

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Especialização em Estatística com Ênfase em Indústria e Mercado

**IDENTIFICAÇÃO DAS VARIÁVEIS QUE MAIS INFLUENCIAM NO
CUSTO ASSISTENCIAL DA SAÚDE SUPLEMENTAR:
Estudo de caso de uma operadora de plano de saúde**

Renata Costa Santos

Belo Horizonte

2014

Renata Costa Santos

**IDENTIFICAÇÃO DAS VARIÁVEIS QUE MAIS INFLUENCIAM NO
CUSTO ASSISTENCIAL DA SAÚDE SUPLEMENTAR:
Estudo de caso de uma operadora de plano de saúde**

Monografia apresentada à Universidade Federal de Minas Gerais, como parte dos requisitos exigidos pelo Programa de Pós-Graduação em Estatística, para obtenção do título de Especialista, com ênfase em Indústria e Mercado.

Orientadora: Professora Sueli Aparecida Mingoti

Belo Horizonte

2014

RESUMO

Este trabalho tem como objetivo identificar as variáveis que mais impactam no custo de saúde de uma operadora de plano de saúde suplementar. O Sistema de Saúde Pública no Brasil não está conseguindo garantir a assistência a toda população brasileira. O modelo atual adotado pelo governo não é suficiente para proporcionar um cuidado digno para os cidadãos e vem mostrando seu esgotamento. Tal fato faz com que as pessoas busquem alternativas, recorrendo assim à contratação de serviços privados de assistência à saúde. Os custos com saúde vêm subindo em proporções assustadoras, fazendo com que as operadoras de planos saúde tenham uma área de gestão de custos eficiente que consiga identificar os “vilões” que impactam nos seus resultados. Tais custos são divididos em dois regimes de atendimento: o ambulatorial e o de internação. Em relação ao regime ambulatorial, as variáveis estudadas serão: Consultas, Exames, Honorários, Materiais, Medicamentos, Pacotes, Procedimentos não Médicos e Taxas. Já no regime de internação as análises serão feitas com base nas seguintes variáveis: Diárias, Exames, Honorários, Materiais, Medicamentos, Pacotes, Procedimentos não Médicos e Taxas. Diante desse cenário, identificar as variáveis e mensurar os pesos das variáveis que compõem o custo assistencial de uma operadora de planos de saúde é de fundamental relevância, visando uma melhor gestão dos custos e norteando os gestores com relação às melhores estratégias de negociação de preço com os prestadores de serviços contratados. Para a realização desse estudo utilizou-se dados históricos dos custos de uma operadora do ramo da saúde que foram analisados através das técnicas de Estatística Multivariada.

Palavras-chave: Gerenciamento de custo; Operadora de saúde; Plano de saúde; Custos assistenciais; Técnicas de estatística multivariada.

LISTA DE ILUSTRAÇÕES

FIGURA 1: Esquema de aplicação da análise de componentes principais.....	16
FIGURA 2: Boxplot das variáveis de custo do regime de atendimento ambulatorial.....	25
FIGURA 3. Gráfico de dispersão: Procedimento não-Médico versus Pacote e Procedimento não médico versus Honorário.....	26
FIGURA 4: Boxplot das variáveis de custo do regime de atendimento de internação.....	28
FIGURA 5. Gráficos de dispersão: Procedimentos não-Médicos versus Diárias e Taxa versus Honorário.....	30

LISTA DE QUADROS

Quadro 1: Conceituação das variáveis utilizadas no estudo.....	21
Quadro 2: Estatística descritiva das variáveis em regime de atendimento ambulatorial.....	24
Quadro 3: Correlações amostrais das variáveis de custo do regime ambulatorial.....	25
Quadro 4: Estatística descritiva das variáveis de custo do regime de internação.....	28
Quadro 5: Matriz de Correlação das variáveis de custo do regime de internação.....	29
Quadro 6: Componentes Principais pela matriz de Covariâncias das variáveis de custo do regime ambulatorial: autovalores e proporção da variância total explicada.....	31
Quadro 7: Coeficientes das componentes principais pela matriz de Covariâncias das variáveis de custo do regime ambulatorial.....	31
Quadro 8: Correlações entre as componentes e as variáveis originais (pela matriz de covariâncias) – regime ambulatorial.....	32
Quadro 9: Componentes Principais pela matriz de Covariâncias das variáveis de custo do regime de internação: autovalores e proporção da variância total explicada.....	33
Quadro 10: Autovetores das componentes principais pela matriz de Covariâncias das variáveis de custo do regime de internação.....	34
Quadro 11: Correlação entre as componentes e as variáveis originais (pela matriz de covariâncias) – regime de internação.....	35
Quadro 12: Componentes Principais pela matriz de Correlação das variáveis do regime ambulatorial.....	36
Quadro 13: Autovetores das componentes principais pela matriz de Correlação das variáveis de custo do regime ambulatorial.....	36
Quadro 14: Correlação entre as componentes e as variáveis originais (pela matriz de correlação) – regime ambulatorial.....	38
Quadro 15: Componentes Principais pela matriz de Correlação das variáveis do regime de internação: autovalores e variância explicada por cada componente.....	38
Quadro 16: Coeficientes das componentes principais pela matriz de Correlação das variáveis de custo do regime de internação.....	39

Quadro 17: Correlação entre as componentes e as variáveis originais (pela matriz de correlação) – regime de internação.....	40
Quadro 18: Análise Fatorial com e sem rotação para o regime ambulatorial.....	41
Quadro 19: Análise Fatorial com e sem rotação para o regime de internação.....	43

LISTA DE GRÁFICOS

Gráfico 1: Evolução do custo mensal por regime de internação.....	22
Gráfico 2: Percentual de representatividade de cada variável no custo no regime ambulatorial.....	23
Gráfico 3: Percentual de representatividade de cada variável no custo assistencial do regime de internação.....	27

SUMÁRIO

1 INTRODUÇÃO.....	9
2 OBJETIVO GERAL.....	10
3 A SAÚDE SUPLEMENTAR NO BRASIL.....	10
4 ANÁLISE MULTIVARIADA.....	13
4.1 Análise de Componentes Principais.....	13
4.2 Análise Fatorial.....	17
5 BANCO DE DADOS.....	21
6 ANÁLISE DOS RESULTADOS.....	23
6.1 Estatística descritiva das variáveis.....	23
6.2 Análise de Componentes Principais.....	31
6.2.1 <i>Análise de componentes principais pela matriz de covariâncias</i>	31
6.2.2 <i>Análise de componentes principais pela matriz de correlação</i>	37
6.3 Análise Fatorial.....	43
7 CONSIDERAÇÕES FINAIS.....	47
REFERÊNCIAS.....	50

1 INTRODUÇÃO

A economia brasileira vem passando por transformações ao longo do tempo, principalmente depois da criação do Plano Real. O controle da inflação, a abertura ao mercado externo e a menor intervenção estatal foram algumas das principais mudanças ocorridas no âmbito econômico. Nos últimos anos, tem-se observado uma maior concessão de crédito e maior poder de compra dos brasileiros.

Isso permite com que produtos antes tidos como produtos de luxo se tornassem mais acessíveis à população. Dentre eles estão os planos de saúde, que atingiram um patamar de crescimento muito significativo até então. Segundo dados da Agência Nacional de Saúde (ANS) hoje já são 50,9 milhões de brasileiros que possuem planos de saúde.

Aliado a esse crescimento está a precariedade nos atendimentos de saúde dos serviços públicos. O Sistema de Saúde Pública no Brasil não está conseguindo garantir a assistência a toda população brasileira. O modelo atual adotado pelo governo não é suficiente para proporcionar um cuidado digno para os cidadãos e vem mostrando seu esgotamento. Tal fato faz com que as pessoas busquem alternativas, recorrendo assim à contratação de serviços privados de assistência à saúde.

Os custos com saúde vêm subindo em proporções assustadoras. O aumento do Rol de serviços obrigatórios ofertados por operadoras de plano de saúde, o avanço da tecnologia para a área médica e o envelhecimento da população são alguns dos principais fatores para esse aumento dos custos. Sendo assim, é necessário que as operadoras de planos saúde tenham uma área de gestão de custos eficiente que consiga identificar os “vilões” que impactam nos seus resultados.

Dentre as metodologias estatísticas existentes, uma que é utilizada na gestão de custos é a Análise Multivariada. Trata-se de um conjunto de técnicas estatísticas que permite a análise simultânea de duas ou mais variáveis.

Uma dessas técnicas é a Análise de Componentes Principais (ACP), que, segundo Mingoti (2007), tem como objetivo principal “explicar a estrutura de variância e covariância de um vetor aleatório, composto por p-variáveis aleatórias, através de combinações lineares das variáveis originais”.

O uso dessa metodologia possibilita a identificação das variáveis responsáveis pelas variações no custo assistencial sem que ocorra perda significativa das informações.

A partir dessa análise é possível identificar também quais as variáveis que possuem características estatísticas semelhantes e agrupá-las em um único grupo (ou fator), o que é feito através da técnica denominada de Análise Fatorial, conforme explica Vicini (2005).

Sendo assim, o objetivo deste trabalho de monografia será o de mensurar os pesos das variáveis que compõem o custo assistencial de uma operadora de planos de saúde visando uma melhor gestão dos custos e norteando os gestores com relação às melhores estratégias de negociação de preço com os prestadores de serviços contratados.

Nesse trabalho será utilizada a base de dados de uma operadora de plano de saúde suplementar contendo os custos assistenciais mensais segregados por diversas variáveis que compõem esse custo (consultas, diárias, exames, honorário, material, medicamento, pacotes, procedimentos não médicos e taxas).

2 OBJETIVO GERAL

O objetivo dessa monografia é identificar, através de técnicas estatísticas de Análise Multivariada, quais as variáveis que mais impactam no custo de uma operadora de plano de saúde suplementar.

De posse dos resultados a operadora poderá redefinir suas estratégias de negociação com os prestadores de serviço bem como melhorar a gestão das variáveis que mais influenciam no custo total.

3 A SAÚDE SUPLEMENTAR NO BRASIL

A Saúde no Brasil está vivendo um cenário de crise. O Sistema Único de Saúde (SUS) tem como obrigação oferecer a todo cidadão qualidade assistencial médica e odontológica, porém não é isso que está ocorrendo no Brasil.

A população, cujo direito à saúde é garantido pela Constituição Federal de 1988, está recorrendo ao sistema privado de assistência médica e odontológica, com o objetivo de fazer valer os cuidados obrigatórios que deveriam ser garantidos pelo Estado. A dificuldade de atendimento nos serviços público, a precariedade dos equipamentos utilizados principalmente nos hospitais e o quadro técnico de profissionais insuficiente para atender à demanda, faz com

que as pessoas procurem serviços médicos e odontológicos privados através das operadoras de planos de saúde.

Segundo Silveira, citado por Martins et al. (2013) os planos de saúde surgiram em meados de 1956 com o aumento do êxodo rural e também por causa da qualidade precária dos serviços públicos de saúde no Brasil desde aquela época.

Porém, foi somente em 1998, com a aprovação da Lei 9.656 que estabelece os princípios e diretrizes para a operação de planos de saúde no Brasil, que esse setor passou a ser regulamentado pelo Governo Federal.

Cerca de dois anos depois foi criada a Agência Nacional de Saúde Suplementar (ANS) que passou a ser o órgão regulamentador desse setor. A partir de então, através das resoluções normativas e operacionais, a ANS vem fiscalizando e regulando as operações dos planos de saúde.

Apesar de ainda recente, é notável o crescimento desse mercado de operação de plano de saúde ao longo dos anos. Segundo a Agência Nacional de Saúde (ANS), já são 50,9 milhões de beneficiários nos planos médico-hospitalares.

Assim como é crescente o número de pessoas que hoje possuem o plano de saúde, os custos para as operadoras também vem aumentando em proporções assustadoras. Segundo Miranda (2003) fatores como envelhecimento da população, aumento da utilização dos serviços, avanço tecnológico na área médica e incorporação de novos procedimentos faz com que haja uma elevação no custo real assistencial dos planos de saúde, chamado pelos especialistas de “inflação médica”. Claro que esses dois últimos fatores trazem enormes benefícios aos usuários dos planos de saúde, não podendo ser considerados somente como agravantes de custo, uma vez que eles proporcionam também uma melhor e mais qualificada assistência médico odontológica.

Ainda nesse raciocínio, Miranda cita também que a elevação dos custos aliada à oscilação do risco em torno de uma média e a incidência dos casos chamados pelas operadoras de “catastróficos” (sendo aqueles imprevisíveis e com custo extremamente alto) distorcem as previsões atuariais de custos, dificultando ainda mais o gerenciamento dos gastos das operadoras com a assistência.

