

**COMBINAÇÃO DE TÉCNICAS PARA
CONSTRUÇÃO DE UM MODELO
DE *CREDIT SCORING*, USANDO
ANÁLISE DISCRIMINANTE**

Marcos dos Santos Dutra
Elenice Biazzi

COMBINAÇÃO DE TÉCNICAS PARA CONSTRUÇÃO DE UM MODELO DE *CREDIT SCORING*, USANDO ANÁLISE DISCRIMINANTE

Marcos dos Santos Dutra*

Elenice Biazii**

RESUMO

Apesar de alguns pesquisadores criticarem o uso da análise discriminante em modelos de *credit scoring*, devido as fortes pressuposições necessárias para correta discriminação dos dados, sua utilização é bastante usual. Este artigo propõe o uso combinado de técnicas capazes de amenizar os problemas de violações das pressuposições da análise discriminante e conseqüentemente reduzir os erros de classificação. Os resultados indicam que o uso combinado das técnicas propostas neste trabalho foi capaz de aumentar a eficiência do modelo em 16,30 pontos percentuais.

Palavras-chave: Análise Discriminante. Credit Scoring. Técnicas. Classificação. Eficiência do Modelo.

1 INTRODUÇÃO

Fisher (1936) desenvolveu a análise discriminante, técnica estatística de análise multivariada que, a partir de características disponíveis de um indivíduo, cria uma regra de classificação que permite inferir a que população ele pertence. Isto permitiu o desenvolvimento dos primeiros modelos de *credit scoring*.

Como afirmam Caoutte, Altman e Narayanan (1999), os modelos tradicionais de *credit scoring* atribuem pesos estatisticamente

predeterminados a alguns atributos dos solicitantes, para gerar um escore de crédito. Se esse escore é favorável, quando comparado a um valor de corte, então a solicitação é aprovada. Assim, os modelos de *credit scoring* são uma ferramenta valiosa para decisões de aprovação ou não de pedidos de crédito, obedecendo à hipótese que o público alvo da carteira de crédito, após a implementação do modelo se mantenha o mesmo que no passado, sobre o qual os procedimentos estatísticos se baseiam.

Caouette, Altman e Narayanan (1999) afirmam que, a maioria dos modelos de *credit scoring* apresentam alguns defeitos, como a violação da normalidade multivariada que pode afetar a validade estatística dos modelos. Entretanto, dificilmente são apresentados métodos, para solucionar estes problemas de violações. O que existe em abundância são discussões acerca dos problemas das metodologias estatísticas, mas de fato, dificilmente é encontrado algum estudo empírico revelando todas as etapas do processo de formação até a aplicação do modelo. A maior razão para isso é a necessidade de sigilo, visto que boas e sofisticadas técnicas trazem vantagem competitivas e, portanto, as instituições que as utilizam procuram não divulgá-las.

*Mestre em Modelagem Matemática e Computacional (CFEDT/MG). Coordenador de Riscos do Banco Intermedium S/A. E-mail: marcosdutra@dppg.cefetmg.br

** PhD In Statistics - University of Warwick. Professora do Centro Federal de Educação Tecnológica de Minas Gerais. E-mail: elenice@dppg.cefetmg.br

Apesar dos problemas de violações da análise discriminante, sua aplicação em modelos de *credit scoring* é bastante usual. Assim, faz-se necessário amenizar ou resolver os problemas de violações das premissas da análise discriminante, caso contrário, pode-se ter um modelo ineficiente.

O presente trabalho justifica-se por tratar de um tema de fundamental importância para as empresas que trabalham com crédito para pessoas físicas, sendo proposta a aplicação de técnicas estatísticas sofisticadas que viabilizem a aplicação da análise discriminante, mesmo quando as pressuposições necessárias para correta discriminação dos dados são violadas. Serão propostas três estratégias para melhor discriminação dos dados. Na estratégia 01 será proposta a transformação de Box-Cox (1964), para obtenção de uma distribuição normal aproximada. Na estratégia 02 será proposta a técnica denominada influência local de Cook (1986), para detectar observações influentes. E finalmente, na estratégia 03 será proposta a análise discriminante quadrática.

