**DRAFT SOBRE ANÁLISE FATORIAL E INDICE SINTETICO**

**CLIENTE: Carlos Benassuly Maués Filho**

**AUTHOR: Mário Diego Rocha Valente**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Será apresentado uma revisão bibliográfica a cerda do Modelo Estatístico Multivariado chamado Análise de Fatores, Análise Fatorial, ou Factor Analysis. Descrevendo a formulação matemática, os métodos para extrair os fatores, sendo utilizado o de Máxima Verossimilhança, os métodos de rotação dos fatores, no caso o Varimax. Para estimar os escores fatorial utilizou-se o de Regressão, em seguida é calculado um Índice Sintético, utilizando uma combinação linear entre o peso fatorial de cada fator, com base no percentual de variância explicada e os escores fatoriais padronizados. Ao final é apresentando a hierarquização dos municípios em estudo.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Modelo Estatístico de Análise Fatorial**

O modelo estatístico usado na análise fatorial explica uma estrutura de correlação existente entre um conjunto de informações, diretamente observado por meio de combinação linear de variáveis, as quais não são diretamente observadas, denominados fatores comuns, acrescidas de componente residual. Um modelo de análise fatorial pode ser apresentado na seguinte forma conforme (**DILLON e GOLDSTEIN*,* 1984**).

Sejam variáveis aleatórias, m fatores comuns; m<p.

 (1)

Em que:

**Xi** = variáveis originais ou observadas;

****= Cargas fatoriais da *i*-ésima variável no *j*-ésimo fator comum e refletem a importância do *j*-ésimo fator na composição da *i*-ésima variável.

**Fj** = variáveis não-observáveis ou variáveis latentes chamadas de fatores comuns;

**** = Fatores específicos que descrevem a variação residual específica da *i-ésima* variável (resíduo que afeta somente X*i)*. É a parte da variável X*i* que não é explicada pelos fatores comuns

Em notação matricial pode-se escrever o modelo como:

 (2)

Onde:

 Vetor de Variáveis Aleatórias Originais ou Observadas

Vetor de Fatores Comuns

Vetor de Cargas Fatoriais

Vetor de Fatores Específicos ou Erros Aleatórios

**Métodos de Estimação dos Fatores**

Há na literatura vários métodos para a estimação de fatores tais como: Method Unweighted Least Squares, Method Generalized Least Squares, Method Maximun Likelihood, Method Principal Axis Factoring, Method Principal Components, Method Alpha Factoring, Method Image Factoring.

Sendo o **Method Unweighted Least Squares** um dos métodos de extração que minimiza a soma das diferenças quadráticas entre a matriz de dados e a matriz de correlação reproduzida, ignorando as diagonais. O **Method Generalized Least Squares** idem ao anterior, mas neste caso a correlação é pesada pelo inverso das suas singularidades, assim como as variáveis com altas singularidades são tomadas com peso menor que aquelas com menor singularidades.

Com relação ao **Method Maximun Likelihood (Máxima Verossimilhança)**, este cria parâmetros estimados como sendo mais prováveis para produzir a matriz de correlação observada, se a amostra pode ser caracterizada por uma distribuição normal multivariada. As correlações são pesadas pelo inverso das singularidades das variáveis, pelo emprego de um algoritmo “iterativo”. (**JOHNSON e WICHERN, 1998**).

O **Method Principal Axis Factoring** parte da matriz de correlação original com os coeficientes de correlações múltiplos colocados na diagonal como estimativas iniciais das *comunalidades*. Estes fatores obtidos são usados para estimar as novas *comunalidades*, que são recolocadas no lugar das velhas na diagonal. As Iterações continuam até a ocorrerem mudanças nas *comunalidades* partindo da primeira até a seguinte, buscando satisfazer o critério de convergência de extração.

