

**Eixo Temático: Inovação e Sustentabilidade em Diferentes Setores**

**ANÁLISE DE CONCESSÃO DE CRÉDITO À PESSOA FÍSICA EM UMA EMPRESA  
COMERCIAL COM A UTILIZAÇÃO DA TÉCNICA DE ANÁLISE  
DISCRIMINANTE**

**ANALYSIS OF CREDIT RISK TO INDIVIDUALS ENTITIES IN A COMMERCIAL  
COMPANY WITH THE USE OF DISCRIMINANT ANALYSIS TECHNIQUE**

Matheus Cancela Bragança, Jorge Ubirajara da Luz dos Santos e Fabiane Tubino Garcia

**RESUMO**

Este estudo tem por objetivo propor um modelo probabilístico de risco de crédito a pessoas físicas, em uma empresa comercial, com a utilização da técnica estatística denominada Análise Discriminante. Quanto à metodologia esta pesquisa classifica-se em quantitativa, bibliográfica e descritiva, e o procedimento escolhido foi o estudo de caso. Utilizou-se uma amostra com 95 clientes que efetuaram compras a prazo, no período de 1/8/2012 a 31/1/2013. As informações dos clientes foram retiradas do banco de dados cadastrais fornecido pela empresa. Primeiramente, identificou-se o perfil desses clientes por meio da técnica de estatística descritiva, e logo para construir o modelo de risco de crédito, utilizou-se como variável dependente a inadimplência e foram testadas quatro variáveis independentes (sexo, idade, renda e estado civil). Os resultados apurados mostraram que o perfil do cliente pessoa física que efetuou compras a prazo são na maioria do sexo feminino (51,58%), casados (54,74%), com média de idade de 51 anos e renda mensal média de R\$1.561,31. Dentre as variáveis explicativas analisadas, a única estatisticamente significativa, com um nível de significância de 10% foi a variável Renda (p-valor = 0,05).

**Palavras-chave:** Crédito, Pessoa Física, Inadimplência, Análise Discriminante.

**ABSTRACT**

This paper is intended to propose a probability model of credit risk to individual entities, in a commercial company, with the use of the statistical method called Discriminant Analysis. In relation to methodology, this research is classified in quantitative, review of literature and descriptive, and the method chosen for it was the case study. For this research it was used a sample of 95 customers who purchased in installments between 1/8/2012 and 31/1/2013. The customers' information was obtained from the database registration provided by the company. First of all, the profile of such customers were identified through descriptive statistics method, and in order to draw up a model of credit risk, it was used – as a dependent variable – the default; and four independent variables were also tested (gender, age, income and marital status). The findings demonstrated that the profile of the individual entity customer – who purchased in installments – are mostly women (51,58%), married (54,74%), average age 51 and average monthly income R\$1.561,31. Among the explanatory variables analyzed, the only one statistically significant – with a level of significance of 10% - was the Variable Income (p-value = 0,05).

**Keywords:** Credit, Individual Entity, Default, Discriminant Analysis.

## INTRODUÇÃO

No atual cenário de desenvolvimento econômico em que o país se encontra é perceptível o aumento do poder aquisitivo da população em todas as classes sociais e também é cada vez maior a oferta dos mais variados produtos e serviços existentes no mercado. Além disso, o governo vem interferindo na economia, com políticas de cortes de impostos em setores específicos e de redução nas taxas de juros, a fim de movimentá-la ainda mais. Esses fatos somados levaram a um novo patamar de consumo, onde, para satisfazer suas necessidades, as pessoas recorrem ao crédito com mais frequência, seja ele oriundo de instituições financeiras ou diretamente do comércio.

Segundo Daroit (2011) a sociedade em geral procura melhor qualidade de vida, tanto individualmente como em família. Dessa forma, a oferta de crédito à pessoa física está sendo incrementada em todo o país (facilitada por taxas de juros atrativas, prazo estendido, e condições para quitação do débito cada vez mais variadas) buscando proporcionar recursos para que as pessoas consumam mais e satisfaçam suas necessidades.

Verifica-se um estudo recente sobre o perfil do consumidor com e sem dívidas no Brasil, divulgado em outubro de 2012, realizado em parceria pela Confederação Nacional de Dirigentes Lojistas (CNDL) e Universidade Federal de Minas Gerais (UFMG) mostra o percentual de inadimplência<sup>1</sup> da população brasileira em relação às classes sociais. Segundo este estudo a classe C é a mais inadimplente e corresponde a quase metade do total, englobando 47% da população nessa situação. Já a classe A, com apenas 2%, corresponde à menor fatia dos inadimplentes no Brasil.

Dados esses fatos o presente trabalho visa estudar o caso de uma empresa comercial, situada na cidade de Santana do Livramento, proporcionando um modelo probabilístico de previsão de risco de crédito, apresentando a seguinte problemática: quais as variáveis significativas que possibilitam identificar o risco de crédito no segmento de pessoas físicas? Destacando a importância de se ter um modelo de risco de crédito, Barth (2004) explica que através desse recurso é possível classificar em “provável adimplente” ou “provável inadimplente” um novo pretendente a crédito.

A relevância da pesquisa é possibilitar que a empresa comercial possa melhor avaliar o crédito que concede em suas vendas a prazo, no segmento de pessoas físicas, a partir da análise de informações fornecidas pelos mesmos durante o preenchimento de uma ficha cadastral elaborada pela empresa. Com isso, este trabalho proporcionará a empresa um modelo de avaliação de crédito objetivo, deixando para trás o modelo subjetivo praticado atualmente, no qual é feita uma análise unicamente intuitiva, através do *feeling* do analista frente às informações fornecidas pelos clientes.

