

A Análise Estatística Multivariada na Previsão de Insolvência de Empresas .

Salvatore Benito Virgillito¹

Rubens Famá²

RESUMO

A Administração do Risco de Crédito serve-se cada vez mais das técnicas Estatísticas de Análise Multivariada Discriminante, com a finalidade de definir parâmetros para a tomada de decisão gerencial.

Descrevem-se neste estudo suas fases, desde a organização dos dados, a identificação das técnicas de estatística necessárias na estrutura de construção de um modelo específico para a análise de risco e finalmente, utilizando-se deste modelo, testar a sua validade verificando a significância estatística do grau de acerto na discriminação das empresas.

O enfoque principal do estudo não é especificamente o índice de acerto nem a simplicidade do modelo de previsão, mas apesar da alta variabilidade da população estudada conseguiram-se resultados satisfatórios. O modelo final é constituído por quatro funções discriminantes (1995 a 1998) que atribuíram ano a ano um valor Z para cada empresa possibilitando a sua caracterização como possuidoras ou não de risco de insolvência.

Palavras Chaves: Análise Multivariada Discriminante, Risco de Crédito, Insolvência, Decisão Gerencial.

ABSTRACT

The Risk of Credit , increasingly utilizes Statistical techniques of Multivariate Discriminant Analysis aiming to define parameters for managerial decision.

This study describes its phases, since the data organization, identification of the necessary statistical techniques while constructing a model for risk analysis and finally, using this model to test the statistical significance and the degree of correct discrimination among the analysed business companies.

¹ Mestre em Administração e Planejamento pela PUC-SP ; Professor Titular no Centro Universitário Assunção UNIFAI; Gerente de Serviços ao cliente da LUFTHANSA A.G.
e-mail: salvatore.virgillito@dlh.de

² Doutor em Administracao pela FEA-USP; Professor Titular na Pós Graduação da FEA PUC SP em Administração e Finanças.
e-mail: rfama@usp.br

The principal scope of this study is not the degree of correct discrimination nor the simplicity of the discrimination model, however, satisfactory results were achieved although the studied population was presenting a high degree of variance.

The final model is settled with four discriminant equations (1995 to 1998) which assigned a Z value to each company in each year describing thus its potential for bankruptcy.

Key Words: Multivariate Data Analysis, Credit Risk, Bankruptcy, Management Decision.

PROBLEMÁTICA E OBJETIVO.

Partindo-se do pressuposto que a Administração Financeira utiliza-se das técnicas de análise tradicionais para comparar o desempenho de empresas no presente com o desempenho das mesmas no passado, e que a estatística pode através de técnicas matemáticas, avaliar a possibilidade de um evento verificar-se ou não, surgem algumas questões.

Poderá a Estatística adicionar algum benefício através de suas técnicas na identificação e até mesmo na prevenção do risco de insolvência de empresas brasileiras?

Quais seriam essas técnicas e em que seqüência aplicá-las?

Altman (1968: e 1977:) utilizou-se da Análise Multivariada para o estudo de risco de insolvência de empresas nos Estados Unidos e posteriormente em 1999 publicou no Brasil seu estudo sobre esse risco em empresas brasileiras.

Estudos desenvolvidos por Famá et al (1998), demostram a eficácia de tais modelos e evidenciam a necessidade de rigor na escolha do método estatístico a ser utilizado.

Desta forma o foco deste trabalho estará na descrição da metodologia estatística, na seqüência de aplicação de suas técnicas com o objetivo de estruturar a construção de modelos análise de risco de crédito e insolvência, utilizando um programa estatístico (NCSS) para computadores de pequeno porte.

METODOLOGIA DO ESTUDO

A construção do modelo de Análise Multivariada no presente estudo, tomará por base a técnica denominada Regressão Logística e a análise *Stepwise* ou *step-by-step*. Determinadas as funções discriminantes, discutir-se-á a sua significância estatística e testar-se-á seu grau de acerto na população de empresas.

Os dados fornecidos pela SERASA provém de empresas de diversos setores da economia nacional do período 1995 a 1998 e constituem-se por 20 índices extraídos de balanços de 57 empresas boas (NÃO FALIDAS legenda Não = 0), e 57 empresas ruins (FALIDAS ou em concordata, legenda Sim = 1), num total de 9120 variáveis.

Os testes estatísticos foram efetuados com a ajuda do software estatístico NCSS Number Cruncher Statistical Software (2000).

ESCOLHA DA TÉCNICA ESTATÍSTICA: ANÁLISE MULTIVARIADA OU REGRESSÃO LOGÍSTICA?

As técnicas estatísticas básicas, analisam variáveis quantitativas e qualitativas de forma unidimensional, enquanto a Análise Multivariada analisa-as de forma multidimensional e podem relacionar variáveis quantitativas com qualitativas explicando suas relações.

Anderson et al. (1998: 244) define a Análise Discriminante Multivariada como “um problema de escolha das variáveis independentes (métricas) que pertençam ou não a um grupo de variáveis dependentes (não métricas)”.

Johnson et al. (1998: 377 e 427) define a Regressão Logística também como “uma metodologia estatística que prediz se uma ou mais variáveis independentes pertencem a um determinado grupo de variáveis dependentes”.

Ambas são técnicas estatísticas apropriadas quando a variável dependente é categórica (empresas falidas e não falidas) e a variável independente é métrica (índices financeiros).

Existem, entretanto, outras técnicas de Análise Discriminante a saber: Regressão Múltipla, Análise Discriminante Múltipla, Análise Multivariada da Variância (MANOVA), Correlação Canônica, Análise de Grupos (Cluster Analysis).

UMA ABORDAGEM ESTRUTURADA NA CONSTRUÇÃO DE MODELOS MULTIVARIADOS.

A escolha do método estatístico deriva do objetivo e dos dados de que se dispõe para a análise. Anderson et al (1998: 20 e 21) sugere uma abordagem que serve como guia na determinação da técnica estatística mais apropriada para cada tipo de estudo.

Seguindo-se o quadro “Seleção da Técnica Estatística de Análise Multivariada”, deve-se considerar em primeiro lugar o tipo de relação entre as variáveis a estudar. A dependência é do tipo simples; os índices financeiros das empresas são as variáveis independentes e o objetivo é determinar sua pertinência a um dos dois grupos de empresas (boas ou falidas). A variável dependente não é métrica mas qualitativa. Dai a

chega-se a indicação de dois métodos aplicáveis a este tipo de estudo: a Análise Discriminante Múltipla, ou a Regressão Logística.