O **Método das Componentes Principais** é usada para obter uma combinação linear não-correlata das combinações das variáveis mensuradas, obtendo-se as soluções dos fatores, ela pode ser usada quando a matriz de correlação é singular. **Method Alpha Factoring** é um método de extração que considera as variáveis na análise como uma amostra do universo potencial de variáveis. Ele maximiza a confiabilidade ou fidedignidade alfa de Cronbach dos fatores. Já o **Method Image Factoring** é um método fatorial de extração desenvolvido por Guttman e está baseado na Teoria de Imagens. A parte comum da variância, chamada de imagem parcial, é definida como uma regressão linear sobre as restantes, preferivelmente que a função dos fatores hipotéticos.

**Métodos de Rotação do Fatores**

Uma ferramenta importante na interpretação de fatores é a rotação fatorial. O termo rotação significa exatamente o que sugere. Especificamente, os eixos de referência dos fatores são rotacionados em torno da origem até que alguma outra posição seja alcançada. As soluções de fatores não-rotacionadas extraem fatores na ordem de sua importância. O primeiro fator tende a ser um fator geral com quase toda a variável com carga significativa, e explica a quantia maior de variância. O segundo fator e os seguintes são então baseados na quantia residual de variância. Cada fator explica porções sucessivamente menores de variância. O efeito final de rotacionar a matriz fatorial é redistribuir a variância dos primeiros fatores para os últimos com o objetivo de atingir um padrão fatorial mais simples e teoricamente mais significativo.

O caso mais simples de rotação é uma rotação ortogonal, na qual os eixos são mantidos a 90º. Também é possível rotacionar os eixos sem manter o ângulo de 90 graus entre os eixos de referência. Procedimento de rotação se chama rotação oblíqua.

Para rotação ortogonal dos fatores existem vários tipos na literatura como: **Varimax** (mais usado) é um método de rotação ortogonal que minimiza o número de variáveis que cada agrupamento terá. Ele simplifica a interpretação dos fatores. **Quartimax** é um método que minimiza o número de fatores necessários para explicar cada variável. Ele simplifica a interpretação das variáveis obtidas. O **Equamax** é também um método que busca uma combinação dos outros (varimax e quartimax). O número de variáveis obtido terá carga fatorial maior e o número de fatores será minimizado.

A rotação varimax é uma das rotações ortogonais mais utilizadas em análise fatorial. Intuitivamente, ela busca soluções nas quais se busca maximizar as correlações de cada variável com apenas um fator.

Sejam as cargas fatoriais rotacionadas. Defina

 e  (3)

Note que  pode ser interpretada como a proporção da Comunalidade de que é explicada pelo fator j. A matriz de rotação será escolhida de sorte a maximizar.

 ,  (4)

O método varimax é um processo em que os eixos de referência dos fatores são rotacionados em torno da origem até que alguma outra posição seja alcançada. Intuitivamente, ela busca soluções nas quais se busca maximizar as correlações de cada variável com apenas um fator. O objetivo é redistribuir a variância dos primeiros fatores para os demais e atingir um padrão fatorial mais simples e teoricamente mais significativo (**REIS, 2001; HAIR, 2005**).

**Métodos de Estimação dos Escores Fatoriais**

Existem na literatura vários métodos de estimação dos escores fatoriais para cada elemento amostral, tais como: o **Método dos Mínimos Quadrados Ponderados**, **Método de Bartlett**, **Método de Anderson Rubin** e o **Método de Regressão**.

Sendo o Método de Regressão o mais usado para estimar os e*scores* dos coeficientes dos fatores. Os e*scores* gerados têm média 0 e variância igual ao quadrado da correlação múltipla entre os e*scores* dos fatores estimados e os valores verdadeiros dos fatores. Os e*scores* devem ser igualados com os fatores ortogonais.

Em relação ao Método de Bartlett usado também para estimação dos escores dos coeficientes dos fatores. Os escores produzidos têm média de zero. A soma dos quadrados de um fator é feita sobre a extensão das variáveis minimizadas. Já o Método de Anderson Rubin é similar ao de Bartlett a diferença está em garantir a ortogonalidade dos fatores estimados. Os escores gerados têm uma média de 0, desvio padrão de 1,0 e são não correlatos.