Outro ponto importante que se deve ressaltar é o baixo número de trabalhos científicos acerca do tema de risco de crédito efetuado em empresas comerciais. Em sua grande maioria, os estudos que abordam o risco de crédito discutem situações de concessão de crédito a pessoas físicas ou jurídicas concedido por bancos ou instituições financeiras. Com isto, verifica-se uma lacuna a ser explorada que trate especificamente de modelos de risco de crédito à pessoa física aplicados em empresas no ramo do comércio.

Dessa forma, esse estudo revela-se pertinente por contribuir com a gestão da empresa no que tange a análise e decisão de concessão do crédito, bem como auxiliar em futuras pesquisas que abordem o tema de risco de crédito na modalidade de pessoa física em empresas comerciais.

## 2 REFERENCIAL TEÓRICO

Nesta etapa do estudo é exposta a teoria na qual será embasada a pesquisa e que auxiliará no desenvolvimento do trabalho abrangendo temas pertinentes como: risco de crédito, análise de crédito no segmento de pessoas físicas, técnicas de análise de crédito e política de crédito.

### 2.1 Risco de Crédito

Em finanças, pode-se definir crédito como forma de financiamento com o objetivo de que transações comerciais entre empresa e cliente possam ocorrer de maneira bem sucedida (SANTOS, 2010). O termo crédito institui, entre as partes de dada operação, uma relação mútua de confiança. Tem sua origem do latim *creditum*, que significa confiança ou segurança na verdade de uma crença, reputação, boa fama (SECURATO, 2007).

Sobre risco de crédito, Oliveira comenta:

Zelar por uma carteira de crédito com baixo índice de inadimplência significa, principalmente, entender os fatores determinantes para a mesma e ainda, compreender atos e sutilezas humanas. Determinar a capacidade de pagamento de uma pessoa é ir muito além do que vender um produto, é acompanhar o histórico, a evolução sempre do tomador (OLIVEIRA, 2010, p. 31-32).

De acordo com Silva (2008) dependendo do contexto em que possa estar inserido, tem-se vários entendimentos para a palavra crédito, mas de maneira específica consiste em entregar um valor presente em troca de uma promessa de pagamento em prazo estipulado.

Santos (2010, p.1) complementa que “uma linha de raciocínio tem predominado entre os autores: crédito refere-se a troca de um valor presente por uma promessa de reembolso futuro, não necessariamente certa, em virtude do fator risco”.

Sobre o risco de crédito Brito e Assaf Neto (2008, p. 19) afirmam que “pode ser entendido como a possibilidade de o credor incorrer em perdas, em razão de as obrigações assumidas pelo tomador não serem liquidadas nas condições pactuadas”.

Conforme Silva (2008) o crédito, em seu sentido restrito, gera risco de que não haja o devido pagamento na data determinada previamente por envolver uma promessa de quitação da dívida, o que não é uma garantia concreta. O mesmo autor faz ainda uma diferenciação entre risco e incerteza, onde expõe que o *risco* existe quando se pode tomar uma decisão baseado em dados históricos e considerando estimativas que se julguem aceitáveis, enquanto *incerteza* ocorre quando não se tem acesso a dados históricos e o tomador de decisões acaba por ter que fazer uma distribuição probabilística objetiva, utilizando-se de sua sensibilidade.

Em relação à concessão de crédito, mesmo que se tenha realizado uma ótima avaliação prévia, não é possível prever com total certeza se o resultado da operação será o esperado e haverá o recebimento do valor acordado no prazo estabelecido. Essa incerteza quanto ao desfecho do processo é que origina o risco na operação de crédito (SECURATO, 2007). O mesmo autor ainda completa que a partir disso, o termo risco de crédito passa a ser utilizado muitas vezes indicando que o final da operação não será o esperado e, conseqüentemente, não haverá o recebimento do resultado da operação de crédito.

Segundo Silva (2008) no comércio e na indústria, o crédito assume o papel de facilitador da venda, possibilitando ao cliente a aquisição de um bem para atender suas necessidades, ao mesmo tempo em que incrementa as vendas do comerciante.

O risco de crédito é dividido em três partes: risco de inadimplência (*default*), risco de exposição e risco de recuperação. O risco de inadimplência refere-se à possibilidade de que aconteça um evento de *default* com o tomador; o risco de exposição advém da falta de certeza

em relação ao valor futuro da operação de crédito; enquanto o risco de recuperação consiste na indeterminação do valor que poderá ser reavido pelo credor no caso de um *default* (BRITO e ASSAF NETO, 2008).

Santos (2010) comenta que uma grande apreensão dos credores é a determinação do risco de inadimplência, sendo que essa pode afetar diretamente a liquidez e a captação de recursos em reflexo de possíveis perdas financeiras.

De acordo com Silva (2008, p. 29):

As razões que levam o cliente ao não-cumprimento da promessa podem estar relacionadas ao seu caráter, a sua capacidade de gerir os negócios, aos fatores externos adversos ou a sua incapacidade de gerar caixa. Mesmo a garantia não devendo ser o fator decisivo para a concessão do empréstimo ou de um financiamento, alguns tipos de operações devem ser respaldados por garantias que equilibrem e compensem as fraquezas relacionadas às demais variáveis implícitas no risco de crédito.

A inadimplência pode ser tratada com um elemento classificador de risco de crédito, visto que tem influência direta na inclusão de juros em operações de crédito. No Brasil, país com elevadas taxas de juros na percepção geral, faz-se pertinente o estudo desse fenômeno, buscando analisar e melhor compreender os fatores que conduzem tais taxas a níveis tão altos (BARTH, 2004).

