
ANÁLISE DE FIDELIZAÇÃO DE CLIENTES NO SEGMENTO DE SEGUROS UTILIZANDO O MÉTODO DE REGRESSÃO LOGÍSTICA

Bruna Moraes de Souza

Danielly Caroline Sandy

Karina Cavalcante Duarte

Virgínia Motta Aguiar

Raquel Cymrot

Universidade Presbiteriana Mackenzie (UPM)

Resumo

Este estudo tem por objetivo identificar, por meio do desenvolvimento de um modelo de regressão logística, os aspectos que caracterizam o perfil do cliente com maior probabilidade de declarar-se fiel à organização a qual possui o seguro do seu automóvel. O instrumento de coleta de dados foi um questionário respondido por moradores de um condomínio-clubes vertical localizado na cidade de São Paulo. A partir do modelo desenvolvido, foi possível afirmar que a fidelização de um cliente está relacionada às variáveis idade e obtenção de outros produtos da mesma seguradora.

Palavras-chave: Fidelização de clientes. Seguro de automóvel. Regressão logística.

1 INTRODUÇÃO

A violência e, em particular, os roubos na cidade de São Paulo têm levado as pessoas a buscar cada vez mais proteção tanto pessoal quanto de bens. Com uma participação de aproximadamente 6% no Produto Interno Bruto (PIB) em 2012 (ESCOLA NACIONAL DE SEGUROS FUNENSEG, 2014), o setor de seguros tem apresentado índices crescentes na economia brasileira desde a década de 1990, pois a estabilidade econômica proveniente da implantação do Plano Real nessa época propiciou o desenvolvimento do mercado de seguros no Brasil, uma vez que a aquisição de um seguro passou a ser uma opção atraente para minimizar os riscos oriundos de eventos inesperados (PINTO, 2006).

A entrada de multinacionais no mercado local, com a vantagem de terem experiência em mercados mais maduros, tornou a competição mais acirrada nessa indústria, de modo a provocar maior preocupação com estratégias e mudança cultural para o atendimento das novas demandas (PINTO, 2006).

Em 2013, o mercado contava com aproximadamente 121 seguradoras e 83 mil corretoras de seguros (ESCOLA NACIONAL DE SEGUROS FUNENSEG, 2014), dados que reforçam a grande competitividade desse setor.

Alcançar a fidelidade dos consumidores constitui uma importante ferramenta para garantir vantagem competitiva às companhias, visto que clientes fiéis tendem a comprar mais frequentemente e em maior quantidade, sendo também menos sensíveis ao aumento de preços e às ofertas dos concorrentes (ALRUBAIEE; AL-NAZER, 2010).

Nesse contexto, surgiu a necessidade de analisar o comportamento dos consumidores para determinar variáveis que influenciam a fidelização dos clientes no segmento de seguros de automóvel, respondendo ao seguinte problema de pesquisa: Existe um conjunto de indicadores que caracterizam a fidelização de clientes no mercado de seguro privado de automóveis?

2 FIDELIZAÇÃO DE CLIENTES

A palavra fidelidade, do latim *fidēlitas*, tem seu significado atribuído à qualidade daquele ou daquilo que é fiel ou leal, o que pode ser associado aos seguintes aspectos: preservação do estado original, firmeza, constância, confiança e perseverança (FERREIRA, 2010).

De acordo com Dick e Basu (1994 apud DONIO; MASSARIO; PASSIANTE, 2006, p. 445), a fidelidade pode ser interpretada como a intensidade da relação de um indivíduo com uma empresa.

No âmbito acadêmico, o conceito e a medição da lealdade em serviços evoluíram de maneira complexa. Em seu nível generalizado, a fidelização é um reflexo de tendências dos vários clientes em relação ao serviço prestado pela empresa, e, primeiramente, sua conceituação foi atribuída ao comportamento de recompra ou intenção de comutação. Entretanto, Bass (1974 apud VELOSO, 2008, p. 27) alerta que é preciso ressaltar que o processo de recompra envolve uma análise pouco precisa, pois inclui partes aleatórias, levando a crer que se trata de um conceito subjetivo e de difícil mensuração.

