

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS EXATAS  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA

Kellen Roberta Rocha da Silva

Orientadora: Professora Sueli Aparecida Mingoti  
PH.D.em Estatística

Belo Horizonte  
Novembro de 2017

KELLEN ROBERTA ROCHA DA SILVA

A ANÁLISE MULTIVARIADA APLICADA NA SELEÇÃO DE FUNDOS DE RENDA VARIÁVEL

Kellen Roberta Rocha da Silva

Orientadora: Professora Sueli Aparecida Mingoti -PH.D.em Estatística

Monografia apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título Especialista em Estatística.

Belo Horizonte  
Novembro de 2017

## RESUMO

A seleção de fundos de investimento é um processo constante no mercado de Previdência Complementar fechado e de grande relevância devido a regulamentação existente para a gestão de recursos dos participantes que precisam ter o risco minimizado devido a necessidade de formação de poupança de longo prazo para fazer frente a época da aposentadoria.

Neste trabalho pretendeu-se, através a análise fatorial, obter uma ferramenta para seleção de fundos de renda variável.

A análise fatorial é um método de estatística multivariada que é muito utilizado em análise de banco de dados numerosos e grande quantidade de variáveis com o intuito de identificar grupos de variáveis mais correlacionadas entre si e construir índices com as variáveis de cada grupo, permitindo uma análise segmenta da informação.

O banco de dados estudado nessa monografia está posicionado em 2016 e conta com mais de 1700 fundos de investimentos em renda variável. Foram avaliadas 11 variáveis relacionadas com rentabilidade, tempo de existência, risco e desempenho do fundo.

Palavras-chave: Investimento, análise fatorial, renda variável, índices, fundos.

## LISTA DE QUADROS

Quadro 1: Descrição das variáveis utilizadas.....	20
Quadro 2: Fundos de Renda Variável, por segmento.....	21
Quadro 3: Estatísticas descritivas – Base de dados Agosto 2016 .....	23
Quadro 4. Correlação de Pearson entre as variáveis .....	23
Quadro 5-Autovalores da matriz de correlação das 9 variáveis .....	25
Quadro 6- Análise Fatorial – matriz de correlação – m=4 – p=9.....	28
Quadro 7: Análise Fatorial – matriz de correlação – m=3 – p=9.....	29
Quadro 8: Análise Fatorial – matriz de correlação – m=4 – p=8.....	30
Quadro 9: Análise Fatorial – matriz de correlação – m=3 – p=8.....	31
Quadro 10: Fator 1 – Comunalidade e Variância Específica .....	32
Quadro 11: Fator 2 – Comunalidade e Variância Específica .....	33
Quadro 12: Fator 3 – Comunalidade e Variância Específica .....	33
Quadro 13: Coeficiente dos fatores dos escores – Fatores via Rotação Varimax.....	34
Quadro 14:cálculo do escore do Fator 1, para o fundo 210FIA.....	35
Quadro 15:Estatísticas Descritivas dos Escores dos Fatores. m=3 .....	37
Quadro 16:Fundos outliers – Dados Originais e Padronizados.....	38
Quadro 17:Análise Fatorial sem fundos com escores discrepantes .....	39
Quadro 18: 5 Melhores e 5 Piores Fundos e suas variáveis originais .....	40
Quadro 19:05 Melhores e 05 Piores Fundos e suas variáveis Padronizadas.....	41
Quadro 20: Escores dos 15 Fundos melhores posicionados, por Fator 1, Fator 2 e Fator 3. ....	41
Quadro 21: Retornos do Fundo 831 em Fev/17. ....	42

## LISTA DE ABREVIATURAS

ANBIMA - Associação Brasileira das Entidades dos Mercados Financeiros e de Capitais,

CVM – Comissão de Valores Mobiliários

EFPC - Entidades Fechadas de Previdência Complementar

PREVIC – Superintendência Nacional de Previdência Complementar

## LISTA DE FIGURAS

Figura 1: Sistema da Previdência Social no Brasil.....	9
Figura 2: Distribuição das classificações em Renda Variável.....	20
Figura 3: Histograma da variável Tempo_Fundo_mes .....	22
Figura4:Gráfico Scree Plot – Autovalores da matriz de correlação das 9 variáveis .....	26
Figura5:Gráfico Scree Plot – Autovalores da matriz de correlação das 8 variáveis .....	30
Figura 6: Histogramas dos Escores dos Fatores 1,2 e 3. ....	36
Figura 7:Box-plots dos fatores 1, 2 e 3.....	36
Figura 8: Histograma dos escores sem outliers .....	37

## LISTA DE TABELAS

Tabela 1- Distribuição de Frequência das observações da variável Tempo_Fundo_mês .....	22
--	----

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>7</b>
<b>2. PROBLEMATIZAÇÃO.....</b>	<b>10</b>
<b>3. JUSTIFICATIVA .....</b>	<b>11</b>
<b>4. OBJETIVO.....</b>	<b>12</b>
4.1    Objetivo Geral.....	12
4.2    Objetivos Específicos .....	12
<b>5. REVISÃO DA LITERATURA .....</b>	<b>12</b>
5.1    Fundos de Investimento em Renda Variável .....	12
5.2    Principais Indicadores de performance utilizados pelo Mercado para avaliação de desempenho .....	13
<b>6. METODOLOGIA.....</b>	<b>16</b>
6.1    Estatística Multivariada .....	17
6.2    Análise Fatorial.....	17
<b>7. BASE DE DADOS .....</b>	<b>19</b>
<b>8. ANÁLISE DESCRITIVA DOS DADOS.....</b>	<b>21</b>
<b>9. APLICAÇÃO DA ANÁLISE FATORIAL.....</b>	<b>24</b>
<b>10. CONSIDERAÇÕES FINAIS .....</b>	<b>43</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>45</b>

## 1. INTRODUÇÃO

A sociedade ao longo do tempo foi desenvolvendo sistemas de proteção para seus membros com o propósito de conceder-lhes condições mínimas de subsistência quando os mesmos, por motivos aleatórios perdem, provisória ou definitivamente a capacidade de se auto sustentar. Um dos mecanismos elaborados para a manutenção da capacidade de subsistência do cidadão é a Previdência.

Os sistemas previdenciários podem diferir de uma sociedade para outra, pois fatores de ordem política, econômica, social e cultural interferem na história de sua formação e desenvolvimento, mas em todas possuem uma função comum: assistir recursos financeiros a população adulta quando afastada de mercado de trabalho, por motivos alheios à sua vontade, como doença, invalidez e idade avançada. (BATICH; p. 33; 2004).

A constituição brasileira, promulgada em 1988, assegura a todo cidadão o direito a previdência, como está disposto no Capítulo II - Dos Direitos Sociais, artigo 6º: “São direitos sociais a educação, a saúde, o trabalho, o lazer, a segurança, a previdência social, a proteção à maternidade e à infância, a assistência aos desamparados, na forma desta Constituição”. (CONSTITUIÇÃO FEDERAL BRASILEIRA, 1988, p. 46).

Entretanto, a Previdência no Brasil tem seu início bem antes da Constituição Brasileira. Pode-se afirmar que os primeiros sinais da previdência no Brasil ocorreram a partir de 1888 com o decreto nº 9.912-A, de 26 de março de 1888, onde foi regulamentado o direito à aposentadoria dos funcionários dos Correios, fixando em 30 anos o tempo de efetivo serviço e idade mínima de 60 anos. Contudo, o decreto nº 4.682, conhecido como a Lei Eloy Chaves, de 24 de janeiro, que determinava a criação de uma Caixa de Aposentadoria e Pensões para empregados de cada empresa ferroviária é considerado como marco inicial da Previdência Social brasileira. Somente a partir de então, é que o Estado começou, gradativamente, a gerenciar o sistema previdenciário Brasileiro.

Outro ponto relevante para a história previdenciária foi a criação do Ministério do trabalho, Indústria e Comércio, em 26/11/1930, com funções de orientar e supervisionar a previdência Social, inclusive como órgão de recursos das decisões das Caixas de Aposentadorias e Pensões – CAPs.

Batich (2004, p. 34) afirma que as transformações econômicas da década de 30 que resultaram no processo de crescimento industrial levaram o Estado a “iniciar um processo de

interferência” nas relações de trabalho devido à presença das classes assalariadas que lutavam e pleiteavam melhores condições de trabalho. Com isso, o Estado expandiu sua interferência tornando-se responsável pela proteção social e determinou que as Caixas de Aposentadorias e Pensões “fossem substituídas por outro tipo de instituição”, instituição que seria formada por categorias de profissões de todo o território nacional.

A administração previdenciária deixou de ser então de responsabilidade de cada CAP, passando para a alçada do estado, que instituiu pela primeira vez na história do seu orçamento de custeio, os recursos necessários para desempenhar as novas tarefas, tornando necessária e legítima sua intervenção sobre os mecanismos de arrecadação e gestão das entidades previdenciárias. (BATICH; 2004; p. 34)

A primeira instituição a ser criada em 1933 foi o Instituto de Aposentadoria e Pensões dos Marítimos - IAPM e a partir dessa data vários outros institutos foram criados. Na mesma década criou-se o Instituto dos Comerciários (1934), o instituto dos bancários (1935), o instituto dos industriários (1938), o instituto dos empregados em transportes e cargas (1938).

Entretanto, no ano de 1960, após 14 anos de discussão, foi promulgada a Lei Orgânica da Previdência Social que instituiu um sistema único de previdência para todos os trabalhadores do setor privado e em 1966 consolidou-se a unificação do sistema com a criação do Instituto Nacional de Previdência Social – INPS.

No início da década de 90 foram extintos os antigos INPS e IAPAS, que deram lugar ao atual Instituto Nacional do Seguro Social (INSS), consolidando a Previdência com foco na prestação de benefícios monetários e dos serviços a eles associados. Nesse momento também foi extinto o INAMPS, que prestava a assistência médica previdenciária, sendo transferidas sua estrutura e suas tarefas para o Sistema Único de Saúde (SUS), que é custeado por impostos.

A Previdência Social no Brasil é garantida a todo cidadão e sua administração é feita pelo governo que tem a responsabilidade de gerir os recursos e garantir os benefícios a todos que atingem os requisitos de elegibilidade.

Art. 1º: A Previdência Social, mediante contribuição, tem por fim assegurar aos seus beneficiários meios indispensáveis de manutenção, por motivo de incapacidade, desemprego involuntário, idade avançada, tempo de serviço, encargos familiares e prisão ou morte daqueles de quem dependiam economicamente. (LEI Nº. 8.213; 1991)

Atualmente, o sistema previdenciário brasileiro está disposto da seguinte maneira:



**Figura 1: Sistema da Previdência Social no Brasil**

Fonte: Guia do Participante; Ministério da Previdência e Assistência Social; 2009

Entretanto, diante das limitações que a previdência social possui e a limitação de recebimento do benefício de aposentadoria, foi propiciado aos segurados da Previdência Social a opção de complementar ou suplementar o valor de seu benefício quando o mesmo atingir todos os requisitos de elegibilidade. Essa opção é denominada previdência complementar.

A Previdência complementar tem por objetivo dar condições aos participantes de possuir a mesma capacidade de pagamento que possuem atualmente, aos trabalhadores e aos familiares quando atingirem o período de gozo da aposentadoria

No curso de suas vidas as pessoas estão sujeitas a alguns infortúnios, tais como: a morte prematura, a invalidez, e ter de enfrentar a velhice em condição financeira desfavorável, comparativamente à usufruída durante o período laboral.

Sistemas públicos universais de previdência e assistência social oferecem proteção para essas e para outras situações, mas em condições que nem sempre suprem as reais necessidades dos indivíduos, em razão de limitações impostas pelos programas governamentais, em particular no tocante à percepção de valores financeiros.

Nesse cenário, são criadas, regulamentadas e prosperam as modalidades privadas de prevenção adicional (complementar) contra os impactos financeiros negativos relacionados ao enfrentamento desses infortúnios, ou seja, programas de adesão voluntária voltados ao atendimento das necessidades de pessoas previdentes, passíveis de serem constituídos e operacionalizados no segmento privado de previdência complementar. (FENAPREVI; 2017)

As Entidades Fechadas de Previdência Complementar - EFPC, conhecidas também como Fundos de Pensão tem por objetivo suplementar o benefício assegurado pela Previdência Social a fim de garantir aos seus beneficiários e a sua família a mesma capacidade econômica que possuía quando se encontrava no exercício de sua atividade profissional.

As EFPC's são criadas a partir da motivação do patrocinador ou instituidor e são restritas aos funcionários ou servidores dos patrocinadores e instituidores.

Recentes escândalos envolvendo os investimentos das Entidades Fechadas de Previdência Complementar, relacionados aos investimentos das mesmas reafirmam que a sociedade de maneira ampla precisa conhecer e entender os investimentos que os chamados fundos de pensão fazem para garantir o pagamento de uma renda complementar ao participante apto.

De acordo com a Superintendência de Previdência Complementar - PREVIC, os fundos de pensão podem investir até 10% do Patrimônio em renda variável. E esse investimento precisa ser pulverizado de forma que haja uma mitigação do risco de concentração. Para isso existe a necessidade de criação de uma carteira que pode ser entendido como conjunto de ativos, que podem ser ações ou fundos de investimentos, mantidos por um investidor.