Este artigo, inicialmente será apresentado uma revisão sobre Análise Discriminante de Fisher e Quadrática. Nas seções 5 e 6 serão descritos o teste de normalidade multivariada de Mardia (1970, 1974, 1975) e a transformação de Box-Cox (1964). Na seção 7 será apresentada a 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 discriminantes 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. Nesse caso, é possível construir s combinações lineares, $s \leq \min(g - 1, p)$, chamadas de funções discriminantes lineares de Fisher definidas por:

$$\hat{e}'_j W \hat{e}_j = 1,$$

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

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

Após a construção das funções discriminantes, para cada elemento amostral com vetor de observações x_j , ter-se-á um vetor com seus escores nessas funções, isto é, $Y'_j = [\hat{e}'_1 x_j, \hat{e}'_2 x_j, \dots, \hat{e}'_s x_j]$. 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 \hat{Y}_i para todo $i = 1, 2, 3, \dots, g$, sendo o correspondente elemento amostral j , classificado na população, cuja distância é a menor.

A regra de classificação consiste em calcular a distância Euclidiana entre os vetores e para todo sendo o correspondente elemento amostral, classificado na população, cuja distância é a menor.

3 FUNÇÃO DISCRIMINANTE QUADRÁTICA

Quando as matrizes de covariâncias Σ_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:

- μ_i = vetor médio da população Π_i .
- Σ_i = matriz de covariâncias da população Π_i .
- p_i = probabilidade a priori de a observação pertencer à população Π_i .
- x = vetor aleatório

A regra de classificação consiste em alocar x em Π_k se:

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

4 COMPARANDO MATRIZES DE COVARIÂNCIA

Como as suposições de igualdade das matrizes de covariância podem ser criticadas, é natural considerar um teste da taxa de verossimilhança para H_0 .

Assim, é obtida a equação para a taxa de verossimilhança,

$$l = \frac{|\hat{\Sigma}|^{-n/2}}{|\hat{\Sigma}_1|^{-n_1/2} |\hat{\Sigma}_2|^{-n_2/2} \dots |\hat{\Sigma}_I|^{-n_I/2}} \quad (6)$$

onde

$$n_i \hat{\Sigma}_i = \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(y_{ij} - \bar{y}_i)' \text{ e } n = \sum_i n_i.$$

Quando H_0 é verdadeiro e n é grande, $-2 \log l$ é aproximadamente distribuído com $\chi^2_{v_I}$, onde: $v_I = \frac{1}{2} d(d+1)(I-1)$. Contudo, uma melhor aproximação qui-quadrado pode ser obtida por uma desprezível modificação de l . Esta modificação, na qual $f_i (= n_i - 1)$, os graus de liberdade associados com Q_i , estabelecido n_p e no qual conduz para um teste não tendencioso, isto é:

$$M = \left(\prod_{i=1}^I |S_i|^{f_i/2} \right) / |S|^{f/2} = \left(\prod_{i=1}^I |Q_i|^{f_i/2} f_i^{df/2} \right) / \left(\sum_{i=1}^I Q_i \right)^{f/2} \prod_{i=1}^I f_i^{df_i/2}, \quad (7)$$

Box (1949) 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)

Os dados de uma população normal multivariada não possuem assimetria ou curtose significativa. Assim, a hipótese de assimetria e curtose não significantes é compatível com a suposição de que os dados foram retirados de uma população normal multivariada. Considere uma amostra aleatória $X = (x_1, x_2, \dots, x_n)$ de tamanho N de uma distribuição p -variáveis e \bar{x} sendo a média de X . A distância Mahalanobis dos casos x_i e \bar{x} pode ser calculada, como:

$$r_i^2 = (x_i - \bar{x})' \Sigma_{p \times p}^{-1} (x_i - \bar{x})$$

onde:

$\Sigma_{p \times p}$ é a matriz de covariância amostral e $i=1, \dots, N$.

