

Aplicação da análise multivariada na infraestrutura e no desempenho das escolas públicas do Ensino Fundamental e Médio pertencentes ao Núcleo Regional de Educação de Paranaíba

Angela Fontana Marques

Faculdade Estadual de Educação Ciências e Letras de Paranaíba, Av. Gabriel Esperidião, s/n, 87703-000, Paranaíba, Paraná, Brasil. E-mail: angelafontanamarques@hotmail.com

RESUMO. Este trabalho tem como técnica principal a aplicação de análise estatística multivariada, para verificar a importância das variáveis referentes ao desempenho e à infraestrutura das escolas públicas estaduais do Ensino Fundamental e Médio coordenadas pelo Núcleo Regional de Educação de Paranaíba (NREP). Para tanto, realizou-se a coleta de dados por meio de questionário aplicado a todas as escolas pertencentes ao NREP, em três segmentos: professores, alunos e escolas, com o objetivo de obter as informações necessárias. Organizaram-se os dados na forma de uma matriz para cada segmento. Primeiramente, utilizou-se a Análise de Componentes Principais para reduzir o número de variáveis e analisar quais variáveis ou quais conjuntos de variáveis explicariam a maior parte da variabilidade total, revelando o relacionamento existente entre eles. Após esta fase, aplicou-se Análise Fatorial rotacionada pelo método varimax para obter os indicadores de maior poder de explicação dos três segmentos e, a partir dessa análise, calculou-se a média entre os indicadores, o que permitiu classificar cada escola em relação a desempenho e infraestrutura.

Palavras-chave: análise multivariada, análise de componentes principais, análise fatorial.

ABSTRACT. Applying multivariate analysis in the infrastructure and performance of public elementary and high schools in the Paranaíba Regional Education Nucleus. The main technique of this work is the application of multivariate statistical analysis, in order to verify the importance of the variables regarding the performance and infrastructure of the state public elementary and high schools under the tutelage of the Paranaíba Regional Education Nucleus (NREP). With this objective, data were collected using questionnaires applied to all NREP schools, in three segments: teachers, students, and schools. Its objective was to raise data and obtain more information. The data were organized into a matrix for each segment. First, the analysis of the main components was used to reduce the number of variables and analyze which variables or sets of variables could explain most of the total variability, revealing the type of relationship that exists among them. Next, the varimax rotation factorial analysis method was used to obtain the greatest indicators in all three segments, and based on this analysis, the average was calculated among the pointers, which allowed us to classify each school in regards to its performance and infrastructure.

Key words: multivariate analysis, principal components analysis, factorial analysis.

Introdução

A Educação brasileira, atualmente, passa por grandes transformações e, por meio de propostas pedagógicas, busca atender aos anseios da sociedade. Apesar da preocupação na busca de soluções para os problemas vividos, escolas públicas ainda encontram grande dificuldade para melhorar sua qualidade e construir um modelo de ensino que consiga atender às principais necessidades educacionais de um indivíduo. A falta de uma política comprometida com resultados, a escassez de recursos públicos e a

ineficiência na utilização destes têm contribuído com o fracasso do ensino brasileiro.

Apesar de toda a preocupação com a qualidade de ensino, os índices do Sistema de Avaliação do Ensino Básico permanecem inalterados, aproximadamente 75% da população é classificada como analfabeta funcional, ou seja, mal sabe ler e/ou escrever e não consegue interpretar. Os investimentos na Educação são cada vez menores. O desinteresse da sociedade pela Educação é cada vez maior e nota-se atualmente a ausência absoluta da família na escola. A escola, hoje, transformou-se

numa 'creche de adolescente', onde os pais deixam seus filhos para que eles, pais, possam trabalhar, ou para que os filhos não fiquem na rua. O poder público transferiu toda a responsabilidade na área social, ética e de cidadania para a escola. Esta perdeu seu objetivo maior, que seria o de proporcionar as mudanças culturais, sociais e políticas por meio do ensino de qualidade, quando teve de conviver com uma 'gama' de problemas, tais como violência, droga, desestrutura familiar, miséria nas famílias, falta de limite e desinteresse, enfim, toda a problemática causada por uma política desumana e cruel, à qual está confinada a sociedade brasileira (INEP, 2006).

O objetivo geral do trabalho é avaliar a infraestrutura e o desempenho das escolas públicas de Ensino Fundamental e Médio do Núcleo Regional de Educação de Paranaíba.

