



UNIVERSIDADE FEDERAL DA PARAÍBA (UFPB)
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS (CCSA)
DEPARTAMENTO DE FINANÇAS E CONTABILIDADE (DFC)
CURSO DE CIÊNCIAS ATUARIAIS (CCA)

CAROLINE DA SILVA LIMA

**ANÁLISE DE SOBREVIVÊNCIA APLICADA AO TEMPO DE PERMANÊNCIA DO
SEGURADO EM UM PLANO DE SAÚDE: UM ESTUDO DE CASO**

JOÃO PESSOA, PB

2019

CAROLINE DA SILVA LIMA

**ANÁLISE DE SOBREVIVÊNCIA APLICADA AO TEMPO DE PERMANÊNCIA DO
SEGURADO EM UM PLANO DE SAÚDE: UM ESTUDO DE CASO**

Trabalho de Conclusão de Curso (TCC II) para o curso de Ciências Atuariais na UFPB, como requisito parcial à obtenção do título de bacharel em Ciências Atuariais.

Área de concentração: Saúde suplementar.

Orientador: Prof. Dr. Luiz Carlos Santos Júnior.

Coorientador: Prof. Me. Marcelo Coelho de Sá.

JOÃO PESSOA, PB

2019

Catologação na publicação
Seção de Catologação e Classificação

L732a Lima, Caroline da Silva.

ANÁLISE DE SOBREVIVÊNCIA APLICADA AO TEMPO DE
PERMANÊNCIA DO SEGURADO EM UM PLANO DE SAÚDE: UM ESTUDO
DE CASO / Caroline da Silva Lima. - João Pessoa, 2019.
50 f.

Orientação: Prof Dr Luiz Carlos Santos Júnior.
Monografia (Graduação) - UFPB/CCSA.

1. Saúde Suplementar. 2. Operadora de Plano de Saúde.
3. Análise de Sobrevivência. 4. Tempo de Permanência do
Segurado no Plano. I. Júnior, Prof Dr Luiz Carlos
Santos. II. Título.

UFPB/CCSA

CAROLINE DA SILVA LIMA

ANÁLISE DE SOBREVIVÊNCIA APLICADA AO TEMPO DE PERMANÊNCIA DO SEGURADO EM UM PLANO DE SAÚDE: UM ESTUDO DE CASO

Monografia apresentada à UFPB, como requisito para a obtenção do título de bacharel em Ciências Atuariais

Caroline da Silva Lima

Caroline da Silva Lima

Monografia aprovada em: 02/05/2019

BANCA EXAMINADORA

Luiz Carlos Santos Júnior

Prof. Dr. Luiz Carlos Santos Júnior

(Orientador)

Universidade Federal da Paraíba (UFPB)

Filipe Coelho de Lima Duarte

Prof. Me. Filipe Coelho de Lima Duarte

Membro avaliador

Universidade Federal da Paraíba (UFPB)

Werton J.C. Rodrigues Filho

Prof. Me. Werton José Cabral Rodrigues Filho

Membro avaliador

Universidade Federal da Paraíba (UFPB)

AGRADECIMENTOS

Em primeiro lugar gostaria de agradecer a Deus, por ter me dado tanta força de vontade para concluir com êxito o presente trabalho. Em segundo lugar a minha família, em especial minha querida mãe Jaqueline Luisa e ao meu querido pai Manoel, por sempre me apoiarem e me lembrarem constantemente que sempre posso obter o melhor, desde que eu me esforce para a realização do meu sonho.

Também gostaria de agradecer ao meu querido amigo e irmão de coração Oscar, que sempre foi solícito em me escutar quando surgia dúvidas a cerca do trabalho e que sempre me deu ideias valiosas para complementação do trabalho, e também entre outros assuntos. Fica registrado que não existem palavras para expressar o tamanho do amor que tenho por ele e toda minha gratidão.

A minha pequena gatinha Magali, por sempre está presente mesmo eu achando que estava sozinha, ela sempre se aproximava e me fazia lembrar de que eu não estava só.

Agradecer também ao meu irmão Gabriel e minha querida cunhada Ayrís, por sempre contribuir com minhas idas a universidade.

Gostaria também de agradecer ao Professor Luiz Carlos, que colaborou grandemente para a realização deste trabalho, sempre muito solícito em me ajudar em todas as dificuldades que surgiram ao decorrer da elaboração do trabalho. E também ao Professor Marcelo por ter contribuído com a elaboração do trabalho.

Também gostaria de agradecer a todos aqueles que não citei e ainda sim torceram sempre por mim, fica aqui registrado minha imensa gratidão a vocês e o meu muito obrigada.

*Dedico este trabalho aos meus queridos
avós Cecília Marinho e Severino
Cavalcante.*

RESUMO

A saúde suplementar movimentou 2,7% do PIB brasileiro em 2018 e seus lucros tem aumentado, apesar da recessão econômica apresentada pelo país e da redução de 1,5 milhões de segurados vinculados aos planos de saúde. Assim, o presente trabalho se propõe a analisar o tempo de permanência do segurado em um plano de saúde, a partir de dados (disponibilizados por uma Operadora atuante no mercado brasileiro que optou por não ser identificada) compostos por 122.381 (segurados e ex segurados) e acompanhados entre 1984 e 2018. A análise foi viabilizada pelo uso dos modelos de sobrevivência para dados censurados (em ambiente R, versão 3.4.0), como o estimador de Kaplan-Meier (1958) para a função de sobrevivência e o modelo de Cox (1972) para a estimação da função de risco (considerando a presença de covariáveis). Dentre os principais resultados, verificou-se que o tempo médio de permanência dos segurados em planos da referida operadora é de 1,974 anos, com desvio padrão de 2,271 e coeficiente de variação igual a 1,15. Observa-se que existe predominância de mulheres, solteiras, jovens, titulares e aderentes ao contrato de individual / familiar. Além disso, constatou-se a significância dos efeitos de quatro covariáveis sobre o tempo de permanência no plano: ser homem (em relação à mulher), ser jovem (em relação à adulto), ser dependente (em relação ao titular) e ser jovem / viúvo (em relação à amasiado) aumentam o risco de saída do plano de saúde analisado.

Palavras-chaves: Saúde Suplementar. Operadora de Plano de Saúde. Análise de Sobrevivência. Tempo de Permanência do Segurado no Plano.

ABSTRACT

Supplementary health accounted for 2.7% of Brazil's GDP in 2018 and its profits have increased, despite the country's economic recession and the reduction of 1.5 million health insurance policyholders. Thus, the present study proposes to analyze the time of permanence of the insured in a health plan, based on data (made available by a Operator in the Brazilian market that opted not to be identified) comprising 122,381 (insured and former insured) and monitored between 1984 and 2018. The analysis was enabled by the use of survival models for censored data (in R environment, version 3.4.0), as the Kaplan-Meier estimator (1958) for the survival function and the model Cox (1972) for the estimation of the risk function (considering the presence of covariates). Among the main results, it was verified that the average time of permanence of the insured in plans of this operator is 1,974 years, with standard deviation of 2,271 and coefficient of variation equal to 1,15. It is observed that there is predominance of women, single, young, titular and adherents to the individual / family contract. In addition, the significance of the effects of four covariates on the length of stay in the plan was verified: being male (in relation to female), being young (in relation to the adult), being dependent (relative to the owner) and being young / widower (in relation to the amassed) increase the risk of leaving the health insurance plan analyzed.

Keywords: Supplementary Health. Health Plan Operator. Survival Analysis. Length of Stay of the Insured in the Plan.

LISTA DE FIGURAS

Figura 1: Taxa de cobertura dos planos privados de assistência médica por unidades da federação (Brasil, 2018).	16
Figura 2: Beneficiários de planos de assistência médica por época de contratação do plano (Brasil, 2008-2018).....	17
Figura 3: Beneficiários de planos de assistência médica por tipo de contratação do plano (Brasil, dez/2018).	18
Figura 4: Número de beneficiários por sexo e grupo etário (Brasil, 2008-2018).....	18
Figura 5: Permanência dos beneficiários.....	31
Figura 6: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados, para a permanência do segurado no plano, Operadora, 2019.....	35
Figura 7: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por sexo, para a permanência do segurado no plano, Operadora, 2019.	36
Figura 8: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por dependência, para a permanência do segurado no plano, Operadora, 2019.....	36
Figura 9: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por estado civil, para a permanência do segurado no plano, Operadora, 2019.	37
Figura 10: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por tipo de contrato, para a permanência do segurado no plano, Operadora, 2019.	38
Figura 11: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por tipo de contrato, para a permanência do segurado no plano, Operadora, 2019.	39
Figura 12: Curva de sobrevivência estimada Cox versus Curva de sobrevivência de Kaplan-Meier.....	40

LISTA DE TABELAS

Tabela 1: Número de segurados por sexo e dependência, Operadora, 2019.	31
Tabela 2: Número de segurados por estado civil e sexo, Operadora, 2019.....	31
Tabela 3: Número de segurados por tipo de contrato e sexo, Operadora, 2019.....	32
Tabela 4: Número de segurados por tipo de contrato e dependência, Operadora, 2019.	32
Tabela 5: Número de segurados por estado civil e dependência, Operadora, 2019.....	32
Tabela 6: Número de segurados por estado civil e tipo de contrato, Operadora, 2019.....	33
Tabela 7: Número de segurados por idade e sexo, Operadora, 2019.	33
Tabela 8: Número de segurados por idade e tipo de contrato, Operadora, 2019.....	33
Tabela 9: Número de segurados por idade e estado civil, Operadora, 2019.	34
Tabela 10: Número de segurados por idade e dependência, Operadora, 2019.....	34
Tabela 11: Estimativas dos parâmetros do modelo ajustado de Cox para o tempo até a saída do plano, Operadora, 2019	40
Tabela 12: Teste de Schoenfeld para proporcionalidade de riscos do modelo de Cox ajustado para o tempo até a saída do plano, Operadora, 2019.....	41

SUMÁRIO

1 INTRODUÇÃO	12
2 REVISÃO DE LITERATURA.....	14
2.1 Saúde suplementar no Brasil.....	14
2.2.1 Breve histórico da saúde suplementar no Brasil.....	14
2.2.2 Breve panorama das operadoras de saúde no Brasil.....	15
2.2 Desligamento de clientes	19
3 METODOLOGIA.....	23
3.1 Tipo de pesquisa, amostra e coleta de dados.....	23
3.2 Análise dos dados.....	23
3.2.1 Caracterização dos dados.....	23
3.2.2 Análise de Sobrevivência	25
4 RESULTADOS	30
4.1 Apresentação dos dados	30
4.2 Análise descritiva.....	30
4.3 Análise do tempo de permanência dos segurados nos planos de saúde.....	34
5 CONCLUSÕES.....	42
REFERÊNCIAS	44
APÊNDICE – SCRIPT DA ANÁLISE EM LINGUAGEM R	48

1 INTRODUÇÃO

De acordo com a Lei 9.656/1998, em seu artigo 1º, definem-se as Operadoras de Planos Privados de Assistência à Saúde como sendo “toda e qualquer pessoa jurídica de direito privado, independente da forma jurídica de sua constituição, que ofereça tais planos mediante contraprestações pecuniárias, com atendimento em serviços próprios ou de terceiros”.

Conforme a Federação Nacional de Saúde Suplementar (FenaSaúde, 2019), as operadoras que compõem a estrutura empresarial do setor de saúde suplementar se classificam em diferentes modalidades de atuação no mercado: medicinas de grupo; seguradoras especializadas em saúde; cooperativas médicas; filantropias; autogestões; odontologias de grupo; cooperativas odontológicas; e, administradoras de benefício.

O referido setor é fiscalizado pela Agência Nacional de Saúde Suplementar (ANS) – criada em 2000 –, que “promove a defesa do interesse público na assistência suplementar à saúde, regula as operadoras setoriais, inclusive quanto às suas relações com prestadores e consumidores, e contribui para o desenvolvimento das ações de saúde no país” (BRASIL, 2019, p. 1). Noutras palavras, a ANS é a agência que regulariza o setor de saúde suplementar, sendo vinculada ao Ministério da Saúde.

Para que se tenha melhor entendimento sobre a dimensão desse setor, destacam-se sua cobertura e seu volume monetário movimentado em 2018: o setor de Saúde Suplementar brasileiro possuía, no referido ano, 47 milhões de beneficiários (com ou sem assistência odontológica), segundo a ANS (2019); movimentou R\$ 180 bilhões, aproximadamente 2,7% do PIB, de acordo com a FenaSaúde (2019).

Em 2016, apesar da perda de mais de 1,5 milhão de clientes – decorrentes da estagnação econômica brasileira e do aumento de desemprego (ABRAMGE, 2016) –, as operadoras de plano de saúde conseguiram ampliar sua receita – o lucro saltou 66% (em relação ao ano anterior) com reajustes ao consumidor bem acima da inflação (GLOBO, 2017).

