ISSN: 2317-0840

Análise Fatorial Exploratória: aplicação em dados de crimes contra as mulheres em 141 municípios do estado de Mato Grosso no ano de 2016

Jaqueline T. Silva[1†], Elkeaeer S. P. Ruvieri[2], Mateus O. Magalhães[3], Névio Lotufo Neto[4], Kuang Houngyu[5]

[1]*Graduanda em Estatística, Universidade Federal de Mato Grosso - UFMT.*
[2]*Graduando em Estatística, Universidade Federal de Mato Grosso - UFMT.*
E-mail: elkeaeer@hotmail.com.
[3]*Graduando em Estatística, Universidade Federal de Mato Grosso - UFMT.*
E-mail: mateusormondes@gmail.com.
[4]*Mestrando em Física Ambiental, Universidade Federal de Mato Grosso - UFMT.*
E-mail: neviolotufo@gmail.com.
[5]*Professor Adjunto do Departamento de Estatística, Universidade Federal de Mato Grosso - UFMT. E-mail:* prof.kuang@gmail.com.

Resumo: *A análise multivariada é um conjunto de técnicas estatísticas que permite a análise e interpretação de conjuntos de dados de natureza quantitativa com grande número de variáveis de forma simplificada. A análise fatorial exploratória (AFE) é uma técnica dentro da análise fatorial cujo objetivo abrangente é identificar as relações subjacentes entre as variáveis medidas. A AFE é uma técnica estatística que estuda correlações entre um grande número de variáveis agrupando-as em fatores. Assim o objetivo principal deste trabalho é apresentar os principais aspectos desta técnica, fazendo-se posteriormente uma aplicação prática sobre os dados de crimes contra mulheres do Estado de Mato Grosso no ano de 2016. Com este trabalho, pretende-se também auxiliar pesquisadores das diversas áreas a utilizarem AFE com maior discernimento teórico e metodológico. Foi possível identificar dois fatores pelo método de Componente Principal, que explicou aproximadamente 80,56% da variância total encontrada, pode-se então afirmar que este método foi eficaz e cumpriu os objetivos propostos que são os de resumo e redução dos dados além de permitir a interpretação dos mesmos.*

Palavras-chave: Análise multivariada; análise componente principal; rotação *Varimax.*

Abstract: *Multivariate analysis is a set of statistical techniques that allows the analysis and interpretation of quantitative data sets with a large number of variables in a simplified way. Exploratory factorial analysis (EFA) is a technique within the factorial analysis whose broad objective is to identify the underlying relationships between the measured variables. AFE is a statistical technique that studies correlations between a large number of variables grouping them into factors. Thus, the main objective of this work is to present the main aspects of this technique, and a practical application on the data of crimes against women in the State of Mato Grosso in the year 2016. This work is also intended to help researchers of diverse areas using AFE with greater theoretical and methodological discernment. It was possible to identify two factors by the Principal Component method, which explained approximately 80.56% of the total variance found, it can be affirmed that this method was effective and fulfilled the proposed objectives that are the ones of summary and reduction of the data besides interpretation.*

Keywords: Multivariate analysis; principal component analysis; Varimax rotation.

[†]Autora correspondente: jaquetrentino@hotmail.com.

Introdução

Crime é um dos problemas que continuamente perturbam a existência da humanidade, sendo ainda mais perturbador no complexo mundo moderno. Tendo começado nos primórdios da civilização como algo simples e desorganizado, atualmente é um problema complexo e organizado. Apesar de o crime ser inevitável em uma sociedade várias medidas preventivas podem e são aplicadas de forma a reduzir a ameaça. A introdução de métodos científicos e técnicos modernos na prevenção de crimes tem se mostrado eficiente. Os métodos multivariados da estatística têm se provado eficientes tendo feito contribuições em muitas explicações criminológicas (GULUMBE et al., 2013).

A análise multivariada é um conjunto de técnicas estatísticas que permite a análise e interpretação de conjuntos de dados de natureza quantitativa com grande número de variáveis de forma simplificada. Por conta da facilidade de análise e interpretação de grande número de variáveis, as técnicas multivariadas têm se tornado muito populares em várias áreas do conhecimento apesar de terem sido desenvolvidas para problemas específicos, dentre as áreas de grande aplicação atualmente estão: agronomia, zootecnia, ecologia, florestal, etc. Dentre as várias possíveis aplicações das técnicas multivariadas estão as taxas de ocorrência de crimes que são classificados em várias categorias como: Assassinato, Roubo, Estupro, Assalto, Arrombamento, Pequenos Furtos, Furtos Veiculares, etc. (NEISSE; HONGYU, 2016).

