



SPOLM 2008

ISSN 2175-6295

Rio de Janeiro- Brasil, 05 e 06 de agosto de 2008.

## **UMA ABORDAGEM ALTERNATIVA DE *CREDIT SCORING* USANDO ANÁLISE DISCRIMINANTE: Eficiência na Concessão de Crédito Para o Segmento de Pessoas Físicas no Brasil**

**Marcos dos Santos Dutra**

Centro Federal de Educação Tecnológica de Minas Gerais  
Avenida Amazonas, 7675, 30.510-000, Belo Horizonte, MG, Brasil.

[marcosdutra@dppg.cefetmg.br](mailto:marcosdutra@dppg.cefetmg.br)

**Elenice Biazi**

Centro Federal de Educação Tecnológica de Minas Gerais  
Av. Amazonas, 7675, 30.510-000, Belo Horizonte, MG, Brasil

[elenice@dppg.cefetmg.br](mailto:elenice@dppg.cefetmg.br)

### **RESUMO**

Neste trabalho é proposta uma abordagem alternativa de *credit scoring*, podendo ser de grande interesse para as instituições financeiras, diante do atual cenário do sistema financeiro brasileiro, que tem apresentado excesso de liquidez, gerado, principalmente, por investimentos externos. Além de tornar mais ágil o processo de concessão de crédito e controlar a inadimplência, esta metodologia possibilita ampliar o nível de concessão de crédito de forma eficaz, resguardando a relação risco-retorno.

Entre as diversas metodologias existentes, neste trabalho optou-se pela metodologia estatística de análise discriminante. Contudo, alguns pesquisadores inviabilizam o uso desta técnica em modelos de *credit scoring* devido às fortes pressuposições dos dados.

Os resultados indicam que o modelo proposto é conveniente e eficiente quando há violações das pressuposições em análise discriminante. Contudo, é fundamental o uso combinado de técnicas poderosas em estatística. A transformação de Box-Cox (1964) e a técnica de influência local de Cook (1986) são algumas delas.

**Palavras Chave:** *Análise Discriminante, Credit Scoring, Excesso de Liquidez*

### **ABSTRACT**

In this paper an alternative approach of *credit scoring* is proposed, that can be of great interest for the financial institutions considering the situation of the Brazilian financial system which has been presenting excess of liquidity generated mainly by external investments.

Therefore, besides turning the process of credit concession more agile and to better control the lack of payment, this methodology makes it possible to enlarge the level of concession of credit in an effective way guarding the relationship risk-return.

In this paper, among the several existing methodologies, the statistical methodology of discriminant analysis was chosen. However, many researchers disregard the use of this technique in models of *credit scoring* due to the strong assumptions of the data.

The results indicate that the proposed model is appropriate and efficient when there are violations of assumptions in discriminant analysis. However, it is fundamental the combined use of powerful techniques in statistics. The Box-Cox (1964) transformation and the technique of local influence of Cook (1986) are some of them.

**Key Words:** *Discriminant Analysis, Credit Scoring, Liquidity Excess*

## 1. INTRODUÇÃO

A expansão das operações de crédito, no Brasil após Plano Real foi abrangente em vários segmentos da economia, entretanto, as operações de crédito para pessoas físicas foram as que mais cresceram nos últimos anos. As informações consolidadas divulgadas pelo Banco Central para cada modalidade de crédito mostram que já no início do ano 1999 o saldo das operações concedidas a pessoas físicas representava um volume superior a 50% do saldo total das operações concedidas a pessoas jurídicas. Apesar desse considerável crescimento das operações de crédito no Brasil, sendo explicado, principalmente, pelo aumento do crédito para pessoas físicas; numa perspectiva internacional, o volume de crédito como proporção do nível de Produto Interno Bruto (PIB) ainda é muito baixo no país. O Banco Central apurou a relação crédito/PIB no ano 2007 em 36,5%, nível muito baixo se comparado com demais países, mesmo em desenvolvimento.

Paralelamente, o sistema brasileiro tem apresentado excesso de liquidez, segundo afirma Carneiro (2007). Isto é gerado, principalmente, por investimentos externos, que de acordo com as informações do Banco Central, no ano de 2003 somavam US\$ 10,1 bilhões, batendo recorde ao somar US\$ 34,616 bilhões em 2007. Diante disso, os bancos têm aumentado o nível de alavancagem e conseqüentemente, são induzidos a trabalhar com uma menor aversão ao risco<sup>1</sup>.

Diante desse cenário, é fundamental que as instituições financeiras além de tornarem mais ágil o processo de concessão de crédito e controlar a inadimplência, sejam capazes de ampliar a concessão de crédito de forma eficaz.

Assim, a gestão do risco de crédito vem alcançando uma posição de destaque nas instituições financeiras, e conseqüentemente existe um maior interesse por modelos de *credit scoring*. Contudo, estes modelos inibem a gestão de riscos, por oferecerem apenas duas opções, rejeição ou aceitação da operação. Não permitindo à instituição financeira o controle do nível de risco, ou seja, ser mais ou menos agressiva na concessão de crédito.