O valor r_{ij} , também conhecido como ângulo Mahalanobis entre os vetores $x_i - \bar{x}$ e $x_j - \bar{x}$, pode ser calculado, como:

$$r_{ij} = (x_i - \bar{x})' \Sigma_{p \times p}^{-1} (x_j - \bar{x}) \quad 9$$

Baseado sobre essas duas *r-medidas*, Mardia (1970) define uma medida de assimetria e curtose multivariadas, respectivamente, como:

$$b_{1p} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N r_{ij}^3 \quad 10$$

$$b_{2p} = \frac{1}{N} \sum_{i=1}^N r_i^4 \quad 11$$

Mardia (1970) apresenta que dado a normalidade multivariada tem-se assintoticamente:

$$A = Nb_{1p}/6 \quad 12$$

segue uma distribuição χ^2_f com:

$$f = p(p+1)(p+2)/6 \quad e,$$

$$B = \frac{b_{2p} - p(p+2)}{[8p(p+2)/N]^{1/2}} \quad 13$$

segue distribuição 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, aumentando a aplicabilidade e o uso de técnicas, baseadas sobre as suposições de normalidade.

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 inde-

pendentes. Exemplos notáveis dessa 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:

$$x^{(\lambda)} = \begin{cases} \frac{x^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(x), & \lambda = 0 \text{ e } x > 0 \end{cases} \quad 14$$

que simultaneamente satisfaz todas as três suposições. O coeficiente λ pode ser estimado pelo método da máxima verossimilhança.

7 INFLUÊNCIA LOCAL

Dado um conjunto de observações, seja $l(\beta)$ o logaritmo da função de verossimilhança correspondente ao modelo postulado, sendo que (β) 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 Ω 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 ω expressa um esquema de pesos, existindo um ponto ω_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 ω varia em Ω . 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 ω_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 em ω_0 , que são as intersecções de com planos, contendo o vetor normal com seu plano tangente em ω_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 por meio do gráfico de $LD(\omega_0 + ah)$ contra $a \in \mathbb{R}$.

A curvatura normal da linha projetada, denotada por C_h , é definida como sendo a curvatura de $(a, LD(\omega_0 + ah))$ em $a = 0$, em que $\alpha(a) = \omega_0 + ah$. Denomina-se C_h curvatura normal da superfície $\alpha(\omega)$ em ω_0 e na direção unitária h .

Cook (1986) mostra que a curvatura normal na direção h pode ser expressa da seguinte forma:

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

Sendo que $\ddot{F} = \Delta^T [I(\hat{\beta})]^{-1} \Delta$, $I(\hat{\beta})$ é a matriz de informação observada sob o modelo postulado, e Δ 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 $\omega = \omega_0$.

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 \ddot{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 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 para comparações dos vetores médias populacionais, 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 pode ser convertido a um valor de F .

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 atraso no pagamento da prestação ou que obtiveram atrasos de até 60 dias. O segundo grupo é chamado "ruim", composto por clientes que obtiveram atrasos superiores há 180 dias.

A amostra é composta por 904 contratos de clientes que tomaram empréstimos em 02/01/2005 a 31/12/2006. Para cada um desses contratos foram obtidas 11 variáveis, as quais foram denotadas por x_1, x_2, \dots, x_{11} .

Aplicando o teste de Mardia (1970,1974,1975) no SPSS 16 por meio da macro de DeCarlo (1997), obtiveram-se as medidas de assimetria e curtose, respectivamente, por:

$$b_{1,p} = 50,68 \quad \text{e} \quad b_{2,p} = 176,41$$

De (14) e (15), com $d=11$ e $f=904$, obtém-se:

$A = 1351,57$ o qual é significativo quando comparado com χ^2_{165} , e

$B = 12,49$ o qual é significativo quando comparado com $N(0,1)$.

