

TÉCNICAS DE ANÁLISE MULTIVARIADA PARA AVALIAÇÃO DAS CONDIÇÕES DE SAÚDE DOS MUNICÍPIOS DO RIO GRANDE DO SUL, BRASIL

Silvio Possoli*

POSSOLI, S. Técnicas de análise multivariada para avaliação das condições de saúde dos municípios do Rio Grande do Sul, Brasil. Rev. Saúde públ., S. Paulo, 18: 288 - 300, 1984.

RESUMO: Estudou-se a estrutura de correlação de variáveis da área de Saúde Pública para a obtenção de um "Índice do Nível de Saúde" para os municípios do Rio Grande do Sul. Utilizando o valor deste índice para cada município formamos 18 grupos homogêneos ("clusters") ordenados de forma decrescente de carência quanto ao nível de saúde. Outros índices foram encontrados: "Índice de Imunização" e "Índice de Não-Assistência Médico-Hospitalar". A variável mortalidade proporcional para menores de 5 anos, do conjunto total das variáveis trabalhadas, foi a que apresentou maior poder discriminativo e de diagnóstico; o peso ao nascer com menos de 2.700g, foi de menor poder diagnóstico.

UNITERMOS: Índice do Nível de Saúde. Municípios do Rio Grande do Sul, condições de saúde. Análise multivariada.

INTRODUÇÃO

Em diversas áreas como: Psicologia, Sociologia, Biologia, Medicina, Educação e Economia, deparamo-nos com observações de várias variáveis para cada elemento de uma amostra (ou população) de indivíduos.

É de interesse examinar as inter-relações entre as variáveis. Estas inter-relações podem ser avaliadas ou pelas covariâncias ou pelos coeficientes de correlação entre as variáveis. Se o número de variáveis é grande deseja-se estruturar e simplificar os dados de maneira a conservar o máximo de informação expressa pelas variáveis originais. Uma solução para este problema é encontrar variáveis hipotéticas que sejam combinações lineares das variáveis observadas e assim mais convenientemente estudadas por seu menor número.

Em muitos problemas, após o encontro destas variáveis hipotéticas, que resumiriam as informações das variáveis originais, deseja-

se unir os indivíduos formando grupos homogêneos.

O ramo da Estatística que trata da análise de dados em várias dimensões, de vários indivíduos é a Análise Multivariada.

O presente trabalho tem por objetivos:

- a obtenção de um Índice do Nível de Saúde, considerando variáveis da área de Saúde Pública, para os municípios do Estado do Rio Grande do Sul, utilizando técnicas estatísticas de Análise Multivariada;
- a hierarquização dos municípios através da formação de grupos homogêneos segundo o Índice do Nível de Saúde.

METODOLOGIA

Dois técnicas são aqui utilizadas para tratar do problema: Análise Fatorial⁴ e Análise de Agrupamento ("Cluster Analysis")³.

* Da Secretaria da Saúde e do Meio Ambiente do Rio Grande do Sul e do Departamento de Estatística da Universidade Federal do Rio Grande do Sul e da Pontifícia Universidade Católica do Rio Grande do Sul - Av. Borges de Medeiros, 1501 - 5º andar - 90000 - Porto Alegre, RS.

Análise Fatorial

Nesta técnica o que se faz é substituir um conjunto inicial (não ordenado) de p-variáveis, X_1, X_2, \dots, X_p , correlacionadas, por um conjunto menor de fatores comuns (ou variáveis hipotéticas) que podem não ser correlacionados (fatores ortogonais) ou correlacionados (fatores oblíquos). Deseja-se determinar um número mínimo de fatores necessários para explicar a maior parte da variância do conjunto original de variáveis. A Análise Fatorial é baseada em um modelo estatístico tratando da explicação da estrutura de covariância das variáveis originais. Alguma variância que não é explicada pelos fatores comuns pode ser descrita pelo termo de 'erro' residual (ou fatores específicos)^{4, 8}.

As idéias básicas de Análise Fatorial surgiram por volta de 1900 por Francis Galton e Charles Spearman entre outros, e originada principalmente a partir dos esforços dos psicólogos em obter uma melhor compreensão da 'inteligência'. Os testes de inteligência contêm uma variedade de questões que dependem de uma maior ou menor habilidade verbal, habilidade matemática, memorização, etc. . . . Esta técnica foi desenvolvida para analisar estes testes assim como para determinar se a 'inteligência' é medida por um único fator principal envolvendo todos os testes ou por diversos fatores, mais limitados, medindo atributos como 'habilidade matemática'. Charles Spearman em seu artigo publicado em 1904 intitulado "General Intelligence Objectively Determined and Measured" apresenta sua teoria sobre um fator de inteligência geral, seu famoso fator "g"².

O modelo de Análise Fatorial assume que cada variável observada é representada como uma função linear de um menor nº de fatores comuns (por serem comuns as várias variáveis) mais uma componente de variação residual (fatores específicos). Assim,

$$X_i = \sum_{j=1}^m \lambda_{ij} f_j + e_i \quad (1)$$

$$i = 1, 2, \dots, p$$

onde: o peso λ_{ij} é a carga fatorial da i-ésima variável no j-ésimo fator comum e refletem a importância do j-ésimo fator na composição da i-ésima variável.

f_j são chamados fatores comuns.

e_i são os fatores específicos, descrevem a variação residual específica da i-ésima variável, resíduo que afeta somente X_i .