A análise fatorial exploratória (AFE) ou *"exploratory factor analysis"* é uma técnica dentro da análise fatorial cujo objetivo abrangente é identificar as relações subjacentes entre as variáveis medidas. A AFE é uma técnica estatística que estuda correlações entre um grande número de variáveis agrupando-as em fatores. Essa técnica permite a redução de dados, identificando as variáveis mais representativas ou criando um novo conjunto de variáveis, bem menor que o original (HAIR et al., 2009; HONGYU, 2018).

Uma situação comum em diversas áreas do conhecimento é aquela na qual, para cada elemento de uma amostra, observa-se um grande número de variáveis. Essas variáveis podem ser, por exemplo, um conjunto de itens de uma escala ou os escores obtidos por um indivíduo em diferentes escalas de avaliação (HONGYU, 2018). Diante de um quadro como esse, o pesquisador enfrenta dois problemas, que podem ser resolvidos por meio da AFE: a) a caracterização dos avaliados, levando-se em conta um conjunto eventualmente grande de variáveis, e b) a descrição da inter-relação dessas variáveis, eventualmente explicitando uma estrutura de interdependência subjacente aos dados (HAIR et al., 2009).

O objetivo principal deste trabalho é apresentar os principais aspectos desta técnica, fazendo-se posteriormente uma aplicação prática sobre os dados de crimes contra mulheres do Estado de Mato Grosso, na qual foca-se nas interpretações dos resultados obtidos pelo método de de análise de componentes principais (ACP). Com este trabalho, pretende-se também auxiliar pesquisadores das diversas áreas a utilizarem AFE com maior discernimento teórico e metodológico.

Materiais e Métodos

O conjunto de dados utilizado neste (artigo) foram obtidos por meio do site da Secretaria de Estado de Planejamento (SEPLAN) de Mato Grosso, proveniente de ocorrências registradas pela Polícia Judiciária Civil do Estado de Mato Grosso referente ao ano de 2016.

O conjunto contém dezoito variáveis que representam tipos de ocorrências registradas de vítimas femininas nos 141 municípios do Estado, são elas: Abandono (X1), Assédio Sexual (X2), Ato Obeseno (X3), Danos Materiais (X4), Embriaguez (X5), Estupro (X6), Estupro (tentado) (X7), Exercício Arbitrário das Próprias Razões (X8), Fato Atípico (Ocorrências Atípicas (X9), Furto (X10), Homicídio (tentado) (X11), Homicídio Doloso (X12), Importunação Ofensiva ao Pudor (X13), Subtração de Incapaz (X14), Suicídio (X15), Tráfico de Mulheres (X16), Vias de Fato (X17) e Violação de Domicílio (X18).

Todas as análises deste artigo foram feitas por meio de rotinas computacionais implementadas no *software* R 3.5.1 (R Development Core Team, 2018) com os pacotes `vegan`, `factoextra` e `psych`.

Pressupostos das Análises Fatoriais Exploratórias

O primeiro passo durante a implementação de AFE (pressupostos) é verificar se a aplicação da análise fatorial tem validade para as variáveis escolhidas, sendo justificado pela pouca quantidade de respondentes da pesquisa. Para isso, dois métodos de avaliação são mais comumente utilizados, o critério de Kaiser-Meyer-Olkin (KMO) e o Teste de Esfericidade de Bartlett (DZIUBAN; SHIRKEY, 1974; DAMÁSIO, 2012; KIRCH, et al., 2017).

O índice de KMO, também conhecido como índice de adequação da amostra, é um teste estatístico que sugere a proporção de variância dos itens que pode estar sendo explicada por uma variável latente, tal índice indica o quão adequada é a aplicação da AFE para o conjunto de dados (HAIR et al., 2009; HONGYU, 2018). O KMO é calculado por meio do quadrado das correlações totais dividido pelo quadrado das correlações parciais, das variáveis analisadas, cuja expressão é dada na Equação 1 (DZIUBAN; SHIRKEY, 1974):

$$KMO = \frac{\sum_{j=1}^{p}\sum_{m=1,m\neq j}^{p} r_{jm}^2}{\sum_{j=1}^{p}\sum_{m=1,m\neq j}^{p} r_{jm}^2 + \sum_{j=1}^{p}\sum_{m=1,m\neq j}^{p} r_{pjm}^2} \tag{1}$$

em que r_{jm}^2 é o coeficiente de correlação linear entre X_j e X_m; r_{pjm}^2 é o coeficiente de correlação parcial amostral entre X_j e X_m, definido como sendo o coeficiente de correlação linear entre os resíduos.

