DISCRIMINAÇÃO DE DUAS POPULAÇÕES MULTIVARIADAS COM BASE EM DESIGUALDADE DE MATRIZES DE DISPERSÕES

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pela Sra.
DORIS ALBINA GÓMEZ TICERÁN
e aprovada pela Comissão Julgadora.

Campinas 12 de Outubro de 1988.

prof. Dr. _____
Orientador

Dissertacão apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do Título de Mestre em ESTATÍSTICA.

DISCRIMINAÇÃO DE DUAS POPULAÇÕES MULTIVARIADAS COM BASE EM DESIGUALDADE DE MATRIZES DE DISPERSÕES

ALUNA : DORIS ALBINA GÓMEZ VICERÁN

ORIENTADOR : PROF. DR. JOSÉ ANTÔNIO CORDEIRO

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E CIÊNCIA DA COMPUTAÇÃO
(IMECC)
UNIVERSIDADE ESTADUAL DE CAMPINAS - (UNICAMP)

CAMPINAS - SÃO PAULO

BRASIL

A mi hijo YAKOV

con todo el amor,
que la distancia
le privó.

AGRADECIMENTOS

Ao prof. Dr. José Antônio Cordeiro, pela orientação e incentivo recebido durante a realização do presente trabalho.

Ao prof. Dr. Armando Infante, pelas valiosas sugestões que tanto nos auxiliaram.

Aos professores, colegas e amigos da U.N.M. de San Marcos de Lima Perú, que me alentaram na iniciação dos meus estudos de mestrado.

Aos meus amigos e colegas do Mestrado, José Luis Llanos, e José Ramos, pelos valiosos auxílios na parte de computação.

Ao meu amigo Mario Tarumoto pelo precioso apoio na datilografia e na correção do português.

Ao meu esposo Mario pela compreensão e a força recebida, que apesar da distância nunca me faltou, e sem as quais seria impossível a culminação dos meus estudos.

Aos meus pais Luis e Guillermina pela paciência e carinho que me dedicaram.

A Eugenia, Myriam, Emma, Rosangela, pela amizade e solidariedade nos meus momentos de saudade. E a todos os meus amigos e colegas que direta ou indiretamente me apoiaram.

Ao CNPq e UNICAMP que financiaram o meu programa de pós-graduação.

INDICE

CAPITULO 3

O DESEMPENHO DO NOVO METODO : SIMULAÇÃO E EXEMPLO

INTRODUÇÃO

Um dos tópicos mais importantes da Estatística Multivariada, é apresentado com o título de Análise Discriminante ou Discriminação e Classificação, cujo interesse principal é alocar um indivíduo(objeto) ou um grupo deles em uma das categorias, grupos ou populações concorrentes. Também, os critérios de discriminação adotados, frequentemente, são utilizados para a redução de dimensão do problema estatístico.

A classificação consiste na identificação da categoria ou grupo ao qual pertence o novo indivíduo, levando em consideração as suas características observadas. Quando tais características são medições numéricas a designação aos grupos é chamada discriminação e a combinação das medições recebe o nome de função discriminante. Mais especificamente, na discriminação tenta-se descrever de maneira gráfica (em duas ou três dimensões) ou algebricamente mediante funções chamadas discriminantes, os aspectos diferenciais dos indivíduos ou objetos de várias populações. Quando a(s) função(ões) de discriminação são lineares o que se tem são combinações das variáveis originais que, escolhidas convenientemente, podem providenciar informação importante. Também, tais combinações simplificam a estrutura da matriz de covariâncias facilitando a interpretação dos dados.

Em muitos casos assume-se a existência de um número finito de populações a uma das quais pertence o indivíduo, sendo que cada população é caracterizada pela distribuição de probabilidade de suas medidas e uma observação é um valor amostral de uma dessas populações. Portanto, o indivíduo a ser classificado é considerado como uma observação aleatória de uma das populações.

Cabe notar que em algumas situações, as distribuições de probabilidade das populações são consideradas conhecidas de início, em outras situações a forma da distribuição de probabilidade das populações é conhecida mas os parâmetros são desconhecidos, então, tomam-se amostras dessas populações para estimar as quantidades desconhecidas.

Os métodos mais utilizados em problemas práticos de discriminação e classificação são o de Fisher (*Fisher, 1936*), o de razão de verossimilhanças e o de Bayes (*Anderson,1958*), que produzem funções de discriminação lineares (caso homoscedástico) ou quadráticas (caso heteroscedástico) nas medições do indivíduo a ser classificado.

No contexto do problema de discriminação e classificação e tomando a idéia de Flury (*Flury,1983*), o objetivo do presente trabalho é propor um novo método de discriminação baseado na comparação das matrizes de covariâncias para duas populações multivariadas de dimensão p, que conduz a uma função discriminante a partir da qual pode-se estabelecer uma regra de classificação.

O método proposto sempre fornece duas combinações lineares das componentes do indivíduo a classificar, que podem ser utilizadas em representações gráficas e redução de dimensão do problema, toda vez que as matrizes de covariâncias entre as duas populações forem diferentes.

A conclusão do presente trabalho é que, comparado com o método de classificação por quociente de verossimilhanças, no caso de populações com distribuições de probabilidade normais de dimensão p=3 o método proposto apresenta vantagems e desvantagens, e, em p=2 eles parecem ser equivalentes. As populações foram simuladas pelo método Monte Carlo.

Sugere-se que o método proposto seja usado como complementar, que podería ajudar numa melhor classificação e num melhor entendimento do problema.

O trabalho está dividido em 4 capítulos. O capítulo 1 contém os resultados de uma pesquisa bibliográfica sobre regras de classificação e as funções de discriminação em populações homoscedásticas. Além disso, apresentamos o novo critério de classificação baseado nas estruturas de covariâncias das populações concorrentes. O problema de discriminação e a nova proposta para populações normais heteroscedásticas são apresentados no capítulo 2. No capítulo 3, usando amostras de populações normais de dimensão p=2 e p=3 geradas por processo de simulação Monte Carlo e com um exemplo de aplicação, avalia-se o desempenho do método de discriminação proposto. As conclusões obtidas do estudo de simulação e do exemplo são apresentadas no capítulo 4.

CAPITULO 1

O PROBLEMA GERAL DA DISCRIMINAÇÃO

1.1 INTRODUÇÃO

Sejam Π_1 Π_2 as duas populações ou classes de objetos e $\underset{\sim}{X}'^{(g)}=(X_1^{(g)},.....,X_p^{(g)})$, g=1,2 um vetor aleatório com valores em $\mathbb{R}^p$ contendo medições dos indivíduos de cada uma das populações. Os valores observados de $\underset{\sim}{X}^{(g)}$ diferem de uma classe à outra e através de suas medidas construiremos uma regra para classificar um novo indivíduo $\underset{\sim}{x}$ de $\mathbb{R}^p$ em uma das duas populações.

Na construção de uma função de discriminação, a suposição de igualdade das matrizes de covariâncias das populações concorrentes(homoscedasticidade) é uma suposição muito forte e às vezes não satisfeita nas aplicações práticas.

Sob a suposição de homoscesdasticidade obtém-se uma função linear, no entanto, quando a suposição de homogeneidade de matrizes de covariâncias não pode ser feita(heteroscedasticidade), a função discriminante contém termo quadrático.

1.2 Regras de Boa Classificação

Em princípio, independentemente da condição de homoscedasticidade ou heteroscedasticidade das matrizes de covariâncias das duas populacões, nos problemas de classificação estamos sujeitos aos possíveis erros de classificação, isto é, alocar um indivíduo a certa população quando na realidade pertence a uma outra população. Portanto, na construção do procedimento de classificação procuramos aquele que minimiza o custo esperado de má classificação.

Podemos pensar de uma observação como um ponto no espaço de dimensão p. Dividimos tal espaço em duas regiões disjuntas R_1 e R_2 ($R_1 \cup R_2 = \mathbb{R}^p$). Se a observação $\underset{\sim}{x}$ pertence a R_1 é classificada como procedente de Π_1 e se pertence a R_2, como procedente de Π_2, ou seja:

Alocar $\underset{\sim}{x}$ em Π_g se $\underset{\sim}{x} \in R_g$ $g=1,2$ (1)

com $R_1 \cup R_2 = \mathbb{R}^p$ e $R_1 \cap R_2 = \emptyset$

As populações Π_1 e Π_2 estão caracterizadas pelas funções de densidade $f_1(\underset{\sim}{x})$ e $f_2(\underset{\sim}{x})$, respectivamente.

Em todo processo de classificação temos associado o custo, dado que, quando alocamos um objeto em Π_1 sendo que ele pertence a Π_2 ou alocamos em Π_2 dado que procede de Π_1, estamos cometendo um erro que tem um determinado custo. Tais custos podem ser medidos em qualquer unidade porque o que importa é o quociente entre eles.

Claramente, um bom procedimento de classificação é aquele que minimiza o custo de má classificação.

Nós iniciaremos agora o procedimento do problema de discriminação e classificação envolvendo duas categorias ou populações, para o qual necessitamos da notação:

$g= 1,2 \quad i=1,2 \quad j=1,2 \quad i \neq j$ (2)

$f_g(\underset{\sim}{x})$ Funções de densidade de probabilidade correspondente à população Π_g.

R_g Regiões de classificação correspondente à população Π_g.

R Denota uma regra de classificação particular.

$C(i/j)$ Custo de classificar erradamente a observação $\underset{\sim}{x}$ da população Π_j em Π_i.

$P(i/j,R)$ Probabilidade condicional de classificar uma observação de Π_j em Π_i, segundo a regra R.

$q_g=P(\underset{\sim}{X} \in \Pi_g)$ Probabilidade a priori de obter uma observação da população Π_g.

$r(j,R)$ Risco ou perda esperada quando uma observação da população Π_j é classificada em Π_i.

$$P(\underset{\sim}{x} \text{ ser classificado corretamente}) = P(g/g,R) = \int_{R_g} f_{g}(\underset{\sim}{x})dx.$$

$$P(\underset{\sim}{x} \text{ ser classificado incorretamente}) = P(i/j,R) = \int_{R_i} f_{j}(\underset{\sim}{x})dx.$$

onde, $dx = dx_1 dx_2 .. dx_p = \prod_{i=1}^{p} dx_i$

$$P(2/1,R) = 1 - P(1/1,R) \qquad P(1/2,R) = 1 - P(2/2,R) \qquad (3)$$

$$r(1,R) = C(2/1)P(2/1,R) \qquad r(2,R) = C(1/2)P(1/2,R)$$

DEFINIÇÃO 1.1 O custo esperado de má classificação (CEM) da regra R é definido como a soma dos produtos da probabilidade de uma observação pertencer a uma determinada população pela perda esperada da mesma. Ou seja:

$$CEM = q_1 r(1,R) + q_2 r(2,R)$$
$$= q_1 C(2/1)P(2/1,R) + q_2 C(1/2)P(1/2,R) q_2 \qquad (4)$$

O procedimento que minimiza (4) é chamado Procedimento de Bayes para q_1 e q_2 dados. Todo procedimento de Bayes é admissível, isto é, na classe de procedimentos R não existe outro melhor que ele.

DEFINIÇÃO 1.2 A Probabilidade total de má classificação (PTMD da regra R é definida como:

$$PTM = P(\underset{\sim}{x} \text{ ser classificado incorretamente em } \Pi_1 \text{ ou } \Pi_2)$$

$$= q_1 P(2/1, R) + q_2 P(1/2, R)$$

$$= q_1 \int_{R_2} f_1(\underset{\sim}{x}) dx + q_2 \int_{R_1} f_2(\underset{\sim}{x}) dx \qquad (5)$$

1.3 PROCEDIMENTOS DE CLASSIFICAÇÃO EM UMA DE DUAS POPULAÇÕES COM DISTRIBUIÇÕES DE PROBABILIDADES CONHECIDAS

Quando as distribuições de probabilidade são conhecidas devemos usar tal informação na construção das regras de classificação.

Vamos voltar ao problema de escolher as regiões R_1 e R_2 tais que o CEM (4) seja mínima.

Dada a regra R, de (4) temos:

$$CEM = C(2/1)P(2/1,R)q_1 + C(1/2)P(1/2,R)q_2$$

$$= q_1\, C(2/1) \int_{R_2} f_1(\underset{\sim}{x})dx \;+\; q_2\, C(1/2) \int_{R_1} f_2(\underset{\sim}{x})dx$$

$$CEM = q_1\, C(2/1) \left\{1 - \int_{R_1} f_1(\underset{\sim}{x})dx \right\} + q_2\, C(1/2) \int_{R_1} f_2(\underset{\sim}{x})dx$$

$$= q_1\, C(2/1) + q_2\, C(1/2) \int_{R_1} f_2(\underset{\sim}{x})dx - q_1\, C(2/1) \int_{R_1} f_1(\underset{\sim}{x})dx$$

$$CEM = q_1\, C(2/1) + \int_{R_1} \left[q_2\, C(1/2)\, f_2(\underset{\sim}{x}) - q_1 C(2/1)\, f_1(\underset{\sim}{x})\right]dx \qquad (6)$$

A expressão (6) é minimizada quando a região R_1 é escolhida de tal maneira que:

$\underset{\sim}{x} \in R_1$, se e só se

$$q_2\, C(1/2)\, f_2(\underset{\sim}{x}) - q_1\, C(2/1)\, f_1(\underset{\sim}{x}) \leq 0$$

ou equivalentemente:

$$\frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} \geq \frac{q_2\, C(1/2)}{q_1\, C(2/1)} \qquad (7)$$

Assim, a regra que minimiza o custo esperado de má classificação é dada pelo seguinte teorema:

TEOREMA 1.1 - Se q_1 e q_2 são as probabilidades a priori de se retirar uma observação da população Π_1 com densidade $f_1(\underset{\sim}{x})$ e da população Π_2 com densidade $f_2(\underset{\sim}{x})$, e se o custo de classificar uma observação de Π_1 como de Π_2 é $C(2/1)$ e uma observação de Π_2 como de Π_1 é $C(1/2)$, então as regiões de classificação R_1 e R_2 definidas por:

$$R_1 : \frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} \geq \frac{q_2\ C(1/2)}{q_1\ C(2/1)}$$

$$R_2 : \frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} < \frac{q_2\ C(1/2)}{q_1\ C(2/1)}$$

minimizam o custo esperado de má classificação.

Outro procedimento que conduz ao resultado (7), é aquele que aloca $\underset{\sim}{x}$ no grupo com maior probabilidade a posteriori, onde consideram-se custos $C(i/j)$ unitários.

Sejam :

$P(\Pi_1/\underset{\sim}{X})$: Probabilidade a posteriori da observação $\underset{\sim}{X}$ pertencer à população Π_1.

$P(\Pi_2/\underset{\sim}{X})$: Probabilidade a posteriori da observação $\underset{\sim}{X}$ pertencer à população Π_2.

E pela Regra de Bayes, temos:

$$P(\Pi_g / \underset{\sim}{X}) = \frac{P(\underset{\sim}{X} \in \Pi_g)\ P(\underset{\sim}{X} / \underset{\sim}{X} \in \Pi_g)}{\sum_{g=1}^{2} P(\underset{\sim}{X} \in \Pi_g)\ P(\underset{\sim}{X} / \underset{\sim}{X} \in \Pi_g)} \qquad g=1,2$$

$$= \frac{q_g\, f_g(\underset{\sim}{x})}{q_1 f_1(\underset{\sim}{x}) + q_2 f_2(\underset{\sim}{x})} \qquad (8)$$

Para um valor observado $\underset{\sim}{x}$ minimizaremos a probabilidade total de má classificação alocando-o na população que apresenta maior probabilidade a posteriori. Ou seja:

Alocar $\underset{\sim}{x}$ em Π_1 quando:

$$\frac{q_1\, f_1(\underset{\sim}{x})}{q_1\, f_1(\underset{\sim}{x}) + q_2\, f_2(\underset{\sim}{x})} \geq \frac{q_2\, f_2(\underset{\sim}{x})}{q_1\, f_1(\underset{\sim}{x}) + q_2\, f_2(\underset{\sim}{x})}$$

$$q_1\, f_1(\underset{\sim}{x}) \geq q_2\, f_2(\underset{\sim}{x})$$

caso contrário, alocar em Π_2 (9)

Apresenta-se também a regra que está baseada na máxima verossimilhança, segundo a apresentação feita por *MARDIA, KENT E BIBBY (1979)*, mediante a definição 1.4. Previamente definiremos o conceito de função de verossimilhança.

DEFINIÇÃO 1.3 Seja $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ uma amostra aleatória de tamanho n da população Π cuja função de densidade de probabilidade é $f(\underset{\sim}{x};\underset{\sim}{\theta}) = f(\underset{\sim}{x})$ onde $\underset{\sim}{\theta}$ é o vetor de parâmetros. A função de verossimilhança para toda a amostra definida por:

$$L(\underset{\sim}{\theta}, \underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n) = \prod_{i=1}^{n} f(\underset{\sim}{x}_i, \underset{\sim}{\theta}) \quad \text{é uma função do parâmetro } \underset{\sim}{\theta}.$$

DEFINIÇÃO 1.4 Sejam Π_1 e Π_2 duas populações p variadas com funções de densidade $f_1(\underset{\sim}{x})$ e $f_2(\underset{\sim}{x})$, a regra de discriminação baseada na máxima verossimilhança aloca $\underset{\sim}{x}$ para uma das populações Π_1 ou Π_2 para a qual $\underset{\sim}{x}$ tem maior função de verossimilhança, ou seja:

Alocar $\underset{\sim}{x}$ em Π_1 quando

$$L_1(\underset{\sim}{x}) \geq L_2(\underset{\sim}{x})$$

ou seja, quando

$$\frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} \geq 1 \qquad (10)$$

caso contrário, alocar em Π_2

Observamos que a regra (10) é equivalente à regra que minimiza o CEM com custos e probabilidades a priori iguais.

Em algumas situações as categorias ou populações estão especificadas de antemão, no sentido de que as distribuições de probalilidade são completamente conhecidas. Em outros casos, só a forma das distribuições de probabilidade são conhecidas e tém-se que estimar os parâmetros populacionais, para o que usamos resultados amostrais de tais populações.

Antes de fazer a apresentação da nova proposta de classificação, é conveniente fazer uma recordação da regra de discriminação linear proposta por *Fisher (1936)*, e apresentaremos também a regra de discriminação para populações normais homoscedásticas.

1.4 MÉTODO DE FISHER PARA DUAS POPULAÇÕES

1.4.1 PARAMETROS CONHECIDOS

Considere Π_1 e Π_2 duas populações em $\mathbb{R}^p$ e $X'^{(g)}=(X_1^{(g)},X_2^{(g)},\ldots,X_p^{(g)})$ um vetor aleatório de Π_g e $f_g(\underset{\sim}{x})$ funções de densidade de probabilidade associados a Π_g, $g = 1,2$.

A técnica de *Fisher (1936)* consiste em determinar a combinação linear das coordenadas de $\underset{\sim}{x}$ que maximiza o quadrado da distância entre as médias das combinações dos dois grupos relativa à sua variância.

Sejam:

$$\underset{\sim}{M}_g = E(\underset{\sim}{X}/\Pi_g) = E(\underset{\sim}{X}^{(g)}) = \begin{bmatrix} m_1^{(g)} \\ \ldots\ldots \\ m_p^{(g)} \end{bmatrix} : \text{Valor esperado da variável aleatória } \underset{\sim}{X}, \text{quando } \underset{\sim}{X} \in \Pi_g.$$

(11)

$V_g = E(\underset{\sim}{X} - \underset{\sim}{M}_g)(\underset{\sim}{X} - \underset{\sim}{M}_g)'$ matriz de covariâncias da população Π_g. $g=1,2$.

com $V_1 = V_2 = V$ positiva definida.

A regra de classificação de Fisher (*Lachenbruch,1975; Johnson e Wichern,1982*) é :

Alocar $\underset{\sim}{x}$ na população Π_1 quando:

$$\underset{\sim}{x}'V^{-1}(\underset{\sim}{M}_1 - \underset{\sim}{M}_2) - \frac{1}{2}(\underset{\sim}{M}_1 + \underset{\sim}{M}_2)'V^{-1}(\underset{\sim}{M}_1 - \underset{\sim}{M}_2) \geq 0$$

caso contrário alocar em Π_2 (12)

1.4.2 PARAMETROS DESCONHECIDOS

Sejam:

$\underset{\sim}{X}_1^{(g)}, \underset{\sim}{X}_2^{(g)}, \ldots\ldots, \underset{\sim}{X}_{n_g}^{(g)}$ amostra aleatória de Π_g

$$\underset{\sim}{X}_i^{(g)} = (x_{i1}^{(g)}, \ldots, x_{ip}^{(g)}) \qquad g=1,2 \qquad i = 1, \ldots, n_g \qquad (13)$$

$\bar{\underset{\sim}{X}}^{(g)} = \frac{1}{n_g} \sum_{i=1}^{n_g} \underset{\sim}{X}_i^{(g)}$ vetor de médias para a amostra do grupo g. É estimador não viesado de $\underset{\sim}{M}_g$.

$$S_g = \frac{1}{n_g - 1} \left\{ \sum_{i=1}^{n_g} \underset{\sim}{X}_i^{(g)} \underset{\sim}{X}_i^{(g)'} - n_g \, \bar{\underset{\sim}{X}}^{(g)} \bar{\underset{\sim}{X}}^{(g)'} \right\}$$

estimador não viesado de V_g.

$S = \frac{(n_1 - 1) S_1 + (n_2 - 1) S_2}{n_1 + n_2 - 2}$ Matriz de covariâncias amostral combinada. Estimador da matriz V.

Substituindo $\underset{\sim}{M}_1$, $\underset{\sim}{M}_2$, V por $\bar{\underset{\sim}{X}}^{(1)}$, $\bar{\underset{\sim}{X}}^{(2)}$ e S em (12), obtemos a seguinte regra de classificação amostral:

Alocar $\underset{\sim}{x}$ em Π_1 quando :

$$\underset{\sim}{x}' S^{-1} (\bar{\underset{\sim}{X}}^{(1)} - \bar{\underset{\sim}{X}}^{(2)}) - \frac{1}{2} (\bar{\underset{\sim}{X}}^{(1)} + \bar{\underset{\sim}{X}}^{(2)})' S^{-1} (\bar{\underset{\sim}{X}}^{(1)} - \bar{\underset{\sim}{X}}^{(2)}) \geq 0$$

caso contrário, alocar em Π_2 (14)

1.5 CLASSIFICAÇÃO EM DUAS POPULAÇÕES MULTIVARIADAS HOMOSCEDASTICAS

Vamos aplicar as regras apresentadas em (1.2) no caso de se terem duas populações normais multivariadas homoscedásticas, onde $\underset{\sim}{M}_g$ $(g=1,2)$ é o vetor de médias da g-ésima população e V, positiva definida, é a matriz de covariâncias comum às duas populações.

Segundo a regra de alocação que minimiza o custo esperado de má classificação, **TEOREMA 1.1**, a região de classificação em Π_1, ou seja em R_1, é o conjunto dos $\underset{\sim}{x}$ para o qual:

$$\frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} = \frac{\left(\frac{1}{2\pi}\right)^{\frac{p}{2}} |V|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_1)'V^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_1)\right\}}{\left(\frac{1}{2\pi}\right)^{\frac{p}{2}} |V|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_2)'V^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_2)\right\}}$$

$$\geq \frac{q_2\, C(1/2)}{q_1\, C(2/1)} \qquad (15)$$

Tomando logaritmo e fazendo simplificações chega-se a:

$$\underset{\sim}{x}'V^{-1}(\underset{\sim}{M}_1-\underset{\sim}{M}_2) - \frac{1}{2}(\underset{\sim}{M}_1+\underset{\sim}{M}_2)'V^{-1}(\underset{\sim}{M}_1-\underset{\sim}{M}_2) \geq \ln\left\{\frac{q_2\, C(1/2)}{q_1\, C(2/1)}\right\} \qquad (16)$$

Onde o primeiro termo do lado esquerdo da desigualdade (16) recebe o nome de *Função Linear Discriminante*.

Quando os parâmetros são desconhecidos, o que geralmente acontece nas aplicações práticas, utilizamos (13) para estimar as quantidades $\underset{\sim}{M}_1$, $\underset{\sim}{M}_2$ e V e apresentamos a seguinte regra de classificação amostral.

Alocar $\underset{\sim}{x}$ em Π_1 quando:

$$\underset{\sim}{x}'S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)}) - \frac{1}{2}(\bar{\underset{\sim}{X}}^{(1)}+\bar{\underset{\sim}{X}}^{(2)})' S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)}) \geq \ln\left[\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right]$$

caso contrário, classificar em Π_2 (17)

A expressão :

$$W=\underset{\sim}{X}'\ S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)}) - \frac{1}{2}(\bar{\underset{\sim}{X}}^{(1)}+\bar{\underset{\sim}{X}}^{(2)})'\ S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})$$

recebe o nome de *FUNÇÃO DE CLASSIFICAÇÃO AMOSTRAL DE ANDERSON.*

Anderson (1958) apresenta o seguinte teorema:

TEOREMA 1.2 - Seja:

$$W=\underset{\sim}{X}'S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})-\frac{1}{2}(\bar{\underset{\sim}{X}}^{(1)}+\bar{\underset{\sim}{X}}^{(2)})'S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})$$

com a notação em (13):

$\bar{\underset{\sim}{X}}^{(g)}$: média da amostra de $\Pi_g \sim N(\underset{\sim}{M}_g, V)$.

S_g : matriz de covariâncias amostral de Π_g.

Então a distribuição de:

$$W=\underset{\sim}{X}'S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})-\frac{1}{2}(\bar{\underset{\sim}{X}}^{(1)}+\bar{\underset{\sim}{X}}^{(2)})'S^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})$$

quando $n_1 \longrightarrow \infty$ e $n_2 \longrightarrow \infty$ é :

$N\left(\frac{1}{2}\alpha^2, \alpha^2\right)$ se $\underset{\sim}{X} \in \Pi_1$

$N\left(-\frac{1}{2}\alpha^2, \alpha^2\right)$ se $\underset{\sim}{X} \in \Pi_2$

onde $\alpha^2=(\underset{\sim}{M}_1-\underset{\sim}{M}_2)'V^{-1}(\underset{\sim}{M}_1-\underset{\sim}{M}_2)$, sendo α chamada distância de Mahalanobis.

Considerando $a^2=(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})'V^{-1}(\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)})$, *Anderson (1973)*, dá a distribuição das expressões:

$$(W-a^2/2)/a \qquad\qquad (W+a^2/2)/a$$

Onde conclui que são N(0,1).

1.6 CRITERIO DE CLASSIFICAÇÃO BASEADO NA ESTRUTURA DE COVARIANCIAS DAS POPULAÇÕES CONCORRENTES

A função de discriminação mais conhecida apresentada na secção (1.4) está baseada no pressuposto de estruturas de covariâncias iguais. No presente item, apresenta-se uma proposta baseada na estrutura de covariâncias diferentes das populações concorrentes.

O entendimento das diferenças das matrizes de covariâncias pode ser conseguido analisando-se as combinações lineares das variáveis $\underset{\sim}{X}^{(g)}$, g=1,2 com quociente extremo de variâncias, isto é, analisando aquelas combinações lineares definidas com os autovetores associados aos autovalores máximo e mínimo da matriz $V_1^{-1}V_2$ (*Flury, 1983*). Ver-se-á que a comparação das matrizes de covariâncias analisando certas combinações lineares, proporciona muito mais informações que a simples decisão de igualdade ou não das matrizes V_1 e V_2 ,(*Flury,1985*).

1.6.1 OBTENÇÃO DAS COMBINAÇÕES LINEARES VIA COMPARAÇÃO DAS MATRIZES DE COVARIANCIAS

O teste da União e Interseção introduzido por Roy (1957) é usado para à obtenção das componentes principais generalizadas (*Flury,1983*). Tais componentes são de muita utilidade para a nova proposta de discriminação, razão pela qual será descrito a seguir.

O teste de Roy está baseado na seguinte consideração:

Sejam $\underset{\sim}{X}^{(1)}$ e $\underset{\sim}{X}^{(2)}$ vetores aleatórios independentemente distribuídos com vetores de médias $\underset{\sim}{M}_1$ e $\underset{\sim}{M}_2$ e matrizes de covariâncias V_1 e V_2 respectivamente, onde as matrizes são positivas definidas.

Dado que o interesse é construir uma regra de discriminação baseada nas estruturas de covariâncias V_1 e V_2, é melhor verificar se a suposição $V_1=V_2$ é adequada. Isto pode ser feito testando a hipótese:

$$H_0: V_1 = V_2$$
$$H_1: V_1 \neq V_2 \qquad (18)$$

A solução de (18) pode-se simplificar analisando o comportamento das variâncias das combinações lineares das variáveis aleatórias originais. Portanto, se definirmos:

$$Y^{(1)} = \underset{\sim}{a}'\underset{\sim}{X}^{(1)}$$
$$Y^{(2)} = \underset{\sim}{a}'\underset{\sim}{X}^{(2)} \qquad \forall\ \underset{\sim}{a} \in \mathbb{R}^p \qquad (19)$$

então, $Y^{(g)} \sim (\underset{\sim}{a}'\underset{\sim}{M}_g\ ,\ \underset{\sim}{a}'V_g\underset{\sim}{a})$ $\qquad g = 1,2$

Para as variâncias das combinações lineares de (19) pode-se construir a seguinte hipótese:

$$H_{0a}: \underset{\sim}{a}'V_1\underset{\sim}{a} = \underset{\sim}{a}'V_2\underset{\sim}{a} \qquad \forall\ \underset{\sim}{a} \in \mathbb{R}^p \qquad (20)$$

Pelo Princípio da União e Interseção de Roy(1957) a hipótese H_o de (18) pode ser escrita como a intersecção do conjunto das hipóteses univariadas de (20), ou seja:

$$H_o: V_1 = V_2 \iff \underset{\sim}{a}'V_1\underset{\sim}{a} = \underset{\sim}{a}'V_2\underset{\sim}{a}, \ \forall \ \underset{\sim}{a} \in \mathbb{R}^p \qquad (21)$$

As matrizes V_1 e V_2 são iguais quando

$$\left[\frac{\underset{\sim}{a}'V_2\underset{\sim}{a}}{\underset{\sim}{a}'V_1\underset{\sim}{a}}\right] = 1$$

ou

$$\max_{\underset{\sim}{a} \in \mathbb{R}^p}\left[\frac{\underset{\sim}{a}'V_2\underset{\sim}{a}}{\underset{\sim}{a}'V_1\underset{\sim}{a}}\right] = \min_{\underset{\sim}{a} \in \mathbb{R}^p}\left[\frac{\underset{\sim}{a}'V_2\underset{\sim}{a}}{\underset{\sim}{a}'V_1\underset{\sim}{a}}\right] = 1 \qquad (22)$$

ou seja, temos que estudar os extremos da função H(a):

$$H(\underset{\sim}{a}) = \frac{\underset{\sim}{a}'V_2\underset{\sim}{a}}{\underset{\sim}{a}'V_1\underset{\sim}{a}} \qquad \underset{\sim}{a} \in \mathbb{R}^p \qquad (23)$$

Dado que a função $H(\underset{\sim}{a})=H(c\underset{\sim}{a})$ para todo $c \in \mathbb{R}$ $(c\neq 0)$, sem perda de generalidade podemos impor $\underset{\sim}{a}'V_1\underset{\sim}{a}=1$ e maximizar a função:

$$H_1(\underset{\sim}{a}) = \underset{\sim}{a}'V_2\underset{\sim}{a} - \lambda\,(\underset{\sim}{a}'V_1\underset{\sim}{a} - 1)$$

onde λ é multiplicador de Lagrange.

Mas, a solução do problema conduz à obtenção dos autovalores e autovetores da matriz $V_1^{-1}V_2$.

Então sem perda de generalidade, os autovalores da matriz $V_1^{-1}V_2$ podem ser representados por $\lambda_1 \geq \lambda_2 \geq \ldots\ldots \geq \lambda_p$, e os autovetores associados por $\beta_1, \beta_2, \ldots, \beta_p$.

Se λ é autovalor de $V_1^{-1}V_2$ associado ao autovetor "$\underset{\sim}{a}$", com $\underset{\sim}{a}'V_1\underset{\sim}{a}=1$, tem-se:

$$\underset{\sim}{a}'V_2\underset{\sim}{a} - \lambda\underset{\sim}{a}'V_1\underset{\sim}{a} = 0$$

$$\underset{\sim}{a}'V_2\underset{\sim}{a} = \lambda \qquad (24)$$

A equação (24) mostra que, se o vetor "$\underset{\sim}{a}$" satisfaz a $(V_2-\lambda V_1)=0$, então a razão de variâncias das combinações lineares $\underset{\sim}{a}'\underset{\sim}{X}^{(1)}$ e $\underset{\sim}{a}'\underset{\sim}{X}^{(2)}$. é "λ", logo vale:

$$(V_2 - \lambda_1 V_1)\,\beta_1 = 0$$

$$\beta_1'V_2\beta_1 - \lambda_1\beta_1'V_1\beta_1 = 0$$

$$\beta_1'V_2\beta_1 = \lambda_1 \qquad \beta_1'V_1\beta_1 = 1 \qquad (25)$$

com :

$$Y_1^{(1)} = \beta_1'\underset{\sim}{X}^{(1)} \qquad Y_1^{(2)} = \beta_1'\underset{\sim}{X}^{(2)}$$

sendo $(Y_1^{(1)}, Y_1^{(2)})$ as combinações lineares com maior quociente de variâncias.