A expressão (1) é usualmente escrita na notação matricial:

$$\underline{X} = \Lambda \underline{f} + \underline{e} \quad (2)$$

onde:

\underline{X} é o vetor das variáveis observadas (X_i)

\underline{f} é o vetor dos fatores comuns (f_j)

\underline{e} é o vetor dos fatores específicos (e_i)

Λ matriz de cargas fatoriais

A matriz de variâncias e covariâncias das variáveis observadas, Σ , pode ser escrita como,

$$\Sigma = \Lambda \Lambda' + \psi \quad (3)$$

supondo fatores não correlacionados de variância unitária, supondo os p-fatores específicos não correlacionados entre si e em relação aos fatores comuns, e que a variância de e_i é igual a ψ_i (variância residual ou especificidade). Temos então que ψ é matriz diagonal $p \times p$ cujos elementos da diagonal são as especificidades de cada variável X_i .

As expressões (2) e (3) são chamadas de Modelo de Análise Fatorial⁹.

No presente trabalho esta técnica é utilizada para a obtenção de índices, relacionados às condições de saúde, que podem pre-

servar de forma relevante a totalidade das informações obtidas por meio das variáveis originais, para os 232 municípios do Estado do Rio Grande do Sul. Os primeiros fatores serão, então, os índices buscados.

A opção selecionada para a obtenção dos fatores foi o Método do Fator Principal sem Interação (PAI). Este método originou-se no trabalho de Hotelling⁷ (1933) e é extremamente relacionado com a Análise de Componentes Principais.

De posse dos índices poder-se-á obter os escores fatoriais (estimativa dos valores dos índices para cada unidade e são funções das variáveis observadas) para cada município.

Análise de Agrupamento ("Cluster Analysis")³

O objetivo desta técnica multivariada é: dado um número de objetos ou indivíduos cada um descrito, por um conjunto de medidas, obter um esquema de classificação para agrupar os indivíduos em um número de classes tal que os indivíduos dentro das classes são similares em algum aspecto e diferente dos das outras classes. O número de classes e as características de cada classe devem ser determinadas. Em suma, esta técnica agrupa elementos de modo a formar grupos homogêneos dentro deles e heterogêneos entre eles.

O critério para formação dos grupos é dado por medidas de similaridades, dissimilaridades ou distâncias que podem ser, por exemplo, distância euclidiana, χ^2 de associação, D^2 de Mahalanobis, distância métrica de Minkowski e medidas de similaridade criadas especialmente para variáveis em diferentes escalas de medida.

Em geral, os dados brutos a serem submetidos à Análise de Agrupamento consistem de uma matriz X , $N \times p$.

As técnicas de Análise de Agrupamento podem ser usadas como redução de dados, visto que reduzem a informação de um conjunto total de N indivíduos para a informação sobre K - grupos, onde K é muito menor de que N .

Não é conveniente usar número muito grande de variáveis para esta técnica, por isso sugere-se utilizar a técnica de Componentes Principais ou Análise Fatorial para reduzir o número de variáveis usando as primeiras componentes principais ou os primeiros fatores.

Everitt³ (1974) apresenta as principais técnicas para Análise de Agrupamento. Neste trabalho utilizou-se apenas o Método de Ward (método aglomerativo da Técnica Hierárquica de "Cluster").

Ward¹³ (1963) propôs que, em qualquer estágio de uma análise, a perda da informação, a qual resulta do agrupamento de indivíduos em conglomerado ("Cluster"), pode ser medida pela soma total dos quadrados dos desvios de cada ponto, da média do conglomerado ao qual ele pertence. A cada passo da análise, a união de cada possível par de grupo é considerada e os dois grupos, cuja união resulta no mínimo aumento na soma dos quadrados dos desvios, são combinados.

A percentagem de perda da informação é dada pela expressão:

$$\frac{\text{SQD dentro}}{\text{SQD total}} \times 100$$

SQD é a soma dos quadrados dos desvios. SQD dentro é calculado entre os indivíduos do grupo que se está considerando e SQD total é calculada entre todos os indivíduos submetidos a Análise de Agrupamento.

Obtidos, então, os índices e os valores destes para os municípios (escores fatoriais), pela Análise Fatorial, a técnica de Análise de Agrupamento será aplicada aos valores de cada um dos índices.

Neste trabalho são apresentados os grupos homogêneos obtidos com base no Índice do Nível de Saúde.

Escolha das variáveis

A escolha das variáveis selecionadas baseou-se em informações extraídas de relatórios, bem como de opiniões de profissionais ligados à área da Saúde Pública.