O KMO pode variar de zero a um. Valores iguais ou próximos a zero indicam que a soma das correlações parciais dos itens avaliados é bastante alta em relação à soma das correlações totais. Nesses casos, possivelmente a análise fatorial será inapropriada (PASQUALI, 1999). Como regra para interpretação dos índices de KMO, valores menores que 0,5 são considerados inaceitáveis, valores entre 0,5 e 0,7 são considerados medíocres; valores entre 0,7 e 0,8 são considerados bons; valores maiores que 0,8 e 0,9 são considerados ótimos e excelentes, respectivamente (HUTCHESON; SOFRONIOU, 1999; PEREIRA, 1999).

Modelo teórico da análise fatorial

O modelo da análise de fatores pode ser escrito conforme Equação 2 (JOHNSON; WICHERN, 2008):

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}Fm + e_i \tag{2}$$

em que X_i é o i-ésimo escore depois dele ter sido padronizado (média 0 e desvio-padrão 1); $i = 1, \ldots, p$; p é o número de variáveis; $a_{i1}, a_{i2}, \ldots, a_{im}$ são as cargas dos fatores para o i-ésimo teste;$F_1, F_2, \ldots, F_m$ são m fatores comuns não correlacionados, cada um com média 0 e variância 1 e e_i é um fator específico somente para o i-ésimo teste que é não correlacionado com qualquer dos fatores comuns e tem média zero.

Os p valores observados X_p são expressos em termos de $p + m$ variáveis aleatórias não observáveis $(F_1, F_2, \ldots, F_m; \epsilon_1, \epsilon_2, \ldots, \epsilon_p)$. Isso distingue o modelo fatorial do modelo de regressão múltipla, no qual as variáveis independentes podem ser observadas, e cujas posições são ocupadas por F no modelo fatorial. Matricialmente, o modelo é expresso pela Equação 3 (JOHNSON; WICHERN, 2008):

$$\boldsymbol{X}_{(p\times 1)} = \boldsymbol{\Lambda}_{(p\times m)}\boldsymbol{F}_{(m\times 1)} + \boldsymbol{\epsilon}_{(p\times 1)} \tag{3}$$

Já o modelo AF é dado por:

$$\begin{aligned} Var\left(X_j\right) &= a_{j1}^2 Var\left(F_1\right) + a_{j2}^2 Var\left(F_2\right) + \ldots + a_{jm}^2 Var\left(F_m\right) + Var\left(\epsilon_j\right) \\ &= a_{j1}^2 + \ldots + a_{jm}^2 + Var\left(\epsilon_j\right) \end{aligned}$$

em que $a_{j1}^2 + a_{j2}^2 + \ldots + a_{jm}^2$ é chamada a comunalidade de X_j (a parte da sua variância que é explicada pelos fatores comuns). A comunalidade não pode exceder 1, sendo necessário que fique compreendida entre $-1 \leq a_{ij} \leq +1$. Pode também ser estabelecido que a correlação entre X_j e $X_{j'}$, seja dado por:

$$r_{jj'} = a_{j1}a_{j'1} + a_{j2}a_{j'2} + \ldots + a_{jm}a_{j'm}$$

Consequentemente, duas variáveis somente serão altamente correlacionadas se elas tiverem altas cargas no mesmo fator (JOHNSON; WICHERN, 2008; HONGYU, 2018).

Análise Fatorial por Componentes Principais

Partindo-se da matriz $\boldsymbol{X}_{(n\times p)}$, realiza-se uma análise de componentes principais, obtendo-se p componentes principais é dado por (JOHNSON; WICHERN, 2008):

$$Z_i = b_{i1}X_1 + b_{i2}X_2 + \ldots + b_{ip}X_p$$

em que $i = 1, \ldots, p$ e b_ij são dados pelos autovetores da matriz de correlações. Por se uma transformação ortogonal, o relacionamento inverso é dado:

$$X_i = b_{i1}Z_1 + b_{i2}Z_2 + \ldots + b_{ip}Z_p$$

O modelo de $m < p$ fatores fica então representado por:

$$X_i = b_{i1}Z_1 + b_{i2}Z_2 + \ldots + b_{im}Z_m + e_i$$

em que e_i é uma combinação linear dos componentes de Z_{m+1} a Z_p.