Neste trabalho é proposta uma metodologia alternativa de *credit scoring* capaz de atender as tendências atuais que induzem a operar com uma menor aversão ao risco, ou seja, maior agressividade e, além disso, resguarda a relação risco-retorno.

Entre as diversas metodologias existentes, neste trabalho optou-se pela metodologia estatística de análise discriminante. Esta metodologia, que apresenta a vantagem de classificar os indivíduos em grupos, de acordo com seu grau de risco, torna-se uma das aplicações mais difíceis em *credit scoring* devido às fortes pressuposições necessárias para correta avaliação dos tipos de dados envolvidos.

No entanto, através da aplicação de técnicas sofisticadas em estatística propõe viabilizar o uso da análise discriminante mesmo quando há violações de pressuposições.

Serão propostas três estratégias para melhor discriminação dos dados. Na estratégia 1 será proposta a transformação de Box-Cox (1964) para obtenção de uma distribuição normal aproximada. Na estratégia 2 será proposta a técnica denominada influência local de Cook (1986) para detectar observações influentes. E finalmente, na estratégia 3 será proposta a análise discriminante quadrática.

Este artigo terá início apresentando uma revisão sobre Análise Discriminante de Fisher e Quadrática. Nas seções 5 e 6 serão descritas o teste de normalidade multivariada de Mardia (1970,1974,1975) e a transformação de Box-Cox (1964). Na seção 7 será apresentada a

---

<sup>1</sup> Um exemplo interessante que mostra a agressividade dos bancos é explicado por Lisboa (2008).

técnica de influência local de Cook (1986). A estatística proposta e algumas outras técnicas estatísticas disponíveis na literatura são analisadas em um exemplo de aplicação financeira e os resultados estão presentes na seção 9. O artigo é concluído com uma discussão na seção 10.

## 2 ANÁLISE DISCRIMINANTE DE FISHER

Fisher (1936) introduziu a idéia de se construir funções discriminante a partir de combinações lineares das variáveis originais. O método envolve a suposição que há  $p$ -variáveis aleatórias,  $g$  populações normais  $p$ -variadas e iguais matrizes de covariâncias populacionais. Neste caso, é possível construir  $s$  combinações lineares,  $s \leq \min(g-1, p)$ , chamadas de funções discriminantes linear de Fisher, e que são definidas por:

$$\hat{Y}_j = \hat{e}'_j X_{p \times 1}, j = 1, 2, \dots, s \quad (1)$$

onde  $\hat{e}'_j$  é o  $j$ -ésimo autovetor correspondente ao  $j$ -ésimo maior autovalor da matriz  $W^{-1}B$  e tal que  $\hat{e}'_j W \hat{e}_j = 1$ , sendo  $W$  e  $B$  chamadas de matrizes soma de quadrados e produtos cruzados dentro (Within) dos grupos e entre (Between) grupos, sendo respectivamente definidas por:

$$W_{p \times p} = \sum_{i=1}^g \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_i)(X_{ik} - \bar{X}_i)' \quad (2)$$

$$B_{p \times p} = \sum_{i=1}^g n_i (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})' \quad (3)$$

Sendo  $X_{ik}$  o vetor de observações do elemento amostral  $k$  que pertence à população  $i$ ,  $\bar{X}_i$  o vetor de médias amostral da população  $i$ ,  $\bar{X}$  o vetor de médias amostral, considerando-se todas as  $n$  observações conjuntamente e  $n_i$  o número elementos pertencentes à amostra da população  $i$ ,  $i = 1, 2, 3, \dots, g$ ,  $\sum_{i=1}^g n_i = n$ . As combinações lineares (1) são as de maior poder discriminante dentro do conjunto de variáveis utilizadas para discriminação.

A primeira função discriminante é a relacionada com o maior autovalor da matriz  $W^{-1}B$  e tem o maior poder de discriminação. A segunda função discriminante é relacionada com o segundo maior autovalor de  $W^{-1}B$  e tem um poder de discriminação menor do que a primeira e maior do que a terceira, e assim sucessivamente.

Após a construção das funções discriminantes, para cada elemento amostral com vetor de observações  $x_j$ , ter-se-á um vetor com os seus escores nestas funções, isto é,  $\hat{Y}'_j = [\hat{e}'_1 x_j \ \hat{e}'_2 x_j \ \dots \ \hat{e}'_s x_j]$ . Além disso, ter-se-á os escores das funções discriminantes aplicadas aos vetores de média amostral observados para cada população, isto é,  $\bar{Y}'_i = [\hat{e}'_1 \bar{x}_i \ \hat{e}'_2 \bar{x}_i \ \dots \ \hat{e}'_s \bar{x}_i]$ ,  $i = 1, 2, \dots, g$ .

A regra de classificação consiste em calcular a distância Euclidiana entre os vetores  $\hat{Y}_j$  e  $\bar{Y}_i$  para todo  $i = 1, 2, 3, \dots, g$ , sendo o correspondente elemento amostral  $j$  classificado na população cuja distância é a menor.

