

Prof. Evandro M Saidel Ribeiro

A2.1 A Análise Fatorial

A2.2 Modelo matemático da análise fatorial

A2.3 Fatores em termos de variáveis

A2.4 Exemplo Clientes de um banco

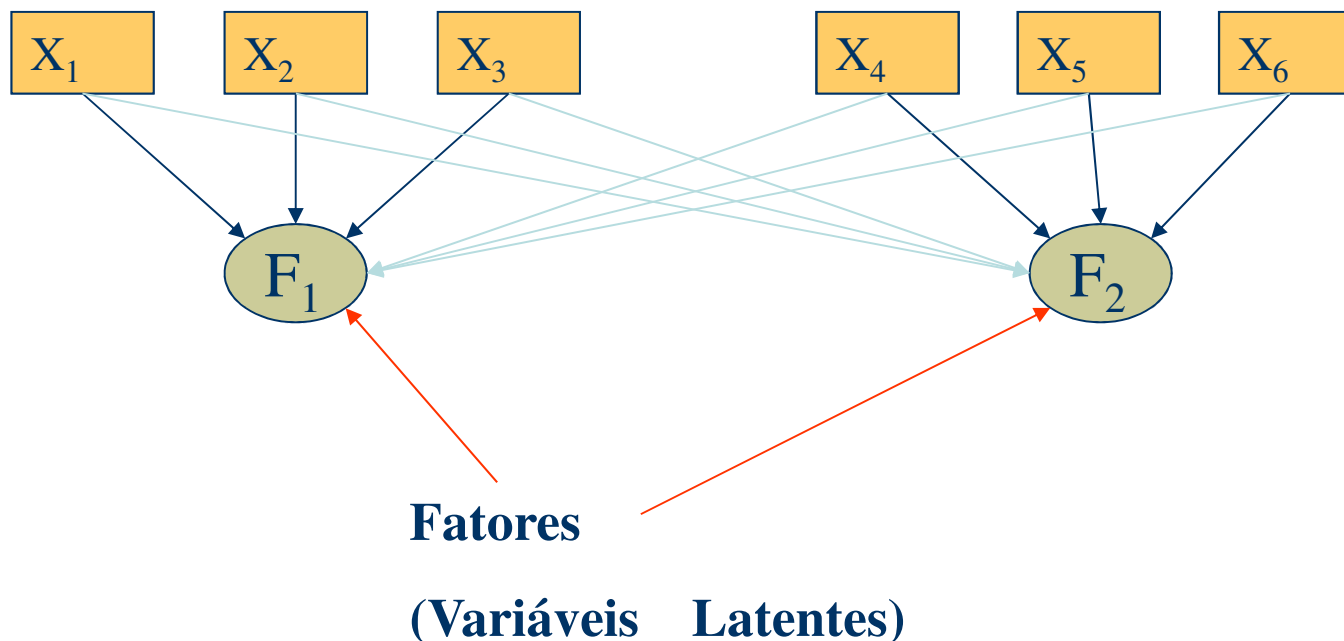
A2.5 Exemplo Completo no STATA (Corrar p96)

Bibliografia:

- J. **Lattin**, J.D. **Carroll**, P.E. **Green**. Análise de dados Multivariados, Cengage Learning, 2011.
- L.P. **Fávero**, P. **Belfiore**, F.L. **da Silva**, B.L. **Chan**, Análise de Dados: Modelagem Multivariada para tomada de decisões, Elsevier, 2009.
- L.J. **Corrar**, E. **Paulo**, J.M. **Dias Filho**, Análise Multivariada para cursos de Administração, Ciências Contábeis e Economia, Atlas, 2007.
- R.A. **Johnson**, Applied Multivariate Statistical Analysis, Prentice Hall, 1992

A2.1 A Análise Fatorial

A Análise Fatorial tem como principal objetivo descrever um **conjunto de variáveis originais** através da criação de um número menor de variáveis (**fatores**). Os fatores são variáveis hipotéticas que explicam parte de variabilidade total dos dados.



A2.1 A Análise Fatorial - Utilização

- Identificar dimensões latentes (fatores) que expliquem correlações entre um conjunto de variáveis.
- Identificar um conjunto menor de variáveis não correlacionadas para substituir um conjunto original de variáveis correlacionadas.
- Identificar em um conjunto maior, um conjunto menor de variáveis que se destacam para uso em um análise multivariada posterior.

A2.1 A Análise Fatorial - Charles Spearman

Análise fatorial: começou com Charles Spearman analisando correlações entre notas de várias disciplinas cursadas por estudantes:

	Cl (X1)	Fr (X2)	In (X3)	Ma (X4)	DT (X5)	Um (X6)
Cl	1	0,83	0,78	0,70	0,66	0,63
Fr	0,83	1	0,67	0,67	0,65	0,57
In	0,78	0,67	1	0,64	0,54	0,51
Ma	0,70	0,67	0,64	1	0,45	0,51
DT	0,66	0,65	0,54	0,45	1	0,40
Mu	0,63	0,57	0,51	0,51	0,40	1

Cl: clássicos
Fr: Francês
In: Inglês
Ma: matemática
DT: Discriminação de tom
Mu: Música

Matriz de correlação (C. Spearman, *Am. J. Psychol.* 15, 201, 1904.)

A2.1 A Análise Fatorial - Charles Spearman

A idéia básica de Spearman era que a correlação entre as variáveis poderia ser explicada pela dependência comum delas **com uma variável latente** que ele chamou de fator, ou seja, no caso das seis disciplinas:

$$\mathbf{x}_i = a_i \mathbf{F} + \boldsymbol{\varepsilon}_i \quad i = 1, 2, \dots, 6$$

Onde:

$\mathbf{x}_i$: variável analisada padronizada (**média zero e desvio-padrão 1**)

a_i : é uma constante (**carga fatorial**)

$\mathbf{F}$: **Fator** (com média zero e desvio-padrão 1) comum a todas as variáveis

$\boldsymbol{\varepsilon}_i$: erro (que é específico de cada variável)

Conclusão de Spearman:

Cada resultado do teste (notas nas disciplinas) é composto de duas partes:

- Uma parte que é comum a todos os testes (disciplinas), contida no fator comum F ao qual foi dado o nome de “inteligência geral”.
- Uma parte que é específica de uma dada disciplina.

A2.2 Modelo matemático da análise fatorial

Caso geral de **p** variáveis $\mathbf{x}_i$ ($i = 1, 2, 3, \dots, p$) **padronizadas** e **m** fatores, o modelo de análise de fatores é:

$$\mathbf{x}_i = a_{i1} \mathbf{F}_1 + a_{i2} \mathbf{F}_2 + a_{i3} \mathbf{F}_3 + \dots + a_{im} \mathbf{F}_m + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, 3, \dots, p$$

Onde:

- a_{ij} são cargas fatoriais. Medem o grau de correlação entre a variável original ($\mathbf{x}_i$) e os fatores.