## 2.2 Análise de crédito para pessoas físicas

De acordo com Securato (2007) todas as informações colhidas do solicitante de crédito e averiguadas através de consultas a agências especializadas e documentos que as comprovem, dão suporte à análise de crédito da pessoa física.

O mesmo autor ainda expõe como ocorrem essas consultas no país:

No Brasil, cada credor utiliza um modelo próprio de ficha cadastral para a obtenção de informações básicas do cliente, recorrendo também a empresas especializadas, tais como: Serasa, e/ou Serviço de Proteção ao Crédito (SPC), e/ou Information Management (SCI), dentre outras; para validar, confirmar e complementar as informações obtidas diretamente do cliente e das fontes por ele indicadas (SECURATO, 2007, p. 33).

Já Santos (2010) afirma que análise de crédito tem como finalidade investigar a idoneidade do cliente e se, financeiramente, ele é capaz de saldar sua dívida. Completa dizendo que duas técnicas são utilizadas pelas empresas a fim de efetuar a análise de crédito: a subjetiva, que baseia-se em decisões e julgamentos humanos; e a objetiva, que tem por base procedimentos estatísticos.

A aplicação das técnicas subjetivas à pessoas físicas, refere-se a compilação e a comparação dos dados constantes na ficha cadastral com parâmetros quantitativos e qualitativos previamente estabelecidos (SECURATO, 2007, p. 34).

Ainda segundo Securato (2007) partindo dos parâmetros empregados no auxílio à concessão de crédito a P.J. surge uma adequação desses parâmetros básicos na concessão de crédito a P.F., que são orientados pelos chamados *C's do Crédito* - que se ajustam às microempresas e P.F. : Caráter; Capacidade; Capital; Colateral e Condições.

Santos (2010) completa dizendo que habitualmente entende-se por *C's do crédito* como o conjunto de elementos que fornecem as informações imprescindíveis a fim de realizar-se a análise da capacidade financeira dos clientes de modo subjetivo.

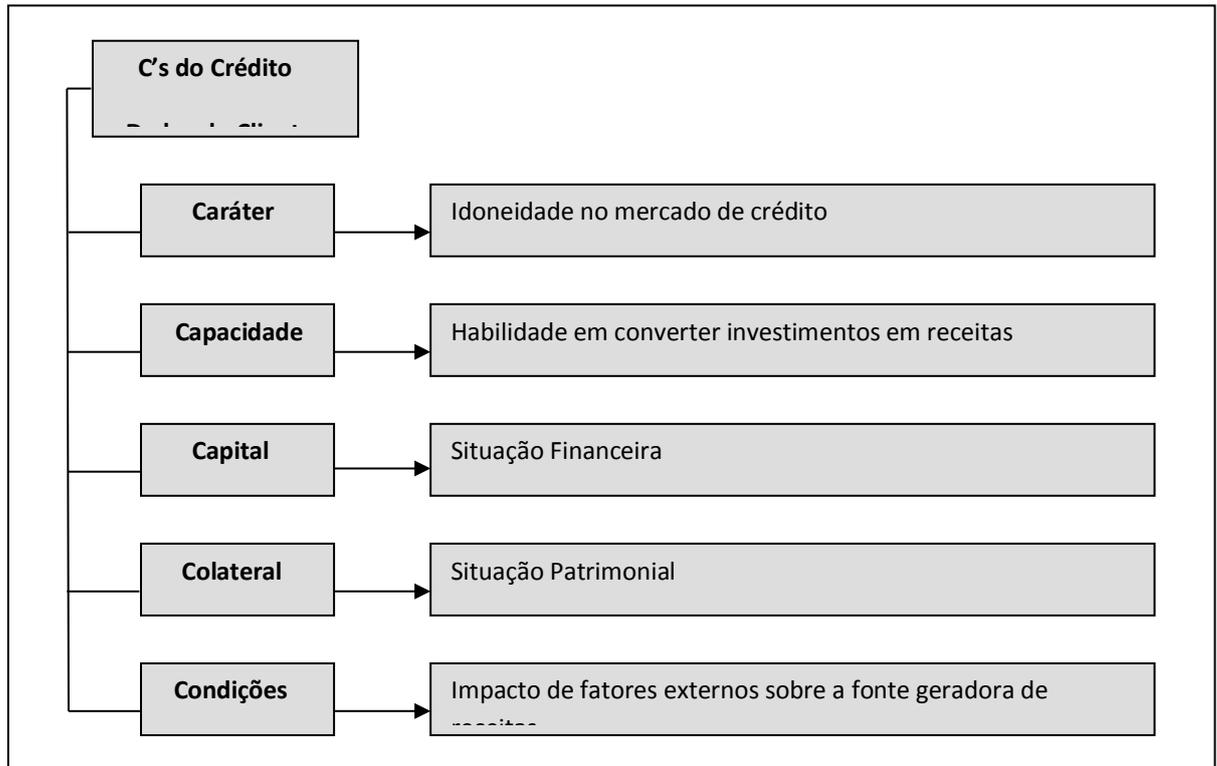


Figura 1 – C's do Crédito  
 Fonte: Santos (2010, p. 30)

Diferentemente da técnica subjetiva que leva muito em consideração a experiência, a intuição e o bom senso do analista de crédito tem-se, como alternativa à análise de crédito, a técnica objetiva, que é totalmente baseada em procedimentos estatísticos (SANTOS, 2010).

Dentre essas modelos destacam-se: o *Credit Scoring*, o *Behaviour Scoring*, o Modelo Misto e a Análise Discriminante.

### 2.3 Análise Discriminante

“É uma técnica estatística que auxilia a identificar quais as variáveis que diferenciam os grupos e quantas dessas variáveis são necessárias para obter a melhor classificação dos indivíduos de uma determinada população” (CORRAR et al, 2009, p. 234).

