

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CAMPUS DE BOTUCATU

**DESENVOLVIMENTO DE APLICATIVO PARA O MÉTODO DE
DISCRIMINAÇÃO DE FISHER E SEU USO NA EXPERIMENTAÇÃO
AGRONÔMICA.**

CARLOS ROBERTO PEREIRA PADOVANI

Dissertação apresentada à Faculdade de
Ciências Agronômicas da UNESP - Câmpus
de Botucatu, para obtenção do título de
Mestre em Agronomia - Área de
Concentração em Energia na Agricultura.

BOTUCATU-SP
Abril/2004

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CAMPUS DE BOTUCATU

**DESENVOLVIMENTO DE APLICATIVO PARA O MÉTODO DE
DISCRIMINAÇÃO DE FISHER E SEU USO NA EXPERIMENTAÇÃO
AGRONÔMICA.**

CARLOS ROBERTO PEREIRA PADOVANI

Orientador: Prof. Dr. Flávio Ferrari Aragon

Dissertação apresentada à Faculdade de
Ciências Agronômicas da UNESP - Câmpus
de Botucatu, para obtenção do título de
Mestre em Agronomia - Área de
Concentração em Energia na Agricultura.

BOTUCATU-SP
Abril/2004

Dedico...

Aos meus queridos pais, Carlos e Silvia, que sempre lutaram para que os filhos pudessem realizar seus sonhos.

A minha noiva Juliana Lara , por estar sempre ao meu lado.

Agradeço...

ao Prof. Dr. Flávio Ferrari Aragon pela valiosa orientação, incentivo e paciência;

aos colegas Ana Helena Neuber de Oliveira e José Higino de Oliveira pelo auxílio na construção do programa computacional;

ao Diretor de Serviço José Roberto Sprandim por possibilitar o uso da tecnologia da FATEC para o desenvolvimento do programa de mestrado;

ao Diretor da FATEC, Prof. Antonio Carlos de Oliveira, por ter propiciado a possibilidade de cursar o programa de mestrado;

aos funcionários da Seção de Pós-graduação pelos serviços prestados;

aos professores do Depto de Bioestatística pelo apoio;

a todos que de certa forma contribuía para a conclusão deste trabalho.

SUMÁRIO

	Página
1. RESUMO.....	1
2. SUMMARY.....	3
3. INTRODUÇÃO.....	5
4. REVISÃO BIBLIOGRÁFICA.....	8
5. DESENVOLVIMENTO METODOLÓGICO.....	18
5.1. Método de Discriminação de Fisher.....	18
5.2. Programa Computacional.....	33
5.3. Manual do Usuário.....	35
5.4. Exemplo de Aplicação.....	39
6. CONCLUSÕES.....	47
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	48
8. APÊNDICE.....	55
A1. Quadro das respostas das características quantitativas do girassol.....	56
A2. Listagem dos procedimentos no MAPLE.....	59
A3. Listagem dos Procedimentos no CLIPPER 5.....	65

1. RESUMO

Nas Ciências Agronômicas, em particular na Energia na Agricultura, encontram-se várias situações em que são observadas diversas variáveis respostas nas parcelas ou unidades experimentais. Nestas situações, um caso de interesse prático à experimentação agronômica é o que considera o estudo das regiões de similaridade entre as parcelas com atenção especial à classificação de novas unidades experimentais.

Um procedimento bastante robusto para o estudo das similaridades em situações multivariadas consiste no método de discriminação de Fisher entre várias populações.

Na literatura Agronômica, pouco se tem encontrado sobre o uso deste procedimento. Entende-se que este fato deve estar relacionado aos procedimentos algébricos e matriciais requeridos na construção do modelo matemático gerador das regiões e, principalmente, pela falta de um programa computacional de fácil manuseio para pesquisadores de áreas aplicadas.

Neste sentido, desenvolveu-se um programa computacional para o método de Fisher, acessível e de fácil manuseio para pesquisadores de áreas aplicadas, complementado com a construção do manual do usuário. Para exemplificar o uso do programa, consideraram-se dados relativos a experimentos desenvolvidos na EMBRAPA,

região de Londrina – PR, envolvendo seis diferentes variedades de girassol (*Helianthus annuus*) e cinco caracteres quantitativos da planta. O uso da discriminação de Fisher possibilitou a construção gráfica das regiões de classificação, segundo as diversidades genéticas do girassol, o que apresentou uma alta porcentagem de retenção de informação da variabilidade associada a uma baixa taxa de classificação errônea.

Palavras-chave: Discriminação de Fisher, regiões de classificação, discriminação gráfica.

APPLICATION DEVELOPMENT FOR THE DISCRIMINATION METHOD OF FISHER AND IT'S USE IN THE AGRONOMICAL EXPERIMENTATION. Botucatu, 2004. 102p. Dissertação (Mestrado em Agronomia/Energia na Agricultura) – Faculdade de Ciências Agronômicas, Universidade Estadual Paulista.

Author: CARLOS ROBERTO PEREIRA PADOVANI

Adviser: FLÁVIO FERRARI ARAGON

2. SUMMARY

In the Agronomical Sciences, particularly in the Energy in the Agriculture, there are several situations in which many answer variables in the experimental parcels or units can be observed. In these situations, a case of practical interest to the agronomical is the study of the region similarities among the parcels, giving special attention to the classification of new experimental units.

A very robust procedure for the discrimination of several multivariate populations is the Fisher's graphic method.

In the Agronomical literature, not much is found about the use of this procedure. This fact may be related to algebraic and matricial procedures required in the construction of the mathematical model generator of the regions and, mostly, by the lack of a friendly computational software for researchers of applied areas.

In this sense, the objective is to develop a computational software for Fisher's Method, which must be accessible and of easy handling for researchers of applied areas, complementing it with the creation of the user's manual and presenting applications for the software in the rational use of energy. Datas related to experiments developed at EMBRAPA, Londrina region PR, are considered, involving six different sunflower variables (*Helianthus annuus*) and five quantitative marks of the plant. Fisher's discrimination use

enable the graphic construction of the classification regions, according to the sunflower genetic diversities, which presented a high percentage of information retention of the variability associated to a low rate of erroneous classification.

Keywords: Fisher discrimination; classification regions; graphic discrimination

3. INTRODUÇÃO

Em pesquisas biológicas, é comum que sejam efetuadas diversas medidas observacionais na mesma parcela ou unidade experimental. O vetor resposta (p -dimensional), composto de p características ou variáveis aleatórias observadas nas parcelas de cada uma das g populações em estudo, apresenta uma estrutura de variação intra e intervariáveis que pode ser dada pela matriz positiva definida, de ordem p , das variâncias e covariâncias. Essa estrutura do vetor de observação não deixa de refletir a situação biológica, que é um sistema altamente integrado, onde as características e/ou componentes estão inteiramente relacionados por meio de suas dependências e interdependências.

A Análise Discriminante, proposta por Sir Ronald Aymer Fisher (Fisher, 1936), o "Arquiteto da Análise Multivariada" segundo Rao (1964), consiste em construir uma função linear das p variáveis aleatórias quantitativas que classifique um novo indivíduo em uma das g populações multivariadas com probabilidade de má classificação minimizada, ou seja, que possibilite a máxima separação entre as populações (Morrison, 1976; Anderson, 1984). Quando o vetor resposta envolve variáveis mistas (quantitativas e qualitativas), é mais comum utilizar na análise discriminante a regressão logística, as árvores de classificação ou as redes neurais artificiais (Barros & Artes, 2003).

A construção matemática da função discriminante linear de Fisher envolve tópicos de Álgebra Matricial e Análise Numérica - desde o uso dos operadores de Lagrange, autovalores e autovetores de uma matriz, até a resolução de sistemas lineares, entre outros (Graybill, 1969). Quanto maior for o número de variáveis observadas nas parcelas, maiores dificuldades operacionais apresentam-se no contexto matemático, tanto no cálculo numérico como no matricial.

Pimentel Gomes (1984) previa que a análise multivariada, principalmente a MANOVA (Análise de Variância Multivariada), mostrava-se bastante promissora para a Experimentação Agronômica. Porém, alertava que o desenvolvimento de procedimentos alternativos aos convencionais estavam associados à evolução da tecnologia computacional. Esta premissa, embora verdadeira, não se mostra suficientemente simples frente ao despreparo dos usuários no manuseio de programas computacionais ("pacotes") sofisticados e/ou especializados.

Tornero (1990) e Messetti (2000), conhecendo situações da experimentação zootécnica e agronômica onde esse fato se configura, desenvolvem programas computacionais de fácil manuseio e mostram aplicações para alguns procedimentos multivariados que tratam de relações em um conjunto de variáveis, especificamente para o estudo das dissimilaridades pela análise de agrupamento ("Cluster Analysis") e análise de componentes principais ("Principal Component Analysis").

Particularmente, quanto ao problema relacionado com a classificação de um indivíduo em uma de várias populações, a técnica da discriminação de Fisher, sem perda de generalidade da proposta inicial e com a retenção de grande parte da informação de variação dos dados, torna-se um procedimento bastante interessante como um método alternativo e simples de alocação de indivíduos. O método exige pouco dos tópicos avançados de matemática e pode ser utilizado sem a necessidade da distribuição multinormal de probabilidades dos dados. Para o estabelecimento do critério de alocação dos indivíduos, com base em medidas de um número p de características do mesmo, em uma das populações distintas, são constituídas as regiões de discriminação máxima entre os grupos e os eixos discriminantes são as funções de alocação dos indivíduos nas respectivas regiões.

Porém, não pode ser desconsiderado que a qualidade da discriminação, quanto à precisão dos resultados, está diretamente ligada à quantidade de

retenção da informação embutida na função linear geradora dos eixos discriminantes (Lachembruch,1967).

O procedimento matemático do método de Fisher para geração dos eixos discriminantes consiste basicamente na redução do espaço dimensional das p características observadas nos dados, em sistemas ortogonais com dimensões menores. Do ponto de vista prático, quando a redução torna-se possível ao sistema bidimensional ou no máximo tridimensional, a representação dos eixos discriminantes pode ser realizada por meio gráfico, o que propicia aos pesquisadores das áreas aplicadas um subsídio visual interpretativo das respostas biológicas e de fácil entendimento prático.

No contexto dos objetivos da análise discriminante, podem ser destacadas: a discriminação entre populações e a alocação ou classificação de um novo indivíduo na população de maior afinidade biológica.

Neste sentido, objetiva-se no presente estudo, elaborar um programa computacional em linguagem de alto nível, para a obtenção dos eixos discriminantes, que seja de fácil acesso e simples manuseio para pesquisadores das áreas biológicas e utilizar este *software* em um exemplo de aplicação na agronomia envolvendo caracteres quantitativos de várias espécies de girassol (*Helianthus annuus*) originárias da América do Sul.

4. REVISÃO BIBLIOGRÁFICA

Pretende-se, nesta revisão, abordar situações práticas em que os dados coletados pelo pesquisador envolvem múltiplas medidas efetuadas sobre uma mesma unidade experimental (parcela). Neste contexto multivariado para a análise dos dados, procura-se destacar a técnica de classificação ou discriminação de unidades amostrais em assuntos relacionados à área de ciências agroflorestais e outras afins.

Jolicoeur & Mosimann (1960) aplicam a análise dos componentes principais para o estudo da variação da largura, comprimento e altura de carapaças de tartarugas. Discutem, com pioneirismo biológico, aspectos numéricos e geométricos inseridos no primeiro componente principal da matriz de correlação, o qual descreve com muita consistência e eficácia a forma e o tamanho dos animais expresso no volume da carapaça. Uma interpretação mais detalhada das funções morfométricas estabelecidas nos eixos descritores permitem estudar o crescimento da espécie com maior acurácia que o modelo curvilíneo de resposta.

Crocci (1979) apresenta uma extensa revisão sobre a obtenção da probabilidade de má-classificação baseando-se na função discriminante linear de Fisher para

duas populações, ilustrando a construção da função e a estimação das taxas de erro considerando 12 caracteres quantitativos mensurados em duas espécies de abelhas (*Partamona testacea* e *Partamona pseudomusarum*). A comparação numérica das probabilidades de má classificação mostrou que o procedimento de Wald, usando a distância generalizada de Mahalanobis, foi a que apresentou a menor taxa de classificação errônea.

Everitt (1979), preocupado com o crescimento das técnicas multidimensionais e seu largo uso por pesquisadores das áreas aplicadas, em muitas situações com impropriedades de utilização, realiza uma ampla discussão dos procedimentos multivariados e compara algumas análises de interdependência (aquelas que consideram as relações das variáveis entre si) com as de dependências (relação de uma ou mais variáveis com as outras). Destacam-se, nesta abordagem, as análises de agrupamento, componentes principais, de fatores, discriminante, de variância e os métodos gráficos.

Curi & Padovani (1980) apresentam uma revisão teórica da análise de discriminação ou classificação e complementam o estudo considerando quatro exemplos biológicos. No primeiro, são envolvidas cinco espécies de psitacíformes consideradas sob oito características biológicas; no segundo, 10 espécies de columbíformes com 10 características; no terceiro, três espécies de abelhas com 12 indicadores morfométricos e, no último, dois grupos bovinos com quatro variáveis observadas.

Smith (1980) discute a diferença entre a análise de variância (ANOVA) e a análise de componentes principais observando que, embora as duas técnicas estabeleçam a partição da soma de quadrados dos desvios totais em partes ortogonais, a primeira utiliza um critério de divisão externo, definido pelas hipóteses para as quais o experimento foi planejado. Na segunda, esse critério é determinado somente pela variação interna dos dados. Mostra ainda que uma análise mista, considerando componentes principais e análise de variância, pode ser apropriada e oportuna para vários conjuntos de dados, principalmente dados provenientes de delineamentos experimentais. Os testes das hipóteses relativas ao modelo experimental podem ser avaliados a partir dos dados transformados pelos primeiros componentes principais.

