

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CÂMPUS CURITIBA
CURSO DE ESPECIALIZAÇÃO DE ENGENHARIA DA CONFIABILIDADE

TIAGO EDELMO DE LIZ ESTACIO

ANÁLISE DA CONFIABILIDADE DE SERVIÇOS DE COBRANÇA VIA
CALL CENTER

MONOGRAFIA DE ESPECIALIZAÇÃO

CURITIBA

2016

TIAGO EDELMO DE LIZ ESTACIO

ANÁLISE DA CONFIABILIDADE DE SERVIÇOS DE COBRANÇA VIA
CALL CENTER -

Monografia apresentada ao Programa de Especialização em Engenharia da Confiabilidade da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do título de Especialista.

Orientador: Dr. Carlos Henrique Mariano

CURITIBA

2016

TERMO DE APROVAÇÃO

TIAGO EDELMO DE LIZ ESTACIO

ANÁLISE DA CONFIABILIDADE DE SERVIÇOS DE COBRANÇA VIA CALL CENTER

Monografia de conclusão do Curso de especialização de Engenharia da Confiabilidade, Turma CEEC 2014 da Universidade Tecnológica Federal do Paraná, aprovado pela seguinte banca examinadora:

Professor Dr. Carlos Henrique Mariano
DAELT – UTFPR – Câmpus Curitiba
Orientador

Professor Dr. Emerson Rigoni
DAELT – UTFPR – Câmpus Curitiba
Banca

Professor Dr. Marcelo Rodrigues
DAELT – UTFPR – Câmpus Curitiba
Banca

AGRADECIMENTOS

Agradeço primeiramente a Deus pela proteção e sabedoria durante o curso.

Ao meu orientador Professor Carlos Mariano, pela dedicação e paciência a mim prestada.

A todos os professores da UTFPR e profissionais da Reliasoft que contribuíram na conclusão da minha especialização.

A minha família e amigos, pelo apoio no decorrer da especialização.

RESUMO

ESTACIO, Tiago. **Análise da confiabilidade de serviço de cobrança via Call Center**. Monografia (especialização em Engenharia da Confiabilidade) -Programa de Pós-Graduação, Universidade Tecnológica Federal do Paraná. Curitiba, 2016.

Esta monografia apresenta um estudo sobre a aplicação de métodos da engenharia da confiabilidade em determinada estratégia de central de call center, chamada de campanha. Para a análise descritiva dos dados foi considerado o gráfico “*box-plot*”, identificando assim, os dados com discrepância chamados de “*outliers*”. Ainda, foi aplicado os conceitos de análise de dados de vida (ADV) em três conjuntos de dados: (1) sem dados discrepantes, (2) com dados de discrepância moderada e (3) conjunto completo. Traz como resultado, a confiabilidade da campanha e a influência dos “*outliers*” no estudo.

Palavras-chaves: Confiabilidade, *Box-plot*, *outliers*, discrepância, dados de vida.

ABSTRACT

ESTACIO, Tiago. **Análise da confiabilidade de serviço de cobrança via Call Center**. Monografia (especialização em Engenharia da Confiabilidade) -Programa de Pós-Graduação, Universidade Tecnológica Federal do Paraná. Curitiba, 2016.

The present monograph demonstrates the application of reliability engineering methods applied in call center strategies, commonly named as campaign. For descriptive data analysis was considered the "box-plot" graphs, thus identifying the data discrepancy called "outliers". The life data analysis concepts (ADV) was applied in three datasets: (1) without outliers, (2) data with moderate discrepancy and (3) complete dataset, with the purpose to evaluate the campaign reliability and the "outliers" influence in the study.

Keywords: reliability, Box-plot, outliers, discrepancy, data analysis.

LISTA DE SIGLAS

ADV	Análise de dados de vida
CDC	Código de defesa do consumidor
CIC	Ciclos
MLFV	Função de verossimilhança
SPC	Serviço Nacional de Crédito

LISTA DE FIGURAS

Figura 1 - Exemplo de "script" de Cobrança.....	18
Figura 2 - Exemplos de distribuições contínuas	20
Figura 3 - Definições do gráfico Box-Plot	22
Figura 4 - Gráficos "box-plot" de quantidade de ligações por pagamento	25
Figura 5 - Teste de aderência para o conjunto de dados sem as observações discrepantes.....	29
Figura 6 - Teste de aderência para o conjunto de dados sem as observações com discrepância alta.	29
Figura 7 - Teste de aderência para o conjunto de dados com as observações com discrepância alta.	30

LISTA DE GRÁFICOS

Gráfico 1 - Recuperação de credito.....	11
Gráfico 2 - Comportamento da variável “ <i>Campaign</i> ”	23
Gráfico 3 - Comportamento da variável " <i>Campaing</i> " por escolaridade	24
Gráfico 4 - Comportamento da variável " <i>Campaing</i> " por profissão.....	24
Gráfico 5 - Conjunto contendo dados discrepantes	26
Gráfico 6 - Conjunto não contendo dados discrepantes	26
Gráfico 7 - Conjunto não contendo dados com discrepância alta	27
Gráfico 8 - Gráfico de dispersão do conjunto de dados sem as observações discrepantes.....	32
Gráfico 9 - Gráfico de dispersão do conjunto de dados sem as observações de alta discrepância.....	33
Gráfico 10 - Gráfico de dispersão do conjunto de dados com as observações de alta discrepância.....	33
Gráfico 11 - Confiabilidade do conjunto de dados sem as observações discrepantes	34
Gráfico 12 - Confiabilidade do conjunto de dados sem as observações de alta discrepância.....	34
Gráfico 13- Confiabilidade do conjunto de dados com as observações de alta discrepância.....	35

LISTA DE TABELAS

Tabela 1 - Exemplo de tabulação dos dados no programa Weibull ++	28
Tabela 2 - Parâmetros da distribuição Weibull para o conjunto de dados sem as observações discrepantes.....	31
Tabela 3 - Parâmetros da distribuição Weibull para o conjunto de dados sem as observações de discrepância alta	31
Tabela 4 - Parâmetros da distribuição Weibull para o conjunto de dados com as observações de discrepância alta	31

SUMÁRIO

1	INTRODUÇÃO	11
1.1	TEMA	12
1.2	DELIMITAÇÃO DO TEMA	12
1.3	OBJETIVOS	13
1.3.1	OBJETIVO GERAL.....	13
1.3.2	OBJETIVOS ESPECÍFICOS.....	13
1.4	JUSTIFICATIVA	13
1.5	PROCEDIMENTOS METODOLÓGICOS	14
1.6	ESTRUTURA DO TRABALHO	15
2	REFERENCIAL TEÓRICO	16
2.1	TRATAMENTO DE CLIENTES INADIMPLENTES	16
2.2	ENGENHARIA DA CONFIABILIDADE	19
2.3	ANÁLISE DE DADOS DE VIDA.....	19
2.4	ANÁLISE DO GRÁFICO BOX-PLOT	21
2.5	CONSIDERAÇÕES FINAIS.....	22
3	RESULTADOS E DISCUSSÃO	23
3.1	ANÁLISE EXPLORATÓRIA.....	23
3.2	RESULTADOS DA ANÁLISE DADOS DE VIDA.....	28
4	CONCLUSÃO	36
	REFERÊNCIAS	37

1 INTRODUÇÃO

Muitas empresas de crédito ou prestadoras de serviços, sofrem com seus maus pagadores em todo mundo. No Brasil o número de devedores vem aumentando ano após ano, segundo o SPC (Serviço de proteção ao crédito) esse número chega a 57 milhões de inadimplentes. NÚMERO de devedores no Brasil é maior (R7,2015).

Uma das estratégias das empresas para diminuir seu número de inadimplentes é utilizar os serviços de “*Call Center*” para realizar a negociação de pagamento. Antigamente esse tipo de serviço era “agressivo” e “truculento”, mas com a evolução do código do consumidor em 1990, as empresas no Brasil tiveram que se adaptar e aplicar novas formas de abordagem ao cliente (CÓDIGO DE DEFESA DO CONSUMIDOR).

Com aumento de inadimplentes e a diminuição da recuperação de crédito dos mesmos, como demonstra o Gráfico1, cresceu o número de empresas especializadas no serviço de cobrança em conjunto com as diferentes formas de abordagem do cliente, chamadas de campanhas.