## 2. PROBLEMATIZAÇÃO

As entidades de Previdência Complementar geralmente assumem a obrigação de garantir uma renda mensal aos seus participantes. Esse benefício por diversas vezes é calculado atuarialmente de acordo com algumas premissas atuariais.

Chan, Silva e Martins (2006) afirmam que, no Brasil, as premissas atuariais são classificadas em três grupos. As premissas econômicas, premissas biométricas e outras premissas. As premissas econômicas são: taxa de inflação de longo prazo, ganho real dos investimentos, escala de ganhos salariais, indexador dos benefícios, teto de benefício do sistema público e custeio administrativo.

Pode-se então notar que para cumprir com o benefício oferecido ao participante, as entidades precisam investir no mercado de capitais e obter ganho real nesses investimentos. e por esse motivo fazem parte dos Investidores institucionais.

Por serem investidores institucionais, os mesmos possuem algumas restrições no momento de aplicar os valores no mercado de capitais. Eles devem priorizar títulos públicos e papéis mais líquidos, priorizando as empresas que são consideradas *blue chips*<sup>1</sup>.

---

<sup>1</sup> O termo **blue chip** é internacionalmente empregado nos mercados de ações para referenciar ações de empresas conhecidas, de grande porte, nacional e internacional, com comprovada liquidez e lucratividade..

Porém, verifica-se que no Brasil há uma tendência de redução da taxa de juros, o que provoca uma necessidade de investimentos mais variados e uma gestão eficaz dos mesmos, o que provoca maior adesão ao mercado de renda variável.

Partindo do cenário da importância da Previdência Privada, principalmente pelo amplo debate acerca da situação previdenciária social brasileira e sua relevância para a economia nacional, e por existir legislações específicas para os fundos de pensão quanto a limites para aplicações nas diversas modalidades de investimentos, questiona-se como problema de pesquisa: como selecionar fundos de investimentos em renda variável?

### **3. JUSTIFICATIVA**

Devido à necessidade de captação de recursos, cada vez mais as entidades fechadas de Previdência Complementar começaram a investir no Mercado de Capitais como forma de garantir melhores retornos para seus ativos.

Mas esta tendência de atuação no mercado de capitais requer também maior preparação e conhecimento, pois mercado de capitais é bastante amplo e complexo. E devido à complexidade existente na economia nacional e internacional e por consequência no mercado de capitais além do fato de que as flutuações do mercado econômico-financeiro são praticamente instantâneas, as Entidades geralmente optam por operar no mercado de capitais através de fundos de investimento.

## **4. OBJETIVO**

### **4.1 Objetivo Geral**

Propor, através da aplicação de técnicas de Estatística Multivariada, um indicador auxiliar para seleção de fundos de investimento em renda variável dos Fundos de Pensão afim de otimizar o retorno do segmento em renda variável.

### **4.2 Objetivos Específicos**

- ✓ Verificar a evolução histórica das entidades fechadas de previdência complementar no Brasil;
- ✓ Construir um índice para seleção de fundos de investimento em Renda Variável de acordo com os objetivos dos investidores.

## **5. REVISÃO DA LITERATURA**

### **5.1 Fundos de Investimento em Renda Variável**

Os fundos de investimento podem ser entendidos como fundos de aplicações em títulos e valores mobiliários, que se encontra dividido em unidades de participações (cotas) que são adquiridas pelos participantes destes fundos, denominado investidores.

De acordo com a instrução da Comissão de Valores Mobiliários, CVM 555 de 14 de dezembro de 2014, em seu artigo 108, os fundos de investimento e os fundos de investimentos em cotas são classificados em Fundo de Renda Fixa; Fundo de Ações; Fundo Multimercado; e Fundo Cambial.

Os fundos de investimento em Renda Variável são fundos que possuem oscilação e risco considerável e que não tem qualquer indexador previamente definido.

## 5.2 Principais Indicadores de performance utilizados pelo Mercado para avaliação de desempenho

O acompanhamento da performance dos investimentos que uma pessoa possui ou que deseja possuir é uma das principais atribuições dos investidores, principalmente de investidores profissionais.

De acordo com a instrução CVM 554:

Investidores Profissionais: são (i) instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil; (ii) companhias seguradoras e sociedades de capitalização; (iii) entidades abertas e fechadas de previdência complementar; (iv) pessoas naturais ou jurídicas que possuam investimentos financeiros em valor superior a R\$ 10.000.000,00 e que, adicionalmente, atestem por escrito sua condição de investidor profissional Mediante termo próprio; (v) fundos de investimento; (vi) clubes de investimento, cuja carteira seja gerida por administrador de carteira autorizado pela CVM; (vii) agentes autônomos de investimento, administradores de carteira, analistas e consultores de valores mobiliários autorizados pela CVM, em relação a seus recursos próprios; e (viii) investidores não residentes. (CVM 554, 2014)

De acordo com Assaf Neto (2007), na avaliação de carteiras com elevado número de ativos, as fases de estudo podem ser definidas em três etapas: análise dos títulos, análise das carteiras e seleção da carteira.

Ao longo do tempo, surgiram vários indicadores de performance. Conforme VARGA, (2001), na avaliação de investimento, é muito comum a utilização das estatísticas dos modelos CAPM<sup>2</sup>, Teoria de Markowitz<sup>3</sup> e APT<sup>4</sup>. “Entre as estatísticas de avaliação de performance mais conhecidas estão o índice de Sharpe (IS), apresentado em Sharpe (1966), e o coeficiente Beta. A seguir descrevemos alguns dos índices mais comuns nessa área de estudo.

### (1) Coeficiente de Beta e o Índice de Sharpe

O cálculo do Beta é realizado a partir das oscilações do ativo

---

<sup>2</sup> CAPM se refere ao *capital asset pricing model*, desenvolvido por Sharpe (1964).

<sup>3</sup> Teoria de Markowitz se refere a seleção e avaliação de carteiras de investimento, basicamente demonstra como os investidores podem utilizar o princípio da diversificação para buscar melhorias na relação risco e retorno de suas carteiras de investimentos. (*Portfolio Selection*, 1952)

<sup>4</sup> APT se refere à *arbitrage princig theory*, desenvolvida por Ross (1976).

$$Beta = \frac{Cov(R_i, R_M)}{\sigma^2(R_M)}$$

$Cov(R_i, R_M)$  é a covariância entre os retornos do ativo  $i$  e da carteira de mercado ou *benchmark*;

$\sigma^2(R_M)$  é a variância da taxa de retorno do mercado

O coeficiente Beta então retrata como o ativo se posiciona no mercado. Quando o beta de um investimento for igual a 1, tem-se que o investimento se movimenta na mesma direção que a carteira de mercado. Se um ativo tem um beta de 1,3, uma valorização média do mercado de 10% determina uma expectativa de rentabilidade de 13% no ativo.

Índice de Sharpe.

O Índice de Sharpe tem como principal característica a possibilidade de mensurar o risco e o retorno do investimento. Ele mede qual é a relação entre o retorno excedente ao ativo livre de risco e a volatilidade, ou seja, é uma medida de avaliação da relação risco-retorno de larga aplicação.

A equação de IS é:

$$IS = \frac{m_j - m_b}{\sigma_j}$$

onde  $m_j$  é retorno esperado do fundo,  $m_b$  é o retorno de um ativo livre de risco ou de uma carteira de referência (*benchmark*) e  $\sigma_j$  é o desvio padrão ou volatilidade do retorno do fundo.

Temos então que, o índice reflete a relação entre o retorno e o risco, então, quanto maior o índice de Sharpe for, mais retorno ele agregará.

(2) Índice de Treynor

O Índice Treynor pode ser compreendido como a média do rendimento da carteira, deduzido a taxa livre de risco, sobre o seu respectivo beta. Trata-se então de um índice de

performance, que divide o retorno da carteira pelo seu respectivo beta em relação a um índice de referência sendo calculado como:

$$Treydor = \frac{\text{Retorno}}{\text{Risco}} = \frac{(\text{Retorno Medio} - \text{Retorno Medio do ativo livre de risco})}{\text{beta do ativo sendo estudado}}$$

O índice de Treynor é adequado para avaliar um ativo ou carteira em relação a uma referência, ou em relação a outra carteira maior já existente.

De acordo com ASSAF Neto (p. 245, 2007), quanto mais elevado apresentar-se o índice de Treynor, mais elevado é o retorno da carteira por unidade de risco assumido.

### (3) Índice de Sortino

O índice de Sortino pode ser compreendido como uma variação do índice de Sharpe. A medida de avaliação é identificada pela seguinte equação:

$$Sortino = \frac{\text{Retorno da carteira} - \text{Retorno mínimo aceitável}}{\text{medida de risco de não atingir o retorno mínimo aceitável}}$$

O índice de Sortino pode ser compreendido como uma comparação entre os retornos apresentados pela carteira e o retorno mínimo aceitável. Então, valores positivos significam que o ativo conseguiu uma performance melhor que o aceitável.

### (4) Alpha de Jensen

O Alpha de Jensen pode ser compreendido como uma medida do retorno comparativamente aos seus valores esperados. Sua equação é:

$$\text{Alpha de Jensen} = (RM_x - RM_{RF}) - \beta * (RM_{BM} - RM_{RF})$$

onde:

$RM_x$  é o retorno médio do investimento sendo estudado;

$RM_{RF}$  é o retorno médio do investimento livre de risco ( $RF = risk\ free$ );

$RM_{BM}$  é o retorno médio do benchmark usado;

$\beta$  é a correlação entre taxa de rentabilidade do fundo e a taxa de rentabilidade do benchmark vezes o desvio padrão da taxa de rentabilidade do fundo dividida taxa de rentabilidade do benchmark.

O índice *alpha de Jensen* efetua uma comparação entre os retornos apresentados por um ativo esperados. Então, valores positivos significam que o ativo conseguiu uma performance melhor que o esperado.

#### (5) VaR – *Value at Risk*

Pode-se entender o *Value at Risk*, ou valor em risco, como sendo o valor monetário das perdas no valor atual a que uma carteira está sujeita, a um determinado nível de confiança e dentro de um horizonte de tempo. O VaR sintetiza a perda máxima esperada, medida em valores monetários, dentro de determinado intervalo de tempo e dada uma probabilidade de ocorrência.

$$P(\Delta X_t \leq VaR) = \alpha$$

onde  $\alpha$  é o nível de significância (ou  $(1 - \alpha)$  é o nível de confiança) adotado,  $0 < \alpha < 1$ ,  $\Delta X_t$  é a variação no valor da carteira a preço  $X_t$  e VaR é o valor em risco para o horizonte de tempo  $t$ .

A notação  $P(\cdot)$  indica a probabilidade de um evento.

Considerando que o Var é uma medida para a perda, então quanto menor, melhor.

## 6. METODOLOGIA

Para Cervo, Bervian (2002, p. 23), metodologia é o processo pelo qual o pesquisador procura descrever os processos necessários para chegar a resposta das questões abordadas na proposta de pesquisa. Para que um pesquisador desenvolva um bom trabalho é necessário que ele seja baseado em outros estudos, pois “a época do empirismo passou”. A metodologia fornece ferramentas que permitem realizar pesquisas tendo outros trabalhos como referência, onde podemos concordar ou discordar, mas sempre confrontando as ideias, afim de se chegar a uma nova questão.

## 6.1 Estatística Multivariada

As técnicas de Estatística Multivariada podem ser utilizadas na coleta e análise de dados quando se dispõe de várias variáveis medidas simultaneamente e correlacionadas entre si.

Sua aplicação ocorre em diversas áreas de conhecimento, possibilitando analisar tanto os dados quantitativos como dados qualitativos. Existem várias técnicas de Estatística Multivariada sendo que algumas se prestam a construção de índices que são combinações lineares das variáveis originais. Nessa monografia utilizou-se a Análise Fatorial que será descrita a seguir.

## 6.2 Análise Fatorial

O método de análise fatorial tem como objetivo explicar a variabilidade total de um conjunto de variáveis originais em termos de um conjunto limitado de variáveis aleatórias não observáveis, inicialmente. Trata-se, desta forma, de uma técnica para reduzir o número de variáveis em análise.

Segundo Corrar et al. (2007 p. 02) a análise multivariada é um conglomerado de métodos estatísticos que viabilizam a análise conjunta de várias variáveis de forma simultânea.

A análise multivariada incluirá tanto as técnicas de múltiplas variáveis quanto técnicas verdadeiramente multivariadas, uma vez que acreditamos que o conhecimento de técnicas de múltiplas variáveis é um primeiro passo essencial na compreensão da análise multivariada. (HAIR et al, 2009, p.27)

Para tanto, são identificados os padrões de correlações ou de covariâncias entre as variáveis originais, reunindo-as em agrupamentos distintos, que passam a ser as novas variáveis de análise. A análise fatorial é um método muito útil e tem sido aplicado em diversas áreas como, em estudos aplicados de ciências sociais ou em pesquisas que utilizam a aplicação de questionários para diversos fins, bem como em análises de mercado. Também é típica a aplicação desta metodologia em estudos promovidos por empresas para identificar os perfis de clientes com os quais se relaciona.