Ainda conforme Corrar et al (2009) essa análise tem como principal característica utilizar informações oriundas de variáveis independentes tentando conseguir um valor de uma variável dependente afim de atingir a classificação pretendida.

Para Barth (2004, p. 21) “na análise discriminante baseada em uma função discriminante linear, constrói-se uma função pela combinação linear das variáveis discriminantes”.

Hair et al. (2005, p. 209) explica que “a análise discriminante envolve determinar uma variável estatística, a combinação linear das duas (ou mais) variáveis independentes que discriminarão melhor entre grupos definidos *a priori*”. Segundo o mesmo autor, a função discriminante pode ser determinada por meio da seguinte equação:

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + \dots + W_nX_{nk} ,$$

Onde:

$Z_{jk}$  = escore Z discriminante da função discriminante j para o objeto k;

a = intercepto;

$W_i$  = peso discriminante para a variável independente i;

$X_{ik}$  = variável independente i para o objeto k.

“A leitura da função discriminante leva-nos a entender como um conjunto de variáveis influencia simultaneamente no comportamento de um elemento (CORRAR; PAULO; DIAS FILHO, 2009, p. 236)”.

Hair et al. (2005) pondera sobre as chances de elementos de um grupo serem mal classificados ilustrando na Figura 2 as distribuições de escores discriminantes entre os grupos A e B.

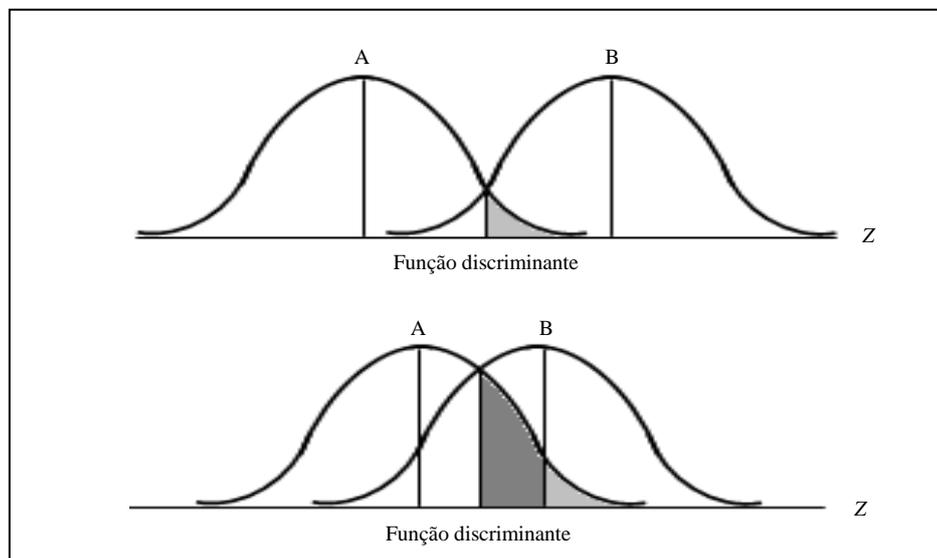


Figura 2 - Representação univariada de escores Z discriminantes  
Fonte: Hair et al. (2005a, p. 209)

Na Figura 2, observa-se que diagrama de cima exibe as distribuições de escores discriminantes em uma função onde A e B estão claramente separados, à medida que o diagrama abaixo mostra essa mesma distribuição mas para uma função onde há um discriminante relativamente pobre entre os grupos. Tem-se, ainda, representada a possibilidade de errônea classificação dos objetos dos grupos nas áreas sombreadas (HAIR et al., 2005, p. 209).

### 3 METODOLOGIA

Em relação ao problema abordado, esta pesquisa classifica-se em quantitativa, bibliográfica e descritiva.

Nesta pesquisa o procedimento escolhido é o estudo de caso que, de acordo com Yin (2010), se trata de uma pesquisa empírica que apura um fenômeno contemporâneo o qual esteja inserido no contexto real, principalmente quando não se têm claramente caracterizados

os limites entre os fenômenos e o contexto. O trabalho baseia-se em um estudo de caso, pois será realizado na empresa comercial Augusto Leonel Fernandes, e, dessa forma, os resultados serão relevantes especificamente para a referida empresa.

No presente estudo a população a ser pesquisada corresponde aos 5.115 clientes do segmento pessoa física de uma empresa comercial situada no município de Santana do Livramento, que efetuaram compras a prazo, no período de 1º de agosto de 2012 a 31 de janeiro de 2013.

O dimensionamento da amostra aleatória probabilística foi calculado para uma população finita, utilizando-se um erro de amostragem de 10%, com um nível de confiança de 95% e com proporção de elementos desfavoráveis e favoráveis de 0,5. Sendo assim, o tamanho de amostra ideal é 95 clientes.

Como critério de seleção da amostra, foi aplicada a Amostragem Aleatória Simples (AAS), onde a probabilidade de os indivíduos de determinado universo serem selecionados é exatamente a mesma. Na Amostragem Aleatória Simples, conforme Bolfarine e Bussab (2005, p.16), “cada unidade elementar é sorteada com igual probabilidade, individualmente, sem estratificação, e com um único estágio e seleção aleatória”.

As informações necessárias para a realização desta pesquisa foram retiradas do banco de dados cadastrais fornecidas pela empresa. Estas informações referem-se a dados pessoais e profissionais dos clientes pessoas físicas que efetuaram compras a prazo.