- $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_m$ são os **m** fatores não correlacionados, cada um com média zero e variância unitária, ou seja,

$$\text{Var}(\mathbf{F}_i) = 1 \text{ e } \text{Corr}(\mathbf{F}_i, \mathbf{F}_j) = 0 \text{ (} i \neq j \text{)}.$$

- $\boldsymbol{\varepsilon}_i$: é um fator específico para a i -ésima variável, tem média zero e não é correlacionado com qualquer um dos fatores comuns, ou seja, $\text{Corr}(\mathbf{F}_i, \boldsymbol{\varepsilon}_j) = 0$

A2.2 Modelo matemático – A variância de x_i

$$x_i = a_{i1} F_1 + a_{i2} F_2 + a_{i3} F_3 + \dots + a_{im} F_m + \epsilon_i$$

$$\text{Var}(x_i) = a_{i1}^2 \underbrace{\text{Var}(F_1)}_1 + a_{i2}^2 \underbrace{\text{Var}(F_2)}_1 + \dots + a_{im}^2 \underbrace{\text{Var}(F_m)}_1 + \text{Var}(\epsilon_i)$$

$$\text{Var}(x_i) = a_{i1}^2 + a_{i2}^2 + \dots + a_{im}^2 + \text{Var}(\epsilon_i)$$

Comunalidade de x_i : h_i^2 , parte da variância relacionada aos fatores comuns (soma dos quadrados das cargas fatoriais)

Especificidade: parte da variância **não relacionada** aos fatores comuns

$$h_i^2 = a_{i1}^2 + a_{i2}^2 + \dots + a_{im}^2$$

A2.2 Modelo matemático – Correlações

$$\mathbf{x}_i = a_{i1} \mathbf{F}_1 + a_{i2} \mathbf{F}_2 + a_{i3} \mathbf{F}_3 + \dots + a_{im} \mathbf{F}_m + \varepsilon_i$$

Correlação entre $\mathbf{x}_i$ e $\mathbf{x}_j$:

$$\text{Corr}(\mathbf{x}_i \mathbf{x}_j) = r_{ij} = a_{i1} a_{j1} + a_{i2} a_{j2} + \dots + a_{im} a_{jm}$$

Portanto a correlação entre duas variáveis é alta se elas tem altas cargas no mesmos fatores.

Correlação entre $\mathbf{x}_i$ e $\mathbf{F}_j$: $\text{Corr}(\mathbf{x}_i \mathbf{F}_j) = a_{ij}$ (carga fatorial)

Como na análise de componentes principais, é interessante obter um número de fatores menor que o número de variáveis. Veremos então o procedimento para realizar uma análise fatorial.

A2.2 Modelo matemático: AF X CP

Semelhança: ambas buscam uma estrutura mais simples de variáveis (menor número de variáveis).

Diferenças:

- **CP** são combinações lineares das variáveis originais.
- Na **AF**, as variáveis originais são combinações lineares dos fatores.
- Na análise de **CP**, busca-se explicar a variância total dos dados. **Em AF, busca-se explicar as correlações** (ou covariância) entre variáveis.

Uma das maneiras de encontrar os fatores é por componentes principais

A2.2 Análise Fatorial – x_i explicada em termos de F_j

Na AF, após analisar a correlação entre variáveis, é feito um agrupamento das variáveis num número menor de variáveis (denominadas fatores)

Modelo matemático: Spearman - As variações em x_i podem ser explicadas a partir de um conjunto de fatores F_j

$$x_i = a_{i1} F_1 + a_{i2} F_2 + a_{i3} F_3 + \dots + a_{im} F_m + \varepsilon_i$$

- a_{ij} são as cargas fatoriais que medem o grau de correlação entre a variável original e os fatores.
- a carga fatorial ao quadrado representa o quanto da variação de uma variável é explicado pelo fator:

a_{ij}^2 : variação de x_i explicado pelo fator F_j

A2.3 Análise Fatorial – F_i escrito em termos de x_i

Os fatores, por sua vez, poderiam ser estimados por uma combinação linear das variáveis originais.

$$F_j = b_{j1}x_1 + b_{j2}x_2 + \dots + b_{ji}x_i$$

Sendo.

- F_j os fatores comuns não relacionados entre si;
- b_{ji} os **coeficientes dos escores fatoriais** e
- x_i as variáveis originais envolvidas no estudo.

O **escore fatorial** é um número resultante da multiplicação dos coeficientes b_{ji} pelo valor das variáveis originais .

Considerando a expressão para $\mathbf{F}_j$:

$$\mathbf{F}_j = b_{j1}\mathbf{x}_1 + b_{j2}\mathbf{x}_2 + \dots + b_{ji}\mathbf{x}_i$$

podemos obter o escore fatorial para todas as observações

Ou seja, para uma determinada observação, o **escore fatorial** é um número resultante da multiplicação dos coeficientes b_{ji} pelo valor das variáveis originais desta observação.

A2.4 Exemplo: Clientes de um banco

Na **Análise de Componentes Principais** vimos um exemplo:

$\mathbf{X} = \begin{bmatrix} \text{SSE} & \text{OC} \\ 7 & 6 \\ 1 & 4 \\ 3 & 4 \\ 3 & 5 \\ 6 & 6 \\ 4 & 5 \end{bmatrix}$	Matriz Covariância $\mathbf{S} = \begin{bmatrix} 4,8 & 1,8 \\ 1,8 & 0,8 \end{bmatrix}$ Matriz Correlação $\mathbf{R} = \begin{bmatrix} 1 & 0,919 \\ 0,919 & 1 \end{bmatrix}$
	Variância Total = $\text{Tr}(\mathbf{S}) = 4,8 + 0,8 = 5,6$ Autovalores de $\mathbf{S}$ $\lambda_1 = 5,491$ $\lambda_2 = 0,109$
	Autovetores $\mathbf{e}_1 = \begin{bmatrix} 0,9336 \\ 0,3583 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} -0,3583 \\ 0,9336 \end{bmatrix}$

Componentes Principais:

$$\text{CP}_1 = 0,9336 X_1 + 0,3583 X_2$$

$$\text{CP}_2 = -0,3583 X_1 + 0,9336 X_2$$

Correlações entre CP e x

	SSE	OC
CP1	0,9985	0,9387
CP2	-0,0540	0,3446

A2.4 Exemplo: Clientes de um banco

No STATA: **factor SSE OC, pcf**

```
Factor analysis/correlation
Method: principal-component factors
Number of obs      =      6
Retained factors   =      1
Number of variables =      1
```

Variância em cada Fator

Proporção de variância explicada por cada fator

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	1.91856	1.83712	0.9593	0.9593
Factor2	0.08144	.	0.0407	1.0000

```
LR chi2(1) = 8.35 Prob>chi2 = 0.0038
```

Pesos e correlações entre variáveis e fatores

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Uniqueness
SSE	0.9794	0.0407
OC	0.9794	0.0407

(1-Comunalidade)
Quanto maior esta medida menor é a contribuição da variável para o modelo fatorial

A2.4 Exemplo: Clientes de um banco

Depois de executar o comando **factor SSE OC, pcf**

- Proceder com a rotação: Comando **rotate** para obter um padrão mais claro. O método padrão é o Varimax que produz fatores ortogonais.
- Após a rotação podem ser obtidos os **scores fatoriais** com o comando **predict**. Deve-se digitar predict e depois o nome dos fatores, por exemplo:
predict fatorA fatorB

A2.5 Exemplo Completo no STATA (Corrar p96)

As seguradoras são agentes que possuem uma forte influência na economia do país, na medida que atenuam os impactos negativos das fatalidades sofridos por indivíduos ou empresas. Elas merecem especial atenção quanto à sua avaliação econômico-financeira por parte dos órgãos reguladores.

