

# ANÁLISE FATORIAL EXPLORATÓRIA NA AVALIAÇÃO DE PACIENTES COM DOENÇA DE PARKINSON OU DISTÚRPIO COMPORTAMENTAL DE SONO

Elizabete Cristina Röpck <sup>1</sup>  
Jéssica Maria Magalhães Moreno <sup>2</sup>  
Eber Lucas Fernandes Saucedo <sup>3</sup>  
Kuang Hongyu <sup>4</sup>

**RESUMO:** Este artigo em Estatística Multivariada por meio da aplicação da Análise Fatorial Exploratória (AFE) teve como objetivo reduzir o número de variáveis do banco de dados obtidos no site UCI Machine Learning Repository doado por Hlavnicka et al (2017A) num número menor de fatores. A coleta de dados resultou da aplicação da Escala Unificada da Doença de Parkinson (UPDRS) em pacientes com doença de Parkinson ou distúrbio comportamental de sono da categoria avaliação motora, totalizando 80 pacientes avaliados. O resultado da AFE agrupou as 25 variáveis em 3 fatores distintos os quais foram nomeados como: Fator 1 “Toques, movimentos e rigidez de membros superior e inferior direito”; Fator 2 “Tremor e movimentos de membros superior e inferior esquerdo”, e Fator 3 “Postural”. Esses fatores explicaram de forma efetiva e satisfatória os dados originais.

**Palavras-chave:** análise paralela; doença de parkinson; teste de esfericidade; análise de componentes principais; carga fatorial.

## EXPLORATORY FACTOR ANALYSIS IN THE EVALUATION OF PATIENTS WITH PARKINSON'S DISEASE OR SLEEP BEHAVIOR DISORDER

**ABSTRACT:** This article in Multivariate Statistics through the application of Exploratory Factor Analysis (EFA) aimed to reduce the number of variables in the database obtained from the UCI Machine Learning Repository donated by Hlavnicka et al (2017A) on a smaller number of factors. Data collection resulted from the application of the Unified Parkinson's Disease Scale (UPDRS) in patients with Parkinson's disease or sleep behavior disorder in the motor assessment category, totaling 80 patients evaluated. The EFA result grouped the 25 variables into 3 distinct factors, which were named as: Factor 1 "Touches, movements and stiffness of the upper and lower right limbs"; Factor 2 "Tremor and movements of the upper and lower left limbs", and factor 3 "Postural". These factors effectively and satisfactorily explained the original data.

**Keywords:** parallel analysis; Parkinson's disease; sphericity test; principal component analysis; factor loading.

---

<sup>1, 2, 3</sup> Graduanda(o) no Departamento de Estatística, Instituto de Ciências Exatas e da Terra – ICET/UFMT (elizabeteropck@gmail.com, jessicamariamoreno@outlook.com, eberlucas01@outlook.com)

<sup>4</sup> Professor Adjunto do Departamento de Estatística, ICET/UFMT (prof.kuang@gmail.com)

## 1. INTRODUÇÃO

A Doença de Parkinson (DP) é a segunda doença neurodegenerativa mais prevalente, superada apenas pela Doença de Alzheimer (DA). Sua incidência é estimada em cerca de 0,3% na população em geral, aumentando para 1% entre indivíduos com mais de 60 anos e alcançando 3% naqueles com 80 anos ou mais (Draoui et al., 2020).

A DP é caracterizada, segundo as diretrizes da Sociedade Internacional de Distúrbios do Movimento de Parkinson (MDS), pela presença de sinais motores fundamentais, como bradicinesia, tremor em repouso, rigidez ou instabilidade postural. No entanto, sintomas não motores (SNM), como perda do olfato, estados depressivos, distúrbios do sono e constipação, definem a fase prodrômica da condição. Essa fase pode começar até duas décadas antes do diagnóstico clínico formal da DP (Berg et al., 2015; Postuma et al., 2014).

Para rastrear a progressão da DP e avaliar a eficácia do tratamento medicamentoso, é amplamente utilizada a Escala Unificada de Avaliação da Doença de Parkinson (UPDRS), desenvolvida em 1987. Esta escala é composta por 42 itens distribuídos em quatro categorias: I - atividade mental, comportamento e humor; II - atividades de vida diária (AVD); III - avaliação motora; e IV - complicações decorrentes da terapia medicamentosa. Uma pontuação mais alta indica um maior impacto da doença (Fahn; Elton, 1987; Hughes et al., 1992).

A validação de um instrumento requer uma análise cautelosa da capacidade de agrupamento dos itens (Hoss, Caten, 2010). A Análise Fatorial Exploratória (AFE) permite a divisão da escala em fatores e a identificação das questões mais representativas do instrumento. A AFE é uma técnica estatística que investiga as correlações entre muitas variáveis, agrupando-as em fatores.

Essa técnica possibilita a redução de dados, identificando as variáveis mais representativas ou criando um conjunto de variáveis menor que o original (Hair et al., 2005). O objetivo deste estudo é explorar os fatores fundamentais que caracterizam os pacientes com Doença de Parkinson (DP), por meio da abordagem estatística multivariada da AFE.