Para a realização do estudo, foram selecionadas quatro variáveis: Sexo ( $X_1$ ), Idade ( $X_2$ ), Renda ( $X_3$ ) e Estado Civil ( $X_4$ ). A escolha dessas variáveis se deu através de análise do formulário solicitado pela empresa, que contém as informações pessoais dos clientes no momento do preenchimento do cadastro. Dentre os dados fornecidos no cadastro, apenas essas quatro variáveis se mostraram importantes para a pesquisa, e, dessa forma, foram escolhidas para serem analisadas.

Através dessas quatro variáveis selecionadas e com o uso de técnicas estatísticas o estudo identificou o potencial de cada cliente pessoa física vir a fazer parte do grupo ‘inadimplente’ ou do grupo ‘adimplente’ em uma venda a crédito.

As técnicas utilizadas para a análise dos dados foram: Estatística Descritiva, Estimação de Parâmetros por Intervalo de Confiança para a média populacional e a técnica multivariada denominada de Análise Discriminante.

A técnica de estatística descritiva permitiu descrever o perfil do cliente pessoa física que realiza compras a prazo na empresa. As variáveis qualitativas (Sexo e Estado Civil) foram apresentadas por meio de tabelas de frequências para facilitar o entendimento desses dados. E, para variáveis quantitativas (Idade e Renda) foram calculadas as seguintes medidas descritivas: Amplitude de Variação, Média Aritmética e Desvio Padrão.

Na sequência, para inferir os dados para a população em estudo, utilizou-se a técnica de estimação de parâmetros por intervalo de confiança da média populacional, para as variáveis Idade e Renda. Para isso aplicou-se um nível de confiança de 95%.

Após, foi utilizada a técnica de análise discriminante com o intuito de propor o modelo de risco de crédito na modalidade pessoa física para a empresa comercial em estudo.

Segundo Corrar et al. (2009) essa técnica permite observar quais variáveis promovem a diferenciação dos grupos e quantas dessas variáveis imprescindíveis para que se possa classificar indivíduos de um universo da melhor maneira possível.

Já para as variáveis Sexo e Estado Civil, por se tratarem de variáveis qualitativas originalmente não mensuráveis, foram utilizadas variáveis *dummies*, denominadas de variáveis binárias que podem tomar um de dois valores, em geral 0 ou 1, isto é, servem para descrever qualquer evento que tenha apenas dois resultados possíveis (HILL, 2003).

Dentro da análise discriminante, a primeiro passo foi utilizar a técnica de análise de ausência de *outliers*, para verificar a existência ou não de valores atípicos que interferissem no

resultado final. Logo após, foi realizada a análise de presença de multicolinearidade, visando evitar a utilização de variáveis com informações semelhantes. O terceiro passo foi realizar a análise de regressão, para verificar a influência de cada uma das variáveis independentes (sexo, estado civil, idade e renda) sobre a variável dependente (inadimplência). E, por fim, foi verificada e calculada a capacidade preditiva do modelo.

As análises estatísticas foram realizadas com a utilização do software *SPSS 13.0 for Windows (Statistical Package for Windows)*.

## **4 RESULTADOS E DISCUSSÕES**

Neste capítulo serão apresentados os resultados obtidos através de cada uma das técnicas já mencionadas e que possibilitaram a realização da pesquisa.

### **4.1 Análise Descritiva**

Na análise descritiva procurou-se identificar o perfil dos clientes pessoas físicas da empresa estudada. Para isto, foram elaboradas tabelas de frequência para as variáveis Sexo e Estado Civil (Tabelas 1 e 2) e foi realizado o cálculo de medidas descritivas para as variáveis Idade e Renda (Tabelas 3 e 4). A medida descritiva de tendência central calculada para estas variáveis foi média aritmética e as medidas de dispersão foram amplitude de variação e desvio padrão.

Tabela 1 – Informações relativas à variável Sexo (n=95)

<b>Sexo</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Masculino	46	48,42
Feminino	49	51,58
<b>Total</b>	<b>95</b>	<b>100</b>

Fonte: Elaborada pelo autor

Verifica-se na Tabela 1 que a maioria dos clientes é do sexo feminino (51,58%), que são representados por 49 clientes da amostra. Já os clientes do sexo masculino, com 46 representantes, atingiram um percentual de 48,42% dos pesquisados.

Tabela 2 – Informações relativas à variável Estado Civil (n=95)

<b>Estado Civil</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Solteiros	23	24,21
Casados	52	54,74
Divorciados	11	11,58
Viúvos	09	9,47
<b>Total</b>	<b>95</b>	<b>100</b>

Fonte: Elaborada pelo autor

Na Tabela 2, observa-se que com relação ao estado civil 54,74% dos clientes da amostra são casados, 24,21% solteiros, 11,58% divorciados e 9,47% viúvos.

Tabela 3 – Medidas Descritivas das variáveis Idade e Renda

	n	H	Vlr. Mín.	Vlr. Max.	$\bar{x}$	S
Idade	95	60	19	79	51	15,21
Renda	95	15.800,00	200,00	16.000,00	1561,31	1841,04

Fonte: Elaborada pelo autor

Nota: n= tamanho da amostra; H = amplitude (valor máximo – valor mínimo); Vlr. Min. = valor mínimo da amostra; Vlr. Max. = valor máximo da amostra; X = média da amostra; S = desvio padrão da amostra.

Os valores da variável Renda estão descritos em Reais.

Observa-se, na Tabela 3, uma grande amplitude entre os clientes da amostra tanto na variável idade quanto na variável renda, e isto se dá pela grande diferença que há entre os valores mínimos e máximos das respectivas variáveis. Também, percebe-se que a idade média dos clientes da amostra é 51 anos e o valor da renda média mensal dos pesquisados é de R\$1.561,31.