Ambas as técnicas são muito utilizadas pois satisfazem as premissas básicas da maioria das amostras dos objetos de estudo.

No caso da Análise Discriminante Múltipla por exemplo, a premissa é que os dois grupos tenham a mesma matriz de covariância, ou seja, que as distribuições de frequência dos dois grupos de variáveis tenham comportamento similar à curva normal.

No caso da Regressão Logística, deve-se tomar o cuidado para que dentro de cada um dos grupos não haja variáveis altamente correlacionadas, ou seja, onde a variação de uma das variáveis possa ser totalmente explicada a partir da variação de outra variável.

No estudo realizado por Famá et al (1998) a escolha recaiu sobre a Regressão Logística, pois este método não requer que os dados tenham fortes correlações “a priori”.

Altman (1968), utiliza-se da metodologia da Análise Discriminante Multivariada, para a elaboração de um modelo genérico construído a partir de amostras de empresas de várias indústrias dos Estados Unidos e do Brasil. Inicia com a escolha e análise tradicional dos índices de empresas, introduz o conceito de Análise Discriminante Múltipla como base para a obtenção do modelo final, desenvolve e testa o seu modelo, representando os resultados de forma gráfica e analisa os resultados obtidos.

ETAPAS PARA A CONSTRUÇÃO DO MODELO DE ANÁLISE DE INSOLVÊNCIA.

Ao iniciar-se a construção de modelos de análise multivariada, deve-se levar em conta as seguintes etapas:

- Objetivo da pesquisa

Avaliar diferenças entre os grupos pela observação da classificação atual.

- Desenhar o modelo

Considerar o tamanho da população e selecionar as variáveis.

- Verificar as premissas básicas

Normalidade das variáveis independentes (T-test).

- Estimar as funções discriminantes

Método Stepwise determinando o grau de acerto do modelo.

- Interpretar as funções encontradas

Quantas funções e seu grau de significância (F-test)

- Validação do Modelo

Testar uma amostra secundária e analisar os resultados

Anderson et al. sugere (figura 1 pg. 21) uma abordagem para a escolha da técnica estatística para a construção de modelos baseada no tipo de variáveis analisadas.

PREPARANDO A CONSTRUÇÃO DO MODELO E VERIFICANDO PRÉ-REQUISITOS.

Procederam-se a diversos ajustes nesta população como sugerido por Altman (1968).

No primeiro teste “*stepwise*” para obtenção das variáveis com maior grau de discriminação e maior probabilidade de constituírem o modelo final, utilizou-se a população inteira.

Para estudar a possibilidade da existência da mesma matriz de covariância (por ser este o pressuposto básico da Análise Discriminante Multivariada), optou-se pelo cálculo do momento de covariância de Pearson e em seguida, pela verificação e validação os resultados com a ajuda do software estatístico NCSS (2000).

A covariância da amostra é dada pela medida da associação linear entre as variáveis observadas e define-se como sendo a média do produto dos desvios de suas respectivas médias:

$$S_{12} = \frac{1}{n} \sum_{j=1}^n (X_{j1} - \bar{X}_1)(X_{j2} - \bar{X}_2) \text{ e por analogia pode-se escrever que a covariância de}$$

uma amostra de p variáveis é dada por :

$$S_{ik} = \frac{1}{n} \sum_{j=1}^n (X_{ji} - \bar{X}_i)(X_{jk} - \bar{X}_k) \text{ onde } i=1,2,\dots,p \text{ e } k=1,2,\dots,p$$

O Coeficiente de correlação de Pearson, por sua vez, mede a associação linear entre duas variáveis e é dada por:

$$R_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}}\sqrt{S_{kk}}} = \frac{\sum_{j=1}^n (X_{ji} - \bar{X}_i)(X_{jk} - \bar{X}_k)}{\sqrt{\sum_{j=1}^n (X_{ji} - \bar{X}_i)^2} \sqrt{\sum_{j=1}^n (X_{jk} - \bar{X}_k)^2}}$$

O coeficiente de correlação da amostra R_{ik} pode também ser interpretado como sendo o coeficiente de covariância da amostra (Johnson R.e Wichern DW. 1998: 9).

Para verificar se nossa população ou alguma parte dela tem a mesma matriz de covariância, extraímos uma amostra aleatória de 4 empresas e dois índices (variáveis), e através da Estatística Descritiva encontramos a Matriz da Médias, a Matriz de

Variância e Covariância e finalmente a Matriz de Correlação (Coeficiente de Momento de Pearson).

Então as Médias entre as variáveis de Endividamento e de Exigibilidades da amostra acima são:

$$\bar{X}_1 = \frac{1}{4} \sum_{j=1}^4 X_{j1} = \frac{1}{4}(518+124+36+55) = 183,25 \quad \bar{X}_2 = \frac{1}{4} \sum_{j=1}^4 X_{j2} = \frac{1}{4}(68+86+100+82) = 84$$

daí podemos escrever a matriz das médias que será: $\begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} = \begin{bmatrix} 183,25 \\ 84 \end{bmatrix}$

A variância da amostra é dada por:

$$S_{11} = \frac{1}{4} \sum_{j=1}^4 (X_{j1} - \bar{X}_1)^2 = \\ = \frac{1}{4} [(518-183,25)^2 + (124-183,25)^2 + (36-183,25)^2 + (55-183,25)^2] = 38.424,68$$

$$S_{22} = \frac{1}{4} \sum_{j=1}^4 (X_{j2} - \bar{X}_2)^2 = \frac{1}{4} [(68-84)^2 + (86-84)^2 + (100-84)^2 + (82-84)^2] = 130$$

$$S_{12} = \frac{1}{4} \sum_{j=1}^4 (X_{j1} - \bar{X}_1)(X_{j2} - \bar{X}_2) = \\ = \frac{1}{4} [(518-183,25)(68-84) + (124-183,25)(86-84) + (36-183,25)(100-84) + (55-183,25)(82-84)] = -1883,50$$

Como $S_{12} = S_{21}$ então, $S_n = \begin{bmatrix} 38424,68 & -1834,50 \\ -1834,50 & 130 \end{bmatrix}$

O momento de correlação de Pearson será então:

$$R_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}} \sqrt{S_{kk}}} = \frac{S_{12}}{\sqrt{S_{11}} \sqrt{S_{22}}} = \frac{-1834,50}{\sqrt{38424,68} \sqrt{130}} = -0,84273$$

Por demonstrar uma alta correlação negativa (próxima a -1), há indícios de que a população total de variáveis em questão, não terá a mesma matriz de covariância, pois outras variáveis poderão apresentar correlações positivas, como demonstra o exame gráfico das correlações de outras variáveis do ano de 1995 (figura 2 pg.22).