O passo seguinte é procurar as combinações lineares que têm o maior quociente de variâncias entre todas as combinações lineares $\underset{\sim}{a}'\underset{\sim}{X}^{(1)}$ e $\underset{\sim}{a}'\underset{\sim}{X}^{(2)}$ que estão não correlacionadas com $Y_1^{(1)}$ e $Y_1^{(2)}$ respectivamente. Ou seja, deve-se maximizar $H(\underset{\sim}{a})$ tal que:

$$H(\underset{\sim}{a}) = \frac{\underset{\sim}{a}'V_2\underset{\sim}{a}}{\underset{\sim}{a}'V_1\underset{\sim}{a}} \qquad \text{razão das variâncias} \qquad \forall\ \underset{\sim}{a} \in \mathbb{R}^p$$

com as restrições:

$$\begin{aligned} &(i) \quad \underset{\sim}{a}'V_1\underset{\sim}{a} = 1 \\ &(ii) \quad \mathrm{Cov}(\underset{\sim}{a}'\underset{\sim}{X}^{(1)}, Y_1^{(1)}) = \underset{\sim}{a}'\ V_1\underset{\sim}{\beta}_1 = 0 \\ &(iii) \quad \mathrm{Cov}(\underset{\sim}{a}'\underset{\sim}{X}^{(2)}, Y_1^{(2)}) = \underset{\sim}{a}'\ V_2\ \underset{\sim}{\beta}_1 = 0 \end{aligned} \qquad (26)$$

Usando multiplicadores de Lagrange e a condição $\underset{\sim}{a}'(V_2 - \lambda_1 V_1)\underset{\sim}{\beta}_1 = 0$, só se tem que maximizar:

$$H_2(\underset{\sim}{a}) = \underset{\sim}{a}'V_2\underset{\sim}{a} - \lambda(\underset{\sim}{a}'V_1\underset{\sim}{a} - 1) - 2v_1(\underset{\sim}{a}'V_1\underset{\sim}{\beta}_1 - 0) \qquad (27)$$

onde o vetor de derivadas parciais é :

$$\frac{\partial H_2(\underset{\sim}{a})}{\partial\ \underset{\sim}{a}} = 2\ V_2\underset{\sim}{a} - 2\ \lambda\ V_1\underset{\sim}{a} - 2v_1V_1\underset{\sim}{\beta}_1 \qquad (28)$$

e o vetor "$\underset{\sim}{a}$" que maximiza (27) deve satisfazer à equação (28) igual a zero. Depois de alguns cálculos chega-se a :

$$(\ V_2 - \lambda\ V_1\ )\ \underset{\sim}{a} = 0 \qquad (29)$$

A equação (29) mostra que os novos valores " λ " e "$\underset{\sim}{a}$" procurados satisfazem à mesma equação $(V_2-\lambda V_1)=0$, que a solução anterior λ_1 e β_1. Dado que λ_1 e seu correspondente autovetor não satisfazem (ii) e (iii) de (26) , os candidatos óbvios são λ_2 e β_2 que satisfazem :

$$\begin{aligned} &(V_2-\lambda_2 V_1)\,\beta_2 = 0 \\ &\beta_2' V_1 \beta_2 = 1 \quad \beta_2' V_2 \beta_2 = \lambda_2 \\ &\beta_2' V_1 \beta_1 = 0 \quad , \quad \beta_2' V_2 \beta_1 = 0 \end{aligned} \tag{30}$$

e se tem

$$Y_2^{(1)} = \beta_2' \underset{\sim}{X}^{(1)} \qquad Y_2^{(2)} = \beta_2' \underset{\sim}{X}^{(2)}$$

Que são o segundo par de combinações lineares com maior quociente de variâncias entre todas as combinações não correlacionadas com $Y_1^{(1)}$ e $Y_1^{(2)}$.

O processo de obtenção das combinações lineares continua e pode-se assumir que os primeiros q pares de combinações lineares $(1 \le q < p)$ podem ser definidas como:

$$Y_i^{(1)} = \beta_i' \underset{\sim}{X}^{(1)} \qquad Y_i^{(2)} = \beta_i' \underset{\sim}{X}^{(2)}$$

onde β_i é o i-ésimo autovetor da matriz $V_1^{-1}V_2$ associado ao autovalor λ_i , satisfazendo a:

$$(V_2 - \lambda_i V_1)\,\beta_i = 0 \tag{31}$$

« 1 »
$$\begin{cases} \mathrm{Var}(Y_i^{(1)}) = \beta_i' V_1 \beta_i = 1 \\ \mathrm{Var}(Y_i^{(2)}) = \beta_i' V_2 \beta_i = \lambda_i \end{cases}$$

« 2 »
$$\begin{cases} \mathrm{Cov}(Y_i^{(1)}, Y_j^{(1)}) = \beta_i' V_1 \beta_j = 0 \\ \mathrm{Cov}(Y_i^{(2)}, Y_j^{(2)}) = \beta_i' V_2 \beta_j = 0 \end{cases} \qquad 1 \le i < j \le q$$

Agora desejam-se as (q+1)-ésimas combinações lineares com quociente máximo de variâncias entre todas as combinações lineares $\underset{\sim}{a}'\underset{\sim}{X}^{(1)}$ e $\underset{\sim}{a}'\underset{\sim}{X}^{(2)}$ que estão não correlacionadas com $Y_i^{(1)}$ e $Y_i^{(2)}$, $1 \leq i \leq q$, sujeito a $\underset{\sim}{a}'V_1\underset{\sim}{a}=1$, ou seja, maximizar:

$$H(\underset{\sim}{a}) = \frac{\underset{\sim}{a}'V_2\underset{\sim}{a}}{\underset{\sim}{a}'V_1\underset{\sim}{a}} \qquad \text{razão das variâncias} \qquad \forall\ \underset{\sim}{a} \in \mathbb{R}^p$$

com as restrições:

$$\begin{aligned} &i) && \underset{\sim}{a}'V_1\underset{\sim}{a} = 1 \\ &ii) && \mathrm{Cov}(\underset{\sim}{a}'\underset{\sim}{X}^{(1)}, Y_i^{(1)}) = \underset{\sim}{a}'V_1\beta_i = 0 \qquad 1 \leq i \leq q \\ &iii) && \mathrm{Cov}(\underset{\sim}{a}'\underset{\sim}{X}^{(2)}, Y_i^{(2)}) = \underset{\sim}{a}'V_2\beta_i = 0 \end{aligned} \qquad (32)$$

Pela condição $(V_2 - \lambda_i V_1)\beta_i = 0$ de (31), (iii) de (32) pode ser omitida e só se vai maximizar a função :

$$H_{q+1}(\underset{\sim}{a}) = \underset{\sim}{a}'V_2\underset{\sim}{a} - \lambda(\underset{\sim}{a}'V_1\underset{\sim}{a} - 1) - 2\sum_{i=1}^{q} v_i \underset{\sim}{a}'V_1\beta_i \qquad (33)$$

com λ e v_i $(1 \leq i \leq q)$ multiplicadores de Lagrange.

O vetor de derivadas parciais igual a zero é :

$$\frac{\partial H_{q+1}(\underset{\sim}{a})}{\partial \underset{\sim}{a}} = 2V_2\underset{\sim}{a} - 2\lambda V_1\underset{\sim}{a} - 2\sum_{i=1}^{q} v_i V_1\beta_i = 0 \qquad (34)$$

Pré-multiplicando (34) por β_j' $(1 \leq j \leq q,\ i \neq j)$ e pelas condições (ii) e (iii) de (32), os dois primeiros termos desaparecem, e por « 2 » de (31) a última soma também desaparece.

Dado que $\beta_1' V_1 \beta_1 = 1$ (por « 1 » de (31)) conclui-se que $v_j = 0$. De tal maneira que a equação (34) conduz à :

$$(V_2 - \lambda V_1)\underset{\sim}{a} = 0 \qquad (35)$$

a qual mostra que "λ" e "$\underset{\sim}{a}$" procurados satisfazem à mesma equação (29) que as soluções prévias λ_i e β_i ($1 \leq i \leq q$). Dado que os "q" maiores autovalores e seus autovetores associados não satisfazem às condições (ii) e (iii) de (32) pela suposição « 1 » de (31), os seguintes candidatos são λ_{q+1} e β_{q+1}, com os quais definimos as (q+1)-ésimas combinações lineares:

$$Y_{q+1}^{(1)} = \beta_{q+1}' \underset{\sim}{X}^{(1)} \qquad Y_{q+1}^{(2)} = \beta_{q+1}' \underset{\sim}{X}^{(2)}$$

satisfazendo a :

$$\text{« 1 »} \quad \begin{cases} \text{Var}(Y_{q+1}^{(1)}) = \beta_{q+1}' V_1 \beta_{q+1} = 1 \\ \text{Var}(Y_{q+1}^{(2)}) = \beta_{q+1}' V_2 \beta_{q+1} = \lambda_{q+1} \end{cases} \qquad (36)$$

$$\text{« 2 »} \quad \begin{cases} \text{Cov}(Y_{q+1}^{(1)}, Y_i^{(1)}) = \beta_{q+1}' V_1 \beta_i = 0 \\ \text{Cov}(Y_{q+1}^{(2)}, Y_i^{(2)}) = \beta_{q+1}' V_2 \beta_i = 0 \end{cases} \qquad 1 \leq i \leq q$$

Dado que V_1 e V_2 são matrizes simétricas definidas positivas , pode-se continuar com o procedimento e no último passo λ_p e β_p serão usados para definir as últimas combinações lineares $Y_p^{(1)}=\beta_p' X^{(1)}$ e $Y_p^{(2)}=\beta_p' X^{(2)}$ satisfazendo às mesmas condições de (31) com $1 \leq i \leq p$.

Os resultados anteriores,são resumidos no teorema apresentado a seguir (*Flury, 1983*).

TEOREMA 1.3.-Dados os vetores aleatórios $X^{(1)}$ e $X^{(2)}$ independentes, com parâmetros de locação e dispersão M_g e V_g g=1,2 respectivamente, sendo que V_1 e V_2 são positivas definidas, definem-se as combinações lineares :

$$Y_i^{(1)} = \beta_i' X^{(1)}$$

$$Y_i^{(2)} = \beta_i' X^{(2)} \qquad 1 \leq i \leq p$$

usando os autovetores $B=(\beta_1,.....,\beta_p)$ da matriz $V_1^{-1}V_2$.
Então :

i) $\quad Var(Y^{(1)}) = I_p$

$$Var(Y^{(2)}) = \Lambda \quad \text{com}$$

$$I_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_p \end{bmatrix}$$

λ_i i-ésimo autovalor de $V_1^{-1}V_2$, isto é, as combinações lineares, $Y'^{(1)} = (Y_1^{(1)},.....,Y_p^{(1)})$, $Y'^{(2)} = (Y_1^{(2)},.....,Y_p^{(2)})$ estão não correlacionadas em ambos os grupos.

ii) $(Y_i^{(1)}, Y_i^{(2)})$ é o i-ésimo par de combinações lineares com maior quociente de variâncias de todas as combinações lineares não correlacionadas com as anteriores.

As combinações lineares $(Y_i^{(1)}, Y_i^{(2)})$ $1 \leq i \leq p$ denominam-se Componentes Principais Generalizadas. (*Flury, 1983*).

Dado que a nossa atenção está no caso V_1 ser diferente de V_2, então, a solução de (22) conduz à escolha das combinações lineares $(Y_1^{(1)}, Y_1^{(2)})$ e $(Y_p^{(1)}, Y_p^{(2)})$ com quocientes de variâncias λ_1 e λ_p, respectivamente. Usaremos tais combinações lineares para propor um novo critério de discriminação.

No caso $V_1 = V_2$, as combinações lineares não são de utilidade para a construção da regra de discriminação, porque em tal situação todos os autovalores da matriz $V_1^{-1}V_2$ valem um. A mesma observação é válida no caso de proporcionalidade entre V_1 e V_2.

Nas aplicações práticas a igualdade entre V_1 e V_2 é facilmente avaliada com o maior e menor autovalores da matriz $S_1^{-1}S_2$.

1.6.2 UMA PROPOSTA DE DISCRIMINAÇÃO

Mesmo que os resultados apresentados em (1.6.1) não sejam essencialmente novos, eles são pouco conhecidos em aplicações práticas.

As combinações lineares $\underset{\sim}{Y}^{(1)}$ e $\underset{\sim}{Y}^{(2)}$ não são correlacionadas em ambos os grupos, e para compreender as diferenças entre V_1 e V_2, é mais fácil analisar as variâncias de tais combinações.

Segundo a demonstração do **TEOREMA 1.3**, toda informação a respeito das diferenças em variabilidade entre os vetores $\underset{\sim}{X}^{(1)}$ e $\underset{\sim}{X}^{(2)}$ está expressa nos autovalores, "λ_j" afastados do valor um, daí o interesse especial nas combinações lineares:

$$Y_i^{(1)} = \underset{\sim}{\beta}_i' \underset{\sim}{X}^{(1)} \qquad Y_i^{(2)} = \underset{\sim}{\beta}_i' \underset{\sim}{X}^{(2)} \qquad i=1,p \qquad (37)$$

e não nas $Y_k^{(g)}$ $g=1,2$, $k \neq i$, que não contribuam para estabelecer as diferenças entre V_1 e V_2, isto é, descartam-se todas as combinações lineares que têm quociente de variâncias perto de um.

Em muitas aplicações práticas tem-se observado que os autovalores extremos diferem marcadamente de " um ", sendo que os outros estão próximos de " um ", (*Flury, 1985*). De um modo geral, se alguns autovalores afastam-se de "um", os que mais o fazem são os extremos. Portanto, concentram-se as atenções somente nas combinações lineares com QUOCIENTES DE VARIÂNCIAS EXTREMOS, sendo as combinações:

POPULAÇÃO	QUOCIENTE DE VARIANCIAS	
	Máximo	Mínimo
Um	$Y_1^{(1)}=\beta_1'X^{(1)}$	$Y_p^{(1)}=\beta_p'X^{(1)}$
Dois	$Y_1^{(2)}=\beta_1'X^{(2)}$	$Y_p^{(2)}=\beta_p'X^{(2)}$

(38)

respectivamente.

Segundo as suposições iniciais para $X^{(1)}$ e $X^{(2)}$, os parâmetros populacionais das combinações de (38) são :

população Um:

$$E(Y_1^{(1)})= \beta_1' E(X^{(1)}) = \beta_1' M_1$$
$$E(Y_p^{(1)})= \beta_p' E(X^{(1)}) = \beta_p' M_1$$

$$Var(Y_1^{(1)})= \beta_1' V_1 \beta_1 = 1$$
$$Var(Y_p^{(1)})= \beta_p' V_1 \beta_p = 1$$

$$Cov(Y_1^{(1)}, Y_p^{(1)})= \beta_1' V_1 \beta_p = 0$$

(39)

população Dois:

$$E(Y_1^{(2)})= \beta_1' E(X^{(2)}) = \beta_1' M_2$$
$$E(Y_p^{(2)})= \beta_p' E(X^{(2)}) = \beta_p' M_2$$

$$Var(Y_1^{(2)})= \beta_1' V_2 \beta_1 = \lambda_1$$
$$Var(Y_p^{(2)})= \beta_p' V_2 \beta_p = \lambda_p$$

$$Cov(Y_1^{(2)}, Y_p^{(2)})= \beta_1' V_2 \beta_p = 0$$

Para simplificar o desenvolvimento posterior considere-se a notação:

$$\underset{\sim}{Y} = \begin{bmatrix} Y_1 \\ \\ Y_p \end{bmatrix} = \begin{bmatrix} Y_{max} \\ \\ Y_{min} \end{bmatrix} \in \mathbb{R}^2$$

$$\underset{\sim}{Y}^{(g)} = \begin{bmatrix} Y_1^{(g)} \\ \\ Y_p^{(g)} \end{bmatrix} = \begin{bmatrix} Y_{max}^{(g)} \\ \\ Y_{min}^{(g)} \end{bmatrix} \in \mathbb{R}^2$$

$$\underset{\sim}{\mu}^{(1)} = \begin{bmatrix} \underset{\sim}{\beta}_1' \underset{\sim}{M}_1 \\ \\ \underset{\sim}{\beta}_p' \underset{\sim}{M}_1 \end{bmatrix} = \begin{bmatrix} \mu_1^{(1)} \\ \\ \mu_2^{(1)} \end{bmatrix} \in \mathbb{R}^2 \qquad (40)$$

$$\underset{\sim}{\mu}^{(2)} = \begin{bmatrix} \underset{\sim}{\beta}_1' \underset{\sim}{M}_2 \\ \\ \underset{\sim}{\beta}_p \underset{\sim}{M}_2 \end{bmatrix} = \begin{bmatrix} \mu_1^{(2)} \\ \\ \mu_2^{(2)} \end{bmatrix} \in \mathbb{R}^2$$

$$\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}$$

e tem-se a situação:

duas populações bivariadas Π_1 e Π_2 com parâmetros de locação $\underset{\sim}{\mu}^{(1)}$ e $\underset{\sim}{\mu}^{(2)}$, parâmetros de dispersão Σ_1 e Σ_2 e com funções de densidade de probabilidade $f_1(\underset{\sim}{y})$ e $f_2(\underset{\sim}{y})$.

Deseja-se usar tal informação para classificar o indivíduo $\underset{\sim}{x} \in \mathbb{R}^p$. As outras suposições com respeito aos custos de má classificação e as probabilidades a priori são as mesmas que as dadas para (1.2) do capítulo 1.

As regiões de classificação que minimizam o custo esperado de má classificação, são dadas pelo

TEOREMA 1.4.- Se q_1 e q_2 são as probabilidades a priori das populações Π_1 e Π_2 com funções de densidade $f_1(\underset{\sim}{y})$ e $f_2(\underset{\sim}{y})$, e se o custo de classificar uma observação de Π_1 como de Π_2 é $C(2/1)$ e uma observação de Π_2 como de Π_1 é $C(1/2)$, então as regiões de classificação R_1 e R_2 definidas por:

$$R_1 : \frac{f_1(\underset{\sim}{y})}{f_2(\underset{\sim}{y})} \geq \frac{q_2\, C(1/2)}{q_1\, C(2/1)}$$

$$R_2 : \frac{f_1(\underset{\sim}{y})}{f_2(\underset{\sim}{y})} < \frac{q_2\, C(1/2)}{q_1\, C(2/1)}$$

minimizam o custo esperado de má classificação.

De onde obtemos a seguinte regra:

$\underset{\sim}{x}$ é alocado em Π_1 se :

$$\frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} \geq \frac{q_2\, C(1/2)}{q_1\, C(2/1)} \qquad (41)$$

caso contrário é alocado em Π_2

---

onde $\underset{\sim}{Y} \in \mathbb{R}^2$ é o vetor transformado segundo:

$$\underset{\sim}{Y} = \begin{bmatrix} \beta_1' \underset{\sim}{X} \\ \beta_p' \underset{\sim}{X} \end{bmatrix} \in \mathbb{R}^2$$

β_1 e β_p são os autovetores associados aos autovalores máximo e mínimo λ_1 e λ_p da matriz $V_1^{-1}V_2$.

CAPITULO 2

PROBLEMA DE DISCRIMINAÇÃO NO CASO NORMAL HETEROSCEDASTICO

2.1 INTRODUÇÃO

Quando o conjunto de variâncias e covariâncias de uma população normal multivariada não é o mesmo que o conjunto de variâncias e covariâncias da outra população normal multivariada, pode-se aplicar a teoria apresentada para matrizes de covariâncias iguais, só que $\ln(f_1(\underset{\sim}{x})/f_2(\underset{\sim}{x}))$ é uma função quadrática.

Neste capítulo serão apresentados, o procedimento usado até agora e a nova proposta de discriminação , considerando as populações Π_1 e Π_2 normais multivariadas com diferentes estruturas de covariâncias, isto é, $N(\underset{\sim}{M}_1, V_1)$ e $N(\underset{\sim}{M}_2, V_2)$, onde $\underset{\sim}{M}_g$ é o vetor de médias da população g, e V_g positiva definida, a correspondente matriz de covariâncias, g=1,2.

A densidade de Π_g é:

$$f_g(\underset{\sim}{x}) = f_g(\underset{\sim}{x}, \underset{\sim}{M}_g, V_g) = \frac{1}{(2\Pi)^{p/2}|V_g|^{1/2}} \exp\left\{-\frac{1}{2}(\underset{\sim}{x} - \underset{\sim}{M}_g)' V_g^{-1}(\underset{\sim}{x} - \underset{\sim}{M}_g)\right\} \qquad (1)$$

$$g = 1,2$$

E foi visto no capítulo 1 que as regiões R_1 e R_2 que minimizam o custo esperado de má classificação são encontradas de acordo com o quociente de densidades.

2.2 CONSTRUÇÃO DA REGRA DE CLASSIFICAÇÃO QUADRATICA

Neste caso as regiões R_1 e R_2 que minimizam o custo esperado de má classificação consideram a razão:

$$\frac{f_1(\underset{\sim}{x})}{f_2(\underset{\sim}{x})} = \frac{(2\pi)^{p/2}\,|V_2|^{1/2}\,\exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_1)'V_1^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_1)\right\}}{(2\pi)^{p/2}\,|V_1|^{1/2}\,\exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_2)'V_2^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_2)\right\}} \qquad (2)$$

$$= \frac{|V_2|^{1/2}}{|V_1|^{1/2}}\exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_1)'V_1^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_1)+\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_2)'V_2^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_2)\right\}$$

pelo teorema 1 de (1.1) a região R_1 está definida por:

$$R_1:\ \frac{|V_2|^{1/2}}{|V_1|^{1/2}}\exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_1)'V_1^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_1)+\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_2)'V_2^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_2)\right\} \geq \left[\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right]$$

Aplicando o logaritmo tem-se:

$$\ln\left[\frac{|V_2|^{1/2}}{|V_1|^{1/2}}\right]-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_1)'V_1^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_1)+\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{M}_2)'V_2^{-1}(\underset{\sim}{x}-\underset{\sim}{M}_2) \geq \ln\left[\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right] \qquad (3)$$

ou, equivalentemente:

$$-\frac{1}{2}\,\underset{\sim}{x}'(V_1^{-1} - V_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(V_1^{-1}\underset{\sim}{M}_1 - V_2^{-1}\underset{\sim}{M}_2) + C \geq \ln\left\{\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right\} \qquad (4)$$

$$\text{onde } C = \ln\left\{\frac{|V_2|^{1/2}}{|V_1|^{1/2}}\right\} - \frac{1}{2}\,(\underset{\sim}{M}_1' V_1^{-1}\underset{\sim}{M}_1 - \underset{\sim}{M}_2' V_2^{-1}\underset{\sim}{M}_2)$$

e com o resultado obtido, formula-se o seguinte teorema:

TEOREMA 2.1 - Se Π_g, $g = 1,2$ tem a densidade (1), as regiões que minimizam o custo esperado de má classificação são dadas pelos valores de $\underset{\sim}{x}$ que satisfazem às seguintes desigualdades:

$$R_1: -\frac{1}{2}\underset{\sim}{x}'(V_1^{-1} - V_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(V_1^{-1}\underset{\sim}{M}_1 - V_2^{-1}\underset{\sim}{M}_2) + C \geq \ln\left\{\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right\}$$

$$R_2: -\frac{1}{2}\,\underset{\sim}{x}'(V_1^{-1} - V_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(V_1^{-1}\underset{\sim}{M}_1 - V_2^{-1}\underset{\sim}{M}_2) + C < \ln\left\{\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right\}$$

De onde surge diretamente a regra de classificação para as duas populações normais multivariadas com diferente estrutura de covariâncias, que é a seguinte :

Alocar $\underset{\sim}{x}$ na população Π_1 se :

$$-\frac{1}{2}\,\underset{\sim}{x}'(V_1^{-1}-V_2^{-1})\underset{\sim}{x}+\underset{\sim}{x}'(V_1^{-1}\underset{\sim}{M}_1-V_2^{-1}\underset{\sim}{M}_2)+C\geq \ln\left\{\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right\}$$

caso contrário alocar $\underset{\sim}{x}$ em Π_2 (5)

Para obter o resultado anterior foi suposto que as duas distribuições normais têm parâmetros conhecidos. Mas, em muitas aplicações desta teoria, os parâmetros não são conhecidos e têm que ser estimados com base nos resultados amostrais das populações Π_1 e Π_2. Substituindo assim, as quantidades amostrais $\bar{\underset{\sim}{X}}^{(1)}$, $\bar{\underset{\sim}{X}}^{(2)}$, S_1, S_2 em $\underset{\sim}{M}_1$, $\underset{\sim}{M}_2$, V_1, V_2 respectivamente, conforme as expressões (13) do capítulo 1, temos então, a regra de classificação quadrátrica amostral para populações com distribuição normal e diferentes matrizes de variâncias e covariâncias, que é a seguinte:

Alocar a observação $\underset{\sim}{x}$ em Π_1 se:

$$\hat{R}_1:-\frac{1}{2}\,\underset{\sim}{x}'(S_1^{-1}-S_2^{-1})\underset{\sim}{x}+\underset{\sim}{x}'(S_1^{-1}\bar{\underset{\sim}{X}}^{(1)}-S_2^{-1}\bar{\underset{\sim}{X}}^{(2)})+C\geq \ln\left[\frac{q_2\ C(1/2)}{q_1\ C(2/1)}\right]$$

caso contrário alocá-lo em Π_2 com

$$\hat{C}=\ln\left[\frac{|S_2|^{1/2}}{|S_1|^{1/2}}\right]-\frac{1}{2}\,(\bar{\underset{\sim}{X}}^{(1)},S_1^{-1}\bar{\underset{\sim}{X}}^{(1)}-\bar{\underset{\sim}{X}}^{(2)},S_2^{-1}\bar{\underset{\sim}{X}}^{(2)}) \qquad (6)$$

Que será chamada de Regra de Classificação Quadrática Amostral de Anderson.

2.3 CONSTRUÇÃO DE UM NOVO MÉTODO DE DISCRIMINAÇÃO

2.3.1 PARAMETROS CONHECIDOS

Considerando os mesmos supostos iniciais de haver duas populações normais multivariadas Π_1 e Π_2 com estrutura de covariâncias diferentes e usando os resultados apresentados nos itens (1.6.1) e (1.6.2) do capítulo anterior, isto é, tem-se:

O vetor $\underset{\sim}{X}'^{(g)} = (X_1^{(g)}, X_2^{(g)}, \ldots\ldots, X_p^{(g)})$ $\quad g = 1,2$

sob Π_1 tem distribuição $N_p(\underset{\sim}{M}_1, V_1)$ e
sob Π_2 tem distribuição $N_p(\underset{\sim}{M}_2, V_2)$

E quer-se estabelecer um critério para classificar um indivíduo com medições $\underset{\sim}{x} \in \mathbb{R}^p$ como pertencente a Π_1 ou Π_2. Então, são propostos os seguintes passos:

1º) Transformar as variáveis $\underset{\sim}{X}^{(g)}$ normais multivariadas em univariadas através das combinações lineares:

$$Y^{(g)} = \underset{\sim}{a}' \underset{\sim}{X}^{(g)}, \qquad g = 1,2$$

E tomar aquelas com quocientes extremos de variâncias, que são:

$$\underset{\sim}{Y}^{(1)} = \begin{bmatrix} Y_{max}^{(1)} \\ \\ Y_{min}^{(1)} \end{bmatrix} = \begin{bmatrix} \beta_1 \underset{\sim}{X}^{(1)} \\ \\ \beta_p \underset{\sim}{X}^{(1)} \end{bmatrix} \in \mathbb{R}^2$$

$$\underset{\sim}{Y}^{(2)} = \begin{bmatrix} Y_{max}^{(2)} \\ \\ Y_{min}^{(2)} \end{bmatrix} = \begin{bmatrix} \beta_1 \underset{\sim}{X}^{(2)} \\ \\ \beta_p \underset{\sim}{X}^{(2)} \end{bmatrix} \in \mathbb{R}^2$$

com distribuições $N_2(\underset{\sim}{\mu}^{(1)}, \Sigma_1)$ e $N_2(\underset{\sim}{\mu}^{(2)}, \Sigma_2)$ respectivamente, onde:

$$\underset{\sim}{\mu}^{(1)} = \begin{bmatrix} \mu_1^{(1)} \\ \mu_2^{(1)} \end{bmatrix} \qquad \underset{\sim}{\mu}^{(2)} = \begin{bmatrix} \mu_1^{(2)} \\ \mu_2^{(2)} \end{bmatrix}$$

(7)

$$\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \qquad \Sigma_2 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}$$

λ_1 e λ_p : máximo e mínimo autovalores da matriz $V_1^{-1}V_2$ e
β_1 e β_2 : autovetores associados a λ_1 e λ_p conforme o teorema 1.3.

2º) Dado que $\underset{\sim}{Y} = \begin{pmatrix} \beta_1' \\ \beta_p \end{pmatrix} \underset{\sim}{X}$ tem distribuição $N_2(\underset{\sim}{\mu}^{(1)}, \Sigma_1 = I_2)$

sob Π_1 e distribuição $N_2(\underset{\sim}{\mu}^{(2)}, \Sigma_2)$ sob Π_2, o vetor $\underset{\sim}{x} \in \mathbb{R}^p$ será bem alocado em Π_1 e Π_2, sempre que $\underset{\sim}{Y}$ o for em uma das duas populações transformadas.

Mas, a densidade transformada f_g é:

$$f_g(\chi, \underset{\sim}{\mu}^{(g)}, \Sigma_g) = \frac{1}{(2\Pi)^{p/2} |\Sigma_g|^{1/2}} \exp\left\{-\frac{1}{2}(\chi - \underset{\sim}{\mu}^{(g)})' \Sigma_g^{-1} (\chi - \underset{\sim}{\mu}^{(g)})\right\} \tag{8}$$

E as regiões R_1 e R_2 que minimizam o custo esperado de má classificação consideram a razão:

$$\frac{f_1(\chi, \underset{\sim}{\mu}^{(1)}, \Sigma)}{f_2(\chi, \underset{\sim}{\mu}^{(2)}, \Sigma)} = \frac{(2\Pi)^{p/2} |\Sigma_2|^{1/2} \exp\left\{\frac{1}{2}(\chi - \underset{\sim}{\mu}^{(1)})' \Sigma_1^{-1} (\chi - \underset{\sim}{\mu}^{(1)})\right\}}{(2\Pi)^{p/2} |\Sigma_1|^{1/2} \exp\left\{-\frac{1}{2}(\chi - \underset{\sim}{\mu}^{(2)})' \Sigma_2^{-1} (\chi - \underset{\sim}{\mu}^{(2)})\right\}}$$

E pelo teorema 1.4 tem-se:

$$R_1: \frac{|\Sigma_2|^{1/2}}{|\Sigma_1|^{1/2}} \exp\left\{-\frac{1}{2}(\chi - \underset{\sim}{\mu}^{(1)})' \Sigma_1^{-1} (\chi - \underset{\sim}{\mu}^{(1)}) + \frac{1}{2}(\chi - \underset{\sim}{\mu}^{(2)})' \Sigma_2^{-1} (\chi - \underset{\sim}{\mu}^{(2)})\right\}$$

$$\geq \frac{q_2 \, C(1/2)}{q_1 \, C(2/1)}$$

Aplicando logaritmo:

$$R_1: \; -\frac{1}{2}\,\underset{\sim}{x}'(I_2 - \Sigma_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(I_2\underset{\sim}{\mu}_1^{(1)} - \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) - \frac{1}{2}\,(\underset{\sim}{\mu}^{(1)\prime}\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + \frac{1}{2}\,\ln\frac{|\Sigma_2|}{|\Sigma_1|} \geq \ln\left\{\frac{q_2 C(1/2)}{q_1 C(2/1)}\right\}$$

(9)

$$R_2: \; -\frac{1}{2}\,\underset{\sim}{x}'(I_2 - \Sigma_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(I_2\underset{\sim}{\mu}^{(1)} - \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) - \frac{1}{2}\,(\underset{\sim}{\mu}^{(1)\prime}\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + \frac{1}{2}\,\ln\frac{|\Sigma_2|}{|\Sigma_1|} < \ln\left\{\frac{q_2 C(1/2)}{q_1 C(2/1)}\right\}$$

Resultado que pode ser expresso na forma do seguinte teorema:

TEOREMA 2.2 - Se Π_g, $g = 1, 2$ tem densidades (7), as regiões R_1 e R_2 que minimizam o custo esperado são dadas pelos valores $\underset{\sim}{y}$ que satisfazem às seguintes desigualdades:

$$R_1: \; -\frac{1}{2}\,\underset{\sim}{x}'(I_2 - \Sigma_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(I_2\underset{\sim}{\mu}^{(1)} - \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + d \geq \ln\left\{\frac{q_2 C(1/2)}{q_1 C(2/1)}\right\}$$

$$R_2: \; -\frac{1}{2}\,\underset{\sim}{x}'(I_2 - \Sigma_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(I_2\underset{\sim}{\mu}^{(1)} - \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + d < \ln\left\{\frac{q_2 C(1/2)}{q_1 C(2/1)}\right\}$$

onde: $\quad d = \frac{1}{2}\ln|\Sigma_2| - \frac{1}{2}\,(\underset{\sim}{\mu}^{(1)\prime}\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)})$

de onde surge a proposta de classificação para duas populações normais com diferente estrutura de covariâncias, que é a seguinte:

Alocar $\underset{\sim}{x}$ na população Π_1 quando:

$$-\frac{1}{2}\underset{\sim}{Y}'(I_2 - \Sigma_2^{-2})\underset{\sim}{Y} + \underset{\sim}{Y}'(I_2\underset{\sim}{\mu}^{(1)} - \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + d \geq \ln\left[\frac{q_2 C(1/2)}{q_1 C(2/1)}\right]$$

caso contrário alocar em Π_2 (10)

---

com $\underset{\sim}{Y} = \begin{bmatrix} \beta_1' \\ \beta_2' \end{bmatrix} \underset{\sim}{X}$

2.3.2 PARAMETROS DESCONHECIDOS

Para apresentar os resultados em (2.3.1) foi admitido que as duas distribuições tem parâmetros conhecidos, mas, em muitas aplicações práticas os parâmetros têm que ser estimados com base nos resultados amostrais das populações Π_1 e Π_2.