As variáveis selecionadas para a elaboração do estudo foram:

- V01 – Mortalidade Proporcional para Menores de 5 anos
- V02 – Mortalidade Proporcional para Maiores de 50 anos (Coeficiente de Swaroop-Uemura)
- V03 – Coeficiente de Mortalidade Infantil Tardia
- V04 – Coeficiente de Mortalidade Infantil
- V05 – % Crianças com Peso ao Nascer Menor 2.700 gramas
- V06 – % partos Não-Hospitalares
- V07 – Mortalidade Proporcional por Doenças Infecciosas e Parasitárias
- V08 – Coeficiente de Mortalidade Materna
- V09 – % de óbitos Mal-Definidos
- V10 – % Menores de um ano vacinados – Vacina Sabin (3ª dose)
- V11 – % Menores de um ano vacinado – Vacina Tríplice (3ª dose)
- V12 – Coeficiente Geral de incidência de Sarampo
- V13 – Coeficiente de Incidência de Sarampo, Menores de 1 ano
- V14 – % Menores de um ano vacinados – Vacina contra Sarampo
- V15 – Coeficiente Incidência de Tuberculose
- V16 – Coeficiente Incidência de Tétano Neo-Natal

A variável Expectativa de Vida ao Nascer pela não viabilidade de obtenção foi desconsiderada. Também variáveis como Coeficiente de Mortalidade Materna e Coeficiente de Incidência de Tétano Neo-Natal, devido ao reduzido número de municípios que apresentavam valores não nulos destas variáveis, foram desconsideradas.

Obtenção dos dados

Os dados para o cálculo das variáveis, em cada município do Estado do Rio Grande do Sul, foram obtidos na Secretaria de Saúde do Estado para o ano de 1979* (último ano onde havia disponibilidade de utilização dos

dados para a maioria das variáveis).

Foram também utilizadas estimativas, com base no registro de Nascimento** e no Censo de 1970⁵, para o número de crianças menores de um ano; e estimativa para a população total de cada município com base nos dados preliminares do Censo de 1980⁶.

Após a coleta e organização dos dados foi utilizado o subprograma "FACTOR" do SPSS ("Statistical Package for Social Sciences")¹⁰.

RESULTADOS E ANÁLISE

Matriz de Correlação

Partindo da matriz de dados brutos, com os valores referentes às 14 variáveis, já especificadas, para cada município, obteve-se a matriz de correlação utilizando-se o Coeficiente de Correlação de Pearson (Tabela 1). Na análise de matriz de correlação, consideraram-se como "boa" correlação aqueles coeficientes maiores ou iguais a 0,35 em valor absoluto (embora, a um nível $\alpha = 0,01$, com 230 graus de liberdade, são estatisticamente significantes todos os coeficientes maiores do que $|\pm 0,168|$).

Na Tabela 2 utilizou-se a seguinte convenção: Os coeficientes de correlação maiores ou iguais $\pm 0,35$ aparecerão na matriz caso contrário aparecerá "0".

A distribuição dos bons coeficientes ilustra quão mais indicadoras ou diagnósticas são certas variáveis em relação às outras dentro de universo total de dados.

Observando a Tabela 2, nota-se que as variáveis:

- V01 – Mortalidade Proporcional para Menores de 5 anos
- V02 – Mortalidade Proporcional para Maiores de 50 anos (Coeficiente de Swaroop-Uemura)
- V03 – Coeficientes de Mortalidade Infantil Tardia
- V04 – Coeficiente de Mortalidade Infantil

* Dados inéditos

** Da Secretaria da Saúde e Meio Ambiente do R.S.

T A B E L A 1

Matriz de correlação para as 14 variáveis - Municípios do Rio Grande do Sul - 1979

Var.	V01	V02	V03	V04	V05	V06	V07	V09	V10	V11	V12	V13	V14	V15
V01	1,000													
V02	-0,758	1,000												
V03	0,580	-0,418	1,000											
V04	0,620	-0,418	0,893	1,000										
V05	0,151	-0,152	0,141	0,199	1,000									
V06	0,036	-0,100	-0,174	-0,261	0,128	1,000								
V07	0,524	-0,476	0,434	0,325	0,063	0,080	1,000							
V09	-0,014	-0,005	-0,090	-0,125	0,186	0,415	-0,101	1,000						
V10	-0,124	0,216	-0,216	-0,160	-0,200	-0,301	-0,272	-0,177	1,000					
V11	-0,122	0,213	-0,218	-0,163	-0,201	-0,294	-0,272	-0,178	0,998	1,000				
V12	0,054	-0,016	0,062	0,076	0,020	-0,133	0,054	-0,090	0,048	0,055	1,000			
V13	0,104	-0,115	0,165	0,158	0,123	-0,137	0,179	-0,100	-0,073	-0,068	0,610	1,000		
V14	-0,173	0,184	-0,267	-0,233	-0,187	-0,205	-0,219	-0,165	0,802	0,796	0,004	-0,058	1,000	
V15	0,129	-0,218	0,325	0,367	0,352	-0,189	0,192	-0,107	-0,147	-0,153	-0,067	0,129	-0,054	1,000

T A B E L A 2

Matriz de correlação resumo para as 14 variáveis - Municípios do Rio Grande do Sul - 1979