No modelo de análise fatorial ortogonal, os grupos, chamados de Fatores, são compostos por variáveis correlacionadas entre si, e não correlacionadas com variáveis que fazem parte de outros fatores. Cada fator gera uma combinação linear das variáveis originais e seus valores

numéricos (chamados de Escores) podem ser usados para as análises estatística de interesse do pesquisador.

As equações do modelo de análise fatorial realizada pela matriz de correlação das variáveis originais, são dadas a seguir:

$$\begin{aligned} Z_1 &= l_{11}F_1 + l_{12}F_2 + \cdots + l_{1m}F_m + \varepsilon_1 \\ Z_2 &= l_{21}F_1 + l_{22}F_2 + \cdots + l_{2m}F_m + \varepsilon_2 \\ &\quad \vdots \quad \quad \quad \vdots \\ Z_p &= l_{p1}F_1 + l_{p2}F_2 + \cdots + l_{pm}F_m + \varepsilon_p \end{aligned}$$

Neste sentido, os fatores ( $F_j$ ) se portam como variáveis explicativas de um modelo de regressão linear e os coeficientes ( $l_{ij}$ ) como seus coeficientes. Esses coeficientes são chamados de cargas fatoriais e representam a correlação da variável  $Z_i$  com o correspondente fator aleatório  $F_j$ ,  $m$  é o número de fatores.

Os erros ( $\varepsilon_j$ ) podem ser entendidos como a parte da variável resposta que não explicada pelos fatores,  $i=1,2,\dots,p$ .

No entanto, a princípio, o número de fatores e suas composições, são desconhecidos e precisam ser de algum modo identificados na análise estatística dos dados.

É importante destacar que para se operacionalizar a estimação do modelo são necessárias algumas suposições. São elas:

- Todos os fatores devem possuir média igual a zero;
- Todos os fatores devem ser não correlacionados e ter variâncias iguais a 1;
- Todos os erros do modelo devem ter médias iguais a zero;
- Todos os erros devem ser não correlacionados entre si, mas sem a necessidade de ter a mesma variância;
- Os vetores dos Erros e dos Fatores são independentes.

É importante ressaltar que a Análise Fatorial é um método que busca explicar com o menor número possível de novas variáveis (fatores), a maior parte da variabilidade total das variáveis aleatórias que fazem parte do estudo, restando ainda o erro, que corresponde à variabilidade referente aos fatores que não foram incluídos no modelo.

Chamamos de comunalidade a parcela da variabilidade da variável original  $Z_i$  que é devida a inclusão dos “m” fatores no modelo, enquanto a variância específica corresponde a parcela devida ao erro ( $\varepsilon_i$ ). Estes são dois conceitos importantes para a avaliação da qualidade do ajuste do modelo. Quanto maior os valores de comunalidades (e consequentemente menores os de variância específica) melhor será a qualidade de ajuste do modelo de análise fatorial.

## 7. BASE DE DADOS

A base de dados em estudo utilizada no desenvolvimento desta monografia, tem como data de referência o mês de agosto de 2016. Trata-se apenas de fundos de renda variável e com classificação da Associação brasileira das Entidades dos Mercados Financeiro e de Capitais - ANBIMA (2016).

A classificação técnica criada pela ANBIMA em 2015 busca agregar as diversas opções de carteira oferecidas via fundos de investimento, conforme os atributos do portfólio. Assim, ela leva em consideração as classes de ativos que compõem essa carteira, estratégias de gestão, riscos e prazos. Ao dividir os fundos em tipos ou categorias que expressam objetivos e políticas de investimento similares, a classificação viabiliza a comparabilidade entre os fundos e desses com outras opções de investimento ou *benchmarks*. Além disso, auxilia no processo de seleção e decisão de investimento. Assim, a classificação possibilita a comparação entre fundos ou mesmo entre esses e outras opções de investimento.

Dentro do seguimento de fundos de renda variável em ações, temos 03 níveis, sendo o Nível 1 a modalidade de classificação de acordo com a CVM.

Art. 108. Quanto à composição de sua carteira, os fundos de investimento e os fundos de investimento em cotas, classificam-se em:

I – Fundo de Renda Fixa;

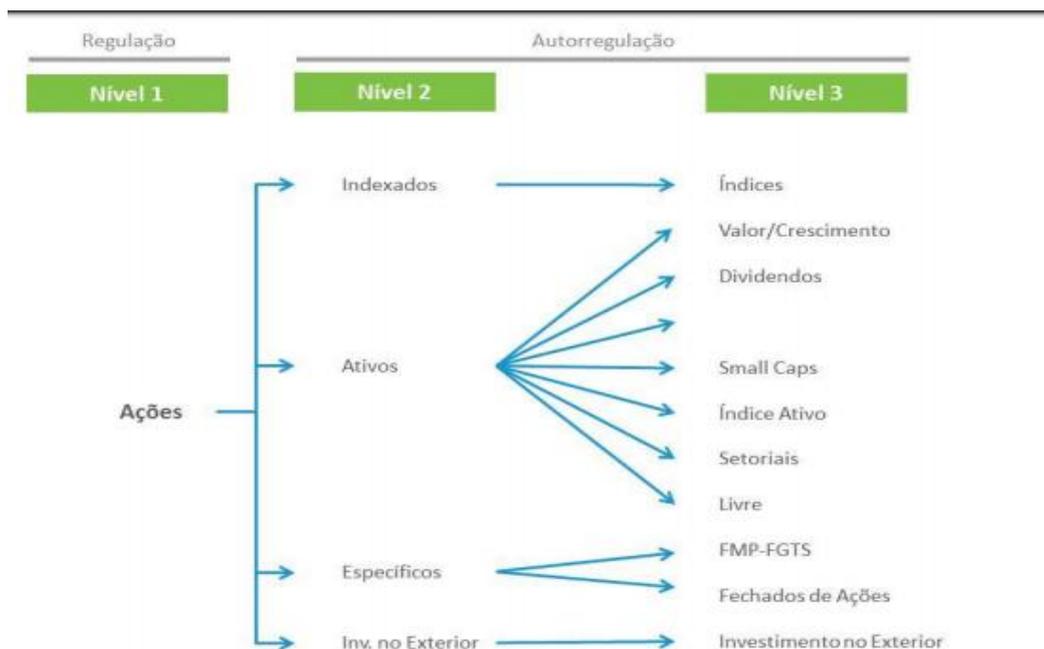
II – Fundo de Ações;

III – Fundo Multimercado; e

IV – Fundo Cambial.

(Instrução CVM Nº 555, DE 17 DE DEZEMBRO DE 2014)

E o próprio mercado, através da ANBIMA, autorregulou-se e desenvolveu os níveis 2 e 3 conforme Figura 2.



**Figura 2: Distribuição das classificações em Renda Variável.**  
**Fonte: ANBIMA; 2015**

Foram utilizadas 1642 observações (fundos de renda variável), para as variáveis descritas no Quadro 1.

**Quadro 1: Descrição das variáveis utilizadas**

VARIÁVEL	DESCRIÇÃO
Nome	Indica o nome registrado junto a CVM para o fundo de investimento
Classificação ANBIMA	agrupa fundos de investimento com as mesmas características, identificando-os pelas suas estratégias e fatores de risco
Empresa Gestora	Pessoa jurídica devidamente registrada como administrador de carteira de valores mobiliários junto a CVM para a prestação de serviços, em caráter profissional, discricionário ou não, de assessoria na escolha para aplicação de recursos em Patrimônio
Tempo_Fundo_mes	Tempo de atividade do fundo de investimento, em mês
Sharpe_310816_1ano6_1ano	Índice de Sharpe de 1 ano posicionado em 31 de agosto de 2016
Beta0816_60meses	Coefficiente Beta do Fundo, em agosto de 2016, utilizando 60 meses anteriores
SortinoMAR=RF_310816-1ano0816-1ano	Índice Sortino de 1 ano posicionado em 31 de agosto de 2016
Varição_Cta_X-1(1s)	Varição da cota entre os períodos de fevereiro e agosto do ano de 2015
Varição_Cta_X-1(2s)/X(1s)	Varição da cota entre o agosto de 2015 e fevereiro de 2016
Varição_Cta_X	Varição semestral da cota entre os períodos de fevereiro e agosto de 2016
Sharpe_290216-1ano	Índice de Sharpe de 1 ano posicionado em 29 de fevereiro de 2016.
VAR 95%_1dia_1ano	VaR de 1 dia, com 95% de confiança. O <i>Value at Risk</i> , com horizonte de 1 ano
VAR 95%_1dia_2anos	VaR de 1 dia, com 95% de confiança. O <i>Value at Risk</i> , com horizonte de 2 anos

**Fonte:** Elaborada pela autora

As observações, conforme segmentação AMBIMA, foram agrupadas conforme Quadro 2.

**Quadro 2: Fundos de Renda Variável, por segmento**

<b>Classificação Observações por segmento</b>	<b>Quantidade</b>
Ações Dividendos	76
Ações FMP – FGTS	79
Ações IBOVESPA Ativo	1
Ações Indexados	51
Ações Índice Ativo	247
Ações Invest. no Exterior	83
Ações Livre	769
Ações Setoriais	40
Ações <i>Small Caps</i>	38
Ações Sustentabilidade/Governança	27
Ações Valor/Crescimento	125
Fechados de Ações	26
Fundos de Mono Ação	47
Off Shore Renda Variável	33
<b>Total Geral</b>	<b>1642</b>

Na fase de análise de consistência da base, desprezou-se os fundos com idade inferior a 1,5 anos de existência ou aqueles que não possuíam informações divulgadas para qualquer variável da base. Do total de observações existentes, 1840, desprezou-se então 198 fundos que não tinham as informações para as variáveis em estudo, perfazendo então uma utilização de 89,24% de todos os fundos relacionados na AMBIMA em agosto/2016.

Diante do exposto, torna-se oportuno a utilização da metodologia de análise multivariada para se analisar os dados obtidos no presente estudo.

## **8. ANÁLISE DESCRITIVA DOS DADOS**

A estatística descritiva tem como principal característica a sumarização dos dados. Uma forma de apresentar os dados é através da tabela de distribuição de frequências na qual os dados são agrupados em classes para melhor visualização.

Na Tabela 1 tem-se a distribuição de frequências (absoluta e acumulada) da variável Tempo\_Fundo\_mes. Essa variável refere-se ao tempo de existência mensal do fundo de investimento em Renda variável na data de 31 de agosto de 2016.

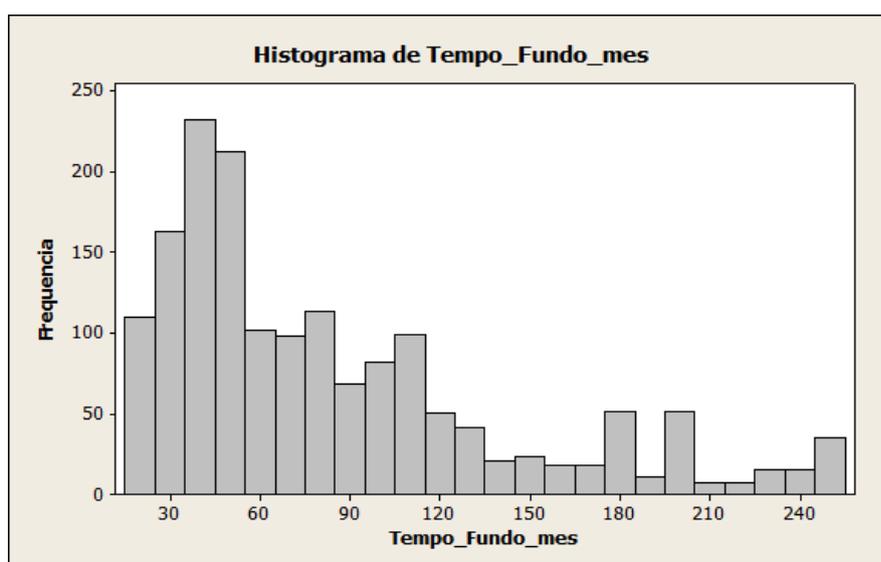
**Tabela 1- Distribuição de Frequência das observações da variável Tempo\_Fundo\_mês**

Intervalo de Classes (meses)	Frequência (Fi)	Fi%	Frequência acumulada (FAC)	FAC%
19  -- 48	610	37,15	610	37,15
48  -- 77	343	20,89	953	58,04
77  -- 106	245	14,92	1198	72,96
106  -- 135	173	10,54	1371	83,5
135  -- 164	61	3,71	1432	87,21
164  -- 193	79	4,81	1511	92,02
193  -- 222	65	3,96	1576	95,98
222  -- 251	66	4,02	1642	100

Tamanho da amostra=N=1642

Na Tabela I observa-se que 37,17% dos fundos tem até 48 meses (intervalo aberto) de lançamento, o que evidencia o aumento das opções de investimento em renda variável nos últimos 4 anos.