Como as seguradoras assumem riscos de seus clientes elas estão em direta exposição aos mais diversos riscos do mercado e nesse negócio ganha quem melhor equilibrar sua carteira de produtos (riscos/ganhos).

Neste exemplo de AF foram calculados 15 indicadores financeiros (slide seguinte) para 107 empresas seguradoras designadas por S001 até S107, para o ano de 2001.

A2.5 Exemplo Completo no STATA (Corrar p96)

Índices de Estrutura de Capital:

- Índice de Captações → **ICAP** = PTL/ATT
- Índice de Endividamento → **IEND** = $(PCD+ELP)/PTL$
- Índice de Recursos Próprios em Giro → **IRPG** = $(PTL-IMO-IVD-RLP)/ACL$
- Índice de Imobilização de Recursos → **IIMR** = $(IMO+IVD)/PTL$

Índices de Rentabilidade:

- Índice de Sinistralidade → **ISIN** = SRT/PGN
- Índice de Colocação do Seguro → **ICOL** = DCM/PGN
- Índice de Despesas Administrativas → **IDAD** = DAD/PRT
- Índice de Lucrativ. sobre Prêmio Ganho → **ILPG** = LLQ/PGN
- Índice de Retorno sobre o PL → **IRPL** = LLQ/PTL

Índices de Alavancagem:

- Índice de Solvência Prêmios → **PRPL** = PRT/PTL
- Índice de Alavancagem Líquida → **IALI** = $(PRT+PCC+PCD)/PTL$

Índices de Liquidez:

- Índice de Liquidez Corrente → **ILCO** = $ACL/(PCC+PCD)$
- Índice de Liquidez Geral → **ILGE** = $(ACL+RLP)/(PCC+PCD+ELP)$

Índices Operacionais:

- Índice Combinado → **ICOM** = $(SRT+DCM+DAD)/PGN$
- Índice Combinado Ampliado → **ICOA** = $(SRT+DCM+DAD)/(PGN+RFC)$

ACL → Ativo Circulante
 ATT → Ativo Total
 DAD → Despesa Administrativa
 DCM → Despesa Comercial
 ELP → Exigível a Longo Prazo
 IMO → Imobilizado
 IVD → Investimento e Diferido
 LLQ → Lucro Líquido
 PCC → Provisão comprometida Circulante
 PCD → Passivo Circulante-Demais
 PCP → Provisão Comprometida
 PGN → Prêmio Ganho
 PRT → Prêmio Retido
 PTL → Patrimônio Líquido
 RFC → Resultado Financeiro
 RLP → Realizável a Longo Prazo
 SRT → Sinistro Retido

A2.5 Exemplo Completo no STATA (Corrar p96)

Nas avaliações, as empresas seguradoras são colocadas num rating que se fundamenta em indicadores financeiros.

No entanto, quando colocado diante de uma série de indicadores financeiros,

- como decidir os pesos para cada indicador?
- como avaliar todos os indicadores conjuntamente?
- como definir quais influenciam o resultado?

Para remover a subjetividade nas respostas utiliza-se a correlação entre os diversos indicadores e aplica-se a AF.

Através da Análise Fatorial definiremos os indicadores que, em conjunto, explicam grande parte da variação que ocorre em todos os demais indicadores.

A2.5 Exemplo Completo no STATA (Corrar p96)

Sequência de comandos no STATA:

ds, has(type numeric)	# Lista de variáveis
corr ICOM-ILGE	# Matriz de correlação
factor ICOM-ILGE, pcf	# Análise Fatorial
screeplot, ytitle(Autovalor) xtitle(Fator)	

Alguns comandos Postestimation.

estat kmo	# Análise KMO
rotate	# Rotação, varimax é o default
predict f1 f2 f3	# Obtenção de coef. scores fatoriais

Para uma lista mais detalhada ver o arquivo:

<http://www.stata.com/manuals13/mvfactorpostestimation.pdf>

A2.5 Exemplo Completo: Correlação 15 indicadores

```
. corr ICOM-ILGE
(obs=107)
```

	ICOM	ICOA	ICAP	IEND	IRPG	IIMR	ISIN	ICOL
ICOM	1.0000							
ICOA	-0.0701	1.0000						
ICAP	0.2963	0.0381	1.0000					
IEND	-0.1269	-0.0597	-0.5661	1.0000				
IRPG	0.1904	0.0286	0.1422	-0.1421	1.0000			
IIMR	-0.1519	-0.0142	0.1075	0.0478	-0.4019	1.0000		
ISIN	-0.0294	0.1606	0.0576	-0.0175	0.0233	-0.1494	1.0000	
ICOL	-0.0506	-0.1560	-0.1167	0.0482	-0.0676	0.0944	-0.4238	1.0000
IDAD	0.9281	-0.0159	0.3314	-0.1507	0.1788	-0.1650	-0.0361	-0.0493
ILPG	0.9938	-0.1216	0.3138	-0.1308	0.2010	-0.1579	-0.0687	-0.0277
IRPL	-0.0088	-0.5462	-0.0829	-0.0567	-0.0271	-0.2774	-0.1347	0.1029
PRPL	-0.1339	-0.1309	-0.5483	0.4896	-0.0103	-0.0641	-0.1854	0.2734
IALI	-0.1134	-0.1101	-0.7789	0.5627	0.0405	-0.1322	-0.1759	0.2847
ILCO	0.3597	0.0425	-0.0744	-0.2513	0.1448	-0.0829	-0.0810	0.0357
ILGE	0.3738	0.0383	-0.0552	-0.2539	0.1584	-0.1048	-0.0694	0.0272
	IDAD	ILPG	IRPL	PRPL	IALI	ILCO	ILGE	
IDAD	1.0000							
ILPG	0.9464	1.0000						
IRPL	-0.0395	0.0194	1.0000					
PRPL	-0.1853	-0.1297	0.1482	1.0000				
IALI	-0.1487	-0.1126	0.0836	0.7612	1.0000			
ILCO	0.4122	0.3596	0.0174	-0.2263	0.1106	1.0000		
ILGE	0.4242	0.3743	0.0215	-0.2229	0.0928	0.9885	1.0000	