A definição de lealdade inclui tanto o comportamento de recompra quanto a disposição para atitude positiva em relação a um fornecedor. Posteriormente, além do comportamento e da atitude, a lealdade passou a abranger também o aspecto cognitivo, refletindo as crenças do consumidor no tocante à marca e exclusividade do provedor de serviço. Pesquisas de *marketing* já consideram que a fidelização é um construto multidimensional, que compreende outros aspectos, destacando-se os associados à cognição, variando de acordo com o indivíduo (JONES; SASSER JR., 1995).

A satisfação é um fator que antecede a fidelização (LARÁN; ESPINOZA, 2004). Para muitas empresas, é o único parâmetro para acompanhar a base de clientes fiéis da organização, todavia a satisfação pessoal é uma régua inexata, e, para alcançar a fidelização, deve-se ir além do item satisfação (JONES; SASSER JR., 1995).

A fidelização em seguros comporta-se de forma muito peculiar perante os demais tipos de serviço. O fato de a venda do seguro possuir um intermediador que, por regulamentação é o corretor de seguros, dificulta um relacionamento mais próximo entre seguradora e cliente, de modo que a seguradora deve se preocupar em fidelizar não só o cliente, mas também o corretor de seguros.

O fato de haver muitas companhias oferecendo produtos com baixa diferenciação sugere, de acordo com as estratégias genéricas de Porter (2005), apelo pela liderança em custo. Nesse tipo de estratégia, é comum as empresas enxugarem os custos para oferecer preços menores, de modo a garantir maior competitividade no mercado.

Outra evidência da estratégia de liderança em custo é o fato de o seguro ser um produto que o cliente adquire sem a intenção de utilizar, havendo baixa interação entre cliente e empresa, o que dificulta uma entrega de serviço com maior valor agregado.

Dessa forma, a tendência é que a escolha de aquisição ou renovação seja baseada em heurísticas de imitação e recomendação. Na perspectiva do detentor da apólice de seguro, não há custo de saída, deixando-o confortável em relação à mudança de fornecedor, e tais aspectos justificam a dificuldade das empresas em fidelizar seus clientes.

3 REGRESSÃO LOGÍSTICA

A análise de regressão propicia um maior entendimento a respeito da relação existente entre duas ou mais variáveis. Quando a variável de saída que se deseja obter tem somente dois resultados possíveis, o modelo de regressão logística é o mais apropriado para a estimação de uma resposta.

A regressão logística é um caso particular de regressão na qual a variável resposta Y é dicotômica, ou seja, pode assumir apenas dois resultados mutuamente excludentes, como bom ou mau, sucesso ou fracasso, aceitar ou rejeitar, entre outros (HAIR et al., 2009).

Na regressão logística, em geral, atribui-se o código um ao resultado de interesse, denominado sucesso, que representaria a presença de uma característica do evento-alvo, e zero ao evento complementar, denominado fracasso. A regressão logística é utilizada quando o objetivo é estabelecer uma relação entre uma variável resposta dicotômica e uma ou mais variáveis explicativas que podem ser tanto qualitativas como quantitativas. Quando de interesse de uma aplicação, a regressão logística pode ser utilizada mesmo quando a resposta não é originalmente binária, bastando para isso dicotomizar a variável resposta (HOSMER; LEMESHOW; STURDIVANT, 2000). A representação matemática de uma regressão linear múltipla é dada pela Equação 1:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n + \varepsilon \quad (1)$$

onde Y é a variável dependente; x_1, x_2, \dots, x_n , as variáveis independentes; $\beta_0, \beta_1, \beta_2, \dots, \beta_n$, os coeficientes das variáveis independentes; e ε , o erro ou resíduo da regressão.

Os métodos empregados na regressão logística seguem os mesmos princípios gerais usados na regressão linear, apesar de apresentarem respostas e suposições de modelos diferentes (FARHAT, 2003). De maneira mais ampla, a regressão logística atribui, à variável resposta, um intervalo compreendido entre zero e um, que pode ser interpretado como a probabilidade de ocorrência de determinado evento. Com base na regra de decisão, resultados superiores a 0,5 são assumidos como possível sucesso e inferiores como possível fracasso (DIAS FILHO; CORRAR, 2007). Tal relação se deve à transformação da variável resposta em razão de chance (probabilidade) e a uma transformação logarítmica, como apresentado a seguir. Seja $g(x_i)$ uma função linear simples, isto é (Equação 2):

$$g(x_i) = \beta_0 + \beta_1x_i \quad (2)$$

onde β_0 e β_1 são os coeficientes das variáveis independentes; e x_i , as variáveis independentes.