O histograma da Figura 3 mostra a distribuição de frequência de fundos de acordo com as classes. As classes iniciais, até 59 meses possui a maior quantidade de fundos.

**Figura 3: Histograma da variável Tempo\_Fundo\_mes**

Como medidas de dispersão importantes tem-se: desvio padrão e coeficiente de variação. Já como medida de posição tem-se a média, mínimo, máximo e mediana. As medidas descritivas são apresentadas no Quadro 3. O tempo de existência de fundo tem por média 83,35

meses, com mínimo de 19 meses. A variável VAR 95%\_1dia\_2anos é a variável com desvio padrão e coeficiente de variação mais elevados.

**Quadro 3: Estatísticas descritivas – Base de dados Agosto 2016**

Variáveis	Média	Desvio Padrão	CoefVaria.	Mínimo	Mediana	Máximo
Tempo_Fundo_mes	83,35	57,47	68,96	19,0	65,02	251,00
Sharpe_310816_1ano	0,65	1,556	240,06	0,0	0,50	60,26
Beta0816_60meses	0,43	0,497	114,95	0,0	0,40	2,26
SortinoMAR=RF_310816-1ano	1,21	5,801	478,20	0,0	0,83	232,59
Variação_Cta_X-1(1s)	-0,04	0,198	-446,07	-1,0	-0,06	6,55
Variação_Cta_X-1(2s)/X(1s)	-0,06	0,138	-225,25	-1,0	-0,04	0,95
Variação_Cta_X	0,30	0,845	280,52	-1,0	0,26	31,65
Sharpe_290216-1ano	1,06	4,37	412,02	0,0	1,01	77,19
VAR 95%_1dia_1ano	2,43	1,92	79,03	0,0	2,04	42,87
VAR 95%_1dia_2anos	7,99	228,00	2853,83	0,0	2,02	9240,85

N=1642 para todas as variáveis

O coeficiente de correlação de Pearson (1925) é um indicador da força de uma relação linear entre duas variáveis intervalares, sendo definido o intervalo de -1 a 1. No Quadro 4, são apresentadas as correlações entre as variáveis analisadas nessa monografia.

**Quadro 4. Correlação de Pearson entre as variáveis**

	Tempo_Fundo_mes	Sharpe_310816_1ano	Beta0816_60meses	SortinoMAR=RF_310816-1ano	Variação_Cta_X-1(1s)	Variação_Cta_X-1(2s)/X(1s)	Variação_Cta_X	Sharpe_290216-1ano	VAR 95%_1dia_1ano	VAR 95%_1dia_2anos
Sharpe_310816_1ano	-0,069									
Beta0816_60meses	<b>0,709</b>	0,041								
SortinoMAR=RF_310816-1ano	-0,03	<b>0,982</b>	0,07							
Variação_Cta_X-1(1s)	-0,076	-0,077	-0,078	-0,102						
Variação_Cta_X-1(2s)/X(1s)	-0,303	0,005	-0,45	-0,056	0,105					
Variação_Cta_X	0,053	0,239	0,164	0,243	-0,067	-0,301				
Sharpe_290216-1ano	-0,024	-0,007	-0,028	-0,006	<b>0,802</b>	0,016	-0,012			
VAR 95%_1dia_1ano	0,283	0,473	<b>0,561</b>	<b>0,504</b>	-0,105	-0,452	0,247	-0,018		
VAR 95%_1dia_2anos	-0,016	-0,007	-0,017	-0,002	-0,011	-0,023	-0,036	0,008	-0,024	

De acordo com o coeficiente de correlação de Pearson, observamos que as variáveis Tempo\_Fundo\_Mes Beta0816\_60m são bem correlacionadas, assim como as variáveis

Variação\_Cta\_x-1(1s) comSharpe\_290216-1ano e Sharpe\_310816\_1ano com SortinoMAR=RF\_310816-1 ano (correlações acima de 0,70). Já a variável VAR 95%\_1dia\_1ano é moderadamente correlacionada com Sharpe\_310816\_1ano, Beta0816\_60meses e SortinoMAR=RF\_310816-1 ano.

Quando observamos a correlação da variável VAR 95%\_1dia\_2anos vemos que é muito baixa com todas as outras variáveis e por este motivo a mesma será analisada separadamente, ou seja, não fará parte da análise fatorial. A variável VAR 95%\_1dia\_2anos sofre grande influência da instabilidade política ocorrida no Brasil no período do afastamento e posteriormente o *Impeachment* da então Presidente da República, Dilma Rousseff ocorrido em agosto/2016.

## 9. APLICAÇÃO DA ANÁLISE FATORIAL

De modo simplificado, pode-se dizer que a aplicação da análise fatorial na base de dados escolhida para desenvolvimento deste trabalho percorreu os passos descritos a seguir.

### **Passo 1. Estimação do Número de Fatores (m)**

A avaliação preliminar da existência de correlação entre as variáveis visa possibilitar uma identificação de possíveis associações entre cada par de variáveis, e também a previsão dos possíveis grupos, no caso em que o número de variáveis não é muito grande. Se as correlações apresentadas entre todas as variáveis originais forem baixas, este é um indicativo de que o ajuste do modelo fatorial não será adequado.

A estimação do número (**m**) de fatores pode ser realizada pelo método de análise de componentes principais (PEARSON, 1901) citado por MINGOTI (2005). Nesse método, os autovalores são extraídos da matriz de correlações sendo **m** estimado pelo número de autovalores que explicam a maior parte da variabilidade total das variáveis. A variância total é a soma das variâncias de cada variável. No caso em que essas estão padronizadas, a variância total é igual ao número de variáveis que fazem parte da análise. Cada autovalor representa uma parte dessa variância. Sendo assim, a razão entre o valor numérico do autovalor e a variância total, representa a importância do autovalor e da dimensão (componente principal) que será

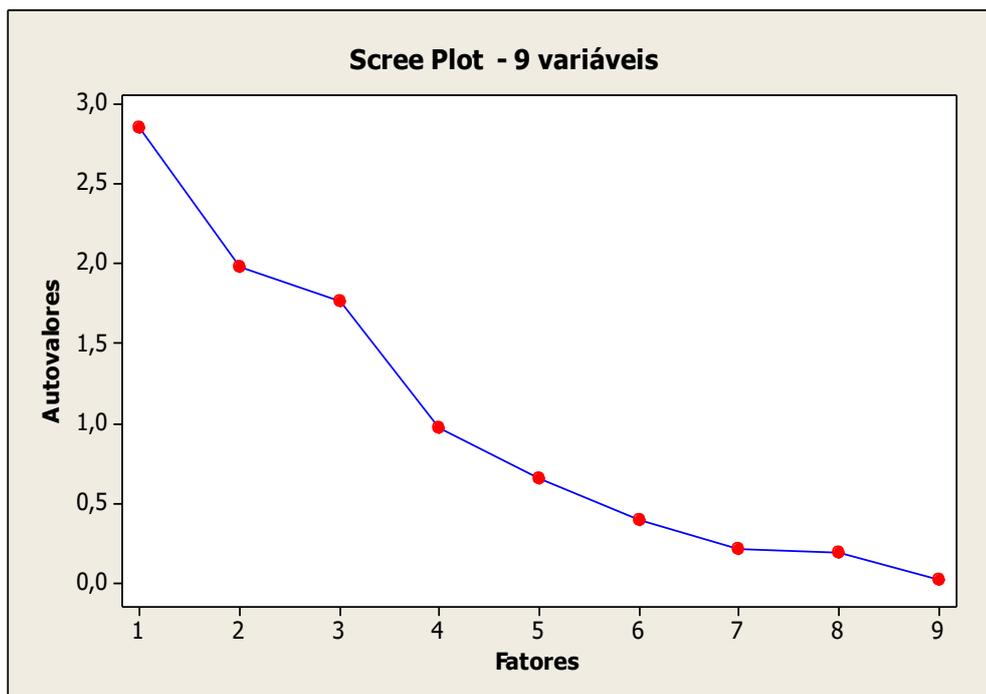
obtida a partir dele, em termos de variância total. Em geral, essa proporção é multiplicada por 100 tendo-se então o percentual de explicação.

A escolha pode ser facilitada por alguns critérios preestabelecidos, como a regra de Kaiser (1958), citado por MINGOTI (2005) que orientam a escolha a partir do número de fatores com autovalores maior que “1” e “0.7”, respectivamente, e também pela observação do gráfico “*ScreePlot* (Cattel, 1966), que apresenta uma curva dos autovalores com sua ordem de importância, de modo que a escolha geralmente é definida por um ponto de inflexão da curva, que demarca os autovalores consideravelmente mais elevados dos demais.

No Quadro 5 apresentam-se os autovalores da matriz de correlação com os respectivos percentuais de variância explicada e na Figura 4 *Scree Plot* correspondente, considerando-se as 9 variáveis de análise. De acordo com a regra de Kaiser (1958), citado por MINGOTI (2005) estima-se o valor de **m** como o número de autovalores que são maiores que 1. No Quadro 5 tem-se 3 autovalores maiores que 1, no entanto o 4º. autovalor está muito próximo de 1 e acima de 0,7. Dessa forma, serão ajustados os modelos com **m=4** e **m=3** para avaliação.

**Quadro 5-Autovalores da matriz de correlação das 9 variáveis**

Identificação	Autovalores	Proporção Explicada	Proporção Acumulada
1	2,847	0,316	0,316
2	1,9733	0,219	0,536
3	1,7645	0,196	0,732
4	0,969	0,108	0,839
5	0,649	0,072	0,911
6	0,3925	0,044	0,955
7	0,2046	0,023	0,978
8	0,1855	0,021	0,998
9	0,0145	0,002	1



**Figura 4: Gráfico *Scree Plot* – Autovalores da matriz de correlação das 9 variáveis**

### **Passo 2. Estimação das Cargas Fatoriais**

A definição do melhor procedimento de estimação das cargas fatoriais depende da base de dados que se está utilizando, sendo o de análise de componentes principais um dos mais utilizados por não depender intrinsecamente da distribuição dos dados ao contrário do método de máxima verossimilhança que tem como premissa que o vetor de variáveis aleatórias tenha distribuição normal multivariada. Nesta monografia, utilizou-se o método de Componentes Principais que foi introduzida por Karl Pearson em 1901 e tem como principal objetivo “explicar a estrutura de variância e covariância de um vetor aleatório, composto de  $p$ -variáveis aleatórias através da construção de combinações lineares das variáveis originais” (MINGOTI, 2005, p.59).

Nesse método, a estimativa da carga fatorial  $l_{ij}$  é a correlação da variável  $Z_i$  com a componente principal de ordem  $j$ , ou seja a correlação da combinação linear gerada pelo  $j$ -ésimo autovetor que é correspondente ao  $j$ -ésimo autovalor,  $j=1,2,\dots, 9$ .

### Passo 3. Identificação da Composição dos Fatores e Análise de Comunalidades

A partir dos resultados do passo 2, parte-se para identificação das variáveis que fazem parte de cada fator, interpretação dos fatores, e avaliação da qualidade de ajuste do modelo de análise fatorial ajustado aos dados, ou seja, a análise do percentual de variância total explicada, das comunalidades e variâncias específicas de cada variável (HAIR, et al., 2009).

Nessa fase se a composição dos fatores e as comunalidades forem adequadas a solução pode ser considerada como final. Caso haja problemas na identificação dos fatores, ou seja, correlação relevante de alguma variável com mais de um fator, o que fere a suposição de ortogonalidade dos fatores, deve-se usar o recurso da rotação ortogonal dos fatores na tentativa de melhorar a visualização da composição desses. Existem vários métodos para realizar a rotação ortogonal sendo que dentre eles o Varimax (Kaiser, 1958) citado por MINGOTI (2005) é um dos mais utilizados. Para um valor  $m$  fixo de número de fatores, a rotação ortogonal não altera os valores de comunalidades e variância total explicada, obtidas no modelo sem rotação.

A partir dos fatores construídos é possível então gerar uma combinação linear das variáveis para cada grupo (Fator) e obter os escores (valores numéricos desses) para cada elemento da amostra e analisá-los de acordo com os objetivos do estudo.

Matematicamente, a equação para o cálculo do escore do fator  $F_j$  para o  $k$ -ésimo elemento amostral, é dada por:

$$\hat{F}_{jk} = w_{j1}Z_{1k} + w_{j2}Z_{2k} + \dots + w_{jp}Z_{pk}$$

sendo  $Z_{ik}$  o valor observado da variável padronizada  $Z_i$  para o  $k$ -ésimo elemento amostral e  $w_{ji}$  o peso de ponderação de cada variável no fator  $F_j$ .

De acordo com HAIR et al. (2009 p.101), o escore fatorial é uma medida composta construída para cada observação de cada fator obtido na análise fatorial e o escore de cada observação é obtido através da conjunção dos pesos fatoriais e os valores da variável original. Os escores fatoriais são padronizados para que tenham média 0 e desvio-padrão 1

Para a estimação dos pesos existem diversos métodos estatísticos, mas um dos métodos mais comuns é a regressão linear (MINGOTI, 2005).