No cálculo do desvio padrão da variável Idade verificou-se uma dispersão de 15,21 anos e para a variável Renda o valor desta medida foi de 1.841 reais. Com isso verifica-se que os dados possuem um grau de dispersão ou afastamento significativo dos valores observado em torno do valor central calculado (média) de cada variável analisada. O desvio padrão indica também que a amostra é heterogênea (bastante variabilidade).

#### 4.2. Estimação de Parâmetros

Na sequência, para inferir os dados amostrais para a população em estudo, utilizou-se a técnica de estimação de parâmetros por intervalo de confiança para a média populacional ( $\mu$ ) das variáveis idade e renda.

Tabela 4 – Estimação da média populacional por Intervalo de Confiança

Variáveis	$\bar{x}$	Estimação por I.C. (95%)	
		Limite Inferior	Limite Superior
Idade	51	47,95	51,15
Renda	1.561,31	1.186,28	1.936,35

Fonte: Elaborada pelo autor

Nota: X= média da amostra. Os valores da variável Renda estão descritos em reais.

Na Tabela 4, verifica-se que a idade da população está contida no intervalo entre 47,95 e 51,15 anos. Em relação à renda, observa-se que os rendimentos mensais dos clientes estão inseridos no intervalo entre R\$1.186,28 e R\$1.936,35. Percebe-se ainda que as médias das duas variáveis estão incluídas nos espaços compreendidos entre os limites inferiores e limites superiores.

#### 4.3 Análise Discriminante

Nesta seção são apresentadas as análises iniciais para utilização da técnica multivariada denominada análise discriminante.

##### 4.3.1 Análise de ausência de Outliers

Segundo Corrar, Paulo e Dias Filho (2009) a presença de outliers (valores atípicos) prejudica fortemente a análise discriminante, então foram realizados testes de outliers para identificar, transformar e, quando for o caso, eliminar da amostra antes da aplicação da análise

discriminante. A presença de tais valores (outliers) pode produzir tendenciosidade no modelo de ajustamento dos dados (FIELD, 2005, P.95).

Os resultados mostraram que a variável Renda (n=95) possui a presença de outliers, ou seja, de valores discrepantes, conforme a Figura 3.

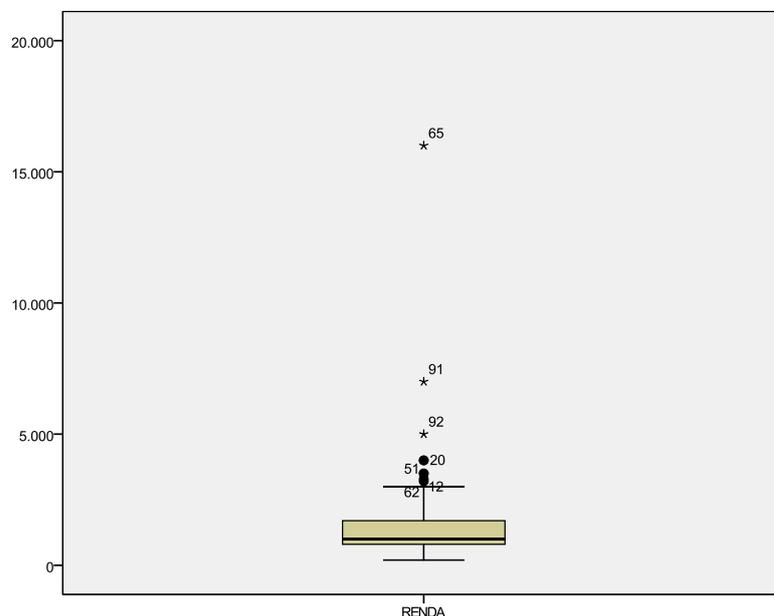


Figura 3 – Gráfico da Variável Renda  
 Fonte: Elaborado pelo autor  
 \*Outliers

No Figura 3, verifica-se que três valores da variável Renda são outliers e como a presença de outliers prejudica na análise discriminante, optou-se por eliminar este valores da amostra. Com isso passou-se a analisar uma amostra de tamanho n = 92 clientes.

#### 4.3.2 Análise de presença de multicolinearidade

Conforme Corrar, Paulo e Dias Filho(2009), a multicolinearidade acontece quando há semelhança entre as informações contidas em mais de uma variável independente do modelo que sirvam para elucidar o mesmo tema. Os mesmos autores afirmam que “duas ou mais variáveis independentes altamente correlacionadas levam a dificuldades na separação dos efeitos de cada uma delas isolada sobre a variável dependente [...]”.

Nesta análise a ferramenta estatística utilizada para verificar a ausência de multicolineariedade entre as variáveis independentes (Sexo, Idade, Renda e Estado Civil) foi o coeficiente de correlação de Pearson. Para isso, utilizou-se um nível de significância de 1% (Tabela 5).

Tabela 5 – Resultados do coeficiente de correlação de Pearson para a multicolinearidade

	<b>Sexo</b>	<b>Idade</b>	<b>Renda</b>	<b>Estado Civil</b>
<b>Sexo</b>				
C.C.	1	- 0,042	0,006	0,156
Significância		0,694	0,953	0,137
<b>Idade</b>				
C.C.	- 0,042	1	- 0,025	0,103
Significância	0,694		0,810	0,329
<b>Renda</b>				
C.C.	0,006	- 0,025	1	- 0,089
Significância	0,953	0,810		0,397
<b>Estado Civil</b>				
C.C.	0,156	0,103	- 0,089	1
Significância	0,137	0,329	0,397	

Fonte: Elaborada pelo autor

Nota: C.C.= coeficiente de correlação

Os resultados desta análise demonstraram que estatisticamente e a um nível de significância de 1% as variáveis independentes não estão correlacionadas entre si (p - valor > 1%). Com isso, testou-se no modelo de análise discriminante as quatro variáveis independentes (Sexo, Idade, Renda e Estado Civil).

