

## Análise de componentes principais dos indicadores de qualidade de vida

**Bruno Rogério Locatelli dos Santos      João Vitor Teodoro**

Universidade Federal da Grande Dourados – FACET, UFGD  
79.804-970, Dourados, MS

E-mail: brunolocatelli@ufgd.edu.br, joaoteodoro@ufgd.edu.br

**Ana Carolina Torelli Marquezini Faccin**

Universidade Federal da Grande Dourados – FCH, UFGD  
79.804-970, Dourados, MS

E-mail: caroltorelli.faccin@gmail.com

### RESUMO

Esse trabalho teve por objetivo propor um modelo para representar os indicadores sociais: Esperança de vida ao nascer em 2011 (anos), população subnutrida em 2005 (%), população com acesso à água potável em 2008 (%), população com acesso à rede sanitária em 2008 (%) e calorias consumidas em 2005 (Kcal/dia/habitante); os dados populacionais: densidade demográfica em 2010 (habitantes/km<sup>2</sup>) e taxa média anual do crescimento da população entre 2005 e 2010 (%); e os dados econômicos: produto interno bruto per capita em 2010 (US\$) e gastos públicos com saúde em 2007 (% do PIB), os dados em questão foram extraídos da base do Instituto Brasileiro de Geografia e Estatística (IBGE).

Foram empregados os métodos multivariados para análises de componentes principais utilizando o pacote estatístico SAS [2]. Na análise de componentes principais, que são combinações lineares das variáveis em estudo, foram utilizadas ferramentas com metodologias multivariadas [2], com o objetivo de reduzir a dimensionalidade do problema de 9 para 2 dimensões.

Para as nove variáveis, foram estudados 137 países, cuja inclusão foi dependente da disponibilidade de informações. Para simplificação, as variáveis serão representadas pelas siglas: Esperança de vida ao nascer (EVN); População subnutrida (SUB); População com acesso à água potável (AGU); População com acesso à rede sanitária (SAN); Calorias consumidas (CAL); Densidade demográfica (DEM); Taxa média anual do crescimento da população (CRE); Produto interno bruto per capita (PIB) e Gastos públicos com saúde (SAU).

As variáveis apresentadas estão em unidades de medida diferentes, e por esta razão, os dados foram transformados de forma a apresentar médias 0 e desvios padrão 1.

Análise de componentes principais é um dos métodos multivariados mais simples, seu objetivo é tomar p variáveis e encontrar uma combinação linear para produzir índices que sejam não correlacionados na ordem de sua importância, e que descreva a variação nos dados [1].

Tabela 1 – Coeficientes de correlação entre as variáveis

	CAL <sup>1</sup>	SUB <sup>1</sup>	AGU <sup>1</sup>	SAN <sup>1</sup>	DEM <sup>2</sup>	CRE <sup>2</sup>	PIB <sup>3</sup>	SAU <sup>3</sup>
EVN <sup>1</sup>	<b>0.7599</b>	<b>-0.6914</b>	<b>0.7801</b>	<b>0.8076</b>	0.1840	-0.5959	0.5871	0.5329
CAL <sup>1</sup>		<b>-0.8486</b>	<b>0.7555</b>	<b>0.7471</b>	0.0975	-0.5676	0.6644	0.6269
SUB <sup>1</sup>			<b>-0.7624</b>	<b>-0.6929</b>	-0.0288	0.5662	-0.4142	-0.4202
AGU <sup>1</sup>				<b>0.8324</b>	0.1595	-0.6705	0.4878	0.5238
SAN <sup>1</sup>					0.1361	-0.6647	0.5236	0.5254
DEM <sup>2</sup>						-0.1247	0.0710	0.0139
CRE <sup>2</sup>							-0.2917	-0.4485
PIB <sup>3</sup>								0.6386

1 – Indicadores sociais; 2 – Dados populacionais; 3 – Indicadores econômicos

A análise de componentes principais apresentou um primeiro componente (PCR1), que explica 60,38% da variação dos dados e um segundo componente (PCR2), que explica 11,51% da variação. Assim os dois componentes juntos explicam 71,89% da variabilidade dos dados.

Tabela 2 – Coeficientes de correlação entre as variáveis originais e os dois primeiros componentes principais

Variável	Correlação com PCR1	Correlação com PCR2
EVN <sup>1</sup>	<b>0.8865</b>	0.0862
CAL <sup>1</sup>	<b>0.9133</b>	-0.0948
SUB <sup>1</sup>	<b>-0.8317</b>	0.0269
AGU <sup>1</sup>	<b>0.8967</b>	0.1117
SAN <sup>1</sup>	<b>0.8923</b>	0.0804
DEM <sup>2</sup>	0.1557	<b>0.8966</b>
CRE <sup>2</sup>	<b>-0.7320</b>	-0.1669
PIB <sup>3</sup>	<b>0.6874</b>	-0.2725
SAU <sup>3</sup>	<b>0.6997</b>	-0.3054

O primeiro componente principal PCR1 está altamente associado aos indicadores sociais, pois, há alta correlação entre estas variáveis e PCR1, e tem associação razoavelmente alta com as variáveis de indicadores econômicos e crescimento da população (CRE).

Valores altos para as variáveis, esperança de vida ao nascer (EVN), calorias consumidas (CAL), população com acesso à água potável (AGU), população com acesso à rede sanitária (SAN), produto interno bruto per capita (PIB), gastos públicos com saúde (SAU), valores baixos para população subnutrida (SUB) e taxa média anual do crescimento da população (CRE), fazem com que PCR1 apresente valores altos. Ou seja, um país com alto valor de PCR1 apresenta ótimos índices sociais, bom investimento em saúde e baixo crescimento populacional, assim podemos denominar PCR1 por qualidade de vida.

O segundo componente principal PCR2 está associado à variável densidade demográfica (DEM), assim, valores altos para DEM também tornam PCR2 altos e valores baixos para densidade demográfica propiciam PCR2 baixos, podendo ser chamado de densidade demográfica.

Este modelo pode auxiliar na decisão e instituição de políticas públicas voltadas ao melhoramento das condições de vida da população, especialmente em países que se mostrarem carentes.

**Palavras-chave:** *Análise multivariada, Componentes principais, Demografia.*

## Referências

- [1] B.F.J. Manly, “Métodos Estatístico Multivariados: uma introdução”; tradução Sara Ianda Carmona – 3.ed.- Bookman, Porto Alegre, 2008
- [2] Statistical Analysis System. SAS/STAT user'guide, version 8.0 ed. Cary: 1999. v.1, 943p.