Além dos fatores citados, a má utilização do plano também é uma variável importante no impacto do custo assistencial. Observa-se no mercado as seguintes situações: beneficiários que utilizam o plano de forma indiscriminada realizando procedimentos que às vezes nem são indicados ou ainda fazendo consultas com várias especialidades médicas para tratar de um

problema específico, há aqueles que não utilizam e que futuramente terão custos elevados em função da não realização de serviços preventivos de saúde e ainda aqueles que são portadores de doenças crônicas, sendo esses últimos os casos mais preocupantes para a operadora.

A questão é que, não somente para as operadoras, mas para as empresas que contratam os planos de saúde para os seus funcionários e dependentes e também para os que contratam os planos individuais para seus familiares, sofrem com o custo crescente, uma vez que eles são repassados na mensalidade dos planos, seja pelo reajuste dos contratos empresariais seja pelo índice indicado pela ANS. Dessa forma, a assistência médica tem se tornado um benefício/serviço insustentável financeiramente para os contratantes.

Diante disso, é necessário aprimorar a gestão dos gastos com saúde, utilizando ferramentas eficazes que possam apoiar na tomada das melhores decisões e das negociações com os prestadores de serviço, de forma que se gaste menos e da melhor forma possível, a fim de garantir além da satisfação do beneficiário. É importante então conhecer a fundo as variáveis que compõem esses gastos para que seja possível minimizar o custo e otimizar a prestação de serviço. Martin citado por Martins e outros complementa que,

As informações para o uso do custo por procedimento médico são requeridas para diferentes objetivos, dentre estes o estabelecimento de preços de cada procedimento médico; a estimativa de custos para negociação de pacotes; a estimativa de rentabilidade das especialidades e procedimentos médicos; a determinação das margens brutas associadas às especialidades e procedimentos médicos; e, o confronto de faturamento e custo do procedimento médico. (MARTINS, 2010, p.60)

As técnicas estatísticas de Análise Multivariada irão auxiliar na identificação dessas variáveis que constituem os custos da operadora e também em agrupá-las de acordo com as suas características semelhantes (correlações próximas) para que os objetivos citados acima sejam melhor atendidos.

4 ANÁLISE MULTIVARIADA

No universo empresarial, são frequentes as situações em que os gestores se veem obrigados a tomarem decisões estratégicas e muitas das vezes sem embasamentos técnicos para tal. A não utilização de ferramentas adequadas pode prejudicar a escolha das melhores decisões, trazendo prejuízo aos negócios. Essas decisões envolvem sempre muitas variáveis, sendo de extrema importância entender as relações existentes entre elas e também identificar aquelas de maior expressão (ou relevância).

A Análise Estatística Multivariada consiste em um conjunto de técnicas estatísticas sendo que algumas delas podem ser utilizadas na gestão de custos. Segundo Pereira (1999) citado por Smaka (2010), “trata-se de qualquer abordagem analítica que considera o comportamento de mais de duas variáveis simultaneamente” (SMAKA, 2010, p.28).

Os métodos de estatística multivariada são utilizados, segundo Mingoti (2007) com o objetivo de propiciar um melhor entendimento do assunto estudado utilizando a construção de índices ou de variáveis alternativas que simplifiquem os dados originais apresentados.

A utilização dessa metodologia possibilitará a identificação das variáveis responsáveis pelas variações no custo assistencial sem que ocorra perda significativa das informações.

4.1 Análise de Componentes Principais

Uma das técnicas da Análise Multivariada é a Análise de Componentes Principais (ACP), que, segundo Mingoti, tem como objetivo principal “explicar a estrutura da variância e covariâncias de um vetor aleatório, composto por p variáveis aleatórias, através de combinações lineares, utilizando como base, as variáveis originais” (MINGOTI, 2007, p.59).

Ainda nessa mesma linha, Miota (2004) citado por Smaka (2010) define a Análise de Componentes Principais como um “método para transformar variáveis correlacionadas em outro grupo de variáveis não correlacionadas” (SMAKA, 2010, p.32).

Nesse mesmo sentido Neto e Moita afirmam que,

Este método permite a redução da dimensionalidade dos pontos representativos das amostras, pois, embora a informação estatística presente nas n -variáveis originais seja a mesma dos n componentes principais, é comum obter em apenas 2 ou 3 das primeiras componentes principais mais que 90% desta informação. (NETO; MOITA, 1997, p.468.)

A ACP possui propriedades importantes, tais como:

- Cada componente é uma combinação linear das variáveis originais;
- São não correlacionadas
- São estimadas com o objetivo de conservar o máximo de informação em relação à variação total contida nos dados.

As componentes principais são combinações lineares de p variáveis originais, $X_1, X_2, X_3, \dots, X_p$, sendo $\mathbf{X}=[X_1, X_2, X_3, \dots, X_p]'$ o vetor aleatório p dimensional.

As componentes principais são estimadas a partir da observação de uma amostra aleatória de tamanho n do vetor aleatório \mathbf{X} de interesse.

O início da construção das componentes principais amostrais se dá com o cálculo da matriz de covariâncias amostral de vetor aleatório \mathbf{X} (denominada de $\mathbf{S}_{p \times p}$) e a extração dos seus autovalores e autovetores normalizados correspondentes. Denota-se matriz de covariância amostral como

$$\widehat{Cov}(\mathbf{X}) = S_{p \times p} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \cdots & \hat{\sigma}_{1p} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \cdots & \hat{\sigma}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{p1} & \hat{\sigma}_{p2} & \cdots & \hat{\sigma}_{pp} \end{bmatrix}$$

sendo, $\hat{\sigma}_{ij} = S_{ij} = \frac{1}{1-n} \sum_{k=1}^n (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j)$, $i \neq j$ que é um estimador não viciado da covariância populacional entre as variáveis X_i e X_j . Para $i = j$, $S_{ij} = S_{ii} = S_i = \hat{\sigma}_i$, que é a variância amostral da variável X_i , $\hat{\mu}_i = \bar{X}_i$, $\hat{\mu}_j = \bar{X}_j$ são respectivamente as médias amostrais de X_i e X_j e são estimativas não-viciadas das médias populacionais μ_i, μ_j .

A j -ésima componente principal é então estimada como:

$$\hat{Y}_j = \hat{e}_j' \mathbf{X} = \hat{e}_{j1} X_1 + \hat{e}_{j2} X_2 + \cdots + \hat{e}_{jp} X_p \quad j=1,2,\dots,p$$

Sendo $\hat{e}_j' = [e_{j1} \ e_{j2} \dots \ e_{jp}]$ o autovetor normalizado correspondente ao autovalor

$\hat{\lambda}_j$, $j=1,2,\dots,p$, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p$, da matriz de covariâncias amostral $\mathbf{S}_{p \times p}$, e

$\mathbf{X} = [X_1, X_2, X_3, \dots, X_p]'$ o vetor aleatório em estudo, conforme explicado por Mingoti (2007).

A análise da importância de cada componente estimada se dá com o cálculo da proporção da variância total explicada pela componente, sendo a razão entre a sua variância e a variância total, conforme fórmula denotada a seguir:

$$\frac{\hat{\lambda}_j}{\sum_{j=1}^p \hat{\lambda}_j}, \quad j = 1, 2, \dots, p$$

Quanto maior for a proporção da variância total explicada pela componente maior é a sua importância. Se a variância explicada pelas primeiras componentes conseguirem explicar uma considerável parte da variância total, toda a análise pode ser limitada aos vetores aleatórios composto por essas primeiras componentes.

Ainda sobre a questão da escolha da quantidade de componentes, Ferreira (1996) citado por Smaka diz que

Não existe uma resposta definitiva sobre o número de componentes a ser retido. Os aspectos que devem ser considerados incluem a quantidade da variação amostral explicada, o tamanho relativo dos autovalores e a interpretação subjetiva das componentes. (SMAKA, 2010, p.43)

A variância total, segundo Mingoti (2007) é definida como a soma de todas as variâncias amostrais das p variáveis, ou melhor, como sendo a soma dos elementos da diagonal principal da matriz $\mathbf{S}_{p \times p}$. (ou a soma de todos os autovalores da matriz $\mathbf{S}_{p \times p}$).

Uma questão muito importante a ser considerada nessa metodologia, é que, se houver uma discrepância entre as variâncias das variáveis estudadas, os coeficientes estimados nas componentes principais serão afetados e por sua vez a componente ficará dominada numericamente pelas variáveis de maiores variâncias. Isso pode ocorrer quando as variáveis possuem escalas ou unidades de medidas diferentes não podendo ser comparadas dessa forma.

Esse efeito pode ser minimizado fazendo a padronização das variáveis, transformando-as de maneira que essas variáveis agora passem a ter média zero e variância igual a 1.

Dessa forma a análise de componentes principais é estimada considerando-se a matriz de covariâncias do vetor aleatório $\mathbf{Z} = (Z_1, Z_2, \dots, Z_p)'$, em que $Z_i = \frac{X_i - \hat{\mu}_i}{\hat{\sigma}_i}$, e a matriz de correlação do vetor aleatório \mathbf{X} , $\hat{\mu}_i$ e $\hat{\sigma}_i$ representam a média e o desvio-padrão amostrais da variável aleatória, X_i , $i=1,2,\dots,p$, respectivamente. Observa-se que a matriz de covariâncias do vetor aleatório \mathbf{Z} é a matriz de correlação do vetor aleatório \mathbf{X} , dessa forma a análise de componentes principais é realizada pela matriz de correlação do vetor aleatório \mathbf{X} que contém as informações das variáveis sem padronização. Como no caso da matriz de covariâncias, as componentes principais são estimadas a partir da observação de uma amostra aleatória de tamanho n do vetor aleatório \mathbf{X} de interesse. As variáveis originais são padronizadas para que as componentes sejam construídas. A matriz de correlação amostral $\mathbf{R}_{p \times p}$ é denotada por:

$$\mathbf{R}_{p \times p} = \begin{bmatrix} 1 & \hat{\rho}_{12} & \hat{\rho}_{13} & \dots & \hat{\rho}_{1p} \\ \hat{\rho}_{21} & 1 & \hat{\rho}_{23} & \dots & \hat{\rho}_{2p} \\ \hat{\rho}_{31} & \hat{\rho}_{32} & 1 & \dots & \hat{\rho}_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}_{p1} & \hat{\rho}_{p2} & \hat{\rho}_{p3} & \dots & 1 \end{bmatrix}$$

sendo, $\hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}} = \frac{\hat{\sigma}_{ij}}{\hat{\sigma}_i\hat{\sigma}_j}$

Os passos para a construção das componentes principais amostrais pela matriz de correlação amostral $\mathbf{R}_{p \times p}$ do vetor aleatório \mathbf{X} segundo Mingoti (2007) são:

- 1- Cálculo da matriz de correlação amostral $\mathbf{R}_{p \times p}$;

- 2- Cálculo dos autovalores da matriz $\mathbf{R}_{p \times p}$;
- 3- Cálculo dos autovetores correspondentes normalizados.

As componentes principais estimadas serão:

$\hat{Y}_j = \hat{e}'_j Z = \hat{e}_{j1} Z_1 + \hat{e}_{j2} Z_2 + \dots + \hat{e}_{jp} Z_p$, $j=1,2,\dots,p$ sendo $\hat{e}'_j = [e_{j1} \ e_{j2} \dots \ e_{jp}]$ o autovetor normalizado correspondente ao autovalor $\hat{\lambda}_j$ da matriz de correlação amostra $\mathbf{R}_{p \times p}$

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

O esquema descrito na Figura 1 ilustra o método de construção das componentes principais.

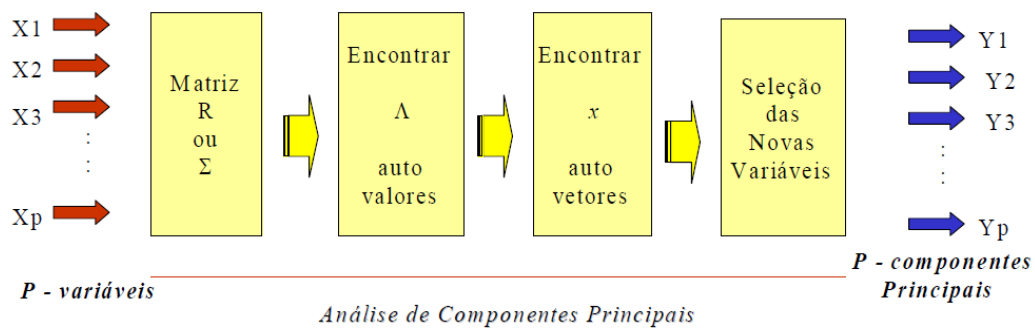


Figura 1: Esquema de aplicação da análise de componentes principais

Fonte: VICINI, Lorena (2005, p.30)

Da mesma forma como na estimação das componentes principais pela matriz de covariâncias, na análise considerando a matriz de correlação, a variância total explicada pela componente é dada por:

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

sendo p a soma das variâncias de todas as p variáveis padronizadas, o que é equivalente à soma de todos os p autovalores da matriz $\mathbf{R}_{p \times p}$.