## 2. MATERIAIS E MÉTODOS

### 2.1 Banco de dados

O conjunto de dados utilizado neste estudo foi obtido no site UCI Machine Learning Repository que contém uma coleção de bancos de dados usados para aprendizado de máquina (Dua; Graff, 2019). Os dados foram doados por Hlavnicka et al (2017A) e analisados originalmente por Hlavnicka et al (2017B). A pesquisa foi realizada com 130 pacientes sendo 30 pacientes com doença de Parkinson (DP) precoce não tratada, 50 indivíduos com distúrbio comportamental do sono de movimento rápido dos olhos (RBD) e 50 controles saudáveis.

De acordo com Hlavnicka (2017B), todos os pacientes foram submetidos a avaliação clínica por um neurologista profissional bem treinado e com experiência em distúrbios do movimento.

Neste estudo foram utilizados apenas os dados dos pacientes com DP e RBD (que apresentam alto risco de desenvolver DP ou outras sinucleinopatias), com base na referência à avaliação neurológica usando UPDRS III (Tabela 1), as variáveis avaliadas são pontuadas em uma escala de classificação de 0 a 4 e pontuações mais altas mostram maior gravidade da doença. Na tabela 1, estão relacionadas as 25 variáveis utilizadas na análise deste estudo.

**Tabela 1. Descrição das 25 variáveis**

Variável	Descrição da Variável
X1	Fala
X2	Expressão Facial
X3	Tremor em Repouso – cabeça
X4	Tremor em Repouso - Membro Sup. Direito (MSD)
X5	Tremor em Repouso – MID
X6	Tremor de Ação ou Postural – MSD
X7	Tremor de Ação ou Postural – MSE
X8	Rigidez – Pescoço
X9	Rigidez – MSD
X10	Rigidez – MSE
X11	Rigidez - Membro Inferior Direito (MID)
X12	Rigidez - Membro Inferior Esquerdo (MIE)
X13	Toques com os Dedos – MSD
X14	Toques com os Dedos – MSD
X15	Movimentos da Mão – MSD
X16	Movimentos da Mão – MSE
X17	Movimentos Alternados Rápidos – MSD
X18	Movimentos Alternados Rápidos – MSE
X19	Agilidade das Pernas – MID
X20	Agilidade das Pernas – MIE
X21	Levantar-se da Cadeira
X22	Postura
X23	Marcha
X24	Estabilidade Postural
X25	Bradícinesia e Hipocinesia Corporal

Todas as análises deste artigo foram feitas por meio de rotinas computacionais implementadas no software R 4.2.3 (R Development Core Team, 2023) com o pacote “corrplot”, “factoextra” e “psych”.

## 2.2 Pressupostos da análise fatorial exploratória

A análise fatorial exploratória (AFE) é uma técnica multivariada utilizada para investigar a estrutura subjacente de um conjunto de dados e procura identificar uma quantidade relativamente pequena de fatores que representam o comportamento conjunto de variáveis originais interdependentes utilizando coeficientes de correlação para agrupar variáveis e gerar fatores (Damásio, 2012; Favero, 2018).

A AFE é frequentemente empregada quando deseja entender como variações observadas estão relacionadas entre si e se podem ser agrupadas em fatores latentes que capturam essas relações. Na análise fatorial exploratória, os dados são submetidos a uma análise estatística que busca identificar a relação entre variáveis e organizá-las em fatores. (Damásio, 2012; Hongyu, 2018).

Na AFE é importante verificar se faz sentido usar essa técnica estatística para as variáveis que escolhemos. Para fazer isso, usamos dois métodos comuns: o critério de Kaiser-Meyer-Olkin (KMO) e o Teste de Esfericidade de Bartlett. Esses métodos nos ajudam a decidir se a AFE é adequada para nossos dados (Dziuban; Shirkey, 1974; Damásio, 2012; Hongyu, 2018).

Para verificar a adequação global propriamente dita da extração dos fatores, devemos recorrer à estatística KMO e ao teste de Bartlett. O KMO, é uma medida estatística que indica quanto nossos dados estão relacionados com uma variável subjacente. Se o valor do KMO for alto, significa que podemos usar a Análise Fatorial Exploratória com confiança em nossos dados (Damásio, 2012; Hongyu, 2018; Favero, 2018).

A estatística KMO fornece a proporção de variância considerada comum a todas as variáveis na amostra em análise, ou seja, que pode ser atribuída à existência de um fator comum. Essa estatística varia de 0 a 1, e, enquanto valores mais próximos de 1 indicam que as variáveis compartilham um percentual de variância bastante elevado (correlações de Pearson altas), valores mais próximos de 0 são decorrentes de correlações de Pearson baixas entre as variáveis, o que pode indicar que a análise fatorial será inadequada.

De acordo com Vicini, et al. (2018) os valores críticos e interpretação das faixas de adequação do KMO são definidos como: valores menores que 0,5 são considerados inaceitáveis, e 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. Com isso, a medida aceitável de adequação da amostra precisa ser maior que 0,5.

O coeficiente KMO é dado pelo quadrado das correlações totais dividido pelo quadrado das correlações parciais das variáveis analisadas, por meio da equação (1) (Dziuban; Shirkey, 1974):

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

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.