Var.	V01	V02	V03	V04	V05	V06	V07	V09	V10	V11	V12	V13	V14	V15
V01	1,000													
V02	-0,758	1,000												
V03	0,580	-0,418	1,000											
V04	0,620	-0,418	0,893	1,000										
V05	0	0	0	0	1,000									
V06	0	0	0	0	0	1,000								
V07	0,524	-0,476	0,434	0	0	0,415	1,000							
V09	0	0	0	0	0	0	0	1,000						
V10	0	0	0	0	0	0	0	0	1,000	0,998				
V11	0	0	0	0	0	0	0	0	0,998	1,000	0,998			
V12	0	0	0	0	0	0	0	0	0	0	1,000	0,610		
V13	0	0	0	0	0	0	0	0	0	0	0,610	1,000		
V14	0	0	0	0	0	0	0	0	0,802	0,796	0	0	1,000	
V15	0	0	0	0,367	0,352	0	0	0	0	0	0	0	0	1,000

apresentam o total máximo de 4 bons coeficientes de correlação.

A variável V07 - Mortalidade Proporcional por Doenças Infecciosas e Parasitárias apresenta 3 bons coeficientes de correlação.

Ainda, estas 5 variáveis estão altamente correlacionadas entre si, sendo a variável Mortalidade Proporcional para Menores de 5 anos (V01) a que apresenta os mais altos valores.

Nota-se, portanto, que as variáveis que medem mortalidade parecem ser medidas discriminantes de um conjunto de dados de Saúde dos municípios do nosso Estado.

A variável V05, % crianças com Peso ao Nascer Menor que 2.700g, para nosso Estado, é a menos discriminante ou diagnóstica

com relação ao conjunto total das variáveis.

Análise Fatorial

O subprograma "FACTOR" fornece as raízes características da matriz de correlação bem como a matriz de cargas fatoriais. Para a obtenção dos Fatores utilizou-se o método de rotação VARIMAX, que foi selecionado por proporcionar uma melhor interpretação dos Fatores, visto que o objetivo da rotação dos fatores é obter uma matriz de cargas mais facilmente interpretáveis ou mais identificáveis com relação a natureza das variáveis observadas^{4,9}.

Obtiveram-se 5 Fatores associados às raízes características maiores do que a unidade (Tabela 3).

TABELA 3

Fatores associados às raízes características maiores do que a unidade

Fator	Raízes Características	% da variância total explicada pelo fator	% acumulada da variância total explicada
1	4,03910	28,9	28,9
2	2,58977	18,5	47,3
3	1,58183	11,3	58,6
4	1,31756	9,4	68,1
5	1,19608	8,5	76,6

TABELA 4

Matriz fatorial rotada VARIMAX

Var.	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5
V01	0.92103	0.00894	0.07557	0.02767	0.04347
V02	-0.82066	0.05968	-0.17034	-0.02438	-0.05523
V03	0.71867	-0.18664	-0.33504	0.02073	0.30222
V04	0.69573	-0.12155	-0.37307	0.02116	0.38765
V05	0.06022	-0.08847	0.34422	0.11454	0.78468
V06	0.04274	-0.22615	0.78943	-0.10272	-0.15659
V07	0.68215	-0.21774	-0.01013	0.10526	-0.07345
V09	-0.05131	-0.09142	0.74576	-0.06451	0.14214
V10	-0.11713	0.95991	-0.13468	-0.00379	-0.08205
V11	-0.11591	0.95802	-0.12952	0.00353	-0.08832
V12	0.02908	0.04254	-0.06167	0.89904	-0.06652
V13	0.10331	-0.06100	-0.08922	0.87722	0.12334
V14	-0.14545	0.88071	-0.07100	-0.01939	-0.04997
V15	0.15417	-0.08599	-0.24073	-0.04347	0.76439

Nota-se que 76,6% da variância total das 14 variáveis observadas é explicada apenas por 5 fatores, representando um alto grau de conservação de informação original.

A matriz fatorial rotada VARIMAX (Ta-

bela 4) é então obtida e a partir das cargas fatoriais (correlação entre as variáveis e os novos fatores), tem-se a seguinte composição dos Fatores:

Fator 1

Variáveis originais com as mais altas correlações (cargas fatoriais) com o Fator 1

Variável	Carga fatorial
V01 – Mortalidade Proporcional Menores de 5 anos	0,92103
V02 – Mortalidade Proporcional Maiores de 50 anos	-0,82066
V03 – Coeficiente Mortalidade Infantil Tardia	0,71867
V04 – Coeficiente Mortalidade Infantil	0,69573
V07 – Mortalidade Proporcional por Doenças Infec. e Parasitárias	0,68215

O Fator 1 é uma componente que agrega variáveis que indicam o nível de saúde de um município, representando a parte comum mais importante das variáveis analisadas e pode ser denominada “Índice do Nível de Saúde”. As mais importantes (mais altas cargas fatoriais) variáveis agregadas por este índice são todas de mortalidade e observa-se que as informações contidas nestas variáveis, antes de mais nada, expressam a condição de saúde de um município. A principal variável deste índice é a Mortalidade Proporcional para Menores de 5 anos. Em texto para discussão sobre Indicadores para a Monitorização do Progresso Visando a Saúde, publicado pela Organização Mundial de Saúde (1980)¹¹, encontra-se o seguinte comentário sobre esta variável:

“Usar somente a taxa de mortalidade infantil (menores de 1 ano) pode não chamar a atenção de forma suficiente para a elevada mortalidade das crianças mais velhas. Os pro-

blemas da má nutrição, em particular, podem não receber atenção suficiente como um fator causal, em especial naqueles países onde a criança de dois anos de idade é mais vulnerável a ela”.