Para a determinação dos escores fatoriais, estimou-se através de algum Método de Extração a matriz de escores fatoriais após algum Método de Rotação da estrutura fatorial inicial, usando especificamente o **Método de Regressão**. Para cada fator , o i-ésimo escore fatorial extraído é definido por , expresso da seguinte forma (**DILLON; GOLDSTEIN, 1984**):

 (5)

Em que:

**:** são os coeficientes de regressão estimados para os n escores fatoriais comuns;

**:** são as *n* observações das p variáveis observadas.

A variável  não é observável, mas pode ser estimada por meio das técnicas de análise fatorial, utilizando-se a matriz de observações do vetor **x** de variáveis observáveis. Em notação matricial, a equação 2, torna-se:

 (6)

Na equação 6, **F** é a matriz da regressão estimada a partir dos n escores fatoriais e que pode ser afetada tanto pela magnitude quanto pelas unidades de medida das variáveis **x**. Para contornar este tipo de problema, substitui-se a variável **x** pela variável padronizada **w**, dada pela razão entre o desvio em torno da média e o desvio-padrão de **x**, como a seguir:

 (7)

Com esses valores, modifica-se a equação 11 para gerar a equação 12.

 (8)

Na equação 7, a matriz de pesos  , com a **q** colunas e **p** coeficientes de regressão padronizados, substitui **b,** dado que as variáveis estão padronizadas em ambos os lados da equação. Pré-multiplicando ambos os lados da equação 7 pelo valor , em que n é o número de observações e é a matriz transposta de w, obtém-se:

 (9)

A matriz se constitui na matriz de variáveis intercorrelacionadas ou matriz de correlação entre as observações da matriz **x**, designada por **R.** A matriz representa a correlação entre os escores fatoriais e os próprios fatores, denotada por . Reescrevendo a equação 8, tem-se que:

 (10)

Se a matriz **R** for não-singular, pode-se pré-multiplicar ambos os lados da equação 9 pela inversa de **R**, obtendo-se:

 (11)

Substituindo o vetor  na equação 7, obtém-se o escore fatorial associado a cada observação, como a seguir:

 (12)

**Viabilidade da Análise Fatorial**

Para aferir a viabilidade da aplicação de uma análise fatorial a um conjunto de dados, é necessário calcular algumas medidas iniciais. As principais medidas aplicadas são: **Matriz Anti-Imagem, *Teste de Esfericidade de Bartlett, Teste de Kaiser-Meyer-Olkin (KMO) e* Measure of Sampling Adequacy (MSA)*.***

Uma das premissas de uma análise fatorial é que exista uma estrutura de dependência clara entre as variáveis envolvidas. No modelo estudado, essa estrutura é expressa através da matriz de covariância ou de correlação. A existência de tal estrutura implica que uma variável pode, dentro de certos limites, ser prevista pelas demais. Para verificar esse fato, pode-se calcular os coeficientes de correlação parcial entre os pares de variáveis, eliminado o efeito das demais variáveis.

### **Matriz Anti-Imagem**

A matriz anti-imagem é construída com esses coeficientes com sinais invertidos, muitas vezes, coloca-se na diagonal principal dessa matriz os indicadores, MAS (Measure of Sampling Adequacy) (**BARROSO e ARTES, 2003**).

### **Teste de Esfericidade de Bartlett**

Esse teste avalia a significância geral da matriz de correlação, ou seja, testa se todas as variáveis oriundas de diversos setores possuem uma possível relação em comum (***Dillon; Goldstein, 1984; Reis, 2001***).

O teste de Bartlett testa as seguintes hipóteses nulas

: R=I ou : 



ou (13)



Em que |R| é o determinante da matriz de correlação amostral,  é a variância explicada por cada fator, n é o número de observações e p é o número de variáveis. A estatística tem uma distribuição assintótica de  com (0,5p(p-1)) graus de liberdade.