A medida de adequação da amostra (MAA) é uma estatística amplamente utilizada na AFE e em análises estatísticas multivariadas. Ela desempenha um papel fundamental na determinação da adequação dos dados para a aplicação da AFE medindo a proporção da variância dos dados que pode ser explicada por fatores subjacentes.

Em outras palavras, o MAA ajuda a determinar se as variações observadas estão inter-relacionadas o suficiente para incluir a eliminação de fatores latentes que foram explicados. A MAA é representada por um valor que varia de 0 a 1. Quanto mais próximo o valor for de 1, mais adequados os dados são para a análise fatorial exploratória e valores mais próximos de 0 indicam que os dados são inadequados para análise por meio da AFE.

O Teste de Esfericidade de Bartlett é uma técnica estatística usada em análise fatorial para avaliar se as correlações entre as observações analisadas em um conjunto de dados são significativas o suficiente para justificar a aplicação da análise fatorial e determinar se os dados possuem uma estrutura fatorial subjacente.

Se o teste de Bartlett indicar que as correlações entre as variáveis são significativas, isso sugere que os dados têm uma estrutura subjacente que pode ser explorada por meio da análise fatorial. O teste de esfericidade de Bartlett verifica o quanto a matriz de covariância difere de uma matriz identidade, ou seja, se as variáveis não estão independentes entre si (Field, 2005).

Além disso, esse teste verifica a importância global de todas as correlações presentes em uma matriz de dados (Hair et al., 2009). O resultado do teste de esfericidade de Bartlett, realizado com nível de significância ( $p < 0,05$ ), sugere que a matriz pode ser fatorada. Isso ocorre porque a hipótese nula de que uma matriz de dados é semelhante a uma matriz identidade é rejeitada (Tabachnick; Fidell, 2007).

De forma geral, os resultados dos testes de KMO e de esfericidade de Bartlett tendem a concordar, seja aceitando ou rejeitando a possibilidade de fatoração da matriz de dados, conforme apresentado por Dziuban; Shirkey (1974).

### 2.3 Modelo da análise fatorial

O modelo de análise de fatores é dado pela equação (2):

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + e_i \quad (2)$$

em que  $X_i$  é o  $i$ -ésimo escore após realizada a padronização (com média 0 e desvio-padrão 1);  $i = 1, \dots, p$  onde  $p$  representa o número de variáveis;  $a_{i1}, a_{i2}, \dots, a_{im}$  são as cargas dos fatores para o  $i$ -ésimo teste;  $F_1, F_2, \dots, F_m$  indica  $m$  fatores comuns não correlacionados, cada um com média 0 e variância 1; e  $e_i$  é um fator exclusivo para o  $i$ -ésimo teste, este é indicado quando não há correlação com nenhum dos fatores comuns e possui média zero (Johnson; Wichern, 2007; Hongyu, 2018).

Os valores de  $p$  observados  $X_p$  são expressos em relação de  $p + m$  os quais são variáveis não observáveis ( $F_1, F_2, \dots, F_m; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ ). Isso é o que diferencia o modelo fatorial do modelo de regressão múltipla, onde as variáveis independentes são observadas, já no modelo fatorial, essas posições são ocupadas por  $F$ . Matricialmente temos (Johnson; Wichern, 2007):

$$\mathbf{X}_{(p \times 1)} = \mathbf{\Lambda}_{(p \times m)} \mathbf{F}_{(m \times 1)} + \boldsymbol{\varepsilon}_{(p \times 1)} \quad (3)$$

Enquanto isso, o modelo de Análise Fatorial é dado pela equação (4):

$$\begin{aligned} \text{Var}(X_j) &= a_{j1}^2 \text{Var}(F_1) + a_{j2}^2 \text{Var}(F_2) + \dots + a_{jm}^2 \text{Var}(F_m) + \text{Var}(e_j) \\ &= a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2 + \text{Var}(e_j) \end{aligned} \quad (4)$$

em que  $a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2$  é chamada a comunalidade de  $X_j$  (refere-se à porção de sua variância que é explicada pelos fatores compartilhados). A comunalidade não deve ser maior que 1, é preciso que  $-1 \leq a_{ij} \leq +1$ . Também a correlação de  $X_j$  e  $X_j$ , pode ser representada da seguinte forma:

$$r_{jj'} = a_{j1}a_{j'1} + a_{j2}a_{j'2} + \dots + a_{jm}a_{j'm} \quad (5)$$

Assim, podemos dizer que duas variáveis têm uma relação forte quando ambas têm um peso importante no mesmo fator. (Neisse; Hongyu, 2016; Hongyu, 2018).

### 2.4 Retenção de Fatores

Diversos métodos e critérios de seleção de fatores foram desenvolvidos ao longo do tempo (Damásio, 2012). Entre esses métodos, o critério de Kaiser-Guttman (Guttman, 1954; Kaiser, 1960) é amplamente utilizado, também conhecido como o critério do autovalor (*eigenvalue*), maior do que 1,0 ( $\lambda_i > 1$ ) oferece uma avaliação rápida e objetiva do número de

fatores a serem retidos, e cada fator retido possui um autovalor que reflete a quantidade de variância explicada por esse fator (Damásio, 2012).