Sejam:

$\underset{\sim}{X}_1^{(1)}, \underset{\sim}{X}_2^{(1)}, \ldots, \underset{\sim}{X}_{n1}^{(1)}$ com distribuição $N(\underset{\sim}{M}_1, V_1)$ e

$\underset{\sim}{X}_1^{(2)}, \underset{\sim}{X}_2^{(2)}, \ldots, \underset{\sim}{X}_{n2}^{(2)}$ com distribuição $N(\underset{\sim}{M}_2, V_2)$

a partir das quais obtém-se as amostras:

$\underset{\sim}{Y}_1^{(1)}, \underset{\sim}{Y}_2^{(1)}, \ldots, \underset{\sim}{Y}_{n1}^{(1)}$ com distribuição $N(\underset{\sim}{\mu}^{(1)}, \Sigma_1)$ e

$\underset{\sim}{Y}_1^{(2)}, \underset{\sim}{Y}_2^{(2)}, \ldots, \underset{\sim}{Y}_{n2}^{(2)}$ com distribuição $N(\underset{\sim}{\mu}^{(2)}, \Sigma_2)$

onde:

$$\underset{\sim}{Y}_i^{(g)} = \begin{pmatrix} \underset{\sim}{b}_1' \underset{\sim}{X}_i^{(g)} \\ \underset{\sim}{b}_p' \underset{\sim}{X}_i^{(g)} \end{pmatrix} \in \mathbb{R}^2 \qquad \begin{matrix} i=1,\ldots,n_g \\ g=1,2 \end{matrix} \qquad (11)$$

$\underset{\sim}{b}_1$ e $\underset{\sim}{b}_p$ são autovetores associados aos autovalores

r_1 e r_p da matriz $S_1^{-1}S_2$. S_1 e S_2 foram obtidas conforme (13) do capítulo 1.

Fazendo

$$\hat{\underset{\sim}{\mu}}^{(1)} = \bar{\underset{\sim}{Y}}^{(1)} = \begin{pmatrix} \underset{\sim}{b}_1' \bar{\underset{\sim}{X}}^{(1)} \\ \underset{\sim}{b}_p' \bar{\underset{\sim}{Y}}^{(1)} \end{pmatrix} \in \mathbb{R}^2$$

$$\hat{\underset{\sim}{\mu}}^{(2)} = \bar{\underset{\sim}{Y}}^{(2)} = \begin{pmatrix} \underset{\sim}{b}_1' \bar{\underset{\sim}{X}}^{(2)} \\ \underset{\sim}{b}_p' \bar{\underset{\sim}{Y}}^{(2)} \end{pmatrix} \in \mathbb{R}^2 \qquad (12)$$

$$\hat{\Sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

$$\hat{\Sigma}_2 = \begin{pmatrix} r_1 & 0 \\ 0 & r_p \end{pmatrix}$$

$$\hat{d} = \frac{1}{2} \ln\left(r_1 r_p\right) - \frac{1}{2} \left[\bar{\underset{\sim}{Y}}^{(1)\prime} \bar{\underset{\sim}{Y}}^{(1)} - \bar{\underset{\sim}{Y}}^{(2)\prime} \hat{\Sigma}_2^{-1} \bar{\underset{\sim}{Y}}^{(2)} \right]$$

Tem-se agora, a seguinte proposta de classificação amostral para populações normais multivariadas com diferente estrutura de covariâncias:

Alocar uma observação $\underset{\sim}{x}$ na população Π_1 quando:

$$\hat{R}_1: \; -\frac{1}{2}\underset{\sim}{x}'(I_2 - \hat{\Sigma}_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(I_2\bar{\underset{\sim}{Y}}^{(1)} - \hat{\Sigma}_2^{-1}\bar{\underset{\sim}{Y}}^{(2)}) + \hat{d} \geq \ln\left\{\frac{q_2 C(1/2)}{q_1 C(2/1)}\right\}$$

e em Π_2 quando:

$$\hat{R}_2: \; -\frac{1}{2}\underset{\sim}{x}'(I_2 - \hat{\Sigma}_2^{-1})\underset{\sim}{x} + \underset{\sim}{x}'(I_2\bar{\underset{\sim}{Y}}^{(1)} - \hat{\Sigma}_2^{-1}\bar{\underset{\sim}{Y}}^{(2)}) + \hat{d} < \ln\left\{\frac{q_2 C(1/2)}{q_1 C(2/1)}\right\} \qquad (13)$$

2.3.3 AVALIAÇÃO DO PROCEDIMENTO DE DISCRIMINAÇÃO

Dado que é importante conhecer o desempenho da função de discriminação proposta em (2.3.1), apresentamos a seguir as tentativas feitas que conduzem à obtenção da sua distribuição com a finalidade de avaliar a probabilidade total de má classificação. Têm-se os seguintes resultados (*Rao e Mitra ,1971*):

LEMA 9.1.2 - Seja $Y \sim N_p(\underset{\sim}{\mu}, I)$. Então a estatística

$$\underset{\sim}{Y}'A\,\underset{\sim}{Y} + 2\,\underset{\sim}{b}'\underset{\sim}{Y} + E$$

tem distribuição $\chi^2(k,\delta)$ se e só se:

$$i)\ A^2 = A$$
$$ii)\ \underset{\sim}{b} \in \mathcal{C}(A),\ E = \underset{\sim}{b}'\underset{\sim}{b}$$
$$\text{com } k = R(A) \text{ e } \delta = (\underset{\sim}{b} + \underset{\sim}{\mu})'A(\underset{\sim}{b} + \underset{\sim}{\mu})$$

TEOREMA 9.2.1 - Seja $\underset{\sim}{Y} \sim N_p(\underset{\sim}{\mu},\Sigma)$, onde Σ pode ser singular. Então a estatística

$$\underset{\sim}{Y}'A\,\underset{\sim}{Y} + 2\,\underset{\sim}{b}'\underset{\sim}{Y} + E$$

tem distribuição $\chi^2(k,\delta)$ se e só se:

$$i)\ \Sigma A \Sigma A \Sigma = \Sigma A \Sigma \quad \text{ou equivalentemente}$$
$$(\Sigma A)^3 = (\Sigma A)^2$$
$$ii)\ \mathcal{C}[\Sigma(A\underset{\sim}{\mu} + \underset{\sim}{b})] \in \mathcal{C}(\Sigma A \Sigma)$$
$$iii)(A\underset{\sim}{\mu} + \underset{\sim}{b})'\Sigma(A\underset{\sim}{\mu} + \underset{\sim}{b}) = \underset{\sim}{\mu}'A\,\underset{\sim}{\mu} + 2\,\underset{\sim}{b}'\underset{\sim}{\mu} + E$$

em tal caso:

$$k = tr(A\Sigma)$$
$$\delta = (\underset{\sim}{b}+A\underset{\sim}{\mu})'\Sigma A \Sigma(\underset{\sim}{b}+A\underset{\sim}{\mu})$$

onde $\mathcal{C}(D)$, é o espaço gerado pelas colunas da matriz D.

Agora considerando as duas populações normais transformadas com parâmetros $\underset{\sim}{\mu}^{(1)}$, $\underset{\sim}{\mu}^{(2)}$, Σ_1 e Σ_2 conhecidos, deseja-se encontrar a distribuição da variável aleatória:

$$U = \underset{\sim}{Y}'(I_2 - \Sigma_2^{-1})\underset{\sim}{Y} - 2\underset{\sim}{Y}'(I_2\underset{\sim}{\mu}^{(1)} - \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + \underset{\sim}{\mu}^{(1)}{}'\underset{\sim}{\mu}^{(1)}$$
$$- \underset{\sim}{\mu}^{(2)}{}'\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)} - \ln(\lambda_1\lambda_p) \qquad (14)$$

$$A = \Sigma_1^{-1} - \Sigma_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_p \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\lambda_1 - 1}{\lambda_1} & 0 \\ 0 & \dfrac{\lambda_p - 1}{\lambda_p} \end{bmatrix}$$

$$B = -\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)} + \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)} =$$

$$= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \mu_1^{(1)} \\ \mu_2^{(1)} \end{bmatrix} + \begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_p \end{bmatrix} \begin{bmatrix} \mu_1^{(2)} \\ \mu_2^{(2)} \end{bmatrix}$$

$$= \begin{bmatrix} -\mu_1^{(1)} \\ -\mu_2^{(1)} \end{bmatrix} + \begin{bmatrix} \mu_1^{(2)}/\lambda_1 \\ \mu_2^{(2)}/\lambda_p \end{bmatrix} = \begin{bmatrix} \mu_1^{(2)}/\lambda_1 & -\mu_1^{(1)} \\ \mu_2^{(2)}/\lambda_p & -\mu_2^{(1)} \end{bmatrix}$$

$$E = \underset{\sim}{\mu}^{(1)\prime}\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)} - \ln(\lambda_1\lambda_p)$$

aplicando os resultados acima, não se verificam as suposições do teorema 9.2.1, pois:

sob Π_1: $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}^{(1)}, I)$:

$$A^2 = \begin{bmatrix} \left[\dfrac{\lambda_1 - 1}{\lambda_1}\right]^2 & 0 \\ 0 & \left[\dfrac{\lambda_p - 1}{\lambda_p}\right]^2 \end{bmatrix} \neq A$$

$B'B \neq E$

e não se pode concluir que U tem distribuição χ^2.

sob Π_2: $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}^{(2)}, \Sigma_2)$:

$$\Sigma_2 A\, \Sigma_2 A\, \Sigma_2 = \begin{pmatrix} \lambda_1 - 1 & 0 \\ 0 & \lambda_p - 1 \end{pmatrix}\begin{pmatrix} \lambda_1 - 1 & 0 \\ 0 & \lambda_p - 1 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{pmatrix}$$

$$= \begin{pmatrix} (\lambda_1 - 1)^2 & 0 \\ 0 & (\lambda_p - 1)^2 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{pmatrix} = \begin{pmatrix} \lambda_1(\lambda_1 - 1)^2 & 0 \\ 0 & \lambda_p(\lambda_p - 1)^2 \end{pmatrix}$$

$$\Sigma_2 A\, \Sigma_2 = \begin{pmatrix} \lambda_1 - 1 & 0 \\ 0 & \lambda_p - 1 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{pmatrix} = \begin{pmatrix} \lambda_1(\lambda_1 - 1) & 0 \\ 0 & \lambda_p(\lambda_p - 1) \end{pmatrix}$$

Com $\Sigma_2 A\, \Sigma_2 A\, \Sigma_2 \neq \Sigma_2 A\, \Sigma_2$.

também não se pode concluir que U tem distribuição χ^2.

Continuando com o estudo para tentar obter a distribuição de U faz-se a seguinte decomposição:

$U = U_1 + U_2 - \ln(\lambda_1 \lambda_p)$ onde:

$$U_1 = \underset{\sim}{Y}' \Sigma_1^{-1} \underset{\sim}{Y} + 2\, \underset{\sim}{Y}'(-\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}) + \underset{\sim}{\mu}^{(1)\prime} \Sigma_1^{-1} \underset{\sim}{\mu}^{(1)} \qquad (15)$$

$$U_2 = \underset{\sim}{Y}'(-\Sigma_2^{-1})\underset{\sim}{Y} + 2\, \underset{\sim}{Y}'(\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) - \underset{\sim}{\mu}^{(2)\prime} \Sigma_2^{-1} \underset{\sim}{\mu}^{(2)}$$

Para se obterem os valores médios e as variâncias e covariâncias tem-se:

sob Π_1: Se $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}^{(1)}, \Sigma_1)$, então com:

$$i)\ A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Longrightarrow A^2 = A$$

$$ii)\ B = -\left[\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}\right] = \begin{bmatrix} -\mu_1^{(1)} \\ -\mu_2^{(1)} \end{bmatrix} \Longrightarrow B \in \mathcal{C}(A)$$

$$E = \underset{\sim}{\mu}^{(1)\prime}\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)} = (\mu_1^{(1)}, \mu_2^{(2)}) \begin{bmatrix} \mu_1^{(1)} \\ \mu_2^{(1)} \end{bmatrix} = B'B$$

tem-se que, sob Π_1:

$$U_1 = \underset{\sim}{Y}'\Sigma_1^{-1}\underset{\sim}{Y} + 2\,\underset{\sim}{Y}'(-\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}) + \underset{\sim}{\mu}^{(1)'}\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)} \sim \chi^2(2,0)$$

e, agora, se:

$$i)\ A = \Sigma_2^{-1} = \begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_p \end{bmatrix} \Longrightarrow$$

$$A^2 = \begin{bmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_p^2 \end{bmatrix} \neq A$$

$$ii)\quad B = \Sigma_2^{-1}\mu^{(2)} = \begin{bmatrix} \mu_1^{(2)}/\lambda_1 \\ \mu_2^{(2)}/\lambda_p \end{bmatrix} \in \mathcal{C}(A)$$

$$B'B = \left[\frac{\mu_1^{(2)}}{\lambda_1}\right]^2 + \left[\frac{\mu_2^{(2)}}{\lambda_p}\right]^2$$

$$E = -\underset{\sim}{\mu}^{(2)'}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)} = -\frac{\left(\mu_1^{(2)}\right)^2}{\lambda_1} - \frac{\left(\mu_1^{(2)}\right)^2}{\lambda_1} \neq B'B$$

e por (i) e (ii) não se pode concluir que, sob Π_1, U_2 tenha distribuição χ^2.

Mas, sob Π_2: $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}^{(2)}, \Sigma_2)$, fazendo:

$$i)\ A = \Sigma_1^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad B = -\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}$$

$$E = \underset{\sim}{\mu}^{(1)}, \underset{\sim}{\mu}^{(1)}$$

$$\Sigma_2 A\ \Sigma_2 A\ \Sigma_2 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix} = \begin{bmatrix} \lambda_1^3 & 0 \\ 0 & \lambda_p^3 \end{bmatrix}$$

$$\Sigma_2 A\ \Sigma_2 = \begin{bmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_p^2 \end{bmatrix}$$

$$\Longrightarrow \Sigma_2 A\ \Sigma_2 A\ \Sigma_2 \neq \Sigma_2 A\ \Sigma_2$$

e, com esse resultado, sob Π_2, ainda não se pode dizer que U_1 tem distribuição χ^2.

Agora, se:

$$i)\ A = -\begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_p \end{bmatrix}$$

$$B = \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)} = \begin{bmatrix} \mu_1^{(2)}/\lambda_1 \\ \mu_2^{(2)}/\lambda_p \end{bmatrix}$$

$$E = -\underset{\sim}{\mu}^{(2)}, \Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}$$

$$\Sigma_z A\ \Sigma_z A\ \Sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}$$

$$\Sigma_z A\ \Sigma_z = -\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}$$

$$ii)\ A\underset{\sim}{\mu}^{(2)} + \underset{\sim}{b} = -\begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_p \end{bmatrix}\begin{bmatrix} \mu_1^{(2)} \\ \mu_2^{(2)} \end{bmatrix} + \begin{bmatrix} \mu_1^{(2)}/\lambda_1 \\ \mu_2^{(2)}/\lambda_p \end{bmatrix} = \underset{\sim}{0}$$

$$\Sigma_z(A\underset{\sim}{\mu}^{(2)} + \underset{\sim}{b}) = 2\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_p \end{bmatrix}\begin{bmatrix} \mu_1^{(2)}/\lambda_1 \\ \mu_2^{(2)}/\lambda_p \end{bmatrix} = \underset{\sim}{0}$$

vemos que $\mathcal{R}(\Sigma_z(A\underset{\sim}{\mu}^{(2)} + \underset{\sim}{b})) \in \mathcal{R}(\Sigma_z A\ \Sigma_z)$

$$iii)\ \delta = (A\underset{\sim}{\mu}^{(2)} + \underset{\sim}{b})'\Sigma_z(A\underset{\sim}{\mu}^{(2)} + \underset{\sim}{b}) = 4\left[\frac{\mu_1^{(2)}}{\lambda_1}\right]^2 = 0$$

por (i) e (ii) tem-se que sob Π_2:

$$U_2 = \underset{\sim}{Y}'\Sigma_2^{-1}\underset{\sim}{Y} + 2\,\underset{\sim}{Y}'(-\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)} \sim \chi^2(2,0)$$

Resumindo:

$$\text{sob } \Pi_1 : U_1 = \underset{\sim}{Y}'\Sigma_1^{-1}\underset{\sim}{Y} + 2\,\underset{\sim}{Y}'(-\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}) + \underset{\sim}{\mu}^{(1)\prime}\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}$$
$$\text{tem distribuição } \chi^2(2,0)$$

$$\text{sob } \Pi_2 : U_2 = \underset{\sim}{Y}'\Sigma_2^{-1}\underset{\sim}{Y} + 2\,\underset{\sim}{Y}'(-\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}$$
$$\text{tem distribuição } \chi^2(2,0) \qquad (16)$$

mas, não se conseguem obter as distribuições de

$$\underset{\sim}{Y}'\Sigma_1^{-1}\underset{\sim}{Y} + 2\,\underset{\sim}{Y}'(-\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)}) + \underset{\sim}{\mu}^{(1)\prime}\Sigma_1^{-1}\underset{\sim}{\mu}^{(1)} \qquad \text{e}$$
$$\underset{\sim}{Y}'\Sigma_2^{-1}\underset{\sim}{Y} + 2\,\underset{\sim}{Y}'(-\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}) + \underset{\sim}{\mu}^{(2)\prime}\Sigma_2^{-1}\underset{\sim}{\mu}^{(2)}$$

sob Π_1 e Π_2, respectivamente.

Pelo exposto, é muito difícil avaliar a probabilidade total de má classificação (PTM) da regra proposta.

2.3.4 ESTIMAÇÃO DA PROBABILIDADE TOTAL DE MÁ CLASSIFICAÇÃO

Mas no entanto, para dar prosseguimento ao estudo de avaliação, serão usados os valores amostrais $\bar{\underset{\sim}{X}}^{(1)}$, $\bar{\underset{\sim}{X}}^{(2)}$, S_1, S_2, $S_1^{-1}S_2$, $\bar{\underset{\sim}{Y}}^{(1)}$, $\bar{\underset{\sim}{Y}}^{(2)}$, na avaliação do desempenho da função classificadora amostral, calculando a Razão de Erro Real, (*Johnson e Wichern, 1982*).

A Razão de Erro Real (RER), é definida como:

$$RER = q_1 \int_{\hat{R}_2} f_1(\underset{\sim}{y})\, dy + q_2 \int_{\hat{R}_1} f_2(\underset{\sim}{y})\, dy \qquad (17)$$

onde $dy = dy_1 dy_2$,

$\hat{R}_1$ e $\hat{R}_2$ são as regiões de classificação amostral construídas através das amostras de tamanho n_1 e n_2 das populações, usando $\bar{\underset{\sim}{Y}}^{(1)}$, $\bar{\underset{\sim}{Y}}^{(2)}$, $\hat{\Sigma}_1$, $\hat{\Sigma}_2$ para estimar os parâmetros na expressão(9).

Abaixo se fará uma apresentação resumida dos estimadores da Razão de Erro Aparente .

2.3.4.1 Primeiro Estimador (Método de Re-substituição)

Existe um tipo de avaliação da função de classificação que não depende da forma das populações de origem, mas pode ser calculado para qualquer procedimento de classificação, e é chamado Razão de Erro Aparente (REA).

A REA é definida como a fração de observações das amostras que são mal classificadas pela função de classificação amostral. Este estimador é fácil de calcular, mas tende a subestimar a RER, a menos que n_1 e n_2, tamanhos amostrais sejam grandes.

Tomam-se n_1 observações de Π_1 e n_2 observações de Π_2, constrói-se a função de classificação amostral e avalia-se cada observação das amostras na função. Obtém-se então, a seguinte tabela:

POPULAÇÃO VERDADEIRA	DECISÃO ESTATÍSTICA		
	Π_1	Π_2	TOTAL
Π_1	n_{1c}	$n_{1m} = n_1 - n_{1c}$	n_1
Π_2	$n_{2m} = n_2 - n_{2c}$	n_{2c}	n_2

(18)

onde:

n_{1c}: nº de ítens de Π_1 classificados corretamente em Π_1.
n_{1m}: nº de ítens de Π_1 classificados incorretamente em Π_2
n_{2c}: nº de ítens de Π_2 classificados corretamentes em Π_2.
n_{2m}: nº de ítens de Π_2 classificados incorretamente em Π_2

A Razão de Erro Aparente é definida por:

$$REA = \frac{n_{1m} + n_{2m}}{n_1 + n_2} \qquad (19)$$

2.3.4.2 Segundo Estimador (Método HO)

Um outro estimador da Razão de Erro Real (RER) é obtido dividindo-se as amostras em dois subconjuntos. As duas sub-amostras do primeiro subconjunto são usadas para construir a função de discriminação amostral, enquanto que as sub-amostras do segundo subconjunto são usadas para avaliar a função.

Neste caso, o estimador da Razão de Erro Real é determinado pela proporção de membros do segundo subconjunto mal classificados, (subconjunto que serve para a avaliação). Este procedimento apresenta dois defeitos principais :

-requer amostras grandes.

-não são usadas todas as observações das amostras para construir a função, e nessa situação perdem-se informações importantes.

Resumindo, se $n_1 = n_{11} + n_{12}$

$n_2 = n_{21} + n_{22}$

$n_{11} + n_{21}$ são usados para construir a função de discriminação, e $n_{12} + n_{22}$ são usados para avaliar a função construída.

2.3.4.3 Terceiro Estimador (Método $\mathcal{U}$ modificado)

Este procedimento foi proposto por *Lachenbruch e Mickey, (1968)* e consiste em tomar amostras de tamanhos n_1 e n_2, de Π_1 e Π_2 respectivamente:

i) Começa-se com as observações de Π_1, omitindo uma observação desse grupo e construindo a função discriminante amostral com as $n_1 - 1$ e n_2 observações.

ii) Classifica-se a observação que é retirada da amostra de Π_1 segundo a função construída no ítem (i).

iii) Repetem-se os passos (i) e (ii) até que todas as observações de Π_1 sejam classificadas.

iv) Repeteu-se o mesmo procedimento com as observações de Π_2.

Seja n_{1m} o nº de observações mal classificadas de Π_1 e n_{2m} o de Π_2. Então, as estimativas de P(1/2) e P(2/1) são dadas por:

$$\hat{P}(2/1) = \frac{n_{1m}}{n_1} \qquad \hat{P}(1/2) = \frac{n_{2m}}{n_2}$$

e o estimador da Razão de Erro Real (RER) é:

$$E(\hat{RER}) = \frac{n_{1m} + n_{2m}}{n_1 + n_2} \qquad (20)$$

O desempenho da Função de Discriminação proposta neste capítulo será avaliada com a Razão de Erro Aparente calculada pelo Método de Re-substituição, por ser o mais conhecido, fácil de trabalhar e supondo que para amostras grandes proporciona uma boa estimativa da probabilidade total de má classificação.

CAPITULO 3

O DESEMPENHO DO NOVO METODO: SIMULAÇÃO E EXEMPLO

3.1. Introdução

No presente capítulo apresenta-se um estudo simulado com respeito ao critério de discriminação proposto no ítem (2.3) do capítulo anterior e uma aplicação prática com dados de *Flury e Riedwyl(1983)*, com a finalidade de mostrar a viabilidade do método.

Na primeira parte será apresentado um estudo baseado em simulações de Monte Carlo onde, com as amostras geradas construiu-se a regra de classificação amostral (13) do capítulo 2.

Tais amostras, além de se classificarem com este método, também são classificadas com (6) do capítulo anterior cuja fórmula consta no pacote estatístico Statistical Analysis System (SAS). O gerador de números aleatórios é o que está implementado no computador **VAX 785/11VMS** da *Universidade Estadual de Campinas (unicamp)*.

Para um melhor entendimento do trabalho ,na seção (3.2) far-se-á uma apresentação da geração das amostras das populações.

Na parte final do capítulo serão apresentados os resultados obtidos da aplicação.

3.2 SIMULAÇÕES

3.2.1 GERAÇÃO DAS DUAS POPULAÇÕES NORMAIS DE DIMENSÃO P

Existem vários geradores de números pseudo-aleatórios. Por facilidade e dada a sua ampla divulgação usamos o gerador RANDU implementado no **VAX 785/11**, o qual gera números pseudo-aleatórios de módulo $2^{31}-1$ que é do tipo:

$$g_{n+1} = k\, g_n \pmod m$$

sugerido por Lehmer. Tal algoritmo aparece no artigo de *Payne, Rabung e Bogyo (1969)*.
O resultado de uma chamada a RANDU é um número uniforme no intervalo (0,1).

Nos programas nº 1 e nº 2 do apêndice (em linguagem FORTRAN), depois da geração das variáveis uniformes, pseudo aleatórias, usou-se o método proposto por *Box-Muller(1958)*, que é uma transformação direta de variáveis uniformes independentes no intervalo (0,1), para variáveis normais independentes com média zero e variância um. Este procedimento inicial de geração de amostras de populações normais padrões é igual em ambos os grupos para qualquer número de variáveis até se chegar às variáveis correlacionadas, usando as transformações que serão apresentadas a seguir e que são muito conhecidas na literatura estatística.

1) A geração de vetores aleatórios normais padronizados de dimensão p (2 ou 3) é feita sob a seguinte suposição :

$$\underset{\sim}{z}_i^{(g)} \sim N(\underset{\sim}{\phi}, I_p) \qquad g=1,2 \qquad i=1,\ldots,n_g \text{ onde cada}$$

$$\underset{\sim}{z}_i^{\prime(g)} = (z_{i1}^{(g)},\ldots,z_{ip}^{(g)}) \tag{1}$$

criando-se assim as matrizes de dados amostrais de ordem $n_g * p$, que será denotada por:

$$\underset{\sim}{Z}^{\prime(g)} = (\underset{\sim}{z}_1^{(g)}, \underset{\sim}{z}_2^{(g)}, \ldots, \underset{\sim}{z}_{n_g}^{(g)}) \qquad g=1,2$$

2) A geração de variáveis independentes em cada população, com vetor de médias diferentes de zero e matriz de covariâncias diferente da identidade é feita através da transformação :

$$\underset{\sim}{w}_i^{(g)} = D^{(g)}\, \underset{\sim}{z}_i^{(g)} + \underset{\sim}{A}^{(g)} \qquad g=1,2$$

onde:

$$\underset{\sim}{w}_i^{(g)} \sim N_p(\underset{\sim}{A}^{(g)}, D^{(g)}\, D^{(g)\prime}) \tag{2}$$

com

$$\underset{\sim}{A}^{(g)} = \begin{bmatrix} a_1^{(g)} \\ \cdot\cdot \\ a_p^{(g)} \end{bmatrix} \qquad D^{(g)} = \begin{bmatrix} v_1^{(g)} & 0 & 0 \\ \cdots & \cdots & \cdots \\ 0 & & v_p^{(g)} \end{bmatrix}_{pxp} \tag{3}$$

3)Depois é introduzida a estrutura de correlações entre as variáveis segundo a transformação :

$$\underset{\sim}{x}_i^{(g)} = R^{(g)} \underset{\sim}{w}_i^{(g)} \qquad g=1,2 \quad , i=1,\ldots,n_g \qquad (4)$$

onde $R^{(g)}$ é um operador que produz uma rotação no plano e no espaço tridimensional segundo p, seja 2 ou 3.

Cabe lembrar que $\underset{\sim}{x}_i^{(g)} \sim N_p(\underset{\sim}{M}_g , V_g)$ com

$$\underset{\sim}{M}_g = R^{(g)} \underset{\sim}{A}^{(g)} \quad e \quad V_g = R^{(g)} D^{(g)} D^{(g)'} R^{(g)'} \qquad (5)$$

Criam-se as matrizes de dados $X^{(g)} = (\underset{\sim}{x}_1^{(g)}, \ldots, \underset{\sim}{x}_{n_g}^{(g)})'$ $g=1,2$ com $\underset{\sim}{x}_i^{(g)} = (x_{i1}^{(g)}, \ldots, x_{ip}^{(g)})$.

Nas presentes simulações foi usada, para o caso p=2 a matriz de rotação:

$$R^{(g)} = \begin{bmatrix} \cos\alpha & -\text{sen}\alpha \\ \text{sen}\alpha & \cos\alpha \end{bmatrix}^{(g)} \qquad (6)$$

e para p=3 a matriz:

$$R^{(g)} = \begin{bmatrix} \cos\alpha & -\text{sen}\alpha & 0 \\ \text{sen}\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}^{(g)}$$

dando origem a:

$$V_g = \begin{cases} \begin{bmatrix} v_1^2\cos^2\alpha + v_2^2\operatorname{sen}^2\alpha & (v_1^2 - v_2^2)\operatorname{sen}\alpha\cos\alpha \\ (v_1^2 - v_2^2)\operatorname{sen}\alpha\cos\alpha & v_1^2\operatorname{sen}^2\alpha + v_2^2\cos^2\alpha \end{bmatrix}^{(g)} & \text{se } p=2 \\ \begin{bmatrix} v_1^2\cos^2\alpha + v_2^2\operatorname{sen}^2\alpha & (v_1^2 - v_2^2)\operatorname{sen}\alpha\cos\alpha & 0 \\ (v_1^2 - v_2^2)\operatorname{sen}\alpha\cos\alpha & v_1^2\operatorname{sen}^2\alpha + v_2^2\cos^2\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}^{(g)} & \text{se } p=3 \end{cases}$$

A transformação usada é a mesma ou diferente nos grupos, dando origem a amostras com a mesma ou diferente estrutura de covariâncias.

Nas simulações com p=3 só foram criadas estruturas de correlações entre as duas primeiras componentes de $\underset{\sim}{x}_i^{(g)} = (x_{i1}, x_{i2}, x_{i3})$. Não obstante com a mesma idéia poderiam ter sido criadas correlações entre as outras ou entre todas elas. O importante é gerar duas amostras provenientes de distribuições normais com estrutura de covariâncias diferentes.

4) Para a execução inicial dos programas nº1 e nº2 os parâmetros de entrada são os seguintes:

$n_1, \ldots, n_g$, tamanho do g-ésimo grupo. Nestas simulações $n_1 = n_2 = 100$.

$v_1^{(g)}, \ldots, v_p^{(p)}$, são os elementos da matriz $D^{(g)}$.

$a_1^{(g)}, \ldots\ldots, a_p^{(g)}$, são os elementos do vetor de médias $\underset{\sim}{A}^{(g)}$, que nas simulações foram fixados como nulos.

$\alpha^{(g)}$,são os ângulos que produzem as rotações nos grupos um e dois criando assim correlações entre as variáveis.

5) Depois de cada simulação de Monte Carlo e usando a teoria desenvolvida temos a seguintes matrizes de dados:

$$\mathcal{X}^{(g)} = \begin{bmatrix} x_{11}^{(g)} & x_{12}^{(g)} \ldots\ldots\ldots\ldots x_{1p}^{(g)} \\ \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{n_g 1}^{(g)} & x_{n_g 2}^{(g)} \ldots\ldots\ldots\ldots x_{n_g p}^{(g)} \end{bmatrix} \qquad \begin{matrix} g = 1,2 \\ \\ p = 2 \text{ ou } 3 \end{matrix}$$

contendo as observações amostrais com as p medições.