Divide-se Z_i por seu desvio padrão $\sqrt{\lambda_i}$ (obtendo-se F_i), a raiz quadrada do correspondente autovalor na matriz de correlações. As equações são calculadas por (JOHNSON; WICHERN, 2008):

$$X_i = \sqrt{\lambda_1}b_{i1}F_1 + \sqrt{\lambda_2}b_{i2}F_2 + \ldots + \sqrt{\lambda_m}b_{im}F_m + e_i$$

Supondo que $a_ij = \sqrt{\lambda_j} b_j i$, é possível demonstrar que X_i é dado:

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}F_m + e_i$$

Procedendo a rotação *varimax*, chega-se que Xi é dado:

$$X_i = g_{i1}F_1^{'} + g_{i2}F_2^{'} + \ldots + g_{im}F_m^{'} + e_i$$

em que $F_i^{'}$ representa o novo i-ésimo fator.Os escores fatoriais são calculados por mínimos quadrados ponderados para estimativas de máxima verossimilhança:

$$\boldsymbol{F^*} = (\boldsymbol{G'G})^{-1}\boldsymbol{G'X'}$$

em que $\boldsymbol{F^*}$ é uma matriz $n \times m$, contendo os valores para os m fatores rotacionados em sua coluna; $\boldsymbol{X}$ é a matriz $n \times p$ de dados originais e padronizados; $\boldsymbol{G}$ é a matriz $p \times m$ das cargas fatoriais rotacionados (JOHNSON; WICHERN, 2008).

Resultados e Discussão

Foi calculada a matriz de correlação entre as variáveis de criminalidades contra as mulheres nos 141 municípios do estado de Mato Grosso apresentaram correlações altas, principalmente entre as variáveis de abandono e estupro (X1 e X6) com correlação de 0,98; abandono e estupro tentado (X1 e X7): 0,96; abandono e furto (X1 e X10): 0,97; assédio sexual e danos materiais (X2 e X4): 0,95; estupro e estupro tentado (X6 e X7): 0,97; estupro e furto (X6 e X10): 0,98; estupro e homicídio tentado (X6 e X11): 0,97; estupro tentado e furto (X12 e X10): 0,98; estupro tentado e homicídio tentado (X7 e X11): 0,95 e furto e homicídio tentado (X10 e X11): 0,97. As menores correlações foram entre a variável fator atípico (X9) com todas as demais variáveis; ato obseno (X3) e homicídio doloso (X12) correlação de 0,59; a variável subtração de incapaz (X15) com todas as demais variáveis; X3 e X16 com correlação de 0,48; X5 e X15 com correlação de 0,43, portanto, os registros dos crimes dessas variáveis não foram muito correlacionados.

Hair et al.(2009) sugerem que a amostra deve ser superior a 50 observações, sendo aconselhável no mínimo 100 casos para assegurar resultados mais robustos. A razão entre o número de casos e a quantidade de variáveis deve exceder cinco para um ou mais. A medida global de adequação amostral do critério de KMO (KaiserMeyer-Olkin) varia entre 0 e 1, quanto mais perto de 1, tanto melhor. Pela Tabela 1, o KMO foi igual 0,92, logo, isso indica que os fatores encontrados na AF conseguem descrever, satisfatoriamente, as variações dos dados originais. Hair et al.(2009) sugerem 0,50 como patamar aceitável.

Outro teste que pode ser avaliado é o Teste de Esfericidade de Bartlett (DZIUBAN; SHIRKEY, 1974) que indica se existe relação suficiente entre as variáveis para aplicação da AF. Para que seja possível a aplicação da AF recomenda-se que o valor-p seja menor que 5%. Na Tabela 1 apresentou valor-$p < 0,001$, logo, pelo teste de esfericidade constata-se que os dados são adequados para a aplicação da AF nas variáveis analisadas, indicando que a matriz é fatorável, rejeitando a hipótese nula de que a matriz de dados é similar a uma matrizidentidade.