No Brasil, resumidamente, os trabalhos que investigaram o setor de saúde suplementar comumente abordam: o funcionamento do setor (OCKÉ-REIS, CARDOSO, 2011); os efeitos econômicos provocados pela alteração de regulação no setor (SOUZA, 2014); a identificação do perfil do usuário dos planos de saúde (KANAMURA, VIANA, 2005; ANDRADE, 2014); os fatores determinantes para o desligamento do usuário dos referidos planos (HÖRBE, 2012). Este, por exemplo, a partir da insatisfação do serviço ofertado pelas operadoras,

determinou a propensão de desligamento do beneficiário do plano de saúde, utilizando-se de modelos de regressão múltipla.

Noutras áreas do conhecimento, é comum a utilização de modelos de sobrevivência para investigação do tempo até a ocorrência de um evento de interesse: tempo de postergação da aposentadoria do servidor público municipal (SANTOS JÚNIOR, 2018); tempo de permanência (sobrevivência) de uma empresa no mercado (PIAIA, JACOBI, VENTURINI, 2018); tempo até a aposentadoria de um trabalhador contribuinte (PORTILHO, 2013); tempo até a ocorrência de insolvência dos bancos privados (ALVES, 2009); tempo até a ocorrência de falha numa produção (BATTISTELLA, 2008); tempo até a morte de um paciente em tratamento (BARROS, MENEZES, 2008); tempo até o desligamento de assinantes de revista (BARROS, 2002).

Em análise de sobrevivência, a variável resposta é o tempo até a ocorrência de um evento de interesse, também denominado tempo de falha, onde a falha representa a ocorrência do evento de interesse; a principal característica de dados de sobrevivência é a presença de censura, que é a observação parcial¹ da resposta, ou seja, refere-se a situações em que, por alguma razão, o acompanhamento do paciente foi interrompido (COLOSIMO; GIOLO, 2006). Sem a presença de censura, as técnicas estatísticas clássicas, como a análise de regressão, poderiam ser utilizadas na análise deste tipo de dados (COLOSIMO; GIOLO, 2006).

Diante do contexto de redução de cobertura e de uma proposição metodológica pouco explorada em pesquisas referentes à saúde suplementar, apresentam-se, a partir de dados dos segurados e ex-segurados de uma Operadora de plano de saúde (que preferiu não ser identificada), os objetivos desta pesquisa.

- De modo geral, analisar o tempo de permanência do segurado no plano de saúde, a partir de modelos de sobrevivência tradicionais.
- De modo específico, descrever o perfil dos beneficiários observados e os efeitos que algumas covariáveis exercem sobre o seu tempo de permanência no plano de saúde.

¹ Isto significa que toda informação referente à resposta se resume ao conhecimento de que o tempo de falha é superior àquele observado.

2 REVISÃO DE LITERATURA

Este capítulo trata de discorrer sobre a área em que a presente pesquisa está inserida – o setor de saúde suplementar – por meio da apresentação de dados e de estudos que ajudem no recorte do problema – o desligamento de seus segurados.

2.1 Saúde suplementar no Brasil

Esta subseção trata de um breve histórico acerca da saúde suplementar no Brasil e de alguns aspectos relativos às operadoras de saúde no Brasil.

2.2.1 Breve histórico da saúde suplementar no Brasil

Em relação ao breve histórico, tem-se que a oferta por parte das operadoras de planos de saúde privados no Brasil teve início no final da década de 1960 (OCKÉ-REIS; ANDREAZZI; SILVEIRA, 2005, p. 1), durante a ascensão financeira do Brasil (BOTARO, 2015):

Desde o fim da década de 1960, um número crescente de trabalhadores passou a ser coberto pelos planos de saúde, seja mediante a celebração de contratos individuais, seja pela adesão a um contrato empresarial ou associativo; a assistência à saúde, assim, se transformava em um bem de consumo no âmbito do mercado de trabalho.

Mesmo após a criação do Sistema Único de Saúde (SUS), disposto no artigo 198 da Constituição Federal de 1988, a demanda pelos serviços privados de saúde permaneceu, principalmente, devido à escassez da oferta e à baixa qualidade dos serviços oferecidos pelo sistema público (OCKÉ-REIS, ANDREAZZI, SILVEIRA, 2005; PAIM *et al.*, 2011).

Nos anos 1990's os planos de saúde atingiram a marca de 38,7 milhões de brasileiros (24,5% da população do país) cobertos por pelo menos um plano de saúde. O aumento dessa celebração foi determinado por quatro fatores, conforme Botaro (2015): crescimento industrial nacional; ampliação da demanda da população por serviços médicos; baixa qualidade na oferta de serviços de saúde no Brasil; o aumento do custo da medicina liberal.

De acordo com Bahia (1999) o crescimento do número de beneficiários, houve o consequente aumento do número de queixas relativas aos serviços prestados e ao setor em

geral, o que explicitou a necessidade de regulamentação do setor, para resguardar beneficiários e operadoras prestadoras dos serviços.

Conforme Bahia (1999), a normatização era questionada, por um lado, pelos consumidores, que estavam insatisfeitos com os preços, inclusive na validade de planos de estabilização e as limitações de cobertura dos planos de saúde. Por outro lado, havia demanda do próprio governo para reformular o setor de seguros que constavam como pontos específicos na agenda de regulamentação dos planos de saúde privados. Assim, resta clara, a época, a necessidade de regulamentação do setor, até mesmo para orientar a ampliação do mercado quanto à precificação, prazo de vigência e cobertura.

Deste modo, estabeleceu-se a legislação estruturante da saúde suplementar:

- Lei nº 9.656 de 1998 – regulamentou setor de planos de saúde;
- Lei nº 9.961, de 2000 – criou a ANS e definiu suas finalidades, estrutura, atribuições, receita e a vinculação ao Ministério da Saúde;
- Decreto nº 3.327, de 2000 – aprovou o Regulamento da Agência Nacional de Saúde Suplementar – ANS, e deu outras providências;
- Lei nº 10.185, de 2001 – dispôs sobre a especialização das sociedades seguradoras em planos privados de assistência à saúde e dá outras providências.

O mercado de saúde suplementar no Brasil funcionou sem órgão específico para a sua fiscalização até a criação da ANS, por meio da Lei 9961/2000 que, segundo Botaro (2015), iniciou a desagregação entre o sistema público e o privado de saúde. Em conformidade com a referida Lei, em seu artigo 3º, a ANS promove a defesa do interesse público na assistência suplementar à saúde.

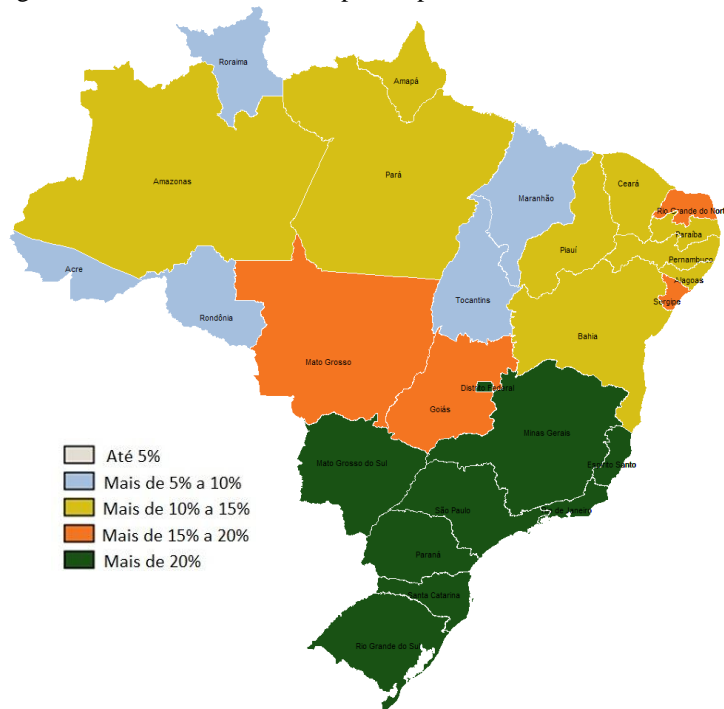
2.2.2 Breve panorama das operadoras de saúde no Brasil

Neste espaço, tratam-se da cobertura, da denominação de contratos (antigos e novos), da relação entre a celebração de novos de contratos e o crescimento econômico, da classificação de novos contratos e da distribuição dos beneficiários (por sexo e faixa etária) que consomem o referido produto.

Os planos de saúde no Brasil são comercializados por seguradoras de saúde, cujo intuito é proteger o segurado dos riscos de doenças. Desta maneira o segurado efetua pagamentos, através de celebração contratual, em que se descreve o valor de mensalidade, o prazo e a cobertura do seu plano. Após a assinatura do contrato o beneficiário terá acesso aos serviços oferecidos pela unidade de saúde que seu plano atende.

Quanto à cobertura dos planos privados no território brasileiro, sua disposição, por unidade da Federação, é exibida na Figura 1.

Figura 1: Taxa de cobertura dos planos privados de assistência médica por unidades da federação (Brasil, 2018).

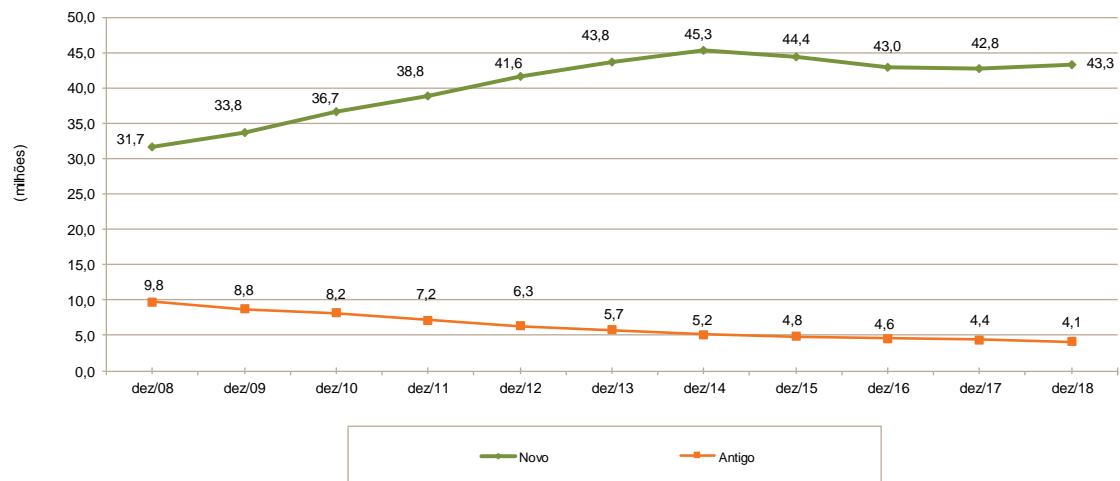


Fonte: SIB/ANS/MS – 12/2018 e população – IBGE/2012.

Constata-se, por meio da visualização da Figura 1, que as regiões sul e sudeste apresentam as mais altas taxas de cobertura, seguidas das regiões centro-oeste, nordeste e norte.

Quanto a sua denominação, os contratos celebrados antes da criação da Lei nº 9.656/98 são chamados de *antigos*, enquanto que os contratos celebrados após a criação da referida lei denominam-se *novos*, conforme a nomenclatura utilizada pela ANS. De acordo com os dados da ANS, até dezembro de 2018 o número de contratos antigos correspondia a aproximadamente 5 milhões (8,69%), enquanto que os novos a mais de 43 milhões (91,31%). Sua distribuição, entre 2008 e 2018, é demonstrada na Figura 2.

Figura 2: Beneficiários de planos de assistência médica por época de contratação do plano (Brasil, 2008-2018).



Fonte: SIB/ANS/MS – 12/2018.

A Figura 2 apresenta o comportamento dos números de novos e antigos contratos entre dezembro de 2008 e de 2018. É possível afirmar que o número de antigos contratos é decrescente e monótono, enquanto que o número de novos contratos (intrinsecamente ligado ao ciclo econômico²) apresenta nos primeiros anos um grande aumento, seguida de leve queda e suave elevação. Esta, relacionada ao pequeno crescimento apresentado pela economia brasileira em dezembro de 2018 (quando comparada a dezembro de 2017), foi impactada positivamente pela geração de oportunidades nos seguintes setores, que mais empregaram no segundo trimestre de 2018 (ABRAMGE, 2018): agropecuária (75.353); saúde (33.335); construção (20.743); atividades administrativas e serviços complementares (20.173); transporte, armazenagem e correio (17.134).

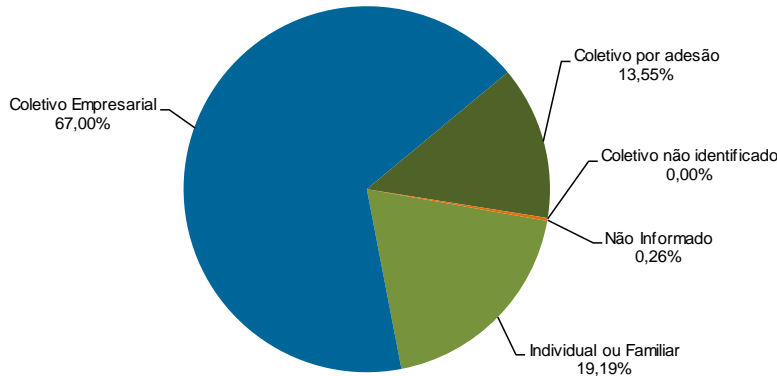
Em relação à classificação dos contratos novos, de acordo com Pauxis (2015), ela é dada por:

- a) Contratos individuais/familiares: são contratos celebrados a pessoa física, onde o beneficiário paga o prêmio integralmente à operadora do plano de saúde.
- b) Contratos coletivos por adesão: são contratos onde a associação profissional ou o sindicato contrata o plano ao beneficiário, ou seja, contratos celebrados a pessoa jurídica onde o beneficiário paga o prêmio integralmente à operadora de plano de saúde.