Outro critério comum é o baseado no gráfico do *screeplot*, também conhecido como teste de Cattell (1966), que identifica o ponto de inflexão no gráfico onde os autovalores exibem uma tendência linear decrescente. Geralmente, os resultados obtidos através da análise fatorial usando os critérios de Kaiser-Guttman ou o método do *screeplot* são consistentes entre si.

Um terceiro critério, o método de análises paralelas (AP) (Horn, 1965), ganhou destaque na literatura internacional e é implementado no software R no pacote "psych". Embora tenha sido desenvolvido originalmente como uma estratégia de retenção de componentes, o método de análises paralelas foi adaptado para ser utilizado no contexto da AFE, sendo considerado um procedimento adequado para determinar o número de fatores a serem retidos (Lorenzo-Seva; Timmerman; Kiers, 2011).

O método da AP envolve uma comparação dos autovalores da amostra de pesquisa com aqueles gerados a partir de uma amostra investigada do mesmo tamanho, retendo apenas os autovalores que excedem os valores gerados na amostra pesquisada.

## 2.5 Procedimentos gerais da análise fatorial

O processo de análise de fatores é composto por três etapas distintas. O primeiro desafio é determinar as cargas dos fatores ( $a_{jk}$ ), onde  $j$  varia de 1 a  $p$  e  $k$  varia de 1 a  $m$  (onde  $m < p$ ). A primeira etapa desse processo envolve a escolha dos métodos de determinação das cargas dos fatores (Hongyu, 2018). Na literatura, existem várias abordagens de estudos para análise fatorial disponíveis, no entanto, neste estudo será aplicado o Componente Principal (CP)

Uma vez selecionado o método para calcular as cargas dos fatores, é importante notar que essas cargas não são exclusivas. Se tivermos fatores provisórios, como  $F_1, F_2, \dots, F_m$  é possível expressar arranjos lineares desses fatores na forma mostrada na Equação (6) (Johnson; Wichern, 2008).

$$\begin{aligned} F'_1 &= d_{11}F_1 + d_{12}F_2 + \dots + d_{1m}F_m \\ F'_2 &= d_{21}F_1 + d_{22}F_2 + \dots + d_{2m}F_m \\ &\vdots \\ F'_m &= d_{m1}F_1 + d_{m2}F_2 + \dots + d_{mm}F_m \end{aligned} \quad (6)$$

É possível criar fatores que não estejam correlacionados entre si e que, ao mesmo tempo, expliquem os dados com a mesma eficácia dos fatores originais. Na realidade, existe uma ampla variedade de soluções para o modelo de análise de fatores (Hongyu, 2018).

## 2.6 Análise fatorial por componentes principais

Iniciando com a matriz  $X_{n \times p}$ , realizamos uma análise de componentes principais, resultando em  $p$  componentes principais, conforme expresso na Equação (7), de acordo com Johnson; Wichern, 2008):

$$Z_i = b_{i1}X_1 + b_{i2}X_2 + \dots + b_{ip}X_p \quad (7)$$

em que  $i = 1, \dots, p$ , e  $b_{ij}$  são determinados pelos autovetores da matriz de correlações. Devido a ser uma transformação ortogonal, a relação inversa é descrita pela Equação (8):

$$X_i = b_{i1}Z_1 + b_{i2}Z_2 + \dots + b_{ip}Z_p \quad (8)$$

O modelo de fatores com  $m < p$  é então expresso pela seguinte equação (9):

$$X_i = b_{i1}Z_1 + b_{i2}Z_2 + \dots + b_{im}Z_m + e_i \quad (9)$$

Nessa expressão,  $e_i$  é uma combinação linear dos componentes que varia de  $Z_{m+1}$  até  $Z_p$ .

Para obter  $F_i$ , cada  $Z_i$  é dividido pelo seu desvio padrão, que é igual a  $\sqrt{\lambda_i}$  (correspondente ao autovalor na matriz de correlações). O cálculo dessas equações é realizado usando a fórmula (10):

$$X_i = \sqrt{\lambda_1}b_{i1}F_1 + \sqrt{\lambda_2}b_{i2}F_2 + \dots + \sqrt{\lambda_m}b_{im}F_m + e_i \quad (10)$$

Assumindo que  $a_{ij} = \sqrt{\lambda_j}b_{ji}$  é possível demonstrar que  $X_i$  é dado pela Equação (11):