Rao *et al.* (1981) utilizam a análise discriminante canônica no estudo da divergência genética de arroz. A discriminação foi realizada no espaço bidimensional gerado pelas duas primeiras variáveis canônicas (87,5% de retenção da variação total dos

dados) construídas considerando nove características biológicas e a matriz de dispersão dos dados.

Curi (1982) discute a aplicação da análise de agrupamento complementada com a ordenação pelos componentes principais em experimentos biológicos. Comenta que a análise de variância multivariada e a análise discriminante de Fisher dependem, para sua aplicação, de pressuposições muito rígidas, como a normalidade dos dados e a homogeneidade das matrizes de covariâncias. Estas pressuposições nem sempre estão presentes em pesquisas clínicas pois a heterogeneidade das unidades amostrais é um aspecto comum e perfeitamente caracterizado, biologicamente, na diversidade individual e de fatores que podem influenciar as variáveis. Nesta linha de aplicação dos procedimentos multivariados, apresenta a ordenação gráfica pelos eixos discriminantes (discute até os três primeiros componentes principais) e como uma alternativa exploratória dos dados observacionais para definição dos agrupamentos e como alternativa indispensável para posterior análise inferencial dos dados.

Misra & Ni (1983) consideram duas espécies de peixes (*Sebastes fasciatus* e *Sebastes mentela*) e várias características morfométricas e, utilizando procedimento estatístico univariado, buscaram a diferenciação das espécies. O estudo isolado para cada um dos caracteres mostrou-se incapacitado para provar a associação morfométrica. A dificuldade foi prontamente contornada quando se utilizou um procedimento multivariado, mais especificamente, a discriminação linear de Fisher. O resultado estatístico mostrou que, na discriminação dos peixes vermelhos, um caracter não diferencia as espécies, mas um único critério envolvendo toda a estrutura de variabilidade permite a separação e, além disso, revela quais foram os indicadores morfométricos responsáveis para a conclusão biológica.

Cruz (1990) enfoca a aplicação de alguns métodos multivariados no melhoramento de plantas, com destaque especial às técnicas de discriminação, dissimilaridade e agrupamentos. Enfatiza que as técnicas multidimensionais são procedimentos altamente eficientes para o melhoramento, pois ao combinar as múltiplas informações de uma unidade experimental permitem a seleção genética e a predição de fenômenos biológicos.

Mallants & Feyen (1990) utilizam a análise dos componentes principais para composição de regiões homogêneas de precipitação no oeste da Bélgica e norte da França. Para a construção dos eixos discriminantes, consideraram-se as precipitações de

três anos de estudo: um declarado seco (1973), outro chuvoso (1977) e, por fim, um moderado (1978). O primeiro eixo, relativo ao primeiro componente principal, reteve 75% da variação total dos dados (três quartos de toda informação) e permitiu dividir a área de estudo em quatro regiões homogêneas de precipitação.

Alvarez *et al.* (1991) utilizam técnicas multivariadas (análise de variância, análise de componentes principais e análise de agrupamento) para estimar o potencial da cultura e a variabilidade genética de girassol (*Helianthus annuus L.*) em 20 populações de diferentes origens geográficas. A descrição pela análise de componentes principais assemelhou-se à análise de agrupamento utilizando-se a técnica não hierárquica com a similaridade baseada na distância euclidiana, resultando em três grupos semelhantes quanto à origem geométrica. Destaca-se nos resultados, um grupo de genótipos da Rússia com mais elevadas porcentagens de óleo e produção de grãos.

Basnet (1993) utiliza a técnica dos componentes principais para estudar as relações entre fatores ligados ao meio ambiente e os padrões de distribuição de árvores em florestas úmidas subtropicais. O estudo desenvolvido em Porto Rico, na Floresta Experimental de Luquilo, considera dados envolvendo variáveis geológicas, edáficas e ambientais. Os eixos discriminantes indicam que a associação entre os fatores edáficos e a geologia local complementados com alterações exógenas estabelecem o padrão de desenvolvimento florestal.

Fonseca (1993) descreve a importância da análise multivariada para os dados experimentais e destaca que, embora o surgimento das técnicas multidimensionais tenha se dado há algum tempo, a utilização em maior escala acompanhou a evolução dos recursos computacionais. Exemplifica seu uso na experimentação agronômica, considerando a discriminação do germoplasma de feijão (*Phaseolus vulgaris L.*) pelos componentes principais.

Demey *et al.* (1994) consideram 529 fazendas de cultivo de arroz inseridas no programa do Sistema de Irrigação do Rio Guarico (Venezuela) e descrevem, utilizando a discriminação gerada pela análise dos componentes principais, a classificação das fazendas em quatro categorias de qualidade de gerenciamento. Para a geração da função discriminante, foram utilizadas 24 variáveis respostas sendo que, entre elas, as mais contributivas para o processo de classificação foram: área territorial da fazenda, taxas e origens de fertilizantes aplicados no solo e tipo de irrigação.

Piassi *et al.*(1995) estudam o desempenho produtivo de três linhagens de aves legorne (ESA, GESA e LD), seus respectivos cruzamentos (ESA x GESA, ESA x LD e GESA x LD) e duas marcas comerciais, por meio de procedimentos multivariados. Oito características de importância econômica, avaliadas durante a fase de produção de ovos, foram comparadas por meio da análise de variância multivariada, segundo o teste estatístico da maior raiz característica de Roy com as respectivas comparações múltiplas entre os diferentes grupos genéticos. A análise discriminante dos grupos apresentou como resultado os melhores desempenhos para as marcas comerciais em relação às aves provenientes da Universidade Federal de Viçosa – MG, além de indicar que as linhagens ESA, GESA e LD necessitam de um trabalho intenso de seleção com base nas características peso médio do ovo, taxa de postura e massa de ovos para atingirem um nível competitivo.

Piassi *et al.* (1995a) consideram oito grupos genéticos de aves de postura observados em relação a características de importância econômica para idade de produção de ovos (peso corporal, consumo alimentar, idade ao primeiro ovo, taxa de postura, peso médio do ovo, massa de ovo, massa de ovo/unidade do tamanho metabólico e viabilidade) avaliadas em dois momentos sucessivos de oito semanas. Foram utilizados vários procedimentos multivariados de análise estatística: MANOVA (análise de variância multivariada), análise de agrupamento utilizando como coeficiente de similaridade a distância generalizada de Mahalanobis e o método de otimização de Tocher considerando a matriz de distância entre pares de genótipos e, finalmente, a análise canônica. Em relação à análise canônica, os dois primeiros eixos discriminantes respondem por mais de 92% da variação total observada em ambos os períodos estudados e a representação bidimensional nos eixos canônicos indicaram dois grupos distintos, com alto grau de divergência genética. O grau de divergência genética entre os seis grupos genéticos remanescentes foi baixo.

Vida *et al.* (1996) consideram a técnica dos componentes principais como suporte da análise de otimização numérica e buscam a resposta ótima para fertilização com os nutrientes NPK, em diferentes níveis de irrigação, de cultivares de trigo direcionados para o processo de panificação. Os eixos canônicos permitem discriminar várias composições entre os níveis de NPK e irrigação quanto à qualidade final do produto na transformação em pão.

Camarano (1997) estuda a divergência genotípica entre populações de girassol considerando quatro experimentos e 11 características biológicas da planta. Na discussão dos dados experimentais duas técnicas multivariadas são utilizadas: a análise de dispersão gráfica por variáveis canônicas e a proximidade dos centróides médios pela distância generalizada de Mahalanobis. Conclui-se que, na comparação dos métodos de estimação da divergência genotípica dos girassóis, as duas técnicas multidimensionais empregadas conduziram aos mesmos agrupamentos.

Vidigal *et al.* (1997) estudam a divergência genética de cultivares da mandioca mensurando 10 características morfoagronômicas e duas relacionadas à qualidade das raízes. A análise multivariada, mais especificamente as variáveis canônicas e a distância generalizada de Mahalanobis, possibilita gerar as regiões de semelhanças entre as cultivares, produzindo condições de classificar genótipos superiores em gerações avançadas e discriminar quais características contribuíram pouco para as alterações genéticas.

Lucio *et al.* (1999) investigam o regime climático da precipitação, temperatura e umidade relativa do ar à superfície na região metropolitana de Belo Horizonte. Para o estudo, utilizaram-se dados coletados durante 30 anos. As seguintes análises estatísticas são consideradas no estudo: modelo quadrático de regressão, análise multivariada (MANOVA), análise discriminante, entre outras. Os resultados mostram que, na caracterização do clima de Belo Horizonte, nenhuma das variáveis consideradas pode ser julgada como de baixa contribuição.

Morgano *et al.* (1999) analisam duas metodologias diferentes de preparação de amostra para determinação da concentração de minerais em sucos de frutas e indicam qual a mais adequada na análise de sete diferentes variedades de suco. Os dados coletados foram submetidos às seguintes técnicas de análise estatística: componentes principais, análise de variância multivariada, agrupamento e a distância euclidiana. A análise de componentes principais proporcionou uma interpretação multivariada interessante dos resultados, a qual mostra que as amostras são agrupadas de acordo com os teores de minerais.

Daher *et al.* (2000) consideram um experimento multivariado delineado em blocos completos casualizados para avaliar clones de capim-elefante (*Pennisetum purpurum Schum.*) quanto à produção de matéria seca, proporção de folhas nas plantas, altura das plantas, diâmetro do colmo e número de perfilhos. O experimento foi

realizado examinando 15 novos clones de capim-elefante e duas cultivares testemunhas (Mineiro e Taiwan A – 146) e os dados observados submetidos a variáveis canônicas, distância generalizada de Mahalanobis e método de agrupamento de Tocher. Os resultados dos procedimentos uni e multidimensionais demonstraram diferenciação entre os clones pelo diâmetro do colmo e perfilhamento.

Fonseca *et al.* (2000) avaliam o desempenho das três principais raças suínas utilizadas nos programas de melhoramento com relação às características reprodutivas e às divergências genéticas entre elas. As informações coletadas foram submetidas às seguintes análises estatísticas: variáveis canônicas, análise de variância multivariada (MANOVA), teste F, média harmônica, teste de Roy e a distância de Mahalanobis. Para uma análise complementar, foi formulada a função linear discriminante de Fisher. Os resultados obtidos demonstraram que as raças Landrace e Large White apresentam semelhanças genéticas quando comparadas à raça Duroc. Apresentam também, as vantagens na obtenção de fêmeas de primeiro cruzamento para posterior acasalamento com machos Duroc.

Messetti & Padovani (2000), observando a variabilidade genotípica de 12 populações de girassol, a partir de características agronômicas das fases vegetativa e reprodutiva, apresentam procedimento multivariado para identificar as similaridades biológicas entre espécies utilizando como indicador de semelhança a distância generalizada de Mahalanobis. São destacados dois aspectos conclusivos ao avanço tecnológico do melhoramento genético: o uso da distância generalizada de Mahalanobis como coeficiente de dissimilaridade e a robustez gráfica da primeira variável canônica na divergência genotípica.

Dantas *et al.* (2001) avaliam o efeito do alumínio (Al), em solução nutritiva, no crescimento e desenvolvimento da parte aérea, raízes e cultivares porta-enxertos de macieira, por meio de sete variáveis, classificando os clones quanto à tolerância ao alumínio. Destacam-se nos procedimentos quantitativos os métodos de análise de variância multivariada, análise de componentes principais e análise de agrupamento hierárquico. A análise multivariada mostra-se mais adequada para classificar os clones quanto ao alumínio; entretanto, os percentuais de redução na matéria seca da parte aérea e no número de folhas são as características que melhor possibilitaram a discriminação dos clones.

Ferreira (2001), avaliando vários caracteres mensurados em palma forrageira (*Opuntia e Nopalea*), utiliza técnicas multivariadas para estudar as divergências

proporcionadas pelos escores das duas primeiras variáveis canônicas e dos três primeiros componentes principais, os quais permitem evidenciar discriminação entre os clones e estabelecer a hierarquia dos caracteres quanto à importância na quantificação da divergência fenética. A revelação de proximidade entre as características avaliadas na planta foram discutidas considerando-se as distâncias de Mahalanobis e a Euclidiana Média.

Carneiro *et al.* (2002) avaliam o comportamento e a divergência genética entre quatro linhas de frango de corte, utilizando-se de técnicas de análise multivariada nos períodos inicial, médio e total de produção. As técnicas de análise multivariada combinam, simultaneamente, informações múltiplas provenientes de uma unidade experimental. Entre os procedimentos utilizados, destacam-se: teste de Roy, análise multivariada, distância de Mahalanobis e variáveis canônicas. Os resultados mostram divergência genética entre as linhas de frango e a não-influência dos números de ovos com o peso corporal e o primeiro ovo.

Ferrão *et al.* (2002) mostram divergências genéticas entre genitores de feijão que se adaptam ou não às condições de inverno e suas combinações híbridas, com base em procedimentos multivariados. Os dados apresentam-se discutidos por técnicas multivariadas: análise de agrupamento, análise de variância, análise de variáveis canônicas e a medida de Mahalanobis. Esses procedimentos multivariados têm-se mostrado promissores, pois permitem combinar um complexo de variáveis na mesma unidade experimental. A divergência genética recebe a influência da temperatura e do estado de melhoramento.

Oshiiwa & Padovani (2001) mostram uma revisão dos principais testes estatísticos multidimensionais e discutem o uso de procedimentos univariados e multivariados em ensaios agronômicos. Destacam que, na utilização da análise multivariada, o nível α de significância está mantido em todas as conclusões (a estrutura da análise está construída sob essa garantia), enquanto, que no conjunto de várias análises univariadas, tal situação não está assegurada. Ou seja, o nível de significância na aglutinação das univariadas é subestimado e, portanto, indicando falsas rejeições da hipótese de nulidade.