Com os dados contidos em $\mathcal{X}^{(g)}$ obtemos os vetores de médias e as matrizes de covariâncias amostrais segundo as fórmulas (13) do capítulo 1 ,chegando a obter $\underset{\sim}{\bar{X}}^{(1)}$, $\underset{\sim}{\bar{X}}^{(2)}$, S_1 e S_2 .

6)Neste ponto introduzimos nos programas n_1^o e n_2^o a subroutina **F02AEF** da biblioteca **NAG.**(Numerical Algorithns Groups).

Para clareza faz-se uma descrição resumida da subrotina **F02AEF** Que calcula todos os autovalores e autovetores de $A\underset{\sim}{b}=\lambda B\underset{\sim}{b}$, ou seja da matriz $B^{-1}A$ usando a redução de *Householder* e o algoritmo *QL.*(*Wilkinson,J.H, e Reinsch,C., 1971*).

Dado que a matriz $B^{-1}A$ é não simétrica, o problema reduz-se primeiro à obtenção de autovalores e autovetores de uma matriz simétrica , usando o método de *Cholesky* para decompor B em matrizes triangulares , B= LL' onde L é triangular inferior.

como $$A \underset{\sim}{b} = \lambda B \underset{\sim}{b} \qquad (7)$$

$$L^{-1}AL^{-1},L'\underset{\sim}{b} = \lambda L^{-1}B \underset{\sim}{b} \qquad \text{implica}$$

$$(L^{-1}AL^{-1},)(L'\underset{\sim}{b}) = \lambda L'\underset{\sim}{b}$$

então, os autovalores de (7) são os de

$$P \underset{\sim}{c} = \lambda \underset{\sim}{c} \qquad \text{onde} \quad \underset{\sim}{c} = L'\underset{\sim}{b} \qquad P=L^{-1}AL^{-1}, \qquad (8)$$

O método de *Householder* é usado para tridiagonalisar a matriz P simétrica e obter os seus autovalores com o algoritmo *QL*.

Um autovetor associado à matriz simétrica P está relacionado com o autovetor $\underset{\sim}{b}$ da matriz original $B^{-1}A$ mediante

$$\underset{\sim}{c} = L'\underset{\sim}{b} \qquad (9)$$

Dado que o autovetor $\underset{\sim}{c}$ é obtido com o algoritmo *QL* e normalizado segundo $\underset{\sim}{c}'\underset{\sim}{c} = 1$,então , os autovalores do problema original são obtidos resolvendo o sistema (9), onde $\underset{\sim}{b}$ é normalizado segundo $\underset{\sim}{b}' B \underset{\sim}{b} = 1$. Ou seja, o uso da subrotina F02AEF permite calcular os autovalores e autovetores da matriz $S_1^{-1}S_2$ com $A = S_2$ e $B=S_1$, onde S_1^{-1} é a inversa da matriz de covariâncias da amostra do grupo um.

A notação usada será $r_1,\ldots\ldots,r_p$ para os autovalores da matriz $S_1^{-1}S_2$, e $\underset{\sim}{b}_1,\ldots\ldots,\underset{\sim}{b}_p$ para os autovetores associados.

3.2.2 GERAÇÃO DE DUAS POPULAÇÕES NORMAIS DE DIMENSÃO DOIS

Com os autovetores $\underset{\sim}{b}_1$ e $\underset{\sim}{b}_p$ associados aos autovalores r_1 e r_p constroem-se as combinações lineares de interesse, não correlacionadas em ambos os grupos , definidas por (11) do capítulo anterior :

$$\underset{\sim}{y}_i^{(g)\prime} = (\underset{\sim}{b}_1,\underset{\sim}{b}_p)'\underset{\sim}{x}_i^{(g)} \qquad i=1,\ldots n_g \quad ,g=1,2$$

$$= (y_{i1}^{(g)},y_{i2}^{(g)}) \qquad \underset{\sim}{b}_1,\underset{\sim}{b}_2 \in \mathbb{R}^p \qquad (10)$$

criando assim as matrizes de dados de ordem $n_g * 2$:

$$\mathbb{Y}^{(g)\prime} = (\underset{\sim}{y}_1^{(g)},\ldots\ldots,\underset{\sim}{y}_{ng}^{(g)})$$

$$\mathbb{Y}^{(g)} = \begin{bmatrix} y_{11}^{(g)} & y_{12}^{(g)} \\ y_{21}^{(g)} & y_{22}^{(g)} \\ \cdots\cdots & \cdots\cdots \\ y_{n_g 1}^{(g)} & y_{n_g 2}^{(g)} \end{bmatrix}$$

$$Y^{(g)} = \begin{bmatrix} x_{11}^{(g)} & x_{12}^{(g)} & \cdots & x_{1p}^{(g)} \\ x_{21}^{(g)} & x_{22}^{(g)} & \cdots & x_{2p}^{(g)} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n_g 1}^{(g)} & x_{n_g 2}^{(g)} & \cdots & x_{n_g p}^{(g)} \end{bmatrix} \begin{bmatrix} b_{11} & b_{p1} \\ b_{12} & b_{p2} \\ \cdots & \cdots \\ b_{1p} & b_{pp} \end{bmatrix}$$

é evidente que :

$$\underset{\sim}{y}_i^{(1)} \sim N_2(\underset{\sim}{\mu}^{(1)}, I) \qquad i=1,\ldots,n_g \quad (11)$$

$$\underset{\sim}{y}_i^{(2)} \sim N_2(\underset{\sim}{\mu}^{(2)}, \Sigma_2)$$

3.2.3 CONSTRUÇÃO DA FUNÇÃO DISCRIMINANTE AMOSTRAL

Com os dados contidos nas matrizes $Y^{(g)}$ constrói-se a função discriminante amostral.

Obtêm-se as estatísticas:

$$\underset{\sim}{\bar{y}}^{(1)} = \begin{bmatrix} \bar{y}_1^{(1)} \\ \\ \bar{y}_2^{(1)} \end{bmatrix} = \frac{1}{n_1} \begin{bmatrix} y_{11}^{(1)} & y_{12}^{(1)} \\ \cdots & \\ y_{21}^{(1)} & y_{22}^{(1)} \\ \cdots & \cdots \\ y_{n_1 1}^{(1)} & y_{n_1 2}^{(1)} \end{bmatrix}' \begin{bmatrix} 1 \\ 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix}$$

$$\underset{\sim}{\bar{Y}}^{(2)} = \begin{bmatrix} \bar{Y}_1^{(2)} \\ \\ \bar{Y}_2^{(2)} \\ \cdot \end{bmatrix} = \frac{1}{n_2} \begin{bmatrix} y_{11}^{(2)} & y_{12}^{(2)} \\ \\ y_{21}^{(2)} & y_{22}^{(2)} \\ \cdots\cdots\cdots\cdots \\ y_{n_2 1}^{(2)} & y_{n_2 2}^{(2)} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix}$$

$$\hat{\Sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = I_2$$

$$\hat{\Sigma}_2 = \begin{bmatrix} r_1 & 0 \\ 0 & r_p \end{bmatrix}$$

necessárias para calcular a função discriminante amostral.

Foram calculados previamente:

a).-Termo independente:

$$XINDT = \underset{\sim}{\bar{Y}}^{(1),} \underset{\sim}{\bar{Y}}^{(1)} - \underset{\sim}{\bar{Y}}^{(2),} \hat{\Sigma}_2^{-1} \underset{\sim}{\bar{Y}}^{(2)} - \ln(r_1 \, r_p)$$

b).-vetor dos coeficientes do termo linear :

$$-2(1_2 \underset{\sim}{\bar{Y}}^{(1)} - \hat{\Sigma}_2^{-1} \underset{\sim}{\bar{Y}}^{(2)})$$

c).-Matriz do termo quadrático:

$$I_2 - \hat{\Sigma}_2^{-1}$$

d).-E logo a Função Discriminante Amostral:

$$D(\underset{\sim}{y}) = \underset{\sim}{y}'(I_2 - \hat{\Sigma}_2^{-1})\underset{\sim}{y} - 2\underset{\sim}{y}'(I_2\bar{\underset{\sim}{y}}^{(1)} - \hat{\Sigma}_2^{-1}\bar{\underset{\sim}{y}}^{(2)}) + XINDT \qquad (12)$$

e).-Assumindo custos de má classificação e probabilidades a priori iguais.

Então, o indivíduo que tem as medições $\underset{\sim}{y}_i^{(g)}$ é classificado como:

$$\begin{cases} \text{do grupo um, se } D(\underset{\sim}{y}_i^{(g)}) \leq 0 \text{ ou} \\ \text{do grupo dois, se } D(\underset{\sim}{y}_i^{(g)}) > 0 \end{cases} \qquad (13)$$

Ou seja, para as n_1 e n_2 observações de cada simulação, $(\underset{\sim}{x}_i^{(g)})$, substituímos os valores de cada observação transformada $(\underset{\sim}{y}_i^{(g)})$:

$$\underset{\sim}{y}_i^{(g)} = \begin{bmatrix} \underset{\sim}{b}_1' \underset{\sim}{x}_i^{(g)} \\ \\ \underset{\sim}{b}_p' \underset{\sim}{x}_i^{(g)} \end{bmatrix} \in \mathbb{R}^2$$

na função (12), obtendo-se assim a tabela (18) do capítulo 2.

Com tais valores calculou-se a proporção de itens classificados incorretamente, ou seja, a Razão de Erro Aparente (REA) com o Método de Re-substituição .

3.2.4 AVALIAÇÃO DA FUNÇÃO DE DISCRIMINAÇÃO AMOSTRAL E APRESENTAÇÃO DOS RESULTADOS EXPERIMENTAIS

Para fazer um melhor julgamento da Função de Discriminação Amostral (12) do ítem anterior, ajusta-se aos conjuntos de dados normais p-variados contidos nas matrizes $\mathbb{X}^{(1)}$ e $\mathbb{X}^{(2)}$, a metodologia implementada no STATISTICAL ANALYSIS SISTEM (SAS), residente no computador VAX/UNICAMP. Portanto, antes de fazer a avaliação, é conveniente apresentar um resumo de tal metodologia.(Programas n^{o} 3 e n^{o} 4 do apéndice).

Com os dados contidos em $\mathbb{X}^{(1)}$ e $\mathbb{X}^{(2)}$ forma-se só um arquivo que contém as variáveis medidas e uma variável classificadora identificando os grupos.
O "PROC DISCRIM" do SAS ajusta uma função discriminante para classificar cada observação em um dos grupos, assumindo distribuição normal com igual ou diferente estrutura de covariâncias nos grupos.

Opcionalmente pedimos um teste de homogeneidade para as matrizes de covariâncias, e se a hipótese de igualdade é rejeitada, a classificação é feita com a função de Discriminação Quadrática Amostral de Anderson, ou seja, com a regra (6) do capítulo 2, em caso contrário com a Função de Discriminação Linear proposta por Fisher-regra (14) do capítulo 1.

Com o método proposto, seja aceitando ou rejeitando o teste de homogeneidade de variâncias, trabalhou-se com (12), só que, é de se esperar que no caso de homoscedasticidade e dado que consideramos médias iguais, vai ser quase impossível a discriminação.

Em cada simulação de Monte Carlo, foram geradas duas matrizes de dados amostrais normais de dimensão dois e três respectivamente, sobre o que devem ser feitas as seguintes considerações:

Dado que, o que interessa é a diferença entre os ângulos e não o quadrante em que estão localizadas as amostras, e tendo confirmado isso através de simulações (ver quadro A1 do apéndice), então para facilitar o trabalho, foi fixado o ângulo de rotação do primeiro grupo em 0^{o}, e só foram feitas rotações no segundo grupo, com ângulos de rotação de 0^{o} a 90^{o}.

Tais condições iniciais para a criação dos grupos estão apresentadas nas colunas "PARAMETROS DE ENTRADA" dos quadros 3.1A até 3.7A para p = 2 e em 3.1B até 3.7B quando p = 3. Por exemplo, para p = 2, na execução 02A temos o seguinte: na população Π_1, $v_1 = v_2 = 1$, $\alpha = 0^{o}$, e na população Π_2, $v_1 = v_2 = 1$, $\alpha = 45^{o}$, ou seja, que as duas amostras foram retiradas de populações com distribuições circulares. Então, simularam-se amostras de populações esféricas, esféricas versus elípticas, elípticas versus elípticas com vários ângulos entre os respectivos eixos principais.

Com a finalidade de avaliar mais especificamente a validade do critério de discriminação proposto, supostamente baseado na diferenças das matrizes de covariâncias, as amostras criadas procedem de populações com vetores de médias iguais a zero. Então, nas simulações supõe-se que as distribuições estão centradas no vetor $\underset{\sim}{0}_{px1}$ e qualquer diferença entre elas está contida nas matrizes de covariâncias.

Para maior aproximação de situações práticas reais, em alguns casos, foram feitas simulações com variâncias iguais e em outros com variâncias diferentes.

Dado que seria impossível esboçarem-se todas as combinações de matrizes de covariâncias, trabalhou-se com algumas delas, de tal maneira que caracterizem situações de interesse.

Por exemplo, nas execuções 01A-02A, mantém-se as duas populações com distribuições circulares, entretanto, nas execuções 07A, 08A, 09A, perturbaram-se as variâncias sob independência em ambas populações, e nas execuções 03A, 10A, 11A, 14A, impõe-se estrutura de correlações no segundo grupo e mantêm-se independência no primeiro.

Com os dados das amostras geradas sob as condições dadas, foram feitos os cálculos expostos nas secções 3.2 e 3.3 do presente capítulo, obtendo-se os resultados apresentados nas colunas "TESTE DE HOMOGENEIDADE ", "PROPOSTA DE DISCRIMINAÇÃO", "CORRELAÇÃO AMOSTRAL", e "SAS" dos quadros 3.1A até 3.4B em p = 2 e p = 3.

Cabe ressaltar que em cada execução o teste de homogeneidade para as matrizes de covariâncias, conduz à rejeição ou aceitação da igualdade das matrizes o qual fica na coluna "TESTE DE HOMOGENEIDADE" dos quadros 3.1A até 3.4B.

Para clareza da interpretação das tabelas 3.1A até 3.4b.1, faz-se, como exemplo, uma pequena descrição das tabelas 3.1A e 3.1A.1.

Na tabela 3.1A apresentam-se quatro execuções onde em todas elas a primeira população é esférica; em tanto que , na segunda população, as duas primeiras execuções são esféricas, e tem-se introduzido estrutura de correlações na terceira execução, e na quarta execução, as variáveis são independentes.

Segundo o teste para as matrizes de covariâncias, do Proc Discrim, nas duas primeiras execuções é aceita a hipótese de homogeneidade e nas terceira e quarta execuções tal hipótese é rejeitada.

Para as mesmas execuções, na parte inferior da tabela apresenta-se o número de elementos mal classificados em cada grupo (n_{1m}, n_{2m}), segundo a proposta e o SAS; com tais valores calculam-se as estimativas das probabilidades de má classificação (REA). Também mostram-se os autovalores máximo e mínimo da matriz $S_1^{-1}S_2$.

Entretanto, a tabela 3.1A.1 refere-se exclusivamente à execução 01A, onde a parte superior contém o número da observação mal classificada segundo ambos os métodos e a parte inferior contém as observações mal classificadas por um dos métodos. Ou seja, que das 92 observações mal classificadas pela proposta e 82 observações mal classificadas com o SAS, (parte A+B ou parte A+C), 70 observações foram erradamente classificadas pelos dois métodos, (parte A).

As conclusões dos resultados experimentais serão apresentadas no último capítulo.

TABELA Nº3.1A

ESTIMATIVAS DAS PROBABILIDADES DE MA CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

EXE	PARAMETROS DE ENTRADA						CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEIDADE
	Pop um			Pop dois					
	v_1	v_2	α	v_1	v_2	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	$V_1=V_2$
01A	1	1	0	1	1	0	0.071	0.089	Aceitar
02A	1	1	0	1	1	45	0.071	0.019	Aceitar
03A	4	4	0	2	4	45	0.071	-0.589	Rejeitar
04A	4	4	0	2	4	90	0.071	-0.089	Rejeitar

EXE	PROPOSTA DE DISCRIMINAÇÃO						SAS		
	REA	n_{1m}	n_{2m}	Autovalores r_1	Autovalores r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
01A	46	35	57	1.33	1.09	1.26	41	42	40
02A	43.5	31	56	1.27	1.14	1.11	42.5	42	43
03A	31.0	42	20	1.27	0.285	4.56	31.0	42	20
04A	32.5	46	19	1.29	0.279	4.64	32.5	46	19

Fonte : Simulações com o Programa nº1. (Ver págs. 80, 81).

TABELLA Nº 3.2A

ESTIMATIVAS DAS PROBABILIDADES DE MA CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

EXE	PARAMETROS DE ENTRADA						CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEIDADE
	Pop um			Pop dois					
	v_1	v_2	α	v_1	v_2	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	$V_1=V_2$
05A	1	1	0	2	0.5	0	0.071	0.089	Rejeitar
06A	1	1	0	2	0.5	60	0.071	0.851	Rejeitar
07A	3	0.5	0	3	0.5	0	0.071	0.089	Aceitar
08A	3	0.5	0	3	0.5	90	0.071	-0.089	Rejeitar

EXE	PROPOSTA DE DISCRIMINAÇÃO						SAS		
	REA	n_{1m}	n_{2m}	Autovalores r_1	r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
05A	24	21	27	5.325	0.272	19.54	24	21	27
06A	28.5	25	32	4.461	0.325	13.72	28.5	25	32
07A	46.0	35	57	1.328	1.092	1.22	41.5	42	40
08A	9.5	8	11	41.230	0.035	1171.32	9.5	8	11

Fonte : Simulações com o Programa nº 1. (Ver págs. 82, 83)

TABELA Nº 3.3A

ESTIMATIVAS DAS PROBABILIDADES DE MA CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

EXE	PARAMETROS DE ENTRADA						CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEIDADE
	Pop um			Pop dois					
	v_1	v_2	α	v_1	v_2	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	$V_1=V_2$
09A	3	0.5	0	0.5	3	0	0.071	0.089	Rejeitar
10A	3	0.5	0	0.5	3	45	0.071	-0.944	Rejeitar
11A	8	0.05	0	8	0.05	60	0.071	-0.93	Rejeitar
12A	8	0.05	0	8	0.05	90	0.071	-.089	Rejeitar

EXE	PROPOSTA DE DISCRIMINAÇÃO						SAS		
	REA	n_{1m}	n_{2m}	Autovalores r_1	Autovalores r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
09A	8	8	8	39.645	0.037	1083.21	8	8	8
10A	13.5	11	16	22.066	0.066	335.35	13.5	11	16
11A	0.0	0	0	21983.5	0.0001	$314.0*10^6$	0.0	0	0
12A	0.0	0	0	29308.4	0.0001	$586.2*10^6$	0.0	0	0

Fonte : Simulações com o Programa nº 1.

TABELA Nº 3.4A

ESTIMATIVAS DAS PROBABILIDADES DE MÁ CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

EXE	PARAMETROS DE ENTRADA						CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEIDADE
	Pop um			Pop dois					
	ν_1	ν_2	α	ν_1	ν_2	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	$\nu_1=\nu_2$
13A	2	0.5	0	8	0.05	0	0.071	0.089	Rejeitar
14A	2	0.5	0	8	0.05	45	0.071	0.999	Rejeitar
15A	2	0.5	0	8	0.05	60	0.071	0.999	Rejeitar
16A	2	0.5	0	8	0.05	90	0.071	-0.089	Rejeitar

EXE	PROPOSTA DE DISCRIMINAÇÃO						SAS		
	REA	n_{1m}	n_{2m}	Autovalores r_1	r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
13A	6	9	3	21.34	0.0109	1958.2	6	9	3
14A	4	7	1	151.76	0.0015	$101.2*10^3$	4	7	1
15A	2	3	1	220.42	0.0011	$200.4*10^3$	2	3	1
16A	0.5	1	0	293.09	0.0008	$366.4*10^3$	0.5	1	0

Fonte : Simulações com o Programa nº 1. (Ver pág. 84).

TABELA Nº 3.1B

ESTIMATIVAS DAS PROBABILIDADES DE MA CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

	PARAMETROS DE ENTRADA								CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEI
	pop. um				pop. dois						
	v_1	v_2	v_3	α	v_1	v_2	v_3	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	
01B	1	1	1	0	1	1	1	0	0.043	0.062	Aceitar
02B	1	1	1	0	1	1	1	45	0.043	-0.013	Aceitar
03B	8	4	2	0	8	4	2	60	0.043	-0.518	Rejeitar
04B	8	4	2	0	8	4	2	90	0.043	-0.062	Rejeitar

	PROPOSTA DE DISCRIMINAÇÃO						SAS		
EXE	REA	n_{1m}	n_{2m}	Autovalores r_1	r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
01B	42	35	49	1.361	0.7258	1.88	45	45	45
02B	40	52	28	1.326	0.7050	1.88	41.5	42	41
03B	30	34	26	3.121	0.2560	12.19	30.5	34	27
04B	28.5	31	26	3.603	0.2199	16.39	25.5	26	25

Fonte : Simulações com o Programa nº 2. (Ver pág. 85).

TABELA Nº 3.2B

ESTIMATIVAS DAS PROBABILIDADES DE MÁ CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

	PARAMETROS DE ENTRADA								CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEI
	pop. um				pop. dois						
	v_1	v_2	v_3	α	v_1	v_2	v_3	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	
05B	1	1	1	0	2	1.5	0.5	45	0.043	0.268	Rejeitar
06B	1	1	1	0	2	1.5	0.5	60	0.043	0.205	Rejeitar
07B	6	4	2	0	4	4	4	0	0.043	0.062	Rejeitar
08B	6	4	2	0	4	4	4	90	0.043	-0.062	Rejeitar

	PROPOSTA DE DISCRIMINAÇÃO						SAS		
EXE	REA	n_{1m}	n_{2m}	Autovalores r_1	r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
05B	31.5	33	30	3.422	0.2800	12.22	26.5	23	30
06B	30.0	32	28	3.464	0.2806	12.34	26.5	23	30
07B	28.5	19	38	4.794	0.3648	13.14	29.5	21	38
08B	29.0	20	38	4.769	0.3875	12.31	28.0	26	30

Fonte : Simulações com o Programa nº 2. (Ver pág. 86, 87).

TABELA Nº 3.3B

ESTIMATIVAS DAS PROBABILIDADES DE MA CLASSIFICAÇÃOCREA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

	PARAMETROS DE ENTRADA								CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEI
	pop. um				pop. dois						
	v_1	v_2	v_3	α	v_1	v_2	v_3	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	
09B	2	0.5	1.5	0	8	0.05	1.5	0	0.043	0.0622	Rejeitar
10B	2	0.5	1.5	0	8	0.05	1.5	45	0.043	0.999	Rejeitar
11B	8	0.05	2.0	0	8	0.05	2.0	60	0.043	0.999	Rejeitar
12B	8	0.05	2.0	0	8	0.05	2.0	90	0.043	-0.062	Rejeitar

	PROPOSTA DE DISCRIMINAÇÃO						SAS		
EXE	REA	n_{1m}	n_{2m}	Autovalores r_1	Autovalores r_p	r_1/r_p	REA (%)	n_{1m}	n_{2m}
09B	7.5	12	3	14.106	0.0090	1567.3	6	10	2
10B	2.5	4	1	117.765	0.0011	107059.4	2	4	0
11B	0	0	0	16936.6	0.0001	$338.7*10^6$	0	0	0
12B	0.5	0	1	225.876	0.0006	$752.6*10^6$	0.5	0	1

Fonte : Simulações com o Programa nº 2. (ver págs. 88)

TABELA Nº 3.4B

ESTIMATIVAS DAS PROBABILIDADES DE MA CLASSIFICAÇÃO (REA) OBTIDAS COM O NOVO METODO DE CLASSIFICAÇÃO PROPOSTO E COMPARADAS COM AS OBTIDAS PELO SAS

	PARAMETROS DE ENTRADA								CORRELAÇÃO AMOSTRAL		TESTE DE HOMOGENEI
	pop. um				pop. dois						
	v_1	v_2	v_3	α	v_1	v_2	v_3	α	$\hat{\rho}^{(1)}$	$\hat{\rho}^{(2)}$	
13B	3	0.5	0.05	0	3	0.5	0.05	0	0.043	0.0622	Aceitar
14B	3	0.5	0.05	0	3	0.5	0.05	9	0.043	0.06887	Rejeitar
15B	3	0.5	0.05	0	3	0.5	0.05	29	0.043	0.9272	Rejeitar
16B	3	0.5	0.05	0	3	0.5	0.05	39	0.043	0.9427	Rejeitar
17B	3	0.5	0.05	0	3	0.5	0.05	45	0.043	0.9448	Rejeitar
18B	3	0.5	0.05	0	3	0.5	0.05	90	0.043	-0.0622	Rejeitar

	PROPOSTA DE DISCRIMINAÇÃO					SAS		
EXE	REA	n_{1m}	n_{2m}	Autovalores r_1	r_p	r_1/r_p	REA	n_{1m}	n_{2m}
13B	42	35	49	1.361	0.7258	1.87	45	45	45
14B	34	44	24	2.311	0.3505	6.59	38	44	32
15B	17.5	23	12	8.919	0.0879	101.48	18.5	23	14
16B	13	16	10	13.794	0.0567	243.28	13.5	16	11
17B	12.5	15	10	16.925	0.0462	366.34	12	15	09
18B	07	06	08	31.819	0.0245	1298.76	8.5	09	08

Fonte : Simulações com o Programa nº 2. (ver pág. 89).

QUADRO 3.1A.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 01A

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2				obs de Π_2 classificada em Π_1			
02	37	82	97	102	125	156	183
04	38	84		104	126	159	184
13	39	85		106	129	161	185
14	48	87		107	130	163	188
16	50	89		111	135	165	190
17	55	90		112	143	166	191
21	59	92		114	145	167	192
30	67	93		115	149	169	197
34	76	96		117	152	172	200
36	78	97		121	155	180	

PROPOSTA (B)			SAS (C)		
de Π_1 em Π_2	de Π_2 em Π_1		de Π_1 em Π_2	de Π_2 em Π_1	
41	101	160	5	99	131
44	105	162	8		
75	116	164	12		
95	119	174	25		
	127	175	35		
	138	176	45		
	140	179	46		
	144	195	56		
	147		62		
	148		79		

TOTAL DE MAL CLASSIFICADAS { PROPOSTA = A + B ; SAS = A + C

QUADRO 3.1A.2 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 04A

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2					obs de Π_2 classificada em Π_1	
02	20	43	65	82	103	146
03	22	46	71	87	104	150
06	23	47	73	90	115	153
08	28	48	74	93	118	159
10	31	50	76	97	120	161
11	33	52	77	100	124	166
12	35	54	78		131	184
14	36	56	79		134	186
16	37	61	80		137	196
17	38	63	81		139	

PROPOSTA (B)		SAS (C)	
de Π_1 em Π_2	de Π_2 em Π_1	de Π_1 em Π_2	de Π_2 em Π_1

TOTAL DE MAL CLASSIFICADAS $\begin{cases} \text{PROPOSTA} = A + B \\ \text{SAS} = A + C \end{cases}$

QUADRO 3.2A.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 05A

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2			obs de Π_2 classificada em Π_1		
02	55	100	101	133	174
14	58		111	142	181
17	63		113	144	182
27	74		116	148	183
28	76		117	152	190
37	77		119	155	194
38	78		121	158	200
47	80		126	162	
48	87		127	169	
50	98		128	170	

PROPOSTA (B)		SAS (C)	
de Π_1 em Π_2	de Π_2 em Π_1	de Π_1 em Π_2	de Π_2 em Π_1

TOTAL DE MAL CLASSIFICADAS $\begin{cases} \text{PROPOSTA} = A + B \\ \text{SAS} = A + C \end{cases}$

QUADRO 3.2A.2 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 07A

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2				obs de Π_2 classificada em Π_1			
02	37	82	98	102	125	159	184
04	38	84		104	126	161	185
13	39	85		106	129	163	188
14	48	87		107	130	165	190
16	50	89		111	135	166	191
17	55	90		112	143	167	192
21	59	92		114	145	169	197
30	67	93		115	149	172	200
34	76	96		117	152	180	
36	78	97		121	156	183	

PROPOSTA (B)			SAS (C)		
de Π_1 em Π_2	de Π_2 em Π_1		de Π_1 em Π_2		de Π_2 em Π_1
41	101	155	05	99	131
44	105	160	08		153
75	116	162	12		
95	119	164	25		
	127	174	35		
	138	175	45		
	140	176	46		
	144	179	56		
	147	195	62		
	148		79		

TOTAL DE MAL CLASSIFICADAS $\begin{cases} \text{PROPOSTA} = A + B \\ \text{SAS} = A + C \end{cases}$

QUADRO 3.4A.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 14A

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2	obs de Π_2 classificada em Π_1
04	158
13	
24	
49	
51	
66	
99	

PROPOSTA (B)		SAS (C)	
de Π_1 em Π_2	de Π_2 em Π_1	de Π_1 em Π_2	de Π_2 em Π_1

TOTAL DE MAL CLASSIFICADAS $\begin{cases} \text{PROPOSTA} = A + B \\ \text{SAS} = A + C \end{cases}$

QUADRO 3.1B.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 01B

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2			obs de Π_2 classificada em Π_1			
01	42	75	103	131	166	198
7	45	79	105	134	167	
12	55	84	106	137	181	
15	59	88	108	139	182	
18	60	89	109	143	184	
19	61	98	110	150	185	
23	62		113	151	187	
25	67		124	152	188	
33	68		126	160	191	
41	70		127	162	195	

PROPOSTA (B)			SAS (C)			
de Π_1 em Π_2		de Π_2 em Π_1	de Π_1 em Π_2		de Π_2 em Π_1	
8	101	149	6	64	107	174
14	102	156	24	69	115	176
22	114	165	28	73	123	197
29	116	172	38	77	128	199
32	120	175	39	82	144	
35	122	183	44	87	145	
48	125	189	46	91	146	
82	140	200	49	93	164	
83	141		54	99	171	
	147		63		173	

TOTAL DE MAL CLASSIFICADAS $\begin{cases} \text{PROPOSTA} = A + B \\ \text{SAS} = A + C \end{cases}$

QUADRO 3.2B.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 05B

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2		obs de Π_2 classificada em Π_1		
4	64	109	168	196
7	66	111	169	197
8	69	113	173	
10	70	126	178	
11	76	129	182	
28	77	132	184	
46	81	143	185	
48	91	148	188	
58		155	191	
63		163	194	

PROPOSTA (B)			SAS (C)		
de Π_1 em Π_2		de Π_2 em Π_1	de Π_1 em Π_2		de Π_2 em Π_1
6	78	104	21		118
13	87	119	27		156
17	88	121	44		157
23	95	131	80		158
26	96	152	89		160
31		153			183
33		159			187
43		166			192
45					
74					

TOTAL DE MAL CLASSIFICADAS { PROPOSTA = A + B ; SAS = A + C

QUADRO 3.2B.2 - OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 07B

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2		obs de Π_2 classificada em Π_1			
1	59	101	123	156	186
3	60	102	124	160	187
18	61	103	125	165	189
25	67	105	127	166	195
36	78	106	133	170	198
37	83	108	134	172	199
40	84	110	139	175	200
50	98	116	141	176	
51		120	144	180	
57		122	148	181	

PROPOSTA (B)		SAS (C)	
de Π_1 em Π_2	de Π_2 em Π_1	de Π_1 em Π_2	de Π_2 em Π_1
15	162	20	151
		41	
		65	

TOTAL DE MAL CLASSIFICADAS { PROPOSTA = A + B ; SAS = A + C

QUADRO 3.3B.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 09B

segundo a PROPOSTA e o SAS (A)

obs de Π_1 classificada em Π_2	obs de Π_2 classificada em Π_1
29	146
32	154
39	
41	
49	
66	
75	
90	
97	
98	

PROPOSTA (B)		SAS (C)	
de Π_1 em Π_2	de Π_2 em Π_1	de Π_1 em Π_2	de Π_2 em Π_1
87	119		
94			

TOTAL DE MAL CLASSIFICADAS $\begin{cases} \text{PROPOSTA} = A + B \\ \text{SAS} = A + C \end{cases}$

QUADRO 3.4B.1 OBSERVAÇÕES MAL CLASSIFICADAS NA EXECUÇÃO 18B

segundo a PROPOSTA e o SAS (A)	
obs de Π_1 classificada em Π_2	obs de Π_2 classificada em Π_1
15	119
45	144
51	146
57	154
59	159
78	166
	169
	182

PROPOSTA (B)		SAS (C)	
de Π_1 em Π_2	de Π_2 em Π_1	de Π_1 em Π_2	de Π_2 em Π_1
		12	
		25	
		67	

TOTAL DE MAL CLASSIFICADAS { PROPOSTA = A + B ; SAS = A + C

3.3 EXEMPLO DE APLICAÇÃO

3.3.1 INTRODUÇÃO

Como aplicação da técnica, usaram-se dados de discriminação de notas de 1000 Francos Suiços verdadeiros e falsos, já utilizadas em análise efetuada por *Flury e Riedwyl (1983)*, relacionados, nos quadro 3.C1 e 3.C2.