Segundo Mingoti (2007), em todos os casos, seja na ACP pela matriz de covariâncias ou pela matriz de correlação das variáveis, o objetivo é resumir a informação das p variáveis em um número de componentes principais k , $k < p$.

O uso dessa metodologia no banco de dados dessa monografia possibilitará a identificação das variáveis responsáveis pelas variações no custo assistencial sem que ocorra perda significativa das informações.

4.2 Análise Fatorial

Complementando a Análise de Componentes Principais, tem-se a Análise Fatorial que também auxilia na avaliação do relacionamento existente entre as variáveis. Essa técnica busca identificar quais são as variáveis que podem ser agrupadas em um único fator (grupo).

O objetivo principal dessa técnica é descrever a estrutura das covariâncias do vetor aleatório \mathbf{X} em um número menor de novas variáveis, chamadas de fatores. A Análise Fatorial se fundamenta na suposição de que as variáveis podem ser agrupadas devido a correlação existente entre elas.

Vicini afirma que,

A Análise Fatorial (AF) é uma técnica que é aplicada para identificar fatores num determinado conjunto de medidas realizadas, sendo utilizada, também, como uma ferramenta na tentativa de reduzir um grande conjunto de variáveis para um conjunto mais significativo, representado pelos fatores. Esse método determina quais variáveis pertencem a quais fatores e o quanto cada variável explica cada fator. (VICINI, 2005, p33)

É necessário primeiramente entender as correlações existentes entre as variáveis em estudo.

Mingoti salienta que,

Em linhas gerais, o que se espera é que as variáveis originais X_i , $i=1,2,\dots,p$ estejam agrupadas em subconjunto de novas variáveis mutuamente não correlacionadas, sendo que a análise fatorial teria como objetivo o encontro desses fatores de agrupamento. (MINGOTI, 2007, p.99)

Assim, as variáveis são divididas em grupos considerando as correlações existentes entre elas, de maneira que as variáveis que foram agrupadas no mesmo grupo possuem alta correlação entre si e baixa correlação com as demais variáveis.

O modelo de análise fatorial ortogonal, segundo Ferreira (2008), segue as seguintes premissas:

- Os fatores existentes são variáveis aleatórias e são desconhecidos possuindo média igual a zero e variância igual a 1. Os fatores são não-correlacionados entre si;

- Os fatores são não correlacionados com os erros do modelo. Os erros aleatórios do modelo são independentes, com média igual a zero, e não precisam possuir o mesmo valor da variância.

Em termos matriciais, os vetores que compõem o modelo de análise fatorial ortogonal são descritos por:

$$\mathbf{Z}_{p \times 1} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{bmatrix}; \boldsymbol{\varepsilon}_{p \times 1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}; \mathbf{F}_{m \times 1} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}; \mathbf{L}_{p \times m} = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pm} \end{bmatrix}$$

sendo:

$\boldsymbol{\varepsilon}_{p \times 1}$ o vetor de erros aleatórios

$\mathbf{F}_{m \times 1}$ o vetor dos fatores

$\mathbf{L}_{p \times m}$ a matriz de cargas fatoriais

No modelo de análise fatorial ortogonal cada variável resposta padronizada é descrita pela equação:

$$Z_i = \left(\sum_{j=1}^m l_{ij} F_j \right) + \varepsilon_i, \quad i = 1, 2, \dots, p.$$

Matricialmente, tem-se o modelo geral dado por: $\mathbf{Z} = \mathbf{L}\mathbf{F} + \boldsymbol{\varepsilon}$.

Cada linha da matriz \mathbf{L} possui as cargas fatoriais relativas à variável Z_i , ou seja, cada linha da matriz tem as correlações da variável Z_i com os fatores F_j , $j=1, 2, \dots, m$.

Para estimação do modelo de análise fatorial é necessário inicialmente estimar o valor de m . Posteriormente estima-se a matriz de cargas fatoriais $\mathbf{L}_{p \times m}$.

A matriz $\mathbf{L}_{p \times m}$ contém as cargas fatoriais sendo que o elemento l_{ij} representa a correlação que a variável Z_i tem com o fator F_j , $j=1, 2, \dots, p$.

O valor de m representa o número de fatores incluídos no modelo. A matriz de covariâncias do vetor $\boldsymbol{\varepsilon}$ é diagonal e dada por:

$$\boldsymbol{\Psi}_{p \times p} = \begin{bmatrix} \psi_1 & & & \\ & \psi_2 & & \\ & & \ddots & \\ & & & \psi_p \end{bmatrix}.$$

As variâncias ψ_i são chamadas de variâncias específicas.

Como mencionado anteriormente no modelo de análise fatorial ortogonal cada variável Z_i é expressa pela seguinte formulação:

$$Z_i = \left(\sum_{j=1}^m l_{ij} F_j \right) + \varepsilon_i, \quad i = 1, 2, \dots, p.$$

Sendo assim, a variância de Z_i pode ser expressa como:

$$\text{Var}(Z_i) = 1 = \left(\sum_{j=1}^m l_{ij}^2 \right) + \psi_i = h_i^2 + \psi_i, \text{ sendo } h_i^2 \text{ chamada de comunalidade da variável } Z_i$$

e ψ_i a variância específica.

O objetivo é construir um modelo de análise fatorial tal que h_i^2 seja elevado e ψ_i seja pequeno para cada variável $Z_i, i=1, 2, \dots, p$.

Os passos para a estimação do número de fatores, conforme Mingoti (2007) são:

- 1- Cálculo da matriz de correlação $\mathbf{R}_{p \times p}$;
- 2- Determinação dos autovalores da matriz $\mathbf{R}_{p \times p}$ ordenando-os de forma decrescente;
- 3- Cálculo da proporção da variância total explicada relativa a cada autovalor.

Dois critérios poderão ser utilizados para a estimação do número de fatores segundo Mingoti (2007):

- Análise da proporção da variância total explicada por cada autovalor, em que permanecerão os m autovalores que representam as maiores proporções de explicação da variância total;
- O número m de autovalores que são maiores que 1 (regra de Kaiser - 1958).

Esses critérios servem apenas como referências para a determinação do número de fatores.

Uma vez definida a quantidade de fatores m , estima-se então as cargas fatoriais. Dentre os diversos métodos de estimação existentes, tem-se o de componentes principais onde não é necessário atender a suposição da normalidade das variáveis e o de máxima verossimilhança que pressupõe que a distribuição das variáveis seja normal multivariada, conforme citado por Ferreira (2008).

Ainda na linha desse autor, no método de componentes principais, para cada um m dos autovalores retidos será encontrado o autovetor normalizado correspondente. As cargas fatoriais são estimadas através da seguinte formulação:

$$\hat{l}_{ij} = \sqrt{\hat{\lambda}_j} \hat{e}_{ji}, \text{ sendo:}$$

\hat{l}_{ij} a correlação da componente principal \hat{Y}_j com a variável resposta Z_i , $\sqrt{\hat{\lambda}_j}$ raiz quadrada do autovalor correspondente a componente principal \hat{Y}_j e \hat{e}_{ji} o coeficiente da variável Z_i na componente principal \hat{Y}_j , $j=1,2,\dots,m$; $i=1,2,\dots,p$.

Os fatores são interpretados através da análise das cargas fatoriais. Quanto maior for a correlação da variável resposta com o fator subjacente F_j , maior será a associação do fator com a variável em questão.

Segundo Ferreira (2008) o próximo passo é avaliar a qualidade do modelo (análise das comunalidades e variâncias específicas). Quanto maior a variância explicada pelos fatores (comunalidade) melhor será o modelo ajustado. Através das análises das estimativas \hat{l}_{ij} , agrupa-se as variáveis Z_i em m fatores, de modo que a correlação das variáveis que compõem o fator deve ser elevada com esse fator e pequena com as variáveis que compõem os outros fatores.

Em algumas situações os fatores não ficam muito bem definidos, ou seja, a variável resposta X_i padronizada pode ficar correlacionada a mais de um fator, não ficando evidente a que fator essa variável realmente pertence.

Uma forma de solucionar essa questão é utilizar a rotação ortogonal dos fatores. Segundo Mingoti (2007) essa técnica realiza transformações nos eixos originais F_j , $j=1,2,\dots,m$ de modo a criar novos eixos F^*_j , $j=1,2,\dots,m$, ortogonais. Um dos métodos de rotação é a Varimax que busca a matriz de rotação ortogonal que maximiza as variâncias das cargas fatoriais em cada coluna da matriz de carga fatoriais $\hat{L}_{p \times m}$.

Vale ressaltar que a rotação ortogonal não altera os valores iniciais de comunalidades e variâncias específicas, não sendo utilizada para melhoria do ajuste do modelo.

Por fim faz-se a interpretação dos fatores construídos.

5 BANCO DE DADOS

Nesse estudo será utilizada a base de dados de uma operadora de plano de saúde, cujo nome não será citado e suas informações deverão ser mantidas em sigilo.

O período de análise considerado para este estudo será de Janeiro de 2013 a Setembro de 2014, por data de competência. Define-se como competência a data em que os atendimentos realizados pelos prestadores de serviço são pagos pela operadora, sendo prestadores as Clínicas Médicas e de Terapias, Laboratórios e Hospitais.

Vale ressaltar que a data de competência não é necessariamente a data em que o serviço foi prestado, uma vez que, contratualmente o prestador de serviço tem até noventa dias para enviar para a Operadora de Plano de Saúde as informações de cobrança dos eventos. Sendo assim uma competência de pagamento pode conter datas de atendimentos de diferentes meses, respeitando o prazo citado acima. Devido a isso, não é possível estudar a sazonalidade temporal, uma vez que os custos em cada competência de pagamento são influenciados pela quantidade de serviços enviados pelos prestadores para o faturamento, considerando que esses prestadores podem acumular contas de até três meses da data em que o evento ocorreu.

Esse trabalho de monografia consiste em identificar os pesos das variáveis que compõem o custo total do plano de saúde, segregadas por regime de atendimento: Ambulatorial e Internação. Define-se como regime ambulatorial os atendimentos relacionados à problemas de saúde de pacientes e que não precisam de internação e nem estar acamados. Como regime de internação entende-se por sua vez o ato de tratar o problema de saúde do paciente em um hospital por um período superior a 1 dia.

Em relação ao regime Ambulatorial, as variáveis estudadas serão: Consultas, Exames, Honorários, Materiais, Medicamentos, Pacotes, Procedimentos não Médicos e Taxas. Já no regime de internação as análises serão feitas com base nas seguintes variáveis: Diárias, Exames, Honorários, Materiais, Medicamentos, Pacotes, Procedimentos não Médicos e Taxas.

O Quadro 1 apresenta as definições das variáveis estudadas por regime de atendimento e a nomenclatura dada na base de dados.

Quadro1: Conceituação das variáveis utilizadas no estudo

NOMENCLATURA DA VARIÁVEL	DEFINIÇÃO
Consulta	Atendimento médico realizado com o objetivo de investigar eventuais doenças existentes, geralmente feitas nos consultórios e nos pronto-atendimentos dos hospitais.
Diária	Serviço cobrado devido a cada dia de permanência de um paciente internado em regime hospitalar.
Exame	Procedimentos diagnósticos e/ou de acompanhamento realizados com o objetivo de averiguar possíveis patologias.
Honorário	Valores pagos aos médicos pelos serviços.
Material	Materiais utilizados nos atendimentos médicos, seja no regime ambulatorial, seja no regime de internação (ex: agulhas, seringas, curativos, etc.)
Medicamento	Produtos farmacêuticos utilizados na cura das patologias ou na prevenção das mesmas.
Pacote	Conjunto de serviços que são agrupados, sendo utilizados para a realização de algum procedimento médico (Ex: Pacote de parto: contempla as diárias, os materiais e os medicamentos que normalmente são utilizados para a realização desse procedimento).
Procedimento não médico	Procedimentos realizados pelos demais profissionais da saúde, não dependendo do profissional médico para a sua realização (ex: fisioterapia, sessão de psicologia, sessão de nutrição, etc.)
Taxas	Valores cobrados na utilização de serviços para a realização de determinados procedimentos, tais como utilização de sala de cirurgia, equipamentos, etc.