A2.5 Exemplo Completo: Autovalores 15 indicadores

factor ICOM-ILGE, pcf

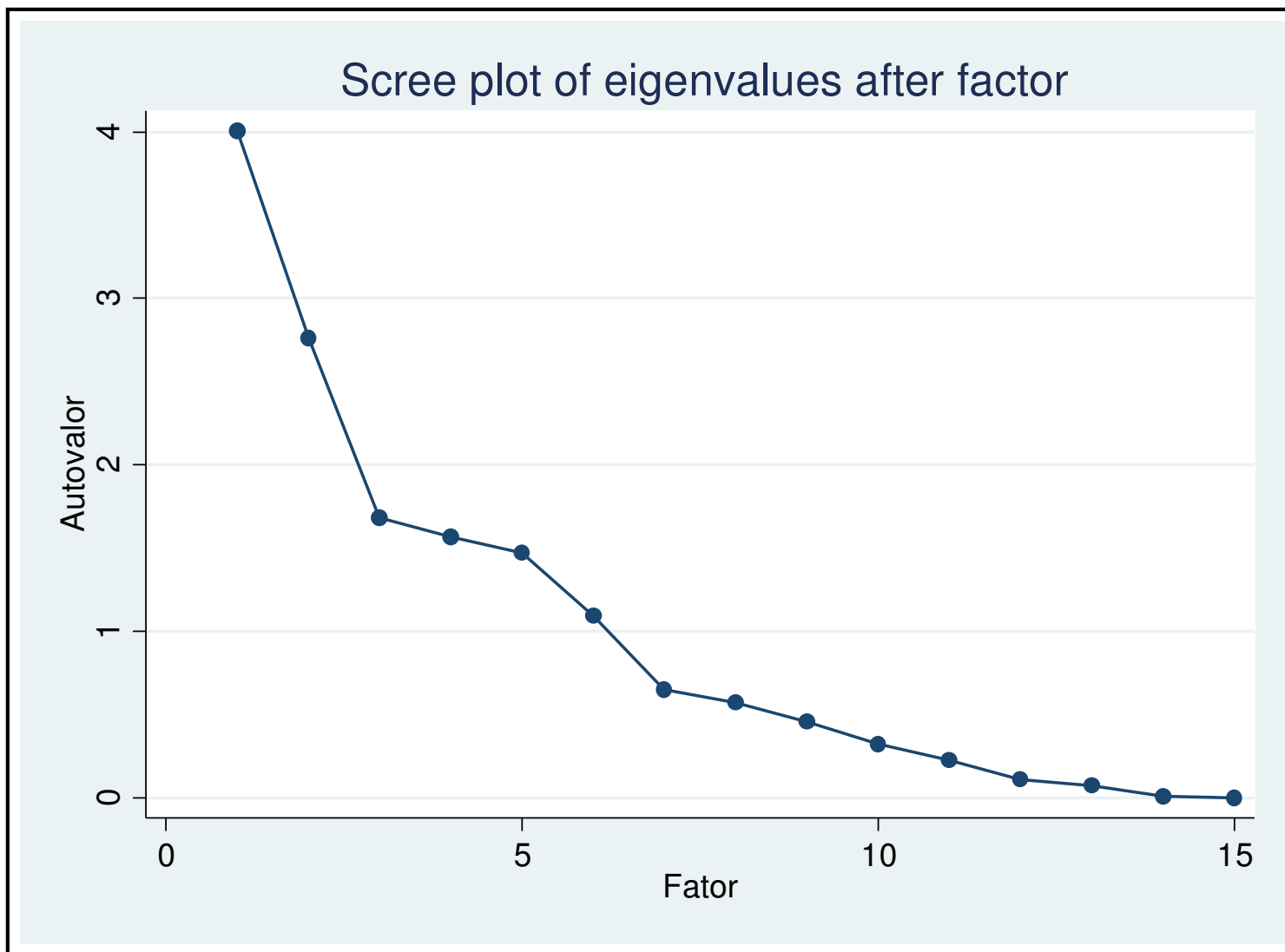
(obs=107)

Factor analysis/correlation Number of obs = 107
 Method: principal-component factors Retained factors = 6
 Rotation: (unrotated) Number of params = 75

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	4.00679	1.24313	0.2671	0.2671
Factor2	2.76366	1.08251	0.1842	0.4514
Factor3	1.68115	0.11474	0.1121	0.5634
Factor4	1.56641	0.09783	0.1044	0.6679
Factor5	1.46858	0.37732	0.0979	0.7658
Factor6	1.09126	0.44104	0.0728	0.8385
Factor7	0.65022	0.07786	0.0433	0.8819
Factor8	0.57237	0.11564	0.0382	0.9200
Factor9	0.45673	0.13586	0.0304	0.9505
Factor10	0.32087	0.09398	0.0214	0.9719
Factor11	0.22688	0.11624	0.0151	0.9870
Factor12	0.11065	0.03821	0.0074	0.9944
Factor13	0.07244	0.06224	0.0048	0.9992
Factor14	0.01020	0.00841	0.0007	0.9999
Factor15	0.00179	.	0.0001	1.0000

A2.5 Exemplo Completo: screeplot 15 indicadores

screeplot, ytitle(Autovalor) xtitle(Fator)



- **Comunalidades.**

Verificar os valores das comunalidades:

se alguma variável apresentar baixo valor de comunalidade ($< 0,7$) talvez haja necessidade de excluí-la do modelo.

$$\text{Comunalidade} = h_i^2 = a_{i1}^2 + a_{i2}^2 + \dots + a_{im}^2 = (1 - \text{unicidade})$$

Então, uma medida alta de unicidade ($> 0,3$) pode indicar a necessidade de excluir a variável do modelo.

A2.5 Exemplo Completo: comunalidade 15 indicadores

Comunalidade: $h_i^2 = a_{i1}^2 + a_{i2}^2 + \dots + a_{im}^2 = (1 - \text{uniqueness})$

Factor loadings (pattern matrix) and unique variances



Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Uniqueness
ICOM	0.8282	0.3249	-0.1675	0.2782	0.2647	-0.0619	0.0291
ICOA	0.0383	-0.2953	0.6809	0.1351	0.2579	0.3326	0.2524
ICAP	0.5639	-0.6110	-0.3422	-0.0214	0.0555	0.2019	0.1473
IEND	-0.5090	0.4334	0.0815	0.3715	0.3155	-0.2130	0.2636
IRPG	0.2919	0.1359	0.1123	0.3046	-0.4216	0.5967	0.2571
IIMR	-0.1682	-0.2626	-0.0544	-0.4331	0.6431	-0.2198	0.2503
ISIN	0.0445	-0.3457	0.3438	0.4751	-0.2405	-0.4081	0.3102
ICOL	-0.1661	0.3780	-0.2893	-0.4433	0.2078	0.4402	0.3124
IDAD	0.8531	0.2926	-0.1107	0.2281	0.2559	-0.0437	0.0549
ILPG	0.8333	0.3413	-0.2142	0.2587	0.2518	-0.0496	0.0105
IRPL	-0.0471	0.2893	-0.5712	-0.1303	-0.5780	-0.2626	0.1679
PRPL	-0.5321	0.6246	-0.0912	0.2328	0.1171	0.1264	0.2345
IALI	-0.4507	0.8078	0.2074	0.0842	0.0314	0.0290	0.0924
ILCO	0.5790	0.4088	0.4583	-0.4746	-0.1664	-0.1358	0.0163
ILGE	0.5944	0.4055	0.4501	-0.4552	-0.1783	-0.1319	0.0233

$$X_i = a_{i1} F_1 + a_{i2} F_2 + a_{i3} F_3 + \dots + a_{i6} F_6 + e_i$$

Observar:

- **KMO** (Kaiser-Meyer-Olkin) Medida de adequação da amostra (MSA), que varia no intervalo entre 0 e 1, sendo atingido o valor 1 quando as variáveis estão perfeitamente correlacionadas.