#### 4.3.3 Análise de Regressão

Nesta etapa foi realizada uma análise de regressão linear múltipla que consiste em determinar uma função matemática que busca descrever o comportamento de determinada variável, denominada dependente, com base nos valores de uma ou mais variáveis, descritas como independentes (CORRAR, PAULO E DIAS FILHO, 2009).

Em uma análise de regressão linear busca-se a melhor relação entre uma variável dependente e uma ou mais variáveis independentes, quantificando a força dessas últimas e utilizando técnicas que possibilitem realizar previsões dos valores da resposta para valores dados do regressor (WALPOLE *et al*, 2009).

Segundo Anderson et al. (2003) na técnica de regressão, a variável que está sendo prevista é chamada de variável dependente, enquanto que a variável ou as variáveis que estão sendo usadas para explicar a variável dependente são chamadas de variáveis independentes.

Para Barrow (2007) um estudo de regressão descreve uma relação causal ajustando uma linha reta aos dados com a finalidade de resumi-los para um melhor entendimento. Este método também permite medir a influência de cada variável explicativa sobre a variável dependente, determinar a significância de cada variável independente, além de realizar previsões e estimativas.

No entanto, para estimar o modelo utilizou-se como variável dependente (Y) situação inadimplente dos clientes que efetuaram compras a prazo e as seguintes variáveis independentes: Sexo (X1), Idade (X2), Renda (X3) e Estado Civil (X4).

O cálculo da análise de variância (ANOVA) foi realizado para verificar se este modelo de regressão é significativo (Tabela 6).

Tabela 6 – Análise de Variância

ANOVA					
	GI	SQ	MQ	F	p – valor
Regressão	4	354	0,089	1,761	p < 0,10*
Resíduo	87	4,374	0,050		
Total	91	4,728			

Fonte: Elaborada pelo autor

Nota: gl = grau de liberdade; SQ = soma dos quadrados; MQ = quadrado médio; F = coeficiente F calculado.

\*Significativo ao nível de 10%.

Verifica-se que na Tabela 6 a significância da variável regressora e do modelo pelo resultado do teste F ou pelo p-valor. Neste caso, com 90% de confiança é possível concluir que pelo menos uma variável independente explica a variável dependente (p < 0,10).

Tabela 7 – Resumo da Regressão

Parâmetros Estimados				
	Coefficientes	Erro Padrão	t	p – valor
Interseção	2,039	0,096	21,245	0,000*
Sexo (X <sub>1</sub> )	- 0,073	0,047	- 1,547	0,125
Idade (X <sub>2</sub> )	0,000	0,002	- 0,080	0,936
Renda (X <sub>3</sub> )	0,00005628	0,001	- 1,981	0,050*
E. Civil (X <sub>4</sub> )	0,040	0,048	0,832	0,408

Fonte: Elaborada pelo autor

Nota: E. Civil = estado civil. \*Significativo ao nível de 10%.

Verifica-se na Tabela 7 que os valores do coeficiente da interseção (p < 0,10) e da variável Renda (p < 0,10) são significativos para estimar o modelo de regressão da função.

Nos resultados o valor do R-Quadrado, medida de aderência, explica 8% da variação total da inadimplência/adimplência. Este valor representa o efeito preditivo que a variável explicativa tem sobre a variável explicada, isto é, corresponde o quanto a situação de inadimplente/adimplente é explicada pela variação no valor da renda.

O modelo de regressão escolhido para estimar a função inadimplente/adimplente contém uma variável independente estatisticamente significativa (o valor da renda mensal dos clientes), representando uma regressão linear simples com um R-Quadrado de 8%. A expressão da função inadimplente/adimplente estimada é:

$$\text{Função estimada} = 2,039 + 0,00005628 (X_3)$$

O coeficiente da variável Renda positivo indica que quando aumenta a renda em uma unidade, a inadimplência tende a aumentar em 0,00005628 unidades. Cabe mencionar que estabelecida e testada a equação de regressão, a mesma poderá explicar o relacionamento entre as variáveis e estimar valores de Y para valores fixados de X.

#### 4.3.4 Capacidade preditiva do modelo

Nesta etapa foi testada a capacidade preditiva do modelo onde foi calculada a média aritmética dos valores discriminantes obtidos através do modelo do grupo de inadimplentes (2,1540, n=5) e a média aritmética dos valores discriminantes obtidos através do modelo do grupo de adimplentes da amostra (2,1102, n = 87). Na sequência calculou-se o ponto de corte

para classificar os grupos por meio do cálculo de uma média ponderada e cujo resultado foi 2,1126. Sendo assim, todos os escores discriminantes obtidos dos clientes maiores que o ponto de corte são considerados inadimplentes e todos os escores menores que o ponto de corte são adimplentes.

Tabela 8 – Tabela de classificação - previsão do modelo

Observado	Estimado		Totais
	Adimplentes	Inadimplentes	
Adimplente	59 (67,8%)	28 (32,2%)	87
Inadimplente	03 (60%)	02 (40%)	05
<b>Totais</b>	62	30	92

Fonte: Elaborada pelo autor

A Tabela 8 mostra os percentuais de acerto do modelo, onde verifica-se que do grupo dos adimplentes o modelo acertou em 67,8% (59 clientes) e do grupo de inadimplentes o modelo estimou corretamente 60% (03 clientes).

## 5 CONSIDERAÇÕES FINAIS

Este estudo teve por objetivo propor um modelo probabilístico de risco de crédito a pessoas físicas, em uma empresa comercial, com a utilização da técnica estatística de Análise Discriminante.