O *Screeplot* apresentado na Figura 1 demonstra os autovalores dos componentes principais e da fatorial, sendo estes gerados pelo método das análises paralelas (AP) de Horn(HORN, 1965). De acordo com Laros (2004), a AP é um procedimento estatístico de

Tabela 1: A estatística KMO e teste de Bartlett

Critério de Kaiser-Mayer-Olkin	0,92
Teste de Esfericidade de Bartlett	
Qui-quadrado aprox.	3367,756
gl	153
valor - p	$< 0{,}001$

simulação Monte-Carlo que consiste na construção aleatória de um conjunto hipotético de matrizes de correlação de variáveis, utilizando como base a mesma dimensionalidade (o mesmo número p de variáveis e o mesmo número n de sujeitos) do conjunto de dados reais.

Pela análise da Figura 1, nota-se que o critério AP indicou que dois fatores foram ideais para ser extraídos para continuação da análise fatorial.

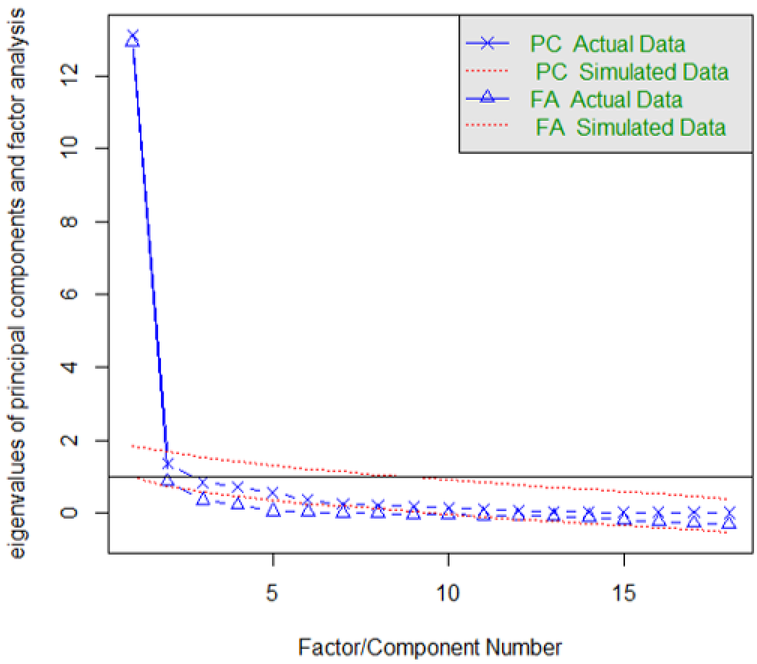

Figura 1: *Screeplot* das análises de componentes principais e fatorial

Pela Tabela 2, apresenta-se os resultados de autovalores e variâncias explicadas e acumuladas pelo método de estimação de Componente Principal. A inspeção nos autovalores (λ_i) demonstra que o primeiro fator consegue captar uma quantidade significativa da variação dos dados: (λ_1) = 10,47. Como o total de variáveis do conjunto original é 18, isso significa que somente o primeiro autovalor explica 10,47/18 = 58,17% da variação. Como o segundo maior autovalor é (λ_2) = 4,03, este explica 22,39% da variância dos dados originais.

Na Tabela 3, apresenta-se as cargas fatoriais e comunalidades do Fator 1 e 2 do método de Componente Principal após a rotação *Varimax*. De acordo com Pallant (2007), o tipo de rotação *Varimax* é o mais comumente utilizado, pois esse método procura minimizar o número de variáveis que apresentam altas cargas em cada fator. As comunalidades representam a proporção da variância para cada variável incluída na análise que é expli-

Tabela 2: Comunalidades e fatores iniciais, extraídos pelo Método do Componente Principal com os seus respectivos autovalores e percentuais de variância explicada

Fator	Autovalor	Variância Explicada(%)	Variância Acumulada(%)
1	10,47	58,17	58,17
2	4,03	22,39	80,53

cada pelos componentes extraídos (SCHAWB, 2007). Por exemplo, X20 apresenta maior comunalidades pelo Componente Principal, logo, os dois fatores extraídos explicam 99% da variância da variável X20.

Tabela 3: Cargas iniciais dos dois primeiros fatores extraídos pelos dois métodos de estimação e pela rotação *Varimax*.