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + e_i \quad (11)$$

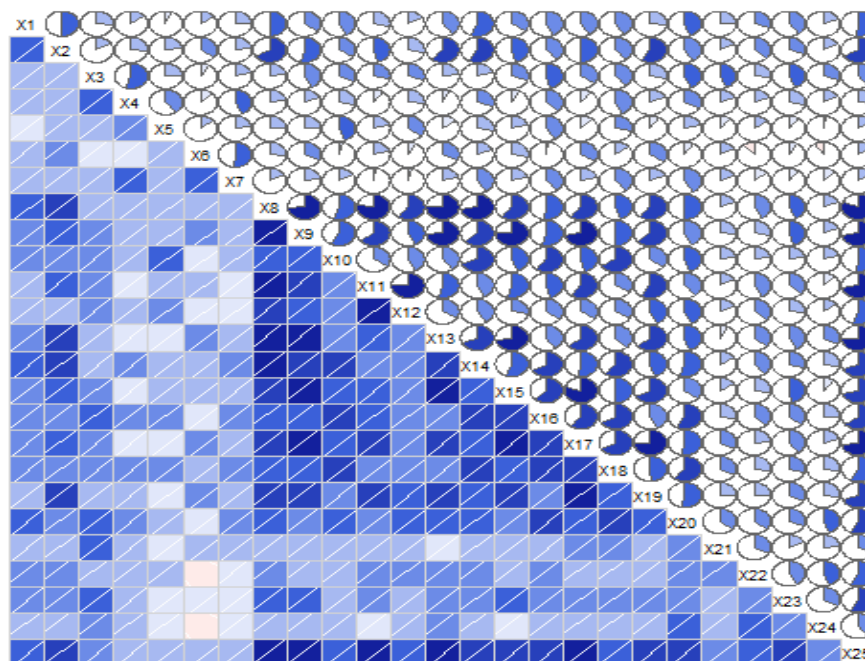
### 3. RESULTADOS E DISCUSSÕES

A correlação é uma medida do grau da relação linear entre duas variáveis métricas, variando entre  $-1$  e  $1$ , um valor próximo desses extremos indica a existência de relação linear entre as duas variáveis analisadas podendo contribuir significativamente para a extração de um único fator. Por outro lado, um valor da correlação muito próximo de  $0$  indica que a relação linear entre as duas variáveis é praticamente inexistente (Favero, 2018).

Portanto, diferentes fatores podem ser extraídos. A matriz de correlação (Figura 1), demonstra os valores das correlações entre as variáveis sendo representadas pelo tom azul e quanto mais a correlação se aproxima de  $1$ , apresenta cor azul mais escura e, quanto mais se aproxima de  $-1$  apresenta cor vermelha mais escura.

Sendo assim, as variáveis de UPDRS III, os itens que apresentaram as maiores correlações foram as variáveis: X8 *versus* X25 (Rigidez – Pescoço *versus* Bradicinesia e Hipocinesia Corporal); X15 *versus* X17 (Movimentos da Mão – Membro Superior Direito (MSD) *versus* Movimentos Alternados Rápidos – MSD); X8 *versus* X9 (Rigidez – Pescoço *versus* Rigidez – MSD); X8 *versus* X11 (Rigidez – Pescoço *versus* Rigidez - Membro Inferior Direito (MID)); X8 *versus* X13 (Rigidez – Pescoço *versus* Toques com os Dedos – MSD e X8 *versus* X14 (Rigidez – Pescoço *versus* Toques com os Dedos – MSD).

**Figura 1. Matriz de correlação das 25 variáveis de Escala Motora Unificada de Avaliação da Doença de Parkinson.**



Para avaliar a aceitabilidade da aplicação da AFE aos dados, foram calculados o valor global de KMO para todas as variáveis obtendo-se o resultado igual a 0,84 e o teste de esfericidade de Bartlett retornou o valor-p do teste muito menor que 0,05, estatisticamente significativo, isso significa que a hipótese nula deve ser rejeitada com uma significância de 5% e a matriz de correlações é diferente da identidade (Tabela 2), aceitando assim, a possibilidade de fatoração da matriz de dados (Matos; Rodrigues, 2019; Hongyu, 2018).

**Tabela 2. Índice KMO e teste de esfericidade de Bartlett**

Fator de Adequação de Kaiser-Meyer-Olkin		0,84
Teste de Esfericidade de Bartlett	Qui-quadrado aprox.	1966.804
	gl	300
Valor-p		<2.095437e-242

Foi verificado ainda, o quanto a amostra é adequada para a realização da análise fatorial por meio do KMO para cada uma das variáveis (MAA), (Tabela 3). O resultado retornou que apenas uma variável apresentou valor inferior a 0,6, o que indica que a amostra é consideravelmente concebível para a aplicação da análise fatorial. Podendo então, prosseguir com a extração de fatores pela AFE.



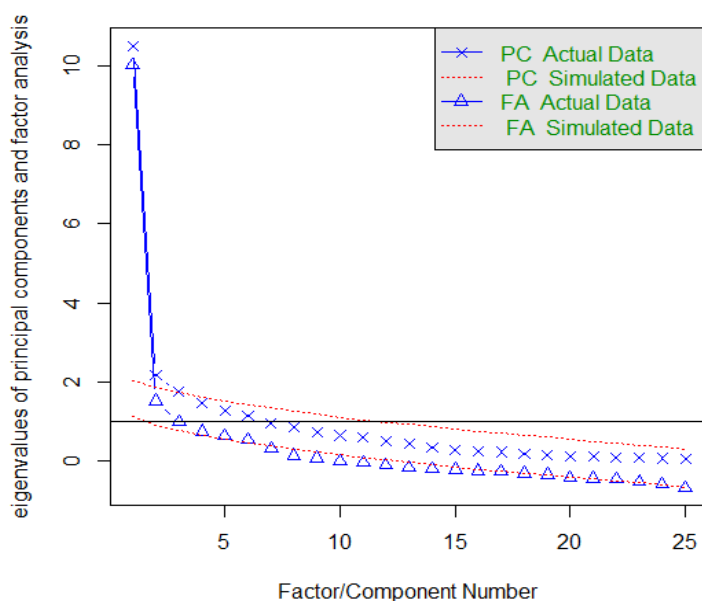
**Tabela 3. Medida de Adequação Amostral entre as 25 variáveis estudadas.**