Fonte: Elaborada pela autora

Na base de dados, essas variáveis estão dispostas seguindo a nomenclatura mostrada acima. Somente a variável consulta é que não está presente na base custo do regime de internação, considerando que esse procedimento não é realizado em pacientes internados.

6 ANÁLISE DOS RESULTADOS

Nesta seção são apresentados os resultados das análises estatísticas realizadas considerando-se os regimes Ambulatorial e Internação.

6.1 Estatística descritiva das variáveis

O estudo foi feito com base nas nove variáveis já descritas na seção 5 segregadas por regime de atendimento (ambulatorial e internação), sendo que cada uma delas possui 21 observações.

Analisando o perfil de custo da operadora observa-se que maior parte dos recursos financeiros da empresa são destinados a pagar os atendimentos do regime ambulatorial, representando 62% do custo total com os atendimentos médicos. Em ambos os regimes é possível ver um aumento crescente no custo da operadora, conforme mostra o Gráfico 1. O aumento do custo em 2014 é maior que 2013, uma vez que os planos de saúde da operadora foram regulamentados no final de 2013, onde havia restrição em relação aos serviços cobertos pelo plano. Vale ressaltar que a carteira de beneficiários está se mantendo estável ao longo do tempo.

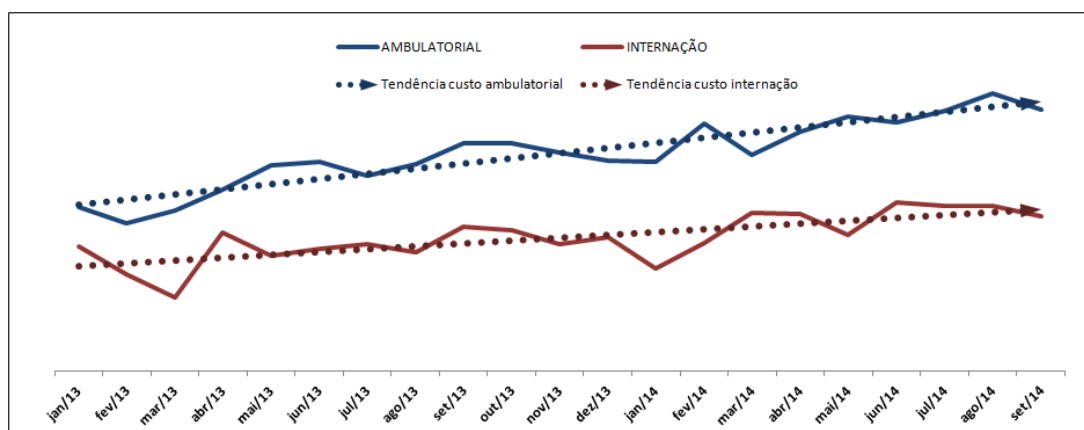


Gráfico 1: Evolução do custo mensal por regime de internação

Em relação às variáveis de custo que compõem o regime de atendimento ambulatorial, observa-se que Exame e Consulta são as de maior representatividade no custo total desse regime, correspondendo a 79% do valor gasto, conforme mostrado no Gráfico 2. Em seguida tem-se os Medicamentos, que consomem 6% desse gasto. Nesse grupo estão os medicamentos utilizados nos tratamentos de Quimioterapia e Radioterapia, sendo esses de alto custo.

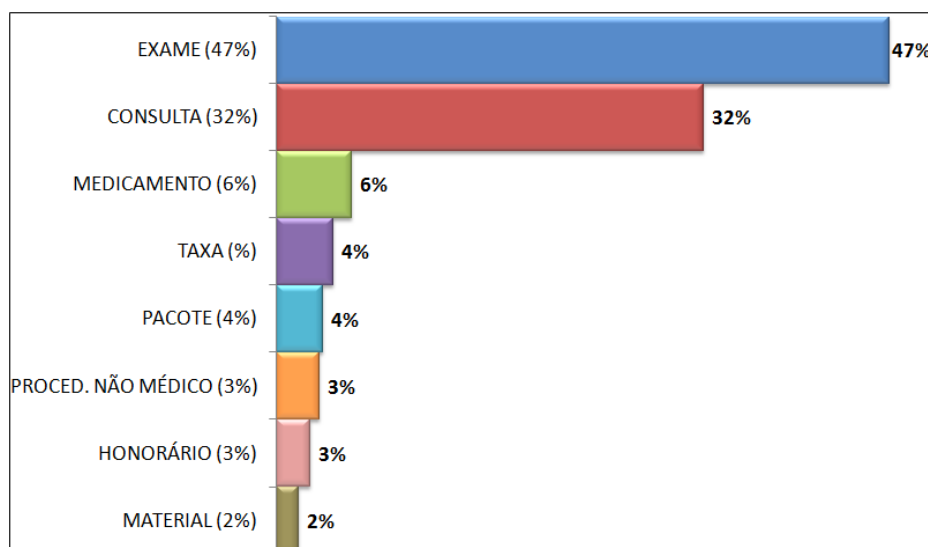


Gráfico 2: Percentual de representatividade de cada variável no custo no regime ambulatorial

O Quadro 2 apresenta as estatísticas descritivas das variáveis de custo que compõem o regime ambulatorial.

As variáveis Exame e Consultas são as que apresentam os maiores custos médios mensais da operadora, R\$ 1.421.567,00 e R\$ 991.380,00, respectivamente, e são também as de maior dispersão. Isso pode ser explicado pelo fato desses serviços serem os mais utilizados pelos beneficiários nesse regime de atendimento. Em média, para cada consulta realizada, são feitos 3,89 exames, segundo pesquisa realizada pela Unidas (União Nacional das Instituições de Autogestão em Saúde).

A variável Material é por sua vez aquela que possui o menor custo médio mensal, R\$ 50.580,00, uma vez que no regime ambulatorial são feitos procedimentos de menor complexidade, não sendo utilizados materiais tão caros nos atendimentos. Porém quando se trata de coeficiente de variação, observa-se que o maior valor é o dessa variável. Mesmo que nesse regime não sejam utilizados os materiais mais caros, a gama de produtos que compõem esse grupo é muito grande e por isso a diferença de preços entre os diversos materiais existentes é significativa.

O mesmo é observado para a variável Medicamento, cujo “leque” de produtos é enorme e ainda, no regime ambulatorial, a utilização dos medicamentos oncológicos (medicamentos caros, conforme já citado) faz com que haja grande variabilidade nos preços pagos para esses itens.

Ainda sob a vertente de variabilidade, a variável Honorário é menor a de menor dispersão, com desvio-padrão de R\$ 13.527,00.

Pode-se notar também uma acentuada discrepância em relação à variância das variáveis estudadas. Os gráficos Box-Plot apresentados na Figura 2 auxiliam na identificação de observações discrepantes em relação às demais (possíveis “outliers”).

Quadro 2: Estatística descritiva das variáveis em regime de atendimento ambulatorial

VARIÁVEL	MÉDIA	DESVIO PADRÃO	COEFICIENTE DE VARIAÇÃO	MÍNIMO	MEDIANA	MÁXIMO
Consulta	991.380,00	177.397,00	17,89	606.960,00	1.010.820,00	1.252.523,00
Exame	1.421.567,00	206.199,00	14,51	1.039.328,00	1.402.895,00	1.861.242,00
Honorário	77.313,00	13.527,00	17,5	48.156,00	75.849,00	104.894,00
Material	50.580,00	30.793,00	60,88	13.510,00	41.695,00	148.559,00
Medicamento	175.173,00	77.165,00	44,05	25.987,00	190.935,00	285.985,00
Pacote	107.988,00	37.165,00	34,42	42985	117.278,00	163.564,00
Procedimento não médico	99.286,00	34.328,00	34,57	47.331,00	107.222,00	147.820,00
Taxa	131.669,00	38.512,00	29,25	97.092,00	115.073,00	230.376,00

As variáveis Material e Taxa apresentaram pontos atípicos em alguns meses de estudo, conforme mostra a Figura 2. Em relação à variável Material esses pontos foram observados nas competências de Julho/2014 e Agosto/2014. Já na variável Taxa os pontos discrepantes foram observados nas competências de Janeiro/13 e Fevereiro/13. Não é possível fazer inferências sobre os motivos para a aparição desses possíveis “outliers” uma vez que o estudo em questão considera a competência de pagamento dos eventos e os prestadores de serviço tem até noventa dias para enviarem as cobranças dos atendimentos para a Operadora, conforme explicado na seção 5.

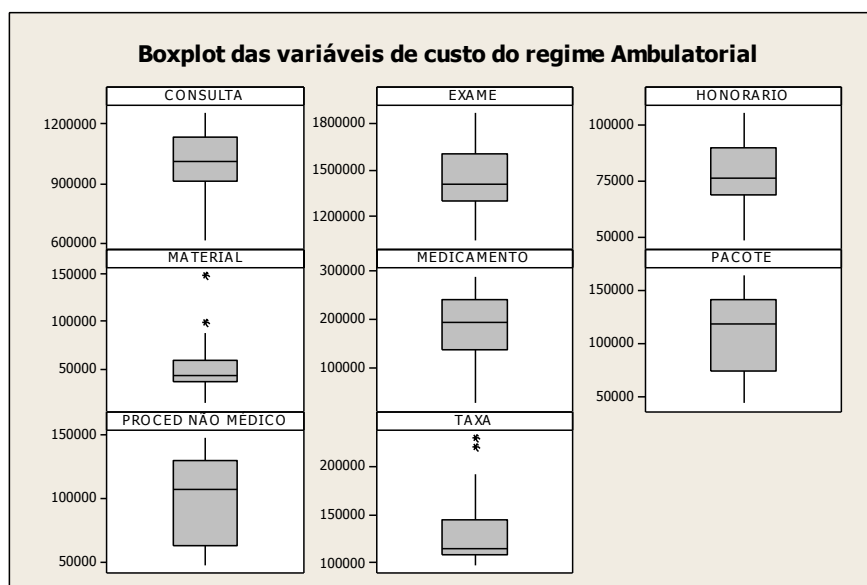


Figura 2: Boxplot das variáveis de custo do regime de atendimento ambulatorial

No Quadro 3, tem-se as correlações amostrais das variáveis do regime ambulatorial. Descritivamente pode ser observado que a maioria das variáveis estão forte e positivamente correlacionadas, destacando as variáveis Procedimento não-Médico e Pacote e Procedimento não-Médico e Exame que possuem as maiores correlações (0,914 e 0,913, respectivamente) e as variáveis Medicamento e Honorário que possuem a menor correlação (0,372), em valor absoluto.

Quadro 3: Correlações amostrais das variáveis de custo do regime ambulatorial

VARIÁVEIS	Consulta	Exame	Honorário	Material	Medicamento	Pacote	Procedimento não médico
Exame	0,825						
Honorário	0,591	0,454					
Material	0,672	0,813	0,392				
Medicamento	0,747	0,727	0,372	0,586			
Pacote	0,730	0,833	0,493	0,667	0,739		
Procedimento não médico	0,753	0,913	0,526	0,725	0,765	0,914	
Taxa	-0,813	-0,747	-0,548	-0,544	-0,853	-0,718	-0,790

Os gráficos de dispersão (ver Figura 3) evidenciam a relação dessas variáveis de maior e menor correlação. Não existe uma explicação de ordem prática que justifique a alta correlação entre as variáveis Procedimento não médico e Pacote, uma vez que o custo da

realização de um desses serviços não intervém no outro. Os pacotes são compostos em sua maioria por diárias, taxas e materiais, não sendo composto por procedimentos não médicos (fisioterapia, fonoaudiologia, psicologia, etc.). Talvez essa alta correlação possa ser explicada pelo gasto muito próximo que se tem com esses dois grupos.