Variável	PC1	PC2	Comunalidades (h^2)
$x1$	**0,92**	0,29	0,93
$x2$	**0,86**	0,40	0,90
$x3$	**0,62**	0,55	0,68
$x4$	**0,88**	0,36	0,91
$x5$	0,46	**0,64**	0,62
$x6$	**0,91**	0,36	0,96
$x7$	**0,94**	0,27	0,96
$x8$	**0,85**	0,44	0,91
$x9$	-0,10	**0,77**	0,60
$x10$	**0,95**	0,30	0,99
$x11$	**0,89**	0,42	0,97
$x12$	0,54	**0,60**	0,65
$x13$	**0,82**	0,39	0,82
$x14$	0,40	**0,52**	0,44
$x15$	**0,93**	0,06	0,88
$x16$	**0,74**	0,19	0,58
$x17$	0,46	**0,77**	0,80
$x18$	**0,79**	0,54	0,91

Usualmente o valor mínimo aceitável é de 0,50 (SCHAWB, 2007). Logo, caso o pesquisador encontre alguma comunalidade abaixo desse patamar a variável deve ser removida da amostra e a análise fatorial deve ser realizada novamente. Além disso, baixa comunalidade entre um grupo de variáveis é um indício de que elas não estão linearmente correlacionadas e, por isso, não devem ser incluídas na análise fatorial (FIGUEIREDO FILHO; SILVA JÚNIOR, 2010).

Pela análise da Tabela 3, é possível verificar as cargas fatoriais, com os coeficientes das colunas representando o relacionamento entre cada uma das variáveis e seus respectivos fatores. Em negrito estão as cargas fatoriais com maior valor para as variáveis. O fator 1 é composto por 13 variáveis positivamente relacionadas, formados pelas as variáveis abandono (X1), assédio sexual (X2), ato obseno (X3), danos materiais (X4), estupro (X6), estupro tentado (X7), exercício arbitrário das próprias razões (X8), furto (X10), homicído tentado (X11), importunação ofensiva ao pudor (X13), suicídio (X15), tráfico de mulheres (X16) e violação de domicílio (X18). Como a maioria dos crimes envolvem vários tipos de violências contra as mulheres o fator 1 será denominado de "violência patrimonial, sexual, fisíca e moral contra as mulheres"(Tabela 3 e Figura 2).

Pela análise da Tabela 3 e Figura 2, nota-se que o fator 2 foi composto por 5 variáveis: embriaguez (X5), fato atípico (ocorrências atípicas) (X9), homicídio doloso (X12), subtração de incapaz (X14) e vias de fato (X17). Os crimes do fator 2 foram todos em relação a morte e violência de mulheres, logo, o nome mais adequado para fator 2 é "mortes intencionais e violência de mulheres" . Assim, visando facilitar a compreensão da análise, elaborou-se o diagrama apresentado na Figura 2, sendo este uma forma de demonstrar a formação dos fatores, principalmente quando o estudo apresenta grande número de variáveis.