“Em alguns países menos desenvolvidos, as mortes antes de cinco anos de idade são responsáveis por até 60% dos óbitos totais, e por níveis tão baixos quanto 3% nos países mais ricos. Este indicador reflete também as taxas mais elevadas de nascimentos e a proporção mais elevada de população nesta faixa etária nos países menos desenvolvidos; este fato é por si só um outro indicador indireto dos níveis de saúde. A proporção do total de mortes em que se encontrem crianças de idade inferior a cinco anos é, portanto, um indicador composto que reflete as altas taxas de mortalidade da criança, altas taxas de nascimento, e o encurtamento da expectativa de vida”.

Fator 2

Variáveis originais com as mais altas correlações (carga fatorial) com o Fator 2

Variáveis	Carga fatorial
V10 – % Menores de 1 ano vacinados - Vacina Sabin	0,95991
V11 – % Menores de 1 ano vacinados - Vacina Triplice	0,95802
V14 – % Menores de 1 ano vacinados - Vacina contra Sarampo	0,88071

As três principais variáveis deste fator sugerem claramente ser este um fator de cobertura vacinal (ou Assistência Preventiva), podendo ser denominado um "Índice de Imunização".

Fator 3

Variáveis originais com as mais altas correlações (carga fatorial) com o Fator 3

Variáveis	Carga fatorial
V06 - % Partos Não-Hospitalares	0,78943
V09 - % Óbitos Mal-definidos	0,74576

Dada a natureza das mais importantes variáveis ligadas ao Fator 3 pode-se interpretá-lo como um "Índice de Não Assistência Médico-Hospitalar". Na variável V09 estão incluídos os óbitos sem assistência médica.

Fator 4

Variáveis originais com as mais altas correlações (carga fatorial) com o Fator 4

Variáveis	Carga fatorial
V12 - Coeficiente Geral de Incidência de Sarampo	0,89904
V13 - Coeficiente de Incidência Sarampo Menores de 1 ano	0,87722

As mais importantes variáveis do Fator 4 dizem respeito ao número de casos de uma doença, sugerindo, portanto, um "Índice de Morbidade". Poderia ser um "Índice de Morbidade de Sarampo", já que considera apenas esta doença.

Fator 5

Variáveis originais com as mais alta correlações (carga fatorial) com o Fator 5

Variáveis	Carga fatorial
V05 - % de Crianças com Peso ao Nascer Menor que 2.700g	0,78468
V15 - Coeficiente de Incidência de Tuberculose	0,76439

Das duas variáveis com as mais altas cargas contidas neste Fator, a primeira indica o estado nutricional da criança (e indiretamente da mãe) e a segunda o coeficiente de incidência por tuberculose, que se apresenta maior na parte da população mais carente. Portanto, variáveis associadas à condição social. Cabe salientar que a interpretação dos Fatores é subjetiva na análise, ou seja, o nome dado aos Fatores depende de um ponto de vista subjetivo, e que em muitas situações

torna-se difícil esta interpretação; e que, embora para cada Fator apresentado tenham sido colocadas apenas as variáveis mais importantes na sua composição, todas as variáveis entram, com coeficientes maiores ou menores, na formação dos Índices.

Passou-se, a seguir, à obtenção dos escores fatoriais de cada componente para cada município. O programa fornece uma matriz onde cada escore é padronizado tendo média igual a 0 e variância igual a 1.

Utilizando o Fator 1, passou-se a trabalhar com os valores do Índice do Nível de Saúde para cada município.

Quanto mais elevado é o valor do escore, mais baixo é o nível de saúde do município, segundo o índice proposto. Assim, os mais altos valores foram observados para os seguintes municípios:

Código	Município	Escore Fatorial
115	Santo Augusto	3,081126
230	Severiano de Almeida	3,013287
183	Cândido Godoi	2,99985
66	Iraí	2,982009
122	São José do Norte	2,473196

Nesta etapa é utilizada a técnica de Análise de Agrupamento com o objetivo de formar grupos homogêneos com relação ao Índice do Nível de Saúde.

Análise de Agrupamento

Considerou-se o algoritmo que procede o agrupamento obtendo uma "árvore hierárquica de vínculos" ("linkage tree"), que começa com tantos grupos quanto as unidades observadas (cada município é um grupo) e termina com um só grupo que inclui todas as unidades.

Quando as unidades observadas vão sendo combinadas em grupos crescentemente heterogêneos, mais e mais o detalhe inicial é perdido. O ponto principal é a escolha de uma linha de corte que indique um conjunto significativo de grupos.