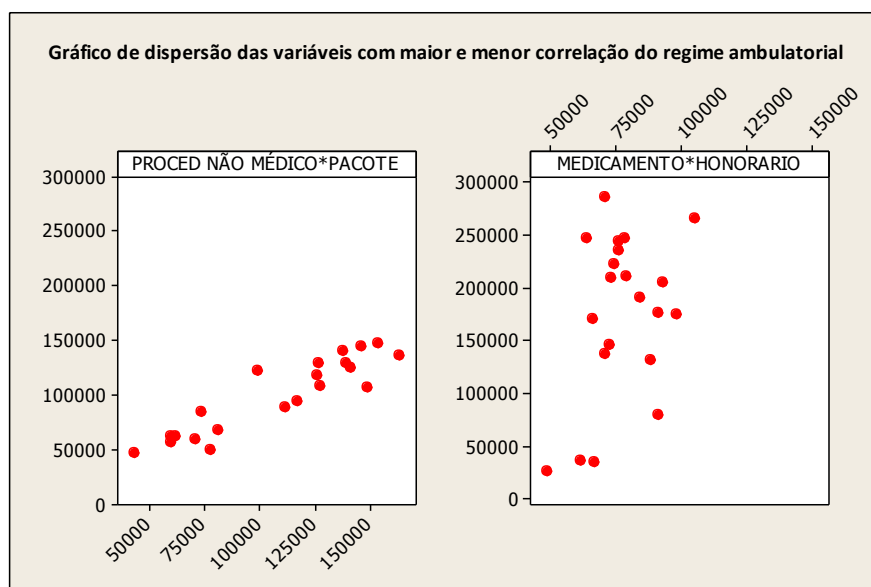


Figura 3. Gráfico de dispersão: Procedimento não-Médicos versus Pacote e Procedimento não médico versus Honorário

Analisando agora a composição do regime de internação, observa-se que a proporção que cada variável representa do custo total está mais distribuída entre as variáveis do que no regime ambulatorial. Na internação, a variável Honorário representa 29% do custo total, seguida de Material que representa 22% e Diária 16%, conforme mostrado no Gráfico 3.

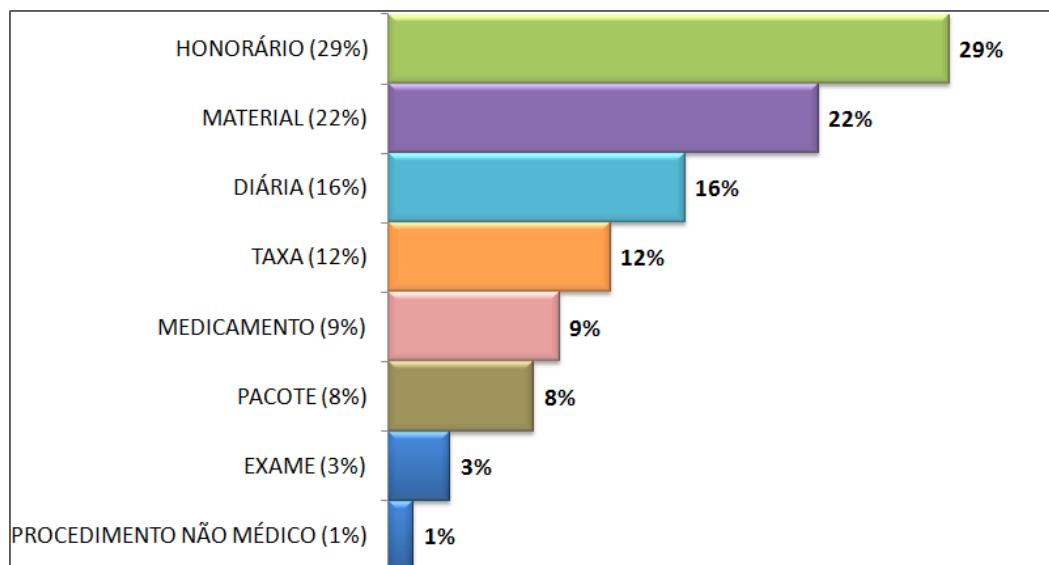


Gráfico 3: Percentual de representatividade de cada variável no custo assistencial do regime de internação

O Quadro 4 apresenta as estatísticas descritivas das variáveis de custo que compõem o regime de internação.

Nesse regime de atendimento, a operadora tem um custo mensal médio de R\$ 547.711,00 com Honorário e de R\$ 419.911,00 com Material (ver Quadro 4). Essas variáveis são também aquelas que possuem maior variabilidade.

A variável Procedimento não-Médico é por sua vez aquela que possui o menor custo médio mensal e menor dispersão, uma vez que nesse regime de atendimento esse serviço não é muito realizado em pacientes internados, à exceção dos casos mais complexos (lembrando que são serviços de fisioterapia, psicologia, fonoaudiologia, etc.).

Assim como na análise do regime ambulatorial, existe uma grande variabilidade entre as variáveis de custo do regime de internação, principalmente em Material, Medicamento e Taxa. Essa variação pode ser explicada pela diferença de custo existente entre esses itens. Como exemplo, tem-se a situação em que um paciente internado devido a problemas de baixa complexidade utiliza materiais mais simples durante o tratamento (ex.: agulha, algodão, curativo, etc.), enquanto que outro paciente, com problemas ortopédicos, precisa de fazer uma cirurgia para se colocar um parafuso, cujo custo é expressivamente superior ao dos materiais simples.

Quadro 4: Estatística descritiva das variáveis de custo do regime de internação

VARIÁVEL	MÉDIA	DESVIO	COEFICIENTE DE VARIACÃO	MÍNIMO	MEDIANA	MÁXIMO
Diária	290.149,00	50.646,00	17,46	186.671,00	309.807,00	359.049,00
Exame	60.620,00	12.782,00	21,09	29.400,00	63.890,00	79.332,00
Honorário	547.711,00	89.000,00	16,25	352.54,00	551.319,00	754.207,00
Material	419.911,00	192.195,00	45,77	104.836,00	369.482,00	885.009,00
Medicamento	166.688,00	76.440,00	45,86	10227	175.695,00	284.677,00
Pacote	142.025,00	34.380,00	24,21	85.910,00	142.704,00	208.113,00
Procedimento não médico	24.605,00	6.331,00	25,73	9.245,00	26.529,00	33.834,00
Taxa	216.989,00	77.099,00	35,53	132.161,00	212.018,00	491.451,00

Através da Figura 4 (gráficos Boxplot), é possível identificar também pontos atípicos nas variáveis Taxa e Medicamento. Assim como nos atendimentos Ambulatoriais, não é possível fazer inferências sobre esses possíveis “outliers” nas contas de internação, considerando que esse estudo considera competência de pagamento dos eventos, conforme problema citado na seção 5.

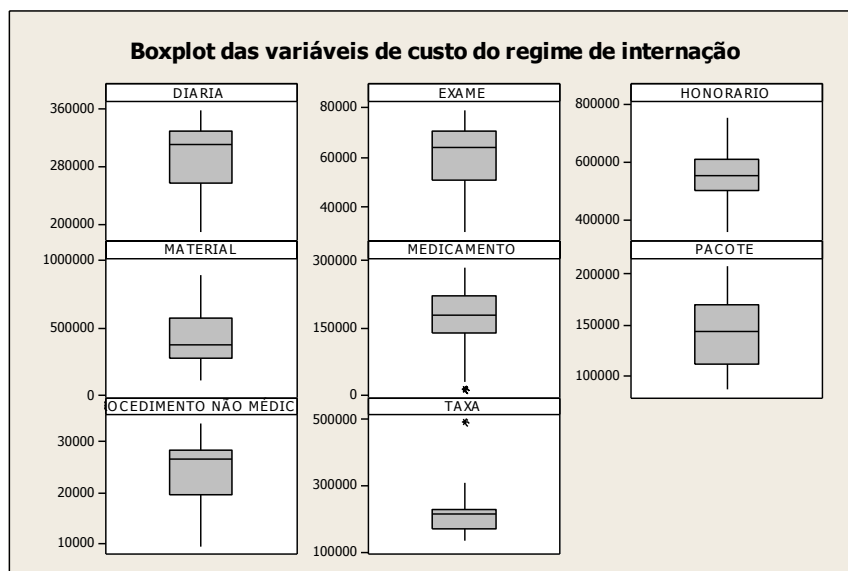


Figura 4: Boxplot das variáveis de custo do regime de atendimento de internação

No Quadro 5, tem-se as correlações amostrais das variáveis do regime de internação. Observa-se que a estrutura das correlações entre as variáveis não tem o mesmo comportamento como mostrado no regime ambulatorial, uma vez que as correlações obtidas na internação não são tão fortes. A maioria das variáveis são positivamente correlacionadas,

destacando-se as variáveis Procedimento não- Médico e Diária e Procedimento não-Médico e Exame, que possuem as maiores correlações (0,775 e 0,755, respectivamente), podendo ser justificada pelo fato de que o paciente quando fica muito tempo internado (utilização de muitas diárias de internação) provavelmente possui um quadro mais grave de saúde necessitando assim de mais cuidados dos profissionais não-Médicos. Como exemplo, tem-se as situações de longos períodos de internação em que são necessárias a realização de sessões de fisioterapia a fim de evitar atrofia dos membros e feridas nos pacientes.

Em relação às variáveis menos correlacionadas (em valor absoluto) tem-se Taxa e Honorário, com correlação positiva de apenas 0,079. Os gráficos de dispersão da Figura 5 mostram a relação dessas variáveis.

Quadro 5: Matriz de Correlação das variáveis de custo do regime de internação

VARIÁVEIS	Diária	Exame	Honorário	Material	Medicamento	Pacote	Procedimento não médico
Exame	0,691						
Honorário	0,390	0,574					
Material	0,146	0,268	0,467				
Medicamento	0,506	0,597	0,586	0,641			
Pacote	0,097	0,127	0,722	0,612	0,326		
Procedimento não médico	0,775	0,755	0,446	0,152	0,376	0,125	
Taxa	-0,167	0,092	0,079	-0,416	-0,536	-0,089	0,135

Os gráficos de dispersão (ver Figura 4) evidenciam a relação dessas variáveis de maior e menor correlação.

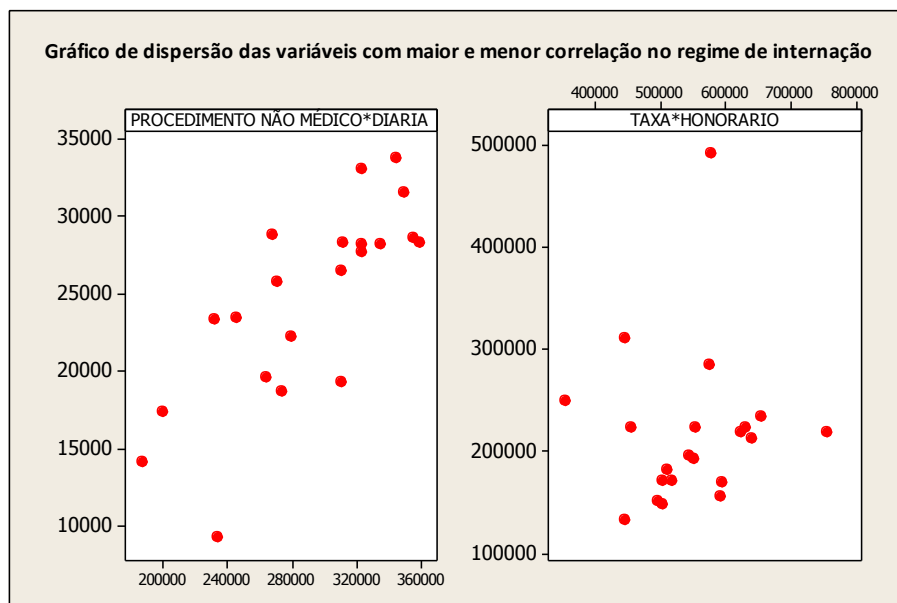


Figura 5. Gráficos de dispersão: Procedimento não-Médico versus Diárias e Taxa versus Honorário

6.2 Análise de Componentes Principais

Nessa seção serão apresentados os resultados da Análise de Componentes Principais construídas a partir das matrizes de Covariâncias e de Correlação.

6.2.1 Análise de componentes principais pela matriz de covariâncias

As componentes principais construídas a partir da matriz de covariâncias permite construir combinações lineares não-correlacionadas das p variáveis originais, sendo possível analisar estatisticamente o comportamento das várias variáveis originais conjuntamente a partir da análise de k componentes principais, $k < p$.

A construção das componentes principais a partir da matriz de covariâncias pode não ser a mais adequada quando há uma discrepância muito acentuada entre as variâncias das variáveis observadas, uma vez que essa discrepância pode afetar os valores dos coeficientes das componentes principais de forma que essas fiquem “dominadas” pelas variáveis de maior variabilidade.

No Quadro 6 estão os resultados apurados através da matriz de covariâncias das variáveis no regime ambulatorial.

Quadro 6: Componentes Principais pela matriz de Covariâncias das variáveis de custo do regime ambulatorial: autovalores e proporção da variância total explicada

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8
Autovalores	74.885.363.264	6.534.903.362	2.479.541.322	459.628.118	360.032.994	246.879.561	91.018.558	58.451.246
Proporção da Variância Total Explicada	0,880	0,077	0,029	0,005	0,004	0,003	0,001	0,001
Proporção acumulada	0,880	0,957	0,986	0,991	0,995	0,998	0,999	1,000

PC1 – PC8 denotam as 8 componentes principais construídas.