Assim, conclui-se que os dados apresentam um afastamento notório da normalidade multivariada. Nesse 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. Os resultados são apresentados na tabela 1.

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

		Grupo	Grupo Previsto		Total
			1	2	
Original	Contagem	1	371,00	67,00	438,00
		2	48,00	418,00	466,00
		%	84,70	15,30	100,00
			10,30	89,70	100,00

87,3% 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	
Original	Contagem	1	350,00	88,00	438,00
		2	62,00	404,00	466,00
		%	79,91	20,09	100,00
			13,30	86,70	100,00

83,4% 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 83,4 % para 87,4%.

O próximo passo 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 $C_{\max} = 4,93$ 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 correspondem às observações influentes.

Na figura 1, é apresentado o gráfico de influência, implementado no software R,

observados que os valores de I_{max} acima de 0.006 se destacam em relação aos demais. Contudo, é observado ainda que os valores de I_{max} abaixo de 0.0018 também se destacam. Nesse caso, serão feitas duas análises:

- consideram-se os valores de acima de 0.006 como pontos influentes, sendo encontrados 67 pontos influentes;
- consideram-se os valores de acima de 0.006 e abaixo de 0.0018 como pontos influentes, sendo encontrados 112 pontos influentes.

Gráfico Influencia Local

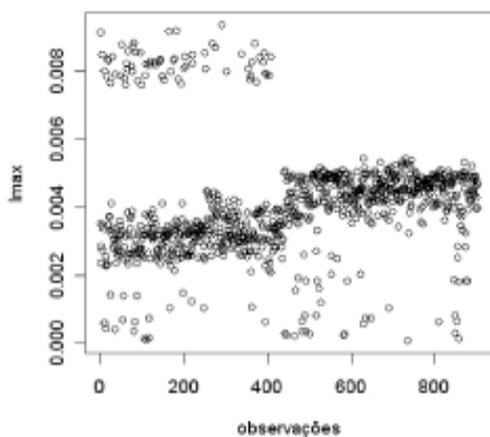


Figura 1: Gráfico de Influência – Caso variáveis explanatórias

Para reanálise dos dados são retirados os 67 e os 112 possíveis pontos influentes, respectivamente.

Após a retirada das 67 e das 112 observações influentes, é importante constatar se houve melhoria na discriminação dos dados. Os resultados são apresentados na tabela 3 e 4.

Tabela 3: Resultado da Função Discriminante de Fisher – retirados 67 pontos influentes

	Grupo	Contagem	Grupo Previsto		Total
			1	2	
Original	1	371,00	-	371,00	
	2	48,00	418,00	466,00	
%	1	100,00	-	100,00	
	2	10,30	89,70	100,00	

94,3% do grupo original classificado corretamente.

Tabela 4: Resultado da Função Discriminante de Fisher – retirados 112 pontos influentes

	Grupo	Contagem	Grupo Previsto		Total
			1	2	
Original	1	351,00	-	351,00	
	2	23,00	418,00	441,00	
%	1	100,00	-	100,00	
	2	5,22	94,78	100,00	

97,1% do grupo original classificado corretamente.

Observando as tabelas 3 e 4, verifica-se uma melhoria significativa quando aplicado a metodologia de influência local de Cook (1986), de 87,3% para 94,3% e de 87,3% para 97,1%, respectivamente.

Diante dos resultados apresentados nas tabelas 3 e 4, 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, nesse caso, a função discriminante quadrática será mais indicada. A seguir é proposta a estratégia 3, caso o teste de Box (1979) rejeite a premissa de igualdade das matrizes de covariâncias.