Em relação à metodologia, esta pesquisa classificou-se em quantitativa, bibliográfica e descritiva, e o procedimento escolhido foi o estudo de caso por se tratar de uma empresa específica. Utilizou-se uma amostra com 95 clientes que efetuaram compras a prazo, no período de 1/8/2012 a 31/1/2013. As informações dos clientes foram retiradas do banco de dados cadastrais da referida empresa.

Com esta pesquisa pode-se identificar o perfil do consumidor que realiza compras a prazo na empresa em estudo. Os resultados apurados mostram que os clientes são em sua maioria do sexo feminino (51,58%), casados (54,74%), com média de idade de 51 anos e renda mensal média de R\$1.561,31.

A variável dependente do estudo foi a situação de inadimplência dos clientes. Para testar o modelo, foram utilizadas as variáveis independentes: Sexo, Idade, Renda e Estado Civil. As informações de cada uma dessas variáveis foram obtidas pela empresa por meio do cadastramento dos clientes P.F. que realizam compras a prazo.

Através da utilização da técnica de análise discriminante, chegou-se ao resultado de que entre as quatro variáveis englobadas no estudo, com um nível de significância de 90%, a única que se mostrou relevante para explicar a inadimplência dos clientes foi a variável Renda (p-valor = 0,050).

Com a análise de regressão, comprovou-se o valor preditivo da variável, ou seja, quanto a situação de inadimplência é afetada quando ocorre alguma alteração no valor da renda.

Por fim, testou-se a capacidade preditiva do modelo proposto, e com os resultados obtidos se evidenciou que de um total de 87 clientes adimplentes, o modelo previu corretamente 59 casos (67,8%). Já entre os inadimplentes, de um total de 5 casos, o modelo fez a previsão correta de 3 (60%).

Sugere-se que a empresa inclua em seu cadastro outras variáveis que possam ser de maior relevância para a análise de concessão de crédito (como por exemplo: número de dependentes, renda familiar, se possui cartão de crédito). Outro ponto importante é que a empresa exija dos clientes o preenchimento da variável Imóvel Próprio, pois ela já está

presente no formulário de cadastramento e pode ajudar no estabelecimento de um modelo probabilístico de inadimplência mais completo.

Como sugestão de outro trabalho, indica-se a possibilidade de utilização de uma amostra diferente de mesmo tamanho, com o intuito de validar e conferir maior grau de certeza ao modelo proposto neste estudo.

Este trabalho é de relevância específica à empresa estudada por se tratar de um estudo de caso, porém a realização de trabalhos semelhantes em outras empresas comerciais da cidade e da região pode revelar, ou não, se os perfis de seus respectivos clientes apresentam similaridades.

Com isso, conclui-se que entre as variáveis explicativas analisadas, a única estatisticamente significativa, com um nível de significância de 10%, foi a variável Renda ( $p$ -valor = 0,05). Dessa forma, com o atual modelo de ficha cadastral adotado, essa variável se mostrou ser a única capaz de auxiliar na determinação da potencialidade de um cliente que realiza suas compras a prazo na empresa vir a se tornar inadimplente, e assim possibilitar aos gestores da empresa que consigam analisar e tomar a melhor decisão no que diz respeito à concessão de crédito.

## Referências Bibliográficas

ANDERSON, D. R. *et al.* **Estatística aplicada à Administração e Economia**. São Paulo: Pioneira Thomson Learning, 2003.

BARROW, M. **Estatística para economia, contabilidade e administração**. São Paulo: Ática, 2007.

BARTH, N. L. **Inadimplência: construção de modelos de previsão**. São Paulo: Nobel, 2004.

BERNARDES, R. S. **Processo de análise de crédito de uma instituição financeira Pública**. Monografia de conclusão de curso em Ciências Contábeis. Novo Hamburgo, Feevale, 2011.

BOLFARINE, H.; BUSSAB, W. **Elementos de Amostragem**. São Paulo, Editora Blucher, 2005.

BRITO, G. A. S.; ASSAF NETO, A. **Modelo de classificação de risco de crédito de empresas**. RCF, São Paulo, v. 10, n. 46, p. 18-29, Jan./Abr., 2008.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. **Análise multivariada: para os cursos de administração, ciências contábeis e economia**. 1. ed. São Paulo: Atlas, 2009

DAROIT, J. **Um estudo sobre a análise da concessão de crédito à pessoa física em instituições financeiras**. Revista Destaques Acadêmicos, CGO/Univates, n. 1, 2011.

FIELD, A. **Descobrendo a estatística usando o SPSS**. 2005. EDITORA ARTMED

HAIR Jr, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Análise multivariada de dados**. 5. ed. Porto Alegre: Bookman, 2005.

HILL, R. Carter. **Econometria**. 2.ed. São Paulo: Saraiva, 2003.

**OLIVEIRA, N. C. Métodos utilizados para análise de crédito de pessoa física nas instituições financeiras e sua relação com o índice de inadimplência.** Monografia de Graduação em Ciências Contábeis, PUC Minas, Belo Horizonte, 2010.

**SANTOS, J. O. Análise de crédito: empresas, pessoas físicas, agronegócio e pecuária.** 3. ed. São Paulo: Atlas, 2010.

**SECURATO, J. R. Crédito: análise e avaliação do risco: pessoas físicas e jurídicas.** São Paulo: Saint Paul Editora, 2007.

**SILVA, J. P. Gestão e análise de risco de crédito.** 6. ed. São Paulo: Atlas, 2008.

**WALPOLE, Ronald. E. et al. Probabilidade e estatística para engenharia e ciências.** São Paulo: Pearson Prentice Hall, 2009.

**YIN, R. K. Estudo de caso: planejamento e métodos.** 4. ed. Porto Alegre: Bookman, 2010.