

# A análise fatorial na identificação de determinantes da mortalidade infantil

Antonio Fernando C. Infantsi<sup>1</sup>; Renan M.V.R. de Almeida<sup>1</sup>; Ronaldo C. Gismondi<sup>2,3</sup>

<sup>1</sup>Programa de Engenharia Biomédica - COPPE/UFRJ  
Caixa Postal 68510 - Rio de Janeiro - RJ - CEP 21945-970

<sup>2</sup>Universidade do Estado do Rio de Janeiro

<sup>3</sup>Ministério da Saúde

E-mail: renan@serv.peb.ufrj.br

**Resumo** - A técnica da Análise Fatorial (AF) é destacada como opção aplicável à redução da dimensionalidade de grandes conjuntos de variáveis, com a vantagem da manutenção do significado de cada variável. Isso é ilustrado na seleção de variáveis independentes a serem tomadas como possíveis determinantes da mortalidade infantil. Esse tipo de metodologia pode ser de grande utilidade em modelagem na área de epidemiologia.

**Abstract** - Factor Analysis is emphasized as an option for dimensionality reduction, when dealing with large numbers of variables, while keeping with their intrinsic meaning. The technique is illustrated in the selection of independent variables, possible determinants of infant mortality. The methodology can be of great help in health modeling studies.

## Introdução

A redução da dimensionalidade de conjuntos de variáveis se constitui em importante etapa nos processos de modelagem em epidemiologia e engenharia de sistemas de saúde. Além de propiciar a seleção de variáveis não-correlacionadas, reduzindo os possíveis erros de redundância e multicolinearidade, permite aumentar a significância estatística do modelo, função inversa do número de variáveis e direta do número de amostras. Com vistas à investigação dos determinantes de mortalidade infantil, uma técnica de redução de dimensionalidade, a Análise Fatorial (AF)<sup>2,3</sup> foi aplicada a um conjunto de 40 variáveis dos setores de produção (16), econômico (13), educacional (6) e moradia e meio ambiente (5).

## Metodologia

O modelo algébrico da análise fatorial pode ser representado por:

$$Z_i = \sum_{p=1}^m a_{ip} F_p + d_i \mu_i$$

onde  $i = 1, 2, \dots, n$ ,  $Z_i$  são as variáveis observadas  $X_i$  normalizadas para média zero e desvio-padrão=1;  $F_p$  são os fatores comuns,  $\mu_i$  são os fatores (ou erros) específicos,  $a_{ip}$  representam os coeficientes dos fatores comuns ou cargas fatoriais (análogos aos coeficientes de regressão em uma regressão linear) e  $d_i$  são os coeficientes dos fatores específicos. Trabalha-se com  $Z_i$ ,  $F_p$  e  $\mu_i$  de médias iguais a zero e variâncias unitárias, pois

todas as variáveis são padronizadas. Pressupõe-se, no modelo, a não existência de correlação entre os fatores comuns, entre os fatores específicos (são independentes) e entre estes e os fatores comuns. A AF procura, assim, explicar cada variável  $X_i$  como uma função linear de fatores comuns não-observados mais um fator que é específico a cada uma das variáveis. Desse modo, a análise fatorial se constitui na determinação dos valores dos coeficientes  $a_{ip}$  e  $d_i$ ,  $i=1,2,\dots,n$  e  $p=1,2,\dots,m$ , com os quais se consiga reproduzir, da melhor maneira possível, a matriz de correlações dos dados originais. O primeiro fator comum explicará o máximo possível da variância do conjunto original dos dados; o segundo fará o mesmo, com a premissa de não ser correlacionado com o primeiro, e assim por diante. No presente estudo, a AF foi adotada por conduzir ao entendimento da relevância de cada variável, tomada isoladamente, em contraposição, por exemplo, à Análise de Componentes Principais, que aponta para combinações lineares. As variáveis em estudo referem-se a 59 municípios, de diferentes regiões brasileiras, e à média em um período de cinco anos (1978-82), como forma de reduzir o erro amostral<sup>2</sup>.

## Resultados

A Tabela 1 exemplifica a redução da dimensionalidade para o "setor de produção", em quatro agrupamentos relacionados aos fatores comuns (cargas fatoriais  $\geq 0,6$  em negrito): estrutura agrária (fator 1), comércio e prestação de serviços (fator 2), indústria (fator 3), número de estabelecimentos comerciais (fator 4). No "setor econômico" (Tabela 2), 13 variáveis foram reduzidas para dois agrupamentos: número de

pessoas com rendimentos de cinco a mais salários mínimos, salários pagos no comércio e em estabelecimentos prestadores de serviços, além do número de automóveis particulares e de trabalho (fator 1); número de pessoas sem rendimentos até aquelas que percebem menos de cinco salários mínimos e as pessoas economicamente ativas (fator 2).