## 3 FUNÇÃO DISCRIMINANTE QUADRÁTICA

Quando as matrizes de covariâncias  $\Sigma_i$  são diferentes, uma Função Discriminante Quadrática é apropriada. A função é dada por:

$$d_i^Q(x) = -\frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) + \ln p_i \quad (4)$$

onde

$\mu_i$  = vetor médio da população  $\Pi_i$ .

$\Sigma_i$  = matriz de covariâncias da população  $\Pi_i$ .

$p_i$  = probabilidade a priori de a observação pertencer à população  $\Pi_i$ .

$x$  = vetor aleatório.

A regra de classificação consiste em alocar  $x$  em  $\Pi_k$  se

$$d_k^Q(x) = \max d_i^Q(x), i = 1, \dots, k \quad (5)$$

#### 4 COMPARANDO MATRIZES DE COVARIÂNCIA

Nesta seção, serão abordados os testes da razão de verossimilhanças, utilizados para comparar a igualdade entre k-matrizes de covariâncias.

A rejeição da hipótese nula é dada quando  $\lambda \leq c$ , em que  $c$  é uma constante requerida de acordo com o tamanho do teste, normalmente conhecida como valor crítico, e  $\lambda$  é denominado como estatística do teste.

$$\lambda = \Lambda^{\frac{N}{2}} = \frac{\max_{\omega} L(\omega)}{\max_{\Theta} L(\Theta)} = \left( \prod_{j=1}^k |S_j|^{n_j/2} N^{pN/2} \right) / \left| \sum_{j=1}^k S_j \right|^{N/2} \prod_{j=1}^k n_j^{pn_j/2} \leq c \quad (6)$$

Sob a hipótese nula, a estatística do teste da razão de verossimilhanças (LRT) pode ser reescrita por

$$\text{LRT} = -2 \log \lambda \quad (7)$$

e possui distribuição qui-quadrado com  $\frac{p(p+1)(k-1)}{2}$  graus de liberdade.

A estatística do teste da razão de verossimilhanças descrita na expressão (6) pode ser modificada, tal que sua distribuição assintótica possa ser mais bem aproximada da distribuição qui-quadrado. Estas modificações propostas, na qual  $f_i (= n_j - 1)$ , os graus de liberdade associado com  $S_j$ , levam a um teste não tendencioso, dado por

$$M = \left( \prod_{j=1}^k |S_j|^{f_j/2} f^{pf/2} \right) / \left| \sum_{j=1}^k S_j \right|^{f/2} \prod_{j=1}^k f_j^{pf_j/2}, \quad (8)$$

onde  $f = N - k$

Modificações propostas por diferentes autores implicam em uma melhoria, principalmente para o caso de amostras pequenas. Box citado por Seber (1984, p.449) propôs duas aproximações para a distribuição de  $M$ , a aproximação qui-quadrado e a aproximação-F.

#### 5 TESTE DE ASSIMETRIA E CURTOSE DE MARDIA (1970, 1974, 1975)

Se  $X_1, X_2, \dots, X_n$  são amostras aleatórias normais multivariadas de tamanho  $n$  no espaço  $p$ -dimensional, o vetor aleatório  $X_j = [X_{j1}, X_{j2}, \dots, X_{jp}]^T$  de dimensões  $p \times 1$  tem densidade normal multivariada dada por

$$f(x) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)\right\} \quad (9)$$

em que

$\Sigma$  é matriz de covariância e  
 $\mu$  é o vetor de médias  $p \times 1$  populacionais.

Assim, sob a hipótese nula de normalidade multivariada, os coeficientes de assimetria e curtose, correspondem a  $\beta_{1p} = 0$  e  $\beta_{2p} = p(p+2)$ , respectivamente. Podendo os mesmos, serem estimados utilizando os seguintes estimadores:

$$\hat{\beta}_{1p} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n g_{ij}^3 \text{ e } \hat{\beta}_{2p} = \frac{1}{n} \sum_{i=1}^n g_{ii}^2 \quad (10)$$

em que

$g_{ij} = (\mathbf{X}_i - \bar{\mathbf{X}})' \mathbf{S}_n^{-1} (\mathbf{X}_j - \bar{\mathbf{X}})$ ,  $\bar{\mathbf{X}} = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j$  é o vetor de médias amostrais e

$\mathbf{S}_n = \frac{1}{n} \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'$  é o estimador viesado de máxima verossimilhança da matriz de covariância amostrais.

O passo seguinte, é aplicar o teste de simetria, cuja estatística é dada por

$$k_1 = \frac{n\hat{\beta}_{1p}}{6} \quad (11)$$

possui distribuição assintótica de qui-quadrado com  $p(p+1)(p+2)/6$ , graus de liberdade sob a hipótese nula  $H_0 : \beta_{1p} = 0$ . Da mesma forma, será aplicado o teste para distribuição mesocúrtica (ou curva normal). A estatística do teste é dada por

$$k_2 = \frac{\hat{\beta}_{2p} - p(p+2)}{\sqrt{8p(p+2)/n}} \quad (12)$$

possui distribuição assintótica normal padrão sob a hipótese nula.