Barroso & Artes (2003) apresentam, em forma de texto didático aplicado às diversas áreas do conhecimento, uma revisão de vários procedimentos multivariados. Destacam-se, entre os tópicos apresentados, a análise de agrupamento, a análise de componentes principais, a análise fatorial e a análise discriminante; esta última, com

especial atenção ao método de Fisher para a discriminação de duas ou mais populações. Consideram um tópico especial sobre detecção de dados aberrantes multivariados envolvendo dois indicadores para identificação de pontos distantes da nuvem dos dados: a distância euclidiana e a distância de Mahalanobis.

Souza *et al.* (2003) empregam a análise multivariada, em particular as técnicas de agrupamento discriminante, para estratificação vertical de florestas ineqüianas considerando dados coletados em 10 parcelas permanentes de 20m x 50m cada, de um experimento instalado na mata da Silvicultura, no município de Viçosa (20°45'S e 42°55'S), estado de Minas Gerais. Consideraram-se na análise da estrutura vertical os dados das alturas totais das árvores amostrais com diâmetro de tronco (dap) igual ou maior que 5,0cm discutidos pela distância euclidiana na técnica de discriminação e, pelo método de ligação complementar, na técnica de agrupamento. Obteve-se como resultado, que as técnicas multivariadas são viáveis para estratificação vertical de floresta ineqüiana quando se utiliza a distribuição de alturas com classes com amplitudes comuns de 1m e, também, que a diversidade de espécies e das estruturas fitossociológicas e paramétrica por meio da estratificação vertical são úteis nas análises estruturadas de florestas ineqüiana.

Martel *et al.* (2003) consideram três técnicas estatísticas multivariada; análise de componentes principais, análise discriminante e análise de agrupamentos, com o objetivo de caracterizar, morfometricamente, raças e populações de pupunha. Foram consideradas pupunheiras (*Bactris gasipaes* Kunth) ao longo dos rios Amazonas e Solimões que apresentam grande variabilidade genética, ainda não totalmente caracterizadas, avaliadas sob os seguintes descritores morfológicos: número de espigas por cacho, comprimento da ráquis, distância morfológica dos frutos, peso dos frutos, adensamento dos frutos, cacho, espessura das cascas, facilidade para descascar os frutos, peso das cascas, textura da polpa, sabor dos frutos, espessura da polpa, peso das sementes e teores de água, óleo e fibras. A análise de agrupamento possibilitou a formação de três grupos de afinidades morfométricas indicados pelas raças Solimões, Putumayo e Pará. Os dois principais eixos discriminantes permitiram a representação bidimensional das 16 populações de pupunha das três raças possibilitando a visualização gráfica da caracterização morfométrica. As três técnicas multivariadas, em conjunto, definem uma diferenciação das raças, mostrando que para a seleção morfométrica das pupunhas destacam-se os seguintes descritores: número de espigas,

comprimento da raquis, peso do fruto, espessura e peso das cascas, facilidade para descascar os frutos, sabor dos frutos, espessura da polpa, distância morfológica dos frutos e peso da semente.

5. DESENVOLVIMENTO METODOLÓGICO

5.1. Método de Discriminação de Fisher

Em estudos biológicos é comum considerar um número elevado de variáveis aleatórias (respostas) correlacionadas entre si. Essa estrutura do vetor de observação não deixa de refletir a estrutura biológica que é um sistema altamente integrado no qual os caracteres e/ou componentes estão internamente relacionados por meio de suas dependências ou interdependências.

Segundo Kendal (1950), classificam-se as técnicas da Análise Multidimensional ou Multivariada em:

- I. Análise de Interdependência: estuda as relações de um conjunto de variáveis entre si.
 - 1) Análise de Agrupamento.
 - 2) Análise de Componentes Principais.
 - 3) Análise de Fatores.
- II. Análise de Dependência: estuda a dependência de uma ou mais variáveis em relação às outras.
 - 1) Análise Discriminante.
 - 2) Análise de Variância.
 - 3) Análise de Medidas Repetidas.
 - 4) Análise de Regressão.
 - 5) Análise de Correlação Canônica.

Neste contexto, Anderson (1984) discutiu vários tópicos teóricos da análise multivariada. Destaca, em especial, a distribuição multinormal de probabilidades que será a base fundamental de todos os testes estatísticos que serão apresentados nos diversos métodos multidimensionais. Continuando a trajetória didática da literatura especializada, seguem-se outros textos com direcionamentos mais práticos, destacando-se Seal (1964), Tatsuoka (1971), Harris (1975), Morrison (1976), Gnanadesikan (1977), entre os mais citados. Johnson & Wichern (1998) apresentam, em 1982, na primeira edição, um enfoque muito interessante que, além de possuir uma teoria consistente e aprofundada e excelentes exemplos práticos como exercícios, contém indicativos computacionais altamente resolutivos para as técnicas multivariadas.

A apresentação de dados observacionais, obtidos a partir de experimentos agronômicos que envolvem mensurações de várias características nas parcelas, em diferentes grupos de estudo, pode ser descrita por meio da estrutura básica de uma tabela de dupla entrada configurando-se, nas linhas, as unidades experimentais e nas colunas as informações observadas nas respostas. Os valores genéricos do vetor aleatório, Y , das respostas observadas em g grupos podem ser representados como na Tabela 1.

Tabela 1. Valores genéricos do vetor resposta Y com p características observadas em g grupos experimentais.

Grupo	Unidade Experimental	Variável Resposta (Característica)			
		V_1	V_2		V_p
1	1	Y_{111}	Y_{112}		Y_{11p}
Λ	Λ	Λ	Λ	Λ	Λ
1	n_1	Y_{1n_11}	Y_{1n_12}		Y_{1n_1p}
Λ	Λ	Λ	Λ	Λ	Λ
g	1	Y_{g11}	Y_{g12}		Y_{g1p}
Λ	Λ	Λ	Λ	Λ	Λ
g	n_g	Y_{gn_g1}	Y_{gn_g2}		Y_{gn_gp}

onde,

$$Y_{hi} = (Y_{hi1}; Y_{hi2}; \Lambda; Y_{hip})$$

representa o vetor resposta correspondente às observações obtidas nas p características para a i -ésima parcela multidimensional do h -ésimo grupo experimental, com $h = 1, \Lambda, g$ (índice de grupo); $i = 1, \Lambda, n_h$ (número de parcelas ou unidades experimentais do grupo h); $j = 1, \Lambda, p$ (número de características ou variáveis respostas).

A caracterização de um procedimento de análise estatística com dados multivariados pode ser entendida como a possibilidade de análise simultânea das variáveis observadas nas parcelas ou unidades experimentais. Embora a idéia esteja apresentada de forma simplista, não é difícil imaginar a complexidade da estrutura de variabilidade envolvida nos dados, pois há que se considerar a variação dentro de cada característica observada e a variação entre as características. A dispersão dos dados observacionais deve ser apresentada por medidas de variabilidade envolvendo variâncias e covariâncias. Esta multiplicidade de características pode ser também estendida à forma de abordagem dos dados observacionais.

Como descrito anteriormente, para cada um dos g grupos experimentais, a estrutura de variabilidade do vetor resposta está apresentada nas dispersões:

a) dentro de variável

$$\text{Var}(V_j) = \text{Var}(Y_{hj}) = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hij} - \bar{y}_{hj})^2,$$

para $j = 1, \Lambda, p$ e $h = 1, \Lambda, g$ (fixo).

b) entre variáveis

$$\text{Cov}(V_j; V_{j'}) = \text{Cov}(Y_{hj}; Y_{hj'}) = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hij} - \bar{y}_{hj})(y_{hij'} - \bar{y}_{hj'}),$$

para $j, j' = 1, \Lambda, p$ ($j \neq j'$) e $h = 1, \Lambda, g$ (fixo).

Comumente, na área agronômica, ocorrem planejamentos em que as mensurações envolvem várias características obtidas na mesma unidade amostral (unidade experimental ou parcela) dificultando o procedimento de cálculos estatísticos e a interpretação dos resultados analíticos. Assim, torna-se recomendável uma análise exploratória dos dados visando resumí-los, porém, considerando ao máximo as informações inerentes às variáveis respostas originais. Tukey (1977) apresenta uma abordagem univariada muito interessante e didática que deve ser considerada na análise exploratória de dados.

No contexto multivariado, para tratar da redução da dimensão do vetor resposta, sem perda substancial das informações do conjunto, existem algumas técnicas de análise multivariada: componentes principais, análise de fatores, discriminação gráfica de Fisher, entre outras. Esses métodos visam gerar um número menor de variáveis auxiliares (hipotéticas) - combinações lineares das variáveis originais - tais que, mesmo em menor número, expliquem adequadamente a interdependência do conjunto original multidimensional dos dados. Além disso, a análise é realizada com o intuito de resumir o padrão de correlação entre as variáveis e muitas vezes é possível chegar a conjuntos de variáveis não correlacionadas mutuamente, levando-os a formar grupos de variáveis. Do ponto de vista prático, o método possibilita, também, a combinação interpretável de variáveis.

Graybill (1969) e Searle (1982) mostraram que a Álgebra de Matrizes pode ser extensivamente utilizada no desenvolvimento de teorias e aplicações estatísticas. Em seus livros, são apresentados vários conceitos e teoremas sobre matrizes sempre com a demonstração matemática seguida de aplicação na teoria estatística. Um conjunto de dados

multivariados (p-dimensional), considerando g grupos experimentais, pode ser apresentado num arranjo matricial na seguinte forma:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{111} & \mathbf{Y}_{112} & \Lambda & \mathbf{Y}_{11p} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \mathbf{Y}_{1n_1 1} & \mathbf{Y}_{1n_1 2} & \Lambda & \mathbf{Y}_{1n_1 p} \\ \mathbf{Y}_{211} & \mathbf{Y}_{212} & \Lambda & \mathbf{Y}_{21p} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \mathbf{Y}_{2n_2 1} & \mathbf{Y}_{2n_2 2} & \Lambda & \mathbf{Y}_{2n_2 p} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \mathbf{Y}_{g11} & \mathbf{Y}_{g12} & \Lambda & \mathbf{Y}_{g1p} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \mathbf{Y}_{gn_g 1} & \mathbf{Y}_{gn_g 2} & \Lambda & \mathbf{Y}_{gn_g p} \end{bmatrix}$$

ou genericamente,

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{hij} \end{bmatrix}$$

onde $\mathbf{Y}$ é de ordem $\mathbf{n} \times \mathbf{p}$, sendo $\mathbf{n} = \mathbf{n}_1 + \Lambda + \mathbf{n}_g$; $\mathbf{h} = 1, \Lambda, \mathbf{g}$ (índice de grupo); $\mathbf{i} = 1, \Lambda, \mathbf{n}_h$ (repetições dentro do grupo) e $\mathbf{j} = 1, \Lambda, \mathbf{p}$ (número de variáveis observadas).

O elemento genérico $\mathbf{Y}_{hij}$ refere-se à j-ésima resposta da i-ésima parcela do h-ésimo grupo.

Uma alternativa interessante para a matriz $\mathbf{Y}$ é sua representação em submatrizes $\mathbf{Y}_h$ associadas ao grupo populacional em consideração, descritas por

$$\mathbf{Y}_h = \begin{bmatrix} \mathbf{Y}_{h11} & \mathbf{Y}_{h12} & \Lambda & \mathbf{Y}_{h1p} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \mathbf{Y}_{hn_h 1} & \mathbf{Y}_{hn_h 2} & \Lambda & \mathbf{Y}_{hn_h p} \end{bmatrix}$$

onde $\mathbf{Y}_h$ é uma matriz de ordem $\mathbf{n}_h \times \mathbf{p}$, para $\mathbf{h} = 1, \Lambda, \mathbf{g}$.

As submatrizes possibilitam representar $\mathbf{Y}$ por:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ M \\ \mathbf{Y}_h \\ M \\ \mathbf{Y}_g \end{bmatrix}$$

onde, nas linhas de $\mathbf{Y}_h$ estão alocados os vetores respostas das unidades do h-ésimo grupo, ou seja, cada vetor-linha da submatriz representa a resposta multidimensional de uma unidade experimental e, cada vetor coluna, as respostas observadas em uma dada variável aleatória.

A proposta pioneira da análise discriminante de Fisher prende-se ao estabelecimento de um critério para a separação de populações (construção de regiões de discriminação); entretanto, este critério também permite ser utilizado na classificação ou indicação de novos indivíduos em uma das populações em estudo.

Em síntese, os principais objetivos da análise discriminante são:

- I. Discriminação: consiste em estabelecer funções das variáveis observadas (funções discriminantes) que sejam responsáveis ou que possam explicar as diferenças entre as g populações.
- II. Alocação ou Classificação: consiste em estabelecer funções das variáveis observadas que permitam alocar ou classificar novos indivíduos em uma das g populações.

A análise discriminante pode ser estabelecida por meio de funções lineares do tipo $\mathbf{t}^T \mathbf{y}$ com algumas propriedades especiais. No contexto da discriminação, não se considera essencial a descrição paramétrica das populações, mas sim, uma regra consistente e sensível para construção de regiões bem definidas das características populacionais. Nesta linha de objetivo, considerando cada conjunto de dados como um grupo, a sugestão de Fisher foi no sentido de buscar combinações lineares $\mathbf{t}^T \mathbf{y}$ de modo que a razão entre grupos e a

soma de quadrados da variação dentro de grupos seja máxima. Diferentemente da análise de componentes principais, a análise discriminante não é afetada pela padronização das variáveis.

Na prática, as funções discriminantes são determinadas com base nas amostras. Uma amostra de cada uma das g populações é coletada e, então, são observadas as p variáveis aleatórias respostas dos indivíduos amostrados com a respectiva população da qual a amostra provém (Método supervisionado). A análise discriminante, juntamente com a análise de agrupamento (Método não supervisionado), têm tido uma ampla utilização como ferramenta de “Data Mining”. Para a construção dos discriminadores não é necessário assumir que as g populações são normais multivariadas. Porém, imprescinde-se da homogeneidade das matrizes de covariância das g populações e que todas sejam de posto completo (a variação dentro dos grupos envolve a matriz conjunta – “pool” dos grupos).