O programa do computador fornece a percentagem de perda de informação para cada etapa de agrupamento. Colocados em um gráfico estes valores, um possível corte é quando a curva cresce mais rapidamente, havendo um "salto" maior entre estes valores.

A linha de corte foi feita na etapa 213, o que corresponde a um valor de 2,12% de perda de informação. Observou-se que a linha de corte é subjetiva, mesmo utilizando critérios como o que foi exposto anteriormente.

Os grupos formados a partir da análise são descritos a seguir, em ordem de grandeza dos escores do índice do Nível de Saúde: o grupo I formado pelo(s) município(o)s com baixos níveis de saúde e sucessivamente até o grupo XVIII formado pelos municípios com os mais altos níveis de saúde.

Grupo I: Santo Augusto, Severiano de Almeida, Cândido Godoi, Iraí

Grupo II: São José do Norte

Grupo III: Mariano Moro, Sobradinho, Santana do Livramento, Humaitá, Frederico Westphalen

Grupo IV: Chiapeta, Portão, Alegrete

Grupo V: Pejuçara, Tramandaí, Quaraí, Estância Velha, Itaqui, Ilópolis, Alvorada, Palmitinho, Camaquã

Grupo VI: Sapucaia do Sul, Tenente Portela, Rosário do Sul, Viamão, Itatiba do Sul, Rio Grande, Uruguaiana, Planalto, Ciríaco, Passo Fundo

Grupo VII: Nova Palma, Faxinal do Soturno, Caiçara, Espumoso, Mata, Tapes, Jaguarão, Tapera, Erechim, Júlio de Castilhos, Sertão, São Valentim, Bom Jesus

Grupo VIII: Soledade, Canoas, São Leopoldo, Guaíba, Seberi, Santiago, Canela, Rodeio Bonito, Ibiraiaras, Porto Lucena, Igrejinha, Lagoa Vermelha, São Francisco de Paula, Bossoroca, Dom Pedrito, Cacequi, Redentora, Jacutinga, Cachoeirinha, Nonoai, Erval Grande, São José do Ouro, Bagé

Grupo IX: Paraí, Três Passos, Tuparendi, Santo Cristo, Santo Ângelo, Macha-

- dinho, Sapiranga, Nova Prata, Maximiliano de Almeida, Caxias do Sul, Palmeira das Missões, Carazinho, Campo Bom, Pedro Osório, Lavras do Sul, Esmeralda, São Borja, Antônio Prado
- Grupo X: Barros Cassal, Arroio do Tigre, Tucunduva, Porto Xavier, São Lourenço do Sul, Butiá, São Jerônimo, Nova Araçá, Novo Hamburgo, Campo Real (Não-Me-Toque), Cruz Alta, Gravataí, Campinas do Sul, Santa Cruz do Sul, Campo Novo, Roque Gonzales, Panambi, Ijuí, Vacaria, Caçapava do Sul
- Grupo XI: Jaguari, Torres, Sarandi, Porto Alegre, São Vicente do Sul, São Sepé, Augusto Pestana, Mostardas, Guaporé, Alpestre, General Câmara, Barrão do Cotegipe, Alecrim, Três Coroas, São Gabriel, São Martinho, Triunfo, Getúlio Vargas, Constantina, Arroio dos Ratos, Tupanciretã, Baracão, Ivoti, Viadutos, Erval Seco, Canguçu
- Grupo XII: Santo Antônio da Patrulha, Boa Vista do Buricá, São Luiz Gonzaga, Osório, Liberato Salzano, Giruá, Independência, Esteio, Victor Graeff, Braga, Encruzilhada do Sul, Santa Maria, Guarani das Missões, Candelária, Gramado, Colorado, Cerro Largo, Casca, Lageado, Arroio Grande, Cachoeira do Sul, Cacique Doble, Farroupilha, Rondinha, São Pedro do Sul, Rolante, Barra do Ribeiro, Santa Vitória do Palmar, São Paulo das Missões, Nova Bréscia, Taquara, Santa Rosa, Vicente Dutra, Dois Irmãos, Fontoura Xavier, Venâncio Aires, São Francisco de Assis, Agudo
- Grupo XIII: Marcelino Ramos, Horizontina, Três de Maio, Salvador do Sul, São Marcos, Flores da Cunha, Coronel Bicaco, Pinheiro Machado, Montenegro, Rio Pardo, Ibirubá, São Sebastião do Caí, Putinga, Paim Filho, Crissiumal, Campinas das Missões, São Nicolau, Pelotas, Cambará do Sul, Tapejara, Ibiaçá, Dona Francisca, Taquari, Dom Feliciano, Marau, Ajuricaba, Gaurama, Chapada, Ronda Alta, Santa Bárbara do Sul, Bento Gonçalves.
- Grupo XIV: Sanaduva, Vera Cruz, Condor, Garibaldi, Santana da Boa Vista, Formigueiro, Restinga Seca, Miraguaí, Encantado, Caibaté, Bom Retiro do Sul.
- Grupo XV: Anta Gorda, Veranópolis, Estrela, Muçum, Catuípe.
- Grupo XVI: Nova Petrópolis, Arvorezinha, Piratini, Arroio do Meio, Selbach, Cruzeiro do Sul, Serafina Corrêa, Herval do Sul, Carlos Barbosa, Aratiba.
- Grupo XVII: Santo Antônio das Missões, David Canabarro, Roca Sales.
- Grupo XVIII: Nova Bassano e Feliz.

