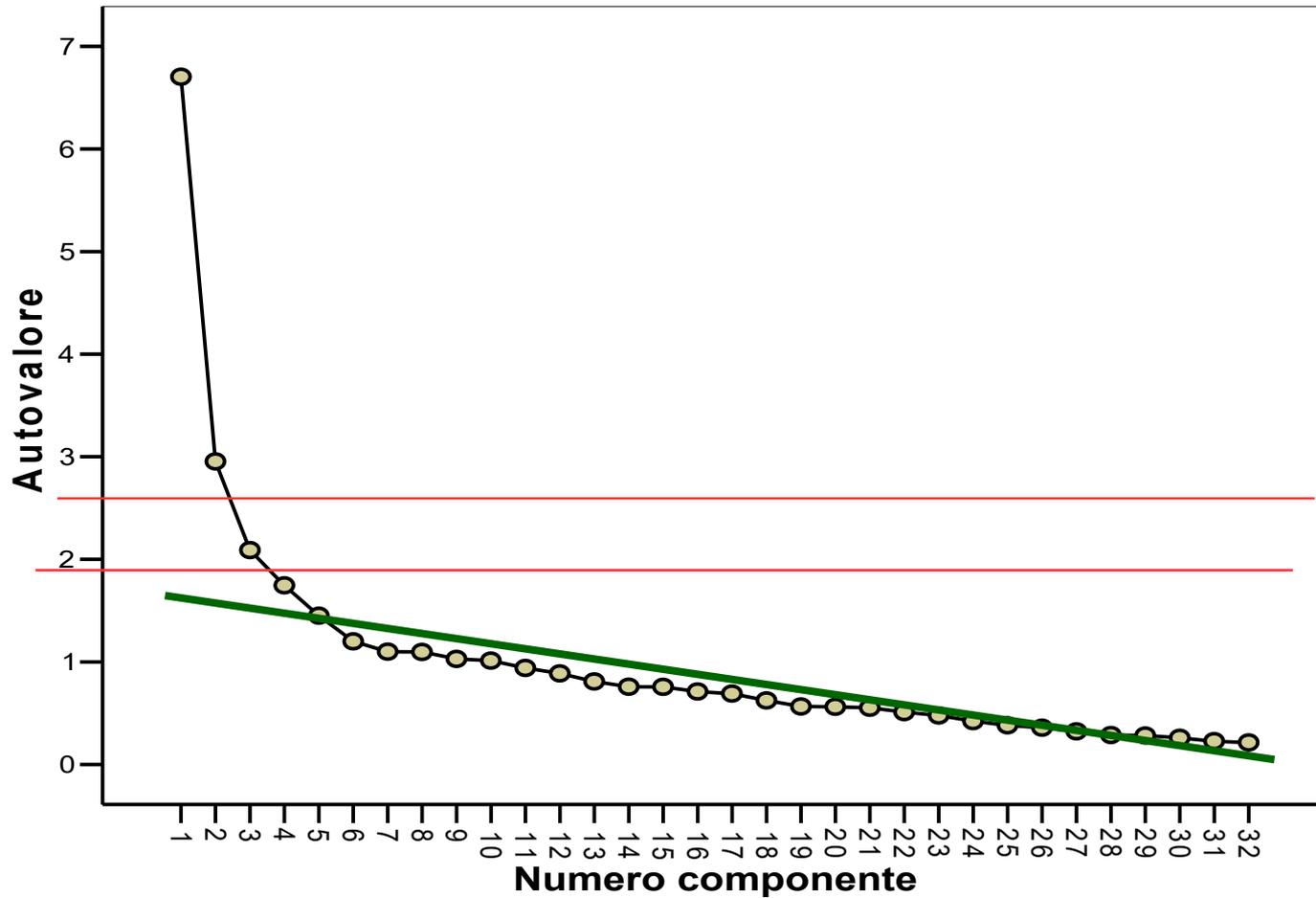
Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	6.703	20.946	20.946	6.703	20.946	20.946
2	2.954	9.232	30.178	2.954	9.232	30.178
3	2.090	6.533	36.711	2.090	6.533	36.711
4	1.746	5.458	42.169	1.746	5.458	42.169
5	1.450	4.532	46.701	1.450	4.532	46.701
6	1.201	3.753	50.454	1.201	3.753	50.454
7	1.100	3.438	53.892	1.100	3.438	53.892
8	1.097	3.429	57.321	1.097	3.429	57.321
9	1.030	3.219	60.540	1.030	3.219	60.540
10	1.013	3.164	63.704	1.013	3.164	63.704
11	.941	2.942	66.645			
12	.888	2.774	69.419			
13	.809	2.529	71.948			
14	.758	2.370	74.318			
15	.757	2.367	76.686			
16	.712	2.227	78.912			
17	.691	2.159	81.071			
18	.625	1.953	83.025			
19	.567	1.771	84.796			
20	.563	1.759	86.555			
21	.553	1.728	88.283			
22	.512	1.599	89.882			
23	.478	1.495	91.377			
24	.423	1.323	92.699			
25	.382	1.193	93.892			
26	.359	1.122	95.014			
27	.323	1.009	96.023			
28	.288	.901	96.924			
29	.283	.883	97.808			
30	.260	.813	98.620			
31	.227	.709	99.329			
32	.215	.671	100.000			

Metodo di estrazione: Analisi componenti principali.