Se na aplicação dos testes, a hipótese nula for rejeitada a um nível nominal  $\alpha$  então, a distribuição dos dados é considerada não-normal.

Caso haja violação da pressuposição de normalidade multivariada, Seber (1984) diz que uma apropriada transformação pode frequentemente produzir um conjunto de dados que segue uma distribuição normal.

Na próxima seção será apresentada a Transformação de Box-Cox (1964) que poderá ser aplicada caso haja violação da normalidade multivariada dos dados.

## 6 TRANSFORMAÇÃO DE BOX-COX (1964)

Box e Cox (1964) têm sugerido uma família de transformações para normalizar observações, para estabilizar a variância, e linearizar a relação entre variáveis dependentes e independentes. Exemplos notáveis desta família de transformações são (a) transformação raiz quadrada para estabilizar a variância e para remover a não-normalidade, (b) transformação raiz cúbica para remover a não-normalidade, e (c) transformação logarítmica para estabilizar a variância e para remover a não-normalidade.

Box e Cox consideram uma família de transformações, sendo definida por

$$x^{(\lambda)} = \begin{cases} \frac{x^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(x), & \lambda = 0 \text{ e } x > 0 \end{cases} \quad (13)$$

que simultaneamente satisfaz todas as três suposições. O coeficiente  $\lambda$  pode ser estimado pelo método da máxima verossimilhança. A máxima verossimilhança estimada de  $\lambda$  maximiza a função verossimilhança  $L(\lambda)$ . Além disso, para testar se a máxima verossimilhança estimada  $\lambda$  é estatisticamente igual a 1, o qual indica uma normalidade, o seguinte teste da taxa verossimilhança tem sido proposto:

$$2[L_{\max}(\lambda) - L_{\max}(1)] \leq \chi_1^2(\alpha) \quad (14)$$

onde  $\chi_1^2(\alpha)$  denota o limite superior do ponto  $100\alpha\%$  de  $\chi^2$  com 1 grau de liberdade.

## 7 INFLUÊNCIA LOCAL

Dado um conjunto de observações, seja  $l(\beta)$  o logaritmo da função de verossimilhança correspondente ao modelo postulado, sendo que  $\beta$  é um vetor  $(p+1) \times 1$  de parâmetros desconhecidos. Perturbações podem ser introduzidas no modelo por um vetor  $\omega^T = (\omega_0, \omega_1, \dots, \omega_n)$  pertencente a um subconjunto aberto  $\Omega$  de  $\mathfrak{R}^n$ .

Supondo que o esquema de perturbação esteja definido, denotado por  $l(\beta \setminus \omega)$  como logaritmo da função de verossimilhança perturbada, o vetor  $\omega$  expressa um esquema de pesos, existindo um ponto  $\omega_0$ , em que  $l(\beta | \omega_0) = l(\beta)$ . Dado que  $\hat{\beta}$  é o estimador de máxima verossimilhança obtido por meio de  $l(\beta)$  e  $\hat{\beta}_\omega$  é o estimador de máxima verossimilhança obtido por meio de  $l(\beta | \omega)$ , o objetivo é comparar  $\hat{\beta}$  e  $\hat{\beta}_\omega$ , quando  $\omega$  varia em  $\Omega$ . Cook (1986) sugere que a comparação entre  $\hat{\beta}$  e  $\hat{\beta}_\omega$  seja feita por afastamento pelo logaritmo da função de verossimilhança  $LD(\omega)$ , expresso da seguinte maneira:

$$LD(\omega) = 2 \left[ l(\hat{\beta}) - l(\hat{\beta}_\omega) \right] \quad (15)$$

Dessa forma,  $LD(\omega)$  contém informação essencial sobre a influência do esquema de perturbação.

A idéia de Cook (1986) é estudar o comportamento da função  $LD(\omega)$  numa vizinhança  $\omega_0$ , que é o ponto em que as duas verossimilhanças são iguais. Para isso, o autor considerou a seguinte superfície geométrica:

$$\alpha(\omega) = \begin{pmatrix} \omega \\ LD(\omega) \end{pmatrix}, \quad (16)$$

que é denominada de gráfico de influência. O estudo de influência local consiste em analisar como a superfície  $\alpha(\omega)$  desvia-se de seu plano tangente em  $\omega_0 (T_0)$ . Essa análise

pode ser feita estudando-se as curvaturas das seções normais da superfície  $\alpha(\omega)$  em  $\omega_0$ , que são as intersecções de  $\alpha(\omega)$  com planos contendo o vetor normal com seu plano tangente em  $\omega_0$ . As curvaturas dessas seções são denominadas curvaturas normais.

A interseção entre a seção normal e o plano tangente  $T_0$  é denominada linha projetada. Essa linha pode ser obtida através do gráfico de  $LD(\omega_0 + ah)$  contra  $a \in \mathfrak{R}$ .

A curvatura normal da linha projetada, denotada por  $C_h$ , é definida como sendo a curvatura de  $(a, LD\{\omega(a)\})$  em  $a=0$ , em que  $\omega(a) = \omega_0 + ah$ . Denomina-se  $C_h$  curvatura normal da superfície  $\alpha(\omega)$  em  $\omega_0$  e na direção unitária  $h$ .