Para tanto, o estudo baseia-se na comparação de variáveis relacionadas ao desempenho e à infraestrutura das instituições de ensino avaliadas, tomando, de forma simultânea, diferentes indicadores como referenciais, a saber: escolas, professores e alunos, pesquisados a fim de classificar as escolas conforme os resultados por eles fornecidos. Na classificação das escolas, serão utilizadas as técnicas da Análise Multivariada, como a Análise Fatorial e a Análise de Componentes Principais.

Componentes principais

Para investigar o relacionamento entre as variáveis que compõem um vetor aleatório de dimensão, pode ser útil transformar o conjunto de variáveis originais em um novo conjunto de variáveis não-correlacionadas, chamadas componentes principais, tendo propriedades especiais em termos de variâncias. Os principais objetivos da técnica são: a) reduzir o número de variáveis; b) analisar quais variáveis ou quais conjuntos de variáveis explicam a maior parte da variabilidade total, revelando o relacionamento existente entre eles (BOUROCHE; SAPORTA, 1982).

Componentes principais populacionais

Algebricamente, as componentes principais são combinações lineares de p variáveis originais: $X_1, X_2, \dots, X_p$. Geometricamente, as combinações lineares representam a seleção de um novo sistema de coordenadas, obtidas por rotação do sistema original com $X_1, X_2, \dots, X_p$ como eixos das coordenadas. Os novos eixos $Y_1, Y_2, \dots, Y_p$ representam as direções com variabilidade máxima e fornecem uma descrição mais simples da estrutura da matriz de covariância (JOHNSON; WICHERN, 1998).

Será visto que, segundo Johnson e Wichern (1998), as componentes principais dependem unicamente da matriz covariância Σ (ou da matriz de correlação ρ) de $X_1, X_2, \dots, X_p$. Seu desenvolvimento não requer que $\underline{x}$ tenha distribuição multivariada normal. Por outro lado, a análise de componentes principais, derivada de populações multivariadas normais, tem interpretação útil em termos de elipsoides de densidade constante.

Seja o vetor aleatório $\underline{X} = [X_1, X_2, \dots, X_p]$ com matriz covariância Σ e seus autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Considere as combinações lineares.

$$Y_1 = \underline{a}'_1 \underline{X} = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p$$

$$Y_2 = \underline{a}'_2 \underline{X} = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

$$Y_p = \underline{a}'_p \underline{X} = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p$$

Então, tem-se:

$$\underline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pp} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \underline{e}_p \underline{X}$$

$$\underline{Y} = \underline{e}_p \underline{C}_p \underline{X}$$

$$V(Y_i) = \underline{a}'_i \Sigma \underline{a}_i$$

$$\text{Cov}(Y_i, Y_k) = \underline{a}'_i \Sigma \underline{a}_k$$

$$i \neq j = 1, 2, 3, \dots, p$$

As componentes principais são das combinações lineares não-correlacionadas, $Y_1, Y_2, \dots, Y_p$, cujas variâncias são as maiores possíveis.

A primeira componente principal é a combinação linear com variância máxima. Isto é, máxima $V(Y_1) = \underline{a}'_1 \Sigma \underline{a}_1$. Fica claro que $V(Y_i) = \underline{a}'_i \Sigma \underline{a}_i$ pode ser aumentada multiplicando-se qualquer $\underline{a}_i$ por alguma constante. Para eliminar essa indeterminação, é conveniente restringir a atenção aos vetores de comprimento da unidade. Portanto, define-se:

- primeira componente = combinação linear $\underline{a}'_1 \underline{X}$ que maximiza $V(\underline{a}'_1 \underline{X})$ sujeito à restrição $\underline{a}'_1 \underline{a}_1 = 1$;

- segunda componente = combinação linear $\underline{a}'_2 \underline{X}$ que maximiza $V(\underline{a}'_2 \underline{X})$ sujeito à restrição $\underline{a}'_2 \underline{a}_2 = 1$;

$$\text{Cov}(\underline{a}'_1 \underline{X}, \underline{a}'_2 \underline{X}) = 0$$

- p-ésima componente = combinação linear $\underline{a}'_p \underline{X}$ que maximiza.

$$\text{Var}(\underline{a}'_p \underline{X}) \text{ sujeito à restrição } \underline{a}'_p \underline{a}_p = 1$$

$$\text{Cov}(\underline{a}'_i \underline{X}, \underline{a}'_k \underline{X}) = 0 \quad i \neq k$$

E, ainda, geometricamente as componentes principais se posicionam ao longo de um hiperelipsoide centrado em $\bar{x}$ (média da amostra) e cujos eixos são dados pelos autovetores da matriz correlação. Os comprimentos desses eixos são proporcionais à raiz quadrada de λ_i , $i = 1, 2, \dots, p$ e $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_i$ são os autovalores da matriz correlação.