² Segundo Abramge (2016), a cada 100 pessoas que ficam desempregadas há uma redução imediata de 17 beneficiários nos planos de saúde. O número aumenta para 30 beneficiários após 6 meses de desligamento, totalizando 47 de 100 beneficiários que perdem o acesso aos planos de saúde após ficarem desempregados.

c) Contratos coletivos empresariais: são celebrados também com pessoa jurídica e o prêmio é parcial ou totalmente pago pela pessoa jurídica contratante à operadora. Quando pago parcialmente, a outra parte do prêmio fica sob a responsabilidade do beneficiário.

Figura 3: Beneficiários de planos de assistência médica por tipo de contratação do plano (Brasil, dez/2018).



Fonte: SIB/ANS/MS – 12/2018.

De acordo com a Figura 3, os contratos coletivos empresariais concentram a maior parte de segurados (67%), seguido do contrato individual/familiar (19,19%) e do coletivo por adesão (13,55%).

Em relação à distribuição dos beneficiários por sexo e faixa etária, a Figura 4 apresenta uma pirâmide etária para dois momentos no Brasil: dezembro de 2008 (a esquerda) e dezembro de 2018 (a direita).

Figura 4: Número de beneficiários por sexo e grupo etário (Brasil, 2008-2018)

Faixa Etária	2008		2018	
	Feminino	Masculino	Feminino	Masculino
80 anos ou mais	1,7%	0,9%	3,0%	1,7%
70 a 79 anos	3,2%	2,2%	4,6%	3,4%
60 a 69 anos	5,1%	4,3%	7,7%	6,7%
50 a 59 anos	10,0%	9,4%	11,1%	10,8%
40 a 49 anos	14,4%	14,5%	14,7%	14,9%
30 a 39 anos	18,8%	18,4%	20,6%	20,2%
20 a 29 anos	20,8%	20,6%	15,0%	14,9%
10 a 19 anos	12,0%	13,3%	10,4%	12,0%
0 a 9 anos	14,0%	16,3%	12,8%	15,4%

Fonte: site da ANS/SIB/MS (www.ans.gov.br/anstabnet) dados: referente à Dezembro/2008

Em conformidade com a Figura 4, vê-se que as faixas compreendidas entre 0 e 29 diminuíram seus percentuais, para ambos os sexos, em 2018 (quando comparado a 2008);

entre 30 e 80 anos ou mais, para ambos os sexos, houve aumento em todos os percentuais. Ou seja, os planos, hoje, apresentam menos jovens e mais adultos e velhos.

É possível, ainda, verificar que para todas as faixas etárias (com exceção daquela entre 20 e 29 anos) observadas em 2018, o percentual de mulheres é superior ao de homens. De acordo com Gomes, Nascimento e Araújo (2007) nota-se que a falta de presença de homens nos serviços de atenção inicial à saúde é menor em relação à mulher. De acordo com um estudo feito pelo Instituto de Estudo de Saúde Suplementar (IESS), coordenada por Novais e Martins (2010), observou-se em dois momentos (2003 e 2008) e através da Pesquisa Nacional por Amostra de Domicílios (PNAD) – a amostra, em 2008, entrevistou 392 mil indivíduos aleatoriamente da população brasileira de 189 milhões – que 53% dos beneficiários dos planos de saúde são do sexo feminino, destacando, ainda, que a tendência é que essa diferença aumente caso a mulher esteja em idade fértil. Segundo o estudo de 2010, 22% dos portadores de planos de saúde são mantidos pelo setor público.

2.2 Desligamento de clientes

Operadoras de planos de saúde, assim como as demais empresas que oferecem serviços e/ou produtos, analisam e se utilizam de mecanismos de retenção de clientes, com o objetivo de maximização de receitas. De acordo com Barros (2002), uma das maneiras de lucrar no mercado é aumentar o tempo de vida de vínculo dos clientes atuais. Reichheld e Sasser (1990 apud BARROS, 2002) afirmam, ainda, que as empresas podem impulsionar seus lucros em quase 100% ao reter 5% a mais de seus clientes.

Mas para obter números positivos no momento de retenção de clientes é necessário verificar seus perfis, ou seja, verificar se o mesmo é rentável para a empresa, de acordo com Chiamulera (2017). Segundo o mesmo, existem, de um lado, variados modelos dinâmicos que são construídos, em especial nas áreas financeira e de telecomunicações, com a finalidade de estimar o risco de um cliente parar de comprar o produto; de outro, modelos para identificar o comportamento em relação anterior ao ato do cancelamento, para que possam ser tomadas medidas remediadoras a tempo de evitá-lo.

Sob a perspectiva do desligamento de clientes, diversos trabalhos foram realizados com o objetivo principal de identificação dos fatores determinantes na manutenção ou perda de clientes por parte das empresas. Com este fim, são aplicados diversos modelos matemático-estatísticos.

Hohgraefe (2015) estimou o risco de cancelamento de clientes de uma empresa de Gestão de Frotas no Brasil (no trabalho não foi informado o nome da empresa por motivos de confidencialidade), por meio do Modelo de Cox, considerando variáveis tempo-dependentes, a partir de um processo de contagem. O modelo estimado selecionou cinco variáveis independentes significativas (no trabalho não foi informado o nome das variáveis por motivos de confidencialidade), sendo duas destas fixas no tempo (X2 e X3) e três tempos-dependentes (Z2, Z4 e Z5). Para ele, são utilizados, tradicionalmente, os modelos de regressão logística, de análise discriminante e de análise de sobrevivência para previsão de cancelamento de clientes.

Barros (2002), por sua vez, analisou o risco de desligamento de clientes de uma empresa que vende, no Brasil, assinaturas de periódicos que tenham realizado a primeira assinatura ou alguma renovação entre os anos de 1994 e 2001, por meio do Modelo de Risco Proporcional de Cox. A amostra, decorrente de amostragem aleatória simples, contempla apenas pessoas físicas e conta com 6.034 assinantes. Os resultados obtidos a partir desta amostra identifica que 50% dos assinantes efetuam o desligamento da assinatura até o 28º mês de vigência da assinatura. A partir do modelo de Cox foi possível concluir que Sexo, Estado Civil, Idade e Percentual do valor pago da última assinatura exercem efeitos sobre o tempo de assinatura de periódicos.

Silva (2012) utilizou a análise de sobrevivência para estimar a inadimplência de clientes correntistas portadores de cartão de crédito de uma instituição financeira no Brasil. A amostra do trabalho contém dados mensais de clientes de maio de 2006 a fevereiro de 2012, um total de 1.418 clientes que contrataram o cartão de crédito em maior de 2006. O autor observou as seguintes variáveis: cliente, tempo de sobrevivência, sexo, limite, enquadramento e renda confirmada, além da taxa de juros Selic, Índice Nacional de Preços ao Consumidor (IPCA) e Taxa de desemprego. O resultados obtidos mostram que cerca de 35% dos clientes se torna inadimplente durante o período observado, e que sexo feminino apresenta maior sobrevivência em relação à inadimplência.

Portilho (2013) verificou através da análise de sobrevivência o cancelamento de contratos de segurados de planos de previdência complementar. Para a análise foi utilizada uma amostragem aleatória com 70 mil participantes, observada entre 3 de janeiro de 2005 e 14 de agosto de 2011. Do total, 1,5% dos participantes tinha valor de contribuição igual a zero e foram excluídos do estudo; complementarmente, 68.968 observações foram utilizadas, sendo que 70% delas serviu para realização da estimação e 30% para validação dos modelos.

As variáveis analisadas no estudo de Portilho (2013) foram: estado civil, faixa etária, sexo, tipo de plano, tipo de pagamento, forma de pagamento, fez aporte, faixa de

contribuição. Dentre os principais resultados, tem-se que: participantes casados ou viúvos tendem a sobreviver mais tempo no plano; a faixa etária que tem maior probabilidade de permanecer no plano são pessoas com mais de 60 anos, assim como participantes de até 19 anos; quanto ao sexo, não mostram diferenças significativas; o tipo de plano mais persistente é o Plano Gerador de Benefício Livre (PGBL); quanto à forma de pagamento o que apresentam mais tempo de sobrevivência são os que participantes que optam por pagamento em débito em conta corrente; no que se refere à faixa de contribuição, os mais persistentes são aqueles que fazem maiores contribuições acima de R\$10.000,00.

Chiamulera (2017) no seu estudo optou por estimar o Valor Vitalício do Cliente (CLV) a partir da análise de sobrevivência. O trabalho teve como base de dados uma empresa que vende roupas e sapatos para lojas físicas e site de vendas (no Brasil). No primeiro momento, a partir do modelo de Cox ajustado ao banco de dados, foram identificadas variáveis associadas ao tempo de vida do cliente. Então a curva de sobrevida dos clientes do banco foi obtida e o CLV estimado.

Para o estudo de Chiamulera (2017) foram utilizadas duas bases de dados, uma com informações de clientes e outra com informações de todas as compras realizadas: a base cadastral foi analisada de Julho de 2010 a Setembro de 2017; já a base transacional, que se refere às compras realizadas, tem informações do período de Janeiro de 2014 a Setembro de 2017. No período de análise 11.283.031 clientes realizaram pelo menos uma compra. Nesse período 90.907 deixaram de realizar compras, o que representa 6,5% da base de clientes analisados.

O modelo final de Chiamulera (2017) apresentou 6 variáveis cadastrais e 18 variáveis transacionais (no trabalho, não foram detalhados os nomes das variáveis por motivos de confidencialidade). Com os resultados obtidos a partir das estimativas é possível verificar a influência de cada uma das variáveis do modelo no risco de deixar de comprar. Verificou-se que os níveis significativos da covariável “tipo de produto” que determinam o risco de desligamento são: ter comprado o produto_2, que aumenta o risco relativo em 81,26%; ter comprado o produto_1, que aumenta o risco em 46,57%; ter comprado produto_10, que aumenta o risco em 15,89%; e ter comprado o produto_20, que aumentam o risco em 15,42%.

Karam (2006) relacionou variáveis explicativas (transacionais e demográficas), a partir de uma regressão logística binária, à probabilidade de cancelamento dos clientes que assinam jornal no estado do Rio de Janeiro, das classes econômicas A e B. Para análise foram analisadas assinaturas canceladas entre Junho de 2004 a Maio de 2005 e assinaturas em Maio de 2005. Com o total de 230.858 assinaturas ativas e 105.524 assinaturas canceladas,

retiraram-se duas amostras aleatórias simples, uma de treinamento com 35.549 casos e para validação uma amostra composta por 4.796 casos. O modelo ajustado apontou que as variáveis que tiveram significância para explicar o desligamento dos clientes do jornal foram: tempo de permanência, forma de pagamento, indicador de reclamação, tipo de assinatura, fonte de venda, faixa etária, SD&W, LTV, região, indicador de participação em ações de fidelização e quantidade de produtos agregados comprados. Após a análise das variáveis citadas concluiu que quanto mais jovem for o cliente, maiores as chances de cancelamento; para a amostra, a variável gênero não influenciou o cancelamento.

De acordo com a revisão realizada, no Brasil, não existem pesquisas que tratam do desligamento de clientes em planos de saúde, em especial sob a ótica da análise de sobrevivência. O presente trabalho visa preencher tal lacuna, contribuindo, principalmente, com a aplicação da referida análise (que será apresentada no capítulo 3), além da apresentação dos fatores determinantes de tal saída.

3 METODOLOGIA

Nesse capítulo são descritos os aspectos metodológicos do presente trabalho, ou seja, detalha-se o caminho para alcance dos objetivos de pesquisa (GIL, 2008).

3.1 Tipo de pesquisa, amostra e coleta de dados

Aqui, classificam-se os procedimentos metodológicos da pesquisa (GIL, 2008): quanto à abordagem, trata-se de uma pesquisa quantitativa (análise dados coletados em formato numérico); quanto à natureza, trata-se de uma pesquisa aplicada (foca na aplicação de teorias); quanto aos objetivos, trata-se de uma pesquisa descritiva (descreve características particulares do grupo estudado e estabelece relações entre diferentes variáveis); quanto aos procedimentos, trata-se de um estudo de caso (estudo sobre um caso específico); quanto à delimitação temporal, trata-se de um estudo retrospectivo (análise fatos que já ocorreram).

Considerando que a população é a totalidade de indivíduos do universo que tenham características em comum, a população dessa pesquisa é constituída pelos segurados (ativos e inativos) vinculados a uma Operadora de saúde (que optou por não ser identificada) dentro do período compreendido entre 1980 e 2018. Como foram disponibilizados dados secundários acerca dos 122.381 segurados, a amostra corresponde à população.