### 9.1. Ajuste do Modelo de Análise Fatorial: Cargas Fatoriais e Comunalidades

Considerando-se as 9 variáveis, as estimativas das cargas fatoriais e comunalidades do modelo com  $m=4$  e sem e com a rotação ortogonal Varimax, estão apresentadas no Quadro 6.

**Quadro 6- Análise Fatorial – matriz de correlação –  $m=4$  –  $p=9$**

<b>Análise Fatorial via Matriz de Correlação: Componentes Principais</b>					
Cargas Fatoriais dos fatores não rotacionados e Comunalidades					
Variáveis	Fator1	Fator2	Fator3	Fator4	Comunalidade
Tempo_Fundo_mes	0,464	0,641	-0,132	-0,345	0,763
Sharpe_310816_1ano	0,644	-0,727	-0,020	-0,164	0,970
Beta0816_60meses	0,662	0,586	-0,170	-0,211	0,856
SortinoMAR=RF_310816-1 ano	0,681	-0,691	-0,019	-0,156	0,966
Variação_Cta_X-1(1s)	-0,274	-0,103	-0,904	-0,003	0,904
Variação_Cta_X-1(2s)/X(1s)	-0,553	-0,427	0,108	-0,415	0,672
Variação_Cta_X	0,458	-0,075	-0,049	0,759	0,794
Sharpe_290216-1ano	-0,156	-0,114	-0,932	0,019	0,905
VAR 95%_1dia_1ano	0,837	0,006	-0,133	-0,075	0,723
Variância	2,847	1,973	1,7645	0,9690	7,554
% Var	0,316	0,219	0,196	0,108	0,839
Cargas Fatoriais dos fatores com rotação Varimax e Comunalidades					
Tempo_Fundo_mes	-0,073	-0,869	-0,030	0,048	0,763
Sharpe_310816_1ano	0,983	0,039	-0,019	-0,053	0,970
Beta0816_60meses	0,066	-0,910	-0,021	-0,153	0,856
SortinoMAR=RF_310816-1 ano	0,979	-0,004	-0,030	-0,078	0,966
Variação_Cta_X-1(1s)	-0,066	0,060	0,945	0,057	0,904
Variação_Cta_X-1(2s)/X(1s)	0,036	0,494	0,037	0,652	0,672
Variação_Cta_X	0,181	0,042	-0,017	-0,871	0,794
Sharpe_290216-1ano	0,016	-0,002	0,951	-0,013	0,905
VAR 95%_1dia_1ano	0,566	-0,565	-0,023	-0,290	0,723
Variância	2,2933	2,1517	1,8025	1,3063	7,554
% Var	0,255	0,239	0,200	0,145	0,839

Como observamos no Quadro 6, algumas variáveis não ficaram bem alocadas em um único fator. No modelo fatorial ortogonal, cargas fatoriais não podem ter grandezas numéricas relevantes próximos em mais de um fator e isso ocorreu em mais de uma variável na solução sem rotação Varimax. Na solução rotacionada esse problema aparece em maior impacto na variável VAR95%\_1dia\_1ano e em menor impacto na variável Variação\_Cta\_X-1(2s)/X(1s). É possível notar que a variável Variação\_Cta\_X é a de maior impacto no Fator 4. Considerando a solução sem rotação esse Fator fica composto apenas por essa variável. Na análise fatorial não é desejável construir fatores compostos por uma única variável já que nesse caso, o mais adequado é que a variável pertinente seja avaliada separadamente. Por essas razões, ajustou-se o modelo com  $m=3$  para comparação. Após o ajuste do modelo com  $m=3$ , algumas variáveis

ainda não ficaram bem alocadas nos fatores, mesmo após a aplicação da rotação ortogonal Varimax. Observou-se ainda que a comunalidade da variável Variação\_Cta\_X foi bem baixa, 0,218, o que induz que a mesma tem pouca participação no modelo e para  $m=3$  não se alocou a qualquer um dos fatores construídos.

**Quadro 7: Análise Fatorial – matriz de correlação –  $m=3$  –  $p=9$**

Cargas Fatoriais dos fatores não rotacionados e Comunalidades				
Variáveis	Fator1	Fator2	Fator3	Comunalidade
Tempo_Fundo_mes	0,464	0,641	-0,132	0,644
Sharpe_310816_lano	0,644	-0,727	-0,020	0,943
Beta0816_60meses	0,662	0,586	-0,170	0,811
SortinoMAR=RF_310816-1 ano	0,681	-0,691	-0,019	0,942
Variação_Cta_X-1(1s)	-0,274	-0,103	-0,904	0,904
Variação_Cta_X-1(2s)/X(1s)	-0,553	-0,427	0,108	0,500
<b>Variação_Cta_X</b>	<b>0,458</b>	<b>-0,075</b>	<b>-0,049</b>	<b>0,218</b>
Sharpe_290216-lano	-0,156	-0,114	-0,932	0,905
VAR 95%_1dia_lano	0,837	0,006	-0,133	0,718
Variância	2,8470	1,9733	1,7645	6,5848
% Var	0,316	0,219	0,196	0,732
Cargas Fatoriais dos fatores rotação Varimax e Comunalidades				
Variáveis	Fator1	Fator2	Fator3	Comunalidade
Tempo_Fundo_mes	-0,100	0,796	0,024	0,644
Sharpe_310816_lano	0,967	-0,086	0,017	0,943
Beta0816_60meses	0,083	0,897	0,016	0,811
SortinoMAR=RF_310816-1 ano	0,969	-0,036	0,028	0,942
Variação_Cta_X-1(1s)	-0,077	-0,075	-0,945	0,904
Variação_Cta_X-1(2s)/X(1s)	-0,111	-0,697	-0,040	0,500
<b>Variação_Cta_X</b>	<b>0,385</b>	<b>0,262</b>	<b>0,025</b>	<b>0,218</b>
Sharpe_290216-lano	0,018	0,002	-0,951	0,905
VAR 95%_1dia_lano	0,607	0,590	0,020	0,718
Variância	2,4280	2,3552	1,8016	6,5848
% Var	0,270	0,262	0,200	0,732

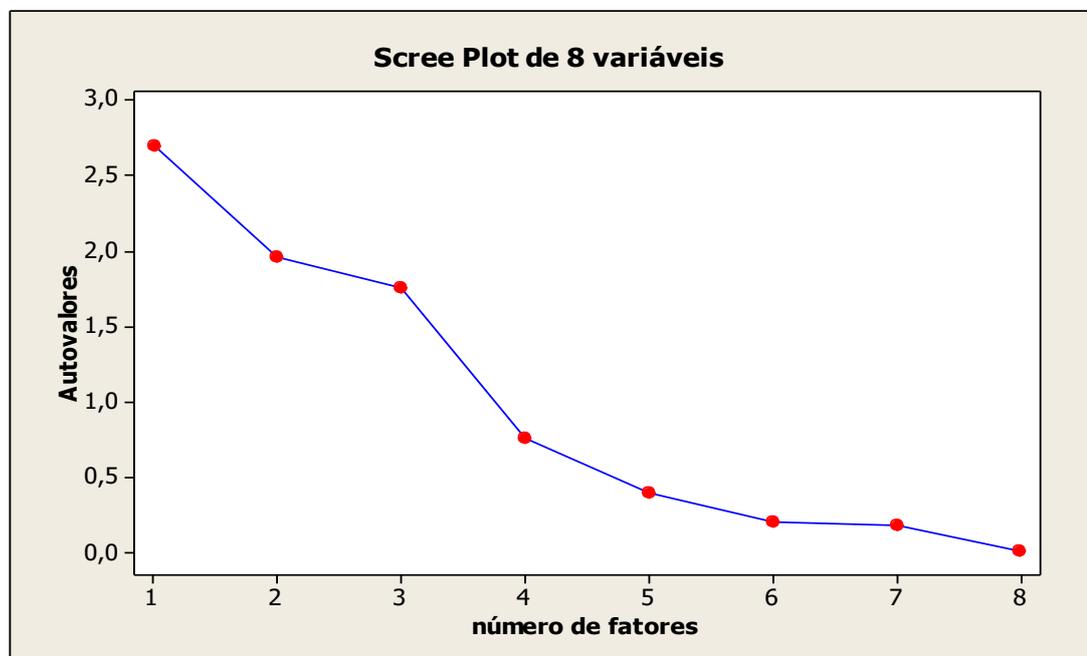
Por este motivo os modelos com  $m=4$  e  $m=3$  sem essa mesma variável foram considerados para avaliação em termos de ajuste e consistência.

Essas novas soluções estão apresentadas nos Quadros 8 e 9 sendo o *Scree-Plot* correspondente a  $p=8$  variáveis apresentado na Figura 5.

Como observamos no Quadro 8, ao rotacionarmos o modelo com 8 variáveis e  $m=4$ , temos problema com a identificação do fator que a variável Var 95%\_1dia\_lano se aloca. Por este motivo aplicamos a técnica para as 8 variáveis com  $m=3$ . Conforme a regra de Kaiser (1958) citado por MINGOTI (2005) assume-se o valor de  $m$  como o número de autovalores que são maiores que 1, por este motivo, tem-se  $m=3$  (Ver Figura 5).

**Quadro 8: Análise Fatorial – matriz de correlação – m=4 – p=8**

Cargas Fatoriais dos fatores não rotacionados e Comunalidades					
Variáveis	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidade
Tempo_Fundo_mes	0,502	0,622	-0,136	-0,476	0,885
Sharpe_310816_lano	0,632	-0,748	-0,039	-0,110	0,973
Beta0816_60meses	0,690	0,565	-0,176	-0,206	0,869
SortinoMAR=RF_310816-1 ano	0,671	-0,714	-0,039	-0,080	0,968
Variação_Cta_X-1(1s)	-0,290	-0,085	-0,901	-0,015	0,904
Variação_Cta_X-1(2s)/X(1s)	-0,538	-0,420	0,105	-0,662	0,914
Sharpe_290216-1ano	-0,174	-0,098	-0,930	0,028	0,906
VAR 95%_ldia_lano	0,848	-0,021	-0,148	0,176	0,773
Variância	2,7017	1,9695	1,7631	0,7579	7,192
% Var	0,338	0,246	0,220	0,095	0,899
Cargas Fatoriais dos fatores rotação Varimax e Comunalidades					
Variáveis	Fator1	Fator2	Fator3	Fator4	Comunalidade
Tempo_Fundo_mes	-0,058	-0,031	-0,937	-0,051	0,885
Sharpe_310816_lano	0,986	-0,020	0,036	0,014	0,973
Beta0816_60meses	0,080	-0,021	-0,860	-0,349	0,869
SortinoMAR=RF_310816-1 ano	0,983	-0,031	0,011	-0,039	0,968
Variação_Cta_X-1(1s)	-0,067	0,945	0,042	0,069	0,904
Variação_Cta_X-1(2s)/X(1s)	0,024	0,037	0,192	0,936	0,914
Sharpe_290216-1ano	0,017	0,951	0,006	-0,020	0,906
VAR 95%_ldia_lano	0,567	-0,022	-0,376	-0,556	0,773
Variância	2,2740	1,8025	1,7997	1,316	7,192
% Var	0,284	0,225	0,225	0,165	10,899

**Figura5:Gráfico Scree Plot – Autovalores da matriz de correlação das 8 variáveis**

Então, pelo fato de que são apenas 3 autovalores maiores do que 1 e que o 4º. autovalor explica apenas 9,5% da variância total, optamos por seguir a análise com apenas 3 autovalores.

A variância total explicada para  $m=3$  ficou elevada (80,4%), sendo a parte não explicada pelo modelo igual a 19,6%.

Com  $m=3$  tem-se boas comunalidades com exceção da variável Variação\_Cta\_X-1(2s)/X(1s) que tem comunalidade abaixo de 50% (valor de 47,6%), o que significa que aproximadamente 50% da variância dessa variável é explicada pelos fatores que não foram inseridos no modelo. Esse fato é decorrência da retirada da variável Variação\_Cta\_X da análise fatorial. As comunalidades restantes estão acima de 65% sendo 4 delas acima de 90%.

### Quadro 9: Análise Fatorial – matriz de correlação – $m=3$ – $p=8$

Cargas Fatoriais dos fatores não rotacionados e Comunalidades				
Variáveis	Fator 1	Fator 2	Fator 3	Comunalidade
Tempo_Fundo_mes	0,502	0,622	-0,136	0,658
Sharpe_310816_lano	0,632	-0,748	-0,039	0,961
Beta0816_60meses	0,690	0,565	-0,176	0,826
SortinoMAR=RF_310816-1 ano	0,671	-0,714	-0,039	0,962
Variação_Cta_X-1(1s)	-0,290	-0,085	-0,901	0,904
Variação_Cta_X-1(2s)/X(1s)	-0,538	-0,420	0,105	0,476
Sharpe_290216-1ano	-0,174	-0,098	-0,930	0,905
VAR 95%_1dia_lano	0,848	-0,021	-0,148	0,742
Variância	2,7017	1,9695	1,7631	6,4343
% Var	0,338	0,246	0,220	0,804
Rotação Fatores: Cargas Fatoriais e Comunalidades				
Rotação Varimax				
Variáveis	Fator1	Fator2	Fator3	Comunalidade
Tempo_Fundo_mes	0,803	0,115	0,023	0,658
Sharpe_310816_lano	-0,039	-0,979	0,020	0,961
Beta0816_60meses	0,907	-0,057	0,016	0,826
SortinoMAR=RF_310816-1 ano	0,012	-0,980	0,030	0,962
Variação_Cta_X-1(1s)	-0,077	0,070	-0,945	0,904
Variação_Cta_X-1(2s)/X(1s)	-0,687	0,056	-0,042	0,476
Sharpe_290216-1ano	0,003	-0,019	-0,951	0,905
VAR 95%_1dia_lano	0,624	-0,593	0,022	0,742
Variância	2,3356	2,2970	1,8017	6,4343
% Var	0,292	0,287	0,225	0,804

## 9.2. Análise da Composição dos Fatores - $p=8$ – $m=3$

A seguir descreve-se a composição dos fatores e como esses podem ser interpretados. A solução considerada é com rotação Varimax.