### **Teste de Kaiser-Meyer-Olkin (KMO)**

Este teste é usado para compara as correlações parciais entre os pares de variáveis sem o efeito das demais, ou seja, testa se duas a duas variáveis possuem algum tipo de relação entre si (***Dillon; Goldstein, 1984; Kaiser, 1970; Reis, 2001***).

 (14)

Em que  é o coeficiente de correlação observado entre as variáveis i e j e  é o coeficiente de correlação parcial entre as mesmas variáveis que é, simultaneamente, uma estimativa das correlações entre os fatores, eliminado o efeito das demais variáveis. Os  deverão assumir valores próximos de zero, uma vez que se pressupõe que os fatores são ortogonais entre si.

Conforme, (Kaiser; Rice, 1974), os valores do teste são classificados da seguinte forma:

Tabela 2. Classificação Geral do Teste de adequabilidade KMO.

|  |  |
| --- | --- |
| **Teste de KMO** | **Classificação** |
| 0,90 |--- 1,00 | Excelente |
| 0,80 |--- 0,90 | Ótimo |
| 0,70 |--- 0,80 | Bom |
| 0,60 |--- 0,70 | Regular |
| 0,50 |--- 0,60 | Ruim |
| 0,00 |--- 0,50 | Inadequado |

Valores do teste abaixo de 0,70 não são inaceitáveis.

### **Measure of Sampling Adequacy (MSA)**

Essa medida é bastante similar ao KMO. Novamente, deseja-se verificar a possibilidade de existir uma estrutura fatorial nos dados. Na verdade, a, MSA será calculada separadamente para cada variável. O objetivo é verificar se uma dada variável pode ser explicada pelas demais (o que é esperado num modelo fatorial). Valores baixos de MSAi () são indícios de que a respectiva variável pode ser retirada da análise sem maiores prejuízos (**BARROSO; ARTES, 2003**).

 (15)

Como medida resumo, pode-se calcular a média dos MASi para ter uma idéia do desempenho do conjunto das variáveis,

 (16)

**Análise Inferencial**

Inicialmente, avaliou-se a viabilidade da análise fatorial a partir da matriz de correlações, no caso a base de dados é composta por 21 variáveis.

**Tabela 1**. Variáveis Originais utilizadas para Realização da Análise Fatorial Exploratória.

|  |  |
| --- | --- |
| **Itens** | **Variáveis** |
| Var1 | População |
| Var2 | Nº de Empregos Formais |
| Var3 | PIB\_(por mil) |
| Var4 | PIB Per Capita (R$ por pessoa) |
| Var5 | Nº de Agências Bancárias |
| Var6 | Incremento do Desmatamento |
| Var7 | Estabelecimentos Agrícolas |
| Var8 | Rebanho Bovino (Cabeças) |
| Var9 | Área Colhida (LP)(hectare) |
| Var10 | Área Colhida (LT)(hectare) |
| Var11 | Educação Básica |
| Var12 | ICMS |

O primeiro passo é um exame visual das correlações, identificando as que são estatisticamente significantes, com isso, verificou-se que, existe um número substancial de correlações maiores que 0,30 (**Gorsuch, 1983**), sugerindo possíveis relações entre as variáveis.

A inspeção da matriz de correlações revela ter coeficiente de correlação de Pearson com valores acima de 0,30. Isso fornece uma base adequada para seguir para o próximo nível, o exame empírico da adequação para a análise fatorial tanto para uma base geral (12 itens) quanto para cada variável.

O próximo passo é avaliar a significância geral da matriz de correlação com o Teste de Esfericidade de Bartlett (**Mingoti, 2005**), neste caso, as correlações em geral são significantes ao nível de 0,0001 (ver Tabela 2). Entretanto, isso testa apenas a presença de correlações não nulas, e não o padrão dessas correlações, assim, para testar se todas as características oriundas de diversos apartamentos possuem uma possível relação em comum, utilizou-se o Teste de Kaiser-Meyer-Olkin (KMO) (**Mingoti, 2005**), ou seja, que vai testar as correlações parciais entre os pares de variáveis sem o efeito das demais, conforme Tabela 2, verificou-se que, neste caso ocupa um intervalo aceitável (abaixo de 0,50) conforme **Kaiser e Rice (1974)**, com um valor de 0,82 classificados como Bom.