A probabilidade de a variável resposta ser classificada como sucesso, em razão de o valor da variável explicativa ser igual a x_i , é dada pela Equação 3 (HOSMER; LEMESHOW; STURDIVANT, 2000):

$$\pi(x_i) = P(Y = 1|x_i) = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \quad (3)$$

No modelo de regressão logística simples, ajusta-se uma função da variável resposta, $g(x)$. Para as estimações dos parâmetros β_0 e β_1 , utiliza-se a função de máxima verossimilhança.

Adota-se o modelo de regressão logística múltipla, conforme Equação 4, quando há mais de uma variável explicativa (qualitativas ou quantitativas).

$$g(x_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} \quad (4)$$

Uma vez estabelecido o modelo de regressão logística, são necessários o teste da significância dos parâmetros e o ajuste do modelo. Dentre alguns possíveis modelos, deve-se escolher o mais simples, desde que tenha um bom ajuste (princípio da parcimônia).

Após a determinação do modelo de regressão logística, faz-se necessário realizar alguns testes que assegurem sua validade. Esses testes permitem reconhecer tanto as variáveis que influenciam diretamente a estimação dos parâmetros da regressão quanto aquelas que não se ajustam bem e, portanto, devem ser descartadas do modelo final.

O nível descritivo do teste, ou valor-p, indica a probabilidade de um valor no mínimo ser tão extremo quanto o valor referente ao do teste que acarreta a rejeição da hipótese nula a partir dos dados obtidos. No caso de o valor-p ser menor que o nível de significância do teste, a hipótese nula será rejeitada (MONTGOMERY; RUNGER, 2012).

4 DESENVOLVIMENTO DA PESQUISA

A pesquisa foi realizada em duas etapas. Em um primeiro momento, realizou-se uma entrevista qualitativa com três profissionais do segmento de seguro de automóveis, com o intuito de fazer um levantamento acerca de como a fidelização é compreendida e explorada no mercado em questão. Em seguida, obtiveram-se indicadores para compor o instrumento de coleta de dados da segunda etapa da pesquisa para maior investigação.

4.1 Entrevista com os profissionais do segmento de seguros

Foram realizadas entrevistas com três profissionais atuantes no mercado de seguros de automóveis.

O profissional A coordena há um ano a área de vendas e gerenciamento da carteira de automóveis de uma corretora de médio porte, tendo experiência de cinco anos em uma seguradora de grande porte no mesmo produto, além de já ter trabalhado anteriormente em outra corretora por aproximadamente dois anos.

O profissional B atua há quatro anos no mercado de seguros e é responsável por gerenciar uma equipe de planejamento e definição de estratégias de mercado em uma seguradora multinacional de grande porte, na qual a maior parte das estratégias delineadas tem foco em produtos massificados, sendo automóvel incluso nesse bloco.

O profissional C atua há dois anos no mercado em questão e é analista sênior de sinistro de automóvel em uma companhia de grande porte.

Quando questionados acerca dos fatores que influenciam a fidelização de clientes no mercado de seguros de automóveis, foram obtidas respostas semelhantes. Os três entrevistados concordam que a fidelização está intimamente ligada à satisfação com o atendimento, o preço do seguro, a qualidade da assistência técnica, a agilidade e a eficiência no atendimento e na solução dos problemas reportados.

O profissional A salientou que o seguro é um serviço que apresenta características peculiares em relação a outros, pois o cliente o adquire com intuito de não utilizá-lo.

Citou ainda que, no tocante à fidelização, no momento de adquirir ou renovar o seguro, os clientes podem ser classificados em dois perfis distintos. Em um deles, o cliente é indiferente à seguradora, uma vez que, no período de vigência da apólice, não houve necessidade de utilizar os serviços disponibilizados pela seguradora, ou seja, como não houve contato com ela, o indivíduo não dispõe de alguns parâmetros necessários para avaliar seu grau de satisfação com o serviço.