A seguir apresenta-se a descrição dos dados usados para construir a função de discriminação amostral, além de fazer a avaliação correspondente pelo cálculo da Razão de Erro Aparente. Faz-se também, uma comparação da estimativa da probabilidade de má classificação com a obtida pelo SAS.

Os programas usados foram: Programa nº 5 e Programa nº 6 escritos em linguagem FORTRAN e SAS respectivamente e estão no apêndice.

3.3.2 DESCRIÇÃO DOS DADOS

Têm-se as medições de seis variáveis em 200 notas de banco, 100 das quais são verdadeiras (vêm de Π_1) e as outras 100 são falsas (vêm de Π_2).

As variáveis medidas foram:

X_1: Comprimento das notas de banco.

X_2: Largura do lado esquerdo das notas.

X_3: Largura do lado direito.

X_4: Largura da margem inferior.

X_5: Largura da margem superior.

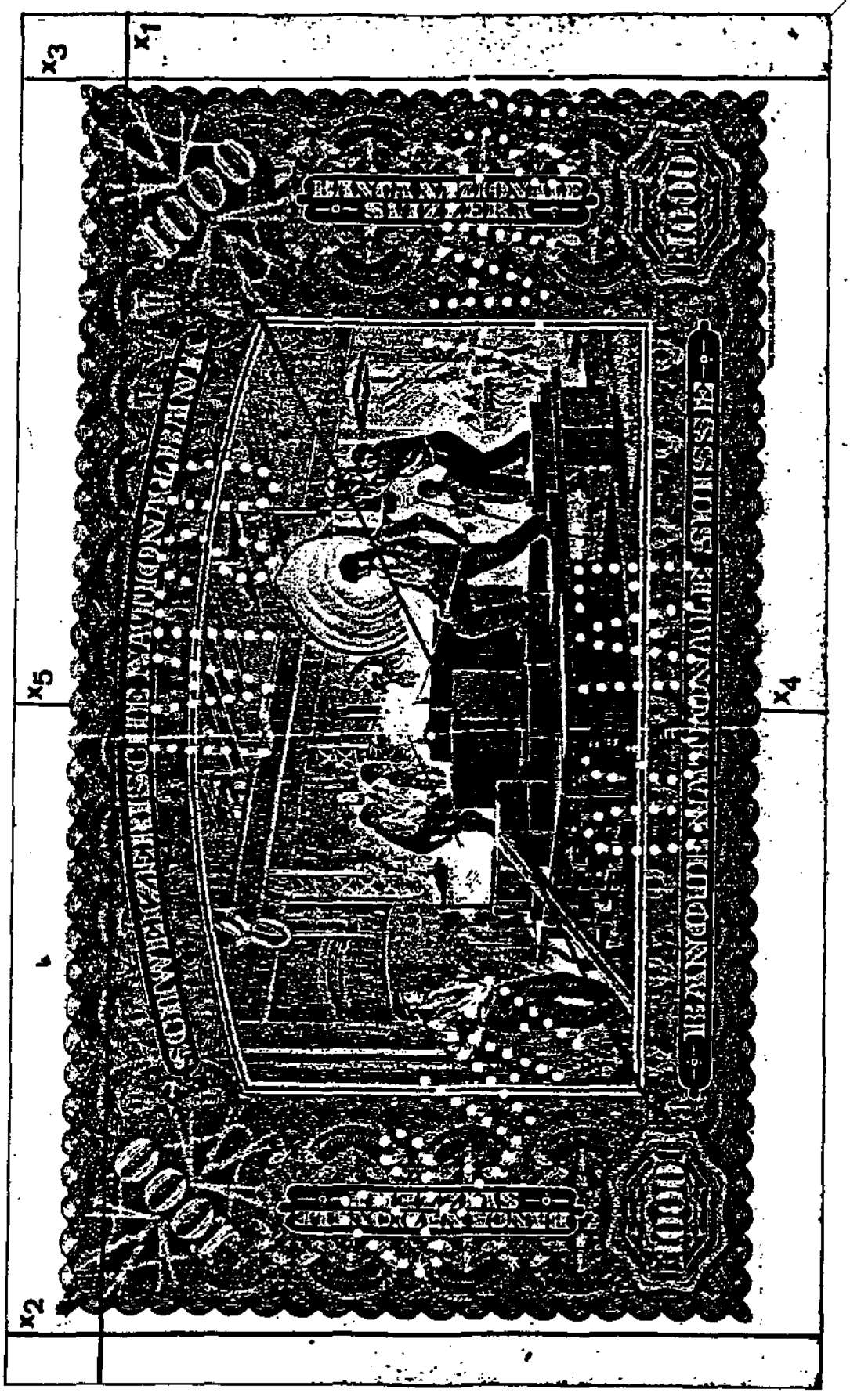
X_6: Comprimento da diagonal medida desde o canto inferior esquerdo até o canto superior direito.

onde: $\underset{\sim}{x}_i^{(g)} = (x_{i1}^{(g)}, x_{i2}^{(g)}, \ldots\ldots\ldots, x_{i6}^{(g)})$

$$g = 1,2 \qquad i = 1,2,\ldots\ldots, n_g$$
$$n_1 = n_2 = 100$$

é o vetor que contém as medições da i-ésima nota no g-ésimo grupo, resultando nas matrizes de dados $\mathbb{X}^{(1)}$ e $\mathbb{X}^{(2)}$ que ficam no quadro 3.C1 e 3.C2.

Abbildung 1



QUADRO 3.C1

MEDIDAS DE NOTAS DE BANCO

VERDADEIRAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
214.8	131.0	131.1	09.0	09.7	141.0	1001
214.6	129.7	129.7	08.1	09.5	141.7	1002
214.8	129.7	129.7	08.7	09.6	142.2	1003
214.8	129.7	129.6	07.5	10.4	142.0	1004
215.0	129.6	129.7	10.4	07.7	141.8	1005
215.7	130.8	130.5	09.0	10.1	141.4	1006
215.5	129.5	129.7	07.9	09.6	141.6	1007
214.5	129.6	129.2	07.2	10.0	141.7	1008
214.9	129.4	129.7	08.2	11.0	141.9	1009
215.2	130.4	130.3	09.2	10.0	140.7	1010
215.3	130.4	130.3	07.9	11.7	141.8	1011
215.1	129.5	129.6	07.7	10.5	142.2	1012
215.2	130.8	129.6	07.9	10.8	141.4	1013
214.7	129.7	129.7	07.7	10.9	141.7	1014
215.1	129.9	129.7	07.7	10.8	141.8	1015
214.5	129.8	129.8	09.3	08.5	141.6	1016
214.6	129.9	130.1	08.2	09.8	141.7	1017
215.0	129.9	129.7	09.0	09.0	141.9	1018
215.2	129.6	129.6	07.4	11.5	141.5	1019
214.7	130.2	129.9	08.6	10.0	141.9	1020
215.0	129.9	129.3	08.4	10.0	141.4	1021
215.6	130.5	130.0	08.1	10.3	141.6	1022
215.3	130.6	130.0	08.4	10.8	141.5	1023
215.7	130.2	130.0	08.7	10.0	141.6	1024
215.1	129.7	129.9	07.4	10.8	141.1	1025
215.3	130.4	130.4	08.0	11.0	142.3	1026
215.5	130.2	130.1	08.1	09.8	142.4	1027
215.1	130.3	130.3	09.8	09.5	141.9	1028
215.1	130.0	130.0	07.4	10.5	141.8	1029

QUADRO 3. C1 (cont.)

MEDIDAS DE NOTAS DE BANCO

VERDADEIRAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
214.8	129.7	129.3	08.3	09.0	142.0	1030
215.2	130.1	129.8	07.9	10.7	141.8	1031
214.8	139.7	129.7	08.6	09.1	142.3	1032
215.0	130.0	129.6	07.7	10.5	140.7	1033
215.6	130.4	130.1	08.4	10.3	141.0	1034
215.9	130.4	130.0	08.9	10.6	141.4	1035
214.6	130.2	130.2	09.4	09.7	141.8	1036
215.5	130.3	130.0	08.4	09.7	141.8	1037
215.3	129.9	129.4	07.9	10.0	142.0	1038
215.3	130.3	130.1	08.5	09.3	142.1	1039
213.9	130.3	129.0	08.1	09.7	141.3	1040
214.4	129.8	129.2	08.9	09.4	142.3	1041
214.8	130.1	129.6	08.8	09.9	140.9	1042
214.9	129.6	129.4	09.3	09.3	141.7	1043
214.9	130.4	129.7	09.0	09.8	140.9	1044
214.8	129.4	129.1	08.2	10.2	141.0	1045
214.3	129.5	129.4	08.3	10.2	141.8	1046
214.8	129.9	129.7	08.3	10.2	141.5	1047
214.8	129.9	129.7	07.3	10.9	142.0	1048
214.6	129.7	129.8	07.9	10.3	141.1	1049
214.5	129.0	129.6	07.8	09.8	142.0	1050
214.6	129.8	129.4	07.2	10.0	141.3	1051
215.3	130.6	130.0	09.5	09.7	141.1	1052
214.5	130.1	130.0	07.8	10.9	140.9	1053
215.4	130.2	130.2	07.6	10.9	141.6	1054
214.5	129.4	129.5	07.9	10.0	141.4	1055
215.2	129.7	129.4	09.2	09.4	142.0	1056
215.7	130.0	129.4	09.2	10.4	141.2	1057
215.0	129.6	129.4	08.8	09.0	141.1	1058

QUADRO 3.C1 (cont.)

MEDIDAS DE NOTAS DE BANCO

VERDADEIRAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
215.1	130.1	129.9	07.9	11.0	141.3	1059
215.1	130.0	129.8	08.2	10.3	141.4	1060
215.1	129.6	129.3	08.3	09.9	141.6	1061
215.3	129.7	129.4	07.5	10.5	141.5	1062
215.4	129.8	129.4	08.0	10.6	141.5	1063
214.5	130.0	129.5	08.0	10.8	141.4	1064
215.0	130.0	129.8	08.6	10.6	141.5	1065
215.2	130.6	130.0	08.8	10.6	140.8	1066
214.6	129.5	129.2	07.7	10.3	141.3	1067
214.8	129.7	129.3	09.1	09.5	141.5	1068
215.1	129.6	129.8	08.6	09.8	141.8	1069
214.9	130.2	130.2	08.0	11.2	139.6	1070
213.8	129.8	129.5	08.4	11.1	140.9	1071
215.2	129.9	129.5	08.2	10.3	141.4	1072
215.0	129.6	130.2	08.7	10.0	141.2	1073
214.4	129.9	129.6	07.5	10.5	141.8	1074
215.2	129.9	129.7	07.2	10.6	142.1	1075
214.1	129.6	129.3	07.6	10.7	141.7	1076
214.9	129.9	130.1	08.8	10.0	141.2	1077
214.6	129.8	129.4	07.4	10.6	141.0	1078
215.2	130.5	129.8	07.9	10.9	140.9	1079
214.6	129.9	129.4	07.9	10.0	141.8	1080
215.1	129.7	129.7	08.6	10.3	140.6	1081
214.9	129.8	129.6	07.5	10.3	141.0	1082
215.2	129.7	129.1	09.0	09.7	141.9	1083
215.2	130.1	129.9	07.9	10.8	141.3	1084
215.4	130.7	130.2	09.0	11.1	141.2	1085
215.1	129.9	129.6	08.9	10.2	141.5	1086
215.2	129.9	129.7	08.7	09.5	141.6	1087

QUADRO 3.C1 (cont.)

MEDIDAS DE NOTAS DE BANCO

VERDADEIRAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
215.0	129.6	129.2	08.4	10.2	142.1	1088
214.9	130.3	129.9	07.4	11.2	141.5	1089
215.0	129.9	129.7	08.0	10.5	142.0	1090
214.7	129.7	129.3	08.6	09.6	141.6	1091
215.4	130.0	129.9	08.5	09.7	141.4	1092
214.9	129.4	129.5	08.2	09.9	141.5	1093
214.5	129.5	129.3	07.4	10.7	141.5	1094
214.7	129.6	129.5	08.3	10.0	142.0	1095
215.6	129.9	129.9	09.0	09.5	141.7	1096
215.0	130.4	130.0	09.1	10.2	141.1	1097
214.4	129.7	129.5	08.0	10.3	141.2	1098
215.1	130.0	129.8	09.1	10.2	141.5	1099
214.7	130.0	129.4	07.8	10.0	141.2	1100

Fonte : *Flury (1983).*

G : Grupo de notas verdadeiras (1).

OBS : Número de observação dentro de cada grupo.

QUADRO 3.C2

MEDIDAS DE NOTAS DE BANCO

FALSAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
214.4	130.1	130.3	09.7	11.7	139.8	2001
214.9	130.5	130.2	11.0	11.5	139.5	2002
214.9	130.3	130.1	08.7	11.7	140.2	2003
215.0	130.4	130.6	09.9	10.9	140.3	2004
214.7	130.2	130.3	11.8	10.9	139.7	2005
215.0	130.2	130.2	10.6	10.7	139.9	2006
215.3	130.3	130.1	09.3	12.1	140.2	2007
214.8	130.1	130.4	09.6	11.5	139.9	2008
215.0	130.2	129.9	10.0	11.9	139.4	2009
215.2	130.6	130.8	10.4	11.2	140.3	2010
215.2	130.4	130.3	08.0	11.5	139.2	2011
215.1	130.5	130.3	10.6	11.5	140.1	2012
215.4	130.7	131.1	09.7	11.8	140.6	2013
214.9	130.4	129.9	11.4	11.0	139.9	2014
215.1	130.3	130.0	10.6	10.8	139.7	2015
215.8	130.4	130.0	08.2	11.2	139.2	2016
214.7	130.6	130.1	11.8	10.5	139.8	2017
214.7	130.4	130.1	12.1	10.4	139.9	2018
214.8	130.5	130.2	11.0	11.0	140.0	2019
214.4	130.2	129.9	10.1	12.0	139.2	2020
214.8	130.3	130.4	10.1	12.1	139.6	2021
215.1	130.6	130.3	12.3	10.2	139.6	2022
215.3	130.8	131.1	11.6	10.6	140.2	2023
215.1	130.7	130.4	10.5	11.2	139.7	2024
214.7	130.5	130.5	09.9	10.3	140.1	2025
214.9	130.0	130.3	10.2	11.4	139.6	2026
215.0	130.4	130.4	09.4	11.6	140.2	2027
215.5	130.7	130.3	10.2	11.8	140.0	2028
215.1	130.2	130.2	10.1	11.3	140.3	2029

QUADRO 3.C2 (cont.)

MEDIDAS DE NOTAS DE BANCO

FALSAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
214.5	130.2	130.6	09.8	12.1	139.9	2030
214.3	130.2	130.0	10.7	10.5	139.8	2031
214.5	130.2	139.8	12.3	11.2	139.2	2032
214.9	130.5	130.2	10.6	11.5	139.9	2033
214.6	130.2	130.4	10.5	11.8	139.7	2034
214.2	130.0	130.2	11.0	11.2	139.5	2035
214.8	130.1	130.1	11.9	11.1	139.5	2036
214.6	129.8	130.2	10.7	11.1	139.4	2037
214.9	130.7	130.3	09.3	11.2	138.3	2038
214.6	130.4	130.4	11.3	10.8	139.8	2039
214.5	130.5	130.2	11.8	10.2	139.6	2040
214.8	130.2	130.3	10.0	11.9	139.3	2041
214.7	130.0	129.4	10.2	11.0	139.2	2042
214.6	130.2	130.4	11.2	10.7	139.9	2043
215.0	130.5	130.4	10.6	11.1	139.9	2044
214.5	129.8	129.8	11.4	10.0	139.3	2045
214.9	130.6	130.4	11.9	10.5	139.8	2046
215.0	130.5	130.4	11.4	10.7	139.9	2047
215.3	130.6	130.3	09.3	11.3	138.1	2048
214.7	130.2	130.1	10.7	11.0	139.4	2049
214.9	129.9	130.0	09.9	12.3	139.4	2050
214.9	130.3	129.9	11.9	10.6	139.8	2051
214.6	129.9	129.7	11.9	10.1	139.0	2052
214.6	129.7	129.3	10.4	11.0	139.3	2053
214.5	130.1	130.1	12.1	10.3	139.4	2054
214.5	130.3	130.0	11.0	11.5	139.5	2055
215.1	130.0	130.3	11.6	10.5	139.7	2056
214.2	129.7	129.6	10.3	11.4	139.5	2057
214.4	130.1	130.0	11.3	10.7	139.2	2058

MEDIDAS DE NOTAS DE BANCO

FALSAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
214.8	130.4	130.6	12.5	10.0	139.3	2059
214.6	130.6	130.1	08.1	12.1	137.9	2060
215.6	130.1	129.7	07.4	12.2	138.4	2061
214.9	130.5	130.1	09.9	10.2	138.1	2062
214.6	130.1	130.0	11.5	10.6	139.5	2063
214.7	130.1	130.2	11.6	10.9	139.1	2064
214.3	130.3	130.0	11.4	10.5	139.8	2065
215.1	130.3	130.6	10.3	12.0	139.7	2066
216.3	130.7	130.4	10.0	10.1	138.8	2067
215.6	130.4	130.1	09.6	11.2	138.6	2068
214.8	129.9	129.8	09.6	12.0	139.6	2069
214.9	130.0	129.9.	11.4	10.9	139.7	2070
213.9	130.7	130.5	08.7	11.5	137.8	2071
214.2	130.6	130.4	12.0	10.2	139.6	2072
214.8	130.5	130.3	11.8	10.5	139.4	2073
214.8	129.6	130.0	10.4	11.6	139.2	2074
214.8	130.1	130.0	11.4	10.5	139.6	2075
214.9	130.4	130.2	11.9	10.7	139.0	2076
214.3	130.1	130.1	11.6	10.5	139.7	2077
214.5	130.4	130.0	09.9	12.0	139.6	2078
214.8	130.5	130.3	10.2	12.1	139.1	2079
214.5	130.2	130.4	08.2	11.8	137.8	2080
215.0	130.4	130.1	11.4	10.7	139.1	2081
214.8	130.6	130.6	08.0	11.4	138.7	2082
215.0	130.5	130.1	11.0	11.4	139.3	2083
214.6	130.5	130.4	10.1	11.4	139.3	2084
214.7	130.2	130.1	10.7	11.1	139.5	2085
214.7	130.4	130.0	11.5	10.7	139.4	2086
214.5	130.4	130.0	08.0	12.2	138.5	2087

QUADRO 3.C2 (cont.)

MEDIDAS DE NOTAS DE BANCO

FALSAS

X_1	X_2	X_3	X_4	X_5	X_6	G-OBS
214.8	130.0	129.7	11.4	10.6	139.2	2088
214.8	129.9	130.2	09.6	11.9	139.4	2089
214.6	130.3	130.2	12.0	09.1	139.2	2090
215.1	130.2	129.8	10.2	12.0	139.4	2091
215.4	130.5	130.6	08.8	11.0	138.6	2092
214.7	130.3	130.2	10.8	11.1	139.2	2093
215.0	130.5	130.3	09.6	11.0	138.5	2094
214.9	130.3	130.5	11.6	10.6	139.8	2095
215.0	130.4	130.3	09.9	12.1	139.6	2096
215.1	130.3	129.9	10.3	11.5	139.7	2097
214.8	130.3	130.4	10.6	11.1	140.0	2098
214.7	130.7	130.8	11.2	11.2	139.4	2099
214.3	129.9	129.9	10.2	11.5	139.6	2100

FONTE : *Flury (1983)*

G : Grupo de notas falsas (2).

OBS : Número de observações dentro de cada grupo.

3.3.3 RESULTADOS OBTIDOS.

Consideramos as duas amostras onde os seus parâmetros populacionais são desconhecidos. As medidas descritivas que são as estimativas dos parâmetros $\underset{\sim}{M}_1$, $\underset{\sim}{M}_2$, V_1 e V_2 obtidas através dos dados do quadro 3.C1 e 3.C2 são as seguintes:

$$\underset{\sim}{\bar{X}}^{(1)} = \begin{bmatrix} 214.9690 \\ 129.9430 \\ 129.7200 \\ 8.3050 \\ 10.1710 \\ 141.5150 \end{bmatrix} \qquad (14)$$

$$\underset{\sim}{\bar{X}}^{(2)} = \begin{bmatrix} 214.8230 \\ 130.3000 \\ 130.1930 \\ 10.5300 \\ 11.1330 \\ 139.4500 \end{bmatrix}$$

Matriz de covariâncias da amostra um: (S_1)

$$\begin{bmatrix} 0.150372 & 0.058198 & 0.057495 & 0.057130 & 0.014246 & 0.005045 \\ 0.058198 & 0.132603 & 0.085952 & 0.056648 & 0.048035 & -0.044810 \\ 0.057495 & 0.085952 & 0.126420 & 0.058179 & 0.029673 & -0.023472 \\ 0.057130 & 0.056648 & 0.058179 & 0.413206 & -0.260460 & 0.000628 \\ 0.014246 & 0.048035 & 0.029673 & -0.260460 & 0.415009 & -0.076023 \\ 0.005045 & -0.044810 & -0.023472 & 0.000628 & -0.076023 & 0.200688 \end{bmatrix}$$

Matriz de covariâncias da amostra dois: (S_2)

$$\begin{bmatrix} 0.124498 & 0.031602 & 0.024045 & -0.100596 & 0.019448 & 0.011619 \\ 0.031602 & 0.064790 & 0.046783 & -0.024041 & -0.011918 & -0.004987 \\ 0.024045 & 0.046783 & 0.088727 & -0.018574 & 0.000134 & 0.034221 \\ -0.100596 & -0.024041 & -0.018574 & 1.281313 & -0.490192 & 0.238489 \\ 0.019448 & -0.011918 & 0.000134 & -0.490192 & 0.404456 & -0.022071 \\ 0.011619 & -0.004987 & 0.034221 & 0.238489 & -0.022071 & 0.311162 \end{bmatrix}$$

Com a sub-rotina F02AEF obtemos os autovalores e autovetores da matriz $S_1^{-1}S_2$ que são os seguintes:

$$r_6 = 0.283892 \qquad r_5 = 0.545285$$
$$r_4 = 0.906826 \qquad r_3 = 1.052638$$
$$r_2 = 1.678315 \qquad r_1 = 6.120406$$

(15)

$\underset{\sim}{b}_6$	$\underset{\sim}{b}_5$	$\underset{\sim}{b}_4$	$\underset{\sim}{b}_3$	$\underset{\sim}{b}_2$	$\underset{\sim}{b}_1$
-0.390511	-1.340457	2.002322	-1.389171	-0.062965	0.977541
-1.194453	3.372635	1.341897	1.030241	0.109882	0.660427
-0.361893	-2.581479	-1.650547	1.910784	1.334965	0.426138
-0.510676	-0.248453	-0.043216	-0.362239	-0.469081	-2.237451
-0.836594	0.023098	-0.746921	-1.320852	0.523342	-1.538391
0.587390	0.626593	0.580347	0.154213	1.934701	-1.059888

As combinações com máximo e mínimo quociente de variâncias são, respectivamente:

$$Y_{max}^{(g)} = 0.9775X_1^{(g)} + 0.6604X_2^{(g)} + 0.4261X_3^{(g)} - 2.2374X_4^{(g)} - 1.5384X_5^{(g)} - 1.0599X_6^{(g)}$$

$$Y_{min}^{(g)} = -0.3905X_1^{(g)} - 1.1944X_2^{(g)} - 0.3619X_3^{(g)} - 0.5107X_4^{(g)} - 0.8366X_5^{(g)} + 0.5874X_6^{(g)} \qquad (16)$$

Com as quais obtém-se a amostra de dados transformados bivariados:

$$\underset{\sim}{Y}_i^{(g)} = \begin{bmatrix} Y_{imax}^{(g)} \\ Y_{imin}^{(g)} \end{bmatrix} \qquad i = 1,\ldots,n_g \qquad g = 1,2 \qquad n_1 = n_2$$

apresentados no quadro 3.D1 e 3.D2 respectivamente.

QUADRO 3. D1

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

VERDADEIRAS (1)

Y_{max}	Y_{min}	GRUPO
167.854767	-217.688538	1
167.783569	-214.512863	1
165.952850	-214.687378	1
167.576447	-214.825089	1
165.625565	-214.159653	1
167.307495	-217.683624	1
168.930963	-214.665741	1
167.574371	-214.717728	1
165.135437	-215.460205	1
166.917587	-217.367889	1
166.142899	-217.519119	1
166.924316	-214.771698	1
168.031525	-217.304062	1
166.622559	-215.518875	1
167.193527	-215.771591	1
166.753922	-214.464447	1
167.400879	-215.198563	1
166.850220	-214.831818	1
166.962891	-216.024719	1
166.196899	-215.777710	1
167.013763	-215.510956	1
168.292557	-216.695526	1
166.730927	-217.328094	1
167.311234	-216.431671	1
168.559814	-215.863037	1
166.508682	-216.727081	1
166.170654	-215.854691	1
164.908661	-216.392639	1

QUADRO 3.D1 (cont.)

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

VERDADEIRAS (1)

Y_{max}	Y_{min}	GRUPO
168.520157	-215.595413	1
167.812378	-213.953873	1
167.172318	-216.104233	1
166.839798	-214.159271	1
168.746628	-216.210968	1
168.233826	-217.117950	1
166.480316	-217.470245	1
165.004501	-216.063522	1
168.102554	-215.951370	1
167.832458	-215.056503	1
168.023331	-215.449692	1
167.313538	-215.105148	1
165.168991	-214.345734	1
166.867050	-216.194656	1
165.505707	-214.847549	1
166.911850	-216.646652	1
166.966614	-215.063385	1
165.600052	-214.677261	1
166.798798	-215.635101	1
167.429443	-215.416336	1
167.678896	-215.468628	1
167.072693	-213.523056	1
169.390259	-214.717361	1
166.385925	-217.204559	1
167.443253	-216.548111	1
168.179855	-216.578064	1
167.398743	-214.535492	1
165.616943	-214.940506	1

QUADRO 3.D1 (cont.)

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

VERDADEIRAS (1)

Y_{max}	Y_{min}	GRUPO
165.613373	-216.800583	1
167.819672	-214.732727	1
167.185638	-216.646042	1
167.376617	-215.999237	1
167.078995	-214.939453	1
168.356079	-215.325317	1
167.247314	-215.822815	1
166.340546	-215.972504	1
165.816376	-216.356689	1
166.787781	-217.737167	1
167.526642	-214.793015	1
165.783142	-215.074402	1
166.562698	-215.072464	1
168.456366	-218.012863	1
164.697617	-216.209213	1
167.280487	-215.810272	1
168.739929	-215.749008	1
167.375626	-215.108932	1
168.399704	-215.211746	1
166.330994	-214.802017	1
166.573914	-216.083176	1
168.337708	-215.497681	1
168.087218	-217.277939	1
167.360153	-214.900620	1
167.088776	-216.278900	1
168.953949	-215.487335	1
165.581055	-215.039520	1
167.591064	-216.517776	1

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

VERDADEIRAS (1)

Y_{max}	Y_{min}	GRUPO
165.493927	-218.292557	1
165.707001	-216.022491	1
167.265701	-215.351227	1
165.723434	-214.872574	1
167.721283	-216.601257	1
166.674103	-215.517288	1
166.566266	-214.804932	1
167.964294	-215.803833	1
167.166382	-214.702515	1
167.315399	-214.854095	1
166.195435	-214.704330	1
166.964752	-215.674271	1
166.214172	-217.171082	1
167.025818	-215.274307	1
165.410736	-216.316467	1
168.383606	-215.360489	1

Fonte : Dados da matriz 3.C1 transformados segundo a proposta (16).

QUADRO 3. D2

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

FALSAS (2)

Y_{max}	Y_{min}	GRUPO
163.157333	-218.903381	2
161.584656	-220.212997	2
165.506454	-218.519470	2
164.323105	-218.743744	2
160.154694	-219.601868	2
163.186005	-218.785233	2
163.939651	-219.316742	2
163.568939	-218.920761	2
163.084503	-219.667847	2
163.155685	-219.639435	2
168.884766	-218.891037	2
168.086778	-219.770584	2
163.870270	-219.913818	2
160.841034	-219.536026	2
163.322693	-219.072510	2
169.064240	-218.750778	2
160.842972	-219.613922	2
160.087524	-219.385818	2
161.726151	-219.461960	2
162.332367	-219.685715	2
162.424698	-219.991013	2
160.874008	-219.964340	2
161.857468	-220.195587	2
163.365692	-219.978577	2
165.188278	-218.325363	2
163.134811	-219.101028	2
164.385696	-219.060364	2
163.144333	-220.271133	2

QUADRO 3.D2 (cont.)

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

FALSAS (2)

Y_{max}	Y_{min}	GRUPO
163.055496	-218.835907	2
162.503906	-219.497391	2
162.606400	-218.381989	2
158.695831	-220.142838	2
162.055666	-219.773788	2
161.623688	-219.688019	2
161.031677	-219.091446	2
159.781769	-219.785019	2
162.221680	-218.830627	2
167.296356	-220.073807	2
161.398209	-219.440155	2
161.297562	-219.319046	2
163.165405	-219.793228	2
163.595123	-218.597549	2
161.537720	-219.007797	2
162.853973	-219.550583	2
162.185455	-218.142776	2
160.942612	-219.851654	2
161.679382	-219.624496	2
167.679474	-220.311737	2
162.694794	-219.227631	2
162.439606	-219.590210	2
160.377609	-219.396027	2
161.352051	-218.780334	2
162.703125	-218.207352	2
160.377686	-219.159424	2
160.976318	-219.745529	2
161.476501	-219.082428	2

QUADRO 3.D2 (cont.)

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

FALSAS (2)

Y_{max}	Y_{min}	GRUPO
161.836365	-218.325790	2
161.623932	-219.127792	2
160.754700	-219.827896	2
168.576157	-220.139877	2
169.965501	-219.220901	2
167.486938	-219.349854	2
161.207794	-219.048141	2
161.129471	-219.696533	2
161.106247	-218.858917	2
162.403534	-220.140320	2
168.303589	-219.800217	2
166.708069	-219.893433	2
163.177383	-218.957108	2
160.942657	-219.092041	2
167.814957	-219.993896	2
160.708054	-219.495865	2
161.383911	-219.840851	2
162.313873	-218.980057	2
162.676966	-218.932755	2
161.265564	-220.177643	2
160.675293	-218.817108	2
162.628342	-219.662796	2
162.820358	-220.536483	2
168.685867	-219.590424	2
162.333389	-219.666394	2
169.437469	-219.292343	2
162.005585	-220.219680	2
163.756104	-219.742386	2

QUADRO 3.D2 (cont.)

TRABALHO DE APLICAÇÃO

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

FALSAS (2)

Y_{max}	Y_{min}	GRUPO
162.434967	-219.252548	2
161.455811	-219.587921	2
167.737701	-219.505920	2
161.751129	-219.023331	2
163.713684	-219.135681	2
161.365723	-218.893539	2
162.538300	-219.856506	2
168.889343	-219.539856	2
162.637848	-219.635483	2
166.686493	-219.742371	2
161.304520	-219.459930	2
163.096160	-220.050247	2
162.874451	-219.468689	2
162.420395	-219.174850	2
161.896881	-220.500793	2
162.157974	-218.686172	2

Fonte : Dados da matriz 3.C2 transformados segundo a proposta (16).

Desta maneira, sob a suposição de normalidade das variáveis originais, obtemos as variáveis normais bivariadas, cujos parâmetros foram apresentados na secção (2.3.1) do capítulo 2.