Gráfico 1 - Recuperação de crédito

Fonte: BOA VISTA SCPC (2015)

1.1 TEMA

Na análise de dados de vida (ADV) ou análise de sobrevivência, aspecto particular da análise de confiabilidade, tudo é baseado em estimativas; o valor real da confiabilidade de um produto nunca será conhecido, somente se todos os produtos já tiverem falhado (COLOSIMO; GIOLO, 2006). Os modelos que fornecem a estimativa de taxa de falha do produto em função do tempo estimam a probabilidade de falha (sobreviver) do produto para uma dada idade, ou para um dado período de tempo. Estes modelos são representações matemáticas dos dados que podem ser tanto funções contínuas como discretas as quais permitem interpolações e algumas extrapolações. Estes modelos como são baseados em distribuições estatísticas são chamados de modelos probabilísticos.

1.2 Delimitação do Tema

A análise de dados de vida é utilizada em vários setores como da produção, manutenção, planejamento, prestação de serviços e etc.

O presente trabalho aplica a ADV em ciclos de uma campanha de cobrança, portanto na área de prestação de serviços.

Estes ciclos representam os números de ligações realizadas para cada cliente até efetuar o debito da dívida.

Como dados de desempenho da central de cobrança das instituições financeiras são sigilosos e de difícil acesso, vamos utilizar dados de uma empresa financeira portuguesa, que os disponibiliza gratuitamente.

O banco de dados possui informações do cliente (idade, profissão, escolaridade, quantidade de ligação efetivas e realização do pagamento).

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Determinar a confiabilidade de uma campanha de “*telemarking*” de cobrança de um banco português, analisando os ciclos de ligações efetuadas para o cliente até efetuar o pagamento.

1.3.2 Objetivos Específicos

- Obter os dados da campanha de “*telemarking*” de cobrança de um banco português.
- Organizar dos dados.
- Aplicar análise exploratória.
- Aplicar os dados no software de análise.
- Analisar os resultados.

1.4 JUSTIFICATIVA

Baseado em conversas com profissionais do ramo, verificar a qualidade da estratégia de uma campanha empregada por uma empresa especializada em serviços de “*Call Center*”, por meio de métodos de análise da confiabilidade não é convencional.

Dado que além do resultado financeiro esperado pode-se avaliar a eficiência do serviço usando estimativas como o tempo médio entre ligações, taxa de falha por ligações, como também as métricas de confiabilidade e não confiabilidade do serviço prestado. Com este conjunto de métricas pode-se oferecer ao tomador de decisões as informações quantitativas importantes para mensurar a qualidade do serviço prestado.

Portanto, implementar novos métodos para tomada de decisões focados em informações quantitativas e qualitativas oriundas da análise da confiabilidade, nesse tipo de serviço, poderão ser bem recebidos pelas empresas que tem como premissa a melhoria constante, com objetivo maior de oferecer maior qualidade no serviço que presta.

1.5 PROCEDIMENTOS METODOLÓGICOS

Os dados foram obtidos de um banco de dados livres, disponível no site: <https://archive.ics.uci.edu/ml/datasets.html>.

Utilizando o programa Microsoft Office Excel será realizada uma análise descritiva dos dados, para serem selecionadas as variáveis de interesse. Com o programa Reliasoft Weibull++, será realizada a análise de vida dos dados utilizando o método da máxima verossimilhança, para ajuste paramétrico dos dados com o objetivo de encontrar a distribuição estatística conhecida que melhor se ajusta ao conjunto de dados.

De posse desta distribuição pode-se analisar as métricas da confiabilidade: distribuição de probabilidade “pdf”, distribuição acumulada de probabilidade “cdf”, taxa de falha “ $h(t)$ ”, t representa quantidade de ligações e a falha seria o não pagamento e a função da confiabilidade “ $R(t)$ ”.

1.6 ESTRUTURA DO TRABALHO

O capítulo 1 constitui a introdução do trabalho contendo a apresentação da proposta, definição do tema, objetivos e motivação para o desenvolvimento do trabalho.

No capítulo 2 será apresentada a fundamentação teórica necessária para o desenvolvimento do trabalho no que diz respeito ao tratamento de clientes inadimplentes, engenharia da confiabilidade, análise de dados de vida, teste da razão da máxima verossimilhança e análise do gráfico “box-plot”.

No capítulo 3 constitui na apresentação e discussão de resultados da análise exploratória e da análise de vida.

No capítulo 4 serão apresentadas as conclusões e assuntos para trabalhos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os métodos utilizados neste trabalho para analisar a confiabilidade da campanha realizada por uma empresa de Call Center de um banco português e o perfil dos clientes com ela abordados.

2.1 TRATAMENTO DE CLIENTES INADIMPLENTES

Com o aumento dos inadimplentes nos últimos anos, as empresas de Call Center especializadas em cobrança buscam aprimorar suas técnicas e estratégias de abordagem de clientes com a preocupação de não infligir qualquer um dos artigos de defesa do consumidor. Analisar o desempenho dessas estratégias é essencial para a continuidade do trabalho.

De acordo com artigo 42 do CDC (CÓDIGO DE DEFESA DO CONSUMIDOR) que se refere à cobrança de dívidas, e determina:

“Na cobrança de débitos, o consumidor inadimplente não será exposto a ridículo, nem será submetido a qualquer tipo de constrangimento ou ameaça”. O capítulo de Infrações Penais no Código, no artigo 71, prevê pena de detenção de três meses a um ano, além de multa, para casos em que se exerça “ameaça, coação, constrangimento físico ou moral, afirmações falsas, incorretas, enganosas ou de qualquer outro procedimento que exponha o consumidor, injustificadamente, a ridículo, ou interfira no seu trabalho, descanso ou lazer”.

No Brasil, a inclusão do inadimplente em Cadastros de Restrição ao Crédito – como o SPC e o SERASA e o Protesto de Títulos – protegem tanto o credor quanto o consumidor: ao restringir ao comprador a possibilidade de nova concessão de crédito, o credor terá, cedo ou tarde, grande possibilidade de receber a dívida. Por outro lado, o consumidor fica impedido de contrair débitos ainda maiores, uma vez que só poderá comprar à vista até quitar as parcelas devidas.

Com essas preocupações as empresas de cobranças criam novas campanhas com novas abordagens ao cliente, chamadas de “*Scripts*”, figura 1.

Nesse trabalho iremos aplicar a metodologia de confiabilidade, para analisar os dados de uma determinada campanha em um determinado período de tempo.

SCRIPT PARA COBRANÇA

Apresentação

Bom dia, sou XXXXXXX da INFORM SYSTEM – SITES, SOLUÇÕES E PESQUISAS e gostaria de falar com a pessoa responsável a pagamento de contas !!

Bom dia Sr quem fala é da INFORM SYSTEM – Sites, Soluções e Pesquisas, o motivo de minha ligação é uma fatura em atraso com vencimento no valor que ainda consta pendente em nossos registros. Gostaria de saber se já foi efetuado este pagamento ou se ainda encontra-se pendente ???

Então gostaria que o Sr estivesse efetuando o PAGAMENTO DO BOLETO por gentileza, e é muito fácil, eu posso te enviar agora por E-MAIL ou por FAX para pagamento hoje em qualquer BANCO, INTERNET, CASAS LOTERICAS ou CORREIOS.

O Sr consegue efetuar este pagamento hoje Sr ?

Sem Objeções: Sr , conto com a sua colaboração para este pagamento ainda hoje para que eu possa passar um parecer para gerencia. Agradeço o Sr pela atenção e tenha um ótimo dia de trabalho para o Sr e sua família.

Objeções

1 - Objeção – Estou sem dinheiro ou previsão: Entendo sua situação Sr, mas todos temos lutas e desafios todos os dias, não é mesmo ? Uns dias melhores, outros não tão bons. Mas tenha certeza de que as coisas vão melhorar, o comercio já está reagindo. O Sr é uma pessoa trabalhadora, dedicada, comprometida com seu trabalho, e não é uma dificuldade que vai desanimar o Sr não. Se fizemos a nossa parte Deus há de abençoar nosso trabalho. Vamos fazer o seguinte: Vou agendar aqui que o Sr pagará a fatura até (3 dias depois), para que eu possa justificar com minha gerencia, até lá sua situação já está melhor, com certeza. Cabeça erguida Sr, e bola pra frente. Conto com o Sr, e agradeço pela atenção e tenha um ótimo dia de trabalho para o Sr e sua família.