No segundo perfil, estão inclusos os clientes que já tiveram algum tipo de contato com a seguradora, seja para assistência ou sinistro. Logo, esses clientes de fato utilizaram o seguro e possuem mais parâmetros para avaliar o serviço prestado, o que o profissional chamou de “momento da verdade”, caracterizado pelo momento em que a seguradora tem a oportunidade de estabelecer um relacionamento cordial e surpreender positivamente o cliente, aumentando consideravelmente as chances de fidelizá-lo.

Outro ponto destacado pelo profissional A foi a importância do corretor de seguros como um possível influenciador do cliente. O seguro de automóvel é um produto cujas características não são conhecidas, em sua totalidade, pelo grande público. Dessa forma, fica a encargo do corretor, como um representante legal do cliente perante as seguradoras, apresentar-lhe propostas que atendam às suas necessidades e auxiliá-lo

tanto no momento da aquisição quanto na renovação do seguro. O profissional ainda afirmou que, com base em sua experiência, quanto menor for o conhecimento do cliente sobre o produto, maior será o poder de influência sobre ele.

De acordo com o profissional B, o produto automóvel não é rentável para as companhias de seguro e, muitas vezes, a receita obtida apenas paga a operação. Entretanto, o seguro automóvel é o “produto de prateleira”, ou seja, é o produto que toda seguradora que atua no mercado de varejo deve ter, por ser o mais conhecido entre os consumidores e ter a maior demanda no segmento. Esse contexto faz com que a fidelização em seguro de automóvel seja importante para as seguradoras do ponto de vista da venda cruzada. Primeiramente, a empresa oferta o seguro automóvel e, depois, oferece outros produtos que sejam mais lucrativos, como previdência privada, seguro-saúde, seguro-garantia, entre outros.

4.2 Metodologia: amostragem, coleta e forma de análise dos dados

Para o desenvolvimento do instrumento de coleta de dados, foram estabelecidas as seguintes dimensões: identificação do pesquisado, identificação do automóvel e da seguradora, relacionamento com o corretor de seguros, relacionamento com a seguradora e aspectos associados ao comportamento de fidelização. Tais dimensões abordaram as variáveis apresentadas nos parágrafos seguintes.

Para a dimensão de identificação do pesquisado, foram explorados aspectos de caracterização da amostra, como idade, escolaridade e renda, considerando a classificação de faixas definidas pelo Instituto Brasileiro de Geografia e Estatística (2013).

Para a dimensão identificação do automóvel e da seguradora, abordaram-se questões que pudessem especificar o automóvel com as características de modelo e ano de fabricação, tempo em que o pesquisado utiliza o serviço de seguro de automóveis, o número de seguradoras com as quais ele já firmou um contrato de seguro, atual seguradora e o tempo de permanência nesta última.

Para a dimensão do relacionamento com o corretor, elaboraram-se questões para avaliar o grau de satisfação do respondente com o atual corretor de seguros, indicando itens como: importância dada pelo corretor ao cliente, disponibilidade do corretor para esclarecer dúvidas, rapidez e eficiência no atendimento dado pelo corretor, importância dada à opinião do corretor e disposição do cliente em aceitar outras opções apresentadas pelo corretor. Para avaliar o grau de satisfação, foi estabelecida uma faixa com seis níveis: muito insatisfeito, insatisfeito, indiferente, satisfeito, muito satisfeito e não sei avaliar, sendo o item escolhido indicado por um “x”.

Para a última dimensão, relacionamento com a seguradora, avaliaram-se o grau de satisfação com a utilização assistência técnica e o grau de importância para alguns itens que podem acompanhar um contrato de seguro, como carro reserva; proteção

para acessórios instalados no veículo por roubo, furto, incêndio; proteção a danos causados em vidros, retrovisores, lanternas e faróis; consertos/pequenos reparos em casa; assistência a danos a passageiros; assistência especializada para mulheres. Nesta última dimensão, também se verificou a utilização de outros produtos da atual seguradora, como previdência privada, seguro de vida, seguro-saúde ou outros, além do nível de satisfação após acionar a seguradora em caso de sinistro.

Para finalizar o questionário, foi apresentado um conjunto de sete afirmativas para testar o comportamento de fidelização por meio de uma resposta binária, concordo ou discordo, a qual contava com os seguintes itens: confiança na seguradora, gostar da seguradora, recomendá-la, pretensão em continuar com ela desde que o preço não seja alterado, pretensão em continuar com a seguradora mesmo que haja um pequeno aumento no valor e pretensão de continuar independentemente do aumento de preço. A pesquisa era concluída com a afirmativa principal: “Considero-me fiel à minha seguradora”.