Analisando os resultados do Quadro 6 pode-se observar que as duas primeiras componentes explicam 95,7% da variância total. Dessa forma serão utilizadas essas duas componentes de forma a identificar as variáveis que mais impactam no custo do regime ambulatorial. O Quadro 7 mostra os coeficientes (autovetores) dessas componentes.

Quadro 7: Coeficientes das componentes principais pela matriz de Covariâncias das variáveis de custo do regime ambulatorial

VARIÁVEIS	PC 1	PC 2
Consulta	0,608	0,745
Exame	0,729	-0,636
Honorário	0,026	0,042
Material	0,089	-0,081
Medicamento	0,226	0,109
Pacote	0,114	-0,062
Proced. não Médico	0,112	-0,095
Taxa	-0,117	-0,086
Proporção da variância total explicada	0,880	0,077

PC1 – PC2 denotam as 2 componentes principais estudadas

A primeira componente explica 88,0% da variância total e refere-se ao índice global de custo ambulatorial da operadora. O coeficiente de maior grandeza numérica é o da variável Exame na ordem de 0,729, seguido do coeficiente de Consulta (0,608). Quanto maior for o custo da operadora com Exame e Consulta, maior será o custo total ambulatorial. O coeficiente da variável Exame se destaca em relação ao demais uma vez que ela é a que

apresenta a maior variância amostral, conforme apresentado na análise descritiva (ver seção 6.1). A equação da primeira componente é apresentada a seguir:

$$\hat{Y}_1 = 0,608 (\text{consulta}) + 0,729 (\text{exame}) + 0,026 (\text{honorário}) + 0,089 (\text{material}) \\ + 0,226 (\text{medicamento}) + 0,114 (\text{pacote}) \\ + 0,112 (\text{procedimentos não médicos}) - 0,117 (\text{taxa})$$

A segunda componente principal explica 7,7% da variância total e é uma comparação entre os custos de Consulta, Honorário e Medicamento com Exame, Material, Pacote, Procedimentos não Médicos e Taxa, sendo dominada pelas variáveis Consulta, cujo coeficiente é da ordem de 0,745, e Honorário (coeficiente igual a 0,636 em valor absoluto). Considerando que a primeira componente já explica 88% da variância total, a segunda componente se torna de menor relevância. Sua equação é dada como:

$$\hat{Y}_2 = 0,745 (\text{consulta}) - 0,636 (\text{exame}) + 0,042 (\text{honorário}) - 0,081 (\text{material}) \\ + 0,109 (\text{medicamento}) - 0,062 (\text{pacote}) \\ - 0,095 (\text{procedimentos não médicos}) - 0,086 (\text{taxa})$$

Analisando a correlação das variáveis originais do regime ambulatorial com as componentes principais identifica-se que a primeira componente possui correlações fortes e positivas com as variáveis originais, principalmente com Exames e Consultas, exceto com a variável Taxa, exceto com honorário cuja correlação é moderada (0,535). O mesmo não é observado na segunda componente, em que as correlações são bem menores do que a primeira e não significativas (ver Quadro 8).

Quadro 8: Correlações entre as componentes e as variáveis originais (pela matriz de covariâncias) – regime ambulatorial

VARIÁVEIS / COMPONENTES	PC1	PC2
Consulta	0,938	0,339
Exame	0,968	-0,249
Honorário	0,535	0,251
Material	0,789	-0,212
Medicamento	0,800	0,114
Pacote	0,840	-0,135
Proced. não Médico	0,896	-0,223
Taxa	-0,830	-0,181

PC1 – PC2 denotam as 2 componentes principais estudadas

Uma análise similar foi realizada para o regime de internação, conforme resultados expostos no Quadro 9, porém serão estudadas agora as três primeiras componentes que explicam 94,8% da variância total. Os coeficientes das componentes são apresentados no Quadro 10.

Quadro 9: Componentes Principais pela matriz de Covariâncias das variáveis de custo do regime de internação: autovalores e proporção da variância total explicada

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8
Autovalores	43.608.850.795	7.998.544.418	5.831.535.357	1.790.047.931	1.124.858.696	194.957.479	38.840.613	10.072.838
Proporção da Variância Total Explicada	0,720	0,132	0,096	0,030	0,019	0,003	0,001	0,000
Proporção acumulada	0,720	0,852	0,948	0,977	0,996	0,999	1,000	1,000

PC1 – PC8 denotam as 8 componentes principais construídas.

A primeira componente refere-se ao índice global de custo com internação da operadora. O coeficiente de maior grandeza numérica é o da variável Material na ordem de 0,908 que praticamente domina a componente. Quanto maior for o custo da operadora com Material, maior será o custo total com internação. Essa representatividade dessa variável se deve aos materiais de alto custo que são frequentemente utilizados na internação, principalmente em cirurgias ortopédicas. Além disso, o coeficiente dessa variável se destaca em relação ao demais uma vez que ela é a que apresenta a maior variância amostral, conforme apresentado na análise descritiva. Essa componente explica 72,0% da variância total e a sua equação é dada como:

$$\hat{Y}_1 = 0,058(\text{diária}) + 0,021(\text{exame}) + 0,241(\text{honorário}) + 0,908(\text{material}) \\ + 0,272(\text{medicamento}) + 0,106(\text{pacote}) \\ + 0,007(\text{procedimento não médico}) - 0,170(\text{taxa})$$

A segunda componente principal explica 13,2% da variância total e é uma comparação entre os custos de Material em relação às demais variáveis, sendo mais dominada por essa variável e em segundo plano pela variável Taxa (em valor absoluto). Pensando em aplicabilidade essa componente se torna importante na análise, uma vez que ela tem um considerável percentual de explicação da variância total e auxilia no entendimento da relação

existente entre o valor gasto com Material e as outras variáveis de uma internação. Sua equação é apresentada a seguir:

$$\hat{Y}_2 = -0,216(\text{diária}) - 0,080(\text{exame}) - 0,791(\text{honorário}) + 0,200(\text{material}) \\ - 0,158(\text{medicamento}) - 0,147(\text{pacote}) \\ - 0,036(\text{procedimento não médico}) - 0,484(\text{taxa})$$

A terceira componente principal também se torna importante no estudo, uma vez que explica 9,6% da variância total. Pode-se dizer que ela é uma comparação entre os custos de Material, Pacote e Taxa em relação às outras variáveis em questão, sendo essa última variável a que possui o maior coeficiente (0,635); em segundo tem-se Medicamento com 0,568 em valor absoluto. Sua equação é dada por:

$$\hat{Y}_3 = -0,318(\text{diária}) - 0,051(\text{exame}) - 0,098(\text{honorário}) + 0,330(\text{material}) \\ - 0,568(\text{medicamento}) + 0,089(\text{pacote}) \\ - 0,021(\text{procedimento não médico}) - 0,635(\text{taxa})$$

Quadro 10: Autovetores das componentes principais pela matriz de Covariâncias das variáveis de custo do regime de internação

VARIÁVEIS	PC 1	PC 2	PC 3
Diária	0,058	-0,216	-0,381
Exame	0,021	-0,080	-0,051
Honorário	0,241	-0,791	-0,098
Material	0,908	0,200	0,330
Medicamento	0,272	-0,158	-0,568
Pacote	0,106	-0,147	0,089
Proced. não Médico	0,007	-0,036	-0,021
Taxa	-0,170	-0,484	0,635
Proporção da variância total explicada	0,720	0,132	0,096

PC1 – PC3 denotam as 3 componentes principais estudadas.

Complementando a análise, o Quadro 11 apresenta a correlação das variáveis originais do regime de internação com as componentes principais. Identifica-se também que somente a primeira componente possui correlações fortes ou moderadas com as variáveis originais, principalmente com Material, Medicamento e Pacote. Assim como no regime ambulatorial, a variável Taxa possui correlação negativa com a primeira componente.

A segunda componente possui correlações não desprezíveis com Honorário (cuja correlação é alta), Exame, Taxa e Procedimentos não-Médicos (essas três com correlação moderada). Já em relação à terceira componente destaca-se a alta correlação existente com Taxa e Diária. Dessa forma pode-se considerar que tanto a segunda, quanto a terceira componente não são pouco informativas, tendo assim expressividade na análise.

Quadro 11: Correlação entre as componentes e as variáveis originais (pela matriz de covariâncias) – regime de internação

VARIÁVEIS / COMPONENTES	PC1	PC2	PC3
Diária	0,239	-0,381	-0,574
Exame	0,350	-0,561	-0,302
Honorário	0,565	-0,794	-0,084
Material	0,986	0,093	0,131
Medicamento	0,742	-0,185	-0,568
Pacote	0,643	-0,382	0,197
Proced. não Médico	0,216	-0,506	-0,251
Taxa	-0,459	-0,562	0,628

PC1 – PC3 denotam as 3 componentes principais estudadas

Conforme observado em ambas as situações, tanto no regime ambulatorial quanto no regime de internação, a discrepância acentuada entre as variâncias das variáveis estudadas influenciou na apuração dos coeficientes das componentes principais havendo dominância de algumas variáveis em relação a outras.

Com o objetivo de melhorar o modelo, foi feita a padronização das variáveis originais de modo que essas variáveis transformadas passassem a ter variâncias iguais a 1, eliminando ao máximo efeito da discrepância de variância sobre os coeficientes das componentes.

Na seção a seguir estão os resultados da Análise de Componentes Principais através da matriz de correlação.

6.2.2 Análise de componentes principais pela matriz de correlação

Como já citado anteriormente, o efeito da variância amostral no cálculo dos coeficientes das componentes principais estavam influenciando grandemente os resultados apresentados na seção 6.2. Fez-se então, a transformação das variáveis originais utilizando a técnica de padronização, sendo que pelo conceito, a matriz de covariâncias das variáveis padronizadas é a matriz de correlação das variáveis originais.

No Quadro 12 estão os resultados obtidos para o regime de atendimento ambulatorial através da matriz de correlação.

Quadro 12: Componentes Principais pela matriz de Correlação das variáveis do regime ambulatorial

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Autovalores	5,8844	0,7318	0,5660	0,3469	0,2054	0,1426	0,0854	0,0374
Proporção da Variância Total Explicada	0,736	0,091	0,071	0,043	0,026	0,018	0,011	0,005
Proporção acumulada	0,736	0,827	0,898	0,941	0,967	0,985	0,999	1,000

Comparando os resultados obtidos para o regime ambulatorial via matriz de correlação com aqueles obtidos via matriz de covariâncias, observa-se que o percentual da variância total explicada pelas duas primeiras componentes foi menor que quando utilizado a matriz de covariâncias. Nessa nova metodologia essas componentes explicam 82,7% da variância total. Sendo assim seriam necessárias pelo menos quatro componentes para se ter 94,1% de variância total explicada. O Quadro 13 contém os coeficientes (autovetores) dessas componentes selecionadas para estudo.

Quadro 13: Autovetores das componentes principais pela matriz de Correlação das variáveis de custo do regime ambulatorial

VARIÁVEIS	PC 1	PC 2	PC 3	PC 4
Consulta	0,370	-0,135	0,071	-0,450
Exame	0,384	0,216	-0,215	-0,002
Honorário	0,254	-0,885	-0,245	0,058
Material	0,326	0,287	-0,600	-0,432
Medicamento	0,353	0,167	0,532	-0,134
Pacote	0,370	0,115	-0,056	0,616
Proced. não Médico	0,388	0,109	-0,077	0,433
Taxa	-0,364	0,127	-0,487	0,152
Proporção da variância total explicada	0,736	0,091	0,071	0,043

Assim como nos demais casos, a primeira componente também refere-se ao índice global (relativo as médias) de custo ambulatorial da operadora. O coeficiente de maior grandeza numérica passa a ser o da variável Procedimentos não-Médicos na ordem de 0,388,

porém todos os demais coeficientes estão muito próximos, não havendo uma dominância significativa de um em relação a outro. Essa componente explica 73,6% da variância total.

A seguir estão as equações matemáticas dessas quatro primeiras componentes. A letra “p” nas equações abaixo indica a padronização das variáveis (subtração do valor amostral da média e divisão pelo respectivo desvio padrão da variável).