Esempi

Grafico decrescente degli autovalori



Scree-test

- Migliore tra i metodi semplici
- Necessita di esperienza, buon occhio e conoscenza teorica
- Meglio se usato insieme a considerazioni teoriche
- Focus sul cambiamento di pendenza (salto degli autovalori)
- Tende a suggerire un numero corretto di fattori
- Logica: i raggruppamenti principali tendono ad emergere in maniera distinta dai raggruppamenti secondari
- I raggruppamenti principali tendono ad essere più affidabili rispetto a quelli secondari (segnale vs. rumore)

Scree-test

Problemi

- Metodo soggettivo
- Può indicare più soluzioni plausibili
- Necessita di esperienza, buon occhio e conoscenza teorica
- Indicazioni chiare sono meno probabili con campioni piccoli e con un rapporto più basso tra variabili e fattori estratti

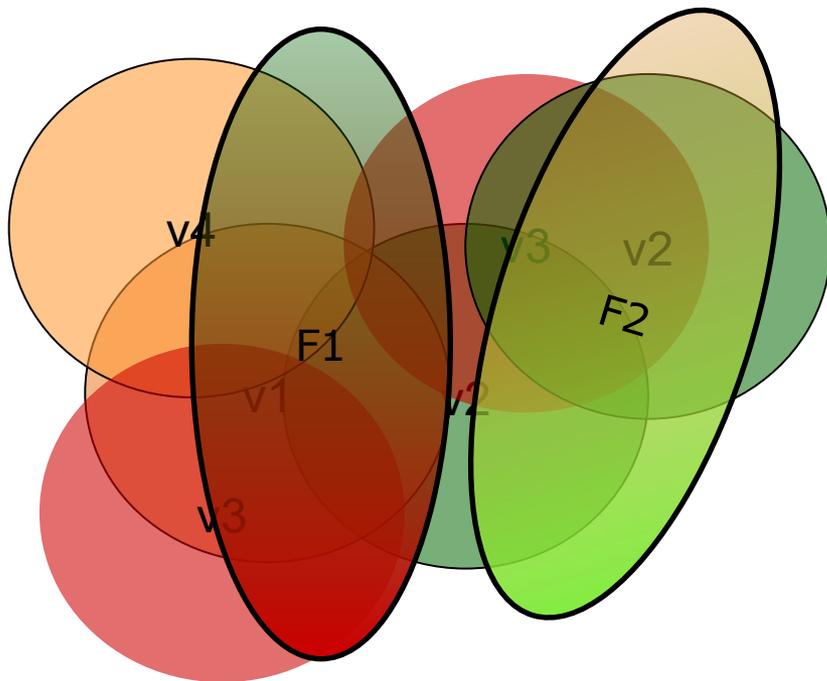
Esempio – Mineigen ?

Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	7.366	14.442	14.442	7.366	14.442	14.442
2	4.077	7.994	22.436	4.077	7.994	22.436
3	3.311	6.493	28.929	3.311	6.493	28.929
4	1.823	3.575	32.504	1.823	3.575	32.504
5	1.743	3.418	35.921	1.743	3.418	35.921
6	1.406	2.756	38.677	1.406	2.756	38.677
7	1.288	2.525	41.202	1.288	2.525	41.202
8	1.244	2.440	43.642	1.244	2.440	43.642
9	1.183	2.320	45.962	1.183	2.320	45.962
10	1.116	2.188	48.150	1.116	2.188	48.150
11	1.108	2.172	50.322	1.108	2.172	50.322
12	1.042	2.044	52.366	1.042	2.044	52.366
13	1.011	1.983	54.349	1.011	1.983	54.349
14	.971	1.904	56.253			
15	.930	1.824	58.077			
16	.903	1.772	59.848			
17	.890	1.745	61.593			
18	.845	1.657	63.251			
19	.835	1.638	64.888			
20	.821	1.610	66.498			
21	.780	1.530	68.028			

Soluzione ridotta

- Qualunque sia il criterio per scegliere il numero di fattori, tenere un numero di fattori ridotti comporta un aumento dell'efficienza ma una perdita di capacità di rappresentazione di tutta la varianza



Se tutti i fattori spiegano tutta la varianza



Un numero ridotto perderà un po' di capacità di spiegazione ma catturerà gli aspetti principali



Lo scopo è di cercare di ottenere un buon equilibrio tra semplicità e rappresentatività
(**modello parsimonioso**)