2 - Objeção – Quero cancelamento: Entendi sua dúvida Sr, eu vou lhe passar o Telefone do Departamento que vai lhe ajudar a sanar esta dúvida, você pode anotar por favor (41) 3207-1700 somente neste numero se trata deste assunto em questão. O meu departamento não trata sobre este assunto, somente tratamos sobre por ventura alguma fatura esteja em atraso. O meu Departamento de Cobranças precisa resolver sobre esta pendência para que eu possa justificar com minha gerencia. Posso contar com a sua colaboração para este pagamento ainda hoje Sr ? Agradeço o Sr pela atenção e tenha um ótimo dia de trabalho para o Sr e sua família.

3 - Objeção - Divergência de valores: Entendi sua dúvida Sr, eu vou acessar neste momento a Tabela de Preços que foi assinada pelo Sr no dia da Afiliação, um momento. Eu verifiquei aqui na tabela assinada pelo Sr e vejo que o valor da fatura está correto, foi isso mesmo, vou lhe enviar agora por e-mail a tabela assinada pelo Sr, pode me confirmar seu e-mail ? O meu Departamento de Cobranças precisa resolver sobre o pagamento desta fatura pendente para que eu possa justificar com minha gerencia. Posso contar com a sua colaboração para este pagamento ainda hoje Sr ? Agradeço o Sr pela atenção e tenha um ótimo dia de trabalho pra você e sua família.

4 - Objeção - Falta de Treinamento ou Suporte: Entendi sua dúvida Sr, eu vou acessar neste momento a Declaração de Treinamento que foi assinada pelo Sr no dia da Afiliação, me dê um momento. Eu verifiquei aqui e vejo que a Declaração de Treinamento está assinada e o Sr marcou com o (X) a opção MUITO BOM, foi isso mesmo, vou lhe enviar agora por e-mail o documento. Independente disso, vou pedir por gentileza para anotar o Telefone do Departamento de Suporte que vai lhe ajudar a sanar esta dúvida e resolver o seu problema, você pode anotar por favor (41) 3207-1700. Somente neste numero se trata deste assunto em questão e vão tirar todas as suas dúvidas. O meu Departamento de Cobranças precisa resolver sobre o pagamento esta fatura pendente para que eu possa justificar com minha gerencia. Posso contar com a sua colaboração para este pagamento ainda hoje Sr ? Agradeço o Sr pela atenção e tenha um ótimo dia de trabalho pra você e sua família.

5 - Objeção - Já cancelei o contrato: Entendi sua dúvida Sr, eu vou acessar neste momento o seu cadastro, me dê um momento. Eu verifiquei aqui e vejo que seu contrato está normal, tudo normal a documentação e não existe nenhum registro ou documento que seu contrato esta cancelado, vou lhe enviar agora por e-mail a documentação que o contrato está ativo conosco. Independente disso, vou pedir por gentileza para anotar o Telefone do Departamento de Suporte que vai lhe ajudar a sanar esta dúvida e resolver o seu problema, você pode anotar por favor (41) 3207-1700. Somente neste numero se trata deste assunto em questão e vão tirar todas as suas dúvidas. E precisamos resolver sobre o pagamento esta fatura pendente para que eu possa justificar com minha gerencia. Posso contar com a sua colaboração para este pagamento ainda hoje Sr ? Agradeço o Sr pela atenção e tenha um ótimo dia de trabalho pra você e sua família.

6 - Objeção – O vendedor mentiu pra mim: Entendi sua dúvida Sr, eu vou acessar neste momento o seu cadastro, me dê um momento. Eu verifiquei aqui e vejo que no dia da afiliação foi realizado a Conferencia Contratual Oral com a pessoa de todos os pontos do contrato, inclusive temos a ligação gravada dando OK em todos os pontos da conferencia da venda do vendedor. Temos também a Declaração de Treinamento de todas as nossas soluções, e o Sr marcou com o (X) a opção MUITO BOM vou lhe enviar agora por e-mail. Também tenho a Tabela de Preços, onde também foi assinada pelo Sr a ciência de todos os valores dos serviços, e também temos o Contrato com todas as clausulas também assinado tudo certinho. Independente disso, vou pedir por gentileza para anotar o Telefone do Departamento de Suporte que vai lhe ajudar a sanar esta dúvida e resolver o seu problema, você pode anotar por favor (41) 3207-1700. Somente neste numero se trata deste assunto em questão e vão tirar todas as suas dúvidas. O meu Departamento de Cobranças precisa resolver sobre o pagamento desta fatura pendente para que eu possa justificar com minha gerencia. Posso contar com a sua colaboração para este pagamento ainda hoje Sr ? Agradeço o Sr pela atenção e tenha um ótimo dia de trabalho pra você e sua família.

Figura 1 - Exemplo de "script" de Cobrança

Fonte: INFORMSYSTEM (2015)

2.2 ENGENHARIA DA CONFIABILIDADE

A Engenharia da Confiabilidade fornece as ferramentas teóricas e práticas que permitem especificar, projetar, testar e demonstrar a probabilidade e a capacidade segundo a qual componentes, equipamentos, produtos e sistemas desempenharão suas funções, por períodos determinados de tempo, em ambientes e condições específicas e sem apresentar falhas. (LEWIS, Elmer E.,1987)

2.3 ANÁLISE DE DADOS DE VIDA

A análise de Dados de Vida utiliza de teorias estatísticas para construir modelos probabilísticos paramétricos e não paramétricos, a partir de dados de falhas. Os modelos fornecem a estimativa de taxa de falha do produto em função do tempo, eles estimam a probabilidade de falha do produto para uma dada idade, ou para um dado intervalo de tempo.

A função taxa de falha também indica qual é a tendência de falha do ativo (item, sistema, produto...) que poder decrescente (período de mortalidade infantil), constante (operação) ou crescente (envelhecimento) (FOGLIATTO, F. S.,2011).

Outro aspecto a respeito dos modelos probabilísticos é que podem ser discretos ou contínuos. Na figura 2 apresentam-se os principais modelos contínuos utilizados na análise da confiabilidade.

O trabalho consiste em, a partir do conjunto de dados ou da amostra estimar os parâmetros das distribuições em questão que em estatística podem ser a média, mediana, moda e erro padrão da população desconhecida. Porém, do ponto de vista da análise da confiabilidade, estes parâmetros ganham outras denominações, por exemplo, o fator de forma ("*Beta*"), fator de escala ("*Eta*") e fator de posição ("*Gama*") (RELIAWIKI,2015). Para estimar esses parâmetros, podemos utilizar o método de estimação pontual, que consiste no cálculo, a partir de uma amostra observada, de valores plausíveis para os parâmetros da distribuição da variável estatística de interesse numa população. E para avaliar essa estimativa, utilizamos intervalos de valores

quem com maior “confiança”, contenha o valor desconhecido do parâmetro de uma população.

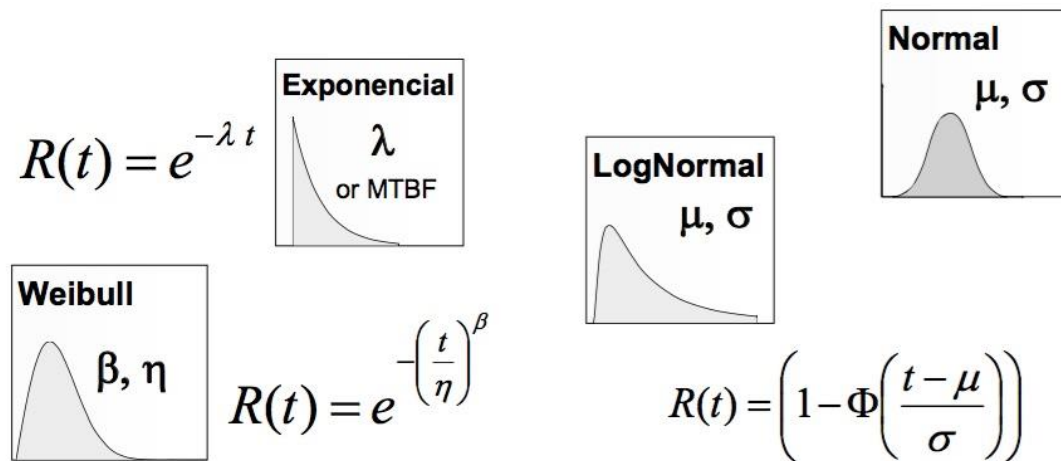


Figura 2 - Exemplos de distribuições contínuas

Fonte: COLOSIMO; GIOLO (2006)

Com a finalidade de decidir entre os modelos probabilísticos, qual é o mais adequado para o conjunto de dados, é utilizado à estatística do teste da razão de verossimilhança (melhor aderência entre os modelos).