$\hat{Y}_1 = 0,370(\text{consulta}_p) + 0,384(\text{exame}_p) + 0,254(\text{honorário}_p) \\ + 0,326(\text{material}_p) + 0,353(\text{medicamento}_p) + 0,370(\text{pacote}_p) \\ + 0,388(\text{procedimento não médico}_p) - 0,364(\text{taxa}_p)$
$\hat{Y}_2 = -0,135(\text{consulta}_p) + 0,216(\text{exame}_p) - 0,885(\text{honorário}_p) \\ + 0,287(\text{material}_p) + 0,167(\text{medicamento}_p) + 0,115(\text{pacote}_p) \\ + 0,109(\text{procedimento não médico}_p) + 0,127(\text{taxa}_p)$
$\hat{Y}_3 = 0,071(\text{consulta}_p) - 0,215(\text{exame}_p) - 0,245(\text{honorário}_p) \\ - 0,600(\text{material}_p) + 0,532(\text{medicamento}_p) - 0,056(\text{pacote}_p) \\ - 0,077(\text{procedimento não médico}_p) - 0,487(\text{taxa}_p)$
$\hat{Y}_4 = -0,450(\text{consulta}_p) - 0,002(\text{exame}_p) + 0,058(\text{honorário}_p) \\ - 0,432(\text{material}_p) - 0,134(\text{medicamento}_p) + 0,616(\text{pacote}_p) \\ + 0,433(\text{procedimento não médico}_p) + 0,152(\text{taxa}_p)$

No Quadro14 estão as correlações das variáveis originais do regime ambulatorial com as componentes principais pela matriz de correlação. O resultado se repete igualmente como ocorrido quando se utilizado a matriz de covariância: somente a primeira componente possui correlações fortes e positivas com as variáveis originais, exceto para a variável Taxa cuja correlação é alta, porém negativa. Essa correlação forte se destaca principalmente com Procedimentos não Médicos e Exames. Porém na segunda componente somente a variável Honorário possui correlação forte (0,757 em valor absoluto). Já as componentes PC3 e PC4 não possuem correlação importante com as variáveis originais. Dessa forma, sob o ponto de vista prático, elas não são relevantes para o estudo.

Quadro 14: Correlação entre as componentes e as variáveis originais (pela matriz de correlação) – regime ambulatorial

VARIÁVEIS / COMPONENTES	PC 1	PC 2	PC 3	PC 4
Consulta	0,898	-0,116	0,053	-0,265
Exame	0,932	0,185	-0,162	-0,001
Honorário	0,615	-0,757	-0,184	0,034
Material	0,791	0,246	-0,451	-0,254
Medicamento	0,855	0,143	0,400	-0,079
Pacote	0,899	0,099	0,042	0,363
Proced. não Médico	0,941	0,093	-0,058	0,255
Taxa	-0,883	0,109	-0,367	0,089

Uma análise similar foi realizada para o regime de internação, utilizado-se a matriz de correlação para estimar as componentes principais. O Quadro 15 mostra os autovalores e a variância explicada por cada componente. De acordo com os resultados apresentados, serão consideradas as quatro primeiras componentes, sendo que elas explicam 91,3% da variância total. Os coeficientes das componentes assim como as respectivas correlações das componentes com as variáveis originais encontram-se respectivamente nos Quadros 16 e 17.

Quadro 15: Componentes Principais pela matriz de Correlação das variáveis do regime de internação: autovalores e variância explicada por cada componente

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Autovalores	3,7713	1,8058	1,2740	0,4530	0,3665	0,1488	0,1207	0,0598
Proporção da Variância Total Explicada	0,471	0,226	0,159	0,057	0,046	0,019	0,015	0,007
Proporção acumulada	0,471	0,697	0,856	0,913	959,000	0,978	0,993	1,000

Quadro 16: Coeficientes das componentes principais pela matriz de Correlação das variáveis de custo do regime de internação

VARIÁVEIS	PC 1	PC 2	PC 3	PC 4
Diária	0,376	0,329	0,278	-0,501
Exame	0,412	0,336	0,034	0,450
Honorário	0,413	-0,062	-0,427	0,059
Material	0,329	-0,456	-0,025	0,215
Medicamento	0,427	-0,200	0,289	0,380
Pacote	0,283	-0,373	-0,505	-0,458
Proced. não Médico	0,366	0,433	0,026	-0,247
Taxa	-0,120	0,447	-0,632	0,285
Proporção da variância total explicada	0,471	0,226	0,159	0,057

Da mesma forma como já citado anteriormente, a primeira componente também refere-se ao índice global (relativo as médias) de custo com internação da operadora. O coeficiente de maior grandeza numérica passa a ser o da variável Medicamento na ordem de 0,427. Da mesma forma que observado no regime ambulatorial, os demais coeficientes calculados possuem valores muito próximos, não havendo uma dominância significativa de um em relação a outro. Em contrapartida, essa componente passa a explicar apenas 47,1% da variância total, sendo que quando utilizado a matriz de covariâncias para o regime de internação esse percentual chegou a 64,4%. A seguir estão as equações matemáticas das quatro primeiras componentes:

$\hat{Y}_1 = 0,376(diaria_p) + 0,412(exame_p) + 0,413(honorário_p) + 0,329(material_p) + 0,427(medicamento_p) + 0,283(pacote_p) + 0,366(procedimento\ não\ médico_p) - 0,120(taxa_p)$
$\hat{Y}_2 = 0,329(diaria_p) + 0,336(exame_p) - 0,062(honorário_p) - 0,456(material_p) - 0,200(medicamento_p) - 0,373(pacote_p) + 0,4433(procedimento\ não\ médico_p) + 0,447(taxa_p)$
$\hat{Y}_3 = 0,278(diaria_p) + 0,034(exame_p) - 0,427(honorário_p) - 0,025(material_p) + 0,289(medicamento_p) - 0,505(pacote_p) + 0,026(procedimento\ não\ médico_p) - 0,632(taxa_p)$
$\hat{Y}_4 = -0,501(diaria_p) + 0,450(exame_p) + 0,059(honorário_p) + 0,215(material_p) + 0,380(medicamento_p) - 0,458(pacote_p) - 0,247(procedimento\ não\ médico_p) + 0,285(taxa_p)$

No que se refere às correlações das componentes com variáveis originais do regime de internação (ver Quadro 17), os resultados se assemelham aos obtidos através da análise da matriz de covariâncias. Identifica-se que somente a primeira componente possui correlações fortes e positivas com as variáveis originais, principalmente com Medicamento, Honorário e Exame. A variável Taxa continuou apresentando uma correlação negativa com a componente, mas de menor grandeza do que na análise via matriz de covariâncias em valor absoluto (ver Quadro 11).

Já a segunda componente possui correlações não desprezíveis com Material, Taxa, Procedimentos não-Médicos e Pacote (com correlações moderadas). Na componente PC3 destaca-se a alta correlação existente com Taxa e moderada com Diária. Porém, a quarta componente não possui nenhuma correlação importante com as variáveis originais, não sendo importante para o estudo do ponto de vista prático.

Quadro 17: Correlação entre as componentes e as variáveis originais (pela matriz de correlação) – regime de internação

VARIÁVEIS / COMPONENTES	PC 1	PC 2	PC 3	PC 4
Diária	0,729	0,442	0,313	-0,337
Exame	0,799	0,451	0,039	0,303
Honorário	0,802	-0,084	-0,482	0,040
Material	0,638	-0,613	-0,028	0,144
Medicamento	0,829	-0,268	0,326	0,256
Pacote	0,549	-0,502	-0,570	-0,308
Proced. não Médico	0,712	0,582	0,029	-0,166
Taxa	-0,234	0,601	-0,713	0,192

A partir de todas as análises apresentadas até o momento, pode-se concluir que as interpretações obtidas para as componentes principais construídas são muito semelhantes, seja pela matriz de covariâncias ou pela matriz de correlação. Entretanto, a padronização das variáveis originais fez com que a influência da discrepância de variabilidade existente entre essas variáveis, em ambos os regimes de atendimentos estudados, fosse minimizada de forma a tornar os valores dos coeficientes das componentes muito próximos. Dessa maneira, a dominância de algumas variáveis deixou de existir, tanto no regime ambulatorial quanto no regime de internação.

Portanto, para análise das variáveis que mais impactam no custo médico hospitalar será utilizada a análise das componentes principais construídas a partir da matriz de

correlação e a primeira componente será aquela utilizada para compreender a relação entre variáveis e os gastos com saúde da operadora em questão.

É importante destacar que, na maioria das vezes, é necessário a aplicação de mais de uma técnica multivariada para se ter a solução para o problema estudado. Sendo assim a ACP será complementada com a Análise Fatorial.

6.3 Análise Fatorial

Considerando as conclusões obtidas na seção anterior, viu-se que é possível utilizar um índice geral de custo da operadora a partir da Análise de Componentes Principais. Porém, é também de interesse da operadora de saúde construir índices separados por grupos de variáveis, que podem ser obtidos a partir do uso da Análise Fatorial.

Para tal, foi realizada a análise fatorial ortogonal usando o método de componentes principais pela matriz de correlação para estimação das cargas fatoriais, tanto no regime ambulatorial quanto no regime de internação.

O Quadro 18 apresenta as cargas fatoriais estimadas da Análise Fatorial para o regime Ambulatorial, sem rotação e com rotação ortogonal Varimax. Os quatro primeiros fatores calculados explicam 94,1% da variância total.

Quadro 18: Análise Fatorial com e sem rotação para o regime ambulatorial

Cargas fatoriais e comunalidade (sem rotação)					
VARIÁVEL	FATOR 1	FATOR 2	FATOR 3	FATOR 4	COMUNALIDADE
Consulta	0,898	-0,116	0,053	-0,265	0,893
Exame	0,932	0,185	-0,162	-0,001	0,929
Honorário	0,615	-0,757	-0,184	0,034	0,986
Material	0,791	0,246	-0,451	-0,254	0,955
Medicamento	0,855	0,143	0,400	-0,079	0,918
Pacote	0,899	0,099	-0,042	0,363	0,951
Proced. não Médico	0,941	0,093	-0,058	0,255	0,963
Taxa	-0,883	0,109	-0,367	0,089	0,934
Parte da variância total devido ao fator	5,8844	0,7318	0,5660	0,3469	7,5292
% Variância total devido ao fator	0,736	0,091	0,071	0,043	0,941

Cargas fatoriais e comunalidade (com rotação ortogonal Varimax)					
VARIÁVEL	FATOR 1	FATOR 2	FATOR 3	FATOR 4	COMUNALIDADE
Consulta	0,660	0,230	0,500	-0,393	0,893
Exame	0,448	0,550	0,626	-0,183	0,929
Honorário	0,200	0,195	0,150	-0,941	0,986
Material	0,236	0,310	0,885	-0,144	0,955
Medicamento	0,836	0,388	0,255	-0,065	0,918
Pacote	0,395	0,797	0,327	-0,229	0,951
Proced. não Médico	0,443	0,733	0,409	-0,250	0,963
Taxa	-0,824	-0,338	-0,210	0,311	0,934
Parte da variância total devido ao fator	2,4616	1,9268	1,8302	1,3106	7,5292
% Variância total devido ao fator	0,308	0,241	0,229	0,164	0,941

Observa-se que os fatores construídos a partir da rotação ortogonal Varimax apresentaram uma melhor distribuição das cargas fatoriais entre os fatores. Além dos fatores estarem melhor definidos, as variáveis respostas (Z_i) estão correlacionadas a somente um fator, e não mais como apresentado na solução sem a rotação. Isso permite com que se tenha uma melhor definição dos grupos de variáveis formadas pelos fatores.

Observa-se que o primeiro fator (F_1), ficou composto pelas variáveis Medicamentos, Taxa e Consulta, sendo que essas duas últimas possuem características semelhantes, uma vez que a cada consulta realizada no pronto atendimento dos hospitais são cobradas taxas referentes a esses atendimentos. O Segundo fator (F_2) ficou composto pelas variáveis Pacotes e Procedimentos não Médicos. Pelo mesmo motivo dito na análise da correlação (custo mensal muito próximo entre as variáveis), essas duas variáveis ficaram no mesmo grupo. Segundo especialista em Saúde da Operadora não existe uma justificativa prática sobre a relação existente entre essas duas variáveis. O terceiro fator (F_3) foi composto pelas variáveis Material e Exame, indicando o vínculo existente entre elas, sendo que sempre são utilizados materiais na realização de qualquer exame e o quarto fator (F_4) é composto somente pela variável Honorário.

O modelo encontra-se bem ajustado, pois apresentou valores altos para as comunalidades (parte da variância de Z_i que é explicada pelos fatores). A menor comunalidade observada foi para a variável Consulta (0,893), mas ainda assim é um valor consideravelmente alto e próximo a um.

Com o objetivo de melhorar a composição dos grupos do ponto de vista prático foi feita a simulação da Análise fatorial com rotação ortogonal considerando apenas 2 fatores ($m=2$). Essa solução foi testada uma vez que, pela Análise de Componentes Principais, observou-se não existe nenhuma correlação significativa das variáveis originais do regime ambulatorial com a terceira e a quarta componente principal estimadas via matriz de correlação.