3.2 Análise dos dados

O tempo de permanência do segurado no plano de saúde é aqui investigado considerando-se: os dados coletados junto à operadora do plano (subseção 3.2.1); os modelos de sobrevivência tradicionais (subseção 3.2.2); o *software* R, versão 3.4.0, com destaque para as funções contidas no pacote *survival* (THERNEAU, 2015).

3.2.1 Caracterização dos dados

Foram disponibilizadas dez (10) informações acerca de cada um dos 122.325 segurados (ativos e inativos): o número de contrato, o número do usuário, a titularidade, o número do plano, a data de nascimento, o sexo, o estado civil, a data de ingresso no plano, a data de exclusão do plano (quando for o caso) e o tipo de contrato. Destas, selecionaram-se as variáveis independentes (covariáveis) exibidas no Quadro 1.

Quadro 1 – Dados sobre os segurados, Operadora, 2018.

Covariáveis	Descrição	Nível	Tipo
Dependência	Código de dependência do usuário	Titular Dependente	Qualitativa
Idade	Idade do usuário na data de saída do plano ou ao término de acompanhamento da pesquisa	Idade < 30 → Jovem Idade ≥ 30 → Adulto	Qualitativa
Sexo	Sexo do usuário	F - Feminino M- Masculino	Qualitativa
Estado Civil	Estado civil do usuário	Amasiado(a) Casado(a) Não Sabe Separado(a) Solteiro(a) Viúvo(a)	Qualitativa
Tipo de Contratação	Tipo de vínculo ao contrato do usuário	Coletivo empresarial Individual familiar Coletivo adesão	Qualitativa

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Das covariáveis apresentadas no Quadro 1, apenas a idade não foi disponibilizada de forma direta. Ela foi calculada pela diferença entre

$$\text{mínimo}(\text{data de término de acompanhamento, data de saída do plano}) - \text{data de nascimento}.$$

Em seguida, em função da característica do modelo de Cox (que exhibe seus resultados como sendo riscos relativos, ou seja, apresenta se risco do grupo explícito é maior ou menor que a do grupo oculto no *output* do modelo), optou-se por segmentar a idade em dois grupos: jovens e adultos.

Assim, as covariáveis apresentadas, em conjunto com o Índice de Falha (que atribui valor 0 a censura e valor 1 a falha), viabilizam a análise de sobrevivência aplicada à variável-resposta intitulada “tempo até a saída do segurado do plano de saúde” (ou “tempo de permanência do segurado no plano de saúde”), definida pela diferença

$$\text{mínimo}(\text{data de término de acompanhamento, data de saída do plano}) - \text{data de admissão no plano}.$$

Isso significa que o referido tempo será analisado levando-se em conta o índice de falha e a presença das cinco covariáveis.

3.2.2 Análise de Sobrevivência

De acordo com Colosimo e Giolo (2006), as partes responsáveis pelo resultado da análise de sobrevivência são os tempos de falha e muito frequentemente a censura.

O tempo até a ocorrência do evento de interesse, ou simplesmente o tempo de falha, constitui a resposta da análise de sobrevivência. No nosso caso em particular, corresponde ao tempo de permanência do segurado num plano (ou o tempo até a saída do segurado).

A censura, que é a observação parcial da resposta – por algum acontecimento o acompanhamento do segurado é interrompido – é o principal aspecto dos dados de sobrevivência. Os tempos de censura devem ser utilizados na análise estatística, pois apesar das informações não serem completas, elas nos fornecem informações de tempo de vida, além do que o não relato das censuras no cálculo estatístico podem provocar resultados viciados (COLOSIMO; GIOLO, 2006). Existem vários tipos de censura, no entanto a única presente no corrente trabalho é a censura à direita do tipo I, aquela que ocorre quando o estudo tem data pré-definida para acabar – então indivíduos que não observaram o evento de interesse durante o estudo são considerados censurados ao término do acompanhamento (BORGES, 2014). Aqui, os segurados que não saíram do plano ao término do acompanhamento, isto é, em dezembro de 2018, foram censurados nessa data.

Como exposto por Colosimo e Giolo (2006), os dados de sobrevivência para o indivíduo i ($i = 1, \dots, n$) pode ser representados pelo par (t_i, δ_i) , sendo t_i o tempo de falha e δ_i a variável indicadora de falha ou censura, ou seja

$$\delta_i = \begin{cases} 1, & \text{se } t_i \text{ é um tempo de falha} \\ 0, & \text{se } t_i \text{ é um tempo censurado.} \end{cases}$$

Na presença de covariáveis medidas no i -ésimo indivíduo tais como, $x_i = (\text{sexo, idade, tipo de contrato, dependência})$, os dados ficam representados por $(t_i; \delta_i; x_i)$.

Para Colosimo e Giolo (2006), a variável aleatória e não negativa T , que representa o tempo de falha, é comumente especificada em análise de sobrevivência pela sua função de sobrevivência ou pela função de taxa de risco.

A função de sobrevivência é definida como a probabilidade de uma observação não falhar até certo tempo t , ou seja, a probabilidade de uma observação sobreviver ao tempo t . Em termos probabilísticos, isto é escrito como $S(t) = P(T \geq t)$. $S(t)$ é uma função como uma escada, ou seja, uma função com degraus nos tempos em que são observados de falha.

Caso existam empates em um determinado tempo t , então cada degrau será multiplicado pelo número de empates.

Para Colosimo e Giolo (2006), a probabilidade da falha ocorrer num intervalo de tempo $[t_1; t_2)$ pode ser expressa em termos da função de sobrevivência como $S(t_1) - S(t_2)$. A taxa de falha no intervalo $[t_1; t_2)$, por sua vez, é definida como a probabilidade de que a falha ocorra neste intervalo (dado que não ocorreu antes de t_1) dividida pelo comprimento do intervalo. Assim, a taxa de falha no intervalo $[t_1; t_2)$ é expressa por:

$$\frac{S(t_1) - S(t_2)}{(t_2 - t_1) S(t_1)} \quad (1)$$

Redefinindo o intervalo como $[t; t + \Delta t)$, a expressão (1) pode assumir a seguinte forma:

$$\lambda(t) = \frac{S(t) - S(t + \Delta t)}{\Delta t \times S(t)} \quad (2).$$

Assumindo Δt bem pequeno, $\lambda(t)$ representa a taxa de falha instantânea no tempo t condicional à sobrevivência até o tempo t . Observe que as taxas de falha são números positivos, mas sem limite superior. A função de taxa de falha de T é então definida como expresso em (3):

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \quad (3).$$

A função crescente indica que a taxa de falha do indivíduo observado aumenta com o transcorrer do tempo. Este comportamento mostra um efeito gradual de envelhecimento. A função constante indica que a taxa de falha não se altera com o passar do tempo. A função decrescente mostra que a taxa de falha diminui à medida que o tempo passa.

Existe, ainda, uma representação simplificada para a função de risco: $\lambda(t) = f(t)/S(t)$, em que $f(t) = [F(t)]'$, ou seja, a função densidade de probabilidade da variável aleatória T ; a função de distribuição acumulada, isto é, $F(t) = 1 - S(t) = P(T \leq t)$, define a probabilidade de uma observação não sobreviver ao tempo t (CARVALHO et al., 2011).

Em seguida, apresentam-se sucintamente: o estimador de Kaplan-Meier (técnica não paramétrica) para a função de sobrevivência; e o modelo de regressão de Cox (semiparamétrico) para estimar a função de risco e a função de sobrevivência do tempo de permanência do segurado no plano (considerando-se o efeito de covariáveis).

3.2.2.1 Estimador de Kaplan-Meier

De acordo com Colosimo e Giolo (2006) e Carvalho *et al.* (2011), o estimador não paramétrico de Kaplan-Meier é utilizado para estimar a função de sobrevivência. Para apresentação da sua expressão geral, consideram-se:

- $t_1 < T_2 < \dots < t_k$, os k tempos distintos e ordenados de falha,
- d_j o número de falhas em $t_j, j = 1, \dots, k$, e
- n_j o número de indivíduos sob risco t_j , ou seja, os indivíduos que não falharam e não foram censurados até o instante imediatamente anterior a t_j .

O estimador de Kaplan-Meier é, então, definido conforme (4)

$$\hat{S}(t) = \prod_{j:t_j < t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j} \right) = \prod_{j:t_j < t} (1 - q_j) = \prod_{j:t_j < t} (p_j). \quad (4)$$

As principais propriedades do estimador são:

- É o estimador de máxima verossimilhança de $S(t)$,
- É não viciado para amostras grandes,
- É fracamente consistente e
- Converge assintoticamente para um processo Gaussiano.

A variância assintótica do estimador Kaplan-Meier, bem como o intervalo de confiança, são apresentados por Colosimo e Giolo (2006).

3.2.2.2 Modelo de regressão de Cox

Segundo Colosimo e Giolo (2006) e Carvalho *et al.* (2011), o modelo de regressão de Cox permite a análise de dados provenientes de estudos de tempo de vida em que a resposta é o tempo até a ocorrência de um evento de interesse, ajustando por covariáveis.

De uma forma genérica, considere p covariáveis de modo que x é um vetor com os componentes $x = (x_1; \dots; x_p)$. A expressão geral do modelo de regressão de Cox, denotado por (5), considera

$$\lambda(t) = \lambda_0(t)g(\mathbf{x}'\boldsymbol{\beta}) \quad (5)$$

em que g é uma função que deve ser especificada, tal que $g(0) = 1$. Este modelo é composto pelo produto de dois componentes, um não paramétrico e outro paramétrico. O componente não paramétrico, $\lambda_0(t)$, não é especificado e é uma função não-negativa do tempo. Ele é usualmente chamado de função de base, pois $\lambda(t) = \lambda_0(t)$ quando $x = 0$. O componente paramétrico é frequentemente usado na forma multiplicativa exposta em (6)

$$g(\mathbf{x}'\boldsymbol{\beta}) = \exp(\mathbf{x}'\boldsymbol{\beta}) = \exp(\beta_1x_1 + \dots + \beta_px_p) \quad (6)$$

em que $\boldsymbol{\beta}$ é o vetor de parâmetros associado às covariáveis.

O suposto básico para utilizar o modelo de regressão de Cox (ou modelo de riscos proporcionais) é que as taxas de falha sejam proporcionais, ou seja, a razão das taxas de falha de dois diferentes indivíduos é constante no tempo:

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0(t) \exp\{\mathbf{x}'_i\boldsymbol{\beta}\}}{\lambda_0(t) \exp\{\mathbf{x}'_j\boldsymbol{\beta}\}} = \exp\{\mathbf{x}'_i\boldsymbol{\beta} - \mathbf{x}'_j\boldsymbol{\beta}\}. \quad (7)$$

Segundo Colosimo e Giolo (2006), o modelo de regressão de Cox é bastante utilizado e a principal razão desta popularidade é a presença do componente não paramétrico, que torna o modelo bastante flexível.

3.2.2.3 Teste de Hipóteses

Nesta seção se apresentam alguns testes com o intuito de comparar estimativas e verificar pressuposto de modelo.

O teste de logrank foi desenvolvido por Mantel em 1966 e serve para comparar estimativas não paramétricas da função de sobrevivência, por grupos (ou níveis) de covariáveis. Seu uso é devido quando a razão das funções de risco dos grupos a serem comparados é aproximadamente constante (as populações que estão sendo observadas tem a propriedade de riscos proporcionais).

Trata-se de um teste aproximado para a igualdade das duas funções de sobrevivência baseada na estatística expressa por (8), conforme Colosimo (2006)

$$T = \frac{[\sum_{j=1}^k (d_{2j} - w_{2j})]^2}{\sum_{j=1}^k (V_j)_2}, \quad (8)$$

sendo d_{2j} o número de mortes do segundo grupo pela causa j , que segue distribuição hipergeométrica, numa tabela de contingência; w_{2j} é a média de d_{2j} ; $(V_j)_2$ é a variância de d_{2j} ; T , sob a hipótese nula $H_0: S_1(t) = S_2(t)$, tem uma distribuição qui-quadrado com 1 grau de liberdade para grandes amostras.

O teste de resíduos de Schoenfeld (1982) verifica o pressuposto de proporcionalidade exigido no modelo de Cox. Para definir tais resíduos, considere que existam $k \leq n$ tempos distintos de falha $t_1 < t_2 < \dots < t_k$. Se o indivíduo i com vetor de covariáveis $x_i = (x_{1i}; x_{2i}; \dots; x_{pi})$ falhar, tem-se para este indivíduo, um vetor de resíduos Schoenfeld $r_i = (r_{1i}; r_{2i}; \dots; r_{pi})$ em que cada componente r_{qi} ($q = 1; \dots; p$) é definido por (9).

$$r_{iq} = x_{iq} - \frac{\sum_{j \in R(t_i)} x_{jq} \exp\{x'_j \hat{\beta}\}}{\sum_{j \in R(t_i)} \exp\{x'_j \hat{\beta}\}}. \quad (9)$$

Para cada uma das p covariáveis consideradas no modelo, tem-se, para cada indivíduo i , um correspondente resíduo Schoenfeld. Como os resíduos são definidos em cada falha, o conjunto de resíduos Schoenfeld é, desse modo, uma matriz com k linhas e p colunas. Cada linha corresponde a um tempo distinto de falha e cada coluna a uma das p covariáveis consideradas no modelo.