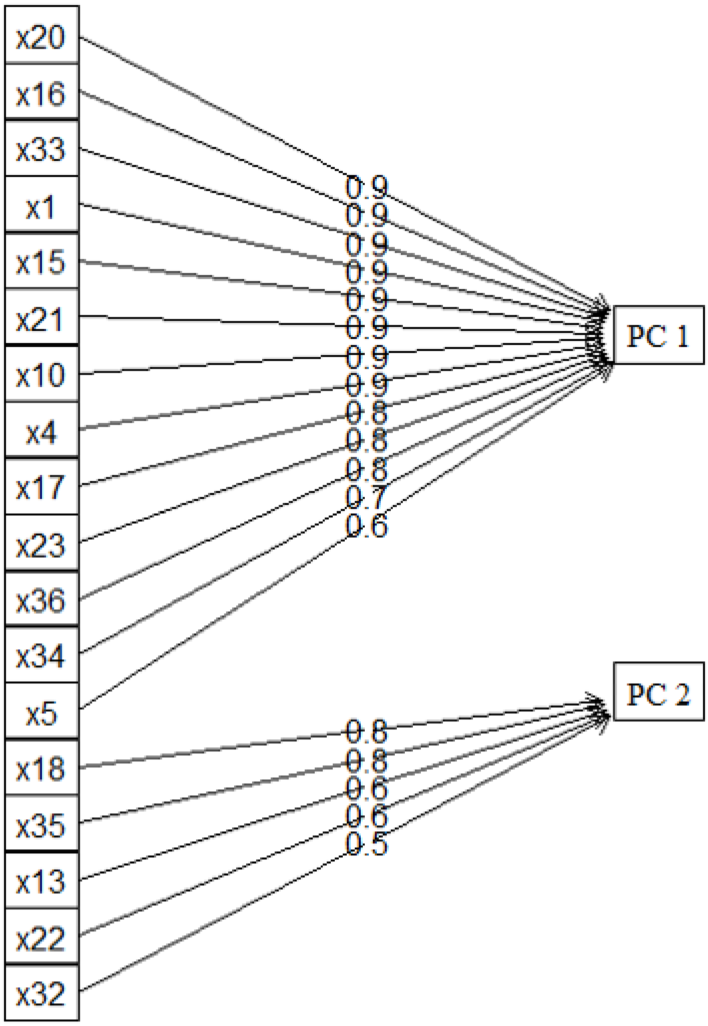

Figura 2: Diagrama da análise fatorial exploratória pelo método de Componente Principal com PC1: fator 1 e PC2: fator 2.

Conclusões

A AFE é uma técnica relativamente complexa que exige dos pesquisadores uma série de decisões para que possa obter uma estrutura fatorial adequada e com essa análise foi possível identificar dois fatores pelo método de Componente Principal, que explicou aproximadamente 80,56% da variância total encontrada, pode-se afirmar que este método foi eficaz e cumpriu os objetivos propostos que são os de resumo e redução dos dados além de permitir a interpretação dos mesmos.

Entretanto, espera-se que as informações aqui presentes possam auxiliar os pesquisadores na condução da técnica de AFE.

Referências bibliográficas

DAMÁSIO, B.F. Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, v.11, n.2, p. 213-228, 2012.

DZIUBAN, C.D.SHIRKEY, E,S. When is a correlation matrix appropriate for factor analysis Some decision rules, *Psychol*, Bull, v.81, p.358-361, 1974.

FIGUEIREDO, F,D. B.; SILVA JÚNIOR, J.A. Visão além do alcance: uma introdução à análise fatorial. *OPINIÃO PÚBLICA*, v.16, n.1, p. 160-185, 2010.

GULUMBE, S. U.; DIKKO, H. G.; BELLO, Y. Analysis of Crime Data using Principal Component Analysis: A case study of Katsina State. *CBN Journal of Applied Statistics*, v.3, n.2, p.39-49, 2013.

HAIR Jr., J.F.; WILLIAM, B.; BABIN, B.; ANDERSON, R.E. *Análise multivariada de dados.* 6.ed. Porto Alegre: Bookman, 2009.

HONGYU, K. Análise Fatorial Exploratória: resumo teórico, aplicação e interpretação. *E&S Engineering and Science*, v.4, n.7, p.88-103, 2018.

HORN, J.L.A rationale and technique for estimating the number of factors in factor analysis. *Psychometrika*, v.30, n.1, p.179-185. 1965.

HUTCHESON, G.D.; SOFRONIOU, N. *The multivariate social scientist: Introductory statistics using generalized linear models.* London: Sage Publications. 1999.

JOHNSON, R.A.; WICHERN, D.W. *Applied multivariate statistical analysis.* 6ª Edição. Madison: Prentice Hall International, p.816, 2008.

KIRCH, J.L.; HONGYU, K.; SILVA, F.L.; DIAS, C.T.S. Análise Fatorial para Avaliação dos Questionários de Satisfação do Curso de Estatística de uma Instituição Federal. *E&S Engineering and Science*, v.6, n.1, p.4-13, 2017.

NEISSE, A. C.; HONGYU, K. Aplicação de componentes principais e análise fatorial a dados criminais de 26 estados dos EUA. E&S Engineering and Science, v.6, n.2, 2016.

PALLANT, J. *SPSS Survival Manual.* Open University Press, 2007.

PASQUALI, L. *Análise fatorial: um manual teórico-prático.* Brasília: EditoraUnB, 1999.

PEREIRA, J.C.R. *Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais.* São Paulo: EDUSP. 1999.

R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2018. Disponível em: https://www.R-project.org/. Acessado em: 18/03/19.

SCHWAB, A. J. *Data analysis and computers II.* 2007.