Considerou-se, para o teste estatístico da igualdade das matrizes de covariância dos grupos experimentais ou populações (homogeneidade na estrutura de variação dos grupos), o teste de Bartlett (Morrison, 1976).

As hipóteses do teste são descritas como:

$$H_0: \Sigma_1 = \Sigma_2 = \Lambda = \Sigma_g \quad (\text{Homogeneidade da estrutura de variação})$$

H_1 : Existe $\Sigma_h \neq \Sigma_{h'}$ para $h, h' \in \{1, 2, \Lambda, g\}$ (Heterogeneidade da estrutura de variação).

Fixado o nível α de significância, sob a veracidade da hipótese nula, a estatística MC^{-1} tem distribuição assintótica qui-quadrado com

$$\frac{1}{2}(g-1)p(p+1)$$

graus de liberdade.

Os valores de M e C^{-1} para o cálculo numérico de MC^{-1} são dados por:

$$M = (n-g)\lambda n|S| - \sum_{h=1}^g (n_h-1)\lambda n|S_h| ,$$

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \left[\sum_{h=1}^g \frac{1}{n_h - 1} - \frac{1}{n - g} \right];$$

onde $n = n_1 + \Lambda + n_h$; S_h é a matriz de covariância do grupo $h = 1, \Lambda, g$ e S a matriz comum de covariância dos grupos definida por:

$$S = \sum_{h=1}^g (n_h - 1) \frac{S_h}{(n - g)}.$$

Se todos os grupos apresentam o mesmo número de unidades experimentais, isto é, $n_1 = \Lambda = n_g = r$, o valor de C^{-1} fica expresso por:

$$C^{-1} = 1 - \frac{(2p^2 + 3p - 1)(g + 1)}{6(p + 1)gr}.$$

Ademais, a aproximação qui-quadrado mostra-se melhor quando g e p não excedem a quatro ou cinco e cada n_h (número de parcelas do grupo) representa 20 ou mais repetições (Box, 1949).

Sob estas considerações, a regra de decisão do teste é a usual, ou seja: se $x_{calc}^2 > x_{tab}^2$, rejeita-se a hipótese de homogeneidade; caso contrário, não há rejeição.

Como descrito anteriormente, o critério de Fisher consiste em encontrar o vetor que maximize a razão

$$\frac{\tilde{\mathbf{t}}' \mathbf{B} \tilde{\mathbf{t}}}{\tilde{\mathbf{t}}' \mathbf{W} \tilde{\mathbf{t}}},$$

onde as matrizes $\mathbf{B}$ e $\mathbf{W}$ referem-se a variação entre e dentro dos grupos, respectivamente. O vetor $\tilde{\mathbf{t}}$ que maximiza a razão será chamado de função discriminante linear canônica de Fisher ou eixo canônico.

Para obtenção do vetor $\tilde{\mathbf{t}}$ pode ser considerada a utilização dos lagrangeanos, ou seja, a maximização da razão pode ser feita pela maximização do numerador

considerando-se que o denominador é constante. Neste sentido, buscar a solução $\tilde{\mathbf{t}}$ significa maximizar $\tilde{\mathbf{t}}' \mathbf{B} \tilde{\mathbf{t}}$ sujeito à restrição

$$\tilde{\mathbf{t}}' \mathbf{W} \tilde{\mathbf{t}} = K ,$$

onde K é um valor constante qualquer.

Isto é, na forma de Lagrange, tem-se

$$L(\mathbf{t}) = \mathbf{t}' \mathbf{B} \mathbf{t} - \lambda (\mathbf{t}' \mathbf{W} \mathbf{t} - K) \text{ e}$$

$$\frac{\partial L(\mathbf{t})}{\partial \mathbf{t}} = 2\mathbf{B} \mathbf{t} - 2\lambda \mathbf{W} \mathbf{t}.$$

Igualando-se a derivada à nulidade, tem-se o sistema linear homogêneo:

$$(\mathbf{B} - \lambda \mathbf{W}) \mathbf{t} = \mathbf{0} ,$$

ou, equivalente,

$$(\mathbf{W}^{-1} \mathbf{B} - \lambda \mathbf{I}) \mathbf{t} = \mathbf{0} ,$$

e, portanto, $\mathbf{t}$ é o autovetor associado a um autovalor de $\mathbf{W}^{-1} \mathbf{B}$ ($\mathbf{t}$ autovetor de $\mathbf{B}$ com respeito à $\mathbf{W}$).

Tomando-se a expressão da derivada igualada ao vetor nulo e pré-multiplicada por $\tilde{\mathbf{t}}'$, tem-se:

$$2\mathbf{B} \mathbf{t} - 2\lambda \mathbf{W} \mathbf{t} = \mathbf{0}$$

$$\mathbf{t}' \mathbf{B} \mathbf{t} - \lambda \mathbf{t}' \mathbf{W} \mathbf{t} = 0 ,$$

mas

$$\mathbf{t}' \mathbf{W} \mathbf{t} = K \text{ (Lagrange),}$$

então

$$\mathbf{t}' \mathbf{B} \mathbf{t} = \lambda K .$$

Logo, o denominador, $\tilde{\mathbf{t}}_1' \mathbf{B} \tilde{\mathbf{t}}_2$, da razão das variações, torna-se máximo quando λ for o maior autovalor de $\mathbf{W}^{-1}\mathbf{B}$. Seguindo procedimento semelhante e acrescentando a restrição que a covariância entre as funções discriminantes seja nula, isto é,

$$\tilde{\mathbf{t}}_1' \mathbf{W} \tilde{\mathbf{t}}_2 = 0,$$

determina-se que $\tilde{\mathbf{t}}_2$ é o autovetor associado ao segundo maior autovalor de

$$\mathbf{W}^{-1}\mathbf{B}.$$

Repetindo-se o procedimento, todas as funções discriminantes associadas aos demais autovetores e autovalores serão estabelecidas.

Para a construção do novo sistema de referência relativo aos eixos discriminantes, os seguintes vetores e matrizes, construídos a partir dos dados observados, são necessários:

I. Vetor das médias amostrais de um dado grupo

$$\bar{\mathbf{Y}}_h' = (\bar{\mathbf{Y}}_{h,1} \quad \bar{\mathbf{Y}}_{h,2} \quad \Lambda \quad \bar{\mathbf{Y}}_{h,p}),$$

onde $\mathbf{h} = 1, \Lambda, \mathbf{g}$ e

$$\bar{\mathbf{Y}}_{h,j} = \sum_{i=1}^{n_h} \frac{\mathbf{y}_{hij}}{n_h},$$

com $\mathbf{j} = 1, \Lambda, \mathbf{p}$.

II. Matriz de covariância amostral de um dado grupo

$$\mathbf{S}_h = \begin{bmatrix} \mathbf{s}_{11}^{(h)} & \mathbf{s}_{12}^{(h)} & \Lambda & \mathbf{s}_{1p}^{(h)} \\ \mathbf{s}_{21}^{(h)} & \mathbf{s}_{22}^{(h)} & \Lambda & \mathbf{s}_{2p}^{(h)} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \mathbf{s}_{p1}^{(h)} & \mathbf{s}_{p2}^{(h)} & \Lambda & \mathbf{s}_{pp}^{(h)} \end{bmatrix} = [\mathbf{s}_{jj}^{(h)}]$$

onde

$$s_{jj}^{(h)} = \begin{cases} \text{var}^{(h)}(Y_j) = \sum_{i=1}^{n_h} \frac{(y_{hij} - \bar{y}_{hj})^2}{n_h} & \text{para } j = j', \\ \text{cov}^{(h)}(Y_j, Y_{j'}) = \sum_{i=1}^{n_h} \frac{(y_{hij} - \bar{y}_{hj})(y_{hj' i} - \bar{y}_{hj'})}{n_h} & \text{para } j \neq j', \end{cases}$$

$k = 1, \Lambda, g$ e $j' = 1, \Lambda, p$.

III. Vetor das médias gerais das variáveis:

$$\bar{\mathbf{y}}' = (\bar{\mathbf{y}}_{..1} \quad \bar{\mathbf{y}}_{..2} \quad \Lambda \quad \bar{\mathbf{y}}_{..p}),$$

onde

$$\bar{\mathbf{y}}_{..j} = \frac{\sum_{h=1}^g \sum_{i=1}^{n_h} y_{hij}}{\mathbf{n}} = \frac{\sum_{h=1}^g \sum_{i=1}^{n_h} y_{hij}}{\left(\sum_{h=1}^g n_h \right)},$$

ou ainda,

$$\bar{\mathbf{y}}_{..j} = \frac{\sum_{h=1}^g n_h \bar{\mathbf{y}}_{h,j}}{\mathbf{n}}, \text{ para } j = 1, \Lambda, p.$$

IV. Matriz comum de covariância de grupos amostrais:

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_{11} & \Lambda & \mathbf{s}_{1p} \\ \mathbf{M} & & \mathbf{M} \\ \mathbf{s}_{p1} & \Lambda & \mathbf{s}_{pp} \end{bmatrix},$$

onde

$$s_{jj} = \frac{\sum_{h=1}^g (n_h - 1) s_{jj}^{(h)}}{(n - g)}; j, j' = 1, \Lambda, p.$$

Matricialmente, tem-se

$$S = \frac{\sum_{h=1}^g (n_h - 1) s_h}{(n - g)},$$

ou equivalente,

$$S = \frac{\mathbf{W}}{(n - g)} \text{ com } \mathbf{W} = \sum_{h=1}^g (n_h - 1) \mathbf{s}_h.$$

O uso da letra **W** é bastante sugestivo, pois **S** avalia a estrutura de variação dentro dos grupos amostrais (“Sample **W**ithin Groups Matrix”).

V. Matriz entre grupos amostrais:

Semelhantemente, a matriz **B** avalia a estrutura de variação entre grupos (“Sample **B**etween Groups Matrix”), cuja definição é dada por

$$\mathbf{B} = \sum_{h=1}^g (\bar{\mathbf{y}}_h - \bar{\mathbf{y}})(\bar{\mathbf{y}}_h - \bar{\mathbf{y}}),$$

e onde

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_{11} & \Lambda & \mathbf{b}_{1p} \\ \mathbf{M} & & \mathbf{M} \\ \mathbf{b}_{p1} & & \mathbf{b}_{p2} \end{bmatrix},$$

$$\text{com } \mathbf{b}_{jj'} = \sum_{h=1}^g (\bar{\mathbf{y}}_{h,j} - \bar{\mathbf{y}}_{..j})(\bar{\mathbf{y}}_{h,j'} - \bar{\mathbf{y}}_{..j'}) \text{ para } j, j' = 1, \Lambda, p.$$

O procedimento de determinação dos eixos discriminantes para o estabelecimento das regiões de classificação, consiste em encontrar combinações lineares das variáveis originais que representem consideravelmente as informações contidas na estrutura de variabilidade dos dados.

Para a determinação das funções discriminantes lineares amostrais, sejam

$$\lambda_1 \geq \lambda_2 \geq \Lambda \geq \lambda_s > 0,$$

os $s \leq \min(\mathbf{g} - \mathbf{1}, \mathbf{p})$ autovalores não-nulos de $W^{-1} B$ e

$$\mathbf{t}_{\sim 1}, \mathbf{t}_{\sim 2}, \Lambda, \mathbf{t}_{\sim s},$$

os correspondentes autovetores padronizados tais que

$$\mathbf{t}_{\sim k}' W \mathbf{t}_{\sim k} = 1$$

e

$$\mathbf{t}_{\sim k}' W \mathbf{t}_{\sim k'} = 0,$$

para $k \neq k'$ e $k, k' \in \{1, 2, \Lambda, s\}$.

O vetor $\hat{\mathbf{t}}_{\sim}$ que maximiza a razão

$$\frac{\mathbf{t}_{\sim}' B \mathbf{t}_{\sim}}{\mathbf{t}_{\sim}' W \mathbf{t}_{\sim}}$$

dado pelo autovetor correspondente ao maior autovalor λ_1 , definirá a combinação linear $\mathbf{t}_{\sim 1}' \mathbf{y}_{\sim}$ chamada de primeira discriminante amostral ou primeiro eixo de discriminação ou primeiro eixo canônico.

A escolha de $\mathbf{t}_{\sim 2}$ correspondente ao segundo maior autovalor, λ_2 , que produz a segunda discriminante linear amostral,

$$\mathbf{t}_{\sim 2}' \mathbf{y}_{\sim}.$$

Continuando, $\mathbf{t}_{\sim k}'$ correspondente a λ_k , determina a k-ésima discriminante amostral com $k \leq s$. (Johnson & Wichern, 1998).

Estabelecidos os eixos discriminantes, os vetores-respostas das variáveis originais são transformados em respostas do novo sistema e seus valores "plotados" no plano cartesiano discriminador. Procedimento semelhante pode ser dispensado aos centróides dos grupos, representado pelo vetor de respostas médias. A partir do diagrama de dispersão dos pontos associado à geometria euclidiana (Rao, 1983) são estabelecidas as regiões de discriminação.

Uma das alternativas para apresentar as regiões de discriminação, estabelecidas a partir dos eixos discriminantes, é o procedimento gráfico. Esta alternativa é bastante interessante pois permite a visualização das regiões de classificação construídas no plano bidimensional, tendo como referência do sistema cartesiano (x,y) os eixos discriminantes: como eixo x, toma-se o primeiro autovetor canônico (aquele responsável pela maior retenção de informação, ou seja, o correspondente ao maior autovalor de $W^{-1}B$) e, como eixo y, o segundo (o responsável pela segunda maior retenção de informação, ou seja, o correspondente ao segundo maior autovalor de $W^{-1}B$).

O procedimento escolhido na presente pesquisa para a delimitação das regiões de classificação envolve os vetores de médias dos grupos e o vetor de médias de todos os dados (média geral). Deve ser também destacado que a construção proposta para as regiões facilita o cálculo freqüentista das probabilidades de má-classificação.