Consequentemente, as componentes principais da amostra do vetor $\underline{x}$ podem ser interpretadas como o resultado de translação da origem do sistema original de coordenadas para $\bar{x}$ e, então, rodam os eixos coordenados até que passem pela nuvem de pontos plotados nas direções de máximas variâncias. Quando os autovalores da matriz correlação forem praticamente idênticos, a variância amostral é homogênea em todas as direções, caracterizando um esferoide; não será possível, neste caso, representar os dados em menos do que dimensões ou fatores, perdendo-se, assim, a parcimônia desejada.

Observa-se, portanto, que os autovalores e os autovetores de correlação são a essência do método das componentes principais. Os autovetores definem as direções da máxima variabilidade e os autovalores especificam as variâncias. Quando os primeiros (e poucos) autovalores são muito maiores que o resto, a maior parte da variância total pode ser explicada em menos do que p dimensões.

Componentes principais amostrais

Na prática, os parâmetros μ e Σ são desconhecidos e devem ser estimados. Suponha que $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ sejam n observações do vetor $\underline{X} \sim (\underline{\mu}, \Sigma)$. As estimativas de $\underline{\mu}$ e Σ são, respectivamente:

$$\hat{\underline{\mu}} = \bar{\underline{x}} = \frac{1}{n} \sum_{i=1}^n \underline{x}_i$$

A j -ésima componente amostral é dada por:

$$\hat{Y}_j = \hat{e}'_j \underline{X} = \hat{e}_{1j} X_1 + \hat{e}_{2j} X_2 + \dots + \hat{e}_{pj} X_p, j = 1, 2, \dots, p$$

em que:

$(\hat{\lambda}_1, \hat{e}_{11}), (\hat{\lambda}_2, \hat{e}_{21}), \dots, (\hat{\lambda}_p, \hat{e}_{p1})$ são os autovalores e autovetores da matriz de covariância amostral Σ .

Propriedades das componentes principais amostrais

$$(1) \text{ var } [\hat{Y}_j] = \hat{\lambda}_j, j = 1, 2, 3, \dots, p$$

$$(2) \text{ cov } [\hat{Y}_i, \hat{Y}_j] = 0, i \neq j$$

$$(3) \sum_{i=1}^p s_i^2 = s_1^2 + s_2^2 + \dots + s_p^2 = \sum_{j=1}^p \hat{Y}_j^2 = \hat{Y}_1^2 + \hat{Y}_2^2 + \dots + \hat{Y}_p^2$$

(4) A proporção da variância total explicada pela j -ésima componente principal estimada é:

$$\frac{\hat{\lambda}_j}{\hat{\lambda}_1 + \hat{\lambda}_2 + \dots + \hat{\lambda}_p},$$

(5) A correlação amostral entre a componente principal $\hat{Y}_j$ e a variável original X_i :

$$r_{\hat{Y}_j X_i} = \frac{\hat{e}_{ij} \sqrt{\hat{\lambda}_j}}{s_i},$$

$$i \neq j = 1, 2, 3, \dots, p$$

Componente principal amostral de variáveis padronizadas

Para um vetor de observações padronizadas $\hat{\underline{z}} = [\hat{z}_1, \hat{z}_2, \dots, \hat{z}_p]$, a matriz covariância será:

$$R = \begin{bmatrix} 1 & \hat{\delta}_{12} & \dots & \hat{\delta}_{1p} \\ \hat{\delta}_{21} & 1 & \dots & \hat{\delta}_{2p} \\ \vdots & \vdots & \dots & \vdots \\ \hat{\delta}_{p1} & \hat{\delta}_{p2} & \dots & 1 \end{bmatrix}$$