Esse teste é apropriado para testar modelos que sejam encaixados, ou seja, em que um dos modelos é submodelo do outro, sendo expresso pela equação (MCCULLAGH; NELDER, 1989):

$$D = -2 \cdot \left(\log(p(y|\hat{\theta})) - \log(p(y|\theta_0)) \right)$$

Sendo que $p(y|\hat{\theta})$ é o máximo do logaritmo natural da função de verossimilhança (MLFV) e $p(y|\theta_0)$ é o MLFV para modelo mais simples sob H_0 . Se o modelo melhor se ajustar ao conjunto de dados, a D (deviance) tem-se distribuição assintótica χ^2_v , onde v é os graus de liberdade. Portanto, a hipótese que o modelo nulo apresenta a log-verossimilhança menor é rejeitada caso, $D > \chi^2_v$.

2.4 ANÁLISE DO GRÁFICO BOX-PLOT

O gráfico de “*Box-Plot*” ou diagrama de caixa é utilizado para identificar pontos fora do padrão dentro do conjunto de dados, chamados de “*outliers*”, acontecimentos fora do comum ocorridos durante a execução da campanha. (MAGALHÃES, M. N.; 2008).

A Representação gráfica através do *Box-Plot* é bastante rica no sentido de informar, entre outras coisas, a variabilidade e simetria dos grupos de clientes pagantes e não pagantes. Também são úteis para detectar, descritivamente, diferenças nos comportamentos dos mesmos. O gráfico “*Box-Plot*” (figura 3), é construído da seguinte forma:

- Calcula-se a mediana, o quartil inferior (Q1) e o quartil superior (Q3);
- Subtrai-se o quartil superior do quartil inferior = (L)
- Os valores que estiverem no intervalo de $Q3+1,5L$ e $Q3+3L$ e no intervalo $Q1-1,5L$ e $Q1-3L$, serão considerados “*outliers*” podendo, portanto, ser aceitos na população com alguma suspeita;
- Os valores que forem maiores que $Q3+3L$ e menores que $Q1-3L$ devem ser considerados suspeitos de pertencer à população, devendo ser investigada a origem da dispersão. Estes pontos são chamados de extremos.

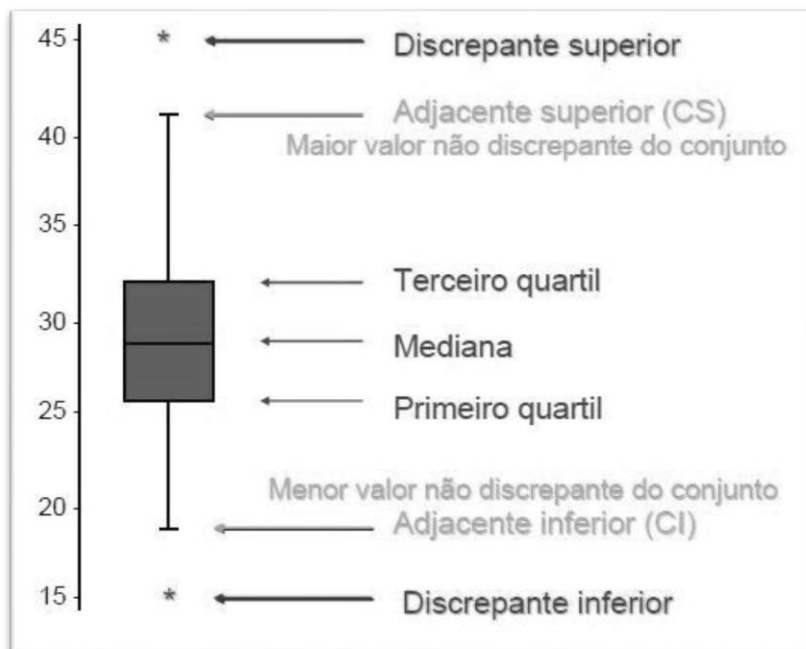


Figura 3 - Definições do gráfico Box-Plot

Fonte: EBAH (2016)

2.5 CONSIDERAÇÕES FINAIS

Nesse capítulo foi apresentado o conceito de tratamento de clientes inadimplentes, confiabilidade e métodos de análise de dados de vida, que auxiliarão para avaliar a qualidade e a confiabilidade da estratégia aplicada pela empresa de cobrança de um banco português.

3 RESULTADOS E DISCUSSÃO

3.1 ANÁLISE EXPLORATÓRIA

O banco de dados inicialmente, era constituído por 16 variáveis distintas e uma variável resposta dicotômica (realização do pagamento ou não).

Para realizar análise de dados de vida, precisa-se de pelo menos uma variável explicativa quantitativa (continua ou não), e a única variável que se encaixa nesse requisito, no caso, seria a “*Campaign*” ou Campanha (Quantidades de ligações efetivas realizadas durante a campanha). Os clientes foram agrupados de acordo com o número de ligações realizadas.

Foram retirados do estudo, os clientes que a empresa já tinha feito algum tipo de contato antes da campanha em estudo, foram identificados com valores diferentes de “-1” na variável “pdays” (Quantidade de contatos anterior a campanha) com objetivo de realizar os estudos somente nos clientes que receberão o tratamento da campanha estudada, sem influência das anteriores.

Objetivo dos tratamentos feitos no banco dados é retirar qualquer influência externa, para podermos avaliar somente a eficácia da determinada estratégia tomada pela empresa de cobrança.

No gráfico 2, pode-se observar o comportamento da variável “*Campaign*” relacionada com a variável resposta.

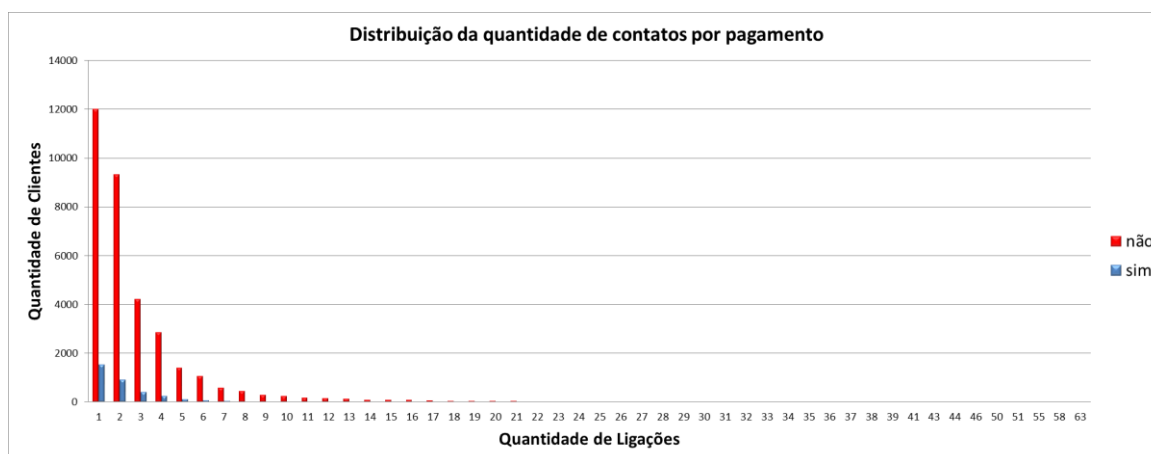


Gráfico 2 - Comportamento da variável “*Campaign*”

Fonte: Autoria Própria

Foi verificada a relação das duas variáveis, “*Campaign*” e a resposta, com as variáveis categóricas: Escolaridade e Profissão, com a intenção de categorizar a variável “*Campaign*”.