Dentro deste encaminhamento prático dado à discriminação de Fisher envolvendo os dois eixos discriminantes, tem-se os seguintes passos para a construção gráfica das regiões e a obtenção das probabilidades de má-classificação.

Em relação a construção gráfica, tem-se:

1. a partir dos vetores $\bar{\mathbf{Y}}_{\sim h}$ ($h = 1, \dots, g$) das médias amostrais, calculam-se os valores numéricos $\mathbf{t}'_1 \bar{\mathbf{Y}}_{\sim h}$ e $\mathbf{t}'_2 \bar{\mathbf{Y}}_{\sim h}$ correspondentes aos dois eixos discriminantes, cujo par $(\mathbf{t}'_1 \bar{\mathbf{Y}}_{\sim h}, \mathbf{t}'_2 \bar{\mathbf{Y}}_{\sim h})$ definirá o centróide do h-ésimo grupo no plano bidimensional;
2. procedimento similar faz-se em relação ao vetor $\bar{\mathbf{Y}}$ da médias gerais, cujo par $(\mathbf{t}'_1 \bar{\mathbf{Y}}, \mathbf{t}'_2 \bar{\mathbf{Y}})$ definirá a origem (centro geral) da região geral de classificação;
3. esta região geral será particionada em g sub-regiões que representarão cada um dos grupos ou populações estudadas;
4. a partir da poligonal fechada definida pelos centróides dos grupos (vértices do polígono), constroem-se as sub-regiões delimitando-as pelas semi-retas que partem da origem (centróide relativo à média geral) passando pelo ponto médio de cada um de seus lados;

5. caracteriza-se cada região pelo seu respectivo centróide. Esta arquitetura de construção possibilitará a classificação de novos elementos em suas respectivas regiões.

Em relação à qualidade da classificação de novos elementos, esta poderá ser avaliada de forma freqüentista considerando-se os seguintes procedimentos:

1. separar inicialmente, por processo casual uma parte das unidades experimentais (neste trabalho optou-se por 20%) de cada grupo;
2. verificar em que região se localiza cada uma das imagens dessas unidades experimentais. Computa-se assim, o percentual de classificação correta, ou seja, aquelas para as quais a classificação pelos eixos discriminantes coincide com a original dos dados.

5.2. Programa Computacional

O programa computacional intitulado por FISHER, do procedimento envolvido na metodologia estatística referente à determinação dos eixos discriminantes de Fisher para a classificação de indivíduos e à discussão da similaridade entre unidades experimentais multivariadas foi desenvolvido em linguagem MAPLE (versão V, release 3) para ambiente WINDOWS 98.

Como a linguagem MAPLE não oferece facilidade para manutenção de arquivos e para emissão de relatórios, foi criado um programa em CLIPPER 5.1 denominado “PROJETO” que permite a execução das seguintes tarefas:

- c) manipulação do banco de dados, possibilitando a introdução das informações via digitação direta ou importação de planilha;
- d) armazenamento dos dados de vários experimentos;
- e) consulta, alteração e inclusão de valores.;
- f) seleção de parte do arquivo para processamento;
- g) consistência dos dados e geração de arquivos de comunicação entre os ambientes CLIPPER e MAPLE;
- h) elaboração de relatório contendo os resultados obtidos pelo programa MAPLE no processamento das informações.

O programa CLIPPER gera um arquivo de comunicação para o MAPLE, e este devolve para o CLIPPER um relatório para visualização e impressão.

O fluxograma relativo ao programa FISHER está apresentado na figura1.

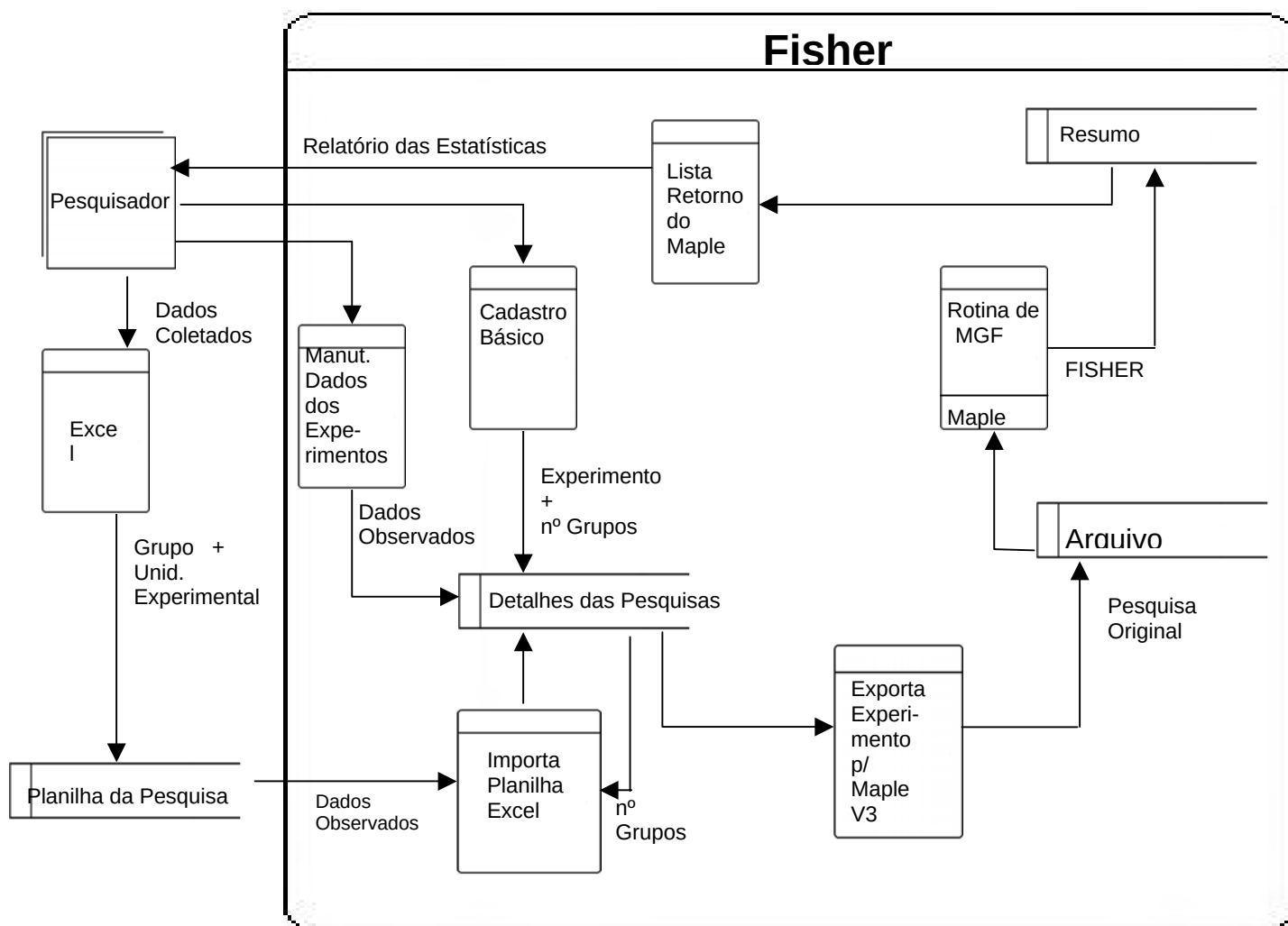


Figura1. Fluxograma do programa FISHER.

5.3. Manual do Usuário

Em relação à operacionalização do sistema computacional desenvolvido no presente estudo, devem ser consideradas pelo usuário as seguintes informações:

<ESC> - Retorna ao menu.

[F1] - 'Help' do sistema.

[F4] - Tabelas do sistema. Fornece a relação dos elementos cadastrados no arquivo associado ao campo.

[F5] - Calculadora. Teclado numérico reduzido.

[Cima] - Retorna ao campo anterior da tela (telas de diálogo). No menu do sistema, permite navegar pelas opções do Módulo em que esta posicionado.

[Baixo] - Avança ao campo seguinte da tela (telas de diálogo). No menu do sistema, permite navegar pelas opções do Módulo em que está posicionado.

[<-] - Move um caracter a esquerda dentro do campo. No menu do sistema, permite navegar pelos Módulos disponíveis. Permite também 'navegar' pelos registros dos arquivos.

[->] - Move um caracter a direita dentro do campo. No menu do sistema, permite navegar pelos Módulos disponíveis. Permite também 'navegar' pelos registros dos arquivos.

Nas rotinas de manutenção de cadastros e tabelas, sempre que informado o código de um elemento já cadastrado, o sistema informará o número de variáveis (p) e grupos (g) questionando se o usuário deseja alterar ou excluir o elemento informado, permitindo 'navegar' pelo cadastro/tabela com o uso das teclas [->] e [<-].

Quando o usuário informar um valor fora dos limites previstos para o campo, o sistema apresentará uma mensagem elucidativa relativa ao mesmo e aguardará que seja pressionada alguma tecla.

Além disso, as rotinas de emissão dos relatórios possibilitam que os mesmos sejam apresentados em vídeo ou impressora. No sistema estão inseridas as seguintes rotinas:

1 - Cadastro Básico dos Experimentos.

O usuário informa o código e nome do experimento, o número de variáveis (p), o número de grupos (g) e o nível de significância (α) desejado no processo inferencial estatístico. Informações fornecidas antes da importação de dados de planilhas Excel ou da digitação das variáveis.

O código do experimento aqui fornecido será utilizado para troca de informações entre a Base de Dados e o MapleV3.

2 - Cadastro das Variáveis do Experimento.

Permite ao usuário digitar - incluir ou alterar - os valores de cada variável diretamente no sistema. Esta rotina somente será utilizada nos casos em que o usuário não tiver os dados disponíveis em planilhas.

Somente os experimentos que já tenham o Cadastro Básico poderão ter seus dados digitados.

O número de variáveis (p) solicitado para cada unidade experimento será o informado no Cadastro Básico.

3 - Importa Dados de Planilha (Formato TXT).

Após o cadastramento Básico do experimento, o usuário poderá importar dados diretamente de uma planilha Excel, observando:

- A planilha deve ter na primeira linha e primeira coluna (célula A1) o literal GRUPOS. As demais células contidas nessa primeira linha serão desprezadas.
- Na primeira coluna da planilha o usuário informará o número do grupo, não havendo necessidade de informá-lo caso seja igual ao da linha anterior.

- Nas colunas, a partir do número 2, estarão os valores relativos a cada variável.
- Caso a planilha tenha mais condições de avaliação do que o informado em Cadastro Básico do Experimento, as condições excedentes serão descartadas.
- A planilha deve obrigatoriamente ser salva em formato TXT, com tabulação, sem a formatação padrão da MICROSOFT.

4 - Listagem para Conferência

Fornece a relação dos dados digitados ou importados do Excel.

5 - Exporta Arquivo para cálculos no MAPLE

Formata um arquivo com os dados do experimento para processamento no Maple.

O arquivo foi estruturado para ser executado em MapleV3, não sendo possível utilizá-lo em outras versões.

6 - Relatório Retorno do Maple

Lista os valores das estatísticas calculadas e os resultados obtidos para cada teste.

Em continuidade ao processo operacional dentro do ambiente MAPLE, tem-se a seguinte seqüência de considerações:

- Após gerar o arquivo de integração com o MAPLE no ambiente Clipper, abra o programa MAPLE e selecione no menu [File] a opção [Open...].
- Na caixa de diálogo que se abre, em [Pastas] selecionar aquela na qual o sistema "Utilização do Método Discriminante de Fisher" foi instalado - o "DEFAULT" é a pasta de nome PROJETO.
- Verificar se o campo [Listar Arquivos do Tipo] está preenchido com <Maple Section - *.MS>. Caso não esteja, selecione essa opção.
- No campo [Nome do arquivo] digitar o código do experimento que está sendo analisado - ou selecionar na barra de rolagem que se apresenta.
- Confirmar a operação para que o arquivo seja aberto, e suas instruções iniciais apareçam na tela.
- No menu [Format], selecionar a opção [Execute Worksheet] que fará com que a rotina de Discriminação seja executada e que o relatório apresentando os resultados obtidos em cada teste seja gerado em arquivo, para posterior impressão no ambiente Clipper.
- Durante a execução do programa [Worksheet] aparecerá na tela alguns processos, até que ao final o comando ">Close();" indicará o término do processo.
- Fechar o programa MAPLE, selecionando no menu [File] a opção [Exit]. O sistema emitirá uma mensagem informando que o arquivo foi executado e questionando se as alterações geradas por essa execução devem ser salvas. Informe <Não> para que o programa se feche.

5.4. Exemplo de Aplicação

Considerou-se, para ilustrar o programa computacional sobre a discriminação de Fisher, dados adaptados de Messetti (2000), relativos a experimentos desenvolvidos por pesquisadores da EMBRAPA, na região de Londrina – PR, no ano agrícola de 1996/97, cujo objetivo destina-se a avaliar o grau de adaptabilidade de variedades de girassol (*Helianthus annuus*) de diferentes origens. No contexto do presente trabalho, foram escolhidas seis variedades: **G1-Corona; G2-Pehuen; G3-Comangir; G4-Talenay; G5-Maribondo; G6-3Grnains**, todas tendo como países de origem a Argentina e Chile (Argentina: G1, G2, G3, G5, G6; Chile: G6) avaliadas sob os seguintes caracteres quantitativos: **altura da planta** (mensurada em metros, do nível do solo até a inserção do capítulo), **diâmetro do caule** (em mm, medição feita a 5 cm do nível do solo); **altura do capítulo** (em cm, medida realizada do nível do solo até a inserção do capítulo); **diâmetro do capítulo** (em cm, realizado no ponto de maturação fisiológica) e **teor percentual de óleo** (análise feita em laboratório).

Os vetores de respostas observados nas unidades experimentais encontram-se no APÊNDICE 1. Para cada uma das variedades foram consideradas 18 parcelas, sendo três destas (aproximadamente 20%), escolhidas por processo randômico (as destacadas por sombreado no final do conjunto de dados) para o estabelecimento freqüentista da probabilidade de má-classificação.