**Fator 1 (F1)** - é composto pelas variáveis:

Tempo Fundo Mes; Beta0816 60 meses, Variação Cta X-1(2s)/X(1s); Var95% 1dia\_lano.

Este fator representa 29,2% da variância total dos dados e basicamente envolve variáveis que representam a performance dos fundos no longo prazo, com pouca variação no longo prazo de seus retornos se comparado com seu mercado. As cargas fatoriais, comunalidades variâncias específicas estimadas das variáveis que compõem o Fator 1 estão apresentadas no Quadro 10.

**Quadro 10: Fator 1 – Comunalidade e Variância Específica**

FATOR 1	Comunalidade	Variância Específica
Tempo_Fundo_mes	0,658	0,342
Beta0816_60meses	0,826	0,174
Varição_Cta_X-1(2s)/X(1s)	0,476	0,524
VAR 95%_1dia_1ano	0,742	0,258

A variável Tempo\_Fundo\_mes é a variável de maior correlação com o Fator 1.

As variáveis com correlações positivas com este fator são: Tempo\_Fundo\_mes; Beta0816\_60meses; VAR 95%\_1dia\_1ano, então os fundos com maiores valores em suas variáveis padronizadas estão associados a maiores valores do fator 1. Já os fundos com maiores valores para a variável padronizada Variação\_Cta\_X-1(2s)/X(1s) estão associados a menores valores do fator 1.

Vale lembrar que para análise faz-se necessário considerar que as variáveis estão padronizadas, ou seja, o valor numérico original de cada variável é subtraído da média amostral da variável e dividida pelo desvio padrão da mesma.

**Fator 2 (F2)** - é composto pelas variáveis:

Sharpe\_310816\_1ano; SortinoMAR=RF\_310816-1 ano.

Nesse Fator tem-se a alocação das variáveis que tratam do risco e retorno dos fundos, ou seja, volatilidade e capacidade de superar as expectativas de rentabilidade do mesmo. Este fator representa 28,7% da variância total dos dados. As comunalidades estimadas foram bem elevadas (acima de 95%), como mostra o Quadro 11

**Quadro 11: Fator 2 – Comunalidade e Variância Específica**

<b>Fator 2</b>	<b>Comunalidade</b>	<b>Variância Específica</b>
Sharpe_310816_1ano	0,905	0,095
SortinoMAR=RF_310816-1 ano	0,962	0,038

As duas variáveis têm alta correlação com o Fator 2 sendo os valores aproximadamente iguais. É importante ressaltar que esse fator tem correlação com a variável VAR 95%\_1dia\_1ano, mas como a correlação desta é maior com o Fator 1, decidiu-se alocá-la no Fator 1 levando-se inclusive em consideração, a natureza da variável que é mais próxima da variável Beta que compõe o Fator 1.

Os fundos com maiores valores para as variáveis Sharpe\_310816\_1ano e SortinoMAR=RF\_310816-1 ano padronizadas estão associados a menores valores do fator 2.

**Fator 3 (F3)** - é composto pelas variáveis:

Varição\_Cta\_X-1(1s), Sharpe\_290216-1ano. As duas variáveis são negativamente correlacionadas com o Fator e tem valores de correlação semelhantes. Valores positivos nesse índice indica grande oscilação. Tais variáveis tratam da volatilidade dos fundos, ou seja, fundos que possuem muita oscilação. Nesse Fator tivemos alocação das variáveis que tratam da volatilidade. Este fator representa 22,5% da variância total dos dados. As comunalidades apresentadas estão no Quadro 12.

**Quadro 12: Fator 3 – Comunalidade e Variância Específica**

<b>Fator 3</b>	<b>Comunalidade</b>	<b>Variância Específica</b>
Varição_Cta_X-1(1s)	0,904	0,096
Sharpe_290216-1ano	0,905	0,095

Os fundos com maiores valores para as variáveis Varição\_Cta\_X-1(1s) e Sharpe\_290216-1anopadronizadas estão associados a menores valores do fator 3.

### 9.3. Cálculo dos Escores dos Fatores e Análise de Resultados

Os coeficientes usados para cálculo dos escores dos fatores descritos na seção 9.2 estão apresentados no Quadro 13, de acordo com a solução  $m=3$ ,  $p=8$  e rotação ortogonal Varimax. O método utilizado para obtenção desses coeficientes foi o de regressão linear.

**Quadro 13: Coeficiente dos fatores dos escores – Fatores via Rotação Varimax**

Rotação Varimax dos Fatores			
Coeficientes dos fatores dos escores			
Variáveis	Fator1	Fator2	Fator3
Tempo_Fundo_mes	0,359845	0,103896	-0,010109
Sharpe_310816_1ano	-0,081357	-0,439174	-0,009597
Beta0816_60meses	0,394778	0,034111	-0,020953
SortinoMAR=RF_310816-1ano	-0,059466	-0,435992	-0,005162
Variação_Cta_X-1(1s)	0,000472	0,005360	-0,524188
Variação_Cta_X-1(2s)/X(1s)	-0,296981	-0,020451	-0,000687
Sharpe_290216-1ano	0,029912	-0,029628	-0,532064
VAR_95%_1dia_1ano	0,235047	-0,223724	-0,020395

As equações para obtenção dos escores para cada fundo da amostra ( $k$  denota o fundo), são dadas a seguir. Em todas as 3 equações as variáveis estão padronizadas, ou seja, os valores numéricos de entrada das variáveis dos fundos nas equações devem estar padronizados pela subtração da média da variável e divisão pelo respectivo desvio-padrão.

$$\begin{aligned}
 F1k = & 0,359845 * Tempo\_Fundo\_mes - 0,081357 * Sharpe\_310816\_1ano + 0,394778 \\
 & * Beta0816\_60meses - 0,059466 * SortinoMARRF_{310816\_1ano} + 0,000472 \\
 & * Variação\_Cta\_X\_1(1s) - 0,296981 * Variação\_Cta\_X\_1(2s)/X(1s) + 0,029912 \\
 & * Sharpe\_290216\_1ano + 0,235047 * VAR\_95\%\_1dia\_1ano
 \end{aligned}$$

$$\begin{aligned}
 F2k = & 0,103896 * Tempo\_Fundo\_mes - 0,439174 * Sharpe\_310816\_1ano + 0,034111 \\
 & * Beta0816\_60meses - 0,43992 * SortinoMAR\_RF\_310816\_1ano + 0,005360 \\
 & * Variação\_Cta\_X\_1(1s) - 0,020451 * Variação\_Cta\_X - 1(2s)/X(1s) - 0,029628 \\
 & * Sharpe\_290216 - 1ano - 0,223724 * VAR\_95\%\_1dia\_1ano
 \end{aligned}$$

$$\begin{aligned}
 F3k = & -0,010109 * Tempo\_Fundo\_mes - 0,009597 * Sharpe\_310816\_1ano - 0,020953 \\
 & * Beta0816\_60meses - 0,005162 * SortinoMARRF\_310816 - 1ano - 0,524188 \\
 & * Variação\_Cta\_X\_1(1s) - 0,000687 * Variação\_Cta\_X\_1(2s)/X(1s) - 0,532064 \\
 & * Sharpe\_290216\_1ano - 0,020395 * VAR\_95\%\_1dia\_1ano
 \end{aligned}$$

Como exemplificação, apresentaremos a seguir o cálculo do escore do Fator 1, para o fundo 210FIA cujos dados são apresentados no Quadro 14, mas para que possamos utilizar os dados, foi necessário padronizá-los.

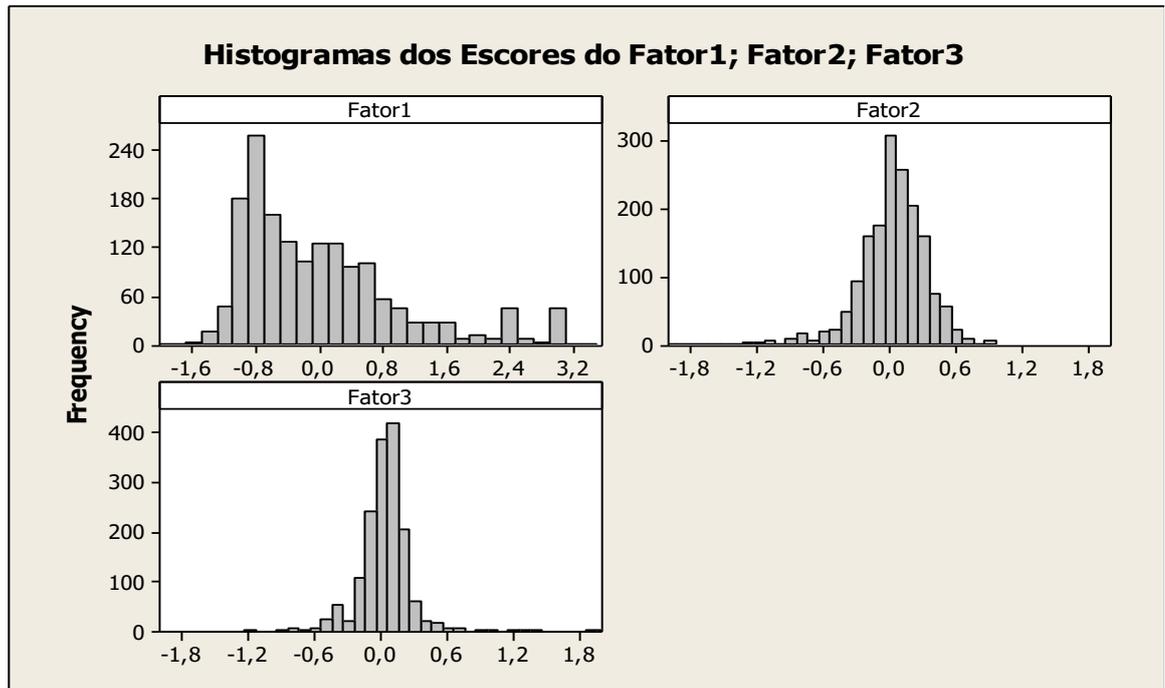
**Quadro 14: cálculo do escore do Fator 1, para o fundo 210FIA**

Variáveis	Valores originais	Média	Desvio Padrão	Valores padronizados
Tempo_Fundo_mes	51,000	83,350	57,470	-0,563
Sharpe_310816_1ano	0,789	0,648	1,556	0,091
Beta0816_60meses	0,000	0,432	0,497	-0,870
SortinoMAR=RF_310816-1ano	1,387	1,213	5,801	0,030
Variação_Cta_X-1(1s)	-0,076	-0,044	0,198	-0,162
Variação_Cta_X-1(2s)/X(1s)	0,012	-0,061	0,138	0,529
Sharpe_290216-1ano	1,187	1,061	4,373	0,029
VAR_95%_1dia_1ano	1,561	2,430	1,920	-0,452

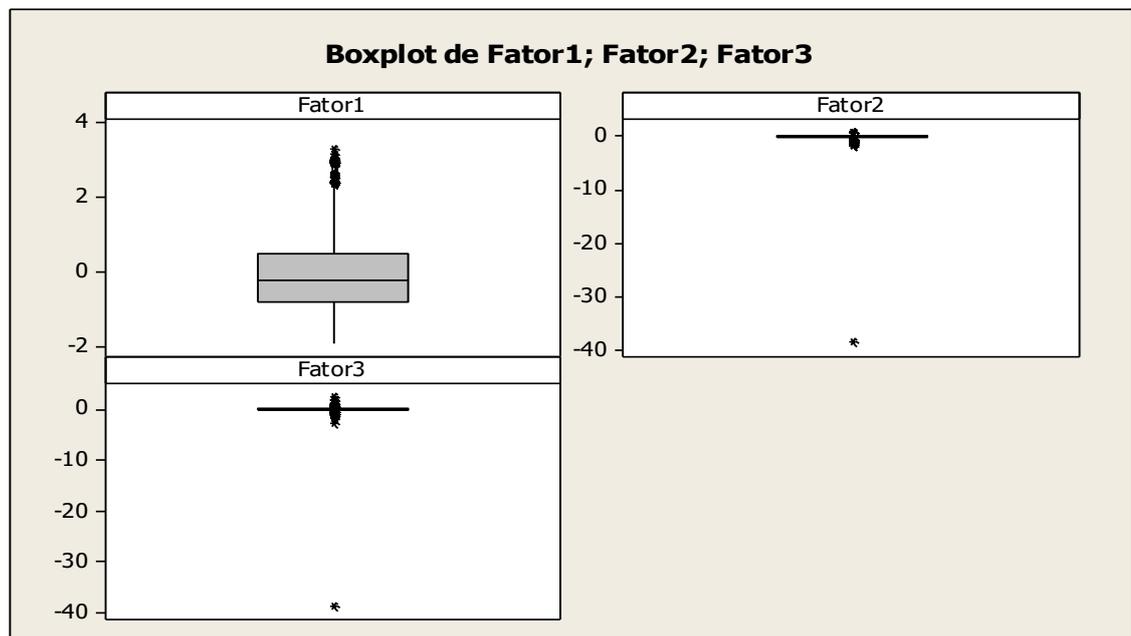
$$\begin{aligned}
 F1 = & 0,359845 * (-0,5629) - 0,081357 * (0,0905) + 0,394778 * (-0,8699) - 0,059466 * (0,0299) \\
 & + 0,000472 * (-0,1621) - 0,296981 * (0,5292) + 0,029912 * (0,0288) + 0,235047 \\
 & * (-0,4522) = -0,818
 \end{aligned}$$

Os escores dos fatores não tem distribuição normal como pode ser visualizado nos histogramas apresentados na Figura 6.