A análise deve ser feita da seguinte maneira:

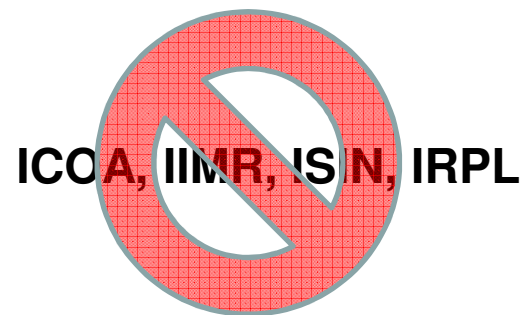
acima de 0,8	- excelente;
0,7 até 0,8	- bom;
0,6 até 0,699	- médio;
0,5 até 0,599	- ruim;
abaixo de 0,5	- inaceitável.

A2.5 Exemplo Completo: KMO 15 indicadores

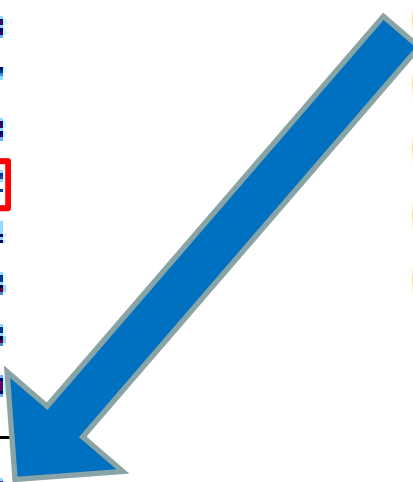
```
. estat kmo
```

Kaiser-Meyer-Olkin measure of sampling adequacy

Variable	kmo
ICOM	0.5314
ICOA	0.2277
ICAP	0.7240
IEND	0.8243
IRPG	0.5147
IIMR	0.4153
ISIN	0.3286
ICOL	0.7098
IDAD	0.6237
ILPG	0.5138
IRPL	0.4852
PRPL	0.7094
IALI	0.6393
ILCO	0.5713
ILGE	0.6020
Overall	0.5689



0.00 to 0.49	unacceptable
0.50 to 0.59	miserable
0.60 to 0.69	mediocre
0.70 to 0.79	middling
0.80 to 0.89	meritorious
0.90 to 1.00	marvelous



A2.5 Exemplo Completo: Correlação 11 indicadores

```
. factor ICOM ICAP-IRPG ICOL-ILPG PRPL-ILGE, pcf
(obs=107)
```

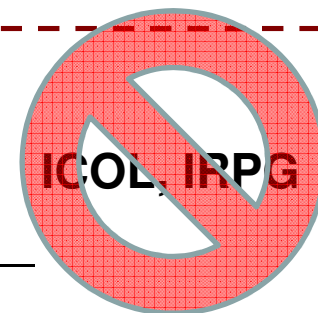
```
Factor analysis/correlation      Number of obs      =      107
Method: principal-component factors  Retained factors    =        3
Rotation: (unrotated)             Number of params    =      30
```

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	3.98197	1.40432	0.3620	0.3620
Factor2	2.57766	1.01592	0.2343	0.5963
Factor3	1.56173	0.56880	0.1420	0.7383
Factor4	0.99293	0.01877	0.0903	0.8286
Factor5	0.97416	0.52172	0.0886	0.9171
Factor6	0.45244	0.19809	0.0411	0.9583
Factor7	0.25434	0.14077	0.0231	0.9814
Factor8	0.11357	0.03698	0.0103	0.9917
Factor9	0.07659	0.06601	0.0070	0.9987
Factor10	0.01058	0.00655	0.0010	0.9996
Factor11	0.00403	.	0.0004	1.0000

```
LR test: independent vs. saturated:  chi2(55) = 1458.55 Prob>chi2 = 0.0000
```