4 RESULTADOS

Neste capítulo se apresentam os dados e os resultados descritivos e de modelagem propostos.

4.1 Apresentação dos dados

Para realizar-se o que foi proposto no corrente trabalho, utilizou-se a base de dados de uma operadora de planos de saúde brasileira. A base de dados é composta por 122.324 segurados (com censura de 22,75%), acompanhados entre os anos de 1984 e 2018.

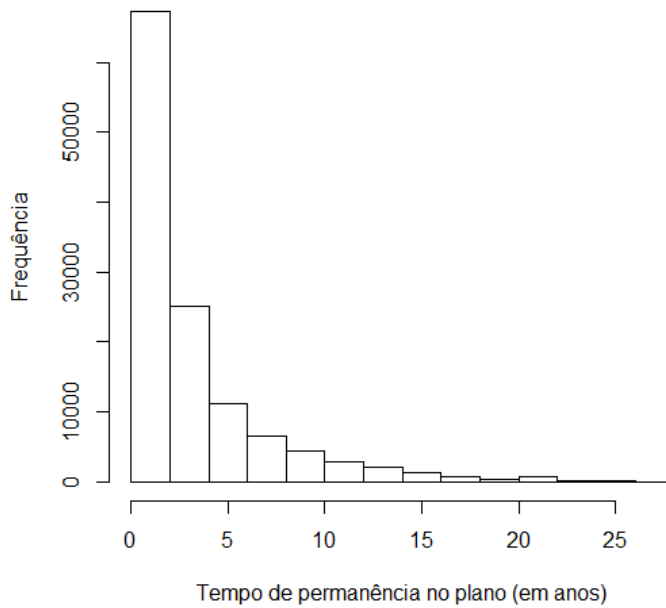
As variáveis para análise foram disponibilizadas pelos arquivos de usuários. Delas, foram selecionadas cinco (5) covariáveis (variáveis explicativas qualitativas), contidas no Quadro 1, que podem exercer efeito significativo sobre a saída do segurado de um plano: dependência, sexo, estado civil, tipo de contrato e idade.

Além destas, tem-se as variáveis fundamentais, aquelas que caracterizam os modelos de sobrevivência: o índice de falha δ (Status do Segurado), que indica se o segurado permaneceu (0) ou saiu (1) do plano; e a variável-resposta denominada tempo de permanência no plano, determinada pela diferença entre a idade ao término do acompanhamento (ou da idade de saída do plano) e a idade de entrada no plano.

4.2 Análise descritiva

A Figura 5 exhibe graficamente informações (em anos) relativas à tendência central e dispersão do tempo de permanência no plano, tais quais: valor mínimo = 0,000; 1º quartil = 0,000; mediana = 1,000; média = 1,974; 3º quartil = 3,000; valor máximo = 13,000; desvio-padrão = 2,271; coeficiente de variação = 1,151%. Percebe-se que o referido tempo se concentra a esquerda, ou seja, os beneficiários observados, em sua maioria, ficam no plano por poucos anos.

Figura 5: Permanência dos beneficiários



Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

Consoante os objetivos do estudo, para fins exploratórios, apresentam-se, das Tabelas 1 a 10, o número de pessoas pertencentes a determinados níveis de covariáveis.

Tabela 1: Número de segurados por sexo e dependência, Operadora, 2019.

Dependência	Sexo	
	F	M
Dependente	20.172 (16,49%)	14.160 (11,57%)
Titular	45.963 (37,57%)	42.029 (34,35%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Por meio da visualização da Tabela 1, é possível verificar que a maior parcela de segurados junto à operadora observada, por sexo e dependência, pertence ao grupo de mulheres titulares (37,57%), enquanto a menor parcela menor corresponde a homens dependentes (11,57%).

Tabela 2: Número de segurados por estado civil e sexo, Operadora, 2019.

Sexo	Estado civil					
	Amasiado	Casado	Não sabe	Separado	Solteiro	Viúvo
F	130 (0,10%)	20.964 (17,13%)	2.609 (2,13%)	1.035 (0,84%)	40.201 (32,86%)	1.196 (0,97%)
M	167 (0,13%)	14.673 (11,99%)	1.487 (1,21%)	427 (0,34%)	39.314 (32,13%)	121 (0,09%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Observando-se a Tabela 2, averigua-se que a maior parcela de beneficiários junto a operadora observada, por estado civil e dependência, é relativa ao grupo de solteiras do sexo feminino (32,86%); o menor número é o de viúvos do sexo masculino (0,09%).

Tabela 3: Número de segurados por tipo de contrato e sexo, Operadora, 2019.

Sexo	Tipo de contrato		
	Coletivo - Adesão	Coletivo - Empresarial	Individual – Familiar
F	4.918 (4,02%)	24.848 (20,31%)	36.369 (29,73%)
M	3.597 (2,94%)	25.880 (21,15%)	26.712 (21,83%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

A Tabela 3 apresenta que, em tipo de contratação por sexo, existe uma predominância de beneficiários do sexo feminino concentrados no plano individual familiar (29,73%); homens em contrato coletivo por adesão concentram menos beneficiários (2,94%). Observa-se que para o tipo de contrato de coletivo empresarial, o sexo masculino está mais presente do que o feminino, podendo indicar que as mulheres estão em menor quantidade no mercado de trabalho em comparação ao homem. Porém, sugere-se verificar se a causa deste fato (em caso afirmativo) ocorre para todas as idades, dado que existe a particularidade da gravidez e o cuidado nos primeiros anos dos filhos.

Tabela 4: Número de segurados por tipo de contrato e dependência, Operadora, 2019.

Dependência	Tipo de contrato		
	Coletivo - adesão	Coletivo - empresarial	Individual – familiar
Dependente	1.505 (1,23%)	21.971 (17,96%)	10.856 (8,87%)
Titular	7.010 (5,73%)	28.757 (23,50%)	52.225 (42,69%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

A observação da Tabela 4, que exhibe o segurado por tipo de contrato e dependência, permite concluir que titulares de plano individual familiar são maioria (42,69%); a minoria é composta dependentes de plano coletivo adesão (1,23%).

Tabela 5: Número de segurados por estado civil e dependência, Operadora, 2019.

Dependência	Estado civil					
	Amasiado	Casado	Não sabe	Separado	Solteiro	Viúvo
Dependente	106 (0,08%)	9.850 (8,05%)	430 (0,35%)	72 (0,05%)	23.749 (19,41%)	125 (0,10%)
Titular	191 (0,15%)	25.787 (21,08%)	3.666 (2,99%)	1.390 (1,13%)	55.766 (45,58%)	1.192 (0,97%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Observando-se na Tabela 5 a disposição de segurados por estado civil e dependência, constata-se que o maior número de beneficiários se encontra no grupo de solteiros e titulares (45,58%); o menor número observado é o de separados e dependentes (0,05%).

Tabela 6: Número de segurados por estado civil e tipo de contrato, Operadora, 2019.

Tipo de contrato	Estado civil					
	Amasiado	Casado	Não sabe	Separado	Solteiro	Viúvo
Coletivo - adesão	8 (0,00%)	1.345 (1,09%)	1.526 (1,24%)	59 (0,04%)	5.565 (4,54%)	12 (0,00%)
Coletivo - empresarial	248 (0,20%)	18.389 (15,03%)	1.710 (1,39%)	414 (0,33%)	29.764 (24,33%)	203 (0,16%)
Individual - familiar	41 (0,03%)	15.903 (13,00%)	860 (0,70%)	989 (0,80%)	44.186 (36,12%)	1.102 (0,90%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Quanto ao agrupamento de segurados por estado civil e tipo de contrato, verifica-se por meio da observação da Tabela 6 que o solteiro atrelado ao contrato individual familiar é a categoria que apresentação maior número de beneficiários (36,12%). Por outro lado, é possível observar que o amasiado com contratação por coletivo por adesão tem o menor número de beneficiários (0,00%). O solteiro que possui contrato do tipo individual familiar é o mais expressivo, podendo indicar que grande parte das pessoas investigadas são jovens.

Tabela 7: Número de segurados por idade e sexo, Operadora, 2019.

Sexo	Idade	
	Adulto	Jovem
F	27.443 (22,43%)	38.692 (31,63%)
M	20.068 (16,41%)	36.121 (29,53%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

A Tabela 7 exhibe, quando se segmenta por idade e sexo, uma maior concentração de usuários no grupo de mulheres jovens (31,63%), enquanto que a menor concentração fica por conta de homens adultos (16,41%).

Tabela 8: Número de segurados por idade e tipo de contrato, Operadora, 2019.

Tipo de contrato	Idade	
	Adulto	Jovem
Coletivo - adesão	3.113 (2,54%)	5.402 (4,42%)
Coletivo - empresarial	24.224 (19,80%)	26.504 (21,67%)
Individual - familiar	20.174 (16,49%)	42.907 (35,08%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

A Tabela 8 apresenta o número de segurados por idade e tipo de contrato: há maior prevalência de jovens vinculados a contratos familiares (35,08%) e menor prevalência de adulto em contratos coletivos de adesão (2,54%).

Tabela 9: Número de segurados por idade e estado civil, Operadora, 2019.

Estado civil	Idade	
	Adulto	Jovem
Amasiado	243 (0,20%)	54 (0,04%)
Casado	27.894 (22,80%)	7.743 (6,33%)
Não_sabe	2.315 (1,89%)	1.781 (1,46%)
Separado	1.301 (1,06%)	161 (0,13%)
Solteiro	14.454 (11,82%)	65.061 (53,19%)
Viúvo	1.304 (1,07%)	13 (0,01%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Na Tabela 9 é apresentada a quantidade de segurados por idade e estado civil. Constata-se que a maior parte pertence ao grupo de jovens e solteiros (53,19%); a menor, jovens e viúvos (0,01%).

Tabela 10: Número de segurados por idade e dependência, Operadora, 2019.

Dependência	Idade	
	Adulto	Jovem
Dependente	8.612 (7,04%)	25.720 (21,03%)
Titular	38.899 (31,80%)	49.093 (40,13%)

Fonte: elaboração própria, a partir dos dados da Operadora, 2019.

Por fim, em idade e dependência, averigua-se que titulares e jovens são maioria (40,13%); adultos e dependentes, minoria (7,04%).

Disto, conclui-se que a maior parte dos segurados observados é composta por mulheres, titulares, jovens, solteiras e aderentes ao contrato de individual / familiar.

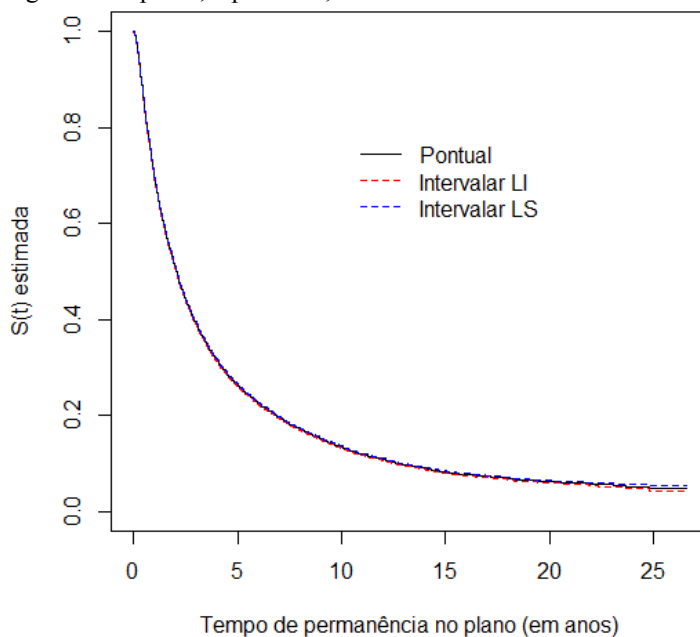
4.3 Análise do tempo de permanência dos segurados nos planos de saúde

Após essa breve descrição dos dados, inicia-se a análise principal do presente trabalho, isto é, a análise de sobrevivência aplicada ao problema da permanência do segurado em um plano de saúde. Aqui, ela é dividida em duas partes: resultados pertinentes ao método de Kaplan-Meier (1958) para estimar não parametricamente a função de sobrevivência (para todo o grupo e por níveis de covariáveis), seguida do teste logrank (para verificação de

igualdade entre as funções estimadas por níveis de covariáveis); e os resultados decorrentes do uso do modelo de Cox (1972) para estimar a função de sobrevivência e a função de risco (considerando o efeito de cinco covariáveis), seguida do teste de Schoenfeld para verificação da proporcionalidade dos riscos.

Em relação ao estimador de Kaplan-Meier para a função de sobrevivência, os resultados são exibidos das Figuras 6 a 10. A Figura 6 permite observar as estimativas (pontual e intervalares) da função de sobrevivência para todo o grupo analisado.

Figura 6: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados, para a permanência do segurado no plano, Operadora, 2019.