Para a realização deste trabalho, foi necessária a predefinição da amostra: um condomínio-clubes vertical localizado na zona leste da cidade de São Paulo. Embora a escolha de condomínio-clubes tenha se realizado por conveniência, tal amostra apresenta características de uma amostragem por conglomerado.

Na amostragem por conglomerado, podem-se identificar possíveis segmentações de mercados e características semelhantes ou diferentes em regiões ou grupos com requisitos comuns. Entretanto, esses grupos ou regiões não são predefinidos (DIAS FILHO; CORRAR, 2007).

Foram coletados dados com 176 questionários, sendo 118 para a elaboração do modelo e 58 para a verificação do ajuste na previsão do modelo. O tamanho da amostra se justifica, pois, na regressão logística, é feita uma aproximação assintótica, e, caso as amostras sejam pequenas, pode-se não obter uma solução a partir dessa análise (FARHAT, 2003).

A coleta dos dados ocorreu em janeiro e fevereiro de 2013. Após essa fase, os dados foram tabulados, parametrizados e consolidados para a análise descritiva e o desenvolvimento do modelo de regressão logística, desencadeando na escolha das variáveis.

Como algumas pessoas se recusaram a responder à pesquisa, a amostragem não pode ser considerada sistêmica, sendo, portanto, não probabilística. Além disso, há características de uma pesquisa quase aleatória (BOLFARINE; BUSSAB, 2005).

Realizou-se uma análise descritiva dos dados coletados, e construíram-se intervalos com nível de confiança de 95% para algumas proporções de interesse. Em seguida, realizaram-se testes de qui-quadrado de independência entre pares de variáveis aleatórias pertinentes, com o intuito de verificar a relação entre elas. Optou-se por testar modelos de regressão logística com as variáveis cujas chances de terem bons índices de ajuste e de poder de previsibilidade fossem maiores.

Para a análise dos resultados, adotaram-se intervalos com confiança de 95%, e, para todos os testes de hipótese realizados, calcularam-se seus respectivos níveis descritivos (valor-p), sendo rejeitadas as hipóteses com $p < 0,05$, igual ao nível de significância adotado.

A análise estatística dos dados foi realizada com o auxílio do programa Minitab®.

5 ANÁLISE DOS RESULTADOS

A pesquisa foi realizada em um condomínio-clubes vertical formado por 28 torres, totalizando 1.792 apartamentos. Os apartamentos são de dois e três dormitórios, e o condomínio-clubes está situado na zona leste de São Paulo, com aproximadamente dez anos de existência.

A pesquisa foi aplicada a 176 indivíduos, dos quais 118 são homens, o que representa 67,04% da totalidade (IC = [60,10;73,99]).

Essa diferença representativa possivelmente se deve ao fato de que, em muitas famílias, o homem é responsável por contratar e administrar serviços que envolvem o carro, o qual é, muitas vezes, um bem compartilhado.

A análise descritiva do perfil do morador de um condomínio vertical classificado como popular faz-se coerente, pois, nesse universo amostral, é possível encontrar pessoas com renda, escolaridade e idade totalmente diferentes, sendo 40% dos respondentes com salário a partir de R\$ 4.746,00 por pessoa e mais de 72% com ensino superior ou pós-graduação. Dentre os pesquisados, mais de 50% possuem automóveis fabricados a partir de 2010, denotando que um carro mais novo tem uma necessidade maior de ser protegido devido ao valor investido e que a maioria das seguradoras mantém um padrão de assegurar frotas de até cinco anos de fabricação.

Quando questionados acerca da quantidade de seguradoras contratadas até o momento, 60 pessoas (34%) responderam ter utilizado apenas uma seguradora, enquanto 56 pessoas (32%) já contrataram duas seguradoras.

Para a amostra em questão, 72 pessoas possuem seguro de automóvel com a seguradora líder do mercado (Porto Seguro), representando uma fatia de 41%: (IC = [33,65;48,17]).

Foi solicitado que atribuíssem notas de 1 a 5, em uma escala de satisfação, à importância que o corretor dá ao cliente. Nesse caso, 81% declaram-se satisfeitos ou muito satisfeitos com a importância recebida do corretor.