Tabela 2. Medidas de avaliação da Adequação da Análise Fatorial para as 9 variáveis significantes.

|  |  |  |
| --- | --- | --- |
| **Estatísticas** | **Coeficientes** | **P-valor** |
| Teste de Esfericidade de Bartlett |  |  |
| Teste de Kaiser-Meyer-Olkin (KMO) |  |  |

Todas essas medidas indicam que o conjunto de variáveis é adequado à análise fatorial, e a análise pode prosseguir para os próximos estágios. Assim, aplicou-se a análise de fatores aos dados de imóveis, utilizando-se o método matemático das componentes principais para extração dos fatores via rotação ortogonal do tipo varimax para obter melhores combinações e usando como critérios para a escolha do número de fatores a extrair, o Critério *Scree Test*, Critério da Raiz Latente e Critério da Percentagem da Variância Explicada.

A Tabela 3, mostra, os autovalores, as percentagens das variâncias explicadas e acumuladas, calculadas com rotação ortogonal, contendo as informações sobre todas as 15 variáveis em estudo, utilizou-se os valores para análise (após-rotação ortogonal) dos 4 primeiros fatores extraídos. Percebeu-se que um modelo com apenas quatro fatores seria suficiente para representar a estrutura de covariância inicial, com **26,32%** de perda de informações e que expliquem **73.68%** da variabilidade total dos dados originais. Sendo que o primeiro fator explica **34.42%**, o seguindo fator explica **17,88%**, o terceiro fator explica **13,83%,** e o quarto fator explica **7.55%**. Com isso, buscou-se identificar as variáveis que mais influenciam em cada fator, ou seja, as que possuem maiores cargas fatoriais.

**Tabela 3**. Resultado dos Autovalores para Extração de Fatores, Componentes e Variância Total explicada pelos fatores com Rotação Ortogonal do Tipo varimax.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Itens** | **Variáveis** | **H²** | **MAS**  **(≥ 0.50)** | **Cargas Fatoriais** | | | **Alpha**  **Cronbach** |
| **Fator** | **Fator** | **Fator**  **3** |
| **1** | **2** |
| 1 | População | 0.96 | 0.83 | 0.97 |  |  | **0.89** |
| 2 | Nº de Empregos Formais | 0.93 | 0.84 | 0.96 |  |  |
| 3 | Nº de Agências Bancárias | 0.80 | 0.80 | 0.94 |  |  |
| 4 | Educação Básica | 0.91 | 0.89 | 0.79 |  |  |
| 5 | Inverso Inc. Desmatamento | 0.77 | 0.52 |  | 0.87 |  |
| 6 | Rebanho Bovino | 0.78 | 0.54 |  | 0.87 |  |
| 7 | Estabelecimentos Agrícolas | 0.99 | 0.59 |  | 0.99 |  |
| 8 | PIB\_(por mil) | 0.80 | 0.89 |  |  | 0.89 |
| 9 | ICMS | 0.68 | 0.92 |  |  | 0.69 |
| **Ʃ Autovalores** | | | | **4.70** | **1.56** | **1.01** | **7.27** |
| **% Variância Explicada** | | | | **52.22** | **17.37** | **11.26** | **80.85** |

**Notas**:

H² = são as medidas de comunalidades;

MAS = medidas de adequação da amostra;

Alpha = coeficiente de correlação alpha de Cronbach;

Com os resultados obtidos na Tabela 3, pode-se observar que o primeiro fator possui pesos mais altos nas variáveis: População, Número de Empregos Formais, Produto Interno Bruto (PIB), Número de Agências Bancárias, Número de Rebanho Bovino, Número de Estabelecimentos com Educação Básica e Volume de Crédito Rural.