Cook (1986) mostra que a curvatura normal na direção  $h$  pode ser expressa por:

$$C_h = 2|h^T \mathbb{F}_h|, \quad (17)$$

Sendo que  $\mathbb{F} = \Delta^T [I(\hat{\beta})]^{-1} \Delta$ ,  $I(\hat{\beta})$  é a matriz de informação observada sob o modelo postulado e  $\Delta$  é a matriz  $(p+1) \times n$  definida por:

$$\Delta = \frac{\partial^2 l(\beta | w)}{\partial \beta \partial w^T} \quad (18)$$

avaliada em  $\beta = \hat{\beta}$  e  $w = \omega_0$ .

O interesse particular está na direção (ou nas direções) que produz(em) maior influência local. Segundo Cook (1986), a direção que produz a maior mudança local na estimativa dos parâmetros é dada por  $l_{\max}$ , que corresponde ao autovetor normalizado correspondente ao maior autovalor da matriz  $\Delta^T I(\hat{\beta})^{-1} \Delta$ . O vetor  $l_{\max}$  é utilizado para identificar as observações que podem estar controlando propriedades importantes na análise dos dados.

Cook(1986), equação (38) e (39), demonstra que para o caso Variáveis Explanatórias, os autovalores de  $\mathbb{F}$  e a curvatura normal são dados respectivamente por

$$e^T e \delta_i / \sigma^2 + \sum_j \hat{\beta}_j^2 s_j^2 / \sigma^2 \quad (19)$$

$$C_{\max} = 2e^T e \delta_{\max} / \sigma^2 + 2 \sum_j \hat{\beta}_j^2 s_j^2 / \sigma^2 \quad (20)$$

## 8 VALIDAÇÃO DO MODELO

Segundo Caouette, Altman e Narayanan (1999, p.214) considera-se um modelo funcional de *credit scoring* válido quando este apresenta uma taxa de sucesso (número de acertos por casos totais) superior a 80%. Neste caso, pode-se considerar o modelo de *credit scoring* desenvolvido válido para o objetivo proposto. Contudo, será apresentado a seguir um método estatístico para verificar a adequação do modelo.

Um dos métodos mais mencionados na literatura para verificar a discriminação para  $g > 2$  populações, com  $n_1, n_2, \dots, n_g$  observações respectivamente e o total de observações  $n = n_1 + n_2 + \dots + n_g$  é o método de Wilk, sendo dado por

$$\Lambda = |W|/|T| \quad (21)$$

A matriz  $W$  é obtida da matriz definida em (2) e a matriz  $T$  é chamada de matriz soma de quadrados e produtos cruzados total sendo obtida pela soma das matrizes (2) e (3).

O Lambda de Wilk ( $\Lambda$ ) pode então ser convertido a um valor de  $F$  usando a seguinte transformação

$$\left( \frac{1 - \Lambda^{1/2}}{\Lambda^{1/2}} \right) \left( \frac{n - g - 1}{(g - 1)} \right) \quad (22)$$

a qual em grandes amostras tem uma distribuição  $F$  com  $2(n - g)$  e  $2(n - g - 1)$  graus de liberdade se as  $(j + 1)$  variáveis não produzem uma melhoria significativa na discriminação entre os grupos.

Após a construção e validação da metodologia alternativa proposta neste trabalho, é possível que o resultado da classificação apresente uma mistura de observações, a qual poderá estar concentrada em dois grupos. Diante disso, é proposto um último procedimento com a finalidade de reclassificar essas observações e consequentemente reduzir os erros de classificação. O procedimento consiste nos seguintes passos:

- 1) Selecionar na base de dados, após a efetivação da transformação de Box-Cox (estratégia 1), os dois grupos que apresentaram maior mistura de observações;
- 2) Construir um modelo de análise discriminante, considerando os dois grupos apresentados no passo 1.
- 3) Reclassificar, usando o modelo de análise discriminante, obtido no passo 2, as observações dos dois grupos classificadas pela metodologia alternativa, que apresentaram maior mistura de observações.

## 9 APLICAÇÃO

Um conjunto de dados reais de uma instituição financeira será usado para ilustração, aplicando as três estratégias apresentadas para uma melhor discriminação dos dados.

O primeiro grupo é denominado “bom” sendo composto por clientes que não obtiveram atrasos ou que obtiveram atrasos até 60 dias. O segundo grupo é denominado “intermediário”, composto por clientes que renegociaram os contratos vencidos e liquidaram o crédito. O último grupo é chamado “ruim”, composto por clientes que obtiveram atrasos superiores há 180 dias.