Com os dados amostrais transformados que estão nos quadros 3.D1 e 3.D2, estimam-se os parâmentros $\underset{\sim}{\mu}^{(1)}$, $\underset{\sim}{\mu}^{(2)}$, Σ_1, Σ_2, por:

$$\hat{\underset{\sim}{\mu}}^{(1)} = \bar{\underset{\sim}{Y}}^{(1)} = \begin{bmatrix} 167.0185 \\ 215.7292 \end{bmatrix} \qquad \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\hat{\underset{\sim}{\mu}}^{(2)} = \bar{\underset{\sim}{Y}}^{(2)} = \begin{bmatrix} 163.0435 \\ -219.4238 \end{bmatrix} \qquad \hat{\Sigma}_2 = \begin{bmatrix} 6.1204 & 0 \\ 0 & 0.2839 \end{bmatrix}$$

$$\widehat{\ln(\lambda_1 \lambda_p)} = \ln(r_1 r_6) = 0.5525$$

A função de discriminação amostral é portanto:

$$D(\underset{\sim}{y}) = \underset{\sim}{y}' \left(\hat{\Sigma}_1^{-1} - \hat{\Sigma}_2^{-1} \right) \underset{\sim}{y} - 2\underset{\sim}{y}' \left(\hat{\Sigma}_1^{-1} \bar{\underset{\sim}{Y}}^{(1)} - \hat{\Sigma}_2^{-1} \bar{\underset{\sim}{Y}}^{(2)} \right) - 2\hat{d}$$

com:

$$\hat{d} = \frac{1}{2} \ln \left(r_1 r_2 \right) - \frac{1}{2} \left(\bar{\underset{\sim}{Y}}^{(1)\prime} \bar{\underset{\sim}{Y}}^{(2)} - \bar{\underset{\sim}{Y}}^{(2)\prime} \hat{\Sigma}_2^{-1} \bar{\underset{\sim}{Y}}^{(2)} \right) \qquad (17)$$

Supondo probabilidades a priori e custos de má classificação iguais, tem-se:

$$D(\underset{\sim}{y}) = 0.8366 Y_{max}^2 - 2.5224 Y_{min}^2 - 280.7582 Y_{max} - 1114.366 Y_{min} - 99504 \qquad (18)$$

---

função que é proposta para classificar uma nova nota com medições $\underset{\sim}{x} \in \mathbb{R}^6$, como verdadeira ou falsa, onde a relação entre $\underset{\sim}{x}$ e χ é:

$$\chi = \begin{pmatrix} \underset{\sim}{b}_1' & \underset{\sim}{x} \\ \underset{\sim}{b}_p' & \underset{\sim}{x} \end{pmatrix} \qquad \begin{matrix} \underset{\sim}{b}_1, \underset{\sim}{b}_p \in \mathbb{R}^6 \\ \chi \in \mathbb{R}^2 \end{matrix}$$

A avaliação de (18), cria Razão de Erro Aparente (REA) para estimar a Probabilidade Total de Má Classificação, já apresentado na secção (2.3.3) do capítulo 2, onde cada observação transformada dos quadros 3.D1 e 3.D2 foi avaliada em (18), isto é foram re-classificadas pelo método proposto.

Paralelamente, foi feita a discriminação pelo "PROC DISCRIM" do SAS e têm-se os seguintes resultados:

	DECISÃO ESTATÍSTICA				
GRUPO	MÉTODO PROPOSTO		SAS		TOTAIS
VERDADEIRO	Π_1	Π_2	Π_1	Π_2	
Π_1	100	0	99	1	100
Π_2	0	100	0	100	100
TOTAIS	100	100	99	101	200

(19)

Segundo (19) pode-se dizer que para os conjuntos de dados dos quadros 3.C1 e 3.C2 , este método é tão bom quanto a Regra de Classificação Quadrática de Anderson com a que classifica o SAS, sendo que o método proposto leva vantagem, já que os conjuntos de notas de banco de dimensão 6 se reduzem à dimensão 2, onde a interpretação é mais fácil, além da representação gráfica.

Para maior detalhes ver os quadros A2, A3, A4, e A5 do apêndice.

Pode-se perguntar se os resultados serão semelhantes se as transformações são obtidas com os autovetores da matriz $S_2^{-1}S_1$ em lugar dos autovetores de $S_1^{-1}S_2$. Aplicando o método para $S_2^{-1}S_1$, tem-se:

$r_6 = 0.163388$ $\quad$ $r_5 = 0.595836$

$r_4 = 0.949989$ $\quad$ $r_3 = 1.102748$

$r_2 = 1.833902$ $\quad$ $r_1 = 3.522463$

e autovetores amostrais associados:

$\underset{\sim}{b}_1$	$\underset{\sim}{b}_2$	$\underset{\sim}{b}_3$	$\underset{\sim}{b}_4$	$\underset{\sim}{b}_5$	$\underset{\sim}{b}_6$
-0.395134	0.048603	1.353993	-2.102674	-1.815269	-0.732920
-0.266953	-0.084819	-1.004152	-1.409150	4.567278	-2.241777
-0.172250	-1.030465	-1.862397	1.733269	-3.495882	-0.679210
0.904406	0.362086	0.353066	0.045382	-0.336460	-0.998486
0.621837	-0.403970	1.287403	0.784356	0.031280	-1.570138
0.429420	-1.493403	-0.150308	-0.609433	0.848543	1.102427

Com as variáveis transformadas:

$$Y_{max}^{(g)} = -0.7329X_1^{(g)} - 2.2418X_2^{(g)} - 0.6792X_3^{(g)} + 0.9585X_4^{(g)} + 1.5701X_5^{(g)} + 1.1024X_6^{(g)}$$

$$Y_{min}^{(g)} = -0.3951X_1^{(g)} - 0.2669X_2^{(g)} - 0.1723X_3^{(g)} - 0.9044X_4^{(g)} - 0.6218X_5^{(g)} + 0.4284X_6^{(g)}$$

Neste caso a tabela de classificação foi:

GRUPO VERDADEIRO	DECISÃO ESTATISTICA: MÉTODO PROPOSTO Π_1	MÉTODO PROPOSTO Π_2	SAS Π_1	SAS Π_2	TOTAIS
Π_1	99	1	99	1	100
Π_2	1	99	0	100	100
TOTAIS	100	100	99	101	200

onde a estimativa da probabilidade total de má classificação foi de 1%.

Portanto para o conjunto de dados do exemplo de aplicação, vê-se que os resultados obtidos com os autovetores da matriz $S_2^{-1}S_1$ são bem semelhantes aos obtidos com os autovetores da matriz $S_1^{-1}S_2$.

CAPITULO 4

CONCLUSÕES DOS RESULTADOS EXPERIMENTAIS E DO EXEMPLO DE APLICAÇÃO

Nos problemas que envolvem regras de classificação é de fundamental importância avaliar o erro de má classificação fornecida por uma função discriminante. Neste estudo além de apresentá-la(função discriminante), ela também foi avaliada mediante simulações de Monte Carlo e um exemplo de aplicação, segundo a apresentação do capítulo anterior.

Agora as conclusões de tais simulações para os casos em duas e três dimensões separadamente, e também as correspondentes ao exemplo proposto.

Com p = 2 observou-se que:

1).-Quando aceitamos a hipótese de homogeneidade das matrizes de covariâncias V_1 e V_2 ;

Segundo o novo critério de discriminação, os autovalores da matriz $S_1^{-1}S_2$ estão próximos do valor "um", e as estimativas das probabilidades de má classificação, ou seja, as Razões de Erro Aparente estão próximos de 0.5, como pode ser observado nas colunas "PROPOSTA DE DISCRIMINAÇÃO" das execuções 01A, 02A, 07A.

Na situação exposta para obter os resultados das colunas "SAS", o pacote estatístico SAS, usa a Função Discriminante de Fisher, com a qual também se obtém probabilidades de má classificação próximos de 0.5.

Apesar das regiões de má classificação geradas com os dois métodos não serem exatamente as mesmas, como pode ser observado nos quadros 3.1A.1, 3.2A.2 em geral há uma alta porcentagem de elementos comuns mal classificados em ambos os métodos, a qual fica expressa na parte (A) dos quadros antes assinalados e com quase o mesmo número de elementos nas partes (B) e (C), portanto as Razões de Erro Aparente são muito parecidas.

Cabe mencionar ainda que, quando a hipótese de homogeneidade é aceita, a tranformação proposta muda a estrutura inicial dos dados em duas distribuições normais bivariadas com matrizes de covariâncias iguais à identidade, como pode ser visto nos gráficos 1 e 2.

2D.-Quando a hipótese de homogeneidade das matrizes é rejeitada;

A referida transformação muda a estrutura inicial dos dados da seguinte maneira: O primeiro grupo transforma-se num conjunto com matriz de covariâncias identidade e o segundo, se transforma em um conjunto com matriz de covariâncias diagonal, cuja primeira variável tem como variância o máximo autovalor e a segunda, o mínimo autovalor da matriz $V_1^{-1}V_2$. (ver gráfico 3 e 4).

Segundo a regra proposta, as estimativas das probabilidades de má classificação decrescem aproximadamente de 32% no quadro 3.1A até 1% no quadro 3.4A, sendo que temos nos quadros 3.2A e 3.3A, 26% e 10% respectivamente, dependendo da relação existente entre os autovalores máximo e mínimo que em geral afastam-se da "unidade".

Nesta situação, para obter os resultados das colunas "SAS" dos quadros apresentados, o sistema estatístico SAS usa a Funçao Quadrática Amostral de Anderson, e observou-se então que as Razões de Erro Aparente são semelhantes às geradas com (12) do capitulo 3, ou seja, com a proposta deste trabalho.

Como era de esperar, as regiões de má classificação são as mesmas em ambos os métodos, como pode ser observado nos quadros 3.1A.2 , 3.2A.1, 3.4A.1, onde nos subconjuntos (B) e (C) não há nenhum elemento.

Caso p=3 observou-se que:

3).-Toda vez que a hipótese de homogeneidade entre V_1 e V_2 é aceita;

Os autovalores máximo e mínimo da matriz $S_1^{-1}S_2$ estão próximos da unidade e as estimativas das probabilidades de má classificação próximas de 0.5 como pode ser observado nas colunas "PROPOSTA DE DISCRIMINAÇÃO" das execuções 01B, 02B, 13B.

Pode-se ressaltar ainda que, para a obtenção dos valores apresentados nas colunas "SAS", usou-se a Função Linear Discriminante de Fisher.

As regiões de classificação geradas pelos dois métodos são diferentes, como pode ser visto no quadro 3.1B.1, mas tem uma alta porcentagem de elementos comuns - parte (A) dos quadros, e o número de observações das partes (B) e (C) são quase iguais dando origem a Razões de Erro Aparente muito próximas.

4).-Quando a hipótese de homogeneidade das matrizes V_1 e V_2 é rejeitada;

Tem-se execuções tais como as 05B e 06B, onde a Função de Discriminação Quadrática Amostral de Anderson discrimina melhor que a proposta, resultado que se obtém comparando as Razões de Erro Aparente respectivas. Entretanto, nas execuções 07B, 14B, 15B, 16B e 18B, este método classifica melhor. Também há situações onde ambos os métodos geram as mesmas regiões de classificação, como pode ser observado nos resultados das execuções 11B, 12B.

Nesta situação, a transformação usada muda a estrutura inicial dos dados em duas distribuições normais bivariadas com as mesmas caracteristicas que no caso p=2, como se pode observar no gráfico 5.

Por outro lado considerando-se o conjunto de dados do exemplo apresentado, podemos dizer que este método classifica tão bem como a função discriminante quadrática de Anderson, com a possibilidade de se poder escolher a matriz $S_1^{-1}S_2$ ou $S_2^{-1}S_1$ melhor condicionada, devendo-se ficar com aquela que conduz a menores estimativas das probabilidades de má classificação, como se pode observar nos resultados obtidos no ítem (3.3) do capítulo 3. Além disso, os dois conjuntos de dados de dimensão seis apresentados nos quadros 3.C1 e 3.C2 foram transformados em outros conjuntos de dimensão dois, quadros 3.D1 e 3.D2, onde a visualização gráfica é bem mais fácil.(ver o gráfico 6).

Das exposições feitas até aqui, pode-se dizer de maneira mais geral, baseados nos resultados apresentados que:

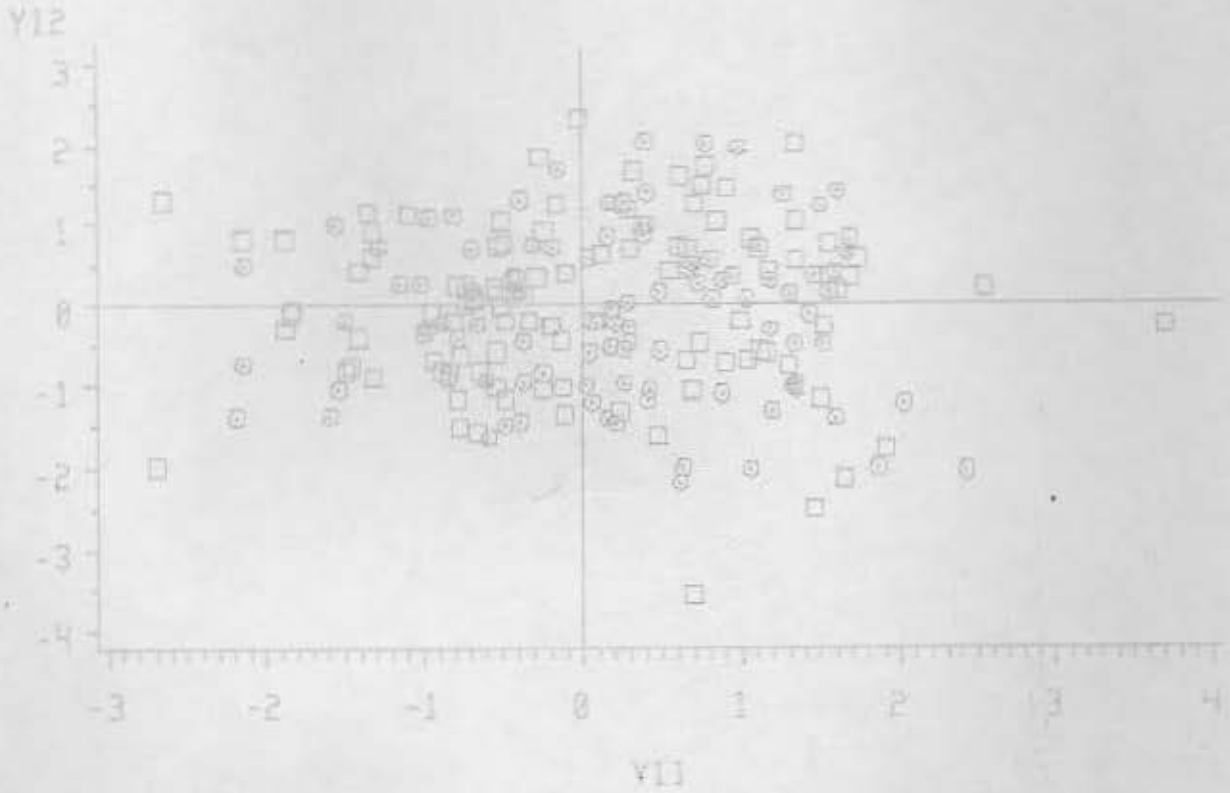
No caso de se terem amostras de duas populações bivariadas, a regra de classificação conduz às mesmas estimativas das probabilidades de má classificação que o SAS, ou seja, os dois métodos discriminam igualmente sempre que a hipótese de homogeneidade das matrizes é rejeitada.

No entanto, quando o número de variáveis medidas em ambos os grupos é três, há situações nas quais os métodos conduzem a estimativas um pouco diferentes para as probabilidades de má classificação.

Vê-se que nos casos estudados, existe um decréscimo significativo na estimativa de probabilidade de má classificação e os métodos se aproximam quando os afastamentos entre as matrizes de covariâncias vai-se acentuando, ou seja, quando a diferença entre os autovalores máximo e mínimo vai aumentando.

Consedere-se o presente trabalho como uma primeira tentativa de abordagem do problema de discriminação e classificação, por outro ponto de vista, para duas populações multivariadas com estrutuuturas de covariâncias diferentes, e que ainda existem incógnitas por responder, tal como a generalização para mais de duas populações.

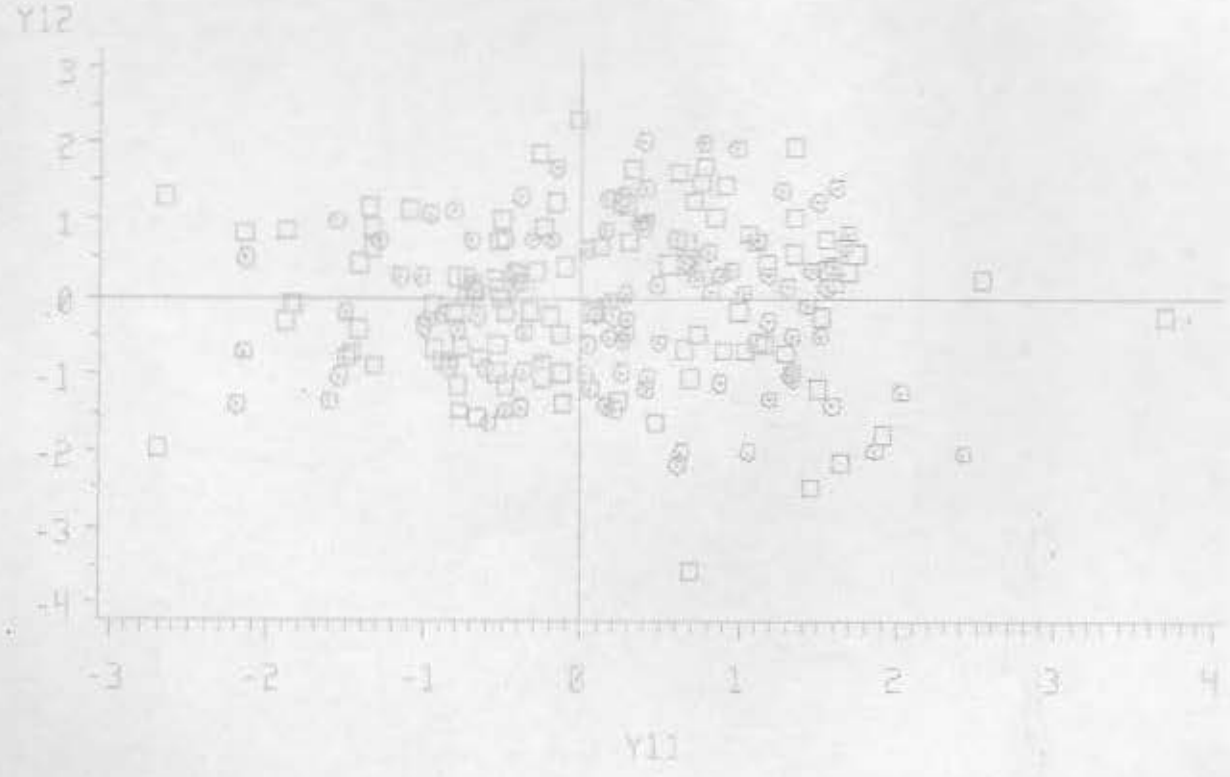
GRAFICO 1
DADOS SIMULADOS TRANSFORMADOS
EXECUCAO NUMERO 01A
O NUMERO DE VAR ORIGINAIS E DOIS



CIRCULO: GRUPO UM
QUADRADO: GRUPO DOIS

GRAFICO 2

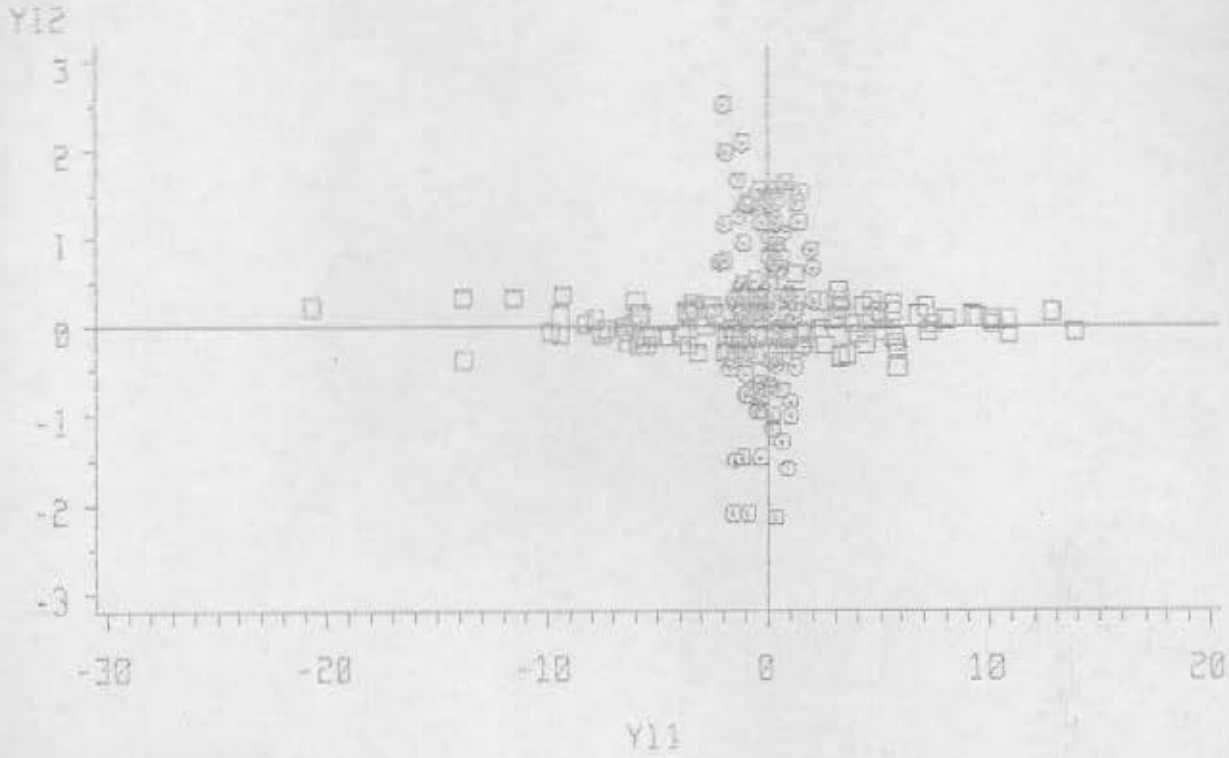
DADOS SIMULADOS TRANSFORMADOS
EXECUCAO NUMERO 87A
O NUMERO DE VAR ORIGINAIS E DOIS



CIRCULO: GRUPO UM
QUADRADO:GRUPO DOIS

GRAFICO 3

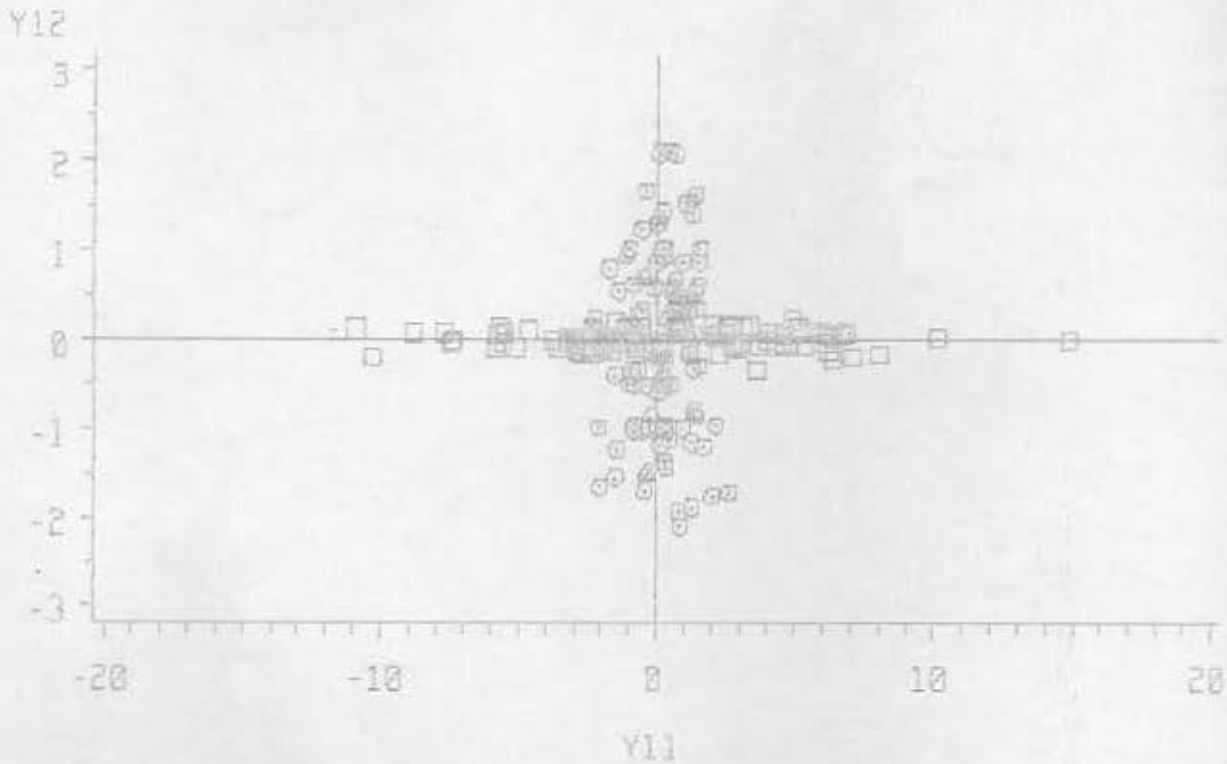
DADOS SIMULADOS TRANSFORMADOS
EXECUCAO NUMERO 09A
O NUMERO DE VAR ORIGINAIS E DOIS



CIRCULO: GRUPO UM
QUADRADO: GRUPO DOIS

GRAFICO 4

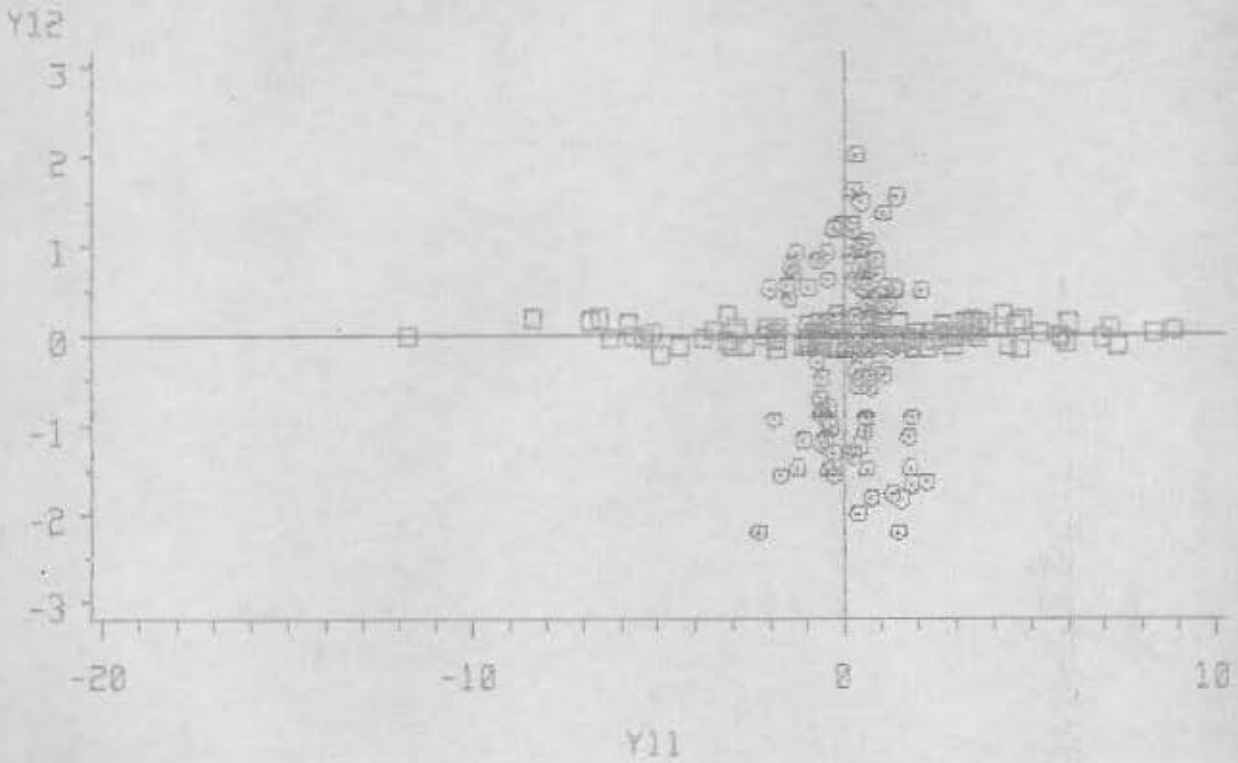
DADOS SIMULADOS TRANSFORMADOS
EXECUCAO NUMERO 13A
O NUMERO DE VAR ORIGINAIS E DOIS