Com a categorização podemos verificar a existência de diferentes padrões de comportamento entre as categorias, assim podendo observar a influência das categorias no sucesso da campanha (pagamento da dívida).

A categorização realizada não teve resultados significativos, isso significa que o comportamento dos clientes é o mesmo independente da profissão ou escolaridade, demonstrado nos gráficos 3 e 4.

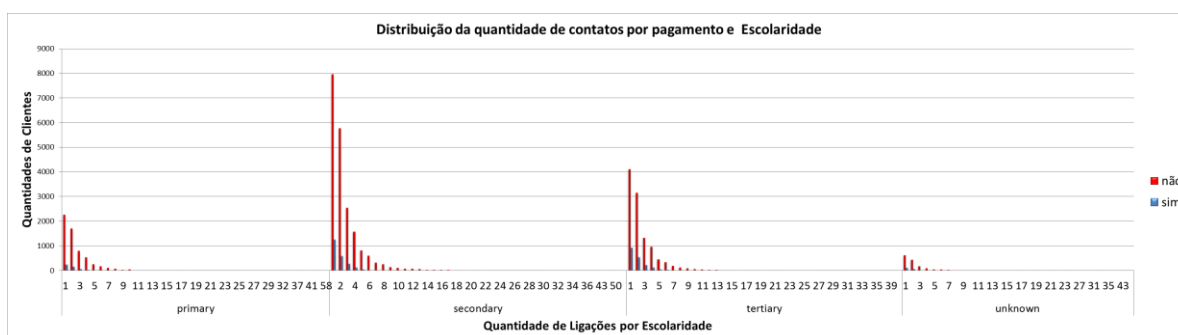


Gráfico 3 - Comportamento da variável "Campaign" por escolaridade

Fonte: Autoria própria

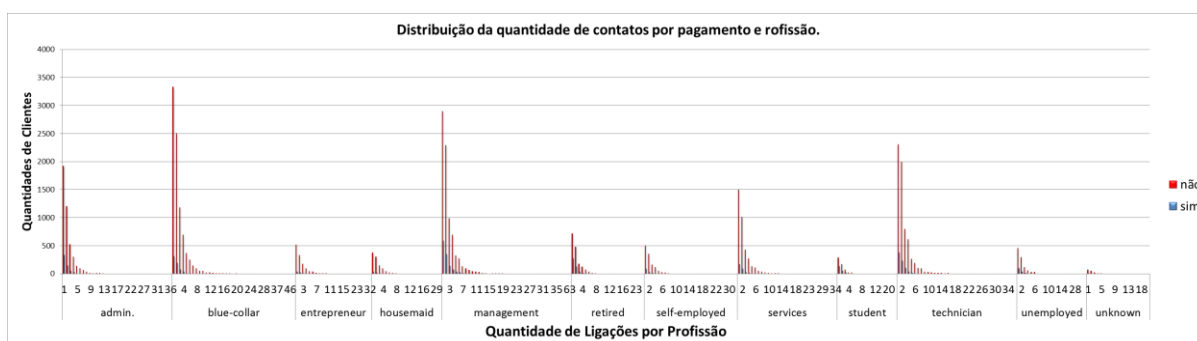


Gráfico 4 - Comportamento da variável "Campaign" por profissão

Fonte: Autoria própria

Analisando o gráfico de “*box-plot*” da “*campanha*” por realização de pagamento, observou-se uma grande quantidade de dados discrepantes “*outliers*”, apresentados no gráfico da esquerda na figura 4. No gráfico da direita na figura 4, apresenta a igualdade das distribuições entre os clientes que realizarão ou não o pagamento de acordo com a quantidade de ligações recebidas.

Sabendo que as duas respostas, sim e não, possuem a mesma distribuição, podemos aplicar o mesmo tratamento, assim facilitando a análise dos resultados.

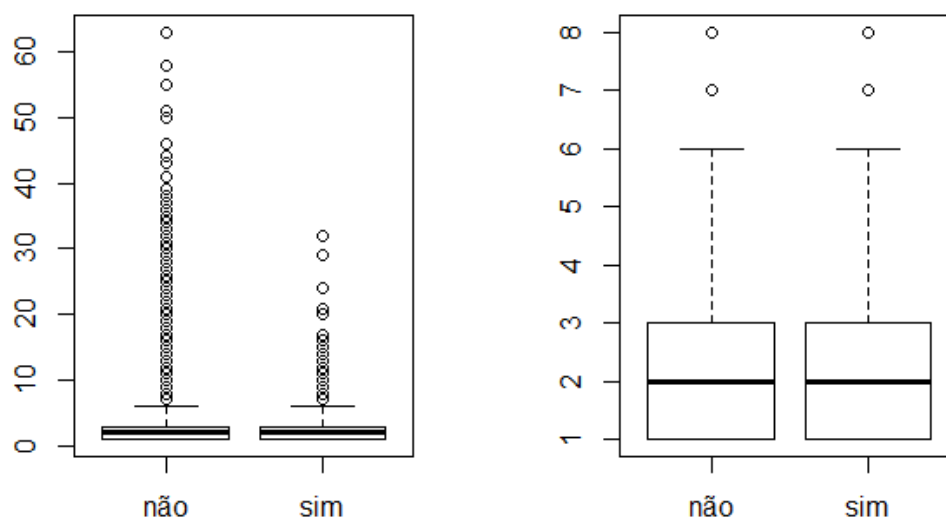


Figura 4 - Gráficos "box-plot" de quantidade de ligações por pagamento

Fonte: Autoria própria

Com a finalidade em analisar a influência dos dados discrepantes nos resultados de confiabilidade, a análise será focada em 3 conjuntos de dados: com os dados discrepantes (Gráfico 5), não contendo os dados discrepantes (Gráfico 6) e não contendo os dados com alta discrepância (Gráfico 7).

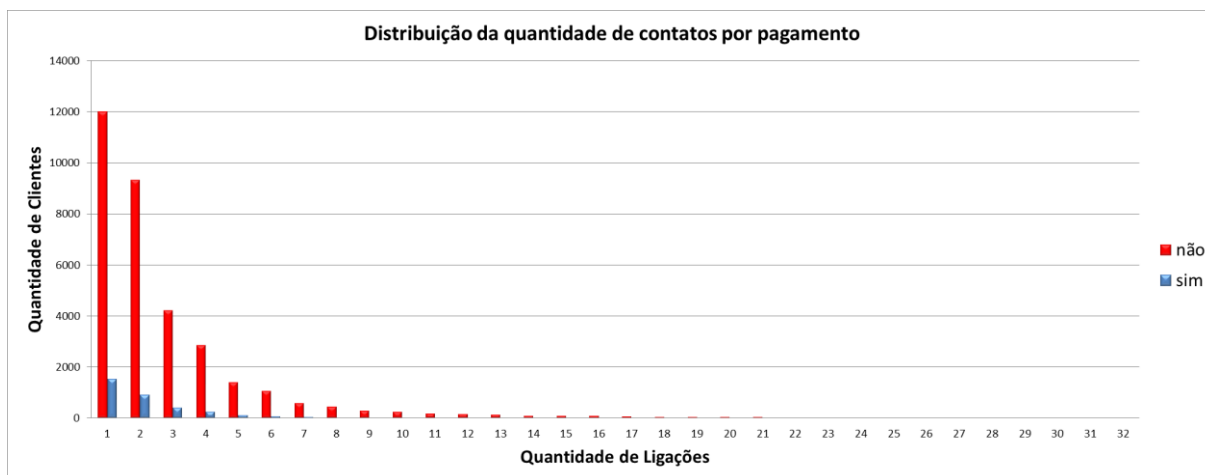


Gráfico 5 - Conjunto contendo dados discrepantes

Fonte: Autoria própria

No Gráfico 5 representa a distribuição dos dados contendo todas as observações, percebemos que a partir da observação de número 8 (8 ligações realizada ao mesmo cliente), a quantidade das demais observações começam a ficar muito pequena, demonstrando as características de observações discrepantes “*outliers*”.