No processo de renovação do seguro de automóvel, o papel do corretor mostrou-se extremamente importante: 85% (IC = [79,33;89,98]) dos pesquisados afirmaram que estão dispostos a mudar de seguradora caso o corretor desta apresente uma opção melhor.

Foram realizados os testes de independência entre a variável declarar-se fiel e as demais variáveis pertinentes. O teste qui-quadrado foi aplicado aos pares de variáveis, e, a seguir, são apresentadas as conclusões obtidas. Nesses itens, considerando um nível de significância de 5%, foram rejeitadas as hipóteses nulas de independência.

A seguir, serão comentadas algumas relações de dependências entre declarar-se fiel e outras variáveis encontradas na análise.

As variáveis declarar-se fiel e renda são dependentes ($p = 0,006$), isto é, proporcionalmente quem não declara a renda declara-se menos fiel, e quem tem renda acima de cinco salários mínimos declara-se mais fiel.

Proporcionalmente, declaram-se mais fiéis os pesquisados que já utilizaram assistência técnica ($p = 0,005$), os satisfeitos com o atendimento da assistência técnica ($p = 0,008$), os satisfeitos com a qualidade da assistência técnica ($p = 0,010$) e os que possuem outros produtos além do seguro de automóvel com a mesma seguradora ($p = 0,000$).

Também proporcionalmente se declararam mais fiéis os pesquisados que gostam da seguradora atual ($p = 0,010$), que a recomendariam ($p = 0,019$), que renovariam com ela mesmo com pequeno aumento ($p = 0,000$) ou independentemente do preço ($p = 0,000$), que afirmaram ter maior idade ($p = 0,001$), que possuem seguro há mais tempo ($p = 0,043$) e que mencionaram que tiveram menor número de seguradoras ($p = 0,001$).

Os profissionais da área que foram entrevistados afirmaram em consenso que a figura do corretor apresenta significativa parcela no processo de renovação do seguro auto, porém as variáveis relacionadas a influência, opinião, atendimento e disponibilidade do corretor em atender o cliente não apresentaram dependência com a variável-chave ser fiel.

Dentre os diversos modelos de regressão logística testados para a predição da variável sou fiel, um modelo que utiliza as variáveis ter outros produtos da mesma seguradora e idade do pesquisado apresentou resultados mais favoráveis. A variável foi denominada sou fiel, entretanto seu significado deve ser interpretado como cliente que se declara fiel.

O modelo encontrado foi:

$$Y_{\text{sou fiel}} = -2,06157 + 1,30321 X_{\text{outros produtos}} + 0,0428373 X_{\text{idade}} \quad (5)$$

A Tabela 1 apresenta os resultados do modelo dessa regressão logística binária.

TABELA 1

Coefficientes, erro padrão do coeficiente, valor Z, nível descritivo P, razão de chances e intervalo com 95% de confiança para a razão de chances

Preditor	Coefficiente	Erro padrão do coeficiente	Z	P	Razão de chances	Limite inferior	Limite superior
Constante	-2,06157	0,805989	-2,56	0,011			
Outros produtos	1,30321	0,446307	2,92	0,004	3,68	1,53	8,83
Idade	0,0428373	0,0215502	1,99	0,047	1,04	1	1,09

Fonte: Elaborada pelas autoras.

Valores maiores que 1 para a razão de chances indicam maior associação entre a ocorrência do fator e a ocorrência da variável resposta. Para a variável outros produtos, uma razão de chances igual a 3,68 indica que haverá um aumento de 268% na chance de sucesso da variável resposta (sou fiel) quando a variável regressora assumir o valor 1, isto é, quando o cliente tiver outros produtos. Propõe-se que, em nova pesquisa, tal variável seja perguntada de forma a mensurar a quantidade de produtos diferentes que o cliente possui da companhia, o que poderia melhorar ainda mais esse modelo. Já para a variável regressora idade, uma razão de chances igual a 1,04 indica que há um aumento de 4% na chance de sucesso da variável resposta (sou fiel) para cada ano a mais na idade do cliente, ou seja, clientes mais velhos têm maior probabilidade de ser fiéis.