Variáveis	MAA
X1	0,86
X2	0,87
X3	0,72
X4	0,61
X5	0,60
X6	0,58
X7	0,61
X8	0,88
X9	0,93
X10	0,86
X11	0,79
X12	0,77
X13	0,87
X14	0,87
X15	0,86
X16	0,84
X17	0,89
X18	0,91
X19	0,90
X20	0,84
X21	0,69
X22	0,80
X23	0,87
X24	0,77
X25	0,93

A AP foi escolhida por reduzir a probabilidade de retenção equivocada de itens, considerar o erro amostral e minimizar a influência do tamanho da amostra e das cargas fatoriais dos itens. Analisando a Figura 2, nota-se que o critério de AP indicou que três fatores foram ideais para ser extraído para dar continuidade a análise fatorial exploratória (Hongyu, 2018).

Em continuação a análise, foi obtido o *Screeplot*, que sugeriu a retenção em cinco fatores, ou seja, autovalores maiores que 1. Analisando os resultados encontrados, decidiu-se pela retenção de três fatores, por verificar pouca variação na porcentagem de variância explicada, a qual é calculada como a razão dos autovalores sobre a soma de todos os autovalores (Matos; Rodrigues, 2019)

Figura 2. *Screeplot* dos autovalores da análise fatorial por componentes principais.



Com os fatores estabelecidos, devemos definir as cargas fatoriais, que são as correlações de Pearson entre as variáveis originais e cada um dos fatores. A Tabela 4, apresenta as cargas fatoriais e comunalidades do Fator 1, 2 e 3 pelo método análise paralelas. As comunalidades representam a proporção da variância para cada variável incluída na análise que é explicada pelos componentes extraídos (Favero, 2017; Hongyu, 2018).

Tabela 4. Cargas fatoriais e a comunalidade para as 25 variáveis.

Variável	Descrição da Variável	Fator 1	Fator 2	Fator 3	Comunalidade
X1	Fala	<b>0,45</b>	0,33	0,14	0,33
X2	Expressão Facial	<b>0,69</b>	0,22	-0,01	0,53
X3	Tremor em Repouso – cabeça	0,16	<b>0,59</b>	0,36	0,50
X4	Tremor em Repouso - Membro Sup. Direito (MSD)	-0,01	<b>0,75</b>	0,10	0,58
X5	Tremor em Repouso – MID	0,10	<b>0,53</b>	0,00	0,29
X6	Tremor de Ação ou Postural – MSD	0,40	0,25	<b>-0,63</b>	0,62
X7	Tremor de Ação ou Postural – MSE	0,14	<b>0,70</b>	-0,37	0,65
X8	Rigidez – Pescoço	<b>0,85</b>	0,17	0,19	0,79
X9	Rigidez – MSD	<b>0,82</b>	0,25	0,06	0,74
X10	Rigidez – MSE	0,45	<b>0,50</b>	0,17	0,49
X11	Rigidez - Membro Inferior Direito (MID)	<b>0,76</b>	0,01	0,29	0,66
X12	Rigidez - Membro Inferior Esquerdo (MIE)	<b>0,48</b>	0,18	0,39	0,41
X13	Toques com os Dedos – MSD	<b>0,85</b>	0,09	0,07	0,74
X14	Toques com os Dedos – MSD	<b>0,66</b>	0,45	0,08	0,64
X15	Movimentos da Mão – MSD	<b>0,82</b>	0,16	-0,01	0,70
X16	Movimentos da Mão – MSE	0,49	<b>0,60</b>	0,24	0,66
X17	Movimentos Alternados Rápidos - MSD	<b>0,81</b>	0,25	0,02	0,72
X18	Movimentos Alternados Rápidos - MSE	0,45	<b>0,65</b>	0,12	0,64
X19	Agilidade das Pernas – MID	<b>0,77</b>	0,17	0,05	0,63
X20	Agilidade das Pernas – MIE	0,43	<b>0,57</b>	0,29	0,58
X21	Levantar-se da Cadeira	0,12	<b>0,44</b>	0,25	0,27
X22	Postura	0,31	0,20	<b>0,63</b>	0,54
X23	Marcha	<b>0,43</b>	0,20	0,42	0,40
X24	Estabilidade Postural	0,13	0,26	<b>0,57</b>	0,41
X25	Bradicinesia e Hipocinesia Corporal	<b>0,82</b>	0,24	0,40	0,89