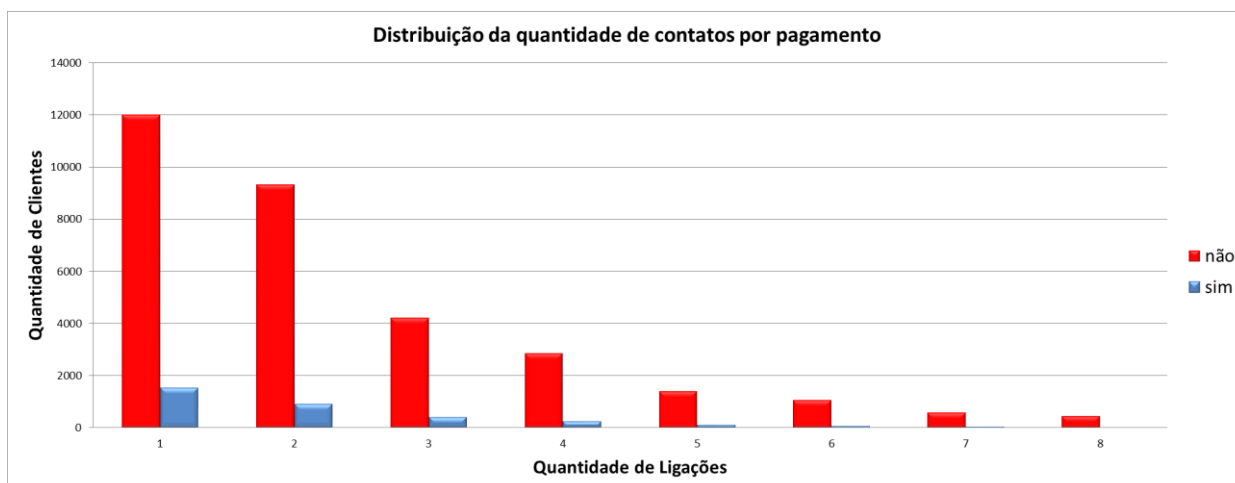


Gráfico 6 - Conjunto não contendo dados discrepantes

Fonte: Autoria própria

No Gráfico 6 demonstra a distribuição dos dados sem os “*outliers*”, dados demonstrados no gráfico “*box-plot*” da esquerda na Figura 4. Podemos observar a igualdade da distribuição das duas respostas, sim e não.

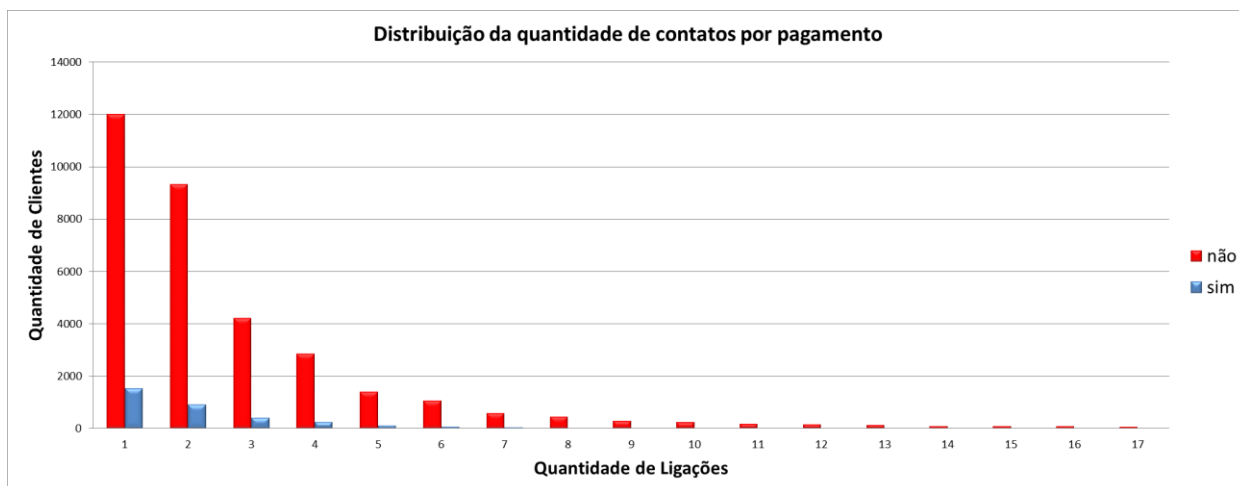


Gráfico 7 - Conjunto não contendo dados com discrepância alta

Fonte: Autoria própria

No Gráfico 7 representa os dados sem as informações de discrepância alta (clientes que tiveram muitas ligações), para podermos estudar suas influências nos resultados finais.

3.2 RESULTADOS DA ANÁLISE DADOS DE VIDA

Utilizando o programa Reliasoft Weibull++, configurado para receber dados livres, pois as ligações são independentes, assim não sofrendo distorção do resultado final. Foram analisados os conjuntos de dados, lembrando que quantidade de ligações serão nomeados de “ciclos” nesse programa, a Tabela 1 demonstra como os conjuntos de dados foram tabulados no programa.

Tabela 1 - Exemplo de tabulação dos dados no programa Weibull ++

	Eixo-X Valor (Cic)	Eixo-Y Valor, (%)	Subconjunto ID 1
1	1	88,74	grupo1
2	2	91,24	grupo2
3	3	91,34	grupo3
4	4	92,12	grupo4
5	5	93,17	grupo5
6	6	93,71	grupo6
7	7	93,57	grupo7
8	8	94,24	grupo8

Fonte: Autoria própria.

A tabulação de dados livres, facilita o manuseio e desempenho do programa Reliasoft Weibull++, pois as 425212 observações distintas foram agrupadas e transformadas em 48 independentes.

Sendo CIC a quantidade de ligações feitas aos clientes, Eixo-Y é a porcentagem de não pagamento para determinado CIC e o campo Subconjunto representa um rotulo para cada CIC.

Utilizando a ferramenta Teste de Aderência do programa Weibull++, considerando o teste da razão da máxima verossimilhança, observou-se, que as distribuições que convergiram aos dados e a classificação pelo valor do teste, como demonstra as figuras 5, 6 e 7.

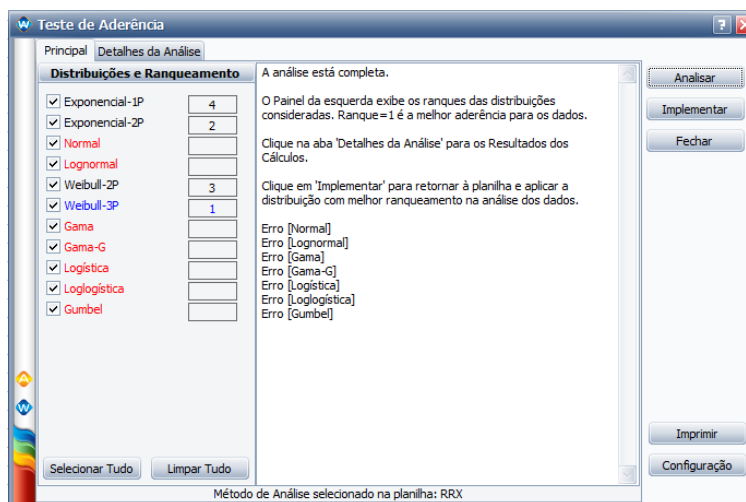


Figura 5 - Teste de aderência para o conjunto de dados sem as observações discrepantes

Fonte: Autoria própria.

Podemos observar que para o conjunto de dados sem as observações discrepantes, somente algumas distribuições convergiram: Weibull 3 e 2 parâmetros, exponencial 2 e 3 parâmetros. Pelo valor do teste da máxima verossimilhança, a distribuição Weibull com três parâmetros que obteve o melhor ajuste.