A j -ésima componente principal das variáveis padronizadas é:

$$\hat{Y}_j = \hat{e}'_j \hat{\underline{z}} = \hat{e}_{1j} \hat{z}_1 + \hat{e}_{2j} \hat{z}_2 + \dots + \hat{e}_{pj} \hat{z}_p$$

em que:

$$(\hat{\lambda}_j, \hat{e}_j) \text{ é o } j\text{-ésimo par autovalor-autovetor de } R,$$

$$\text{com } \hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p \geq 0.$$

A proporção da variância total amostral no caso padronizado, explicada pela j -ésima componente, é dada por:

$$\frac{\hat{\lambda}_j}{p}, \quad i = 1, 2, \dots, p$$

Análise fatorial

A análise fatorial é uma técnica multivariada de interdependência em que todas as variáveis são simultaneamente consideradas, cada uma relacionada com as demais, a fim de estudar a inter-relação existente entre elas, buscando a redução de dados ou a sumarização.

Teste de esfericidade de Bartlett

O teste de esfericidade de Bartlett testa a hipótese de que as variáveis não sejam correlacionadas na população. A hipótese básica diz que a matriz de correlação da população é uma matriz identidade, o que indica que o modelo fatorial é inapropriado.

A estatística do teste é dada por:

$$\chi^2 = - \left[(n-1) - \frac{2p+5}{6} \right] \ln |R|$$

que tem uma distribuição qui-quadrado com

$$v = \frac{p(p-1)}{2} \text{ grau de liberdade, em que:}$$

n = tamanho da amostra

p = número de variáveis

$|R|$ = determinante da matriz de correlação

Medida de Adequação da Amostra de Kaiser-Meyer-Olkin (KMO)

Esta medida é representada por um índice (MAS) que avalia a adequação da análise fatorial, calculado por:

$$MSA = \frac{\sum_{j \neq k} r_{jk}^2}{\sum_{j \neq k} r_{jk}^2 + \sum_{j \neq k} q_{jk}^2}$$

em que: r_{jk}^2 é o quadrado dos elementos da matriz de correlação original (fora da diagonal); q_{jk}^2 é o quadrado dos elementos fora da diagonal da matriz anti-imagem da correlação.

Valores altos (entre 0,5 e 1,0) indicam que a análise fatorial é apropriada, enquanto valores baixos (abaixo de 0,5) indicam que a análise fatorial pode ser inadequada.

Crítérios para determinação do número de fatores

A escolha do número de fatores é uma das tarefas mais importantes de uma AF. Hair Junior et al. (2005) discutem que, se o pesquisador opta por um número muito reduzido, não é possível identificar estruturas importantes existentes nos dados; por outro lado, se o número é excessivo, ele pode vir a ter problemas de interpretabilidade dos fatores. Existem, na literatura, vários critérios que auxiliam na determinação do número de fatores que, invariavelmente, quando empregado em um mesmo conjunto de dados, conduz a resultados diferentes. Como regra geral, o pesquisador deve procurar um

compromisso entre o número de fatores (que, a princípio, deve ser o menor possível) e a sua interpretabilidade. É comum, em situações práticas, simplesmente comparar soluções com um número diferente de fatores e fazer a escolha com base no bom senso do pesquisador. Nesses casos, os critérios apresentados na sequência podem ser utilizados como ponto de partida para a obtenção de uma solução final.

Rotação dos fatores

Os objetivos da extração dos fatores diferem daqueles associados à rotação. A técnica da extração é projetada para maximizar fontes independentes de variância na matriz de correlação. Já o objetivo da rotação consiste em delinear os fatores, ou seja, os fatores são rodados até que correspondam a dimensões separadas, construto investigado. O desenvolvimento dos computadores digitais propiciou o aparecimento de uma série de programas que realizam rotação de eixos, tanto ortogonais quanto oblíquas. Neste trabalho foi empregada a rotação Varimax.

Material e métodos

A metodologia proposta usa como referência básica dados internos do Núcleo Regional de Educação de Paranavaí, questionários (em apêndice) aplicados nas instituições, em três segmentos: o primeiro é direcionado às escolas; o segundo, aos professores e o terceiro aos alunos e também dados do censo disponibilizado pelo Núcleo Regional de Educação de Paranavaí em forma de dados digitais que são coletados através de formulários entregues ao Instituições estaduais para a análise da situação geral das mesmas ao final de cada ano. O objetivo desta coleta de dados foi levantar características sobre a educação tais como: satisfação, comprometimento, opinião, estrutura, desempenho e atitudes da população escolar estudada.