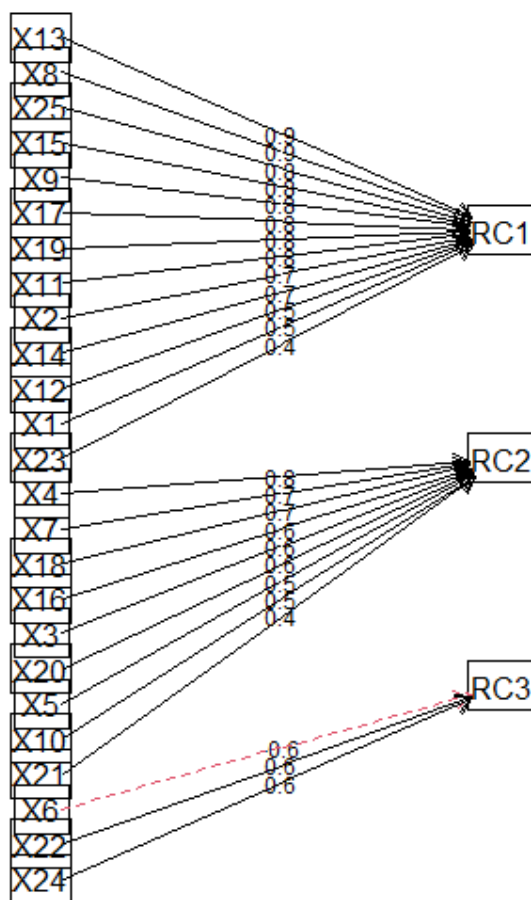
Sendo assim, os fatores foram divididos e nomeados de acordo com a Tabela 4; Figura 2; Fator 1 – Toques, movimentos e rigidez de membros superior e inferior direito que agrupou as variáveis: X13, X8, X25, X15, X9, X17, X19, X11, X2, X14, X12, X1 e X23. O fator 1 agregou as variáveis relacionadas a Rigidez, Movimentos, Agilidade dos membros superior e inferior direito, além das variáveis de marcha, fala e expressão facial durante o desempenho da avaliação.

As variáveis, marcha e fala apresentaram as menores cargas fatoriais 0,43 e 0,45 respectivamente, essas variáveis têm menor contribuição para esse fator quando analisadas em relação as demais variáveis. À medida que a comunalidade aumenta, o modelo fatorial ganha uma maior capacidade de explicação para o item. Geralmente, espera-se que as comunalidades sejam superiores a 0,5 (Tabela 4; Figura 3).

O Fator 2, denominado Tremor e movimentos de membros superior e inferior esquerdo concatenou as variáveis X13, X8, X25, X15, X9, X17, X19, X11, X2, X14, X12, X1 e X23 relacionadas aos movimentos dos membros inferior e superior esquerdo. Este fator reflete as variáveis: Tremor em Repouso, Tremor de Ação ou Postural, Movimentos Alternados Rápidos, Movimentos da Mão, Agilidade das Pernas, Rigidez e Levantar-se da Cadeira (Tabela 4; Figura 3).

O fator 3, chamado de Postural, agrupou as variáveis X6, X22 e X24. Este fator agrupou as variáveis relacionadas a postura, sendo elas: Tremor de Ação ou Postural, Postura e Estabilidade Postural (Tabela 4; Figura 3), sendo observado que a variável X6 apresentou uma correlação negativa.

Figura 3. Diagrama da AFE por CP com RC1: fator 1, RC2: fator 2 e RC3: fator 3.



#### **4. CONCLUSÃO/CONSIDERAÇÕES FINAIS**

De acordo com os resultados, a aplicação da AFE foi efetiva, pois permitiu agrupar as 25 variáveis em 3 fatores de acordo com as características específicas da avaliação motora da Escala Unificada de Avaliação da Doença de Parkinson (UPDRS). Esses 3 fatores explicaram de forma satisfatória os dados originais.