A2.5 Exemplo Completo: Uniqueness 11 indicadores

Factor loadings (pattern matrix) and unique variances

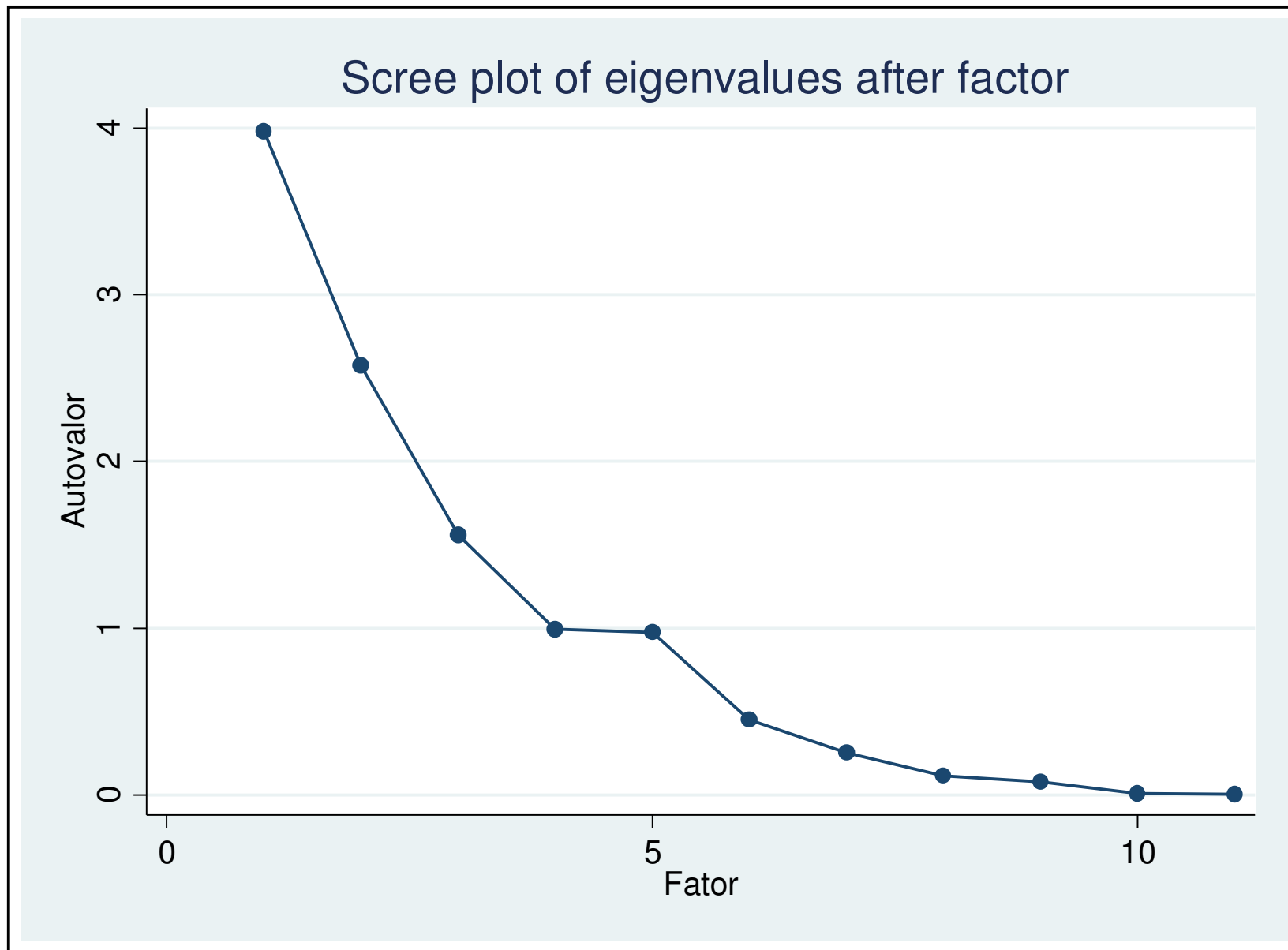


Variable	Factor1	Factor2	Factor3	Uniqueness
ICOM	0.8274	0.3359	0.4127	0.0323
ICAP	0.5746	-0.6542	0.1887	0.2063
IEND	-0.5136	0.4902	0.3526	0.3716
IRPG	0.2706	0.1074	0.0031	0.9152
ICOL	-0.1536	0.3066	-0.0336	0.8813
IDAD	0.8513	0.3099	0.3433	0.0614
ILPG	0.8335	0.3375	0.4181	0.0165
PRPL	-0.5371	0.6109	0.2976	0.2497
IALI	-0.4606	0.8320	-0.0162	0.0953
ILCO	0.5785	0.4586	-0.6580	0.0220
ILGE	0.5931	0.4530	-0.6449	0.0271

Comunalidade: $h^2 = (1 - \text{uniqueness})$



A2.5 Exemplo Completo: screeplot 11 indicadores



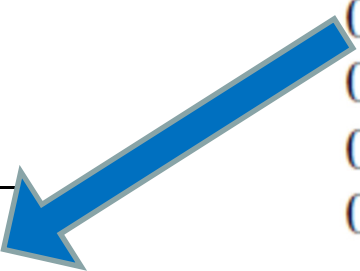
A2.5 Exemplo completo: KMO 11 indicadores

estat kmo

Kaiser-Meyer-Olkin measure of sampling adequacy

Variable	kmo
ICOM	0.6593
ICAP	0.7629
IEND	0.8426
IRPG	0.5369
ICOL	0.4792
IDAD	0.8065
ILPG	0.6249
PRPL	0.7237
IALI	0.6293
ILCO	0.5830
ILGE	0.6074
Overall	0.6732

0.00 to 0.49	unacceptable
0.50 to 0.59	miserable
0.60 to 0.69	mediocre
0.70 to 0.79	middling
0.80 to 0.89	meritorious
0.90 to 1.00	marvelous



A2.5 Exemplo Completo: Autovalores 9 indicadores

```
factor ICOM ICAP IEND IDAD ILPG PRPL-ILGE, pcf
(obs=107)
```

Factor analysis/correlation	Number of obs	=	107
Method: principal-component factors	Retained factors	=	3
Rotation: (unrotated)	Number of params	=	24

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	3.90916	1.39987	0.4344	0.4344
Factor2	2.50928	0.94802	0.2788	0.7132
Factor3	1.56126	1.01763	0.1735	0.8866
Factor4	0.54363	0.28162	0.0604	0.9470
Factor5	0.26201	0.13914	0.0291	0.9761
Factor6	0.12288	0.04603	0.0137	0.9898
Factor7	0.07685	0.06616	0.0085	0.9983
Factor8	0.01069	0.00645	0.0012	0.9995
Factor9	0.00424	.	0.0005	1.0000

LR test: independent vs. saturated: $\chi^2(36) = 1432.46$ Prob> $\chi^2 = 0.0000$

A2.5 Exemplo Completo: Uniqueness 9 indicadores

Comunalidade: $h^2 = (1 - \text{uniqueness})$

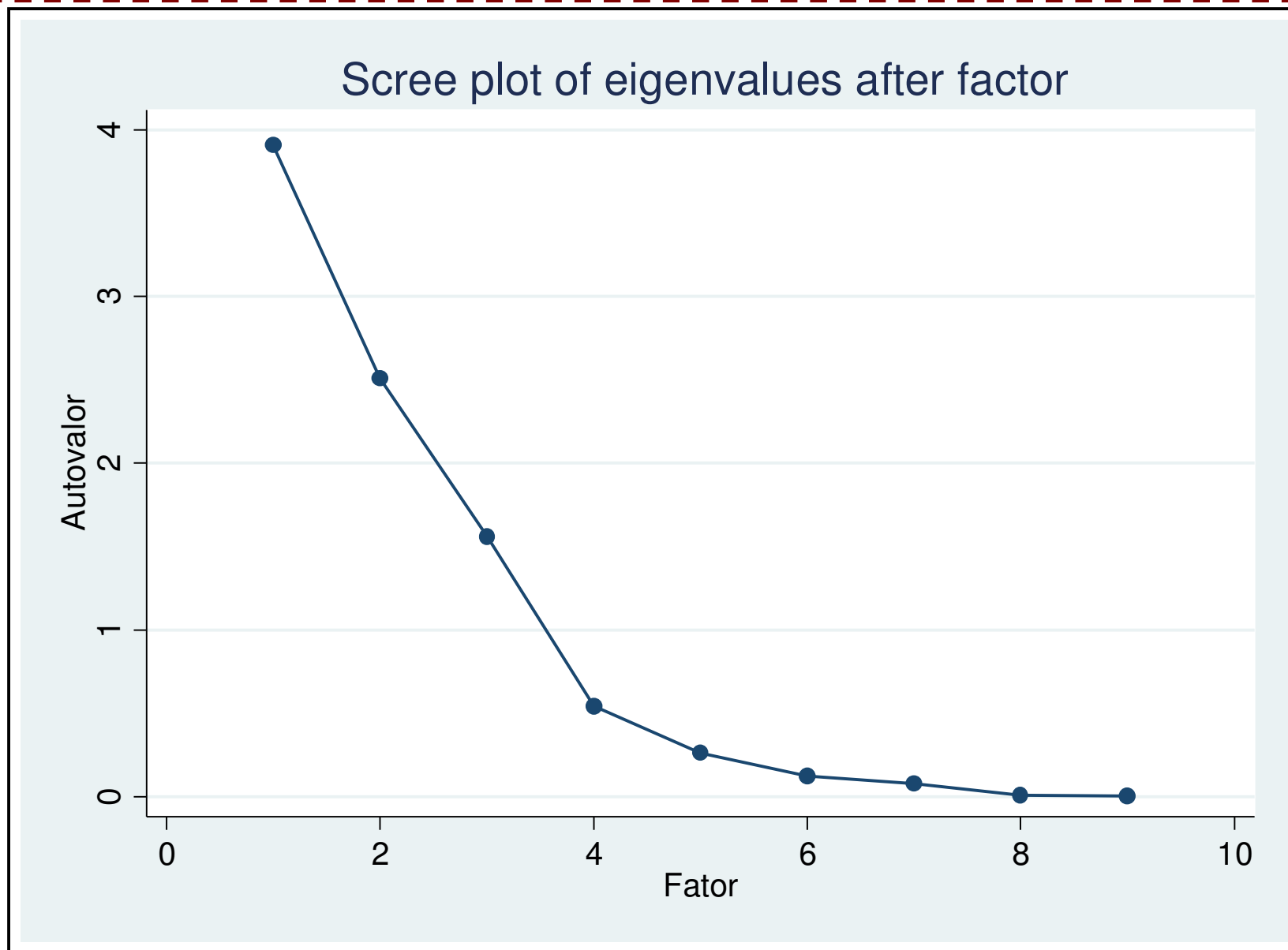
Factor loadings (pattern matrix) and unique variances