As medidas descritivas envolvendo tendência central (média) e variabilidade (variância e covariância) dos caracteres agronômicos nas diferentes variedades de girassol (grupo) estão apresentadas na seqüência.

Tabela1. Vetor de médias segundo variedade e média geral

Variedade	Característica Quantitativa				
	Alt. Planta (v1)	Diâm. Caule (v2)	Alt. Capítulo (v3)	Diâm. Capítulo (v4)	Teor de Óleo (v5)
Corona (G1)	1,3440	20,1707	109,5973	16,5720	33,5580
Pehuen (G2)	1,9833	25,9247	125,2413	19,4193	32,3107
Comangir (G3)	2,3120	28,9720	176,3900	23,7887	34,6053
Talenay (G4)	2,0753	25,9193	157,6553	20,2740	36,2900
Maribondo (G5)	1,3560	20,5993	116,5513	15,2480	49,6060
3Grnains (G6)	1,6040	16,8347	111,3633	12,9927	41,6327
Média Geral	1,7791	23,0701	132,7998	18,0491	38,0005

As matrizes simétricas de dispersão (variabilidade) dos grupos são dadas por:

- Grupo1 (Corona)

$$S_1 = \begin{bmatrix} 0,0784 & -0,01057 & -0,2071 & 0,3917 & -0,4897 \\ & 14,2585 & 16,8429 & 1,0507 & 0,6785 \\ & & 236,0032 & -13,4857 & 2,5500 \\ & & & 5,7820 & -2,2585 \\ & & & & 12,8433 \end{bmatrix}$$

- Grupo2 (Pehuen)

$$S_2 = \begin{bmatrix} 0,1312 & 1,0768 & 5,0807 & 0,3974 & -0,1870 \\ & 17,0221 & 39,1786 & 4,5142 & -3,0428 \\ & & 278,0366 & 3,3928 & -25,3072 \\ & & & 7,2647 & 3,0707 \\ & & & & 12,9961 \end{bmatrix}$$

- Grupo3 (Comangir)

$$S_3 = \begin{bmatrix} 0,1157 & 0,3100 & 2,6771 & 0,2023 & -0,2221 \\ & 17,9140 & 28,0143 & 5,4357 & -3,5642 \\ & & 247,2422 & 1,2714 & -18,9072 \\ & & & 8,6115 & 1,2928 \\ & & & & 14,2491 \end{bmatrix}$$

- Grupo4 (Talenay)

$$S_4 = \begin{bmatrix} 0,0743 & 0,3789 & -2,2135 & 0,0500 & 0,3564 \\ & 13,3594 & -39,9786 & 0,7450 & -1,0642 \\ & & 198,2113 & -7,1785 & -11,4786 \\ & & & 8,6581 & -0,8285 \\ & & & & 13,4108 \end{bmatrix}$$

- Grupo5 (Maribondo)

$$S_5 = \begin{bmatrix} 0,0678 & 0,2782 & 0,2078 & -0,1140 & 0,1079 \\ & 14,9831 & 22,5143 & 0,2050 & 5,3785 \\ & & 223,1411 & 21,7286 & 6,7785 \\ & & & 5,1749 & 0,2928 \\ & & & & 14,6212 \end{bmatrix}$$

- Grupo6 (3Grnains)

$$S_6 = \begin{bmatrix} 0,0914 & 0,0898 & -0,2628 & -0,0806 & -0,5104 \\ & 15,0270 & 27,9643 & -2,5357 & 1,3428 \\ & & 200,2696 & -1,0785 & 15,4286 \\ & & & 6,5010 & -0,9692 \\ & & & & 14,9179 \end{bmatrix}$$

O resultado do teste estatístico de Bartlett ($\chi^2 = 88,9554$, $P = 0,1293$) mostrou que as matrizes de dispersão das variedades são homogêneas. A condição de homogeneidade é uma restrição inferencial para a construção dos eixos canônicos de Fisher.

Na sequência, foram determinadas as matrizes **W** e **B**, relativas às variações dentro dos grupos e entre caracteres, visando a construção dos eixos discriminantes. As matrizes **W** e **B** e os autovalores de $\mathbf{W}^{-1}\mathbf{B}$, ordenados em ordem crescente de magnitude, são descritos a seguir.

- Matriz **W** (Variação dentro de grupos)

$$\mathbf{W} = \begin{bmatrix} 7,8284 & 29,7251 & 73,9503 & 11,8570 & -13,2370 \\ & 1295,9100 & 1323,5040 & 131,8102 & -3,8001 \\ & & 19360,7400 & 65,1009 & -433,1031 \\ & & & 587,8975 & 8,3000 \\ & & & & 1162,5432 \end{bmatrix}$$

- Matriz **B** (Variação entre grupos)

$$\mathbf{B} = \begin{bmatrix} 0,8124 & 7,9708 & 49,7723 & 6,7107 & -7,0913 \\ & 104,4906 & 547,5938 & 86,8574 & -79,5956 \\ & & 3836,9535 & 483,3139 & -310,8592 \\ & & & 75,3658 & -75,4003 \\ & & & & 214,4426 \end{bmatrix}$$

- Autovalores de $\mathbf{W}^{-1}\mathbf{B}$

$$\lambda_1 = 0,375167 ;$$

$$\lambda_2 = 0,144037 ;$$

$$\lambda_3 = 0,025568 ;$$

$$\lambda_4 = 0,012435 ;$$

$$\lambda_5 = 0,000374 ;$$

com $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 0,557581 = \text{tr } \mathbf{W}^{-1}\mathbf{B}$ e $s = 5 = \min(\mathbf{g} - \mathbf{1}, \mathbf{p})$,

o número de raízes características (autovalores) não nulas de $\mathbf{W}^{-1}\mathbf{B}$.

Considerando os dois maiores autovalores, λ_1 e λ_2 , são estabelecidos os seus respectivos autovetores, $\mathbf{t}_1$ e $\mathbf{t}_2$, os quais são denominados de primeiro e segundo eixos discriminantes ou canônicos de Fisher, respectivamente. Os coeficientes das funções lineares dos eixos e suas porcentagens de informação a respeito da variabilidade total dos dados explicadas no procedimento de Fisher estão descritos na Tabela2.

Tabela2. Coeficientes dos eixos canônicos e respectivas porcentagem de informação

Variável	Primeiro Eixo Canônico	Segundo Eixo Canônico
Altura da Planta (V1)	1,728630	2,210510
Diâmetro da Planta (V2)	0,115292	-0,109557
Altura do Capítulo (V3)	0,098147	0,139445
Diâmetro do Capítulo (V4)	0,535945	0,064635
Teor de Óleo (V5)	-0,312706	0,902240
Porcentagem de Explicação	67,28	25,83
Porcentagem Acumulada	67,28	93,11

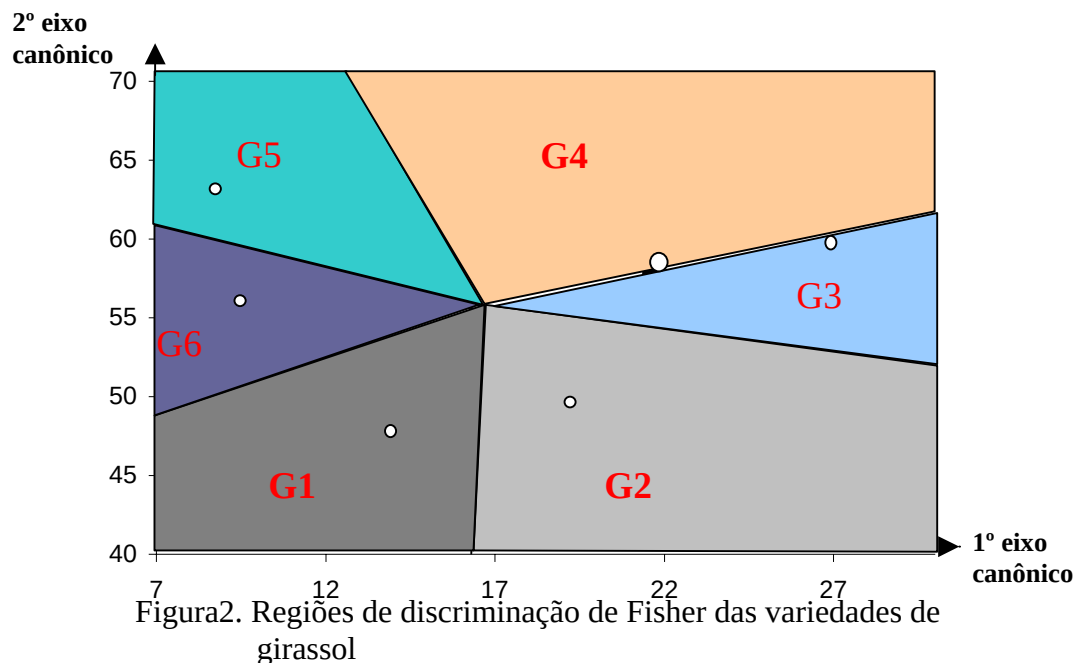
Os resultados mostram que o primeiro eixo canônico responde por 67,28% da retenção, enquanto o segundo eixo, 25,83%. A porcentagem acumulada nos dois eixos é da ordem de 93,11%, o que representa uma alta taxa de retenção da informação a respeito da variabilidade total dos dados, ou seja, o redimensionamento do espaço paramétrico de cinco para dois ocorreu conservando, praticamente, a totalidade da informação agronômica.

Considerando os vetores $\mathbf{t}_1$ e $\mathbf{t}_2$ (os dois eixos), são estabelecidos a partir das médias dos grupos e geral, os centróides das regiões de classificação e o centro geral do sistema, respectivamente. Os valores numéricos dos pares representantes desses pontos no plano cartesiano estão descritos na Tabela3.

Tabela3. Valores numéricos das médias no plano cartesiano dos eixos canônicos

Variedade	Grupo	Eixo Canônico	
		Primeiro	Segundo
CORONA	G1	13,79	47,39
PEHUEN	G2	19,01	49,42
COMANGIR	G3	26,58	59,29
TALENAY	G4	21,57	57,78
MARIBONDO	G5	8,82	62,74
3GRNAINS	G6	9,59	55,63
Geral	Centro	16,56	55,38

Utilizando-se o procedimento da representação gráfica proposto no presente trabalho, têm-se as regiões de discriminação das variedades de girassol (Figura2).



De forma semelhante, em relação às unidades amostrais separadas inicialmente por randomização para obtenção da probabilidade de má-classificação, têm-se os valores numéricos das representações no sistema de eixos canônicos (Tabela4).

Tabela4. Valores numéricos das médias no plano cartesiano dos eixos canônicos e respectivas classificações

Unid. Amostral	Origem	Eixo Canônico		Classificação
		Primeiro	Segundo	
1	CORONA	14,17	48,26	CORONA
2	CORONA	14,62	45,80	CORONA
3	CORONA	12,35	49,32	CORONA
4	PEHUEN	20,87	53,39	PEHUEN
5	PEHUEN	19,35	46,26	PEHUEN
6	PEHUEN	17,76	51,46	PEHUEN
7	COMANGIR	27,58	62,13	TALENAY
8	COMANGIR	25,01	61,23	TALENAY
9	COMANGIR	24,68	58,68	COMANGIR
10	TALENAY	21,49	62,73	TALENAY
11	TALENAY	19,63	56,96	TALENAY
12	TALENAY	19,94	58,59	TALENAY
13	MARIBONDO	10,05	65,31	MARIBONDO
14	MARIBONDO	10,58	64,11	MARIBONDO
15	MARIBONDO	9,71	62,31	MARIBONDO
16	3GRNAINS	9,18	53,05	3GRNAINS
17	3GRNAINS	10,33	54,52	3GRNAINS
18	3GRNAINS	12,14	56,34	3GRNAINS

A representação gráfica das unidades amostrais nas regiões de discriminação de Fisher está mostrada na Figura3.

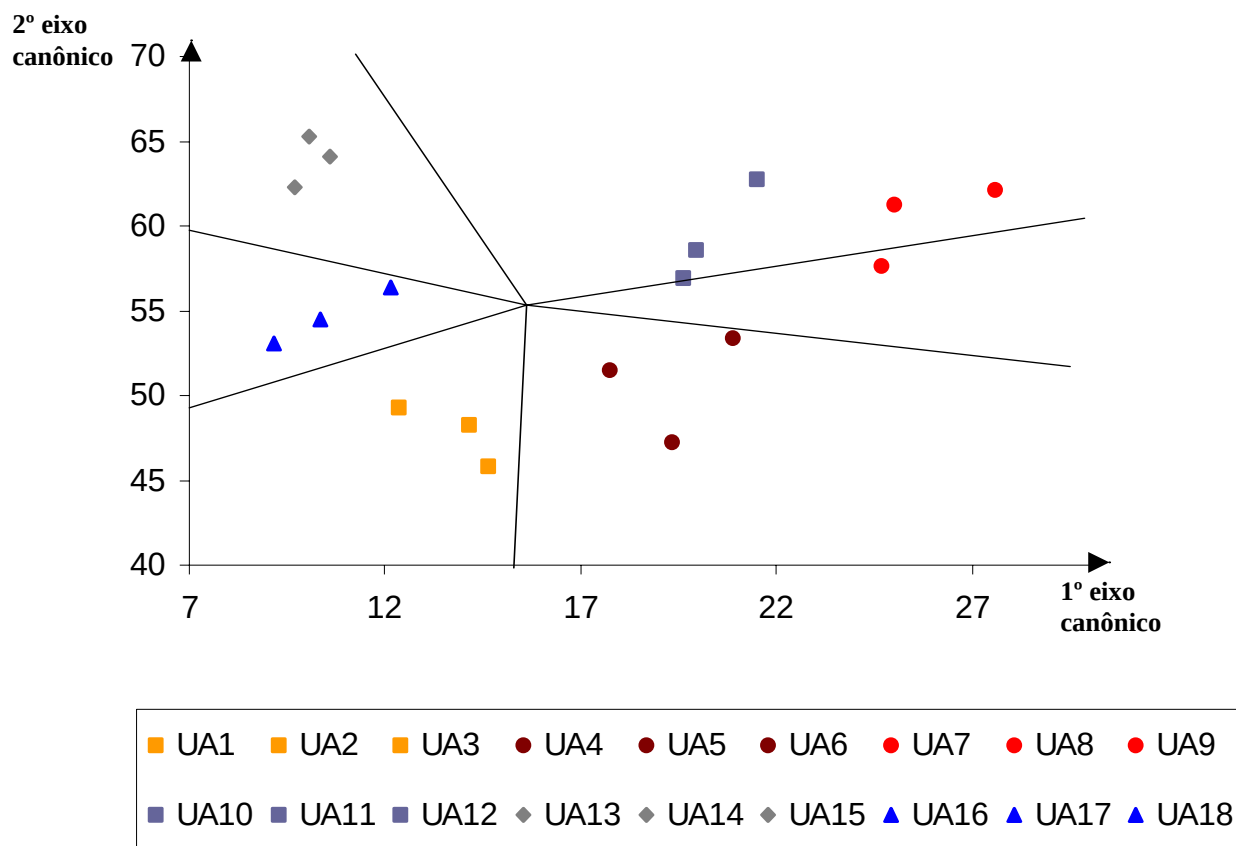


Figura3. Representação gráfica das unidades amostrais nos eixos canônicos.