Portanto, pode-se afirmar que o método foi eficaz, atendeu os objetivos propostos, pois possibilitou a redução dos dados, criando um conjunto de variáveis menor que os dados originais, agrupando os fatores de acordo com a lateralização corporal dos membros em relação aos movimentos que são afetados e caracterizam os pacientes com Doença de Parkinson.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BARLETT, M. S. Tests of significance in factor analysis. *British Journal of Psychology*, v.3, p.77-85, 1950.
- BERG, Daniela et al. MDS research criteria for prodromal Parkinson's disease. *Movement Disorders*, v. 30, n. 12, p. 1600-1611, 2015. Disponível em: <https://doi.org/10.1002/mds.26431>. Acesso em: 15 set. 2023
- CATTELL, R. B. The scree test for the number of factors. *Multivariate Behavioral Research*, v.1, n.1, p.245-276, 1966.
- COSTELLO, A. B.; OSBORNE, J. W. Best practices in exploratory factor analysis: Fourre commendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, v. 10, n.7, p.1-9, 2005.
- DAMÁSIO, B. F. Uso da análise fatorial exploratória em psicologia. *Avaliação Psico-lógica*, v.11, n.2, p. 213-228, 2012.
- DRAOUI, A. et al. Parkinson's disease: from bench to bedside. *Revue neurologique*, v. 176, n. 7-8, p. 543-559, 2020. Disponível em: <https://doi.org/10.1016/j.neurol.2019.11.002>. Acesso em: 25 set. 2023.
- DUA, D.; GRAFF, C. **UCI Machine Learning Repository**. Irvine, CA, 2019. Disponível em: <http://archive.ics.uci.edu/ml>>. Acesso em: 21 ago. 2023.
- DZIUBAN, C.; SHIRKEY, E. When is a correlation matrix appropriate for factor analysis? Some decision rules, *Psychol, Bull*, v.81, p.358-361, 1974. Field, Andy *Discovering Statistics Using SPSS*. Ed. 2, London: Sage. 2005.
- FAHN, S.; ELTON, R. L. Unified Parkinson's disease rating scale. *Recent developments in Parkinson's disease*, p. 153-163, 1987.
- FIELD, A. *Discovering Statistics Using SPSS*. Ed. 2, London: Sage. 2005.
- GUTTMAN, L. Some necessary conditions for common factor analysis, *Psychometrika*, v.19, p.149-162, 1954.
- HAIR, J. F., ANDERSON, R. E., TATHAM, R. L. BLACK, W. C. *Análise multivariada de dados*. A. S. Sant'Anna & A. C. Neto (Trad.). Porto Alegre: Bookman. 2005.
- HAIR, J. F.; WILLIAM, B.; BABIN, B.; ANDERSON, R. E. *Análise multivariada de dados*. 6.ed. Porto Alegre: Bookman, 2009.
- HARMAN, H.H. *Modern Factor Analysis*. 2ª Edição. Chicago: University of Chicago Press, 1967.
- HLAVNIKA, J. et al. Early biomarkers of Parkinson's disease based on natural connected speech. *UCI Machine Learning Repository*. 2017A. Disponível em

<https://archive.ics.uci.edu/dataset/392/early+biomarkers+of+parkinson+s+disease+based+on+natural+connected+speech>, Acesso em 21 ago. 2023.

HLAVNIKA, J. et al. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Sci Rep* 7. 2017B. Disponível em: <<https://doi.org/10.1038/s41598-017-00047-5>>. Acesso em 21 ago. 2023.

HONGYU, K. (2018) Análise Fatorial Exploratória: resumo teórico, aplicação e interpretação -. (Acessos: 3803; Volume: 7, Nº 4, Ano: 2018; DOI: <https://doi.org/10.18607/ES201877599>). Acesso em: 20 ago. 2022.

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.

HUGHES, A. J. et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinicopathological study of 100 cases. *Journal of neurology, neurosurgery & psychiatry*, v. 55, n. 3, p. 181-184, 1992. Disponível em: <https://doi.org/10.1136/jnnp.55.3.181>. Acesso em: 20 set. 2022.

JOHNSON, R.A. e Wichern, D.W. (2007) Análise Estatística Multivariada Aplicada. 6ª Edição, Pearson Prentice Hall, Upper Saddle River

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

KAISER, H.F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, v.23, p.187-200. 1958.

KAISER, H. F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, v.20, p.141-151, 1960.

KASSAMBARA, A, MUNDT, F. `_factoextra`: Extract and Visualize the Results of Multivariate Data Analyses\_. R package version 1.0.7, 2020. Disponível em: <<https://CRAN.R-project.org/package=factoextra>>. Acesso em: 20 jul. 2023.

KIM, J.; MUELLER, C. W. Factor analysis: Statistical methods and practical issues. Beverly Hills, CA: Sage Publications. 1978.

LORENZO-SEVA, U.; TIMMERMAN, M. E.; KIERS, H.A. The hull method for selecting the number of common factors. *Multivariate Behavioral Research*, v.46, n.2, p.340-364, 2011.

MATOS, D. A. S.; RODRIGUES, E. C. Análise fatorial. Brasília: Enap, 2019.

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. 5, N. 2. 2016.

POSTUMA, R. B. et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain*, v. 142, n. 3, p. 744-759, 2019. Disponível em: <https://doi.org/10.1093/brain/awz030>. Acesso em: 15 set. 2023.

R CORE TEAM. **R**: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2023. Disponível em: <<https://www.R-project.org/>>. Acesso em: 20 jul. 2023.

REVELLE, G. `_psych`: Procedimentos para pesquisa psicológica, psicométrica e de personalidade. Universidade Northwestern, Evanston, Illinois. Pacote R versão 2.3.9. 2023. Disponível em: <<https://CRAN.R-project.org/package=psych>>. Acesso em: 20 jul. 2023.

TABACHNICK, B.; FIDELL, L. Using multivariate analysis. Needham Heights: Allyn & Bacon, 2007.

TAIYUN, W.; Viliam, S. R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). 2021. Disponível em: <<https://github.com/taiyun/corrplot>>. Acesso em: 20 jul. 2023.

VELICER, W. F. Determining the number of components from the matrix of partial correlations. *Psychometrika*, v.41, n.3, p.321-327, 1976.

VICINI, L. et al. Técnicas multivariadas exploratórias: teorias e aplicações no software *Statística®*. Santa Maria: Ed. UFSM, 2018. Ebook. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em 20 set 2023..