Variable	Factor1	Factor2	Factor3	Uniqueness
ICOM	0.8316	0.3349	0.4117	0.0268
ICAP	0.5693	-0.6788	0.1965	0.1765
IEND	-0.5105	0.5209	0.3448	0.3492
IDAD	0.8569	0.3079	0.3427	0.0536
ILPG	0.8381	0.3330	0.4177	0.0123
PRPL	-0.5347	0.6035	0.2957	0.2624
IALI	-0.4584	0.8294	-0.0208	0.1015
ILCO	0.5849	0.4540	-0.6601	0.0160
ILGE	0.5985	0.4485	-0.6470	0.0220

$$X_i = a_{i1} F_1 + a_{i2} F_2 + a_{i3} F_3 + \dots + a_{i6} F_6 + e_i$$

A2.5 Exemplo Completo: screeplot 9 indicadores



A2.5 Exemplo Completo: KMO 9 indicadores

```
estat kmo
```

Kaiser-Meyer-Olkin measure of sampling adequacy

Variable	kmo
ICOM	0.6617
ICAP	0.7658
IEND	0.8479
IDAD	0.8137
ILPG	0.6254
PRPL	0.7020
IALI	0.6295
ILCO	0.5822
ILGE	0.6068
Overall	0.6784

0.00 to 0.49

unacceptable

0.50 to 0.59

miserable

0.60 to 0.69

mediocre

0.70 to 0.79

middling

0.80 to 0.89

meritorious

0.90 to 1.00

marvelous



A2.5 Exemplo Completo: Rotação varimax - rotate

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
ICOM	0.9701	-0.0728	0.1635	0.0268
ICAP	0.3034	-0.8375	-0.1733	0.1765
IEND	-0.0047	0.7487	-0.3004	0.3492
IDAD	0.9403	-0.1181	0.2198	0.0536
ILPG	0.9777	-0.0768	0.1607	0.0123
PRPL	-0.0224	0.8246	-0.2391	0.2624
IALI	-0.0687	0.9352	0.1387	0.1015
ILCO	0.2150	0.0028	0.9684	0.0160
ILGE	0.2307	-0.0068	0.9616	0.0220

Factor rotation matrix

	Factor1	Factor2	Factor3
Factor1	0.7455	-0.5007	0.4400
Factor2	0.3435	0.8543	0.3901
Factor3	0.5712	0.1396	-0.8089

A2.5 Exemplo completo: varimax - rotatecompare

```
. estat rotatecompare
```

Rotation matrix — orthogonal varimax (Kaiser off)

Factor loadings

Variable	Rotated			Unrotated		
	Factor1	Factor2	Factor3	Factor1	Factor2	Factor3
ICOM	0.9701	-0.0728	0.1635	0.8316	0.3349	0.4117
ICAP	0.3034	-0.8375	-0.1733	0.5693	-0.6788	0.1965
IEND	-0.0047	0.7487	-0.3004	-0.5105	0.5209	0.3448
IDAD	0.9403	-0.1181	0.2198	0.8569	0.3079	0.3427
ILPG	0.9777	-0.0768	0.1607	0.8381	0.3330	0.4177
PRPL	-0.0224	0.8246	-0.2391	-0.5347	0.6035	0.2957
IALI	-0.0687	0.9352	0.1387	-0.4584	0.8294	-0.0208
ILCO	0.2150	0.0028	0.9684	0.5849	0.4540	-0.6601
ILGE	0.2307	-0.0068	0.9616	0.5985	0.4485	-0.6470

$$X_i = a_{i1} F_1 + a_{i2} F_2 + a_{i3} F_3 + \dots + a_{i6} F_6 + e_i$$

A2.5 Exemplo Completo: Resultado - Fatores

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
ICOM	0.9701	-0.0728	0.1635	0.0268
ICAP	0.3034	-0.8375	-0.1733	0.1765
IEND	-0.0047	0.7487	-0.3004	0.3492
IDAD	0.9403	-0.1181	0.2198	0.0536
ILPG	0.9777	-0.0768	0.1607	0.0123
PRPL	-0.0224	0.8246	-0.2391	0.2624
IALI	-0.0687	0.9352	0.1387	0.1015
ILCO	0.2150	0.0028	0.9684	0.0160
ILGE	0.2307	-0.0068	0.9616	0.0220

Fator 1: ICOM, IDAD, ILPG

Controle das despesas Operacionais

Fator 2: ICAP, IEND, PRPL, IALI

Alavancagem

Fator 3: ILCO, ILGE

Liquidez

A2.5 Exemplo Completo: Resultado - Fatores

Índices de Estrutura de Capital:

- Índice de captações → $ICAP = PTL/ATT$
- Índice de Endividamento → $IEND = (PCD+ELP)/PTL$
- Índice de Recursos Próprios em Giro → $IRPG = (PTL-IMO-IVD-RLP)/ACL$
- Índice de Imobilização de Recursos → $IIMR = (IMO+IVD)/PTL$

Índices de rentabilidade:

- Índice de Sinistralidade → $ISIN = SRT/PGN$
- Índice de Colocação do Seguro → $ICOL = DCM/PGN$
- Índice de Despesas Administrativas → $IDAD = DAD/PRT$
- Índice de Lucratividade sobre Prêmio Ganho → $ILPG = LLQ/PGN$
- Índice de Retorno sobre o PL → $IRPL = LLQ/PTL$

Índices de alavancagem:

- Índice de Solvência Prêmios → $PRPL = PRT/PTL$
- Índice de Alavancagem Líquida → $IALI = (PRT+PCC+PCD)/PTL$