A Figura apresenta a localização geográfica dos grupos de municípios que apresentaram maior carência nas condições de saúde.

Nota-se que as zonas de piores condições de saúde são municípios de minifúndios (norte do Estado), a região da campanha (municípios latifundiários), municípios litorrâneos e alguns municípios da Região Metropolitana de Porto Alegre.

Em geral os municípios da Região Serrana do Estado apresentam-se como os de melhores condições.

Na utilização do teste χ^2 para verificar a associação entre o grupo a qual o município pertence e o tamanho do município (considerando sua população) foi encontrada não significância ($0,10 < p < 0,20$), ou seja, não existe associação. Esta é uma comprovação importante para o tipo de variáveis utilizadas.

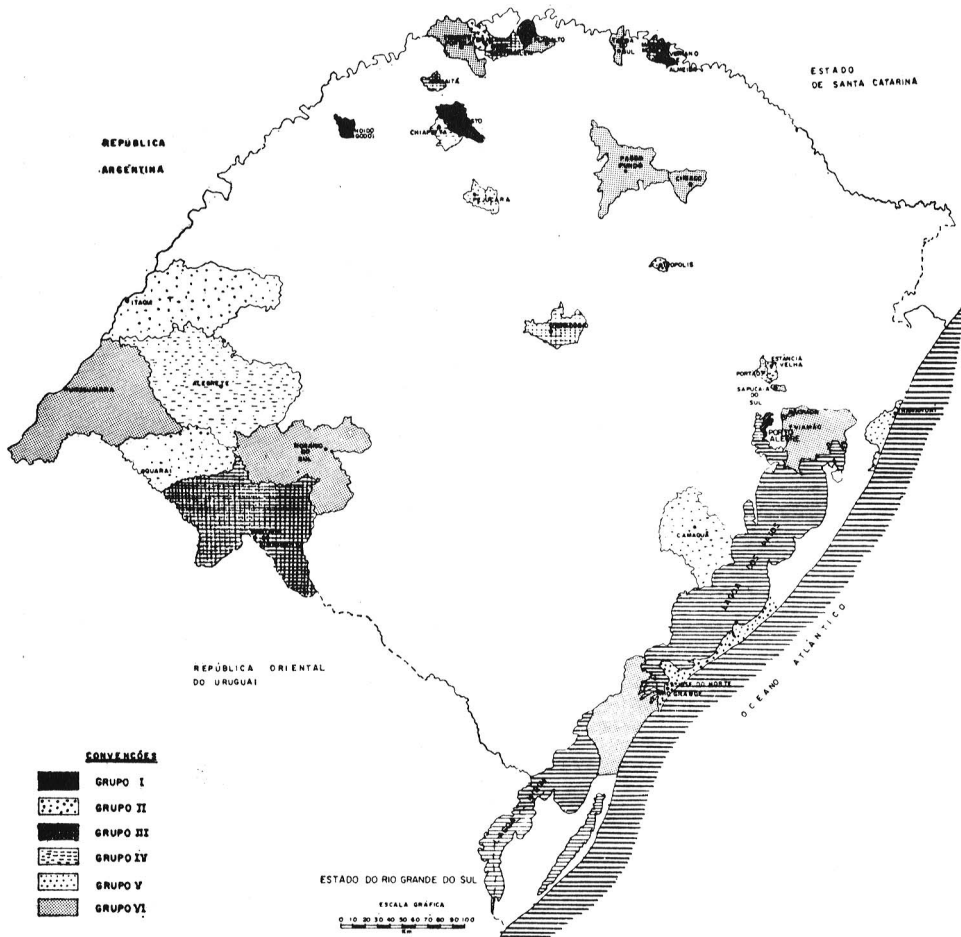


Figura – Localização geográfica dos grupos de municípios com os mais “altos” valores no Índice do Nível de Saúde.

CONCLUSÕES

Utilizando a técnica de Análise Fatorial que simplificou a informação contida nas variáveis observadas, relacionadas com a área de Saúde Pública, obteve-se o índice do Nível de Saúde.

Aos valores deste índice para os municípios do Estado do Rio Grande do Sul, foi aplicada a técnica de Análise de Agrupamento formando dezoito grupos de municípios. Estes grupos são caracterizados por possuírem homogeneidade interna e por serem

diferenciados entre si, ao máximo.