A probabilidade de um cliente considerar-se fiel é:

$$P(\text{sou fiel}) = \frac{e^y}{1 + e^y} \quad (6)$$

Se essa probabilidade for superior a 0,50, o indivíduo será classificado como um cliente que se declara fiel. Por exemplo, para um indivíduo que possui outros produtos com a seguradora e tem 40 anos de idade, tem-se o seguinte resultado:

$$Y_{\text{sou fiel}} = -2,06157 + 1,30321 \times 1 + 0,0428373 \times 40 = 0,9551 \quad (7)$$

Tal cliente terá então a seguinte probabilidade estimada de declarar-se fiel:

$$P(\text{sou fiel}) = 0,7221 > 0,5 \quad (8)$$

Logo, esse cliente será classificado como fiel. A estatística G testa a hipótese de que todos os coeficientes são iguais a zero contra a hipótese de que ao menos um deles é diferente de zero. O *software* forneceu o valor $G = 15,420$ como seu respectivo nível descritivo p igual a 0,000. Portanto, pode-se afirmar que há evidências de que pelo menos um dos coeficientes da regressão é diferente de zero.

O modelo foi submetido aos três diferentes testes de adequação. O método de Pearson apresentou valor- $p = 0,566$; o método de *deviance*, um valor- $p = 0,156$; e o método de Hosmer-Lemeshow, um valor- $p = 0,833$. Portanto, como todos os indicadores foram maiores que 0,05, concluiu-se, ao nível de significância de 5%, que o modelo foi bem ajustado.

Calcularam-se várias medidas de associação entre a variável resposta e as probabilidades de previsão.

Para os dados da amostra usada na escolha do modelo, há 55 indivíduos que declararam ser fiéis à marca contra 63 que se declararam não ser fiéis. Podem-se formar assim $55 \times 63 = 3.465$ pares com valores de respostas diferentes. Com base no modelo, um par é considerado concordante se o indivíduo que se considera fiel à marca tem probabilidade maior de ser classificado como tal do que aquele que se declara não ser fiel. A porcentagem de concordância do modelo foi de 68,7%. Também foram feitos os cálculos das estatísticas de Somers' D, com valor igual a 0,39, Goodman-Kruskal, com valor igual a 0,40, e Kendall's Tau-a, com valor igual a 0,20. Essas estatísticas resumem as tabelas de concordância e discordância dos pares de observações e devem estar compreendidas entre 0 e 1. Valores altos indicam que o modelo tem uma boa capacidade de previsão. Com exceção do teste de Kendall's Tau-a que apresentou índice baixo, os índices apresentaram poder de previsão satisfatório.

A amostra de teste foi utilizada para calcular a probabilidade da variável sou fiel nos 58 indivíduos restantes. Como era conhecida a verdadeira situação quanto à variável sou fiel, pode-se calcular a porcentagem de acerto que foi igual a 69%. Sem realizar nenhuma estatística, a probabilidade de o cliente declarar-se fiel era de 57,95%, faltando 42,05% para completar os 100%. O modelo explicou 11,05% a mais a variável resposta no caso em que não seja feita nenhuma das estatísticas, o que corresponde a 19,07% a mais de melhoria no poder preditivo. Dentre os clientes que se declararam não fiéis, o modelo acertou em 84,21% dos casos, o que representa 45,32% de melhoria no reconhecimento de clientes não fiéis.

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

O modelo obtido teve índices de ajuste e de poder de previsibilidade que atendem às necessidades.

O fato de os clientes que se declaram fiéis possuírem outros produtos da mesma companhia confirma a tática praticada no mercado segurador de que a fidelização em seguro de automóvel é vantajosa do ponto de vista de facilitar a oferta de outros produtos mais lucrativos para a companhia, ou seja, ações de venda cruzada.

A idade também é um fator que o modelo associa à fidelização. Possivelmente, clientes mais velhos estão mais propensos a se declarar fiéis porque tendem a ter um comportamento que visa atenuar o risco de uma escolha inadequada. Por isso, optam por manter o estado atual, ou seja, continuar com a seguradora atual. Entretanto, deveria ser feita uma investigação mais profunda acerca desse fenômeno para confirmar tal hipótese.

O método de regressão logística para previsão de fidelização de clientes, quando desenvolvido a partir de uma base de dados de clientes de apenas uma companhia, pode ser explorado como um recurso para estimativa de fluxo de caixa futuro, considerando que a demanda por um produto ou serviço se constitui por consumidores novos e clientes antigos. Conhecida a proporção desses dois tipos de cliente, a companhia pode efetuar um planejamento mais assertivo do uso de seus recursos. Portanto, o grupo sugere que esse aspecto seja explorado em pesquisas futuras.