O segundo fator possui pesos mais altos nas variáveis: Pessoa Física e Jurídica, Financiamento de Porte Mini e Pequeno. O terceiro fator tem pesos maiores nas variáveis: Financiamento de Porte Médio e Grande. O quarto fator possui pesos mais altos nas variáveis: ICMS e Área Colhida por Lavoura Temporária.

A construção do **IFSPublic** foi realizada em três etapas. Na primeira etapa, empregou-se a técnica Estatística chamada **Análise Fatorial** para extrair fatores e estimar os escores fatoriais a serem usados no cálculo do Índice, que tem a finalidade de hierarquizar os municípios, conforme seu grau de Financiamento.

A segunda etapa consistiu em utilizar a proporção da variância explicada por cada fator em relação à variância total explicada pelo conjunto de fatores comuns, para determinar o Peso associado a cada escore fatorial utilizado na construção do Índice.

A última etapa consistiu em hierarquizar o I**FSPublic** do maior para o menor valor obtido, visando revelar a posição competitiva de cada município nas Unidades da Federação.

Para a definição do Índice de Financiamento Sustentável Público (**IFSPublic**), estimou-se a matriz de escores fatoriais após a rotação ortogonal do tipo Varimax da estrutura fatorial inicial. O escore fatorial, por definição, situa cada observação no espaço dos fatores comuns, com isso para cada fator fi, o i-ésimo escore fatorial extraído é definido pela equação (5).

O I**FSPublic** é definido como a combinação linear dos escores fatoriais e a proporção de variância explicada por cada fator em relação a variância total comum (73.70%).

A expressão matemática utilizada para calcular o Índice é dada por:

Em que:

é a variância explicada por cada fator

é a soma total da variância explicada pelo conjunto de fatores comuns.

EFP é o escore fatorial padronizado.

O escore fatorial foi padronizado (EFP) para se obter valores positivos dos escores originais e permitir a hierarquização das cidades, uma vez que os valores do Índice estão situados entre zero e um.

Em que e são os valores mínimo e máximo observados para os escores fatoriais associados aos municípios

**Tabela 4**. Valores dos Escores Fatoriais Padronizados, índice e o ranking por municípios.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **UF** | **MUNICIPIOS** | **Escore Fatorial Padronizado Escala [0,1]** | | | | **Índice** | **Ranking** |
| **EPF1** | **EPF2** | **EPF3** | **EPF4** |
| RO | ALTA FLORESTA D OESTE | 0,1 | 0,5 | 0,33 | 0,06 | 0,236 | 28 |
| RO | ARIQUEMES | 0,04 | 0,44 | 0,6 | 0,06 | 0,245 | 23 |
| RO | CABIXI | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 874 |
| RO | CACOAL | 0,1 | 0,36 | 0,41 | 0,06 | 0,217 | 47 |
| RO | CEREJEIRAS | 0,1 | 0,18 | 0,33 | 0,11 | 0,164 | 318 |
| RO | COLORADO DO OESTE | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 488 |
| RO | CORUMBIARA | 0,1 | 0,18 | 0,33 | 0,11 | 0,165 | 308 |
| RO | COSTA MARQUES | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 665 |
| RO | ESPIGAO D OESTE | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 431 |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| RO | ITAPUA DO OESTE | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 567 |
| RO | MINISTRO ANDREAZZA | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1208 |
| RO | MIRANTE DA SERRA | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1306 |
| RO | MONTE NEGRO | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 685 |
| RO | SAO FRANCISCO DO GUAPORE | 0,1 | 0,36 | 0,33 | 0,06 | 0,202 | 89 |
| RO | SERINGUEIRAS | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 628 |
| RO | TEIXEIROPOLIS | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1153 |
| RO | THEOBROMA | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 729 |
| RO | URUPA | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1437 |
| RO | VALE DO ANARI | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1678 |
| RO | VALE DO PARAISO | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1131 |
| AC | ACRELANDIA | 0,1 | 0,35 | 0,26 | 0,06 | 0,187 | 166 |
| AC | ASSIS BRASIL | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 828 |
| AC | BRASILEIA | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 445 |
| AC | BUJARI | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1125 |
| AC | CAPIXABA | 0,1 | 0,18 | 0,28 | 0,06 | 0,151 | 1783 |
| AC | PORTO WALTER | 0,19 | 0,18 | 0,33 | 0,06 | 0,2 | 99 |
| AC | RIO BRANCO | 0,05 | 0,26 | 0,63 | 0,06 | 0,211 | 67 |
| AC | RODRIGUES ALVES | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 1588 |
| AC | SANTA ROSA DO PURUS | 0,1 | 0,18 | 0,28 | 0,06 | 0,15 | 1792 |
| AC | SENA MADUREIRA | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 381 |
| AC | SENADOR GUIOMARD | 0,1 | 0,18 | 0,33 | 0,06 | 0,16 | 749 |