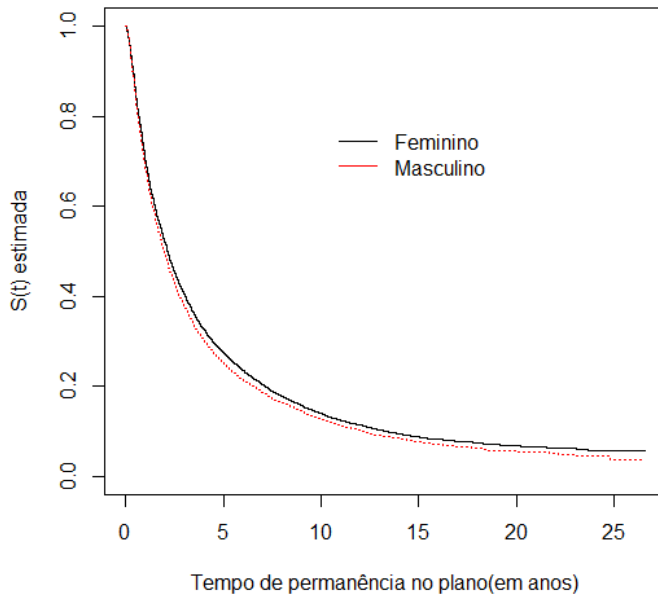


Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

Na Figura 6, observa-se que as estimativas intervalares se sobrepõem à estimativa pontual. Além disso, a sobrevivência do beneficiário na operadora é relativamente pequena: por exemplo, a probabilidade de sobrevivência em $t = 5$ (anos), ou seja, a probabilidade de o segurado permanecer no plano após 5 anos de sua entrada, fica em torno de 30% .

A Figura 7 permite observar as estimativas da função de sobrevivência, para todo o grupo analisado, por nível da covariável sexo, ou seja, de duas estimativas: uma referente ao sexo masculino, a outra ao feminino.

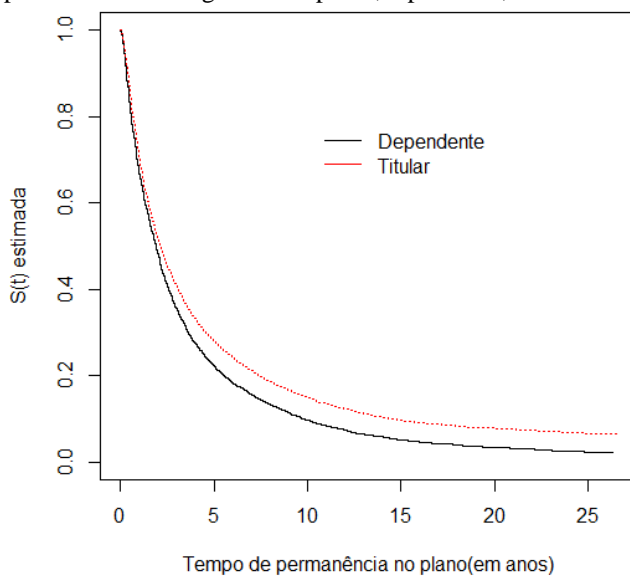
Figura 7: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por sexo, para a permanência do segurado no plano, Operadora, 2019.



Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

Os resultados deduzidos da observação da Figura 7 indicam que as sobrevivências estimadas para o sexo feminino e masculino nos primeiros anos de contratação do plano de saúde, parecem não diferir; com o passar dos anos, a sobrevivência do sexo feminino fica um pouco acima da do sexo masculino, o que pode significar que a permanência da mulher é maior do que a do homem no plano (isto, no entanto, precisa ser verificado por meio do teste logrank).

Figura 8: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por dependência, para a permanência do segurado no plano, Operadora, 2019.

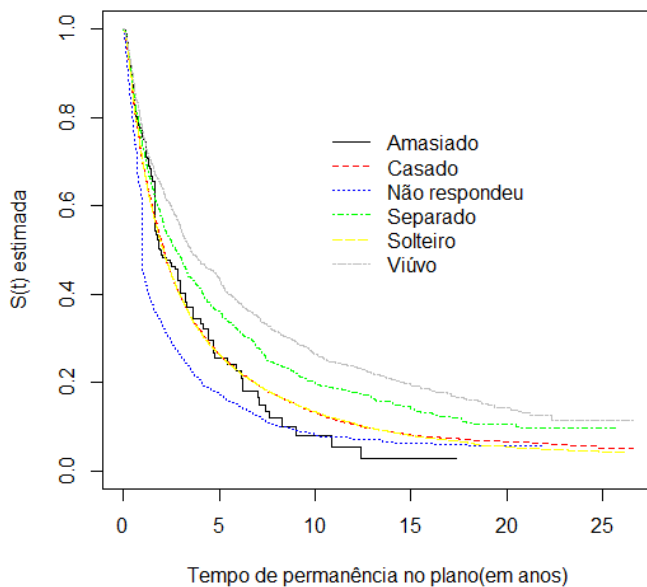


Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

A Figura 8 permite observar as estimativas da função de sobrevivência, para todo o grupo analisado, por nível da covariável dependência, ou seja, de duas estimativas: uma referente a titulares, a outra a dependentes. Observa-se que a curva de sobrevivência estimada para o titular é, de modo geral, superior a do dependente. Essa diferença (ou não) estatística, contudo, precisa ser verificada.

A Figura 9 permite observar as estimativas da função de sobrevivência, para todo o grupo analisado, por nível da covariável estado civil, ou seja, de seis estimativas: amasiado, casado, não respondeu, separado, solteiro e viúvo.

Figura 9: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por estado civil, para a permanência do segurado no plano, Operadora, 2019.



Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

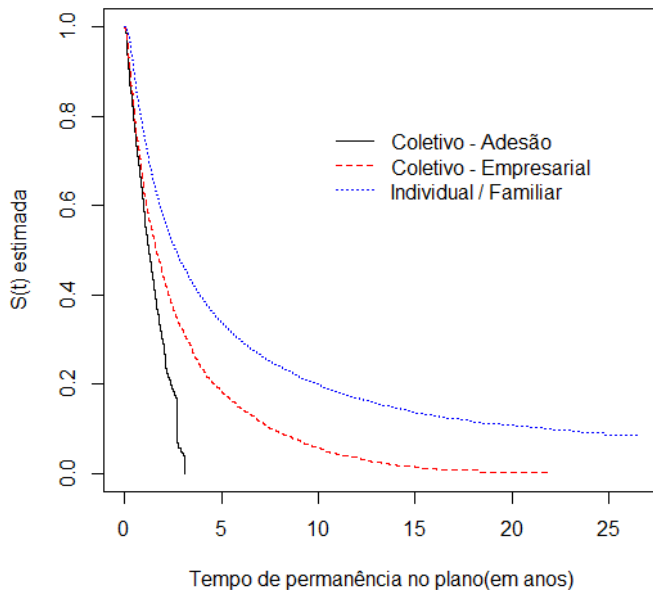
Da visualização da Figura 9, conclui-se que os viúvos representam o grupo que permanecem mais tempo no plano, seguido de beneficiários separados (estado civil, neste caso, pode ter relação de interação com a idade / renda – mas isso não foi investigado –, o que poderia indicar que quanto mais avançada a idade / maior a renda, mais propenso ao casamento e consequentemente maior a probabilidade de permanência no plano); solteiro e casado apresentaram estimativas próximas; o amasiado é o grupo de menor permanência no plano.

Noutros termos, pode-se pensar na possibilidade de viúvo(as) possuírem mais recursos financeiros, decorrentes do acúmulo dos seus proventos com os daqueles que faleceram; o casado, por sua vez, possui sobrevivência menor no plano, visto que seus gastos com a família

é maior. Outra possibilidade seria pensar que o viúvo é mais velho e passou, conseqüentemente, mais tempo vinculado ao plano.

A Figura 10 permite observar as estimativas da função de sobrevivência, para todo o grupo analisado, por nível da covariável tipo de contrato, ou seja, de três estimativas: coletivo – adesão; coletivo – empresarial; individual – familiar.

Figura 10: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por tipo de contrato, para a permanência do segurado no plano, Operadora, 2019.

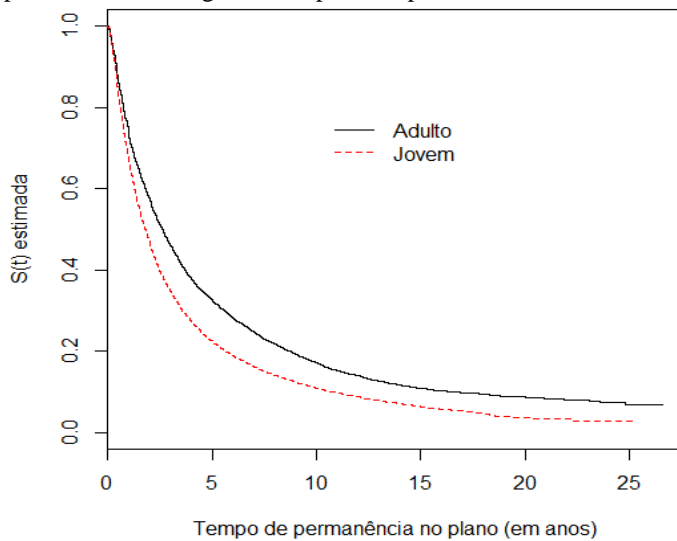


Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

Ao analisar-se a sobrevivência por tipo de contrato, por meio do exposto na Figura 10, nota-se que o contrato individual familiar é o tipo de contrato com maior duração. Complementa-se que quem tem menor expectativa de sobrevivência é o contrato coletivo por adesão (que não chegam nem ao quinto ano de permanência); pode ser que a amostra tenha apresentado algum grande cancelamento; as crescentes taxas de desemprego podem influenciar nessa quantidade.

A Figura 11 permite observar as estimativas da função de sobrevivência, para todo o grupo analisado, por nível da covariável idade, ou seja, de duas estimativas: uma referente a jovens, outra a adultos. É possível observar que a sobrevivência do adulto é maior, quando comparada a do jovem (o teste logrank deve ser realizado), o que faz sentido, já que de um modo geral o jovem é mais saudável e possui menos renda para pagar o plano.

Figura 11: Sobrevivência estimada por Kaplan-Meier, para a totalidade de segurados e por idade, para a permanência do segurado no plano, Operadora, 2019.



Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

Após a demonstração gráfica acerca das funções de sobrevivência estimadas, verifica-se, por meio do teste de *logrank*, se existe igualdade (estatística) entre as curvas estimadas por níveis de covariável utilizado (se a sobrevivência entre homens e mulheres é igual, se a sobrevivência entre titulares e dependentes é igual, etc.). Para esse teste, a hipótese nula é a de que as estimativas de sobrevivências são iguais, enquanto que a regra de decisão é: quando o valor p é menor ou igual a 5%, rejeita-se a hipótese nula; quando o valor p é maior que 5%, não se rejeita a hipótese nula. Após a análise dos dados, verificou-se que os valores p , para as quatro covariáveis (sexo, dependência, estado civil, tipo de contrato e idade) são menores que 5%. Sendo assim, rejeitam-se as hipóteses nulas, isto é, há evidências para rejeitar a hipótese de que as sobrevivências estimadas entre os níveis sejam iguais.

Com o ajuste do modelo de Cox é possível verificar quais covariáveis exercem efeitos significativos sobre o risco de saída do plano. O efeito das variáveis categóricas (que compara o efeito de cada nível em relação ao nível omissivo) é dado pela coluna “exponencial do coeficiente estimado” e a significância desse efeito, pela de “valor p ” (a significância do efeito é auferida quando o valor p é inferior 5%).

O modelo ajustado de Cox selecionado considerou quatro das cinco covariáveis: dependência, estado civil, sexo e idade. Assim, o tipo de contrato foi descartado, já que piorava o ajuste de Cox em relação à sobrevivência estimada por Kaplan-Meier. Seus principais resultados são apresentados na Tabela 11.

Tabela 11: Estimativas dos parâmetros do modelo ajustado de Cox para o tempo até a saída do plano, Operadora, 2019.

Variável - Nível	Coefficiente estimado	Exponencial do coeficiente estimado	Valor p
Dependência - Titular	-0,115985	0,890488	<0,001 ***
Estado civil - Casado	0,046372	1,047463	0,55264
Estado civil - Não respondeu	0,401859	1,494600	<0,001 ***
Estado civil - Separado	-0,111417	0,894566	0,18260
Estado civil - Solteiro	-0,250248	0,778608	0,00138***
Estado civil - Viúvo	-0,266493	0,766061	0,00157***
Sexo - Masculino	0,059772	1,061594	<0,001 ***
Idade - Jovem	0,430404	1,537879	<0,001 ***

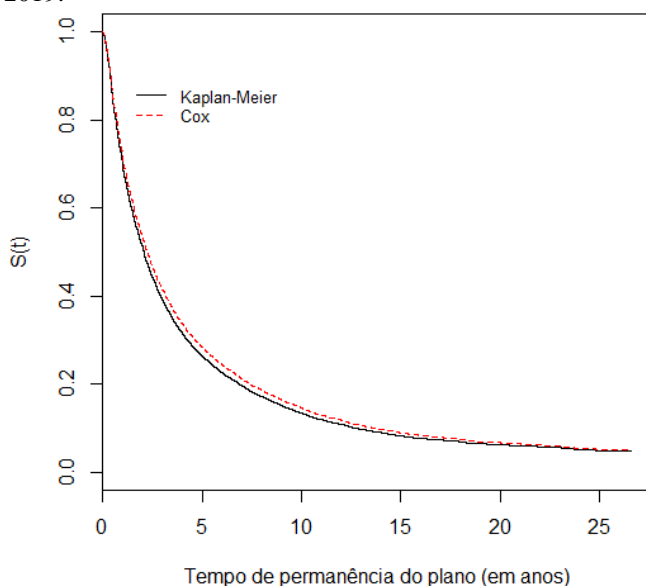
Códigos de significância: a) 0,01 → “***”; b) 0,05 → “**”; c) 0,1 → “*”; d) 1 → “”.

Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

Da Tabela 11, interpreta-se que: quanto à dependência, ser titular reduz o risco de sair do plano (em 10,96%), quando comparado a ser dependente; ser solteiro (em 22,14%) e viúvo (em 23,40%) reduzem o risco de sair do plano, quando comparados à categoria “amasiado”; ser homem, em relação a ser mulher, aumenta o risco de saída do plano em 6,15%; ser jovem, em relação ao adulto, aumenta o risco de saída em 53,78%. Os níveis mencionados foram considerados significativos tomando como referência um nível de significância de 5%.

A função de sobrevivência decorrente do modelo de Cox é confrontada com a estimada por Kaplan-Meier na Figura 12. Quanto mais próxima aquela estiver dessa, melhor o ajuste proporcionado por Cox.

Figura 12: Curva de sobrevivência estimada Cox versus Curva de sobrevivência de Kaplan-Meier, Operadora, 2019.



Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

É possível observar na Figura 12 que o modelo de Cox bem ajustou os dados, superestimando um pouco a sobrevivência.

Tabela 12: Teste de Schoenfeld para proporcionalidade de riscos do modelo de Cox ajustado para o tempo até a saída do plano, Operadora, 2019.

Variável – Nível	rho	chisq	Valor p
Dependência – Titular	-0,01257	14,76	<0,001
Estado civil – Casado	-0,00551	2,87	0,0905
Estado civil - Não respondeu	-0,01134	12,16	<0,001
Estado civil - Separado	-0,00749	5,3	0,0213
Estado civil - Solteiro	-0,00516	2,52	0,112
Estado civil - Viúvo	-0,00751	5,33	0,021
Sexo – Masculino	-0,01628	25,08	<0,001
Idade – Jovem	-0,00968	10,68	0,00108
GLOBAL	NA	120,03	0

Fonte: elaboração própria, a partir dos dados disponibilizados (2019).

No entanto, faz-se a ressalva de que a hipótese de proporcionalidade de riscos foi rejeitada (pelo teste de Schoenfeld), dado que os valores p global e por covariáveis (conforme Tabela 12) foram todos inferiores ao nível de significância de 5%.

Assim, encerra-se a seção de resultados destacando que o modelo ajustado, apesar de não obedecer à premissa de proporcionalidade dos riscos, bem ajustou os dados e identificou os fatores de risco de saída do segurado do plano de saúde analisado: ser homem, jovem, amasiado e dependente.

5 CONCLUSÕES

A importância do desligamento de clientes vem à tona quando nos deparamos com notícias que evidenciam a perda de segurados nos planos de saúde, decorrente da recessão econômica brasileira. Ciente que o impacto que essa variação pode proporcionar sobre a solvência da instituição, optou-se, neste trabalho, por analisar o tempo de permanência do segurado num plano de saúde, a partir de dados disponibilizados por uma seguradora que optou por não se identificar.

A análise exploratória permitiu identificar os segurados da Operadora investigada: a maior parte dos segurados observados é composta por mulheres, titulares, jovens, solteiras e aderentes ao contrato de individual / familiar.

O uso dos modelos de sobrevivência foi motivado pela identificação de uma lacuna em estudos que tratam de desligamento de clientes em operadoras de plano de saúde no Brasil. O comum, no tratamento de desligamento de clientes noutras ramos é o uso de modelos de regressão logístico. Essa “novidade” permitiu considerar informações parciais sobre os segurados (censura), um ganho em termos de tamanho de amostra. Por meio dela, foi possível estimar a funções de sobrevivência por Kaplan-Meier e comparar as estimativas entre os níveis de covariáveis. Além disso, estimou-se o risco de saída do plano, considerando-se os efeitos de quatro variáveis qualitativas, em que todas apresentaram efeitos significativos sobre o referido risco: duas aumentam o risco de saída do plano (ser homem e ser jovem) e duas reduzem o risco de saída do plano (ser titular e solteiro/viúvo).

Dentre as limitações e sugestões, citam-se:

- Não foram encontrados trabalhos brasileiros que tratassem do tempo de permanência do segurados nos planos de saúde, tampouco utilizando-se dos modelos de sobrevivência, o que dificultou o direcionamento das discussões na seção de resultados. Sugere-se a elaboração de trabalho bibliométrico, com foco na revisão internacional;
- Apesar da disponibilização dos dados, observa-se a dificuldade que se tem para acessá-los: as companhias costumam impor limitações em relação ao uso de variáveis importantes, como renda. A ausência dessa variável provavelmente piora a qualidade do ajuste e impossibilita a averiguação de algumas hipóteses de pesquisa.

- Indica-se a continuidade dessa pesquisa, quer seja considerando-se o uso de outros bancos de dados, quer seja utilizando-se de outros modelos de sobrevivência (como os paramétricos) e de outras covariáveis.

Por fim, espera-se que, apesar das limitações, a corrente pesquisa tenha contribuído para o aprimoramento de estudos em saúde suplementar, dada a sua escassez e importância para o bem-estar e para a economia brasileira.

REFERÊNCIAS

- ALVES, Karina Lumena de Freitas. **Análise de Sobrevivência de Bancos Privados no Brasil**. 2009. 83 f. Dissertação (Mestrado) - Curso de Engenharia de Produção, Universidade de São Paulo, São Carlos, 2009. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18140/tde-28102009-103529/en.php>>. Acesso em: 07 abr. 2019.
- ANDRADE, Vanelli Dantas de. **A importância dos planos de saúde conhecerem o perfil dos usuários do plano empresarial**. 2014. 76 f. TCC (Graduação) - Curso de Serviço Social, Universidade Federal do Rio Grande do Norte, Natal, 2014. Disponível em: <https://monografias.ufrn.br/jspui/bitstream/123456789/6500/1/VanelliDA_Monografia.pdf>. Acesso em: 07 abr. 2019.
- ANS. **ANS TABNET: Informações em Saúde Suplementar**. 2019. Beneficiários por UFs, Regiões Metropolitanas (RM) e Capitais. Disponível em: <http://www.ans.gov.br/anstabnet/cgi-bin/tabnet?dados/tabnet_br.def>. Acesso em: 07 abr. 2019.
- BAHIA, Lígia. **Mudanças e Padrões das Relações Público-Privado: Seguros e Planos de Saúde no Brasil**. 1999. 380 f. Tese (Doutorado) - Curso de Saúde Pública, Escola Nacional de Saúde Pública da Fundação Oswaldo Cruz, Rio de Janeiro, 1999. Disponível em: <https://saudecomunista.files.wordpress.com/2014/07/o-mercado-de-planos-de-sac3bade-no-brasil-_ligia-bahia-1.pdf>. Acesso em: 14 fev. 2019
- BARROS, José Eduardo. **Análise de Sobrevivência: Modelo de Risco de desligamento de clientes**. 2002. 124 f. Dissertação (Mestrado) - Curso de Administração, Ufrj, Rio de Janeiro, 2002. Disponível em: <https://www.coppead.ufrj.br/upload/publicacoes/Jose_Eduardo_Barros.pdf>. Acesso em: 25 mar. 2019.
- BARROS, Victor Freitas de Azerêdo; MENEZES, José Elmo de. Aplicação da Teoria da Análise de Sobrevivência no Cálculo do Risco de Morte por Infecção Hospitalar. In: SIMPOETS, Não use números Romanos ou letras, use somente números Arábicos]., 2008, Góiais. p. 81 - 85. Disponível em: <https://www.researchgate.net/publication/307645288_APLICACAO_DA_TEORIA_DA_ANALISE_DE_SOBREVIVENCIA_NO_CALCULO_DO_RISCO_DE_MORTE_POR_INFECCAO_HOSPITALAR>. Acesso em: 07 abr. 2019.
- BATTISTELLA, Patrícia Moises Davila. Universidade Federal de Santa Catarina - centro de ciências agrárias, departamento de ciência e tecnologia de alimentos, programa de pós-graduação em ciência dos alimentos. **Análise de sobrevivência aplicada à estimativa da vida de prateleira de salsicha**. 2008. 115 f. Dissertação (Mestrado) - Curso de Ciência dos Alimentos, Universidade Federal de Santa Catarina, Florianópolis, 2008. Disponível em: <<https://repositorio.ufsc.br/bitstream/handle/123456789/91016/264736.pdf?sequence=1>>. Acesso em: 07 abr. 2019

BOTARO, Nittina. A trajetória do Sistema de Saúde no Brasil:: considerações sobre o cenário atual. **Revista Escrita da História**, Juiz de Fora, v. 1, n. 1, p.157-184, mar 2015. Anual. Disponível em: <<http://www.escritadahistoria.com/revista/index.php/reh/article/view/26/23>>. Acesso em: 14 fev. 2019.

BRASIL. AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR. . **Cobertura assistencial**: série planos de saúde conheça seus direitos. 2005. Disponível em: <http://bvsmms.saude.gov.br/bvs/publicacoes/guia_cobertura_assistencial.pdf>. Acesso em: 02 maio 2017.

BRASIL. AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR. **Dados Gerais**. 2016. Disponível em: <<http://www.ans.gov.br/perfil-do-setor/dados-gerais>>. Acesso em: 26 mar. 2017.

BRASIL. AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR. **Dados Gerais**. 2019. Disponível em: <<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>>. Acesso em: 23 mar. 2019.

BRASIL. Constituição (1998). Lei Nº 9.656, 03 de Junho de 1998. **Planos de Saúde** Brasília, DF: 03 jun. 1998. Disponível em <http://www.planalto.gov.br/ccivil_03/leis/L9656compilado.htm>. Acesso em: 26 mar. 2017

CARVALHO, M. S.; ANDREOZZI, V. L.; CODEÇO, C. T.; CAMPOS, D. P.; BARBOSA, M. T. S; SHIKAMURA, S. E.. **Análise de sobrevivência: teoria e aplicações em saúde**. 2 ed.. Rio de Janeiro: Fiocruz, 2011.

CÉNARIO Saúde. São Paulo: Abramge, v. 3, n. 2, 2018. Trimestral. Disponível em: <https://www.abramge.com.br/portal/files/cenario-saude/cenario_da_saude_ed12.pdf>. Acesso em: 23 mar. 2019.

CÉNARIO Saúde. São Paulo: Abramge, v.II , n. 06, ago. 2016. Semestral. Disponível em: <https://www.abramge.com.br/portal/files/cenario-saude/cenario_da_saude_ed6.pdf>. Acesso em: 23 mar. 2019.

CHIAMULERA, Maximiliano Luis. **Modelagem do valor vitalício do cliente via abordagem de análise de sobrevivência**. 2017. 46 f. TCC (Graduação) - Curso de Estatística, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017. Disponível em: <<https://lume.ufrgs.br/handle/10183/175315>>. Acesso em: 01 abr. 2019.

COLOSIMO, E. A.; GIOLO, S. R. 2006. **Análise de Sobrevivência Aplicada**. 1ª Edição, Editora Edgard Blücher, São Paulo-SP.

DESCONHECIDO. **Quantidade de segurados de planos de saúde diminuiu no Brasil em 2015**. 2015. Disponível em: <<http://g1.globo.com/jornal-nacional/noticia/2015/10/quantidade-de-segurados-de-planos-de-saude-diminuiu-no-brasil-em-2015.html>>. Acesso em: 26 mar. 2017

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6 ed. São Paulo: Atlas, 2008.

GOMES, Romeu; NASCIMENTO, Elaine Ferreira do; ARAÚJO, Fábio Carvalho de. Por que os homens buscam menos os serviços de saúde do que as mulheres? As explicações de homens com baixa escolaridade e homens com ensino superior. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 23, n. 3, p.565-574, mar. 2007. Disponível em:

<<https://www.scielo.org/pdf/csp/2007.v23n3/565-574/pt>>. Acesso em: 23 mar. 2019.

HOHGRAEFE, Mariana Wink. **Estimação do risco de cancelamento de clientes no setor degerenciamento de frotas por meio de modelo de Cox com variáveis tempo-dependentes**. 2015. 41 f. TCC (Graduação) - Curso de Estatística, Ufrgs, Porto Alegre, 2015. Disponível em: <<https://www.lume.ufrgs.br/handle/10183/129444>>. Acesso em: 25 mar. 2019.