A amostra é composta por 1151 contratos de clientes que tomaram empréstimo em 02/01/2005 à 31/12/2006. Para cada um destes contratos, foram obtidas 11 variáveis. Estas variáveis foram denotadas por  $x_1, x_2, \dots, x_{11}$ , sendo que as 9 primeiras variáveis são contínuas ou discretas e as duas últimas são binárias. Diante disso, aplicou-se o teste de Mardia (1970,1974,1975) apenas no conjunto formado pelas variáveis  $x_1, x_2, \dots, x_9$ , segundo um algoritmo computacional proposto por DeCarlo (1997), executável no aplicativo SPSS versão 14, obtendo-se os estimadores dos coeficientes de assimetria e curtose, respectivamente por  $\hat{\beta}_{1,d} = 46,43$  e  $\hat{\beta}_{2,d} = 162,08$ . De (11) e (12), com  $p=9$ , obtém-se:

$$k_1 = 8913,15, \text{ o qual é significativo quando comparado com } \chi_{165}^2, \text{ e}$$

$$k_2 = 76,04, \text{ o qual é significativo quando comparado com } N(0,1).$$

Assim, conclui-se que os dados apresentam um afastamento notório da normalidade. Neste caso, será proposta a estratégia 1: transformação de Box-Cox.

### 9.1 ESTRATÉGIA 1

Com os dados aproximadamente normalizados aplicou-se a Análise Discriminante de Fisher no aplicativo SPSS versão 14. Os resultados são apresentados na tabela 1.



Tabela 1: Resultado da Função Discriminante de Fisher – dados transformados

Grupo		Grupo Previsto			Total
		1	2	3	
Original Contagem	1	231,00	140,00	67,00	438,00
	2	45,00	202,00	-	247,00
	3	16,00	32,00	418,00	466,00
%	1	52,74	31,96	15,30	100,00
	2	18,22	81,78	-	100,00
	3	3,43	6,87	89,70	100,00

73,9% do grupo original classificado corretamente.

Para responder a questão concernente se houve melhoria no resultado aplicando a Transformação de Box-Cox (1964), foi aplicado a Análise Discriminante de Fisher para os dados não transformados. Os resultados são apresentados na tabela 2.

Tabela 2: Resultado da Função Discriminante de Fisher – dados sem transformação

Grupo		Grupo Previsto			Total
		1	2	3	
Original Contagem	1	278,00	58,00	102,00	438,00
	2	60,00	136,00	51,00	247,00
	3	69,00	24,00	373,00	466,00
%	1	63,47	13,24	23,29	100,00
	2	24,29	55,06	20,65	100,00
	3	14,81	5,15	80,04	100,00

68,4% do grupo original classificado corretamente.

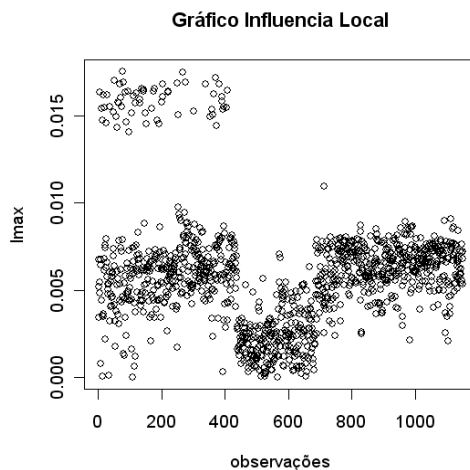
Analisando os resultados apresentados nas tabelas 1 e 2, observa-se uma melhoria significativa quando aplicado a aproximação de Box-Cox (1964) de 68,4% para 73,9%.

O próximo passo da análise é verificar se existem observações que possam estar influenciando o resultado da análise. Diante disso, será aplicada a estratégia 2: influência local.

## 9.2 ESTRATÉGIA 2

Tem-se que a curvatura  $C_{\max} = 2,44$  computada de (20) é relativamente pequena. Contudo, Verbeke e Molenberghs (2000) têm apresentado que, indiferente do tamanho de  $C_{\max}$ , uma inspeção de  $l_{\max}$  é válida. Sendo que os maiores valores de  $l_{\max}$  correspondem às observações influentes.

Na figura 1, é apresentado o gráfico de  $l_{\max}$  versus *observações*, sendo encontradas 67 observações que se destacaram das demais. Fazendo-se uma análise mais precisa, nota-se também que houve certa discriminação dos três grupos e, além disso, existem outras observações que podem estar exercendo influencia indevida no conjunto de dados, sendo verificadas 95 observações supostamente influentes.



**Figura 1:** Gráfico de Influência – Caso variáveis explanatórias

Para reanálise dos dados opta-se pela retirada ou não das observações influentes, neste caso seguiram-se os seguintes procedimentos:

**Procedimento 1:** Aplicação da Análise Discriminante de Fisher após a retirada dos 67 pontos supostamente influentes;

**Procedimento 2:** Aplicação da Análise Discriminante de Fisher após a retirada dos 95 pontos supostamente influentes.

É importante constatar qual dos dois procedimentos acima será mais adequado para melhor discriminação dos dados. Os resultados são apresentados nas tabelas 3 e 4, respectivamente.

TABELA 3: Classificação da Função Discriminante de Fisher – procedimento 1.