CIRCULO: GRUPO UM
QUADRADO:GRUPO DOIS

GRAFICO 5

DADOS SIMULADOS TRANSFORMADOS
EXECUCAO NUMERO 09B
O NUMERO DE VAR ORIGINAIS E DOIS

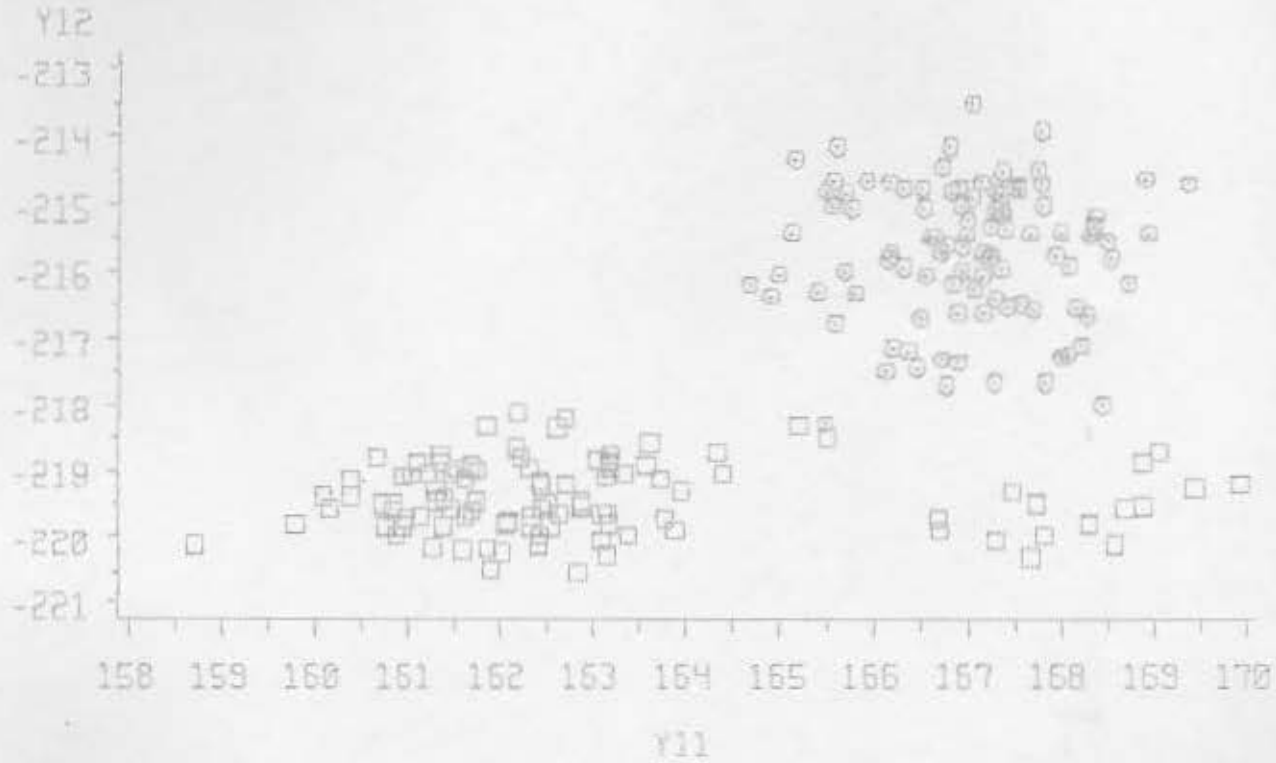


CIRCULO: GRUPO UM
QUADRADO: GRUPO DOIS

GRAFICO 6

DADOS TRANSFORMADOS DAS NOTAS DE BANCO

O NUMERO DE VARIAVEIS ORIGINAIS E SEIS



CIRCULO: GRUPO UM
QUADRADO:GRUPO DOIS

APÊNDICE

```
C       PROGRAMA  1
C       ----------
C
C       PROGRAMA PARA GERAR AMOSTRAS DE DUAS POPULAÇÕES
C       -----------------------------------------------
C       NORMAIS DE DIMENSAO DOIS E  FAZER DISCRIMINAÇÃO.
C       -----------------------------------------------
C         GERA DUAS AMOSTRAS COM DISTRIBUIGAO NORMAL BIVARIADA
C         CALCULA OS VETORES DE MEDIAS AS MATRIZES DE COVARIANCIAS
C
C         USA A SUBRUTINA F02AEF  DA NAG  PARA OBTER OS
C         AUTOVALORES E AUTOVETORES DA MATRIZ INVERS(S1)*S2
C         REALIZA AS TRANSFORMAGOES DAS VARIAVEIS GERADAS USANDO
C         OS AUTOVALORES E AUTOVETORES DA MATRIZ INVERS(S1)*S2.
C
C         REALIZA OS CALCULOS  PARA A FUNGAO DE DISCRIMINAGAO PROPOSTA
C
C         RECLASSIFICA COM A FUNÇÃO DE DISCRIMINAGAO PROPOSTA OS
C         ELEMENTOS AMOSTRAIS GERADOS NO INICIO DO PROGRAMA.
C
      DOUBLE PRECISION R,RR,V1,V2,SOMAX,SOMAY,SOMAZ,XMED,YMED,ZMED,ZMEDA
      DOUBLE PRECISION SOMAX2,SOMAY2,SOMAZ2,VARX,VARY,VARZ,VARZZ,VARXX
      DOUBLE PRECISION A,B,VECP
      DOUBLE PRECISION Z1,W,W1,W1T,W1SOM,W2SOM,W1M,W2M,WMED,VARYY
      DOUBLE PRECISION P2,DIF,RES,DIFL,SWI1,SWI2,TL1,TQA
      DOUBLE PRECISION DISC,AV,VEC,DL,E
      DIMENSION X(2,300),Y(2,300),Z(2,300),TETA(10),L1(300)
      DIMENSION R(2,2),RR(2,2,2),V1(10),V2(10),A1(10),A2(10)
      DIMENSION SOMAX(2),SOMAY(2),SOMAZ(2),XMED(2),YMED(2),ZMED(2)
      DIMENSION ZMEDA(2,2),NTA(2),XT(300,2),YT(300,2),ZT(300,2)
      DIMENSION SOMAX2(2,2),SOMAY2(2,2),SOMAZ2(2,2),ZTT(2,300,2)
      DIMENSION VARX(2,2),VARY(2,2),VARZ(2,2),VARZZ(2,2,2),VARYY(2,2,2)
      DIMENSION A(2,2),B(2,2),VARXX(2,2,2),L2(1,300)
      DIMENSION VEC(2,2),AV(2),DL(2),E(2)
      DIMENSION Z1(300,2),W1(300,2),W(2,300,2),W1T(2,300)
      DIMENSION W1SOM(2),W1M(2),W2M(2),WMED(2,2)
      DIMENSION SWI1(2,2),SWI2(2,2),P2(2),DIF(2,2)
```

```
      DIMENSION TOAC3000),DIFC(2D),MES(300,2D),TLI(3000)
      DIMENSION DISCC300,2D),INDC300,2D),TQA1(3000),VECPC2,2D)
C       PARAMETROS INICIAIS DA SIMULAGAO
C
        TYPE *,'ENTRE COM O NUMERO DE POPULAÇOES'
         ACCEPT *,NP
        TYPE *,'ENTRE COM O NUMERO DE SIMULAÇOES MONTE CARLO'
         ACCEPT *,NMC
        TYPE *,'ENTRE COM DUAS SEMENTES'
         ACCEPT *,S1,S2
C
C      ENTRADA DO TAMANHO DE AMOSTRA E PARAMETROS DE CADA POPULAGAO
C
         DO I=1,NP
         TYPE *,'ENTRE COM O TAMANHO (NTA) DA AMOSTRA DA POPULAÇÃO',I
          ACCEPT *,NTA(I)
         TYPE *,'ENTRE COM OS PARAMETROS V1,V2 DA POPULAGAO',I
          ACCEPT *,V1(I),V2(I)
         type *,'ENTRE COM OS PARAMETROS A1,A2 DA POPULAGAO',I
          accept *,A1(I),A2(I)
         TYPE *,'ENTRE COM O VALOR DE TETAC RADD DA POPULAGAO',I
          ACCEPT *,TETA(I)
         END DO
C
C      COMEÇANDO A SIMULAÇÃO
C
        DO 1000  IMC=1,NMC
C
         DO 5 I=1,2
          DO 5 J=1,2
           DO 5 K=1,2
          VARZZ(I,J,K)=0.0
          VARXX(I,J,K)=0.0
          VARYY(I,J,K)=0.0
5        CONTINUE
C       AMOSTRANDO NTA UNIDADES DE CADA POPULAÇAO
C
      DO 500 IP=1,NP
```

```
C       GERANDO OS VETORES ALEATORIOS N(0,1)
C
         NT=NTA(IP)
        DO ITA=1,NT
         DO I=1,2
          X(I,ITA)=0.0
          Y(I,ITA)=0.0
          Z(I,ITA)=0.0
          XT(ITA,I)=0.0
          YT(ITA,I)=0.0
          ZT(ITA,I)=0.0
         END DO
        END DO
C
           DO ITA=1,NT
            L1(ITA)=1
            L2(1,ITA)=1
            CALL RANDU(S1,S2,U1)
            CALL RANDU(S1,S2,U2)
        X(1,ITA)=SQRT(-2.0*ALOG(U1))*COS(6.283185307*U2)
        X(2,ITA)=SQRT(-2.0*ALOG(U1))*SIN(6.283185307*U2)
C
C       DETERMINACAO DO VETOR Y=D*X+A
C
           Y(1,ITA)=V1(IP)*X(1,ITA)+A1(IP)*L2(1,ITA)
           Y(2,ITA)=V2(IP)*X(2,ITA)+A2(IP)*L2(1,ITA)
          END DO
C
C        DETERMINACAO DO VETOR  Z=R*Y
C        DADO QUE TEMOS DOIS GRUPOS PRECISAMOS  GUARDAR EM DUAS
C        MATRIZES:  ZTT(IP,ITA,I)
           SENO=SIN(TETA(IP))
         COSENO=COS(TETA(IP))
           R(1,1)=COSENO
           R(1,2)=-1.0*SENO
           R(2,1)=SENO
           R(2,2)=COSENO
          DO I=1,2
```

```
      DO ITA=1,NT
        Z(I,ITA)=0.0
       DO K=1,2
         Z(I,ITA)=Z(I,ITA)+R(I,K)*Y(K,ITA)
       END DO
        ZTT(IP,ITA,I)=Z(I,ITA)
      END DO
    END DO
    DO I=1,2
      DO ITA=1,NT
       XT(ITA,I)=X(I,ITA)
       YT(ITA,I)=Y(I,ITA)
       ZT(ITA,I)=Z(I,ITA)
      END DO
    END DO
      DO I=1,2
       DO J=1,2
        RR(IP,I,J)=R(I,J)
      END DO
      END DO
C
C     CALCULO DOS VETORES DE MEDIAS DE X ,Y ,Z EM CADA GRUPO
C
    DO I=1,2
      SOMAX(I)=0.0
       XMED(I)=0.0
      SOMAY(I)=0.0
       YMED(I)=0.0
      SOMAZ(I)=0.0
       ZMED(I)=0.0
     DO ITA=1,NT
       SOMAX(I)=SOMAX(I)+X(I,ITA)*L1(ITA)
       SOMAY(I)=SOMAY(I)+Y(I,ITA)*L1(ITA)
       SOMAZ(I)=SOMAZ(I)+Z(I,ITA)*L1(ITA)
     END DO
      XMED(I)=SOMAX(I)/(1.0*NT)
      YMED(I)=SOMAY(I)/(1.0*NT)
      ZMED(I)=SOMAZ(I)/(1.0*NT)
```

```
      ZMEDAC I , IPO = ZMED( J )
     END DO
C
C    CALCULANDO AS MATRIZES DE COVARIANCIAS DAS VARIAVEIS ORIGINAIS
C
C    a) ZERANDO AS MATRIZES DE COVARIANCIAS
C
      DO I=1,2
       DO J=1,2
         VARX(I,J)=0.0
         VARY(I,J)=0.0
         VARZ(I,J)=0.0
       END DO
      END DO
C
C     b) CALCULANDO AS MATRIZES DE SOMA DE QUADRADOS
C
      DO I=1,2
       DO J=I,2
          SOMAX2(I,J)=0.0
          SOMAY2(I,J)=0.0
          SOMAZ2(I,J)=0.0
        DO K=1,NT
          SOMAX2(I,J)=SOMAX2(I,J)+X(I,K)*XT(K,J)
          SOMAY2(I,J)=SOMAY2(I,J)+Y(I,K)*YT(K,J)
          SOMAZ2(I,J)=SOMAZ2(I,J)+Z(I,K)*ZT(K,J)
        END DO
          SOMAX2(J,I)=SOMAX2(I,J)
          SOMAY2(J,I)=SOMAY2(I,J)
          SOMAZ2(J,I)=SOMAZ2(I,J)
       END DO
      END DO
C    c)CALCULO DAS MATRIZES DE COVARIANCIAS DOS VETORES X  Y  E Z
C
     DO I=1,2
      DO J=I,2
       VARX(I,J)=(SOMAX2(I,J)-1.0*NT*XMED(I)*XMED(J))/(1.0*(NT-1))
       VARY(I,J)=(SOMAY2(I,J)-1.0*NT*YMED(I)*YMED(J))/(1.0*(NT-1))
```

```
            VARZZ(I,J)=(SOMAZZ(I,J)-1.0*NT*ZMED(I)*ZMED(J))/(1.0*(NT-1))
            VARXX(I,J)=VARX(I,J)
            VARYY(I,J)=VARY(I,J)
            VARZZ(I,J)=VARZ(I,J)
           END DO
           DO J=1,2
            VARXX(IP,I,J)=VARX(I,J)
            VARYY(IP,I,J)=VARY(I,J)
            VARZZ(IP,I,J)=VARZ(I,J)
           END DO
          END DO
C
500     CONTINUE
C
C       CALCULANDO  OS AUTOVALORES E AUTOVECTORES
C       DA MATRIZ INV(VARZZ(1)*VARZZ(2)
C       USANDO A SUBROUTINA DA NAG
C
           DO 540 I=1,2
            DO 540 J=1,2
              B(I,J)=VARZZ(1,I,J)
              A(I,J)=VARZZ(2,I,J)
540         CONTINUE
          N=2
          IA=2
          IB=2
          IV=2
          IFAIL=1
          CALL F02AEF(A,IA,B,IB,N,AV,VEC,IV,DL,E,IFAIL)
          TYPE *,'VOLTOU DA SUBROUTINA'
           IF (IFAIL.EQ.0) THEN
               TYPE *,' AUTOVALORES'
               TYPE *,AV(2)
               TYPE *,' '
               TYPE *,AV(1)
               WRITE(90,95)(AV(I),I=1,N)
               TYPE 95,(AV(I),I=1,N)
               WRITE(90,97)((VEC(I,J),J=1,N),I=1,N)
```

```
            TYPE 97,((VEC(I,J),J=1,ND),I=1,ND)
95          FORMAT(12HO EIGENVALUES=/1H,2F18.6)
97          FORMAT(13HO EIGENVECTORS=/(1H,2(2X,F18.6)))
          ELSE
           TYPE 96, IFAIL
           WRITE(90,96)IFAIL
96         FORMAT(25HO ERROR IN F02AEF IFAIL= ,I2)
           STOP
          END IF
C
C       CALCULANDO AS COMBINACOES LINEARES USANDO OS AUTOVETORES
C       OBTEN-SE AS VARIAVEIS DE INTERESSSE
C
C          REDEFININDO A MATRIZ DE AUTOVETORES
C
           VECP(1,1)=VEC(1,2)
           VECP(2,1)=VEC(2,2)
           VECP(1,2)=VEC(1,1)
           VECP(2,2)=VEC(2,1)
C
           TYPE *,'MATRIZ AUTOVETORES ORDENADOS'
             DO I=1,2
           TYPE *,(VECP(I,J),J=1,2)
             END DO
         DO 550 IP=1,2
           NT=NTA(IP)
          DO 510 I=1,NT
           DO 510 J=1,2
             Z1(I,J)=0.0
             W(IP,I,J)=0.0
             Z1(I,J)=ZTT(IP,I,J)
             W1(I,J)=0.0
510       CONTINUE
           DO 511 I=1,NT
             DO 511 J=1,2
             DO K=1,2
              W1(I,J)=W1(I,J)+Z1(I,K)*VECP(K,J)
            END DO
```

```
            W(IP,I,JD)=W1(I,JD)
            W1T(JD,ID)=W1(I,JD)
511       CONTINUE
C
C          CALCULANDO OS VETORES DE MEDIAS DAS VARIAVEIS TRANSFORMADAS
          DO 520 I=1,2
            W1SOM(ID)=0.0
            W1M(ID)=0.0
             DO ITA=1,NT
              W1SOM(ID)=W1SOM(ID)+W1T(I,ITA)*L1(ITA)
             END DO
            W1M(ID)=W1SOM(ID)/(1.0*NTD)
            WMED(I,IPD)=W1M(ID)
520        CONTINUE
550       CONTINUE
C
C        CALCULANDO A FUNCÃO DISCRIMINANTE
C
C         a) Inversa das matrizes de covariancias das transformacoes
C
            DO 595 I=1,2
             DO 595 J=1,2
              SWI1(I,JD)=0.0
              SWI2(I,JD)=0.0
595         CONTINUE
             SWI1(1,1D)=SWI1(1,1D)+1.0
             SWI1(2,2D)=SWI1(2,2D)+1.0
             SWI2(1,1D)=SWI2(1,1D)+1.0/AV(2D)
             SWI2(2,2D)=SWI2(2,2D)+1.0/AV(1D)
           DO 596 I=1,2
            DO 596 J=1,2
            DIF(I,JD)=0.0
            DIF(I,JD)=DIF(I,JD)+SWI1(I,JD)-SWI2(I,JD)
596       CONTINUE
C
C         b)Calculando os termos constantes da fungao
C
         DO  I=1,2
```

```
          W1MC1)=0.0
          W2MC1)=0.0
          W1MC1)=WMEDC1,1)
          W2MC1)=WMEDC1,2)
        END DO
        XIND1=0.0
        XIND2=0.0
        DO I=1,2
          P2C1)=0.0
          DIFLC1)=0.0
          DO J=1,2
            P2C1)=P2C1)+SWI2CI,J)*W2MCJ)
          END DO
          DIFLC1)=W1MC1)-P2C1)
          XIND1=XIND1+W1MC1)*W1MC1)
          XIND2=XIND2+W2MC1)*P2C1)
        END DO
          XX=AVC1)*AVC2)
C
      XINDT=XIND1-XIND2-LOGCXX)
C
C     c)Calculando os termos quadratico e linear
      DO 600 IP=1,2
        NT=NTACIP)
        DO I=1,NT
          INDCI,IP)=0
          DO J=1,2
            W1CI,J)=0.0
            W1TCJ,I)=0.0
            W1CI,J)=WCIP,I,J)
            W1TCJ,I)=W1CI,J)
          END DO
        END DO
        DO I=1,NT
          DISCCI,IP)=0.0
          TQAC1)=0.0
          TLIC1)=0.0
          DO J=1,2
```

```
                RESC I , JD = 0.0
              DO K=1 , 2
                RESC I , JD = RESC I , JD + W1 C I , K D * DI F C K , JD
              END DO
            END DO
          END DO
        DO I =1 , NT
          DO J=1 , 2
           TQAC I D = TQAC I D + RESC I , JD * W1 T C J , I D
          END DO
        END DO
        DO I =1 , NT
            DO J=1 , 2
              TL1 C I D = TL1 C I D + W1 C I , JD * DI FL C JD
            END DO
        END DO
C
C       d) Obtencao da fungao de discriminagao
C          usando os resultados anteriores
         DO I =1 , NT
          DI SCC I , I PD = TQAC I D - 2 * TLI C I D + X1 NDT
          I FC DI SCC I , I PD . LE. 0D THEN
                                    I NDC I , I PD = 1
              ELSE
                   I NDC I , I PD = 2
           END IF
         END DO
600     CONTI NUE
1000    CONTI NUE
        STOP
        END
```

```
C      PROGRAMA 2
C      PROGRAMA PARA GERAR AMOSTRAS  DE DOIS POPULAÇÕES NORMAIS
C      ------------------------------------------------------
C      DE DIMENSÃO TRES   E   FAZER DISCRIMINAÇÃO
C      ------------------------------------------
C      GERA DUAS AMOSTRAS  DE DIMENSAO TRES.
C
C      CALCULA OS VETORES DE MEDIAS E AS MATRIZES DE COVARIANCIAS
C      PARA AS AMOSTRAS GERADAS.
C      USA A SUBROTINA F02AEF DA BIBLIOTECA DA NAG PARA  OBTER OS
C
C      AUTOVALORES E AUTOVETORES DA MATRIZ  INV(S1)*S2
C
C      REALIZA OS CALCULOS CONDUCENTES A  OBTENCAO DA FUNCAO
C      DISCRIMINANTE.
C
C     RECLASSIFICA AS OBSERVACOES AMOSTRAIS  SEGUNDO A PROPOSTA DE
C      DISCRIMINACAO.
C
      DOUBLE PRECISION RR,RRR,SOMAX,SOMAY,SOMAZ,XMED,YMED,ZMED,ZMEDA
        DOUBLE PRECISION SOMAX2,SOMAY2,SOMAZ2,VARX,VARY,VARZ,VARZZ
        DOUBLE PRECISION A,B,SI,R,V,DL,AUTOVE
        DOUBLE PRECISION Z1,W,W1,W1T,W1SOM,W2SOM,W1M,W2M,WMED
        DOUBLE PRECISION P2,DIF,RES,DIFL,SW11,SWI2,TLI,TQA
        DOUBLE PRECISION DISC
        DIMENSION X(3,300),Y(3,300),Z(3,300),TETA(10),L1(300)
        dimension L2(1,300),A1(10),A2(10),A3(10)
        DIMENSION RR(3,3),RRR(2,3,3),V1(10),V2(10),V3(10)
        DIMENSION SOMAX(3),SOMAY(3),SOMAZ(3),XMED(3),YMED(3),ZMED(3)
        DIMENSION ZMEDA(3,2),NTA(2),XT(300,3),YT(300,3),ZT(300,3)
        DIMENSION SOMAX2(3,3),SOMAY2(3,3),SOMAZ2(3,3),ZTT(2,300,3)
        DIMENSION VARX(3,3),VARY(3,3),VARZ(3,3),VARZZ(2,3,3)
        DIMENSION A(3,3),B(3,3),R(3),V(3,3),AUTOVE(3,2)
        DIMENSION Z1(300,3),W1(300,2),W(2,300,2),W1T(2,300)
        DIMENSION W1SOM(2),W1M(2),W2M(2),WMED(2,2),P2(2)
        DIMENSION SWI1(2,2),SWI2(2,2),DIFQ(2,2),RES(300,2)
        DIMENSION TQA(300),DIFL(2),TLI(300)
        DIMENSION DISC(300,2),INDICA(300,2)
```

```
C
C       ENTRADA DE PARAMETROS INICIAIS
C
        TYPE *,'ENTRE COM O NUMERO DE POPULACOES'
         ACCEPT *,NP
        TYPE *,'ENTRE COM O NUMERO DE SIMULACOES MONTE CARLO'
         ACCEPT *,NMC
        TYPE *,'ENTRE COM DUAS SEMENTES'
         ACCEPT *,S1,S2
C
C     ENTRADA DO TAMANHO DE AMOSTRA E PARAMETROS DE CADA POPULAÇAO
      ,DO I=1,NP
      TYPE *,'ENTRE COM O TAMANHO (NTA) DA AMOSTRA DA POPULAÇAO',I
           ACCEPT *,NTA(I)
          TYPE *,'ENTRE COM OS PARAMETROS V1,V2,V3 DA POPULACAO',I
           ACCEPT *,V1(I),V2(I),V3(I)
          TYPE *,'ENTRE COM OS PARAMETROS A1,A2,A3 DA POPULACAO',I
           ACCEPT *,A1(I),A2(I),A3(I)
          TYPE *,'ENTRE COM O VALOR DE TETA(RAD) DA POPULACAO',I
           ACCEPT *,TETA(I)
          END DO
C
C       COMECANDO A SIMULACAO
C
          DO 1000  IMC=1,NMC
C
           DO 5 I=1,2
            DO 5 J=1,3
             DO 5 K=1,3
            VARZZ(I,J,K)=0.0
5           CONTINUE
C         AMOSTRANDO NTA UNIDADES DE CADA POPULACAO
           DO 500 IP=1,NP
           NT=NTA(IP)
C
C         GERANDO OS VETORES ALEATORIOS N(0,I)
           DO ITA=1,NT
            DO I=1,3
             X(I,ITA)=0.0
```

```
      Y(1,ITA)=0.0
      Z(1,ITA)=0.0
      XT(ITA,ID)=0.0
      YT(ITA,ID)=0.0
      ZT(ITA,ID)=0.0
      END DO
    END DO
   DO ITA=1,NT
      L1(ITA)=1
      L2(1,ITA)=1
      CALL RANDU(S1,S2,U1)
      CALL RANDU(S1,S2,U2)
      CALL RANDU(S1,S2,U3)
      CALL RANDU(S1,S2,U4)
  X(1,ITA)=SQRT(-2.0*ALOG(U1))*COS(6.283185307*U2)
  X(2,ITA)=SQRT(-2.0*ALOG(U1))*SIN(6.283185307*U2)
  X(3,ITA)=SQRT(-2.0*ALOG(U3))*COS(6.283185307*U4)
C
C     DETERMINACAO DO VETOR Y=D*X
C
      Y(1,ITA)=V1(IP)*X(1,ITA)+A1(IP)*L2(1,ITA)
      Y(2,ITA)=V2(IP)*X(2,ITA)+A2(IP)*L2(1,ITA)
      Y(3,ITA)=V3(IP)*X(3,ITA)+A3(IP)*L2(1,ITA)
   END DO
C
C     DETERMINAÇÃO DO VETOR  Z=R*Y
C
      SENO=SIN(TETA(IP))
   COSENO=COS(TETA(IP))
      RR(1,1)=COSENO
      RR(1,2)=-1*SENO
      RR(1,3)=0.0
      RR(2,1)=SENO
      RR(2,2)=COSENO
      RR(2,3)=0.0
      RR(3,1)=0.0
      RR(3,2)=0.0
      RR(3,3)=1.0
```

```
      DO I=1,3
       DO ITA=1,NT
          ZC(I,ITA)=0.0
         DO K=1,3
           ZC(I,ITA)=ZC(I,ITA)+RRC(I,K)*YC(K,ITA)
         END DO
          ZTTC(IP,ITA,I)=ZC(I,ITA)
        END DO
       END DO
        DO I=1,3
         DO J=1,3
          RRR(IP,I,J)=RRC(I,J)
         END DO
        END DO
      DO ITA=1,NT
       DO I=1,3
        XTC(ITA,I)=XC(I,ITA)
        YTC(ITA,I)=YC(I,ITA)
        ZTC(ITA,I)=ZC(I,ITA)
       END DO
      END DO
C
C     CALCULO DOS VETORES DE MEDIAS DE X ,Y ,Z PARA CADA AMOSTRA
C
C
       DO I=1,3
          SOMAX(I)=0.0
           XMED(I)=0.0
          SOMAY(I)=0.0
           YMED(I)=0.0
          SOMAZ(I)=0.0
           ZMED(I)=0.0
         DO ITA=1,NT
           SOMAX(I)=SOMAX(I)+XC(I,ITA)*L1(ITA)
           SOMAY(I)=SOMAY(I)+YC(I,ITA)*L1(ITA)
           SOMAZ(I)=SOMAZ(I)+ZC(I,ITA)*L1(ITA)
         END DO
          XMED(I)=SOMAX(I)/(1.0*NT)
```

```
          YMED(I)=SOMAY(I)/(1.0*NT)
          ZMED(I)=SOMAZ(I)/(1.0*NT)
          ZMEDA(I,IP)=ZMED(I)
        END DO
C
C       CALCULANDO AS MATRIZES DE COVARIANCIAS PARA AS AMOSTRAS
C
C        a) ZERANDO AS MATRIZES DE COVARIANCIAS
C
        DO I=1,3
          DO J=1,3
            VARX(I,J)=0.0
            VARY(I,J)=0.0
            VARZ(I,J)=0.0
          END DO
        END DO
C
C       b)CALCULANDO AS MATRIZES DE SOMA DE QUADRADOS
C
        DO I=1,3
          DO J=1,3
            SOMAX2(I,J)=0.0
            SOMAY2(I,J)=0.0
            SOMAZ2(I,J)=0.0
          END DO
        END DO
        DO I=1,3
          DO J=I,3
            DO K=1,NT
              SOMAX2(I,J)=SOMAX2(I,J)+X(I,K)*XT(K,J)
              SOMAY2(I,J)=SOMAY2(I,J)+Y(I,K)*YT(K,J)
              SOMAZ2(I,J)=SOMAZ2(I,J)+Z(I,K)*ZT(K,J)
            END DO
            SOMAX2(J,I)=SOMAX2(I,J)
            SOMAY2(J,I)=SOMAY2(I,J)
            SOMAZ2(J,I)=SOMAZ2(I,J)
          END DO
        END DO
```

```
C            c)CALCULO DAS MATRIZES DE COVARIANCIAS
C
            DO I=1,3
             DO J=1,3
          VARX(I,J)=(SOMAX2(I,J)-1.0*NT*XMED(I)*XMED(J))/(1.0*(NT-1))
          VARY(I,J)=(SOMAY2(I,J)-1.0*NT*YMED(I)*YMED(J))/(1.0*(NT-1))
          VARZ(I,J)=(SOMAZ2(I,J)-1.0*NT*ZMED(I)*ZMED(J))/(1.0*(NT-1))
              VARX(J,I)=VARX(I,J)
              VARY(J,I)=VARY(I,J)
              VARZ(J,I)=VARZ(I,J)
             END DO
            DO J=1,3
             VARZZ(IP,1,J)=VARZ(I,J)
           END DO
            END DO
500      CONTINUE
C
C        REDEFININDO AS MATRIZES DE COVARIANCIAS DAS DUAS POPULAÇÕES
C
           DO 550 I=1,3
            DO 550 J=1,3
              B(I,J)=VARZZ(1,I,J)
              A(I,J)=VARZZ(2,I,J)
550        CONTINUE
C
C       CALCULANDO OS AUTOVALORES E AUTOVETORES DA MATRIZ INV(B)*A
C       USANDO A SUBROUTINA DA NAG
C
         N=3
         TYPE *,'B(I,J)'
        DO I=1,N
         TYPE *,(B(I,J),J=1,N)
        END DO
         TYPE *,'A(I,J)'
        DO I=1,N
         TYPE *,(A(I,J),J=1,N)
        END DO
         IA=3
```

```
      IB=3
      IV=3
      IFAIL=1
       CALL F02AEF(A,IA,B,IB,N,E,V,IV,DL,E,IFAIL)
       TYPE *,'VOLTOU DA SUBRUTINA'
      IF(IFAIL.EQ.0) THEN
       WRITE(80,95)(R(I),I=1,N)
       TYPE 95,(R(I),I=1,N)
       WRITE(80,97)((V(I,J),J=1,N),I=1,N)
       TYPE 97,((V(I,J),J=1,N),I=1,N)
95     FORMAT(12HO EIGENVALUES/1H,3F18.6)
97     FORMAT(13HO EIGENVECTORS/(1H,3(2X,F15.6)))
       ELSE
        TYPE 96,IFAIL
        WRITE(80,96)IFAIL
96      FORMAT(25HO ERROR IN F02AEF IFAIL= ,I2)
        STOP
       END IF
           XX=R(3)*R(1)
          DO I=1,3
           AUTOVEC(I,1)=V(I,3)
           AUTOVEC(I,2)=V(I,1)
          END DO
C
      TYPE *,'MATRIZ ORDENADA DE AUTOVETORES'
       DO I=1,3
          TYPE *,(AUTOVEC(I,J),J=1,2)
      END DO
C
C     CALCULANDO AS COMBINACOES LINEARES USANDO OS AUTOVETORES
C     OBTEN-SE AS VARIAVEIS DE INTERESSSE
C
      DO 200 IP=1,2
         NT=NTAC(IP)
       DO 110 I=1,NT
         DO J=1,3
          Z1(I,J)=0.0
          Z1(I,J)=ZTT(IP,I,J)
```

```
          END DO
         DO 110 J=1,2
           WC1P,I,JD=0.0
           W1CI,JD=0.0
            DO K=1,3
             W1CI,JD=W1C1,JD+Z1C1,KD*AUTOVECK,JD
            END DO
           WC1P,I,JD=W1CI,JD
           W1TCJ,ID=W1CI,JD
110      CONTINUE
         DO 120 I=1,2
           W1SOMCID=0.0
           W1MCID=0.0
            DO ITA=1,NT
             W1SOMCID=W1SOMCID+W1TCI,ITAD*L1CITAD
            END DO
           W1MCID=W1SOMCID/C1.0*NTD
           WMEDCI,IPD=W1MCID
120       CONTINUE
200      CONTINUE
         TYPE 210
210      FORMATC/,10X,'OS VETORES DE MEDIAS TRANSFORMADOS SAO:',/D
         DO I=1,2
         TYPE *,CWMEDCI,IPD,IP=1,2D
         END DO
C
C        OBTENCAO DA FUNCAO DISCRIMINANTE
C
C         a) Inversa das matrizes de covariancias das combinações
C           segundo o objetivo proposto.
C
            do 595 I=1,2
            do 595 J=1,2
              swi1Ci,jD=0
              swi2Ci,jD=0
595           continue
            SWI1C1,1D=swi1C1,1D+1.0
            SWI1C2,2D=swi1C2,2D+1.0
```

```
            SWI1(I,J)=SWI1(I,J)+1.0*P(J)
            SWI2(I,J)=SWI2(I,J)+1.0*P(J)
          DO 220 I=1,2
           DO 220 J=1,2
            difq(I,J)=0.0
            DIFQ(I,J)=difq(I,J)+SWI1(I,J)-SWI2(I,J)
220       CONTINUE
C
C         b)Calculando os termos constantes da função
C
        DO  I=1,2
          W1MC(I)=0.0
          W2MC(I)=0.0
          W1MC(I)=WMEDC(I,1)
          W2MC(I)=WMEDC(I,2)
        END DO
          XIND1=0.0
          XIND2=0.0
        DO I=1,2
          P2C(I)=0.0
          DIFLC(I)=0.0
          DO J=1,2
           P2C(I)=P2C(I)+SWI2(I,J)*W2MC(J)
          END DO
          DIFLC(I)=W1MC(I)-P2C(I)
          XIND1=XIND1+W1MC(I)*W1MC(I)
          XIND2=XIND2+W2MC(I)*P2C(I)
        END DO
C       Termo independente total
C
        XINDT=XIND1-XIND2-ALOG(XX)
C
        TYPE *,'TERMOS INDEPENDENTES'
        TYPE *,XINDT,XIND1,XIND2,ALOG(XX)
C        c)Calculando os termos quadraticos e linear
        DO 300 IP=1,2
            NT=NTAC(IP)
           DO I=1,NT
```

```
            INDICAC I,IP)=0
          DO J=1,2
            W1(I,JD=0.0
            W1TCJ,ID=0.0
            W1(I,JD=WCIP,I,JD
            W1TCJ,ID=W1CI,JD
          END DO
         END DO
        DO I=1,NT
            DISCCI,IPD=0.0
            TQACID=0.0
            TLICID=0.0
            DO J=1,2
              RESCI,JD=0.0
             DO K=1,2
              RESCI,JD=RESCI,JD+W1CI,KD*DIFQCK,JD
             END DO
            END DO
            DO J=1,2
              TLICID=TLICID+W1CI,JD*DIFLCJD
              TQACID=TQACID+RESCI,JD*W1TCJ,ID
            END DO
          END DO
C       dDCalculando a funcao discriminante com os resultados
C         anteriores
        DO I=1,NT
         DISCCI,IPD=TQACID-2*TLICID+XINDT
           IFCDISCCI,IPD.LE.0D THEN
                                  INDICACI,IPD=1
              ELSE
                   INDICACI,IPD=2
           END IF
          END DO
300     CONTINUE
1000    CONTINUE
800      CONTINUE
        STOP
        END
```

PROGRAMA Nº 3

DISCRIMINAÇÃO COM O PROC DISCRIM CSASO

```
OPTIONS LS=80;
TITLE 'ANALISE DISCRIMINANTE';
DATA SIMUL;
TITLE2 'DADOS SIMULADOS EM DUAS VARIAVEIS';
INFILE SIMOO1A.DAT;
INPUT X1 5-15  X2 20-30  ESPEC 47 ;
IF ESPEC=1 THEN TIPO='PRIMEIRA POPULAÇÃO';
           ELSE TIPO='SEGUNDA POPULAÇÃO';
LABEL X1='PRIMEIRA VARIAVEL NORMAL';
      X2='SEGUNDA VARIAVEL NORMAL';
  ESPEC ='POPULAÇÃO DE PROCEDENÇA';
PROC PRINT;
PROC DISCRIM SIMPLE LISTERR POOL=TEST;
CLASS TIPO;
VAR X1-X2;
```