Índices de liquidez:

- Índice de Liquidez Corrente → $ILCO = ACL/(PCC+PCD)$
- Índice de Liquidez Geral → $ILGE = (ACL+RLP)/(PCC+PCD+ELP)$

Índices operacionais:

- Índice Combinado → $ICOM = (SRT + DCM + DAD)/PGN$
- Índice Combinado Ampliado → $ICOA = (SRT + DCM + DAD)/(PGN + RFC)$

Fator 1: ICOM, IDAD, ILPG

Fator 2: ICAP, IEND, PRPL, IALI

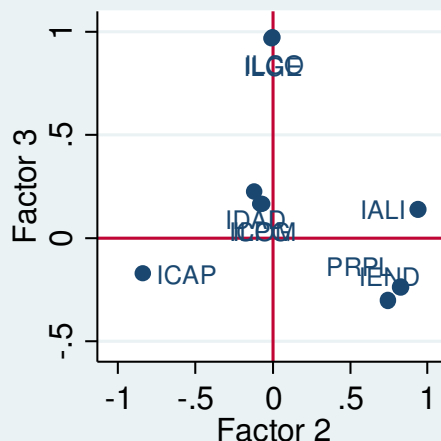
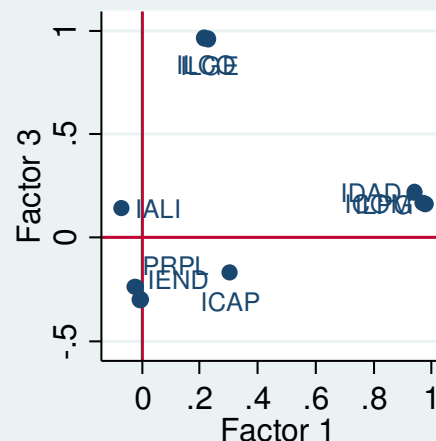
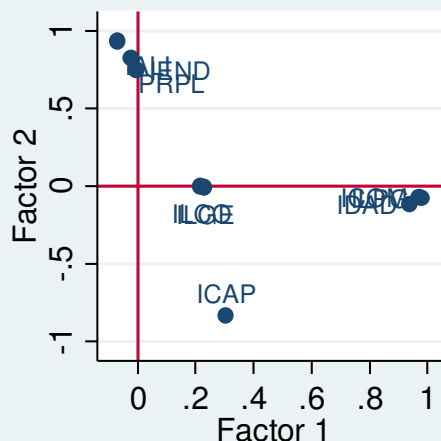
Fator 3: ILCO, ILGE

Ver Páginas 114 – 116 no livro Análise Multivariada

A2.5 Exemplo Completo: loadingplot

loadingplot, factors(3) combined xline(0) yline(0) aspect(1)

Factor loadings



Fator 1: ICOM, IDAD, ILPG
Controle das despesas Operacionais

Fator 2: ICAP, IEND, PRPL, IALI
Alavancagem

Fator 3: ILCO, ILGE
Liquidez

Rotation: orthogonal varimax
Method: principal-component factors

A2.5 Exemplo Completo: Coeficientes do Score Fatorial

$$\mathbf{F}_j = b_{j1}\mathbf{x}_1 + b_{j2}\mathbf{x}_2 + \dots + b_{ji}\mathbf{x}_i$$

```
predict f1 f2 f3
(regression scoring assumed)
```

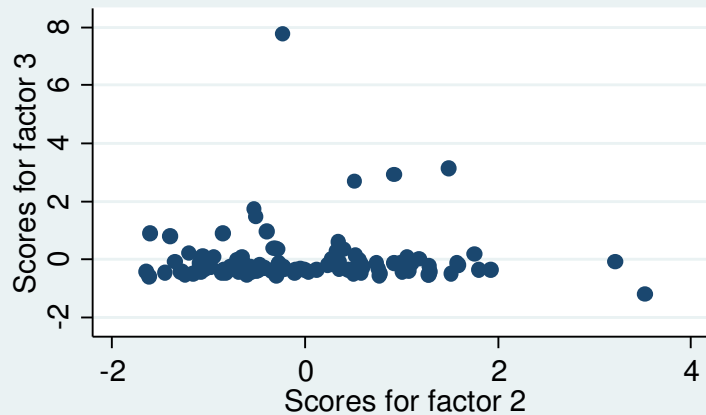
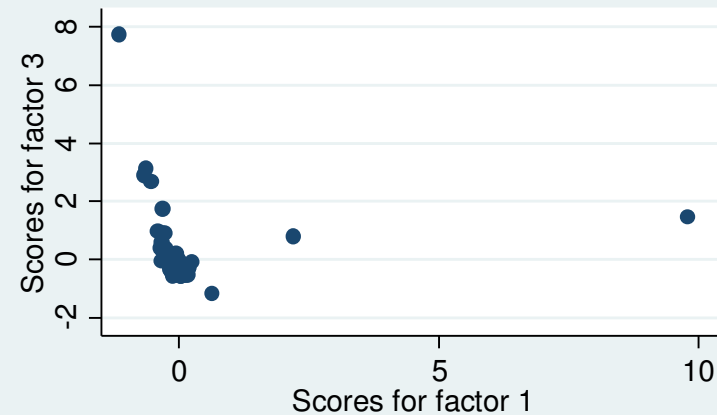
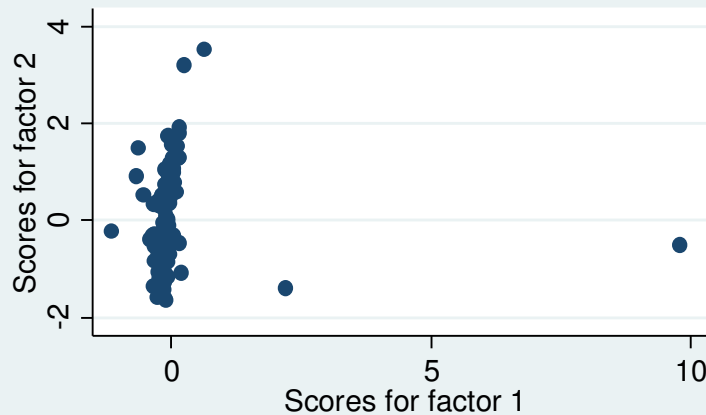
Scoring coefficients (method = regression; based on varimax rotated factors)

Variable	Factor1	Factor2	Factor3
ICOM	0.35505	0.04432	-0.06765
ICAP	0.08751	-0.28645	-0.14326
IEND	0.10009	0.27355	-0.15513
IDAD	0.33093	0.02572	-0.03327
ILPG	0.35822	0.04338	-0.07033
PRPL	0.08884	0.30040	-0.11960
IALI	0.01852	0.33923	0.08812
ILCO	-0.06779	0.02062	0.47840
ILGE	-0.06118	0.01817	0.47230

A2.5 Exemplo Completo: scoreplot

scoreplot, factors(3) combined

Score variables (factor)



Rotation: orthogonal varimax
Method: principal-component factors