Com essas informações, foi possível verificar o número de alunos e professores de cada instituição, podendo ser avaliada a quantidade de questionários que deveria ser usada na coleta de dados, para que a pesquisa pudesse selecionar as características das populações envolvidas. Os dados fornecidos pelo Núcleo Regional de Educação de Paranavaí relatam a existência de 23.942 alunos e 1992 professores. Tais dados não estavam agrupados, por isso, criou-se uma tabela que caracterizava o número de professores e alunos de cada instituição; calculou-se o número equivalente a 40% dos alunos e 20% dos professores das instituições para que,

posteriormente, fossem aplicados os questionários.

O método de análise da pesquisa está estruturado da seguinte forma: escolas, professores e alunos, seguindo a sequência dos passos:

- determinação do perfil da amostra;
- aplicação do teste de esfericidade de Bartlett e cálculo da Medida de Adequacidade da Amostra (MSA) de Kaiser-Meyer-Olkin;
- aplicação da análise fatorial com rotação Varimax; o critério de escolha de fatores utilizado foi o de explicar pelo menos 80% da variabilidade total;
- cálculo dos escores fatoriais para os fatores retidos;
- cálculo dos escores finais únicos, por escola, ponderados pela explicação de cada fator;
- transformação dos escores finais únicos dos professores, por escola, na escala de 0 a 1;
- classificação e identificação das escolas, em ordem crescente, do indicador determinado.

Resultados e discussão

Caracterização da amostra das escolas:

Teste de esfericidade de Bartlett e Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO):

(1) Teste de esfericidade de Bartlett: $\chi^2 = 1938,4$ com $p = 0$, mostrando que existe correlação significativa entre as variáveis.

(2) Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO): $MSA = 0,0662$, como o $MSA < 0,5$, a análise fatorial é inadequada para a amostra utilizada.

Como o tamanho da amostra é pequeno em relação ao número de variáveis, uma alternativa para viabilizar o uso da Análise Fatorial foi transformar grupos de variáveis de mesma natureza em uma única variável, adotando valores médios.

Para a nova matriz de dados, o resultado foi:

(1) Teste de esfericidade de Bartlett: $\chi^2 = 777,2497$ com $p = 0$, mostrando que existe correlação significativa entre as variáveis.

(2) Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO): $MSA = 0,5301$, como o $MSA > 0,5$, a análise fatorial é adequada para a amostra utilizada.

Na sequência, são apresentados os resultados obtidos nos carregamentos de cada componente. Realizou-se a análise fatorial por componentes principais, a qual revelou oito fatores com autovalores maiores do que 1, explicando, respectivamente, 16,39; 11,09; 9,15; 10,35; 8,98; 7,66; 6,67 e 9,85% da variação total dos escores (os oito fatores juntos explicam 80,14%).

Na Figura 1, verifica-se que 17 escolas atingiram o indicador 0,2 ou abaixo deste. Isto significa que 24,24% das escolas pesquisadas desempenham somente 20% ou menos dos requisitos que caracterizam o desempenho e a infraestrutura; 25 escolas aparecem com os indicadores acima de 0,2 até 0,4, representando 39,39% do resultado esperado; 14 escolas com indicadores maiores que 0,4 e menores ou iguais a 0,6, indicando 22,72% de aproveitamento; sete escolas com indicadores que variam com valores maiores que 0,6, mas inferiores ou iguais a 0,8, relatando 10,60% dentre o total de 100%; e duas escolas com os indicadores iguais a 1, representando 3,03% da amostra total que obteve êxito completo em relação às características avaliadas sobre o desempenho e a infraestrutura das 66 escolas. É notório que, no intervalo de 0,8 a 0,9, não se encontra nenhuma escola; as demais escolas apresentam indicadores iguais ou menores que 0,7. Assim, 64 escolas entre o total pesquisado não conseguiram obter mais que os 71% dos itens necessários para classificá-las com sucesso em relação ao desempenho e à infraestrutura.