Grupo		Grupo Previsto			Total
		1	2	3	
Original Contagem	1	256,00	115,00	-	371,00
	2	62,00	185,00	-	247,00
	3	17,00	31,00	418,00	466,00
%	1	69,00	31,00	-	100,00
	2	25,10	74,90	-	100,00
	3	3,65	6,65	89,70	100,00

79,2% do grupo original classificado corretamente.

TABELA 4: Classificação da Função Discriminante de Fisher – procedimento 2.

Grupo		Grupo Previsto			Total
		1	2	3	
Original Contagem	1	251,00	103,00	-	354,00
	2	59,00	185,00	-	244,00
	3	17,00	23,00	418,00	458,00
%	1	70,90	29,10	-	100,00
	2	24,18	75,82	-	100,00
	3	3,71	5,02	91,27	100,00

80,9% do grupo original classificado corretamente.

Observando os resultados obtidos através da aplicação dos procedimentos 1 e 2, os quais comparados com a tabela 1 apresentaram uma melhoria significativa de 73,9% para 79,2% e de 73,9% para 80,9%, respectivamente.

Diante do resultado é possível constatar que o procedimento de diagnósticos em análise de influência proposta por Cook (1986) é eficiente para o modelo proposto. Contudo, é importante verificar se a igualdade das matrizes de covariâncias pode ser contestada, neste

caso a função discriminante quadrática será mais indicada. A seguir é proposta a estratégia 3, caso o testes da razão de verossimilhanças rejeite a premissa de igualdade das matrizes de covariâncias.

### 9.3 ESTRATÉGIA 3

Aplicado o testes da razão de verossimilhanças e convertido a um valor F, devido a Box(1979), conforme apresentado na seção 4, obtiveram-se os valores de 6,29 e 11,16 para os dois procedimentos, respectivamente. Como o P-valor nos dois casos é aproximadamente 0,001, menor que o nível de significância  $\alpha$ , igual a 0,05, é rejeitado a hipótese das matrizes de covariância intra-classe serem iguais nos dois modelos. Para contornar este problema a Função Discriminante Quadrática será mais indicada. Os resultados estão presentes na tabela 5 e 6.

TABELA 5: Classificação da Função Discriminante Quadrática – procedimento 1

		Grupo	Grupo Previsto			Total
			1	2	3	
Original	Contagem	1	326,00	45,00	-	371,00
		2	109,00	138,00	-	247,00
		3	24,00	24,00	418,00	466,00
	%	1	87,87	12,13	-	100,00
		2	44,13	55,87	-	100,00
		3	5,15	5,15	89,70	100,00

81,4% do grupo original classificado corretamente.

TABELA 6: Classificação da Função Discriminante Quadrática – procedimento 2

		Grupo	Grupo Previsto			Total
			1	2	3	
Original	Contagem	1	349,00	5,00	-	354,00
		2	142,00	102,00	-	244,00
		3	17,00	23,00	418,00	458,00
	%	1	98,59	1,41	-	100,00
		2	58,20	41,80	-	100,00
		3	3,71	5,02	91,27	100,00

82,3% do grupo original classificado corretamente.

Aplicado a função discriminante quadrática verifica-se uma melhoria de 79,2% para 81,4% e 80,9% para 82,3%, conforme tabelas 3 e 5; 4 e 6.

O passo seguinte é verificar a adequação do modelo de discriminação. Neste caso, será validado apenas o modelo o qual apresentou melhor desempenho, cujos resultados de classificação estão na tabela 6.

### 9.4 VALIDAÇÃO DO MODELO

Utilizando o método de Wilk, o mesmo apresentou o valor  $F$  aproximado de 126,25. O p-valor calculado apresentou ser menor que o nível de significância  $\alpha$ , igual a 0,05, neste caso, deve-se rejeitar a hipótese nula  $H_0$ : os vetores médios das 3 classes são iguais, em favor da hipótese alternativa  $H_a$ : os vetores médios das 3 classes são diferentes. Portanto, conclui-se que as diferenças entre as três amostras são adequadamente descritas pelas funções discriminantes.

Apesar do resultado final de 82,3% de acerto ser satisfatório, observa-se que 58,20% dos clientes “intermediários” foram classificados como “bons”. Isto significa que, apenas 41,80% dos clientes “intermediários” foram classificados corretamente, o que pode inviabilizar o uso desta metodologia alternativa.

O percentual de aceitação do grupo “intermediário” dependerá da política da empresa, assim, uma grande mistura de clientes “intermediários” classificados como “bons”, poderá até mesmo suplantar os limites de aceitação de clientes “intermediários” estabelecido pela empresa.

Neste caso, utilizaram-se os procedimentos descritos na seção 8 para reclassificar as observações dos dois grupos, bom e “intermediário”, os quais apresentaram uma concentração de observações misturadas.

Após reclassificar os dois grupos, “bom” e “intermediário”, são apresentados na tabela 7 o resultado da classificação do modelo aplicado o uso combinado das metodologias apresentadas neste trabalho.