CUSTOMER LOYALTY ANALYSIS IN INSURANCE SEGMENT USING LOGISTIC REGRESSION METHOD

Abstract

This study aims to identify aspects that characterize the profile of the customer most likely to declare themselves loyal to the organization which insurances their car, through the development of a logistic regression model. The data collection instrument was a questionnaire answered by residents of a vertical condominium club located in the São Paulo city. From the model developed, it was possible to say that the loyalty of a customer is related to age and having other products from the same insurer.

Keywords: Customer loyalty. Auto insurance. Logistic regression.

REFERÊNCIAS

ALRUBAIEE, L.; AL-NAZER, N. Investigate the impact of relationship marketing orientation on customer loyalty: the customer's perspective. *International Journal of Marketing Studies*, Toronto, v. 2, n. 1, p. 155-174, May 2010.

BOLFARINE, H.; BUSSAB, W. O. *Elementos de amostragem*. ABE-Projeto Fisher. São Paulo: Edgard Blücher, 2005.

DIAS FILHO, J. M.; CORRAR, L. J. Regressão logística. In: CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. (Coord.). *Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia*. São Paulo: Atlas, 2007.

- DONIO, J.; MASSARIO, P.; PASSIANTE, G. Customer satisfaction and loyalty in a digital environment: an empirical test. *Journal of Consumer Marketing*, v. 23, n. 7, p. 445-457, 2006.
- ESCOLA NACIONAL DE SEGUROS FUNENSEG. *Tudo sobre seguros. Fatos e indicadores de mercado: estabilidade e crescimento*. São Paulo, 2014. Disponível em: <<http://www.tudosobreseguros.org.br/sws/porta/pagina.php?l=267>>. Acesso em: 17 mar. 2015.
- FARHAT, C. A. V. *Análise de diagnóstico em regressão logística*. 2003. Dissertação (Mestrado em Estatística)—Universidade de São Paulo, São Paulo, 2003.
- FERREIRA, A. B. H. *Dicionário Aurélio da Língua Portuguesa*. 5. ed. Curitiba: Positivo, 2010.
- HAIR, J. F. et al. *Análise multivariada de dados*. 6. ed. Porto Alegre: Bookman, 2009.
- HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X. *Applied logistic regression*. 2. ed. New York: John Wiley & Sons, 2000.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Características demográficas*. Rio de Janeiro: IBGE. Disponível em: <http://biblioteca.ibge.gov.br/visualizacao/livros/liv64529_cap3.pdf>. Acesso em: 9 ago. 2013.
- JONES, T. O.; SASSER JR., W. E. Why satisfied customers defect. *Harvard Business Review*, v. 73, n. 6, p. 88-99, Nov./Dec. 1995.
- LARÁN, J. A.; ESPINOZA, F. S. Consumidores satisfeitos, e então? Analisando a satisfação como antecedente da lealdade. *Revista de Administração Contemporânea*, Curitiba, v. 8, n. 2, p. 51-70, abr./jun. 2004.
- MONTGOMERY, D. C.; RUNGER, G. C. *Estatística aplicada e probabilidades para Engenheiros*. 5. ed. Rio de Janeiro: LTC, 2012.
- PINTO, M. R. F. *Estratégia competitiva no mercado de seguros: estudo de caso na Cia Marítima Seguros S.A.* 2006. Trabalho de Conclusão de Curso (Bacharelado em Ciências Econômicas)—Universidade Federal de Santa Catarina, Florianópolis, 2006.
- PORTER, M. E. *Estratégia competitiva: técnicas para análise de indústrias e da concorrência*. Rio de Janeiro: Campus, 2005.
- VELOSO, A. D. S. *Fidelidade em telefonia celular: proposição e validação de um índice para previsão da fidelidade de clientes*. 2008. Dissertação (Mestrado em Administração de Empresas)—Faculdade de Estudos Administrativos de Minas Gerais, Belo Horizonte, 2008.

Contato

Bruna Moraes de Souza
bruna-msouza@hotmail.com

Tramitação

Recebido em agosto de 2013.
Aprovado em agosto de 2015.