9.3 ESTRATÉGIA 3

Aplicado o teste de Box e convertido a um valor F, obtiveram-se os valores de 26,13 e 24,16 para os dois procedimentos, respectivamente. Como o p-valor nos dois casos são 0,00, menor que o nível de significância igual a 0,05, é rejeitada a hipótese das matrizes de covariância intra-classe serem iguais nos dois modelos.

Para contornar esse problema a Função Discriminante Quadrática será a mais indicada. Os resultados estão presentes na tabela 5 e 6.

Tabela 5: Resultado da Função Discriminante Quadrática – retirados 67 pontos

	Grupo	Contagem	Grupo Previsto		Total
			1	2	
Original	1	364,00	7,00	371,00	
	2	37,00	429,00	466,00	
%	1	98,11	1,89	100,00	
	2	7,94	92,06	100,00	

94,7% do grupo original classificado corretamente.

Tabela 6: Resultado da Função Discriminante Quadrática – retirados 112 pontos

Grupo		Grupo Previsto		Total
		1	2	
Original	Contagem	1	2	
		350,00	1,00	351,00
		1,00	440,00	441,00
%		1	2	
		99,72	0,28	100,00
		0,23	99,77	100,00

99,7% do grupo original classificado corretamente.

Aplicado a função discriminante quadrática verifica-se uma melhoria de 94,3%, para 94,7% e 97,1% para 99,7%, conforme tabelas 3 e 5; 4 e 6.

O passo seguinte é verificar a adequação do modelo de discriminação. Nesse caso, será validado apenas o modelo que 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 296,78. O *p*-valor calculado apresentou ser menor que o nível de significância $\alpha = 0,05$. Nesse 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.

10 DISCUSSÃO

A conclusão que pode ser feita a partir de todos os resultados observados é que o uso combinado da metodologia proposta, neste trabalho, foi capaz de melhorar a classificação dos clientes em 16,30%.

O resultado de 99,70% indica 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. Diante desse resultado, é possível afirmar que – apesar da análise discriminante ser baseada em diversas premissas que geralmente são violadas quando a mesma é aplicada na construção de modelos de *credit scoring* – é possível obter um excelente modelo de *credit scoring*, utilizando a análise discriminante. Contudo, faz-se necessário o uso combinado das técnicas apresentadas neste trabalho. Caso contrário, os resultados obtidos podem distanciar consideravelmente da realidade.

REFERÊNCIAS

- BOX, G. E. P. e Cox, D. R., An analysis of transformations. **Journal of the Royal Statistical Society**, v. 26, p. 211-252, 1964.
- _____. A general distribution theory for a class of likelihood criteria, **Biometrika**, v. 36, p. 317-346, 1949.
- CAOQUETTE, J. B., Altman, E. I. e Narayanan, P. **Gestão do Risco de Crédito**. 1. ed. Qualitymark, Rio de Janeiro, 1999.
- COOK, R. R. Assessment of local influence (with discussion), **Journal of the Royal Statistical Society**, v. 48, p 133-169, 1986.
- DECARLO, L. T. On the Meaning and Use of Kurtosis, **Psychological Methods**, v. 3, p. 292-307, 1997.
- FISHER, R. A. The use of multiple measurement in taxonomic problems. **Annals of Eugenics**, v. 7, p. 179-188, 1936.
- MARDIA, K.V. Measures of multivariate skewness e kurtosis with applications, **Biometrika**, v. 57, p. 519-530, 1970.
- _____. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, **Sankhyā B**, v. 35, p. 115-128, 1974.
- MARDIA, K.V. Assessment of multinormality and therobustness of Hotelling's T^2 test, **Applied Statistics**, v. 24, p. 163-171, 1975.
- SEBER, G. A. F. **Multivariate Observations**. 1 ed. John Wiley & Sons, Canada, 1984.
- Verbeke, G. e Molenberghs, G. **Linear Mixed Models for Longitudinal Data**, 1 ed. Springer, New York, 2000.