PROGRAMA Nº 4

DISCRIMINAÇÃO COM O PROC DISCRIM CSASO

```
OPTIONS LS=80;
TITLE 'ANALISE DISCRIMINANTE';
DATA SIMUL;
TITLE2 'DADOS SIMULADOS EM TRES VARIAVEIS';
INFILE SIMOO1B.DAT;
INPUT X1 5-15  X2 20-30  X3 32-45  ESPEC 60;
IF ESPEC=1 THEN TIPO='PRIMEIRA POPULAÇÃO';
           ELSE TIPO='SEGUNDA POPULAÇÃO';
LABEL X1='PRIMEIRA VARIAVEL NORMAL';
      X2='SEGUNDA VARIAVEL NORMAL';
      X3='TERCEIRA VARIAVEL NORMAL,;
  ESPEC ='POPULAÇÃO DE PROCEDENÇA';
PROC PRINT;
PROC DISCRIM SIMPLE LISTERR POOL=TEST;
CLASS TIPO;
VAR X1-X3;
```

```
C       PROGRAMA  9
C       ----------
C
C
C       PROGRAMA PARA CALCULAR OS VETORES DE MEDIAS, MATRIZES DE
C       --------------------------------------------------------
C
C       COVARIANCIAS DOS DADOS CONSIDERADOS NA APLICAÇÃO.
C       -------------------------------------------------
C
C        ARQUIVO FORO10.DAT CONTEM AS MEDICOES DAS NOTAS VEERDADEIRAS.
C
C        ARQUIVO FORO50.DAT CONTEM AS MEDICOES DAS NOTAS FALSAS.
C
C        CALCULA OS AUTOVALORES E AUTOVETORES DA MATRIZ INV(S1)*S2.
C
C        REALIZA TRANSFORMACOES COM OS AUTOVETORES ASSOCIADOS AOS
C        AUTOVALORES MAXIMO E MINIMO DA MATRIZ INV(S1)*S2.
C
C        APRESENTA-SE A FUNCAO DE DISCRIMINACAO SEGUNDO A PROPOSTA.
C
C        RECLASSIFICA AS OBSERVACOES AMOSTRAIS AVALIANDOAS NA
C        FUNCAO PROPOSTA
C
        double precision somax1,somax2,somay1,somay2
        double precision x1med,x2med,y1med,y2med,somaqx1,somaqx2
        double precision varx1,varx2,a,b,r,v,autove
        double precision swi1,swi2,difq,l1,quay1,quadry1
        double precision quay2,quadry2,s1,s2,difl,vliny1,vliny2
        dimension x1(100,6),x1t(6,100),x2(100,6),x2t(6,100)
        dimension somax1(6),somax2(6),x1med(6),x2med(6),l1(100)
        dimension Y1(100,2),Y2(100,2),Y1T(2,100),Y2T(2,100)
        dimension somay1(2),somay2(2),y1med(2),y2med(2)
        dimension somaqx1(6,6),somaqx2(6,6)
        dimension varx1(6,6),varx2(6,6),A(7,7),B(7,7),R(7),DL(6)
        dimension V(7,7),autove(6,2)
        dimension swi1(2,2),swi2(2,2),difq(2,2)
        dimension s1(2),s2(2),difl(2),vliny1(100),vliny2(100)
```

```
      dimension quadrx1(1000),quadrx2(1000),quay1(100,2),quay2(100,2)
      dimension indica1(1000),indica2(1000)
      dimension discrim1(1000),discrim2(1000)
C
C     LENDO OS  ARQUIVOS DE DADOS
C
      do I=1,100
        read(10,*)(X1(I,J),J=1,6)
      end do
      do I=1,100
        read(50,*)(X2(I,J),J=1,6)
      end do
C
C     Calculando a matriz transposta yt,wt
C
      do I=1,100
       L1(I)=1
        do J=1,6
          X1t(J,I)=X1(I,J)
          X2t(J,I)=X2(I,J)
        end do
      end do
C
C
C     CALCULANDO OS VETORES DE MEDIAS
C
      DO I=1,6
       SOMAX1(I)=0
       X1MED(I)=0
       SOMAX2(I)=0
       X2MED(I)=0
       end do
         do I=1,6
          DO J=1,100
           SOMAX1(I)=SOMAX1(I)+X1T(I,J)*L1(J)
           SOMAX2(I)=SOMAX2(I)+X2T(I,J)*L1(J)
          END DO
           X1MED(I)=SOMAX1(I)/(1.0*100)
           X2MED(I)=SOMAX2(I)/(1.0*100)
         END DO
```

```
C
C
C       CALCULANDO AS MATRIZES DE covariancias
C
C        a)Zerando as matrizes de covariancias
C
            do i=1,6
              do j=1,6
               varx1(i,j)=0.0
               varx2(i,j)=0.0
              end do
             end do
C
C           b)Calculando a matriz  soma de quadrados
C
           DO I=1,6
            DO J=1,6
             SOMAQX1(I,J)=0.0
             SOMAQX2(I,J)=0.0
            end do
           end do
            do I=1,6
             do j=I,6
               DO K=1,100
                SOMAQX1(I,J)=SOMAQX1(I,J)+X1t(I,K)*X1(K,J)
                SOMAQX2(I,J)=SOMAQX2(I,J)+X2t(I,K)*X2(K,J)
               END DO
                SOMAQX1(J,I)=SOMAQX1(I,J)
                SOMAQX2(J,I)=SOMAQX2(I,J)
             END DO
            END DO
C       c) CALCULANDO AS MATRIZES DE COVARIANCIAS
C
C
           DO I=1,6
            DO J=I,6
             VARX1(I,J)=(SOMAQX1(I,J)-1.0*100*X1MED(I)*X1MED(J))/(1.0*99.0)
             VARX2(I,J)=(SOMAQX2(I,J)-1.0*100*X2MED(I)*X2MED(J))/(1.0*99.0)
             VARX1(J,I)=VARX1(I,J)
             VARX2(J,I)=VARX2(I,J)
            END DO
```

```
      END DO

C
C     dREDEFININDO AS MATRIZES DE COVARIANCIAS DAS DUAS POPULAGOES
C       PARA USAR A SUBROUTINA F02AEF DA NAG
C
       DO 550 I=1,6
        DO 550 J=1,6
        A(I,J)=VARX2(I,J)
        B(I,J)=VARX1(I,J)
550     CONTINUE
C
C     CALCULANDO OS AUTOVALORES E AUTOVETORES DA MATRIZ
C     INV(VARX1)*VARX2  USANDO A SUBROUTINA DA NAG
C
      N=6
      TYPE *,'B(I,J)'
       DO I=1,N
        TYPE *,(B(I,J),J=1,N)
       END DO
        TYPE *,'A(I,J)'
       DO I=1,N
        TYPE *,(A(I,J),J=1,N)
       END DO
        TYPE *,'LEU OS DADOS'
        IA=7
        IB=7
        IV=7
        IFAIL=1
      CALL F02AEF(A,IA,B,IB,N,R,V,IV,DL,E,IFAIL)
      TYPE *,'VOLTOU DA SUBRUTINA'
       IF(IFAIL.EQ.0) THEN
        WRITE(33,95)(R(I),I=1,N)
        TYPE 95,(R(I),I=1,N)
        WRITE(33,97)((V(I,J),J=1,N),I=1,N)
        TYPE 97,((V(I,J),J=1,N),I=1,N)
95        FORMAT(12HO EIGENVALUES/1H,6F18.6)
97        FORMAT(13HO EIGENVECTORS/(1H,6(2X,F15.6)))
        ELSE
        TYPE 96,IFAIL
```

```
      WRITE(*,99)IFAIL
 99   FORMAT(25HO ERROR IN F02AEF IFAIL= , I2)
      STOP
      END IF
      XX=R(6)*R(1)
     DO I=1,6
      AUTOVEC(I,1)=V(I,6)
      AUTOVEC(I,2)=V(I,1)
     END DO
C
C
C
C    CALCULANDO AS COMBINACOES LINEARES OBTEM-SE AS
C    VARIAVEIS DE INTERESSE USANDO OS AUTOVETORES
C    MAXIMO E MINIMO
C
C
      DO I=1,100
       DO J=1,2
        Y1(I,J)=0.0
        Y2(I,J)=0.0
       END DO
      END DO
C
      DO I=1,100
       DO J=1,2
        DO K=1,6
         Y1(I,J)=Y1(I,J)+X1(I,K)*AUTOVEC(K,J)
         Y2(I,J)=Y2(I,J)+X2(I,K)*AUTOVEC(K,J)
        END DO
       END DO
      END DO
C
C    Calculando as matrizes transpostas das variaveis transformadas
C
      do I=1,100
        do J=1,2
         Y1T(J,I)=Y1(I,J)
         Y2T(J,I)=Y2(I,J)
        END DO
```

```
      END DO
C
C
C     Calculando o vetor de medias das  variaveis
C     transformadas(Y)
c
      do I=1,2
       somay1(I)=0.0
       y1med(I)=0.0
       somay2(I)=0.0
       y2med(I)=0.0
        do J=1,100
         somay1(I)=somay1(I)+Y1T(I,J)*L1(J)
         SOMAY2(I)=SOMAY2(I)+Y2T(I,J)*L1(J)
        end do
       y1med(i)=somay1(I)/(1.0*100)
       y2med(i)=somay2(I)/(1.0*100)
      end do
c
c
c
c     Inversa das matrizes de covariancias  das combinacoes
c     segundo o nosso objetivo
c
      do 595 I=1,2
        do 595 j=1,2
          swi1(i,j)=0.0
          swi2(i,j)=0.0
595   continue
        swi1(1,1)=swi1(1,1)+1.0
        swi1(2,2)=swi2(2,2)+1.0
        swi2(1,1)=swi2(1,1)+1.0/R(6)
        swi2(2,2)=swi2(2,2)+1.0/R(1)
c
       do 220 I=1,2
        do 220 j=1,2
         difq(i,j)=0.0
          difq(i,j)=difq(i,j)+swi1(i,j)-swi2(i,j)
220    continue
c
```

```
C
C       CALCULANDO A FUNCAO DISCRIMINANTE
C
C       a) Calculando o Termo quadratico em y
C
        do I=1,100
          do J=1,2
           quay1(I,J)=0.0
           quay2(I,J)=0.0
              do K=1,2
               quay1(I,J)=quay1(I,J)+y1(I,K)*difq(K,J)
                quay2(I,J)=quay2(I,J)+y2(I,k)*difq(K,J)
              end do
          end do
        end do
         do i=1,100
            quadry1(I)=0.0
            quadry2(I)=0.0
              do j=1,2
               quadry1(I)=quadry1(I)+quay1(I,J)*y1t(J,I)
               quadry2(I)=quadry2(I)+quay2(I,J)*y2t(J,I)
            end do
        end do
c
c       b)Calculando o Termo LInear em y da Funcao Discriminante
c
                do I=1,2
                 s1(I)=0.0
                 s2(I)=0.0
                difl(I)=0.0
                  do J=1,2
                    s1(I)=s1(I)+swi1(I,J)*y1med(J)
                    s2(I)=s2(I)+swi2(I,J)*y2med(J)
                  end do
                difl(I)=difl(I)+s1(I)-s2(I)
                end do
               type *,'Calculou o vetor diferenga'
               do I=1,2
               type *,difl(I)
```

```
c
c       Calculando o produto 2*y*diferenca
c
            do I=1,100
             vliny1(I)=0.0
             vliny2(I)=0.0
              do J=1,2
                 vliny1(I)=vliny1(I)+2*y1(I,J)*dif1(J)
                 vliny2(I)=vliny2(I)+2*y2(I,J)*dif1(J)
              end do
            end do
c
c       c)Calculando o Termo Independente da fun. Discriminante
          ind1=0.0
          ind2=0.0
          indt=0.0
                ind1=ind1+y1med(1)*s1(1)+y1med(2)*s1(2)
                ind2=ind2+y2med(1)*s2(1)+y2med(2)*s2(2)
           indt=ind1-ind2-log(xx)
            type *,'TERMO INDEPENDENTE'
            type *,indt,ind1,ind2,log(xx)
c
c       f)Calculando a Funcao Discriminante e reclassificando os
c         Elementos das amostras
c
          do I=1,100
           discrim1(I)=0.0
           discrim2(I)=0.0
             discrim1(I)=discrim1(I)+quadry1(I)-vliny1(I)+indt
             discrim2(I)=discrim2(I)+quadry2(I)-vliny2(I)+indt
           if(discrim1(i).le.0) then
                                    indica1(i)=1
               else
                     indica1(i)=2
            end if
            if(discrim2(i).le.0) then
                                     indica2(i)=1
               else
                     indica2(i)=2
```

```
            end if
          end do
C
C
C     IMPRIMINDO E GERANDO OS RESULTADOS
C
      type 10
      write(33,10)
10    format(/,15x,'matriz de covariancias da amostra um')
      do I=1,6
      type 15,(varx1(i,j),j=1,6)
      write(33,15)(varx1(i,j),j=1,6)
15    format(/,15x,6(f10.6,3x))
      end do
       type 20
       write(33,20)
20     format(/,15x,'matriz de covariancias da amostra dois')
       do I=1,6
        type 25,(varx2(i,j),j=1,6)
        write(33,25)(varx2(i,j),j=1,6)
25      format(/,15x,6(f10.6,3x))
        end do
       type 31
       write(33,31)
31    format(/,15x,'matriz dados transformados das amostras um e dois')
        do I=1,100
        type 35,(y1(i,j),j=1,2),(y2(i,j),j=1,2)
        write(34,35)(y1(i,j),j=1,2),(y2(i,j),j=1,2)
35        format(/,10x,100(f15.6,2x))
        end do
         type 30
         write(33,30)
30       format(/,15x,'vetores de medias das variaveis  transformadas')
          do I=1,2
         type 32,y1med(I),y2med(I)
         write(33,32)y1med(I),y2med(I)
32       format(/,2(3X,f15.6))
          end do
      stop
       end
```

PROGRAMA Nº 6

DISCRIMINAÇÃO COM O PROC DISCRIM (SAS)

NOTAS DE BANCO VERDADEIRAS E FALSAS

```
OPTIONS LS=80;
TITLE 'ANALISE DISCRIMINANTE';
DATA SIMUL;
TITLE2 'MEDIDAS DAS NOTAS DE BANCO';
INFILE FLURY.DAT;
INPUT X1 X2  X3  X4  X5  X6   ESPEC 41  OBS 42-44;
IF ESPEC=1 THEN TIPO='PRIMEIRA POPULAÇÃO';
           ELTE TIPO='SEGUNDA POPULAÇÃO';
LABEL X1='COMPRIMENTO DAS NOTAS';
      X2='LARGURA DO LADO ESQUERDO';
      X3='LARGURA DO LADO DEREITO';
      X4='LARGURA DA MARGEM INFERIOR';
      X5='LARGURA DA MARGEM SUPERIOR';
      X6='COMP. DA DIAG. MED DESDE O CIE ATE CID';
  ESPEC ='POPULAÇÃO DE PROCEDENÇA';
    OBS ='NUMERO DE OBSERVAÇÃO';
PROC PRINT;
 PROC DISCRIM SIMPLE LISTERR POOL=TEST;
CLASS TIPO;
VAR X1-X6;
```

QUADRO A1

OBTENÇÃO DE AUTOVALORES E RAZÃO DE ERRO APARENTE PARA DIFERENTES PARAMETROS DE ENTRADA.

DIF ENT ANG	PARAMETROS DE ENTRADA						RESULTADOS AMOSTRAIS				
	POP. UM			POP. DOIS						AUTOVALORES	
	V_1	V_2	θ	V_1	V_2	θ	REA	n_{1m}	n_{2m}	MAX	MIN
15	3	0.5	0	0.5	3	15	0.090	7	11	37.680	0.038
15	3	0.5	15	0.5	3	30	0.090	7	11	37.680	0.038
15	3	0.5	45	0.5	3	60	0.090	7	11	37.680	0.038
15	3	0.5	90	0.5	3	105	0.090	7	11	37.680	0.038
15	3	0.5	180	0.5	3	195	0.090	7	11	37.680	0.038
45	3	0.5	0	0.5	3	45	0.135	11	16	22.060	0.066
45	3	0.5	45	0.5	3	90	0.135	11	16	22.060	0.066
45	3	0.5	90	0.5	3	135	0.135	11	16	22.060	0.066
60	3	0.5	60	0.5	3	120	0.185	14	23	12.564	0.115
60	3	0.5	120	0.5	3	180	0.185	14	23	12.564	0.115
90	3	0.5	0	0.5	3	90	0.480	32	64	1.427	1.016
90	3	0.5	90	0.5	3	180	0.480	32	64	1.427	1.016
90	3	0.5	270	0.5	3	360	0.480	32	64	1.427	1.016
90	3	0.5	180	0.5	3	270	0.480	32	64	1.427	1.016
180	3	0.5	0	0.5	3	180	0.095	8	11	39.646	0.037
180	3	0.5	180	0.5	3	360	0.095	8	11	39.645	0.037

FONTE : Simulações com o Programa nº 1

Dif Ent Ang : Diferença entre ângulos das duas populações.

QUADRO A2

ANÁLISE DISCRIMINANTE

MEDIDAS DAS NOTAS VERDADEIRAS E FALSAS

DISCRIMINANTE ANALYSIS TESTE HOMOGENEITY
OF WITHIN COVARIANCE MATRICES

NOTATION K = NUMBER OF GROUPS

P = NUMBER OF VARIABLES

N = TOTAL NUMBER OF OBSERVATIONS

NCI = NUMBER OF OBSERVATIONS IN THE I'TH GROUPS

$$V = \frac{\prod \left| \text{WITHIN SS MATRIX(I)} \right|^{NCI/2}}{\left| \text{POOLED SS MATRIX} \right|^{N/2}}$$

$$RHO = 1.0 - \left| \text{SUM} \frac{1}{NCI - 1} - \frac{1}{N-K} \right| \frac{2P^2+3P-1}{6(P+1)(K-1)}$$

$$DF = 0.5\,(K-1)(P+1)P$$

UNDER NULL HYPOTHESIS: $-2\,RHO\,LN\left[\frac{N^{PN/2}\,V}{\prod NCI^{PNCI/2}}\right]$,

IS DISTRIBUTED APPROXIMATELY AS CHI-SQUARE(DF)

TEST CHI-SQUARE VALUE = 121.76518909

WITH 21 DF PROB > 0.0001

Since the chi-square value is significant at the 0.1000 level, the within covariance matrices will be used in the discriminant function

---

QUADRO A3

RESULTADOS DA APLICAÇÃO CONSIDERANDO A MATRIZ $S_1^{-1}S_2$

Vetores de médias das variáveis transformadas

$$\bar{Y}^{(1)} = \begin{bmatrix} 167.018450 \\ -215.729188 \end{bmatrix} \qquad \bar{Y}^{(2)} = \begin{bmatrix} 163.043472 \\ -219.423832 \end{bmatrix}$$

Matriz diferença das inversas : $\hat{\Sigma}_1^{-1} - \hat{\Sigma}_2^{-2}$

$$\begin{bmatrix} 0.836612 1331 & 0.0000000000 \\ 0.0000000000 & -2.5224626115 \end{bmatrix}$$

Vetor $\hat{\Sigma}_1^{-1}\bar{Y}^{(1)}$ vetor $\hat{\Sigma}_2^{-1}\bar{Y}^{(2)}$

$$\begin{bmatrix} 167.018450 \\ -215.729188 \end{bmatrix} \qquad \begin{bmatrix} 26.639325 \\ -772.912244 \end{bmatrix}$$

$$\hat{\Sigma}_1^{-1}\bar{Y}^{(1)} - \hat{\Sigma}_2^{-1}\bar{Y}^{(2)} :$$

$$\begin{bmatrix} 140.379125 \\ 557.183086 \end{bmatrix}$$

INDT (Termo independente)

-99504

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NÚMERO OBSERVAÇÃO	DO	DECISÃO ESTATÍSTICA
1	001	-9.361799	Pertence à população = 1
1	002	- 89.070702	Pertence à população = 1
1	003	- 77.696114	Pertence à população = 1
1	004	- 76.234184	Pertence à população = 1
1	005	- 93.809860	Pertence à população = 1
1	006	- 9.245865	Pertence à população = 1
1	007	- 80.132515	Pertence à população = 1
1	008	- 79.546623	Pertence à população = 1
1	009	- 51.947102	Pertence à população = 1
1	010	- 14.196910	Pertence à população = 1
1	011	- 9.889459	Pertence à população = 1
1	012	- 77.280579	Pertence à população = 1
1	013	- 15.897683	Pertence à população = 1
1	014	- 55.116024	Pertence à população = 1
1	015	- 49.278294	Pertence à população = 1
1	016	- 86.727692	Pertence à população = 1
1	017	- 65.071686	Pertence à população = 1
1	018	- 75.321953	Pertence à população = 1
1	019	- 42.629009	Pertence à população = 1
1	020	- 47.286865	Pertence à população = 1
1	021	- 55.969963	Pertence à população = 1
1	022	- 27.675718	Pertence à população = 1
1	023	- 14.564607	Pertence à população = 1
1	024	- 33.444485	Pertence à população = 1
1	025	- 46.751476	Pertence à população = 1

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NÚMERO OBSERVAÇÃO	D(y)	DECISÃO ESTATÍSTICA
1	026	- 25.894316	Pertence à população = 1
1	027	- 45.246223	Pertence à população = 1
1	028	- 27.552956	Pertence à população = 1
1	029	- 53.766724	Pertence à população = 1
1	030	-104.838531	Pertence à população = 1
1	031	- 40.948692	Pertence à população = 1
1	032	- 96.996445	Pertence à população = 1
1	033	- 37.967319	Pertence à população = 1
1	034	- 19.236160	Pertence à população = 1
1	035	- 11.563610	Pertence à população = 1
1	036	- 35.746159	Pertence à população = 1
1	037	- 44.942091	Pertence à população = 1
1	038	- 69.329063	Pertence à população = 1
1	039	- 58.108063	Pertence à população = 1
1	040	- 67.711250	Pertence à população = 1
1	041	- 85.748825	Pertence à população = 1
1	042	- 38.390926	Pertence à população = 1
1	043	- 71.204636	Pertence à população = 1
1	044	- 28.270760	Pertence à população = 1
1	045	- 68.554138	Pertence à população = 1
1	046	- 76.821434	Pertence à população = 1
1	047	- 52.321270	Pertence à população = 1
1	048	- 58.958168	Pertence à população = 1
1	049	- 57.621872	Pertence à população = 1
1	050	-119.942833	Pertence à população = 1

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NUMERO OBSERVAÇÃO	EIXO	DECISÃO ESTATISTICA
1	051	- 77.4688941	Pertence à população = 1
1	052	- 16.1083868	Pertence à população = 1
1	053	- 30.9928820	Pertence à população = 1
1	054	- 30.2783490	Pertence à população = 1
1	055	- 85.2130531	Pertence à população = 1
1	056	- 68.8091568	Pertence à população = 1
1	057	- 21.7076824	Pertence à população = 1
1	058	- 79.1203854	Pertence à população = 1
1	059	- 28.6295576	Pertence à população = 1
1	060	- 43.6885576	Pertence à população = 1
1	061	- 72.3800074	Pertence à população = 1
1	062	- 61.3398400	Pertence à população = 1
1	063	- 48.0142882	Pertence à população = 1
1	064	- 42.7268860	Pertence à população = 1
1	065	- 32.0657708	Pertence à população = 1
1	066	- 7.7373995	Pertence à população = 1
1	067	- 77.1976389	Pertence à população = 1
1	068	- 65.4190037	Pertence à população = 1
1	069	- 67.5913862	Pertence à população = 1
1	070	- 4.0303286	Pertence à população = 1
1	071	- 30.741774	Pertence à população = 1
1	072	- 48.3646889	Pertence à população = 1
1	073	- 49.2341104	Pertence à população = 1
1	074	- 67.6476444	Pertence à população = 1
1	075	- 64.5178296	Pertence à população = 1

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NÚMERO OBSERVAÇÃO	D(y)	DECISÃO ESTATÍSTICA
1	076	- 75.1888622	Pertence à população = 1
1	077	- 40.5352755	Pertence à população = 1
1	078	- 56.5942068	Pertence à população = 1
1	079	- 16.3492933	Pertence à população = 1
1	080	- 73.8199869	Pertence à população = 1
1	081	- 36.7169961	Pertence à população = 1
1	082	- 55.9988283	Pertence à população = 1
1	083	- 65.7309881	Pertence à população = 1
1	084	- 31.6888032	Pertence à população = 1
1	085	- 3.8948633	Pertence à população = 1
1	086	- 39.6166070	Pertence à população = 1
1	087	- 60.6439033	Pertence à população = 1
1	088	- 71.2378233	Pertence à população = 1
1	089	- 29.8950422	Pertence à população = 1
1	090	- 55.2578933	Pertence à população = 1
1	091	- 75.5834500	Pertence à população = 1
1	092	- 48.7269711	Pertence à população = 1
1	093	- 79.7311022	Pertence à população = 1
1	094	- 75.1966322	Pertence à população = 1
1	095	- 77.8648833	Pertence à população = 1
1	096	- 51.5405433	Pertence à população = 1
1	097	- 16.3038433	Pertence à população = 1
1	098	- 62.5470055	Pertence à população = 1
1	099	- 31.5088945	Pertence à população = 1
1	100	- 60.3293500	Pertence à população = 1

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NUMERO OBSERVAÇÃO	D(y)	DECISÃO ESTATISTICA
2	001	24.5136405	Pertence à população = 2
2	002	47.5781582	Pertence à população = 2
2	003	6.6886653	Pertence à população = 2
2	004	14.9446306	Pertence à população = 2
2	005	61.1222597	Pertence à população = 2
2	006	23.0761159	Pertence à população = 2
2	007	22.6666847	Pertence à população = 2
2	008	21.6376608	Pertence à população = 2
2	009	31.2872379	Pertence à população = 2
2	010	30.5341107	Pertence à população = 2
2	011	7.3945309	Pertence à população = 2
2	012	40.6179654	Pertence à população = 2
2	013	26.9337705	Pertence à população = 2
2	014	52.3045320	Pertence à população = 2
2	015	24.8778388	Pertence à população = 2
2	016	6.2857749	Pertence à população = 2
2	017	52.7981765	Pertence à população = 2
2	018	60.4648900	Pertence à população = 2
2	019	42.1422982	Pertence à população = 2
2	020	37.7782082	Pertence à população = 2
2	021	38.5588918	Pertence à população = 2
2	022	54.3815323	Pertence à população = 2
2	023	44.7455664	Pertence à população = 2
2	024	30.7878955	Pertence à população = 2
2	025	5.5766639	Pertence à população = 2

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NUMERO OBSERVAÇÃO	D(X)	DECISÃO ESTATISTICA
2	026	28.5722170	Pertence à população = 2
2	027	17.7573433	Pertence à população = 2
2	028	33.9964060	Pertence à população = 2
2	029	24.6288046	Pertence à população = 2
2	030	35.0027754	Pertence à população = 2
2	031	23.1378838	Pertence à população = 2
2	032	84.3258851	Pertence à população = 2
2	033	40.8861139	Pertence à população = 2
2	034	44.6895822	Pertence à população = 2
2	035	46.9847059	Pertence à população = 2
2	036	67.1104811	Pertence à população = 2
2	037	31.7669889	Pertence à população = 2
2	038	14.9985388	Pertence à população = 2
2	039	45.4041102	Pertence à população = 2
2	040	45.5678852	Pertence à população = 2
2	041	31.3682217	Pertence à população = 2
2	042	17.9809288	Pertence à população = 2
2	043	40.2966474	Pertence à população = 2
2	044	32.3723911	Pertence à população = 2
2	045	23.7711164	Pertence à população = 2
2	046	53.0336111	Pertence à população = 2
2	047	43.7221137	Pertence à população = 2
2	048	15.6366989	Pertence à população = 2
2	049	31.2665667	Pertence à população = 2
2	050	36.2051143	Pertence à população = 2

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NUMERO OBSERVAÇÃO	D(y)	DECISÃO ESTATISTICA
2	051	56.8715506	Pertence à população = 2
2	052	39.9804500	Pertence à população = 2
2	053	20.0206999	Pertence à população = 2
2	054	54.9478334	Pertence à população = 2
2	055	52.0645664	Pertence à população = 2
2	056	41.6334590	Pertence à população = 2
2	057	29.6001187	Pertence à população = 2
2	058	40.5023316	Pertence à população = 2
2	059	55.0918992	Pertence à população = 2
2	060	15.5621108	Pertence à população = 2
2	061	13.2841176	Pertence à população = 2
2	062	10.5723488	Pertence à população = 2
2	063	44.2220356	Pertence à população = 2
2	064	50.0483221	Pertence à população = 2
2	065	43.5009888	Pertence à população = 2
2	066	39.3692274	Pertence à população = 2
2	067	13.6940113	Pertence à população = 2
2	068	14.9555220	Pertence à população = 2
2	069	24.8911947	Pertence à população = 2
2	070	47.6040920	Pertence à população = 2
2	071	14.4465111	Pertence à população = 2
2	072	53.5884599	Pertence à população = 2
2	073	48.0801588	Pertence à população = 2
2	074	32.4094390	Pertence à população = 2
2	075	38.1496810	Pertence à população = 2

QUADRO A4

OBSERVAÇÕES DAS AMOSTRAS DE NOTAS VERDADEIRAS E FALSAS AVALIADAS NA FUNÇÃO DE DISCRIMINAÇÃO PROPOSTA

GRUPO VERDADEIRO	NÚMERO OBSERVAÇÃO	D(y)	DECISÃO ESTATÍSTICA
2	076	50.8553839	Pertence à população = 2
2	077	48.0468829	Pertence à população = 2
2	078	35.0059943	Pertence à população = 2
2	079	36.8576609	Pertence à população = 2
2	080	12.8788863	Pertence à população = 2
2	081	38.7828894	Pertence à população = 2
2	082	12.2959928	Pertence à população = 2
2	083	43.474178	Pertence à população = 2
2	084	28.797161	Pertence à população = 2
2	085	33.747467	Pertence à população = 2
2	086	45.8150114	Pertence à população = 2
2	087	11.6458833	Pertence à população = 2
2	088	38.2470882	Pertence à população = 2
2	089	22.6487733	Pertence à população = 2
2	090	41.0049110	Pertence à população = 2
2	091	36.8943002	Pertence à população = 2
2	092	12.8792881	Pertence à população = 2
2	093	34.7583063	Pertence à população = 2
2	094	14.1789119	Pertence à população = 2
2	095	46.5576697	Pertence à população = 2
2	096	33.2019904	Pertence à população = 2
2	097	31.6333766	Pertence à população = 2
2	098	33.2218828	Pertence à população = 2
2	099	45.187294	Pertence à população = 2
2	100	30.8120775	Pertence à população = 2

QUADRO A5

ANÁLISE DISCRIMINANTE SEGUNDO O SAS

NÚMERO DE OBSERVAÇÕES E PERCENTAGENS DE CLASSIFICAÇÃO

População de Origem	Decisão Estatística Π_1	Π_2	TOTAL
Π_1	99	1	100
	99.00	1.00	100.00
Π_2	0	100	100
	0.00	100.00	100.00
TOTAL	99	101	200
PERCENTAGEM	49.50	50.50	100.00

Fonte: Dados do Plano Piloto analisados com o SAS

REFERENCIA BIBLIOGRAFICA

1.-ANDERSON, T. W. (1958) An Introduction to Multivariate Statistical Analysis. *Wiley & Sons, New York.*

2.-ANDERSON, T. W. (1973) An asymptotic expansion of the distribution of the Studentized Classification Statistics. *The Annals of Statistics.* 1:964-72.

3.-BOX, G. E. P. AND MULLER, M. E. (1958) A Note on the Generation of Random Normal Deviates. *The annals of Mathematical Statistics.* 29:610-11.

4.-FISHER, R. A. (1936) The use of multiple measurement in taxonomic problems. In: Atchley, W. R. (1975) Multivariate Statistical Methods: Among-Groups Covariation. *Dowden, Huchinson, Stroudsburg, Pennsylvania.*

5.-FLURY, B. AND RIEDWYL, H. (1983) Angewandte Multivariate Statistik. *Gustav Fischer, Stattgart. New York.*

6.-FLURY, B. (1983) Some relations between the comparison of covariance matrices and principal component analysis. *Computational Statistics & Data Analysis.* 1: 97-109.

7.-FLURY, B. N. (1985) Analysis of Linear Combinations with Extreme Ratios of Variance. *Journal of the American Statistical Association.* 80:915-922.

8.-JOHNSON, R. A. AND WICHERN, D. W. (1982) Applied Multivariate Statistical Analysis. *Prentice Hall, Glewood Cliffs.*

9. -KRUSKAL, W. H. AND TANUR, J. M. (1978) International Encyclopedia of Statistics. 1:628-35.

10. -LACHENBRUCH, P. A. (1975) Discriminant Analysis. *Hafner Press, New York.*

11. -LACHENBRUCH, P. A. AND MIKEY, M. R. (1968) Estimation of error rates in discriminant analysis. *Technometrics.* 10:1-10.

12. -MARDIA, K. V.; KENT, J. T. AND BIBBY, J. M. (1979) Multivariate Analysis. *Academic Press, London.*

13. -PAYNE, W. H.; RABUNG, J. R. AND BOGYO, T. P. (1969) Coding the Lehmer Pseudo-random Number Generator. *Communications of the ACM.* 12(2):85-86.

14. -RAO, C. R. AND MITRA, S. K. (1971). Generalized Inverse of Matrices and its Applications. *Wiley & Sons, New York.*

15. -ROY, S. N. (1957) Some Aspects of Multivariate Analysis. *Wiley & sons. New York.*

16. -WILKINSON, J. H. AND REINSCH, C. (1971). Handbook for Automatic Computation.
Linear Algebra, Springer-Verlag. Vol II.