TABELA 7: Resultado da Classificação do Modelo Final

	Grupo	Grupo Previsto			Total
		1	2	3	
Original Contagem	1	308,00	46,00	-	354,00
	2	88,00	156,00	-	244,00
	3	15,00	25,00	418,00	458,00
%	1	87,01	12,99	-	100,00
	2	36,07	63,93	-	100,00
	3	3,28	5,46	91,27	100,00

83,5% do grupo original classificado corretamente.

Este último procedimento foi adequado, pois além de aumentar em 1,1% o poder de precisão do modelo, foi possível reduzir em 22,13% os erros de classificação de clientes “intermediários” classificados como “bom”. Contudo, ainda existe um percentual muito elevado de clientes “intermediários” classificados como “bons”.

## 10 DISCUSSÃO

A avaliação de crédito constitui um processo dinâmico, exigindo revisões periódicas, devendo o modelo utilizado ser formatado para suprir em tempo hábil, as mudanças observadas no mercado. Foi diante desta perspectiva que se propôs, neste trabalho, a elaboração de uma metodologia alternativa de *credit scoring*.

Nesse trabalho, ficou evidenciada a importância da inclusão deste grupo “intermediário” diante do atual cenário do sistema financeiro brasileiro o qual tem apresentado excesso de liquidez. Essa metodologia alternativa mostrou ser capaz de controlar a inadimplência e, além disso, possibilita ampliar a concessão de crédito resguardando a relação risco-retorno.

O resultado final de 83,5% de acerto comprova que o modelo de risco de crédito desenvolvido é válido para o objetivo proposto, conforme a funcionalidade do modelo apresentada neste trabalho na seção 8. Os resultados indicam ainda que a aplicação do modelo na prática será capaz de classificar corretamente uma proporção bastante elevada de todas as operações de crédito submetidas à análise. O que comprova que podem ser obtidos excelentes resultados nos modelos de *credit scoring* usando análise discriminante, mas para isso é fundamental que os problemas de violações de pressuposições sejam resolvidos.

Um problema apresentados neste trabalho foi que apesar de serem obtidos excelentes resultados na classificação geral dos grupos, ainda existem 36,07% de clientes do grupo “intermediário” classificado como “bom”, podendo ser considerado um percentual bastante elevado. Este problema pode estar relacionado com a aplicação neste trabalho de um caso de pequena dimensão (1151 clientes e 11 variáveis) se comparado aos casos reais. Outro fato relevante é que apesar desta metodologia alternativa considerar a inclusão de um grupo “intermediário” que pode oferecer uma ampliação na concessão de crédito, podendo também ser um grande potencial de lucro para a instituição, o modelo estatístico teve como objetivo principal classificar operações de crédito. Assim, é recomendado que seja feito paralelamente

algum estudo verificando se esta metodologia de fato trará resultados positivos para aumentar a lucratividade da carteira de crédito.

## 11 REFERÊNCIAS

- [ 1 ] **Box, G. E. P. e Cox, D. R.** (1964), “An analysis of transformations”, *Journal of the Royal Statistical Society*, 26, 211-252.
- [ 2 ] **Caouette, J. B., Altman, E. I. e Narayanan, P.**, “*Gestão do Risco de Crédito*”, 1º ed., Qualitymark, Rio de Janeiro, 2000.
- [ 3 ] **Carneiro, D.D.**, Por que o Brasil vai crescer, mas não muito, *Jornal Estado de São Paulo*, 2007 ([www.estado.com.br/editorias/2007/02/02/eco-1.93.4.20070202.5.1.xml](http://www.estado.com.br/editorias/2007/02/02/eco-1.93.4.20070202.5.1.xml)) Acesso em: 27 fev. 2008.
- [ 4 ] **Cook, R. R.** (1986), “Assessment of local influence (with discussion)”, *Journal of the Royal Statistical Society*, 48, 133-169.
- [ 5 ] **DeCarlo, L. T.** (1997), “On the Meaning and Use of Kurtosis”, *Psychological Methods*, 3, 292-307.
- [ 6 ] **Fisher, R. A. (1936)**, “The use of multiple measurement in taxonomic problems”, *Annals of Eugenics*, 7, 179-188.
- [ 7 ] **Lisboa, C.**, Qualidade do crédito avança nos quatro maiores bancos, *Diário do Comércio*, São Paulo, 2008 ([www.dci.com.br/noticia.asp?id\\_noticia=215188](http://www.dci.com.br/noticia.asp?id_noticia=215188)) Acesso em: 29 fev. 2008.
- [ 8 ] **Mardia, K.V.**(1970), “Measures of multivariate skewness e kurtosis with applications”, *Biometrika*, 57, 519-530.
- [ 9 ] **Mardia, K.V.** (1974), “Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies”, *Sankhyā B*, 35,115-128.
- [10] **Mardia, K.V.** (1975), “Assessment of multinormality and therobustness of Hotelling’s  $T^2$  test”, *Applied Statistics*, 24, 163-171.
- [11] **Seber, G. A. F.**, “*Multivariate Observations*”, 1º ed., John Wiley & Sons, Canada, 1984.
- [12] **Verbeke, G e Molenberghs, G.**, “*Linear Mixed Models for Longitudinal Data*”, 1º ed., Springer, New York, 2000.