Pelos *Box-plots* apresentados na Figura 7 observamos que o fator 1 possui alguns pontos discrepantes, ao passo que os fatores 2 e 3 possuem *outliers*.



**Figura 6: Histogramas dos Escores dos Fatores 1,2 e 3.**



**Figura 7: Box-plots dos fatores 1, 2 e 3**

Para entender melhor os escores gerados pelo modelo, temos as estatísticas descritivas apresentadas no Quadro 15.

**Quadro 15: Estatísticas Descritivas dos Escores dos Fatores. m=3**

**Estatísticas Descritivas com *Outliers*: Escore Fator 1; Escore Fator 2; Escore Fator 3**

Variáveis	Mínimo	1º quartil	Mediana	3º quartil	Máximo	amplitude
Escore Fator 1	-1,916	-0,782	-0,2272	0,483	3,321	5,2375
Escore Fator 2	-38,755	-0,124	0,0332	0,202	0,889	39,6439
Escore Fator 3	-38,881	-0,071	0,0329	0,130	2,712	41,5925

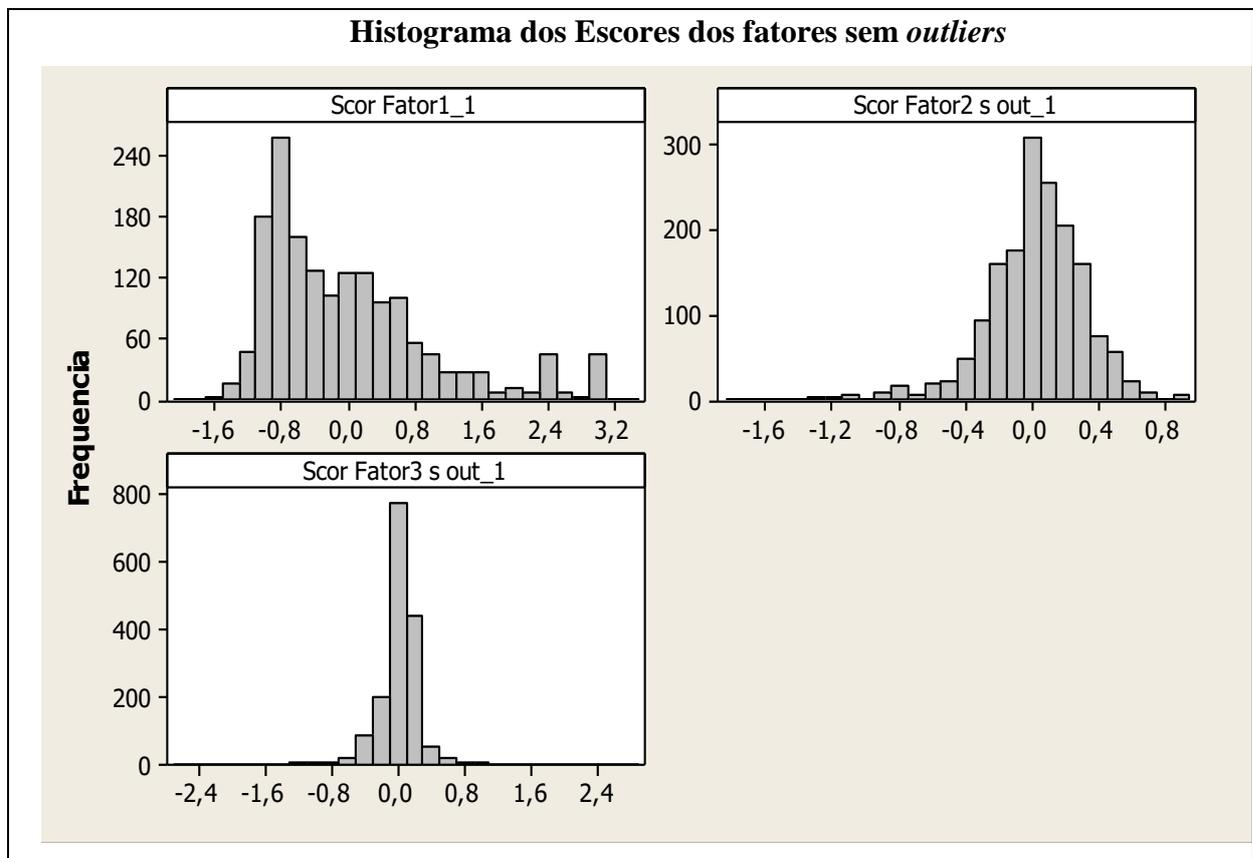
Variáveis	Assimetria
Escore Fator 1	1,23
Escore Fator 2	-35,53
Escore Fator 3	-35,84

**Estatísticas Descritivas sem *Outliers*: Escore Fator 1; Escore Fator 2; Escore Fator 3**

Variáveis	Mínimo	1º quartil	Mediana	3º quartil	Máximo	amplitude
Escore Fator 1	-1,9162	-0,7823	-0,2272	0,4830	3,3214	5,2375
Escore Fator 2	-1,79339	-0,12366	0,03324	0,20210	0,88896	2,68235
Escore Fator 3	-2,55076	-0,07088	0,03287	0,12971	2,71154	5,26230

Variáveis	Assimetria
Escore Fator 1	1,23
Escore Fator 2	-0,97
Escore Fator 3	0,57

Após a retirada dos *outliers* observa-se que tanto a amplitude quanto a assimetria diminuíram consideravelmente, o que mostra que os dados ficaram mais homogêneos.



**Figura 8: Histograma dos escores sem *outliers***

Observamos que após a retirada dos *outliers* os escores dos fatores 2 e 3 apresentou considerável redução na amplitude, com baixo valor de assimetria.

**Quadro 16: Fundos outliers – Dados Originais e Padronizados**

Dados Originais											
Nome	ID	Tempo_Fundo_mes	Sharpe_310816_1ano	Beta_0816_60meses	SortinoMAR=RF_310816-1ano	Variação_Cta_2015	Variação_Cta_2015/2016	Variação_Cta_2016	Sharpe_290216_1ano	VAR%base1dia_confiança95%_280816_1ano	VAR%base1dia_confiança95%_280916_2anos
Bram FIA	347	158,00	0,38	0,95	0,66	-0,09	-0,08	0,32	1,07	2,63	2,66
Ja Centenario FIA	1017	67,00	60,26	2,26	232,59	-0,91	-0,72	8,52	0,43	42,87	33,52
<b>Média dos demais fundos</b>		<b>83,31</b>	<b>0,61</b>	<b>0,43</b>	<b>1,07</b>	<b>-0,04</b>	<b>-0,06</b>	<b>0,30</b>	<b>1,06</b>	<b>2,40</b>	<b>7,98</b>

Dados Padronizados											
Nome	ID	Tempo_Fundo_mes	Sharpe_310816_1ano	Beta_0816_60meses	SortinoMAR=RF_310816-1ano	Variação_Cta_2015	Variação_Cta_2015/2016	Variação_Cta_2016	Sharpe_290216_1ano	VAR%base1dia_confiança95%_280816_1ano	VAR%base1dia_confiança95%_280916_2anos
Bram FIA	347	1,30	-0,17	1,05	-0,10	-0,25	-0,15	0,02	0,00	0,11	-0,02
Ja Centenario FIA	1017	-0,28	38,32	3,68	39,88	-4,36	-4,74	9,72	-0,14	21,06	0,11

Os dados acima evidenciam que o fundo identificado como 347, quando analisamos os dados originais, diverge consideravelmente da média dos demais fundos na variável Tempo\_Fundo\_mes, que apresenta valor bem superior à média dos demais fundos e na variável VAR%base1dia\_confiança95%\_280916\_2anos, sendo inferior à média.

Já o fundo identificado como 1017 diverge da média em praticamente todas as variáveis quando analisamos os dados originais, mas quando analisamos os dados padronizados, temos que as variáveis que mais destoam são: Sharpe\_310816\_1ano, SortinoMAR=RF\_310816-1ano e VAR%base1dia\_confiança95%\_280816\_1ano

Considerando que a construção dos escores leva em consideração as variáveis originais, a presença de *outlier* trará oscilação nos escores, mas o impacto para os fatores é minimizado, pois a construção dos mesmos é feita a partir da padronização das variáveis. Como demonstração, apresentamos o Quadro 17 que traz os resultados do ajuste da análise fatorial sem os fundos responsáveis pelos *outliers* nos escores dos fatores 2 e 3.

A variância total explicada para  $m=3$  com a presença dos *outliers* era de 80,4 e ao retirar os 2 fundos do ajuste do modelo fatorial a variância explicada caiu para 72,4%; contudo a composição dos fatores permaneceu a mesma, alterando um pouco apenas as cargas fatoriais. Nessa nova solução a alocação da variável VAR 95%\_1dia\_1ano para o Fator 1 ficou mais acentuada em relação à solução com a presença dos 2 fundos *outliers*.

### Quadro 17: Análise Fatorial sem fundos com escores discrepantes

<b>Análise Fatorial via Matriz de Correlação: Componentes Principais</b>				
Cargas Fatoriais dos fatores não rotacionados e Comunalidades				
Variáveis	Fator1	Fator2	Fator3	Communalidade
Tempo_Fundo_mes	0,672	-0,368	-0,053	0,589
Sharpe_310816_1ano	-0,636	-0,688	-0,229	0,929
Beta0816_60meses	0,756	-0,513	-0,012	0,834
SortinoMAR=RF_310816-1 ano	-0,599	-0,694	-0,312	0,938
Variação_Cta_X-1(1s)	-0,248	-0,163	0,768	0,678
Variação_Cta_X-1(2s)/X(1s)	-0,733	0,018	-0,004	0,537
Sharpe_290216-1ano	0,134	0,401	-0,699	0,667
VAR 95%_1dia_1ano	0,621	-0,481	-0,012	0,617
Variância	2,7863	1,7725	1,2320	5,7908
% Var	0,348	0,222	0,154	0,724
<b>Análise Fatorial via Matriz de Correlação: Componentes Principais</b>				
Cargas Fatoriais dos fatores rotacionados. Rotação Varimax				
Variáveis	Fator1	Fator2	Fator3	Communalidade
Tempo_Fundo_mes	-0,761	-0,083	0,051	0,589
Sharpe_310816_1ano	0,102	0,953	-0,107	0,929
Beta0816_60meses	-0,913	-0,036	-0,017	0,834
SortinoMAR=RF_310816-1 ano	0,064	0,966	-0,025	0,938
Variação_Cta_X-1(1s)	0,139	-0,008	-0,812	0,678
Variação_Cta_X-1(2s)/X(1s)	0,604	0,399	-0,116	0,537
Sharpe_290216-1ano	0,095	-0,131	0,801	0,667
VAR 95%_1dia_1ano	-0,785	0,017	-0,030	0,617
Variância	2,4358	2,0259	1,3291	5,7908
% Var	0,304	0,253	0,166	0,724

Por esse motivo passamos a análise dos fatores. No fator 1, possuir um escore alto significa que os Fundos estão consolidados no mercado com tempo de mercado, com rentabilidade mais estável. Os Fundos de investimento com altos valores positivos neste fator são fundos que possuem tempo de mercado, mas com rentabilidade. Os Fundos com baixos escores são aqueles que são recentes no mercado, com maior volatilidade.

Já o fator 2 trata da capacidade de geração de retornos positivos, além de ser um bom indicador da capacidade de superar as expectativas de rentabilidade do mesmo. Os Fundos de investimento com altos valores negativos no escore deste fator são fundos que tem desempenho melhor que o *benchmark*, sem ter muita variação em sua performance. Os fundos com maiores

escores positivos no fator 2 são fundos que trazem maior volatilidade dos retornos, já os fundos com menores valores de escores neste fator são fundos mais estáveis

Os Fundos que possuem escores alto valores no Fator 3 são fundos voláteis.