Analisando os demais Fatores foram obtidos outros índices que ficaram assim denominados:

- Índice de Imunização (Fator 2)
- Índice de Não Assistência Médico-Hospitalar (Fator 3)

Este trabalho servirá de grande valia em atuações práticas na área de Saúde Pública do Estado. Assim, por exemplo, foram apre-

sentados os municípios que merecem maior atenção para a melhoria do seu nível de saúde. Sugeriu-se a utilização do Índice do Nível de Saúde como um indicador composto (que guarda em si a informação de todas as variáveis) e por isso mais completo para avaliar a saúde dos municípios, e a maior atenção a variáveis importantes (como Mortalidade Proporcional para Menores de 5 anos) por serem de maior poder de diagnóstico e mais discriminantes que outras menos diagnósti-

cas (como % de Peso ao Nascer Menor que 2.700g.) para nosso Estado.

Ao se apresentar a hierarquização dos municípios do Estado do Rio Grande do Sul, em grupos homogêneos, possibilitará um controle da melhoria dos municípios quanto a saúde e a sua assistência no decorrer dos anos. Este controle e avaliação vem ao encontro às metas mínimas regionais de saúde apresentadas pela Organização Panamericana da Saúde¹².

POSSOLI, S. [Multivariate analysis techniques for the assessment of the health status of the "counties" of State of Rio Grande do Sul, Brazil]. *Rev. Saúde públ.*, S. Paulo, 18: 288-300, 1984.

ABSTRACT: The correlation structure of variables in the Public Health field were studied in order to obtain a "Health Status Index" for the "counties" (municípios) of Rio Grande do Sul State. By using the value of this index for each one of the counties it was possible to built up 18 homogeneous groups (Clusters) which were ordered in the reverse magnitude of their health status index. Additional indices were found: "Imunization Index" and "Non-Medical-Hospital Care Index". The variable proportional mortality for the under 5 year-olds, was the one that showed itself to be the most sensitive diagnostic discriminator among the series of variables that were studied, and the birth weight of less than 2,700g the least sensitive.

UNITERMS: Health Status Index. Counties of Rio Grande do Sul, health conditions. Multivariate analysis.

REFERÊNCIAS BIBLIOGRÁFICAS

1. CAMINHA, J. A. et al. Indicadores de saúde. *Rev. Med. ATM*, Porto Alegre, 9: 25-9, 1974.
2. CHATFIELD, C. & COLLINS, A. J. *Introduction to multivariate analysis*. London, Methuen, 1980.
3. EVERITT, B. *Cluster analysis*. London, Heinemann Educational Books, 1974.
4. FACHEL, J. M. G. Análise fatorial. São Paulo, 1976. [Dissertação de Mestrado - Instituto de Matemática USP]
5. FUNDAÇÃO IBGE. *Censo demográfico: Rio Grande do Sul, 1970*. Rio de Janeiro, 1970. (Recenseamento Geral do Brasil, 89, 1970, Série Regional, v. 1, t. 21).
6. FUNDAÇÃO IBGE. *Censo demográfico - dados gerais, migração, instrução, fecundidade, mortalidade: Rio Grande do Sul, 1980*. Rio de Janeiro, 1982. (Recenseamento Geral do Brasil, 99, 1980, v. 1, t. 4, n. 22).
7. HOTELLING, H. Analysis of a complex of statistical variables into principal components. *J. educ. Psychol.*, 24: 417-41, 490-520, 1933.
8. MAXWELL, A. E. *Multivariate analysis in behavioral research*. London, Chapman and Hall, 1977.
9. MORRISON, D. F. *Multivariate statistical methods*. New York, McGraw-Hill, 1976.

POSSOLI, S. Técnicas de análise multivariada para avaliação das condições de saúde dos municípios do Rio Grande do Sul, Brasil. *Rev. Saúde públ.*, S. Paulo, 18: 288 - 300, 1984.

10. NIE, N. et al. Factor analysis. In: Nie, N. et al. *Statistical Package for Social Sciences*. New York, McGraw-Hill, 1978. p. 468-97.

13. WARD apud EVERITT, B. *Cluster analysis*. London, Heinemann Educational Books, 1974.

11. ORGANIZAÇÃO MUNDIAL DA SAÚDE. *Indicadores para a monitorização do progresso visando a saúde*; Documento do Comitê Executivo da Organização Mundial de Saúde, Genebra, 1980.

12. ORGANIZACION PANAMERICANA DE LA SALUD. *Salud para todos en el año 2.000: plano de acción para la instrumentación de las estrategias regionales*. Washington, D. C., 1982. (Documento Oficial Nº 179).

Recebido para publicação em 14/09/1983
Reapresentado em 28/05/1984
Aprovado para publicação em 06/06/1984

E R R A T A

REVISTA DE SAÚDE PÚBLICA,
18(2), 1984

p. 131, Tabela 6 – última linha

Onde se lê: Sintomas. . .

Leia-se: Demais grupos

18(4), 1984.

p.300, Referência Bibliográfica nº 13

Onde se lê: 13. WARD apud EVERITT, B. *Cluster analysis*. London, Heinemann Educational Books, 1974.

Leia-se: 13. WARD Jr., J. H. Hierarchical grouping to optimize an objective function. *J. Amer. Statist. Assoc.* 58:236-44, 1963.

O arquivo disponível sofreu correções conforme ERRATA publicada no Volume 18 Número 5 da revista.