Analisaram-se as correlações existentes ano a ano entre as duas populações de empresas com a ajuda da matriz de gráficos de correlações do NCSS 2000 das quais exemplificamos a seguir os gráficos de correlação de 4 variáveis de empresas falidas e não-falidas numa matriz de 4×4 do ano de 1995, onde pode-se correlações positivas e negativas ao mesmo tempo e portanto a ausência de normalidade da distribuição das variáveis.

Johnson e Wichern (1998: 199) descrevem que normalidade da distribuição de freqüências de uma população é manifestada por dois indicadores: o primeiro quando pelo menos a metade dos coeficientes de correlação for menor ou igual a 0,50; o segundo quando todas as correlações indicadas tiverem comportamento próximo a uma linha reta com declive o que acontece entre uma variável e si própria, como ocorre na diagonal da matriz, de outra forma não ocorre. Da mesma forma, como já indicava o momento de correlação de Pearson, $-0,84$ está muito afastada de 0,5, o que pode ser interpretada como uma boa indicação para a utilização da Regressão Logística em detrimento da Análise Discriminante Multivariada.

Desejando-se ainda um maior rigor metodológico para verificar a normalidade da população, pode-se optar pelo T-test. Este teste baseia-se no cálculo de Mann-Whitney U test e Kolmogorov-Smirnov para populações contendo duas variáveis (colunas) ou uma variável sendo indexada por um segundo grupo de variáveis como no nosso caso.

O resultado do T-test efetuado com o NCSS 4 variáveis escolhidas Exigibilidades, Liquidez Seca, Rentabilidade do Patrimônio Líquido e Participação de Financiamentos no Ativo Circulante, observou-se entre 184 variáveis de empresas não falidas e 153 variáveis de empresas falidas, a existência de “rejeição à premissa de normalidade” da distribuição (reject normality) para ambas as populações, a das Empresas não-falidas = 0, e a população das Empresas falidas = 1.

O resultado do T-test para a variável Exigibilidades encontra-se no quadro1 pg.26. Notar a rejeição à premissa de normalidade “Reject normality”.

Desta forma, fica ainda mais evidenciado a não normalidade das populações .Por conseguinte esta constatação é a justificativa da opção, neste estudo, pelo método da Regressão Logística ao invés do método de Análise Discriminante Multivariada ADM.

FUNÇÃO DISCRIMINANTE.

A “discriminação” entre dois ou mais grupos definidos “a priori” é obtida estabelecendo-se os “pesos de cada variável que maximize a variância entre os grupos e minimize a variância dentro de cada grupo de variáveis independentes (Matarazzo 1995: 230 a 238).

A função discriminante segundo Johnson e Wichern (1998: 378) e Anerson et al (1998: 241) portanto assume a seguinte forma:

$$Z_{jk} = a + W_1 X_{1k} + W_2 X_{2k} + \dots + W_n X_{nk} \text{ onde,}$$

Z_{jk} = Score Z da função discriminante j, para o objeto k

a = Ponto de intercepto

W_i = Peso de discriminação para a variável independente i

X_{ik} = Variável independente i do objeto (grupo) k

Existindo dois grupos de empresas (figura 3 pg.22) , rentáveis (A) e não rentáveis (B); sejam V_1 e V_2 suas observações (índices); sejam as elipses A e B seus universos; seja Z o eixo determinado pela sua função discriminante que por sua vez é constituída por índices. Se traçarmos uma linha reta através da área de intercessão das duas elipses e projetarmos esta linha num novo eixo Z , segundo Anderson et al (1998: 250) pode-se dizer “que a área de sobreposição entre as duas distribuições univariadas A’ e B’ (representada pela área sombreada) será a menor dentre todas as outras retas possíveis de serem traçadas através da área de sobreposição das duas elipses” (figura 3 pg.22).

O significado geométrico portanto da área de sobreposição projetada no eixo Z é o da discriminação entre os dois grupos, o dos índices das empresas rentáveis e não rentáveis. Quanto menor a área de sobreposição , menor será o número de empresas rentáveis classificadas como não rentáveis e vice versa e por conseqüência, menor a probabilidade de conceder crédito a uma empresa considerada de risco.

CONSTRUÇÃO DO MODELO.

A técnica stepwise é aplicada para a análise do acréscimo da habilidade discriminante acarretada pela inclusão de mais uma variável ao grupo de variáveis já testadas. Através da análise da variância pela inclusão de variáveis passo a passo, verifica-se a possibilidade de construir-se diversos modelos com até 17 variáveis para cada ano

selecionado (1995 a 1998). Pela observação do acréscimo da variância que ocorre quando inclui-se mais uma variável ao modelo, verificamos que, para todos os anos, o número ideal de variáveis não deverá ser superior a 11 pois a partir da inclusão da 12ª variável, não ocorre um aumento significativo no grau de variância. Entretanto nada impede que seja utilizado o modelo com 17 variáveis.

O programa NCSS atribui automaticamente nomes de A até Z (vide descrição dos índices à pg.28) para as variáveis e selecionou as indicadas de A até Q. A= Endividamento, B= Composição das Exigibilidades, C= Imobilização do PL, D= Liquidez Geral, E= Liquidez Corrente, F= Liquidez Seca, G= Prazo Médio de Recebimento de Vendas, H= Prazo Médio de Pagamento de Compras, I= Ciclo Financeiro, J= Rentabilidade do Patrimônio Líquido Médio, K= Rentabilidade das Vendas, L= Giro do Ativo, M= Rentabilidade do Ativo, N= Participação de Financiamentos no Ativo, O= Endividamento Oneroso, P= Participação de Financiamentos no Endividamento, Q= Participação de Financiamento no Ativo Circulante.

TESTE DE SIGNIFICÂNCIA (F-test).

A significância estatística é o resultado desejado de qualquer estudo. O F-test é um teste estatístico de probabilidade de que a função discriminante neste caso, resulte como tendo “significância estatística”. Entretanto devido à sua complexidade o F-test é, na maioria das vezes, omitido nas pesquisas.