HÖRBE, Christine de Vasconcellos. **Comportamento do consumidor após falhas em serviços: uma pesquisa com usuários de planos de saúde**. 2012. 116 f. Dissertação (Mestrado) - Curso de Administração, Universidade do Vale do Rio dos Sinos, São Leopoldo, 2012. Disponível em:

<<http://repositorio.jesuita.org.br/bitstream/handle/UNISINOS/4375/21b.pdf?sequence=1&isAllowed=y>>. Acesso em: 07 abr. 2019.

<http://cnseg.org.br/fenasaude/sobre-o-setor/>

JACKSON, Christopher (2016). flexsurv: A Platform for Parametric Survival Modeling in R. *Journal of Statistical Software*, 70(8), 1-33. doi:10.18637/jss.v070.i08

KANAMURA, Alberto Hideki; VIANA, Ana Luiza D'Ávila. Gastos elevados em plano privado de saúde: com quem e em quê. **Saúde**, Desconhecido, v. 5, n. 41, p.815-820, jun. 2007. Disponível em: <<http://www.revistas.usp.br/rsp/article/view/32316/34483>>. Acesso em: 01 jun. 2017.

KARAM, Karine de Almeida. **Regressão Logística: Um modelo de Risco de Cancelamento de Clientes**. 2006. 100 f. Dissertação (Mestrado) - Curso de Administração, Puc, Rio de Janeiro, 2006. Disponível em:

<https://s3.amazonaws.com/academia.edu.documents/34008216/dissertacao_de_mestrado.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1553480520&Signature=5wAzGoJskTNfH%2FzfcaP4HuKvL3w%3D&response-content-disposition=inline%3B%20filename%3DDissertacao_de_mestrado.pdf>. Acesso em: 25 mar. 2019.

LEAL, Rodrigo Mendes; MATOS, João Boaventura Branco de. PLANOS DE SAÚDE: UMA ANÁLISE DOS CUSTOS ASSISTENCIAIS E SEUS COMPONENTES. **Forum**, São Paulo, v. 49, n. 4, p.447-458, dez. 2009. Disponível em:

<<http://bibliotecadigital.fgv.br/ojs/index.php/rae/article/view/36417/35191>>. Acesso em: 02 maio 2017.

NASCIMENTO, Alexandre Costa. **Inadimplência rompe contratos de planos de saúde**. 2012. Disponível em: <<http://www.gazetadopovo.com.br/economia/inadimplencia-rompe-contratos-de-planos-de-saude-2brbb5s5ahj7pqfob06i7ndji>>. Acesso em: 28 mar. 2017.

NOVAIS, Marcos; MARTINS, Carina Burri. **Perfil dos Beneficiários de Planos e SUS e o Acesso a Serviços de Saúde – PNAD 2003 e 2008**. 2010. Textos para Discussão. Disponível em: <<http://documents.scribd.com/s3.amazonaws.com/docs/2r5eyi6jsw3oglgw.pdf>>. Acesso em: 23 mar. 2019

OCKÉ-REIS, Carlos Octávio; ANDREAZZI, Maria de Fátima Siliansky de; SILVEIRA, Fernando Gaiger. O mercado de planos de saúde no Brasil: uma criação do estado?. **Ipea**: Instituto de Pesquisa Econômica Aplicada, Rio de Janeiro, v. -, n. -, p.01-20, jun. 2005. Disponível em: <http://repositorio.ipea.gov.br/bitstream/11058/1953/1/TD_1094.pdf>. Acesso em: 05 mar. 2019.

OCKÉ-REIS, Carlos Otávio; CARDOSO, Simone de Souza. A regulamentação dos preços dos planos individuais de saúde. **Revista de Economia Política**, Não Informado, v. 31, n. 3, p.455-470, jun. 2011. Disponível em: <<http://www.scielo.br/pdf/rep/v31n3/08.pdf>>. Acesso em: 07 abr. 2019.

Painel de precificação planos de saúde 2015. Rio de Janeiro: Ans, 2016. Disponível em: <http://www.ans.gov.br/images/stories/Materiais_para_pesquisa/Perfil_setor/Foco/painel_precificacao2015_completo.pdf>. Acesso em: 12 maio 2017

PIAIA, Rafael; JACOBI, Luciane Flores; VENTURINI, Jonas Cardona. Estudo da sobrevivência de empresas de Santa Maria - RS no período de agosto de 2008 a agosto de 2013. **Revista Thema**, [s.l.], v. 15, n. 4, p.1200-1216, 3 out. 2018. Instituto Federal de Educacao, Ciencia e Tecnologia Sul-Rio-Grandense. <<http://dx.doi.org/10.15536/thema.15.2018.1200-1216.998>>

PORTILHO, Carolina Marques. **Estimação da Persistência de Segurados de Planos de Previdência Privada Via Modelos de Sobrevivência**. 2013. 61 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Pontifca Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013. Disponível em: <http://www.dbd.puc-rio.br/pergamum/tesesabertas/1021478_2013_completo.pdf>. Acesso em: 01 abr. 2019.

SANTOS JUNIOR, Luiz Carlos. **Análise de sobrevivência aplicada a premissas atuariais: o caso da previdência pública municipal de Cabedelo/PB**. 2018. 162 f. Tese (Doutorado) - Curso de Biometria, Bioestatística, Universidade Estadual Paulista, Botucatu, 2018. Disponível em: <https://repositorio.unesp.br/bitstream/handle/11449/181622/santosjunior_lc_dr_bot.pdf?sequence=3&isAllowed=y>. Acesso em: 25 abr. 2019.

SILVA, Sandra Almeida. **Estudo de risco de crédito em operações de cartão de crédito usando variáveis macroeconômicas e técnicas de análise de sobrevivência**. 2012. 93 f. Dissertação (Mestrado) - Curso de Administração, Universidade Presbiteriana Mackenzie, São Paulo, 2012. Disponível em: <<http://tede.mackenzie.com.br/jspui/handle/tede/590>>. Acesso em: 01 abr. 2019.

SOUZA, Rodrigo Mendes Leal de. **O mercado de saúde suplementar no brasil: regulação e resultados econômicos dos planos privados de saúde**. 2014. 296 f. Tese (Doutorado) - Curso de Políticas Públicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em: <http://www.ie.ufrj.br/images/pos-graduacao/ppge/Tese_RodrigoMendesLeal-v-25ago2014.pdf>. Acesso em: 07 abr. 2019.

THERNEAU, T. M. (2015). *_A Package for Survival Analysis in S_*. version 2.38, <URL: <https://CRAN.R-project.org/package=survival>>.

APÊNDICE – SCRIPT DA ANÁLISE EM LINGUAGEM R

```
require(survival) #requerer o pacote survival
```

```
#####ABRIR OS DADOS
```

```
dados=read.table("camed.txt",h=T) #serve pra ler o banco de dados; precisa que você informe anteriormente onde está localizado o arquivo.
```

```
attach(dados) #separa as informações por colunas (por variáveis)
```

```
dim(dados) #dimensão da matriz
```

```
head(dados) #ver as 6 primeiras linhas
```

```
names(dados) #ver nome das variáveis
```

```
dep=Dependência
```

```
sex=Sexo
```

```
est=Estado_civil
```

```
tip=Tipo_contratação
```

```
ida=Idade
```

```
cens=Falha
```

```
tempo=Tempo
```

```
ida1=ifelse(Idade<30,"Jovem", "Adulto")
```

```
dados1=data.frame(dep, sex,est,tip,ida1,cens,tempo) #criar novo banco de dados
```

```
summary(dados1) #resumo - estatística descritiva das variáveis
```

```
#####ANÁLISE DESCRITIVA (EXPLORATÓRIA)
```

```
hist(tempo, main="",xlab="Tempo de permanência no plano (em anos)", ylab="Frequência")
```

```
#gerar histograma do tempo de permanência no plano
```

```
table(dep,sex)
```

```
table(dep,tip)
```

```
table(sex,tip)
```

```
table(dep,est)
```

```
table(sex,est)
```

```
table(tip,est)
```

```
table(sex,ida1)
```

```
table(tip,ida1)
```

```
table(est,ida1)
```

```
table(dep,ida1)
```

```
#####NÃO PARAMÉTRICO
```

```
#a) Kaplan-Meier
```

```
#Geral
```

```
ekm=survfit(Surv(tempo,cens)~1,conf.type="plain")
```

```
summary(ekm)
```

```
X11()
```



```

plot(ekm)
plot(ekm,main="",col=c("black","red","blue"),xlab="Tempo de permanência no plano (em
anos)",ylab="S(t) estimada")
legend(10,0.8,col=c("black","red","blue"),c("Pontual", "Intervalar LI", "Intervalar
LS"),lwd=1,bty="n",lty=c(1,2,2))
#Geral por sexo
ekm1=survfit(Surv(tempo,cens)~sex,conf.type="plain")
summary(ekm1)
X11()
plot(ekm1,main="",col=c("black","red","blue"),xlab="Tempo de permanência no plano(em
anos)",ylab="S(t) estimada",lty=c(1,3))
legend(10,0.8,col=c("black","red","blue"),c("Feminino", "Masculino"),lwd=1,bty="n")

#Geral por Dependência
ekm2=survfit(Surv(tempo,cens)~dep,conf.type="plain")
summary(ekm2)
X11()
plot(ekm2,main="",col=c("black","red","blue"),xlab="Tempo de permanência no plano(em
anos)",ylab="S(t) estimada",lty=c(1,3))
legend(10,0.8,col=c("black","red","blue"),c("Dependente", "Titular"),lwd=1,bty="n")

#Geral por Estado Civil
ekm3=survfit(Surv(tempo,cens)~est,conf.type="plain")
summary(ekm3)
X11()
plot(ekm3,main="",col=c("black","red","blue","green","yellow","grey"),xlab="Tempo de
permanência no plano(em anos)",ylab="S(t) estimada",lty=1:6)
legend(10,0.8,col=c("black","red","blue","green","yellow","grey"),c("Amasiado", "Casado", "
Não respondeu", "Separado", "Solteiro", "Viúvo"),lwd=1,bty="n",lty=1:6)

#Geral por Tipo de contratação
ekm4=survfit(Surv(tempo,cens)~tip,conf.type="plain")
summary(ekm4)
X11()
plot(ekm4,main="",col=c("black","red","blue"),xlab="Tempo de permanência no plano(em
anos)",ylab="S(t) estimada",lty=c(1:3))
legend(10,0.8,col=c("black","red","blue"),c("Coletivo - Adesão", "Coletivo -
Empresarial", "Individual / Familiar"),lwd=1,bty="n",lty=1:3)

#Geral por ida1
ekm5=survfit(Surv(tempo,cens)~ida1,conf.type="plain")
summary(ekm5)
X11()
plot(ekm5,main="",col=c("black","red"),xlab="Tempo de permanência no plano (em
anos)",ylab="S(t) estimada",lty=c(1:2))
legend(10,0.8,col=c("black","red"),c("Adulto", "Jovem"),lwd=1,bty="n",lty=1:2)

```

```
#Logrank --> Quando se rejeita H0, há evidências de que os grupos comparados são diferentes.
survdiff(Surv(tempo,cens)~sex,rho=0)#Há evidências que apontam para a diferença entre
grupos comparados.
survdiff(Surv(tempo,cens)~dep,rho=0)#Há evidências que apontam para a diferença entre
grupos comparados.
survdiff(Surv(tempo,cens)~est,rho=0)#Há evidências que apontam para a diferença entre
grupos comparados.
survdiff(Surv(tempo,cens)~tip,rho=0)#Há evidências que apontam para a diferença entre
grupos comparados.
survdiff(Surv(tempo,cens)~ida1,rho=0)#Há evidências que apontam para a diferença entre
grupos comparados.
```

#####SEMIPARAMÉTRICO

```
#Ajustes das funções de sobrevivência (não paramétrica e paramétricas)
```

```
ekm<-survfit(Surv(tempo,cens)~1)
```

```
time<-ekm$time
```

```
ajuste.km<-ekm$surv
```

```
#Ajuste do modelo de Cox (função de risco)
```

```
fit3=coxph(Surv(tempo,cens)~dep+est+sex+ida1,data=dados1,x=T,method="breslow")
```

```
summary(fit3)
```

```
#Função de sobrevivência
```

```
Ht3=basehaz(fit3,centered=F)
```

```
tempos3=Ht3$time
```

```
H03=Ht3$hazard
```

```
S03=exp(-H03)
```

```
#Pedir sobrevivências juntas (não paramétrica e semiparamétrica)
```

```
cbind(time,ajuste.km,S03)
```

```
#Curvas de sobrevivência estimada COX x curva de sobrevivência de kaplan meier
```

```
X11()
```

```
plot(ekm, conf.int=F, xlab="Tempo de permanência no serviço público", ylab="S(t)")
```

```
lines(c(0,time),c(1,S03), lty=2,col="red")
```

```
legend(1,0.9,lty=1:2,c("Kaplan-Meier","Cox"),bty="n",cex=0.8,col=c("black","red"))
```

```
#Teste de Schoenfeld de riscos proporcionais
```

```
cox.zph(fit3) #todas as covariáveis rejeitam H0
```