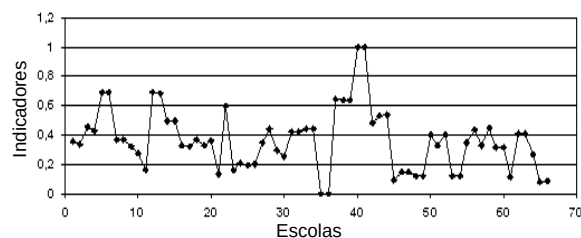


Figura 1. Indicador de aproveitamento por escola.

Para a caracterização da amostra dos professores, tem-se:

Teste de Esfericidade de Bartlett e Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO):

(1) Teste de esfericidade de Bartlett: $\chi^2 = 3415,7$ com $p = 0$, mostrando que existe correlação significativa entre as variáveis.

(2) Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO): $MSA = 0,7219$, como o $MSA > 0,5$, a análise fatorial é adequada para a amostra utilizada.

A aplicação da análise fatorial por componentes principais revelou 17 fatores, explicando, respectivamente, 5,98; 6,97; 5,48; 4,41; 3,95; 5,08; 4,04; 3,88; 4,26; 10,37; 0,71; 3,96; 3,77; 4,51; 4,52; 3,94 e 4,22% da variação total dos escores (os 17 fatores juntos explicam 80,05%).

De acordo com a Figura 2, 17 escolas tiveram como indicador valores iguais ou menores que 0,2. Isto revela que, dos itens contidos no questionário

aplicado aos professores, 25,75% das escolas resumem seu desempenho e infraestrutura em 20% ou menos; outras 29 escolas, estimando 43,39%, têm um indicador de aproveitamento acima de 0,2, com oscilação até o valor 0,4, explicando que de 20 a 40% das características necessárias foram evidenciadas pelos professores. Com indicadores maiores que 0,4, mas menores ou iguais a 0,6, verificam-se 15 escolas, ou seja, 22,72% das instituições pesquisadas; portanto, nestas escolas, na opinião dos professores, somente são satisfeitos entre 40 e 60% dos itens estabelecidos no questionário. Entre os indicadores acima de 0,6 até 0,8, encontram-se três escolas, representando 4,54% do total; o aproveitamento delas, conforme indicado pelos professores, é de 60 a 80%. Apenas duas escolas aparecem acima dos 90% de aproveitamento, segundo os professores.

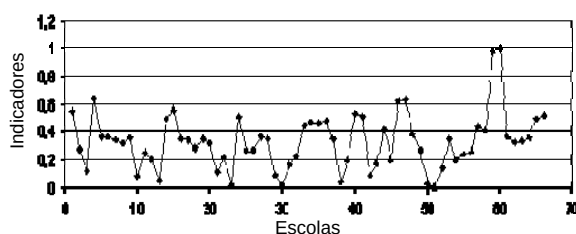


Figura 2. Indicador de aproveitamento dos professores por escola.

Já na caracterização da amostra por escolas, tem-se:

Teste de esfericidade de Bartlett e Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO):

(1) Teste de esfericidade de Bartlett: $\chi^2 = 88.269,0$ com $p = 0$, mostrando que existe uma correlação significativa entre as variáveis.

(2) Medida de Adequacidade da Amostra de Kaiser-Meyer-Olkin (KMO): $MSA = 0,9360$, como o $MSA > 0,5$, a análise fatorial é adequada para a amostra utilizada.

A análise fatorial por componentes principais revelou 15 fatores, explicando, respectivamente, 11,03; 3,95; 6,68; 3,90; 3,79; 3,91; 3,76; 3,87; 6,30; 5,89; 4,44; 6,39; 6,04; 4,09 e 6,28% da variação total dos escores (os 15 fatores juntos explicam 80,29%).

Na Figura 3, existem sete escolas que aparecem como indicador valores iguais ou menores que 0,2, isso evidencia que dos itens contidos no questionário aplicado aos alunos 10,60% das escolas resumem seu desempenho e infra-estrutura de 20% ou menos, 12 escolas, estimando 18,18% tem um indicador de aproveitamento acima de 0,2 ao valor 0,4, situando que das características necessárias foram delineados pelos professores de 20 a 40%. Se verificar os indicadores maiores que 0,4, mas menores ou iguais que 0,6 estará neste grupo 36

escolas, atribuindo 54,54% das escolas pesquisadas, portanto, observa-se que este grupo está o maior número de escolas pesquisadas e as opiniões dos alunos são satisfeitas entre os itens estabelecidos no questionário de 40 a 60%. Encontram-se entre os indicadores acima de 0,6 até 0,8, um conjunto com nove escolas, representando 13,63% do total, o aproveitamento destas escolas indicado pelos alunos é de 60 a 80%, e no grupo avaliado entre os valores maiores que 0,8 e menores ou iguais a 1, estão indicadas duas escolas que aparecem com um parecer acima dos 90% de aproveitamento, citados pelos alunos.