Este teste pode também ser utilizado de forma a estimar quais variáveis devem tomar parte do modelo para depois formá-lo através do método de Regressão Logística. A significância estatística pode ser verificada através das medidas lambda de Wilks ,

Hotellings-Lawey e o critério de Pillai e o método de Roy. Todos os métodos estão disponíveis na maioria dos softwares estatísticos. O NCSS, no teste MANOVA estima os valores de F para todos os testes constantes na figura 4 pg 23.

DETERMINAÇÃO DO VALOR CRÍTICO DE Z.

Entretanto faz-se necessário definir o valor de Z crítico, também chamado “valor de corte”, com o qual cada empresa deve ser comparada para determinar sua pertinência a este ou aquele grupo, minimizando assim os erros de classificação.

Para grupos com o mesmo tamanho, Anderson (1998) sugere que: $Z_{CE} = \frac{Z_A + Z_B}{2}$ onde

, Z_{CE} = Valor crítico de Z para grupos de mesmo tamanho

Z_A = Mediana do Grupo A

Z_B = Mediana do grupo B

Relacionam-se na tabela 4 os valores médios de Z para empresas falidas e não-falidas, de 1995 a 1998, com o respectivo Z crítico.

O valor de Z é na verdade o índice de correlação entre todas as variáveis analisadas no ano em referência, representa o desempenho da empresa sendo assim o referencial para a discriminação entre as empresas.

MATRIZ DE CLASSIFICAÇÃO.

Altman (1968: 601), discute ainda que se deve testar a significância do percentual das classificações corretas.

Para tal, Anderson et al (1998: 267 e 268), sugere que seja utilizado uma Matriz de classificação, a qual foi testada sobre o índice de acerto da equação discriminante obtida para o ano de 1995 ilustrados na tabela 4 pg. 27.

A porcentagem total de classificação correta é dada por $\left[\frac{(41+43)}{114} \right] \times 100 = 73,68\%$

Note-se que com o índice de acerto (73,68%) para 1995 com a amostra de validação, é ligeiramente maior que o índice encontrado no primeiro teste (feito sem ajustes das altas variâncias inicialmente encontradas na população), que foi de 69,70%.

Segundo Anderson et al (1998: 270), deve-se utilizar a Matriz de Classificação para verificação da significância do índice de acerto e compará-lo com a chance matemática de 50% quando simplesmente o observador jogando uma moeda honesta ao ar, teria chances iguais de classificar uma empresa como Falida ou Não Falida pela obtenção de cara ou coroa.

O índice obtido deve segundo Anderson ser ao menos $\frac{1}{4}$ (um quarto) maior do que a chance matemática para dois grupos que é de:

$$C_{prob} = (N_A)^2 + (1 - N_B)^2 \text{ onde:}$$

N_A = Proporção Grupo A em relação ao tamanho total dos Grupos

N_B = Proporção Grupo B em relação ao tamanho total dos Grupos

desta forma teremos que, $C_{prob} = (0,5)^2 + (1 - 0,5)^2 = 50\%$

Conclui-se que com os dados do presente estudo não seria significativa obter-se um índice de acerto abaixo de: $50\% \times 1,25 = 62,5\%$.

TESTE “T” PARA A AMOSTRA SECUNDÁRIA.

Da mesma forma que feito por Altman (1968: 601) e sugerido por Anderson et al (

1998: 270), testou-se a significância do índice de acerto: $t = \frac{p - 0,5}{\sqrt{\frac{0,5 \times (1 - 0,5)}{N}}}$ onde:

p = porcentagem classificada corretamente

N = tamanho da amostra,

portanto, $t = \frac{0,7368 - 0,5}{\sqrt{\frac{0,5 \times (1 - 0,5)}{114}}} = 5,05$, desta forma obteve-se a significância de

5,05 referente ao índice de acerto de 73,68% para a função discriminante do ano de 1995.

Anderson et al (1998: 270) sugere ainda outro teste de significância para verificar-se a validade do índice de acerto da função discriminante que denomina de Estatística “Q”.

Através deste teste, verifica-se a significância do índice de acerto de Altman, sendo que

o valor encontrado para “Q” $Q = \frac{[N - (nK)]^2}{N(K - 1)}$ onde:

N = Tamanho total do grupo

n = total das observações classificadas corretamente

K = número de grupos

que no caso da função discriminante para 1995 e o tamanho total do grupo resulta em:

$$Q = \frac{[114 - (84 \times 2)]^2}{114(2-1)} = 7,89 \quad , \text{ por ser } Q > t \quad , \quad (7,89 > 5,05) \text{ verifica-se então}$$

empiricamente que existe significância estatística no índice de discriminação da função.

Os índices de discriminação obtidos após a aplicação das funções discriminantes foram :

69,7% de acerto para 1995, 69,79% para 1996, 72,5% para 1997 e 76,67% para 1998.

Os gráficos obtidos na pagina 24 e 25, foram extraídos diretamente do NCSS. Note-se o valor de Z crítico representado pela linha horizontal em negrito acima da qual estariam as empresas falidas representadas pelo triângulo vermelho e abaixo as não falidas representadas pelo círculo amarelo.

Traçada a linha paralela ao eixo das Empresas no valor de Z crítico (0,2 para 1995, 0,16 para 1996, 0,3 para 1997 e 0,36 para 1998) a cada ano, obtém-se a visualização das classificações corretas e das classificações incorretas. À medida em que se aproximava a data de corte 1998, o grau de acerto do modelo aumentava diminuindo a área de sobreposição como mostram as figuras dos gráficos de Z das empresas de 1995 a 1998.

Na figura 6, note-se que, em 1995, os valores encontrados para Z, que reflete as características das empresas, dificultariam a classificação correta. Pode-se atribuir o baixo índice de acerto (69,70%) à duas componentes: a primeira delas, como já dito anteriormente, por tratarem-se de empresas de ramos de atividades diferentes e magnitudes diferentes de ativos, e a segunda a antecedência ao ano de corte 1998 quando estas teriam então uma situação financeira melhor caracterizada.

Note-se na figura 7 que o valor de Z crítico em 1996 foi de 0,16 e a melhora no grau de

acerto pela melhor definição na classificação de empresas como as de número 64,65 e 66 que em 1995 permaneciam numa área muito próxima ao valor de Z crítico sendo consideradas de risco naquele ano.