Porém os resultados apresentados não foram satisfatórios, pois houve redução no percentual da variância total explicada, passando de 0,941 para 0,827, e redução também nas comunalidades, principalmente na variável Material, cuja comunalidade na primeira simulação foi de 0,955 e passou para 0,687. Além disso o primeiro fator (F_1) foi composto por praticamente todas as variáveis, à exceção de Honorário que ficou isolada no segundo fator (F_2).

O fato então de se usar $m=4$ traz um ganho para a análise uma vez houve a redução de 8 variáveis para apenas 4, de forma a explicar 94,1% da variância total dos dados.

Semelhante ao apresentado para o regime ambulatorial, o mesmo foi feito para o regime de internação. O Quadro 19 mostra os resultados da Análise Fatorial para esse regime.

Quadro 19: Análise Fatorial com e sem rotação para o regime de internação

Cargas fatoriais e comunalidade (sem rotação)					
VARIÁVEL	FATOR 1	FATOR 2	FATOR 3	FATOR 4	COMUNALIDADE
Diária	0,729	0,442	0,313	-0,337	0,939
Exame	0,799	0,451	0,039	0,303	0,936
Honorário	0,802	-0,084	-0,482	0,040	0,885
Material	0,638	-0,613	-0,028	0,144	0,805
Medicamento	0,829	-0,268	0,326	0,256	0,931
Pacote	0,549	-0,502	-0,570	-0,308	0,973
Proced. não Médico	0,712	0,582	0,029	-0,166	0,874
Taxa	-0,234	0,601	-0,713	0,192	0,961
Parte da variância total devido ao fator	3,7713	1,8058	1,2740	0,4530	7,3041
% Variância total devido ao fator	0,471	0,226	0,159	0,057	0,913

Cargas fatoriais e comunalidade (com rotação ortogonal Varimax)					
VARIÁVEL	FATOR 1	FATOR 2	FATOR 3	FATOR 4	COMUNALIDADE
Diária	0,941	-0,063	0,217	0,054	0,939
Exame	0,724	-0,099	-0,135	0,619	0,936
Honorário	0,366	-0,765	-0,136	0,384	0,885
Material	-0,026	-0,615	0,447	0,476	0,805
Medicamento	0,356	-0,291	0,519	0,671	0,931
Pacote	0,028	-0,983	0,077	-0,017	0,973
Proced. não Médico	0,905	-0,122	-0,122	0,158	0,874
Taxa	0,024	0,009	-0,979	-0,055	0,961
Parte da variância total devido ao fator	2,4915	1,5317	1,5317	1,2389	7,3041
% Variância total devido ao fator	0,311	0,255	0,191	0,155	0,913

Os quatro primeiros fatores calculados explicam 91,3% da variância total.

Assim como no regime Ambulatorial, os fatores construídos a partir da rotação ortogonal Varimax apresentaram melhores resultados também no regime de internação.

O primeiro fator (F_1) ficou composto pelas variáveis Diária, Procedimento não Médico e Exame, o que justifica na prática o fato de quanto mais tempo internado o paciente ficar, mais complicado é o seu estado de saúde e mais exames e cuidados não médicos são exigidos, semelhante à explicação dada em relação à correlação dessas variáveis (ver seção 6.2.2). O Segundo fator (F_2) ficou composto pelas variáveis Pacote, Honorário e Material, guardando o vínculo existente entre essas variáveis de que o pagamento feito aos médicos (honorário) pela realização de determinados procedimentos está, em sua maioria, vinculado a um pacote de procedimentos e a materiais utilizados na execução desses serviços. O terceiro fator (F_3) foi

composto pela variável Taxa e o quarto fator (F_4) é composto somente pela variável Medicamento.

O modelo também encontra-se bem ajustado, com valores altos para as comunalidades. A menor comunalidade observada foi para a variável Material (0,805), mas ainda assim é um valor consideravelmente alto e próximo a um.

Vale ressaltar que a correlação das componentes principais com as variáveis originais (ver Tabelas 14 e 17) são os mesmos aos das Tabelas 18 e 19 para a estimativa das cargas fatoriais sem rotação, uma vez que a estimação foi feita utilizando o método de componentes principais.

Com o mesmo objetivo apresentado no regime Ambulatorial (melhorar a composição dos grupos do ponto de vista prático) foi feita a simulação da Análise fatorial com rotação ortogonal no regime de Internação, considerando 3 fatores ($m=3$). Através da Análise de Componentes Principais, observou-se não existe nenhuma correlação significativa das variáveis originais do regime internação com a quarta componente principal estimada via matriz de correlação, por isso a sugestão de simular com $m=3$.

No regime de internação houve também uma redução no percentual da variância total explicada pelos fatores, passando de 0,913 para 0,856. Analisando as comunalidades observa-se também uma redução, mas todas elas continuaram apresentando valores altos (próximos a 1). Em relação aos novos fatores formados, a única alteração ocorrida em relação ao resultado apresentado utilizando 4 fatores, foi que a variável Taxa (que estava sozinha no fator F_3) foi agrupada à variável Medicamento (que também estava sozinha, porém no fator F_4), porém, essas duas variáveis não guardam nenhuma relação prática para permanecerem agrupadas. Dessa forma a simulação feita considerando os quatro fatores ainda é a melhor solução para a Análise Fatorial no regime de Internação.

A Análise Fatorial se mostrou eficiente tanto para o regime ambulatorial quanto para o regime de internação, conseguindo separar os grupos de acordo com a realidade do mercado.

Tal análise se torna importante, principalmente para as negociações de tabelas com os prestadores, podendo criar formas de negociações de acordo com os grupos de variáveis formados e apresentados nessa seção.

7 Considerações finais

A Análise de Componentes Principais mostrou ser uma técnica muito útil para este estudo de caso de uma operadora de plano de saúde, cujo objetivo era identificar as variáveis que mais impactam no custo total do plano. A partir de sua utilização foi possível resumir os dados das 8 variáveis originais em um número menor de variáveis (2 a 3 componentes) sem perda de informação relevante do conjunto de dados original.

Essa técnica, estimada primeiramente pela matriz de covariâncias das variáveis originais, mostrou que 95,7% da variação total foi explicada pelas 2 primeiras componentes no regime ambulatorial e no regime de internação as 3 primeiras componentes explicaram 91,3% da variância total.

Porém como a matriz de covariâncias leva em consideração as variâncias das variáveis, a discrepância acentuada entre as variâncias dessas variáveis influenciou na apuração dos coeficientes das componentes principais havendo dominância de algumas variáveis em relação a outras. No caso do regime ambulatorial a primeira componente ficou dominada pela variável Exame, cujo valor do coeficiente foi de 0,729, e no regime de internação pela variável Material com coeficiente de 0,908.

Com o objetivo de melhorar o modelo evitando que a variável original de maior variância dominasse a componente principal, foi feita a estimação das componentes principais através da matriz de correlação em ambos os regimes de atendimento.

Os resultados obtidos nessa segunda estimativa se mostraram mais satisfatórios. A padronização das variáveis originais fez com que a influência da discrepância de variabilidade existente entre essas variáveis, em ambos os regimes de atendimentos estudados, fosse minimizada de forma a tornar os valores dos coeficientes das componentes muito próximos.

As variáveis de maior impacto no custo total da operadora foram escolhidas através da primeira componente principal, critério utilizado levando em consideração que essa é a componente que mais explica a variância total dos dados e em suma representa o índice global do custo.

Conforme apresentado ao longo desse estudo as variáveis do regime ambulatorial que mais impactam no custo da operadora são Procedimentos não-Médicos, Exames e Consultas. Observa-se que, pelo Gráfico 2 (página 23), as variáveis que apareciam como mais representativas eram somente Exame e Consulta. Com os resultados da análise das componentes principais é necessário também dar atenção para a variável Procedimentos não-

Médicos, mesmo que hoje ela não esteja no topo dos procedimentos mais representativos do custo (conforme Gráfico 2). Diante disso é necessário e muito importante avaliar as estratégias de negociação com os prestadores desses serviços a fim de obter melhores resultados financeiros, e assim melhor gestão do custo ambulatorial, lembrando que o custo desse regime representa 62% do custo total da operadora.

Em relação regime de internação observou-se que as variáveis mais representativas no custo foram Medicamento, Honorário e Exame. Esse resultado não é muito comum, quando se comparado ao mercado da saúde suplementar e também quando se analisa o Gráfico 3, onde as variáveis Honorário, Material e Diária eram as variáveis que mais representavam no custo. Normalmente as variáveis que mais impactam são Material e Medicamento, itens muito caros quando utilizados em pacientes internados na alta complexidade. A Operadora explicou que como houve uma reformulação dos produtos de saúde, até outubro de 2013, alguns tipos de internação não eram autorizadas pela operadora. Dessa forma a utilização de medicamentos e materiais mais caros era restrita, fazendo com que os outros itens que compõem a internação se destacassem.

Complementando a análise de componentes principais foi realizada também a análise fatorial com o propósito de construir índices separados por grupos de variáveis. Essa técnica se mostrou muito eficiente para ambos os regimes de atendimento, separando grupos de variáveis que estrategicamente poderão ser negociados em conjunto, trazendo assim redução no custo da operadora. Cabe salientar que com o uso da Análise Fatorial foi possível resumir os dados originais das 8 variáveis estudadas nessa monografia, em apenas 4 grupos, considerando a correlação que as variáveis tinham entre si, obtendo-se uma boa qualidade de ajuste do modelo, como apresentado nas páginas 41 e 43.

Portando, o estudo se mostrou eficiente de acordo com os objetivos propostos, permitindo identificar as variáveis que mais impactam no custo da operadora e os pesos de cada uma na composição do custo, utilizando as técnicas de componentes principais e análise fatorial como ferramentas de gerenciamento de custo e subsídios para negociações com os prestadores.

REFERÊNCIAS

Agência Nacional de Saúde. **Caderno de saúde suplementar de Setembro de 2014.**

Disponível em

<http://www.ans.gov.br/images/stories/Materiais_para_pesquisa/Perfil_setor/Caderno_informacao_saude_suplementar/2014_mes09_caderno_informacao.pdf> Acesso em 6 de outubro de 2014.

FERREIRA, D.F. **Estatística multivariada.** Lavras. Editora Ufla, 2008. 662 p.

MACÊDO, David. **Determinação do custo assistencial em planos de saúde através de simulação.** Disponível em

<http://www.abepro.org.br/biblioteca/ENEGEP2007_TR590445_0022.pdf> Acesso em 28 de setembro de 2014.

MARTINS, Ana Virgínia Feitosa. **Estratégias para redução de custos em operadoras de plano de saúde.** Disponível em

<<http://www.revista.ufpe.br/ricontabeis/index.php/contabeis/article/viewFile/344/320>> Acesso em 18 outubro 2014.

MINGOTI, Sueli Aparecida. **Análise de dados através de métodos de estatística multivariada:** uma abordagem aplicada. Belo Horizonte: Ed. UFMG, 2005. 295p.

MIRANDA, Cláudio da Rocha. **Gerenciamento de Custos em Planos de Assistência à Saúde.** Disponível em

<http://www.ans.gov.br/portal/upload/biblioteca/TT_AS_20_ClaudioMiranda_GerenciamentodeCusto.pdf> Acesso em 28 de setembro de 2014.

MOITA, Graziella Ciaramella; NETO, José Machado Moita. **Uma introdução à análise exploratória de dados multivariados.** Disponível

em<<http://www.scielo.br/pdf/qn/v21n4/3193.pdf>> Acesso em 28 de setembro de 2014.

NETO, J. M. M.; MOITA, G. Ciaramella. **Uma introdução à análise exploratória de dados multivariado.** 1997. Disponível em: <<http://www.scielo.br/pdf/qn/v21n4/3193.pdf>> Acesso em 28 setembro 2014.

União Nacional das Instituições de Autogestão em Saúde – UNIDAS. **Pesquisa Nacional da Unidas 2012.** Disponível em<http://www.unidas.org.br/uploads/pesquisa_2012.pdf> Acesso em 1 de novembro de 2014.

SMAKA, Carlos. **Aplicação da Análise Multivariada na identificação de fatores que influenciam no custo de um plano de saúde.** Curitiba. Dissertação de Mestrado da Universidade Federal do Paraná, 2010. Disponível em <http://www.ppgmne.ufpr.br/arquivos/diss/229.pdf> Acesso em 12 de maio de 2014.

VICINI, Lorena. **Análise multivariada: da teoria à prática.** Disponível em http://72.29.69.19/~nead/disci/agr_preci/doc/mod4/3.pdf Acesso em 28 de setembro de 2014.