Quando classificamos os fundos usando os escores de acordo com a interpretação de cada fator, obtemos a disposição dos mesmos conforme Quadro 17, ou seja, o Quadro 17 foi construído com os 15 fundos melhor classificados pelo Fator 1.

Se utilizarmos apenas o Fator 2 como critério de classificação, por ser o fator que está mais relacionado a retorno, ressalta-se que a análise fatorial foi realizada pela matriz de correlação assim, tem-se o Quadro 18.

**Quadro 18: 5 Melhores e 5 Piores Fundos e suas variáveis originais**

5 PIORES FUNDOS CLASSIFICADOS PELO FATOR 2									
ID	Escore Fator2	Tempo Fundo_mes	Sharpe_310816_1ano	Beta0816_60mes	SortinoMA R=RF_310816-1ano	Varição_Cta_X-1(2s)/X(1s)	Varição_Cta_X-1(1s)	Sharpe_290216_1ano	VAR%base1dia - confiança95%_280816_1ano
1478	0,8890	227	0,00	0,32	0,00	-0,26	-0,59	0,00	0,00
1083	0,8873	245	0,00	0,95	0,00	-0,12	-0,10	1,26	0,00
1219	0,8677	227	0,00	0,88	0,00	0,11	-0,12	0,00	0,00
942	0,8620	238	0,00	0,81	0,00	-0,09	-0,08	1,39	0,00
894	0,8550	227	0,00	0,60	0,00	-0,09	-0,20	0,00	0,00
5 MELHORES FUNDOS CLASSIFICADOS PELO FATOR 2									
ID	Escore Fator2	Tempo Fundo_mes	Sharpe_310816_1ano	Beta0816_60mes	SortinoMA R=RF_310816-1ano	Varição_Cta_X-1(2s)/X(1s)	Varição_Cta_X-1(1s)	Sharpe_290216_1ano	VAR%base1dia - confiança95%_280816_1ano
1236	-1,7934	24	1,47	0,00	3,47	0,46	0,05	1,13	13,17
1348	-1,6971	64	2,89	1,42	6,12	-0,57	-0,08	0,56	8,61
586	-1,3308	27	2,57	0,00	4,98	-0,56	0,57	0,59	4,74
519	-1,3139	47	2,14	0,00	3,86	-0,20	0,42	0,31	6,96
460	-1,2621	37	0,56	0	1,29	-0,43	0,95	0,45	11,11

**Quadro 19:05 Melhores e 05 Piores Fundos e suas variáveis Padronizadas**

5 PIORES FUNDOS CLASSIFICADOS PELO FATOR 2									
ID	Escore Fator2	Tempo_Fundo_mes	Sharpe_310816_1ano	Beta0816_60mes	SortinoMAR=RF_310816-1ano	Varição_Cta_X-1(2s)/X(1s)	Varição_Cta_X-1(1s)	Sharpe_290216_1ano	VAR%base1dia_confiança95%_280816_1ano
1478	0,8890	2,499612	-0,42	-0,22	-0,21	-1,12	-3,81	-0,24	-1,27
1083	0,8873	2,812814	-0,42	1,05	-0,21	-0,40	-0,28	0,05	-1,27
1219	0,8677	2,499612	-0,42	0,90	-0,21	0,81	-0,43	-0,24	-1,27
942	0,8620	2,691013	-0,42	0,76	-0,21	-0,22	-0,17	0,08	-1,27
894	0,8550	2,499612	-0,42	0,34	-0,21	-0,23	-0,97	-0,24	-1,27
5 MELHORES FUNDOS CLASSIFICADOS PELO FATOR 2									
ID	Escore Fator2	Tempo_Fundo_mes	Sharpe_310816_1ano	Beta0816_60mes	SortinoMAR=RF_310816-1ano	Varição_Cta_X-1(2s)/X(1s)	Varição_Cta_X-1(1s)	Sharpe_290216_1ano	VAR%base1dia_confiança95%_280816_1ano
1236	-1,7934	-1,03262	0,53	-0,87	0,39	2,53	0,82	0,02	5,59
1348	-1,6971	-0,33661	1,44	1,99	0,85	-2,68	-0,12	-0,11	3,22
586	-1,3308	-0,98042	1,24	-0,87	0,65	-2,63	4,60	-0,11	1,20
519	-1,3139	-0,63241	0,96	-0,87	0,46	-0,79	3,50	-0,17	2,36
460	-1,2621	-0,80641	-0,05	-0,87	0,01	-1,93	7,33	-0,14	4,52

Os fundos que apresentaram melhores classificações apresentam elevados valores para as variáveis que compõe o fator 2, e, surpreendentemente, são fundos com menor tempo de existência se observamos os fundos com piores classificações.

**Quadro 20: Escores dos 15 Fundos melhores posicionados, por Fator 1, Fator 2 e Fator 3.**

Classificação pelo Fator 1		Classificação pelo Fator 2		Classificação pelo Fator 3	
ID	Escore Fator1	ID	Escore Fator2	ID	Escore Fator3
831	3,3214	1478	-0,2877	1291	1,1695
299	3,1633	1083	-0,1543	467	-0,3655
133	3,0596	1219	-0,0764	1236	-0,3745
169	3,0440	942	-0,1773	833	-0,3040
724	3,0364	894	-0,0944	78	-0,3978
1120	3,0275	1192	-0,0878	830	-0,3972
239	3,0108	1248	-0,1360	142	-0,4229
134	3,0001	1220	-0,1135	1065	-0,4193
240	2,9996	255	-0,1386	186	-0,4211
725	2,9933	1221	-0,1582	828	-0,4201
277	2,9813	1414	-0,1608	542	-0,3460
278	2,9807	1290	-0,1637	829	-0,3460
279	2,9805	565	-0,1651	541	-0,3449
531	2,9758	557	-0,1279	751	-0,4207
778	2,9755	625	-0,1619	1291	-0,4180

Os melhores fundos bem classificados por um fator, não necessariamente serão classificados como bom em outro fator da análise, devido a característica de convergência das variáveis de acordo com suas correlações.

No Mês de fevereiro/17, colheu-se mais uma amostra dos retornos dos fundos e o fundo apontado como o melhor pelo fator 1 apresentou estabilidade nos retornos, principalmente a longo prazo. Esta característica foi apontada pelo fator, comprovando assim a assertividade do índice para este objetivo

**Quadro 21: Retornos do Fundo 831 em Fev/17.**

ID	Data da Últ Cotação	Retorno 1 mês em %	Retorno 3 meses em %	Retorno 6 meses em %	Retorno 12 meses em %	Retorno 24 meses em %
831	31/01/2017	14,13	25,13	6,14	22,97	83,53

## 10. CONSIDERAÇÕES FINAIS

Através do trabalho desenvolvido nessa monografia conclui-se que, por meio da Análise Fatorial, identificamos 3 grupos de variáveis sendo que cada grupo tem a característica de geração de um índice que pode ser usado para realizar a classificação de fundos. O primeiro grupo é formado por variáveis que evidenciam maior tempo de mercado, retornos mais expressivos e com menor possibilidade de perda financeira. Fundos de investimento bem posicionados neste grupo, ou seja, com maiores valores neste índice tendem a gerar maior credibilidade para os investidores, principalmente para os investidores institucionais que precisam prestar contas aos participantes de seus investimentos e dos retornos alcançados, ou seja, o primeiro grupo é formado por variáveis que evidenciam maior tempo de mercado, retornos mais expressivos e com menor possibilidade de perda financeira.

Já o segundo grupo é constituído por variáveis que retratam a capacidade de gerar retornos maiores que seus *benchmarks*, por isso os fundos com baixos valores neste índice tendem a ter desempenho superior ao seu *benchmark*, mas sem grandes oscilações. São fundos que conseguem gerar retornos superiores à sua referência, o que o mercado de investimento descreve como a capacidade de gerar alfa de Jensen e assim são mais estáveis do ponto de vista de investimento. Estes fundos conseguem ter períodos retornos positivos superiores a períodos de retornos negativos. Resumidamente, o segundo grupo é constituído por variáveis que retratam a capacidade de gerar retornos maiores que seus *benchmarks*.

O terceiro grupo é composto por variáveis que caracterizam oscilação. Fundos bem posicionados (escores menores) neste índice são fundos com muita oscilação, ou seja, fundos com possibilidade de ter rentabilidades elevadas, mas também trajetória mais incerta. Tais fundos podem ser considerados mais especulativos por esta característica, então o terceiro grupo é composto por variáveis que caracterizam oscilação.

O modelo de análise fatorial então teve sucesso em atender o objetivo deste trabalho, pois com sua utilização foi possível definir índices que podem auxiliar na seleção de fundos de investimento em renda variável.

Importante destacar que dentro de cada índice podemos segregar fundos com desempenho bom, mediano e ruim. O modelo de análise fatorial então teve sucesso em atender o objetivo deste trabalho, pois com sua utilização foi possível definir índices que podem auxiliar

na seleção de fundos de investimento em renda variável, podendo inclusive ser utilizado pelos Fundos de Pensão a fim de otimizar o retorno do segmento em renda variável.

Considerando que as Entidades de Previdência necessitam otimizar o retorno dos investimentos, principalmente no segmento de renda variável, e que tais entidades devem preferir fundos de investimento que sejam mais consolidados no mercado, o fator 1 atende então, o objetivo geral deste trabalho.

## REFERÊNCIAS BIBLIOGRÁFICAS

ASSAF NETO, A., **Mercado Financeiro**, 7º ed São Paulo, Atlas, 2007.

ASSAF NETO, A.; **Finanças Corporativas e valor**, São Paulo, Atlas, 2003b.

BATICH; M. **Previdência do Trabalhador: uma trajetória inesperada**, São Paulo. 2004.

BRASIL, **Lei das Sociedades por ações**, 29º ed São Paulo, Atlas, 2002.

BRASIL. Ministério da Previdência e Assistência Social (MPAS). **Livro Branco da Previdência Social** – Brasília: MPAS/GM, 2002. Disponível em:<[http://www.previdencia.gov.br/arquivos/office/3\\_081014-104854-755.pdf](http://www.previdencia.gov.br/arquivos/office/3_081014-104854-755.pdf)>acesso em 15 nov. 2016

BRASIL. Ministério da Previdência e Assistência Social (MPAS). **Panorama da Previdência Social** – Brasília: MPAS/GM, 2008. Disponível em:<[http://www.previdencia.gov.br/arquivos/office/3\\_090126-092058-729.pdf](http://www.previdencia.gov.br/arquivos/office/3_090126-092058-729.pdf)>acesso em 02 Fev 2017.

BRASIL; Constituição (1988) **Constituição Federativa da Brasil**. Brasília: Senado, 1988.

CERVO, A. L. BERVIAN, P. A. **Metodologia científica**. 5. ed. São Paulo: Prentice Hall, 2002.

CHAN, B. L.; SILVA; F. L.; MARTINS, G. A. Fundamentos da previdência Complementar: Da Atuária à Contabilidade. São Paulo, Atlas FIECAFI/USP, 2006

Comissão de Valores Mobiliários, **CVM**, Brasília, disponível em:<<http://www.cvm.gov.br/legislacao/instrucoes/inst554.html>> Acesso em 19 ago. 2016.

Comissão de Valores Mobiliários, Brasília, **CVM**, disponível em:<<http://www.cvm.gov.br/legislacao/instrucoes/inst555.html>> Acesso em 19 ago. 2016.

CORRAR, L. J; PAULO, E.; DIAS FILHO, J.M.a (Coords.). **Análise multivariada: para os cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2007.

FARIAS, A., ORNELAS, J. R. H. Finanças e Sistema Financeiro Nacional para Concurso: **Questões Resolvidas de Concursos do Banco Central**, Tesouro Nacional, BNDES, CVM, CEF e BB, dentre outros. São Paulo, Atlas, 07/2015. [Minha Biblioteca]

FEDERAÇÃO NACIONAL DE PREVIDÊNCIA PRIVADA E VIDA, **Mercado**. Disponível em:<<http://fenaprevi.org.br/fenaprevi/previdencia-aberta/>> Acesso em 02 out. 2017

HAIR,J.F., BLACK, W.C.; BABINB.J.;ANDERSON, R.E.;TATHAM,R.L. **Multivariate data analysis .UpperSaddle River**, NJ: Prentice hall,2009.

LEI Nº 8.213, de 24 de Julho de 1991, disponível em:<[http://www.planalto.gov.br/ccivil\\_03/leis/L8213cons.htm](http://www.planalto.gov.br/ccivil_03/leis/L8213cons.htm)> Acesso em 12 de dez. de 2016.

<[http://anpad.org.br/periodicos/arq\\_pdf/a\\_419.pdf](http://anpad.org.br/periodicos/arq_pdf/a_419.pdf)> acesso em 30 de out. de 2016

MINGOTI, S. A. **Análise de Dados Através de métodos de Estatística multivariada**: uma abordagem aplicada. Belo Horizonte: Editora UFMG, 2005.