Em 1997 (figura 8), o quadro fica ainda mais claro para outras empresas que ainda permaneciam em situação de indefinição quanto à sua correta classificação.

Note-se que em 1997 a empresa de número 64 já não existia mais, verificando-se portanto a observação que esta empresa possuía, já em 1996, risco de não conseguir continuar suas operações.

Em 1998, o índice de acerto define-se em 76,67% sendo que a área de sobreposição ou de incerteza, nas proximidades do eixo Z crítico apresentou uma menor concentração.

Note-se na figura 9, um pequeno número de empresas boas sendo classificadas como de risco pela diversidade das indústrias envolvidas na análise e, pela proximidade dos seus respectivos valores de Z, ao valor de Z crítico (0,36).

A utilização de quatro funções discriminantes e a determinação do valor de Z crítico estabelecem um parâmetro adequado para a previsão de insolvência. A aproximação do valor de Z da empresa e sua manutenção por demonstrações de resultados financeiros em seqüência deverá alertar o analista a respeito deste risco.

CONSIDERAÇÕES FINAIS.

As funções discriminantes encontradas são perenes no tempo? Certamente não.

Elisabetsky (1976: 184) aponta para esta dificuldade quando sugere a implantação de seu modelo no Banco Comercial: “(...) um problema que surge na implantação do

modelo é do caráter único apresentado pela equação: a equação Z obtida a partir de dados do ano A só pode ser usada com índices em moeda dessa data(...); (...) *ocorre que o banco só poderá obter os balanços do ano A (para compor as amostras e desenvolver a equação do ano A) em meados do anos A+1, até esta data estará utilizando a equação do ano A-1. Mais ainda, as empresas que falirem durante o ano A, já não publicarão Balanços nesse ano(...).*

Note-se na tabela 5 pg. 27, o que aconteceria com o índice de acerto se continuássemos a aplicar a função obtida para 1995 aos anos sucessivos; Uma considerável perda em pontos percentuais do poder discriminante que aproximaria de forma arriscada as decisões do analista à simples chance matemática (50%) de conceder crédito pelo método “cara e coroa”. A partir desta constatação, evidencia-se o risco de conceder crédito a partir de modelos discriminantes compostos por uma única função (para referência, ilustram-se na página 29 as quatro funções discriminantes obtidas), e ainda, a necessidade de atualização constante da função e finalmente o rigor na escolha da metodologia para a construção de modelos de previsão de insolvência de empresas.

BIBLIOGRAFIA

ALTMAN, Edward. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Backruptcy. *Journal of Finance*, Boston, v.23, p. 586 - 609. 1968.

ALTMAN, Edward; CAOUILLE, J.B; NARANAYAN, P. *Gestão do Risco de Crédito*. Rio de Janeiro: Qualitymark – Serasa, 1999.

ASSAF NETO, Alexandre. Estrutura e Análise de Balanços: um enfoque econômico-financeiro. São Paulo: Atlas, 2001.

ANDERSON, Rolph; BLACK, William C.; TATHAM, Ronald; HAIR Jr, Joseph. Multivariate Data Analysis. New Jersey: Prentice Hall, 1998.

BEAVER, William. Financial Ratios as Predictors of Failure. Journal of Accounting Research, Chicago, Supplement of Selected Studies, p. 77-111. 1966.

BREALEY, Richard; MYERS, Stewart. Principles of Corporate Finance. 5.ed. Chapter 25 Analyzing Financial Performance. New York: McGraw-Hill, 1986.

ELISABETSKY, Roberto. Um Modelo Matemático para Decisões de Crédito no Banco Comercial. São Paulo, 1976. Dissertação (Mestrado) - Instituto Politécnico da USP.

FAMÁ, Rubens; BRUNI, A; MURRAY, A. Modelos Brasileiros Preditivos de Risco de Crédito: um estudo exploratório atual sobre suas eficácias. TEMA, São Paulo: Faculdades Teresa Martin, n. 32, jan. / jun. 1998.

GITMAN, Lawrence J. Princípios de Administração Financeira. São Paulo: Harbra, 1997.

HINTZE, Jerry. NCSS 2000 Number Cruncher Statistical Software, Kaysville, Uta.

JOHNSON, Richard; WICHERN, Dean.W. Applied Multivariate Statisical Analysis, 4.ed. New Jersey: Prentice Hall, 1998.

MATARAZZO, Dante. Análise Financeira de Balanços. São Paulo: Atlas, 1995.

MATIAS, Alberto. Contribuição às Técnicas de Análise Financeira: um modelo para a concessão de crédito. São Paulo, 1978. (Trabalho apresentado ao Departamento de Administração da FEA-USP).

PAULA LEITE, Helio de; HOPP, João Carlos. O Mito da Liquidez. Revista de Administração de Empresas, São Paulo, out./dez. 1989.

ROSS, Stephen; WESTERFIELD, Randolph; JAFFE, Jeffrey. Administração Financeira. São Paulo: Atlas, 1995.

SICSÚ, Abraham. Análise Discriminante. São Paulo, 1975. (Mestrado em Matemática) – Universidade de São Paulo, São Paulo.

ANEXOS

FIGURAS

Figura 1 - Seleção da Técnica Estatística de Análise Multivariada. Fonte Anderson et al. (1995: 21-22)