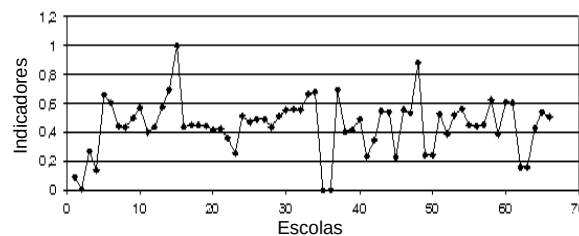


Figura 3. Indicador de aproveitamento dos alunos por escola.

As escolas são qualificadas em relação aos três segmentos pesquisados, determinando seus respectivos indicadores individuais e após a classificação pelos indicadores médios, que são a soma dos indicadores das escolas, professores e alunos e o resultado dividido por três.

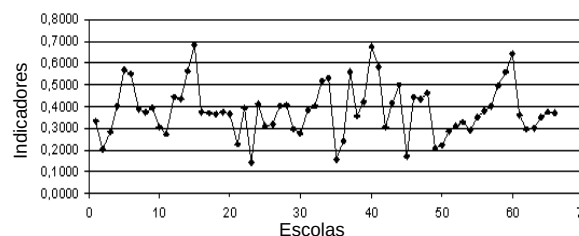


Figura 4. Classificação das escolas pelas médias obtidas nos três segmentos.

Portanto, por meio dos indicadores dos utilizados, dados que constam nas Figuras 1, 2 e 3, pode-se construir a Figura 4, que apresenta o resultado da média dos indicadores das escolas, professores e alunos, transportando por meio destes indicadores médios o aproveitamento das escolas. Percebe-se que relacionando o desempenho e a infraestrutura, três escolas estão agrupadas com valores iguais ou menores que 0,2, admitindo que 4,45% das escolas desenvolvem um rendimento igual ao inferior a 20%, outro grupo com 13 escolas aparecem com indicadores superiores a 0,2 até 0,3, demonstrando que 19,96% trabalham com margem de superior a 20% ou igual a 30% dos itens necessários, 24 escolas definem 36,36%, este é o

maior grupo, descrevendo uma explicação maior que 0,3 até 0,4, representando um percentual de aproveitamento acima de 30 até 40%, 15 escolas estão localizadas no grupo com indicadores maiores que 0,4 até 0,5, destacam-se neste grupo 22,72% das escolas pesquisadas, oito escolas refletem valores maiores que 0,5 até 0,6, descrevendo 12,12% do montante e três escolas estão relacionadas aos indicadores maiores que 0,6 e inferiores a 0,7, estas escolas representam 4,54% da amostra total identificada. Observa-se que das 66 escolas, nenhuma teve aproveitamento melhor que 70%.

Conclusão

Após as técnicas aplicadas e a análise do resultado dos três segmentos, no cálculo do indicador médio de cada escola verifica-se que somente 11 tiveram aproveitamento final em relação ao desempenho e à infraestrutura, com percentual acima de 50%; destas que atingiram os indicadores acima de 0,5, a que teve a primeira colocação, ou seja, o maior indicador, obteve 68,54%. Nenhuma escola conseguiu aproveitamento superior a 70% do resultado esperado, restando, assim,

55 escolas com indicativo menor que 50%. Os dados analisados levavam em conta a opinião das escolas, dos professores e dos alunos.

Referências

- BOUROCHE, J. M.; SAPORTA, G. **Análise de dados**. Rio de Janeiro: Zahar, 1982.
- INEP-Instituto Nacional de Estudos e Pesquisa Educacionais Anísio Teixeira. **Dados referentes ao ano de 2005**. Disponível em: <http://www.inep.gov.br>. Acesso em: 8 dez. 2006.
- HAIR JUNIOR, J. F. **Análise multivariada de dados**. Porto Alegre: Bookman, 2005.
- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4. ed. New Jersey: Prentice Hall, 1998.

Received on March 14, 2008.

Accepted on September 18, 2008.

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.