**REFERÊNCIAS BIBLIOGRÁFICAS**

ANDERSON, T. W. **An Introduction Multivariate Analysis**. New York: John Wiley, 2003.

BARROSO, L.P.; ARTES, R. **Análise Multivariada**. São Paulo: IME-USP, 2003.

DILON, W.R.; GOLDSTEIN, M. **Multivariate Analysis**: methods and applications. New York: John Wiley & Son, 1984.

DRAPER, N.R.; SMITH, H. **Applied Regression Analysis**. New York: John Wiley & Sons, 1981.

ELIAN, Silva. N. **Análise de Regressão**. São Paulo: IME, 1998.

GREENE, W. H. **Econometric Analysis**. 3.ª Ed. Prentice Hall. New Jersey, 1997.

GORSUCH, R. L. **Factor Analysis**. New Jersey: Lawrence Erlbaum, 1983.

GUJARATI, Dalmodar. N. **Econometria Básica**. São Paulo: Makron Book, 2000.

HAIR JR, J.F., ANDERSON, R.E., TATHAM, R.L., BLACK, W.C. **Análise Multivariada de Dados**. 5. ed. Porto Alegre: Bookman, 2005.

HAIR JR, J.F., ANDERSON, R.E., TATHAM, R.L., BLACK, W.C. **Multivariate Data Analysis**. New York: Prentice Hall, 1998.

JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 4. ed. Nova Jersey: Prentice Hall, 1998.

KAISER, H. F. **The Varimax Criteion for Analaytic Rotation in Factor Anaysis**. Psychometrika, 1958.

KAISER, H. F.; RICE, J. **Little Jiffy**, mark IV. Education and Psychological Measurement, 1974.

NETER, J.; WASSERMAN, W. **Applied linear statistical models**. Illinois: Richard D. Irwin, 1974.

REIS, E. **Estatística Multivariada Aplicada**. 2. ed. Lisboa: Silabo, 2001.

RENCHER, A. **Methods of Multivariate Analysis**. 2 ed. New York: John Wiley & Son, 2002.

ROYSTON, J. B. Some techniques for assessing multivariate based on the Shapiro-Wilk W. **Applied Statistics**, London, v. 32, n. 2, p. 121-133, 1983.

SANTOS, A. C.; FERREIRA, D. F. **Definição do tamanho amostral usando simulação Monte Carlo para o teste de normalidade baseado em assimetria e curtose**: II. Abordagem multivariada. Ciência e Agrotecnlogia, Lavras, v. 24, n. 1, p. 62-69, 2003.

MAHALANOBIS, P. C. On **the Generalized Distance in Statistics**. Proceedings of the National Institute of Science of India, 1936.

MALHOTRA, N. K. **Marketing Research**: An Applied Orientation. New Jersey: Prentice Hall, 2001.

MINGOTI, S.A. **Análise de Dados Através de Métodos de Estatística Multivariada**: Uma Abordagem Aplicada. Belo Horizonte: UFMG, 2005