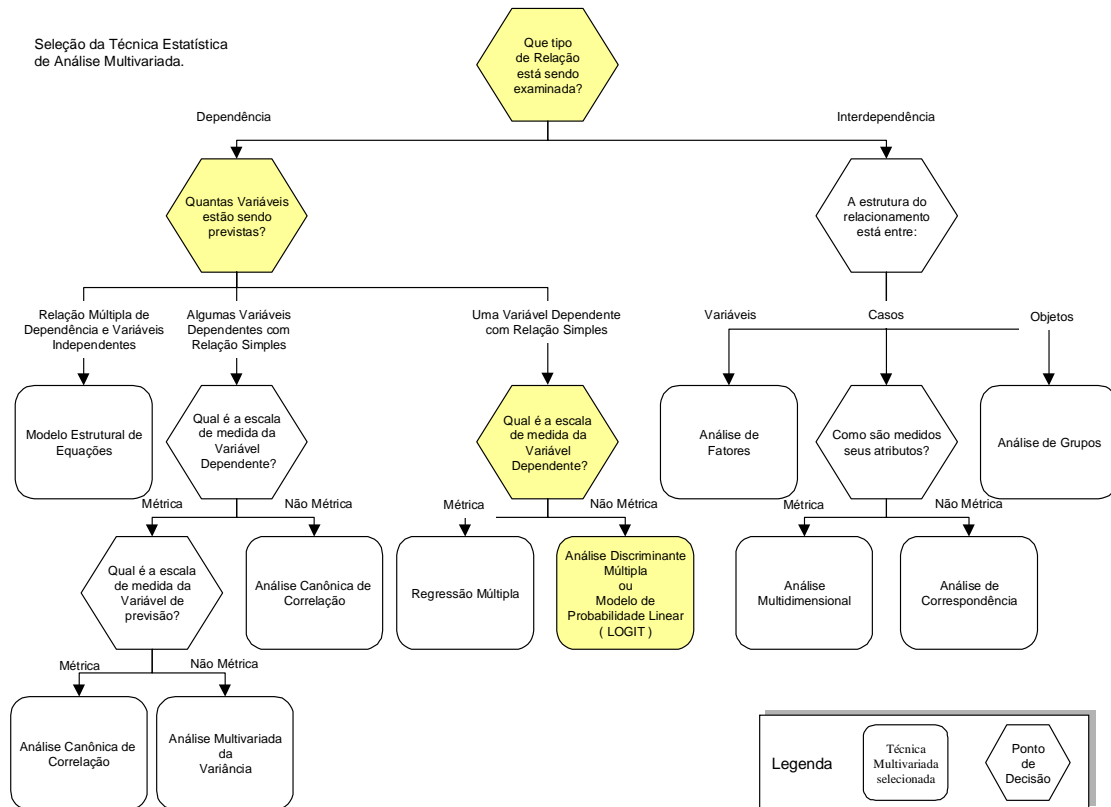


Figura 2 Estudo Gráfico da Correlações de 4 índices do ano de 1995.

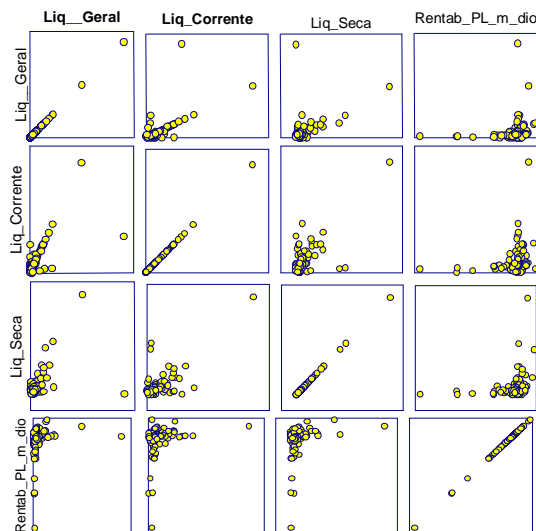


Figura 3 - Representação Gráfica de Discriminação entre dois grupos. Fonte Anderson et al. (1995: 251).

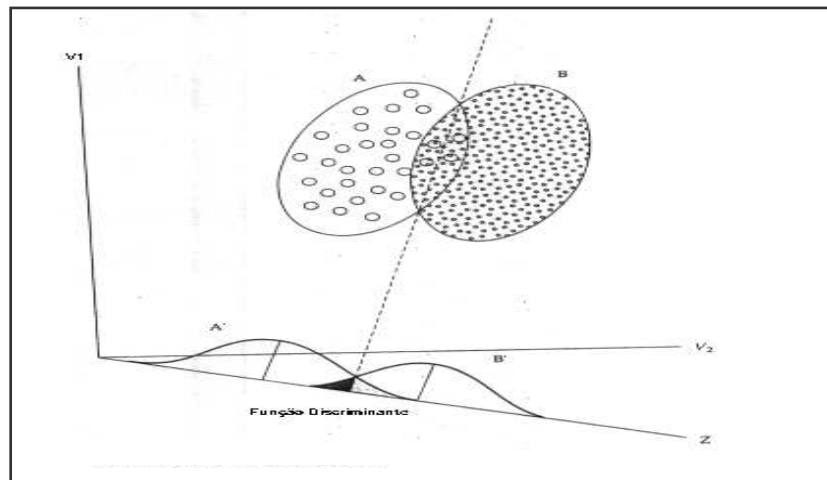


Figura 4 - Teste de Análise Multivariada da Variância 1995. Fonte: teste MANOVA NCSS 2000 com as variáveis do ano 1995.

MANOVA Report							
Page/Date/Time	1 14/04/01 18:40:27						
Database	C:\Meus documentos\P.U.C\Bal_1995.S0						
Response	Endividament,Exigibilidades,Imob_PL,Liq_Geral,Liq_Corrente,Liq_Seca,PM_Rec_Vendas,PM_Pag_Cpompras,Ciclo_Financeiro,Rentab_PL_m_dio,Rentab_Vendas,Giro_Ativo,Rentab_Ativo,Particip_Financ_Ativo,Endiv_Oneroso,Partic_Financ_Endividamento,Partic_Financ_Ativo_Circulante						
Expected Mean Squares Section							
Source	DF	Term Fixed?	Denominator Term	Expected Square			
A: Fal____Conc_	1	Yes	S(A)	S+sA			
S(A)	97	No		S			
Note: Expected Mean Squares are for the balanced cell-frequency case.							
MANOVA Tests Section							
Term(DF)	Test Statistic	Test Value	DF1	DF2	F-Ratio	Prob Level	Decision (0.05)
A(1):Fal____Conc_							
Wilks' Lambda		0.746525	17	81	1.62	0.078596	Accept
Hottelling-Lawley Trace		0.339541	17	81	1.62	0.078596	Accept
Pillai's Trace		0.253475	17	81	1.62	0.078596	Accept
Roy's Largest Root		0.339541	17	81	1.62	0.078596	Accept
Endividament	682179.746918		1	97	1.83	0.179353	Accept
Exigibilidades	49.884173		1	97	0.11	0.741698	Accept
Imob_PL	60423.112337		1	97	0.39	0.534557	Accept
Liq_Geral	12.097280		1	97	2.55	0.113500	Accept
Liq_Corrente	17.890190		1	97	3.94	0.049939	Reject
Liq_Seca	5.599837		1	97	2.39	0.125363	Accept
PM_Rec_Vendas	93.529746		1	97	0.09	0.760939	Accept
PM_Pag_Cpompras	5613.748700		1	97	1.32	0.252921	Accept
Ciclo_Financeiro	3324.771390		1	97	0.96	0.330435	Accept
Rentab_PL_m_dio	8209.203209		1	97	6.83	0.010376	Reject
Rentab_Vendas	350.231766		1	97	2.34	0.129120	Accept
Giro_Ativo	0.120274		1	97	0.06	0.811937	Accept
Rentab_Ativo	2666.062426		1	97	10.89	0.001352	Reject
Particip_Financ_Ativo	2124.706477		1	97	9.07	0.003318	Reject
Endiv_Oneroso	126308.021390		1	97	3.29	0.072905	Accept
Partic_Financ_Endividamento	3292.522876		1	97	5.61	0.019834	Reject
Partic_Financ_Ativo_Circulante	7946.315508		1	97	13.80	0.000341	Reject

Figura 5- Visualização gráfica de Z crítico para grupos de tamanhos iguais. Fonte ANDERSON et al. (1998: 266).