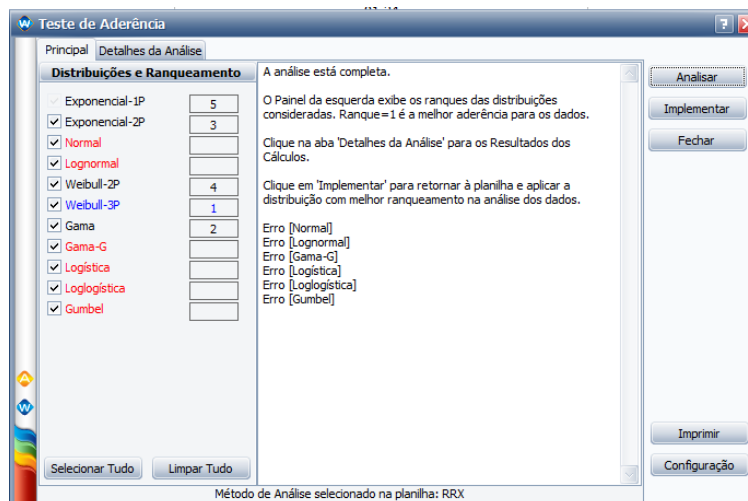


Figura 6 - Teste de aderência para o conjunto de dados sem as observações com discrepância alta.

Fonte: Autoria própria.

Ao analisar o quadro teste de aderência dos dados sem as observações com discrepância alta, somente algumas distribuições convergiram: Weibull 3 e

2 parâmetros, exponencial 2 e 3 parâmetros e a Gama. Pelo valor do teste da máxima verossimilhança, a distribuição Weibull com três parâmetros novamente que obteve o melhor ajuste.

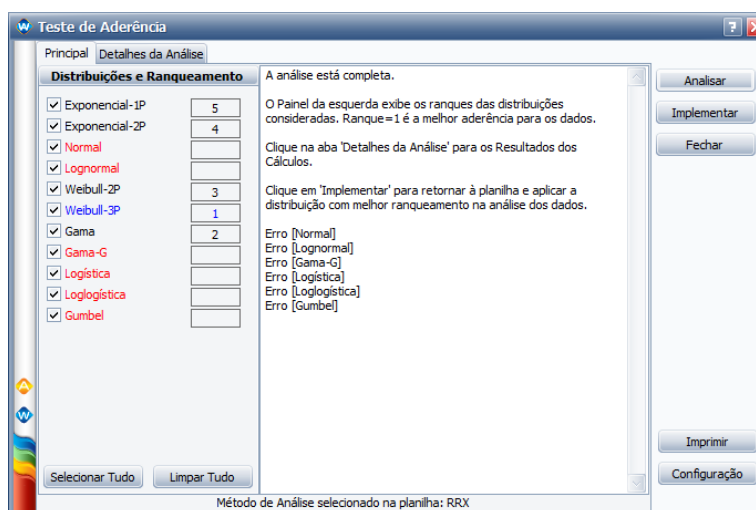


Figura 7 - Teste de aderência para o conjunto de dados com as observações com discrepância alta.

Fonte: Autoria própria.

Comparando os três testes de aderência (Figura 5, 6 e 7), identificamos a distribuição Weibull com 3 parâmetros pelo teste da máxima verossimilhança, é a mais adequada para os 3 conjuntos de dados. Com esse fato, facilitando a comparação entre eles, assim podendo revelar a influência dos “outliers” nos resultados.

Os parâmetros das distribuições propostas para oss conjuntos de dados, estão representadas pelas Tabelas 2, 3 e 4.

Tabela 2 - Parâmetros da distribuição Weibull para o conjunto de dados sem as observações discrepantes

Início Weibull-3P
Beta=0,121095940683264
Eta=0,00151245894244033
Gama=0,0425
Terminado Weibull-3P

Fonte: Autoria própria.

Tabela 3 - Parâmetros da distribuição Weibull para o conjunto de dados sem as observações de discrepância alta

Início Weibull-3P
Beta=0,122380863466314
Eta=0,00163990714148553
Gama=0,03
Terminado Weibull-3P

Fonte: Autoria própria.

Ao analisar as tabelas 2 e 3, observamos que os dados sem discrepância muito alta, "outliers", as duas distribuições possuem os mesmos parâmetros (com uma diferença irrelevante), assim contendo a mesma forma de curva de probabilidade. Com essa igualdade entre as distribuições demonstra que os dados com discrepância moderada não influenciam nos resultados finais.

Tabela 4 - Parâmetros da distribuição Weibull para o conjunto de dados com as observações de discrepância alta

Início Weibull-3P
Beta=0,0679430355124516
Eta=1,71244098218657E-06
Gama=0,8299999999999999
Terminado Weibull-3P

Fonte: Autoria própria.

Na Tabela 4 verificamos a influência dos dados com alta discrepância, “*outliers*”, obtendo parâmetros com diferenças relevantes, comparados com os parâmetros das duas distribuições apresentadas nas tabelas 2 e 3. Com essas diferenças, a curva de probabilidade obtém mudanças e alterando as probabilidades.

Nos Gráficos 8, 9 e 10, são apresentadas as curvas de probabilidade de falha desejada com a curva Weibull desejada.

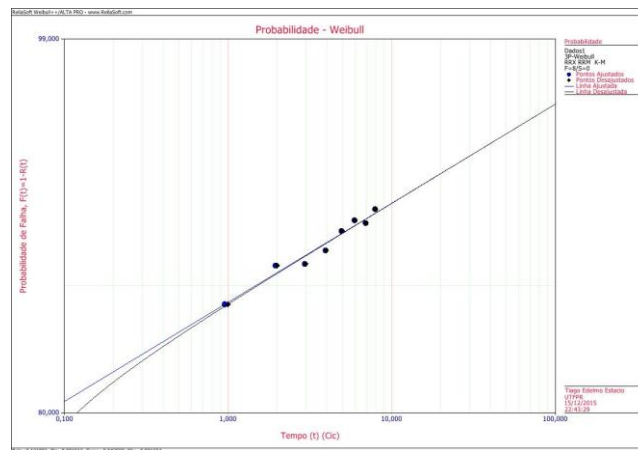


Gráfico 8 - Gráfico de dispersão do conjunto de dados sem as observações discrepantes

Fonte: Autoria própria.

No gráfico 8 a linha preta representa a distribuição Weibull proposta anteriormente para os dados sem observações discrepantes e os pontos azuis são os dados realizados. Podemos observar que a distribuição proposta está bem próxima de todos os pontos azuis, demonstrando a adequação da distribuição para os dados.

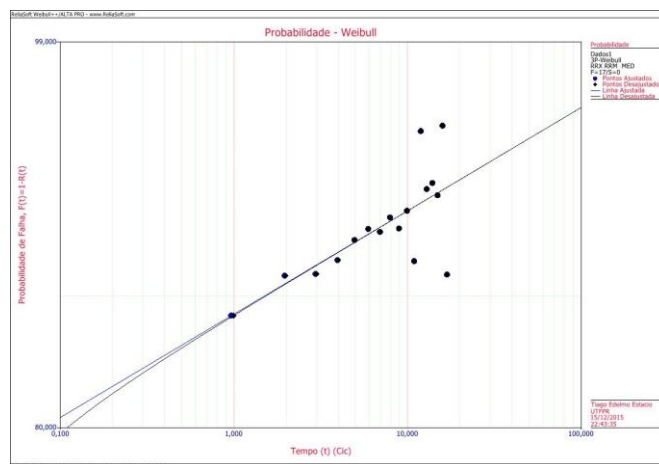


Gráfico 9 - Gráfico de dispersão do conjunto de dados sem as observações de alta discrepância

Fonte: Autoria própria.

No gráfico 9, utilizando da mesma interpretação do Gráfico 8, podemos observar que a distribuição Weibull proposta está bem próxima de todos os pontos azuis, demonstrando a não influência das observações de relevância moderada.

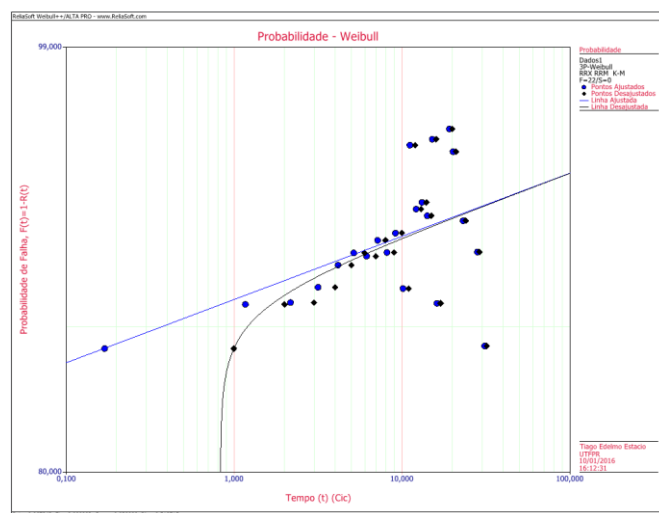


Gráfico 10 - Gráfico de dispersão do conjunto de dados com as observações de alta discrepância

Fonte: Autoria própria.