Os resultados descritos nas Tabela4 e Figura2 indicam uma probabilidade freqüentista de classificação errônea da ordem de 11,11% (duas amostras do G3-Comangir classificadas erroneamente em G4-Talenay), revelando uma alta consistência do procedimento gráfico na discriminação das variedades de girassol.

6. CONCLUSÕES

O desenvolvimento do presente trabalho permite apresentar as seguintes conclusões:

- O programa **FISHER** oferece aos usuários das áreas aplicadas, subsídios interessantes para busca de objetivos agronômicos, referentes à discriminação de grupos e alocação de novas unidades em suas respectivas populações.
- O programa **FISHER** resultante do procedimento metodológico foi desenvolvido de maneira auto-explicativa, com a finalidade de fácil manuseio por qualquer pesquisador da área biológica.
- Com relação ao exemplo ilustrativo, o procedimento gráfico mostrou-se bastante eficiente na discriminação das variedades de girassol, com alta taxa de retenção de informação no sistema redimensionado (bidimensional) e baixa porcentagem de classificação errônea de novos elementos.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ALVAREZ, M.P.; MANCUSO,N.; FRUTOS,E. Genetic divergence among open-pollinated populations of sunflower (*Helianthus annuus,L.*). Proceeding of 11 International Sunflower Conference, Mar del Plata: 230-236,1991.

ANDERSON,T.W. An introduction to multivariate statistical analysis. 2nd ed. John Wiley, New York,45lp., 1984.

BARROSO, L.P. ARTES, R. Análise multivariada . Minicurso do 10º Simpósio de Estatística Aplicada à Experimentação Agronômica e 48º Reunião Anual da Sociedade Internacional de Biometria,UFLA-Lavras/MG,156 p..2003.

BASNET, K. Controls of environmental factors on pattern of montane rain forests in Puerto Rico. Tropical Ecology, 34(1):51-63, 1993.

BOX, G.E.P. A general distribution theory for a class of likelihood criteria. Biometrika, 36: 317-346, 1949.

CAMARANO, L.F. Estudo da divergência genotípica entre populações de girassol (*Helianthus annuus* L.). Dissertação de Mestrado em Agronomia (Área de Concentração: Genética e Melhoramento de Plantas). Universidade Federal de Goiás, Goiânia, 57p., 1997.

CARNEIRO,P.L.S; FONSECA, R; PIRES,A.V; TORRES FILHO,R.A; TORRES, R.A; PEIXOTO,J.O; LOPES,P.S; EUCLYDES, R.F. Estudo da divergência entre linhagens de matrizes de frangos de corte por meio de análise multivariada. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 54:(1) Fevereiro de 2002.

CROCCI, A.J. Problemas de classificação em uma das várias populações. Estudo da probabilidade de má classificação, baseado na função discriminante linear. Dissertação (Mestrado em Medicina -Área de Concentração: Bioestatística). Faculdade de Medicina de Ribeirão Preto USP, Ribeirão Preto, SP, 64p.,1979.

CRUZ, C.D. Aplicação de algumas técnicas multivariadas no melhoramento de plantas. Tese (Doutorado em Agronomia - Área de Concentração: Genética e Melhoramento de Plantas) Escola Superior de Agricultura Luiz de Queiroz -USP, Piracicaba, 92p., 1990.

CURI, P.R.; PADOVANI,C.R. Função discriminante linear: aplicação em experimentos biológicos. Naturalia,5:59-77,1980.

CURI, P.R. Análise de agrupamento complementada com ordenação pelos componentes principais e análise de variância multivariada. Aplicação em experimentos biológicos. Tese de Livre-Docência. Departamento de Bioestatística – IBBMA/UNESP, Botucatu –SP, 107p, 1982.

DAHER, R.F.; VÁZQUEZ, H. M.; PEREIRA, A.V. FERNANDES,A.M. Introdução e avaliação de clones de capim-elefante (*Pennisetum purpureum* Schum.) em Campos dos Goytacazes, R.J. Revista Brasileira de Zootecnia, Vol. 29, Nº 5: 1296 – 1301, 2000.

DANTAS, A.C.M; FORTES, G.R.L; SILVA, J.B; NEZI, A.N; RODRIGUES, A.C. Tolerância ao alumínio em porta-enxertos somaclonais de macieira cultivados em solução nutritiva. Pesquisa Agropecuária Brasileira, Vol. 36, Nº 4:615-623, 2001.

DEMEY, J.R.; ADAMS,M.; FREITES, H. Use of principal components analysis to classify farms. Agronomia Tropical Maracay,44 (3):475-497,1994.

EVERITT, B. S. Unresolved problems in clusters analysis. Biometrics,35:169-181,1979.

FERRÃO,M.A.G.; VIEIRA,C.; CRUZ,C.D.; CARDOSO,A.A. Divergência genética em feijoeiro em condições de inverno tropical. Pesquisa Agropecuária Brasileira, Vol. 37, Nº 8:1089-1098, 2001

FERREIRA,C.A.Utilização de técnicas multivariadas na avaliação da divergência genética entre clones de palma forrageira (*Opuntia* e *Nopalca*). Dissertação (Mestrado em Agronomia - Área de Concentração: Biometria) Universidade Federal Rural de Pernambuco - Recife, PE,63p., 2001.

FISHER, R.A. The use multiple measurements in toxonomic problems, Annals of Eugenics,7: 179 - 188,1936.

FONSECA, J.R. Emprego da análise multivariada na caracterização de germoplasma de feijão (*Phaseolus vulgaris*,L.). Tese (Doutorado em Agronomia - Área de Concentração: Fitotécnica) Escola Superior de Agricultura de Lavras, Lavras, MG, 123p., 1993.

FONSECA, R; PIRES, A.V; LOPES, P.S; TORRES, R.A; EUCLYDES, R.F. Estudo da divergência genética entre raças suínas utilizando técnicas de análise multivariada. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, Vol. 52, Nº 4:403-409, 2000.

GNANADESIKAN,R. Mehods for statistical data analysis of multivariate observations. John Wiley, New York, 311p.,1977,

GRAYBILL, F.A. Introduction to matrices with applications in statistics. Wadsworth Publishing Company, Belmont, 372p, 1969.

HARRIS, R.J. A primer of multivariate statistics. Academic Press, New York, 251p., 1975.

JOHNSON, R.A.; WICHERN, D.W. Applied multivariate statistical analysis. 4th ed Prentice-Hall International, New Jersey, 642p., 1998.

JOLICOUER, P.; MOSIMANN, J.E. Size and shape variation in the painted turtle. A principal component analysis. Growth, 24: 339- 354, 1960.

KENDALL, M.G. Factor analysis. Journal the Royal Statistical Society, Serie B, 12: 60-94, 1950.

LACHEMBRUCH, P.A. An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis. Biometrics, 23: 639-645, 1967.

LUCIO, P.S.; TOSCAO, E.M.M.; ABREU, M.L. Caracterização de séries climatológicas pontuais via análise de correspondência . Estudo de caso: Belo Horizonte – MG (Brasil). Revista Brasileira de Geofísica, Vol. 17, Nº 2-3: 193-207, 1999.

MALLANTS, D.; FEYEN, J. Defining homogeneous precipitation regions by means of principal components analysis. Journal of Applied Meteorology, 29(9): 892 – 901, 1990.

MARTEL, J.H.I.; FERRAUDO, A.S.; MÔRO, J.R.; PERECIN, D. Estatística multivariada na discriminação de raças amazônicas de pupunheiras (*Bactris gasipaes* Kunth) em Manaus (Brasil). Revista Brasileira de Fruticultura, 25(1): 115-118, 2003.

MESSETTI, A.V.L. Estudo da semelhança de genótipos de girassol (*Helianthus annuus* L.) com uso da distância generalizada de Mahalanobis na análise de agrupamento. Dissertação

(Mestrado Agronomia - Área de Concentração: Energia na Agricultura). Universidade Estadual Paulista, Botucatu, SP, 86p., 2000.

MESSETTI, A.V.L.; PADOVANI, C.R. Análise de agrupamento: estudo da semelhança de genótipos de girassol (*Helianthus annuus* L.) através da distância generalizada de Mahalanobis. Anais da 45a Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria (RBRAS), São Carlos - SP:210-212,2000.

MORGANO, M.A; QUEIROZ,S.C.N; FERREIRA,M.M.C. Determinação dos teores de minerais em sucos de frutas por espectrometria de emissão óptica em plasma indutivamente acoplado (ICP-OES).Ciência e Tecnologia de Alimentos, Vol. 19, Nº 3:344-348, 1999.

MORRISON, D.F. Multiple statistical methods. 2nd ed. Mac-Graw-Hill Kogakusha, LTD, Tokyo, 415p., 1976.

NISRA,R.K.; NI,I.H. Distinguishing beaked redfishes (deepwater redfish *Sebastes mentella* and *Sebastes fasciatus*) by discriminant analysis (with covariance) and multivariate analysis of covariance. Fish.Res., Vol. 40, Nº 9:31-35, 1983.

OSHIIWA, M.; PADOVANI,C.R. O uso de procedimento estatístico univariado e multivariado em ensaios agrônômicos com delineamento em blocos completos casualizados. Energia na Agricultura, Vol.17, Nº 2:58-68,2001.

PIASSI, M.; SILVA, M.A.; REGAZZI, A.Z.; TORRES, R.A.; SOARES,P.R.; CARNEIRO, M.C.. Avaliação de diferentes grupos genéticos de aves de postura, usando-se análise de variância multivariada. Revista da Sociedade Brasileira de Zootecnia.Vol.24, Nº 3;453-461,1995

PIASSI, M.; SILVA, M.A.; REGAZZI, A.Z.; TORRES, R.A.; SOARES,P.R.; JÚNIOR, R.A.A.T. Estudo da divergência entre oito grupos de aves de postura, por meio de técnicas de análise multivariada. Revista da Sociedade Brasileira de Zootecnia. 24(5);715-727,1995a.

PIMENTEL GOMES, F. A estatística moderna na pesquisa agropecuária. Piracicaba: POTAFOS, Piracicaba, 47p.,1984.

RAO,A.V.; PRASAD, A.S.R.; SAIKRISHNA, T; SESHU, D.V.; SIRINIVASAN, T.E. Genetic divergence among some brown planthopper resistant rice varieties. The Indian J. of Genet. and Plant Breed, 41: 179-185,1981.

RAO, C.R. Sir Ronald Aylmmer Fisher. The architect of multivariate analysis. Biometrics,20: 286-300, 1964

RAO, C.R. Multivariate analysis - some reminiscences on the origin and development. Sankyhã: 45, série B: 284-299,1983.

SEAL, H.L. Multivariate statistical analysis for biologists. Methuen and Co Ltd, London, 209p.,1964,

SEARLE, S.R. (1982) Matrix algebra useful for statistics. John Wiley, New York, 438p

SMITH, R.H. A geometrical comparisom of principal component analysis with analysis of variance. The Statistician,29(2):91-96,1980.

SOUZA, D.R.; SOUZA, A.L.; GAMA, J.R.V.; LEITE, H.G. Emprego de análise multivariada para estratificação vertical de florestas inequiiâneas. Revista Arvore,27(1):59-63,2003.

TATSUOKA, M.M. Multivariate analysis: Techniques for educational and psychological reseach. John Wiley, New York, 309p., 1971.

TORNERO, M.T.T. Utilização de componentes principais em análise de agrupamento, de classificação e de variância de experimentos biológicos. Relatório Final de Pesquisa da Comissão Permanente de Regime de Trabalho.Departamento de Bioestatística, IB/UNESP. Botucatu, 89p., 1990.

TUKEY, J.W. Exploratory data analysis. Addison - Wesley, Reading, 506p., 1977.

VIDA, G.; BEDO, Z.; JOLANKAI, M.; LANG, L. Use of principal components analysis in the selection of optimum agronomics treatments for winter wheat (*Triticum aestivum* L.) varieties with various baking qualities. Acta Agronomica Hungarica, 44(1):1-9, 1996.

VIDIGAL, M.C.G., VIDIGAL FILHO, P.S.; AMARAL Jr, A.T. BRACCINI, A.L. Divergência genética entre cultivares de mandioca por meio de estatística multivariada. Bragantia, 56(2):263-271, 1997.

APÊNDICE

APÊNDICE 1

QUADRO1. Vetores de resposta das características quantitativas do girassol.