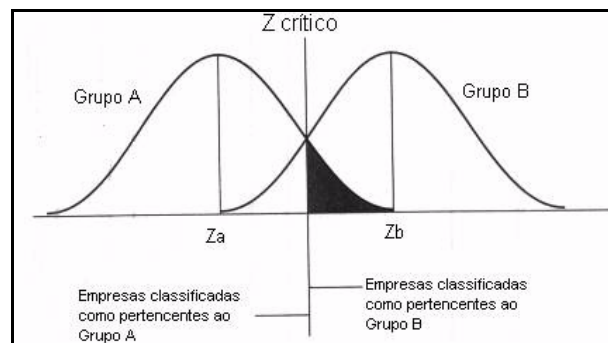


Figura 6- Discriminação gráfica das empresas pelo valor de Z crítico em 1995.

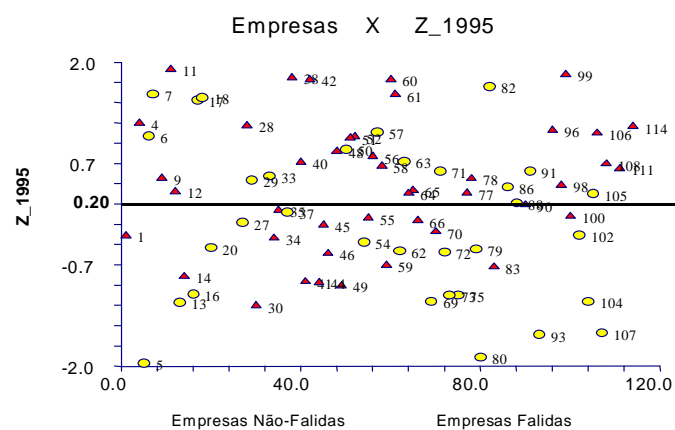


Figura 7- Discriminação gráfica das empresas pelo valor de Z crítico em 1996.

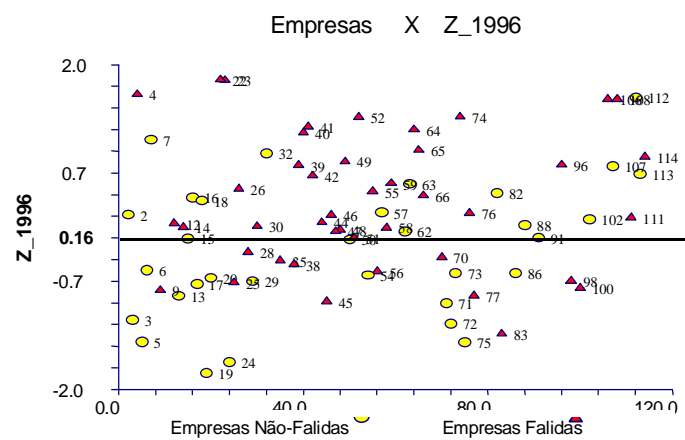


Figura 8 - Discriminação gráfica das empresas pelo valor de Z crítico em 1997.

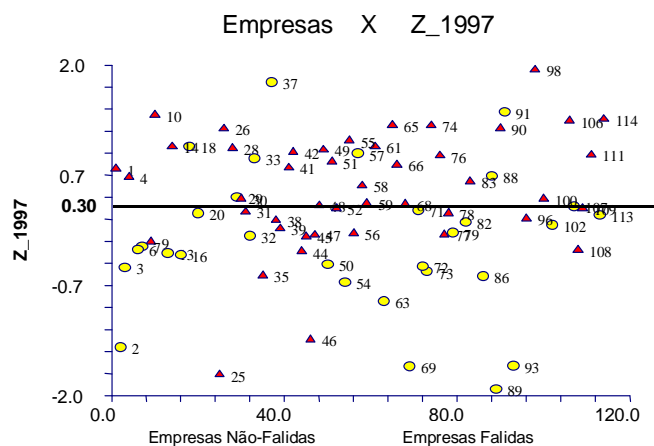
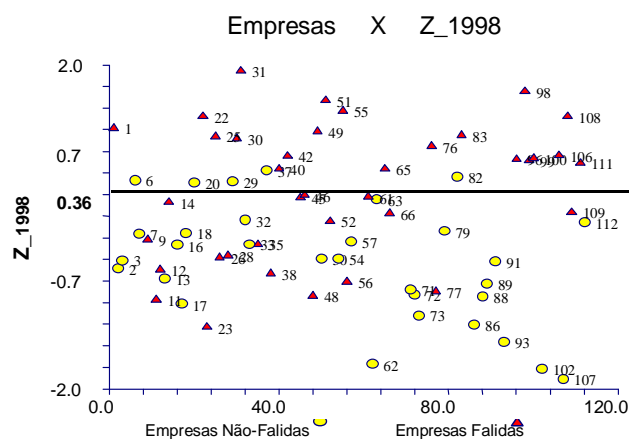


Figura 9 - Discriminação gráfica das empresas pelo valor de Z crítico em 1998.



QUADROS

Quadro 1 - Teste T para verificação da normalidade da distribuição. Variável Exigibilidades.