A distribuição Weibull proposta para o conjunto de dados com todas as informações, demonstrada pelo Gráfico 10, observamos a distorção nos primeiros pontos, apresentando a influência das observações com discrepância

alta, pois a curva de probabilidade muda drasticamente para poder se alinhar com os “outliers”, prejudicando as interpretações corretas dos dados, influenciando na predição de eventos ocorridos e previsão de eventos futuros.

Analisando os gráficos de confiabilidade de acordo com os ciclos (Gráficos 11,12 e 13), observou-se que a confiabilidade da campanha diminui à medida que o número de contatos com o cliente aumenta.

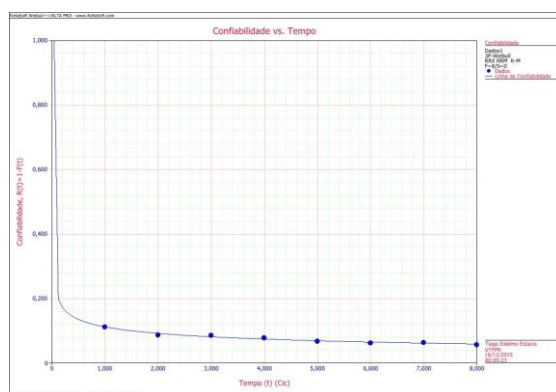


Gráfico 11 - Confiabilidade do conjunto de dados sem as observações discrepantes

Fonte: Autoria própria.

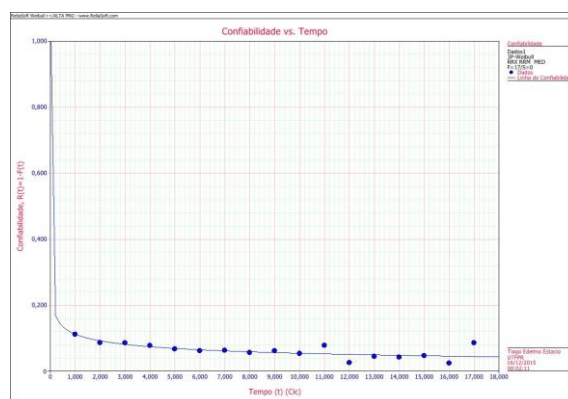


Gráfico 12 - Confiabilidade do conjunto de dados sem as observações de alta discrepância

Fonte: Autoria própria.

Nos Gráficos 11 e 12 a semelhança entre as duas curvas de confiabilidade fica clara, ressaltando a igualdade dos parâmetros das

distribuições Weibull proposta para os dados sem e com as observações de discrepância moderada, comprovando a não influência das informações das mesmas.

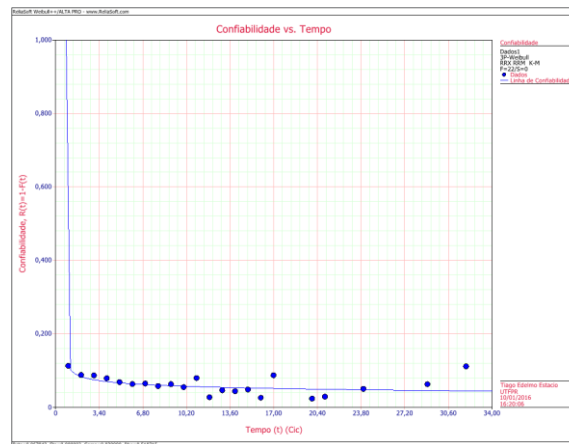


Gráfico 13- Confiabilidade do conjunto de dados com as observações de alta discrepância

Fonte: Autoria própria.

A curva de confiabilidade para os dados completos, representada pelo Gráfico 13, tem um formato diferente das curvas apresentadas anteriormente, demonstrando a influência das observações com elevada discrepância. Com essa influência, as informações de confiabilidade da campanha podem ser equivocadas.

4 CONCLUSÃO

As análises realizadas nesta monografia tiveram como objetivo analisar a confiabilidade da campanha de cobrança de um Banco Português.

O resultado final das análises mostrou uma confiabilidade da campanha é muito baixa, abaixo de 20% para uma ligação com o cliente e diminuindo com o aumento do número de tentativas (ligações), mas não é possível concluir se o resultado foi ruim ou bom. Para se obter esse diagnóstico deve-se comparar os valores da confiabilidade com outras campanhas.

Com a análise exploratória utilizando o gráfico “*box-plot*” e com os resultados dos dados de vida, percebeu-se que a retirada os dados com alta discrepância, provocou alterações no modelo probabilístico e seus parâmetros, logo a presença destas discrepâncias influencia muito os resultados e podem prejudicar a análise.

Observou-se também que retirando esses dados com uma discrepância elevada, com representatividade no conjunto inteiro baixa, sendo 0,3% do total, os conjuntos com ou sem os dados de discrepância moderada, obtiveram os mesmos resultados. Demonstrando a Influência dos dados discrepantes “*outliers*”, resultados e nas suas interpretações.

Uma das vantagens de retirar os dados discrepantes do seu modelo, é a diminuição do tamanho do seu banco de dados, assim melhorando a velocidade de processamento dos programas utilizados para realizar as análises e facilitando o manuseio do banco de dados e impedindo falsas interpretações.

Sugere-se para trabalhos futuros, utilizar as técnicas apresentadas nessa monografia, em um banco de dados que contenha o valor devedor do cliente e o custo operacional da campanha, para se realizar uma análise de risco para verificar a eficácia da campanha em termos monetários.

REFERÊNCIAS

BOA VISTA SCPC. **Recuperação de crédito cai 1,5% no acumulado em 12 meses, constata Boa Vista SCPC**, Disponível em:<<http://www.boavistaservicos.com.br/tag/inadimplencia/>>. Acesso em:01 out 2015.

CHARNET,R.; **Análise de Modelos de regressão Linear: com aplicações**. São Paulo: Unicamp,1999

COLOSIMO, E. A.; GIOLO, S. R. **Análise de Sobrevivência Aplicada**. São Paulo: Edgard Blucher, 2006.

CÓDIGO DE DEFESA DO CONSUMIDOR. Lei 8.078 de 11/09/90. Brasília, Diário Oficial da União, 1990.

FOGLIATTO, F. S. **Confiabilidade e manutenção industrial**. Rio de Janeiro: Elsevier: ABEPRO.2011.

EBAH. **Box plot e diagrama ramos e folhas** <<http://www.ebah.com.br/content/ABAAABPqEAI/box-plot-diagrama-ramos-folhas>>. Acesso em: 02 Fev 2016.

INFORMSYSTEM. **SCRIPT PARA COBRANÇA**, Disponível em:<<http://www.informsystem.com.br/franquias/php/clientes/documentos/SCRIPT-PARA-COBRANCA.pdf>>. Acesso em:05 out 2015.

LEWIS, Elmer E. **Introduction to Reliability Engineering**. New York: John Wiley, 1987

MAGALHÃES, M. N. **Noções de Probabilidade e Estatística/ Marcos Nascimento Magalhães, Antonio Calos Pedroso de lima – 6 ed. rev., 3ªreimpr.**-São Paulo: Editora da Universidade de São Paulo, 2008.

MCCULLAGH, P; NELDER, J. **Generalized Linear Models, Second Edition**. Chapman & Hall/CRC, New York,1989

RELIAWIKI, **The Weibull Distribution**, Disponível em:<http://reliawiki.org/index.php/The_Weibull_Distribution>. Acesso em: 02 Fev 2016.

R7. **Número de devedores no Brasil é maior do que toda a população do Estado de São Paulo**, Disponível em:<<http://noticias.r7.com/economia/numero-de-devedores-no-brasil-e-maior-do->

que-toda-a-populacao-do-estado-de-sao-paulo-18082015>. Acesso em:01 out 2015.