A aplicação da AF ao “setor educacional” resultou, também, em apenas dois fatores (Tabela 3): pessoas com escolaridade de oito anos ou mais

Tabela 1 - Análise Fatorial, Grupo de Produção

Fator Variável	Fator	
	1	2
Nºestabel.industriais ,83 ,18	,11	,05
Valor produção industrial ,71 -,45	-,10	,04
Nºpessoas ocup.est.indus. ,92 -,19	-,12	-,07
Receita prest.serviços ,14 -,06	-,15	,76
Receita venda comércio ,27 -,24	-,03	,64

Tabela 2 - Análise Fatorial, Variáveis do Setor

Fator Variáveis	Econômico	
	1	2
Núm.pessoas rend.mensal<1 sal.mín. ,19 -,83		-
Núm.pess.rend.mensal 1 a <3 sal.mín. ,17 ,86		-
Núm.pess.rend.mensal 3 a <5 sal.mín. ,31 ,85		-
Núm.pessoas sem rendimentos ,46 -,74		-

Tabela 3 - Análise Fatorial, Grupo de Variáveis Educacionais

Variáveis	Fator	
	1	2
Número pessoas com 1º grau completo ,16 -,91		
Número pessoas alfabetizadas ,29 ,83		
Núm.pessoas 8 ou + anos de estudos ,94 ,14		
Núm.pessoas 11 ou + anos de estudos ,94 ,11		
Núm.pessoas 2º grau completo ,90 -,24		
Núm.pessoas curso superior completo ,91 ,13		

(fator 1); pessoas alfabetizadas ou com o 1º grau completo (fator 2). Por outro lado, as variáveis do “setor de condições de moradia e meio-ambiente” concentraram-se em um único fator (Tabela 4).

## Discussão e Conclusão

Tabela 4 - Análise Fatorial, Grupo de Variáveis Condições de Moradia e Meio Ambiente

Fator Variável	
1	
Número moradias c/algum sist.de esgoto ,70	
Número domicílios c/ilumin. elétrica ,77	
Número aparelhos telefônicos ,77	
Número moradias com água encanada ,88	
Número moradias com poço ou nascente -,69	

Os resultados da Análise Fatorial evidenciaram associações entre as variáveis em estudo, que permitiriam a seleção de possíveis determinantes da mortalidade infantil, relacionados aos setores de “produção”, “econômico”, “educacional” e de “condições de moradia e meio-ambiente”. Algumas das variáveis indicadas pelos fatores comuns têm sido identificadas na literatura como determinantes de morbidade e da mortalidade infantil. Por exemplo, Victora e col.<sup>4</sup> apontaram variáveis relacionadas à estrutura agrária (Fator 1, setor de produção); Benício e col.<sup>1</sup> e Silva e col.<sup>5</sup> variáveis de “educação” e “renda”.

Associando-se os resultados encontrados pela Análise Fatorial a critérios de robustez, completude e confiabilidade nos dados, pode ser possível identificar um sub-conjunto de variáveis (determinantes), que então podem ser utilizadas em modelos preditores de mortalidade infantil, ao mesmo tempo em que se preserva a informação do conjunto original de variáveis.

## Referências

- BENICIO, M.H.D.; MONTEIRO, C.A.; SOUZA, J.M.P.; CASTILHO, E.A.; LAMONICA, I.M.R. Análise multivariada de fatores de risco para o baixo peso ao nascer em nascidos vivos do município de São Paulo. SP (Brasil). Rev. Saúde Públ., S.Paulo, v.19, p. 311-320, 1985.
- GISMONDI, R.C. Sistema computadorizado para modelagem interativa de determinantes de saúde - SIMIS. Tese de mestrado, PEB/COPPE. UFRJ, 1991.
- MAINOUS, A.G. Factor analysis as a tool in primary care research. *Family Practice*, v.10, n.º 3, p. 330-336, 1993.
- VICTORA, C.G., GRASSI, P.R.; SCHMIDT, A.M. Situação de saúde da criança em área da Região Sul do Brasil, 1980-1992: tendências temporais e distribuição espacial. Rev.Saúde Pública, 28(6): 423-432, 1994.
- SILVA, L.C.; GONZÁLEZ, G.; FARIÑAS, H.; HERRERA, L. Evaluación de la mortalidad infantil según condiciones higienicosociales en el municipio. Un enfoque multivariado. *Rev. Cub. Admin. Salud*, v. jul/set, p.243-254, 1985.

## Agradecimentos

Ao CNPq e FAPERJ, que financiaram parte da presente pesquisa.