Estadística Descritiva

Variável	Cont.	Média	Desvio Padrão	Standard Error	95% LCL of Mean	95% UCL of Mean
Fal___Conc_=0	184	83.26087	21.79531	1.606771	80.09069	86.43105
Fal___Conc_=1	153	82.80392	18.88785	1.526992	79.78706	85.82079

Note: T-alpha (Fal___Conc_=0) = 1.9730, T-alpha (Fal___Conc_=1) = 1.9757

Teste de Premissa

Teste	Valor	Probabilidade	Decisão (5%)
Skewness Normality (Fal___Conc_=0)	-7.2671	0.000000	Reject normality
Kurtosis Normality (Fal___Conc_=0)	4.2728	0.000019	Reject normality
Omnibus Normality (Fal___Conc_=0)	71.0680	0.000000	Reject normality
Skewness Normality (Fal___Conc_=1)	-5.3726	0.000000	Reject normality
Kurtosis Normality (Fal___Conc_=1)	2.8801	0.003976	Reject normality
Omnibus Normality (Fal___Conc_=1)	37.1601	0.000000	Reject normality

TABELAS

Tabela 1 - Cálculo da Covariância entre o Índice de Endividamento e Exigibilidades

Empresas n.º	Endividamento	Exigibilidades
2	$X_{11} = 518.00$	$X_{21} = 68.00$
4	$X_{12} = 124.00$	$X_{22} = 86.00$
5	$X_{13} = 36.00$	$X_{23} = 100.00$
7	$X_{14} = 55.00$	$X_{24} = 82.00$

Tabela 2 - Escolha das Variáveis do Modelo.

Tamanho	Variância	Erro Médio	Maslow	Variáveis do Modelo
1	0.124514	0.4724185	-0.007309	Q
2	0.167635	0.4630303	-2.686061	MQ
3	0.177772	0.4626179	-1.785961	HMQ
4	0.191000	0.4613158	-1.221290	DEMQ
5	0.202262	0.4605499	-0.443238	DEHMQ
6	0.212509	0.4600627	0.444987	ADEJMQ
7	0.223437	0.4593628	1.259274	ACDIJMQ
8	0.234124	0.4587183	2.099665	ACDIJMOQ
9	0.240650	0.4593185	3.391541	ACDHIJMOQ
10	0.246474	0.4601463	4.759692	ACDEHJKMOQ
11	0.249327	0.4619063	6.450126	ACDEFHJKMOQ
12	0.251497	0.4639122	8.214707	ABCDEFHJKMOQ
13	0.252495	0.4663219	10.106402	ABCDEFHJKMOQ
14	0.252833	0.4689832	12.069707	ABCDEFHIJKMNOQ
15	0.253250	0.4716683	14.024454	ABCDEFHIJKMNOPQ
16	0.253467	0.4744665	16.000862	ABCDEFHGIJKMNOPQ
17	0.253475	0.4773838	18.000000	ABCDEFHGIJKLMNOPQ

Tabela 3 - Valores de Z crítico de 1995 a 1998.

Ano	Valor de Z Não Falidas	Valor de Z Falidas	Valor Crítico de Z
1995	-0.44702	0.634076	0,20
1996	-0.56538	0.637917	0,16
1997	-0.43995	0.7897327	0,30
1998	-0.80259	1.124111	0,36

Tabela 4 - Matriz da Classificação percentual da Discriminação entre Empresas.

	Grupo Atual	Grupo Predito	Total	% Classificação Correta
	Não Falidas = 0	Falidas = 1		
Não Falidas = 0	41	16	57	71,93%
Falidas = 1	14	43	57	75,44%
Total	55	59	114	73,68%

Tabela 5 - Aplicação do Modelo de 1995 para as Empresas operantes de 1996 a 1998

Classificações Corretas						
Modelo do ano	Aplicação do Modelo de 1995 para os anos de 1996 a 1998					
Ano	%acerto	%acerto	Corretas	Erradas	Total	Dif. Pontos %
1995	69,70%	69,70%	79	35	114	0,00
1996	69,79%	61,40%	70	44	114	- 8,39
1997	72,53%	59,65%	68	46	114	- 12,88
1998	76,67%	54,38%	62	52	114	- 22,29

As funções discriminantes obtidas

Z_1995 = 1,533957 + 0,002118663***B** - 0,03193196***C** - 0,002692934***D** + 2,768697* **E** - 2,604233* **F** + 0,1528927***G** + 0,004968031***H** + 0,005367486* **I** + 0,0008229673***J** + 0,05480665* **K** + 0,03976471* **L** + 0,07612021* **M** - 0,1802415* **N** - 0,02074066***O** + 0,002899373* **P** - 0,001983856***Q** + 0,04353809***R**

Z_1996 = 4,799199 + 0,002994071***B** - 0,05614699* **C** - 0,006709214* **D** + 3,785284***E** - 0,864722* **F** - 0,6974394* **G** + 0,01666887* **H** - 0,002145869* **I** + 0,001005979* **J** - 0,002139698* **K** + 0,04909386* **L** + 0,130444* **M** - 0,02149716* **N** - 0,03361323* **O** + 0,003236873* **P** + 0,01757129* **Q** + 0,02488947* **R**

Z_1997 = 1,037193 + 0,002671816***B** - 0,01437995***C** - 0,006875481***D** - 0,08000931* **E** - 0,05968606* **F** - 2,418277* **G** + 0,0267394* **H** + 0,00620077* **I** - 0,004064261* **J** - 0,004320121* **K** + 0,01871338* **L** + 0,2540574* **M** - 0,02604646* **N** - 0,08695164* **O** + 0,003515554* **P** + 0,05713542* **Q** + 0,003349302* **R**

Z_1998 = 3,274066 + 0,00111997* **B** - 0,0452389* **C** - 0,002793578* **D** + 0,9960911* **E** - 1,272084* **F** - 2,474569* **G** + 0,02661187* **H** + 0,003359224* **I** + 0,003859246* **J** - 0,02224002* **K** + 0,01769009* **L** + 0,3665507* **M** + 0,03701901* **N** + 0,04053191* **O** - 0,008583037* **P** - 0,004793949* **Q** + 0,01122298***R**

Descrição dos Índices

Índice	Código Índice
Endividamento	B
Exigibilidades	C
Imobilização do P.L	D
Liquidez Geral	E
Liquidez Corrente	F
Liquidez Seca	G
Prazo Médio recebimento Vendas	H
Prazo Médio Pagamento Compras	I
Ciclo Financeiro	J
Rentabilidade P.L	K
Rentabilidade Vendas	L
Giro do Ativo	M
Rentabilidade do Ativo	N
Partic. Financiamentos no Ativo Total	O
Endividamento Oneroso	P
Partic. Financiamentos no Endividamento	Q
Partic. Financiamentos no Ativo Circulante	R