Variedade	Parcela	Alt. Planta	Diâm. Caule	Alt. Capítulo	Diâm. Capítulo	Teor de Óleo
CORONA	1	1,36	21,18	95,22	16,62	34,68
CORONA	2	1,50	24,72	84,08	20,50	32,56
CORONA	3	1,45	16,23	104,72	15,62	30,12
CORONA	4	1,77	14,28	112,54	20,00	28,16
CORONA	5	1,12	15,24	113,45	14,45	32,69
CORONA	6	0,96	25,01	113,78	15,18	31,14
CORONA	7	1,45	23,45	114,89	16,78	28,48
CORONA	8	1,68	22,47	114,23	17,25	28,78
CORONA	9	1,15	21,55	135,42	12,26	36,58
CORONA	10	0,94	23,45	125,69	17,25	39,57
CORONA	11	1,75	24,78	136,78	18,25	35,00
CORONA	12	1,24	16,77	90,25	19,56	35,60
CORONA	13	1,51	18,24	94,24	17,12	37,24
CORONA	14	1,33	18,45	103,78	14,20	36,78
CORONA	15	0,95	16,74	104,89	13,54	35,99
CORONA	Amostra 1	1,40	20,85	112,40	16,70	34,02
CORONA	Amostra 2	1,27	19,20	108,45	17,83	31,94
CORONA	Amostra 3	1,22	19,35	108,94	16,03	36,04
CORONA	V. Mínimo	0,9400	14,2800	84,0800	12,2600	28,1600
CORONA	V. Máximo	1,7700	25,0100	136,7800	20,5000	39,5700
CORONA	Média	1,3440	20,1707	109,5973	16,5720	33,5580
CORONA	D. Padrão	0,2801	3,7762	15,3599	2,4046	3,5820
PEHUEN	1	1,41	23,00	105,32	16,00	32,81
PEHUEN	2	1,71	24,24	104,48	17,45	33,47
PEHUEN	3	1,58	19,00	102,84	20,45	32,37
PEHUEN	4	1,43	21,00	100,65	20,99	35,35
PEHUEN	5	1,61	24,00	102,56	19,14	36,51
PEHUEN	6	2,00	22,89	138,45	16,01	26,84
PEHUEN	7	1,98	23,45	140,78	17,58	31,78
PEHUEN	8	2,32	24,56	135,98	16,23	31,86
PEHUEN	9	2,42	25,68	136,45	25,00	36,32
PEHUEN	10	2,18	26,45	142,35	17,65	32,41
PEHUEN	11	1,99	28,89	128,32	19,14	36,01
PEHUEN	12	2,06	30,12	131,73	20,25	26,12
PEHUEN	13	2,28	31,25	142,23	21,54	30,78
PEHUEN	14	2,22	31,56	136,75	20,32	26,05
PEHUEN	15	2,56	32,78	129,73	23,54	35,98
PEHUEN	Amostra 4	2,10	27,40	141,01	20,35	34,10
PEHUEN	Amostra 5	1,74	24,03	110,15	24,00	32,29
PEHUEN	Amostra 6	1,95	23,72	127,17	18,39	34,17
PEHUEN	V. Mínimo	1,4100	19,0000	100,6500	16,0000	26,0500
PEHUEN	V. Máximo	2,5600	32,7800	142,3500	25,0000	36,5100
PEHUEN	Média	1,9833	25,9247	125,2413	19,4193	32,3107
PEHUEN	D. Padrão	0,3622	4,1260	16,6717	2,6951	3,6053

QUADRO2. Vetores de resposta das características quantitativas do girassol.

Variedade	Parcela	Alt. Planta	Diâm. Caule	Alt. Capítulo	Diâm. Capítulo	Teor de Óleo
COMANGIR	1	2,15	33,45	156,21	26,56	36,09
COMANGIR	2	2,20	34,12	159,32	27,78	31,40
COMANGIR	3	2,15	35,16	200,12	26,65	33,10
COMANGIR	4	2,20	34,11	189,45	19,78	35,17
COMANGIR	5	2,56	25,96	187,21	20,54	30,78
COMANGIR	6	2,34	26,45	174,23	19,98	33,45
COMANGIR	7	2,67	27,56	166,45	21,65	32,84
COMANGIR	8	2,58	32,78	198,12	22,45	31,58
COMANGIR	9	2,74	28,51	187,12	26,98	30,29
COMANGIR	10	2,78	32,54	198,21	27,79	35,36
COMANGIR	11	1,69	23,38	158,55	20,39	31,41
COMANGIR	12	1,88	23,45	174,23	23,10	34,64
COMANGIR	13	2,20	25,45	161,25	25,43	39,32
COMANGIR	14	2,65	25,21	172,25	24,30	41,39
COMANGIR	15	1,89	26,45	163,13	23,45	42,26
COMANGIR	Amostra 7	2,35	30,22	187,30	24,15	36,10
COMANGIR	Amostra 8	2,32	30,42	168,34	24,06	38,13
COMANGIR	Amostra 9	2,16	25,35	175,80	21,35	34,12
COMANGIR	V. Mínimo	1,6900	23,3800	156,2100	19,7800	30,2900
COMANGIR	V. Máximo	2,7800	35,1600	200,1200	27,7900	42,2600
COMANGIR	Média	2,3120	28,9720	176,3900	23,7887	34,6053
COMANGIR	D. Padrão	0,3402	4,2325	15,7239	2,9348	3,7755
TALENAY	1	2,60	24,58	144,91	20,37	36,46
TALENAY	2	1,98	26,78	145,65	19,29	40,64
TALENAY	3	2,19	27,45	146,79	24,11	38,39
TALENAY	4	1,85	28,56	147,19	19,26	30,86
TALENAY	5	2,43	29,87	150,45	18,46	38,12
TALENAY	6	2,09	29,17	151,79	17,21	37,24
TALENAY	7	1,71	20,23	184,49	24,24	36,00
TALENAY	8	1,89	25,12	160,49	15,04	39,92
TALENAY	9	1,83	24,23	161,19	23,30	33,70
TALENAY	10	2,17	19,45	163,45	20,27	42,87
TALENAY	11	2,00	20,44	186,42	16,26	32,08
TALENAY	12	1,70	24,56	175,45	18,24	32,74
TALENAY	13	1,98	28,45	149,98	23,78	33,26
TALENAY	14	2,45	30,45	148,69	22,26	39,76
TALENAY	15	2,26	29,45	147,89	22,02	32,31
TALENAY	Amostra 10	2,09	28,17	164,72	20,97	40,87
TALENAY	Amostra 11	2,04	19,54	150,62	19,17	35,85
TALENAY	Amostra 12	2,45	28,12	148,96	18,16	38,03
TALENAY	V. Mínimo	1,7000	19,4500	144,9100	15,0400	30,8600
TALENAY	V. Máximo	2,6000	30,4500	186,4200	24,2400	42,8700
TALENAY	Média	2,0753	25,9193	157,6553	20,2740	36,2900
TALENAY	D. Padrão	0,2727	3,6548	14,0748	2,9425	3,6752

QUADRO3. Vetores de resposta das características quantitativas do girassol.

Variedade	Parcela	Alt. Planta	Diâm. Caule	Alt. Capítulo	Diâm. Capítulo	Teor de Óleo
MARIBONDO	1	1,10	15,89	99,12	15,41	43,39
MARIBONDO	2	1,04	16,45	110,67	14,71	52,56
MARIBONDO	3	1,09	17,45	106,79	17,01	48,95
MARIBONDO	4	1,64	18,21	105,48	13,78	46,35
MARIBONDO	5	1,31	19,47	106,78	13,26	45,38
MARIBONDO	6	1,09	20,12	110,36	12,95	53,61
MARIBONDO	7	1,82	28,14	117,89	13,65	54,13
MARIBONDO	8	1,31	25,56	149,15	18,87	44,38
MARIBONDO	9	1,68	21,13	109,78	17,34	54,65
MARIBONDO	10	1,17	22,47	110,12	16,52	49,60
MARIBONDO	11	1,31	26,89	111,45	12,18	52,22
MARIBONDO	12	1,09	21,79	143,78	18,86	52,75
MARIBONDO	13	1,59	15,49	117,25	14,00	46,54
MARIBONDO	14	1,57	19,78	138,54	17,53	52,34
MARIBONDO	15	1,53	20,15	111,11	12,65	47,24
MARIBONDO	Amostra 13	1,47	21,23	123,15	16,72	51,13
MARIBONDO	Amostra 14	1,40	18,97	127,32	16,45	49,07
MARIBONDO	Amostra 15	1,33	19,17	123,85	14,98	47,92
MARIBONDO	V. Mínimo	1,0400	15,4900	99,1200	12,1800	43,900
MARIBONDO	V. Máximo	1,8200	28,1400	149,1500	18,8700	54,6500
MARIBONDO	Média	1,3560	20,5993	116,5513	15,2480	49,6060
MARIBONDO	D. Padrão	0,2605	3,8706	14,9351	2,2748	3,8238
3GRNAINS	1	1,87	18,49	111,87	15,99	40,16
3GRNAINS	2	1,88	13,45	110,25	11,81	45,15
3GRNAINS	3	1,91	17,18	135,94	16,89	37,21
3GRNAINS	4	1,30	25,15	152,78	9,32	47,87
3GRNAINS	5	1,33	14,12	105,98	14,52	44,36
3GRNAINS	6	1,39	13,79	104,65	10,98	41,13
3GRNAINS	7	1,99	17,49	102,28	9,16	44,42
3GRNAINS	8	1,68	25,17	103,64	10,71	36,83
3GRNAINS	9	1,92	16,24	105,99	14,54	39,45
3GRNAINS	10	1,26	18,97	108,23	16,80	48,28
3GRNAINS	11	1,27	12,89	99,12	13,41	38,39
3GRNAINS	12	1,28	13,45	110,67	12,71	44,56
3GRNAINS	13	1,29	14,45	106,79	15,01	40,95
3GRNAINS	14	1,74	15,21	105,48	11,78	38,35
3GRNAINS	15	1,95	16,47	106,78	11,26	37,38
3GRNAINS	Amostra 16	1,74	17,89	107,32	10,97	39,33
3GRNAINS	Amostra 17	1,56	24,12	115,42	11,76	40,85
3GRNAINS	Amostra 18	1,88	22,46	120,38	13,62	40,99
3GRNAINS	V. Mínimo	1,2600	12,8900	99,1200	9,1600	36,8300
3GRNAINS	V. Máximo	1,9900	25,1700	152,7800	16,8900	48,2800
3GRNAINS	Média	1,6040	16,8347	111,3633	12,9927	41,6327
3GRNAINS	D. Padrão	0,3024	3,8766	14,1489	2,5499	3,8628

APÊNDICE 2

Listagem dos procedimentos no MAPLE

```

> with(linalg):
> with(stats):
> ### Variáveis de inicialização
> ### tcond => número de condições
> ### tgp => número de grupos
> ### Telm => matriz com número de elementos p/ grupo
> ### A => matriz com os dados apurados
> tcond:=3:
> tgp:=3:
> Telm := matrix(1,3,[6,6,6]):
> Ret_Clip := RETORNO:
> Niv_Conf := 0.95:
>
A:=matrix(18,3,[6,7,9,8,8,10,11,15,22,17,12,13,18,24,20,19,22,17,7,5,10,8,9,9,13,16,20,16,11
,14,16,22,21,20,25,16,4,1,11,6,6,8,30,42,50,45
> ,38,35,25,30,32,27,34,32]):
> readlib(write):
> open(Ret_Clip):
>
#####
> ### calcula media das condições => medcol[i]
#####
> medcol := array(1..tgp,1..tcond):
> writeln(MediaColunas):
> deslin := 0:
> for l from 1 to tgp do:
>   for j from 1 to tcond do medcol[l,j] := 0:
>     for i from 1 to Telm[1,l] do nlin := i + deslin:
>       medcol[l,j] := medcol[l,j] + A[nlin,j] od:
>       medcol[l,j] := evalf(medcol[l,j] / Telm[1,l],6) :
>       writeln(l,j,medcol[l,j]) od:
>     deslin := deslin + Telm[1,l]:
>   od:
> totem := deslin:
>
#####
> ### calc matriz de variância/covariância => matval[j,j]
#####
> matval := array(1..tgp,1..tcond,1..tcond):
> deslin := 0:
> for l from 1 to tgp do:
>   for j from 1 to tcond do matval[l,j,j]:=0:
>     for i from 1 to Telm[1,l] do nlin := i + deslin:

```

```

>          matval[l,j,j] := matval[l,j,j] + A[nlin,j]^2 od :
>          matval[l,j,j] := ( matval[l,j,j] - Telm[1,l] * medcol[l,j]^2 ) / (Telm[1,l]-1) od:
>          for i from 2 to tcond do:
>              for j from 1 to i-1 do matval[l,i,j] := 0:
>                  for k from 1 to Telm[1,l] do nlin := k + deslin:
>                      matval[l,i,j] := matval[l,i,j] + A[nlin,i] * A[nlin,j] od:
>                      matval[l,i,j] := evalf(( matval[l,i,j] - Telm[1,l] * medcol[l,i] * medcol[l,j] ) /
(Telm[1,l]-1),6):
>                      matval[l,j,i] := matval[l,i,j] od od:
>                  deslin := deslin + Telm[1,l] od:
>          writeln(MatrizCovariancia):
>          for l from 1 to tgp do:
>              for i from 1 to tcond do:
>                  for j from 1 to tcond do writeln(l,i,j,matval[l,i,j]) od od od:
>
#####
> ###      calculo da matriz Spool => matspl[j,j]
#####
> matspl := array(1..tcond,1..tcond):
> mataux := array(1..tcond,1..tcond):
> writeln(MatrizSpool):
> for i from 1 to tcond do:
>     for j from 1 to tcond do:
>         matspl[i,j] := 0:
>         mataux[i,j] := 0:
>         for k from 1 to tgp do:
>             matspl[i,j] := matspl[i,j] + (Telm[1,k]-1)*matval[k,i,j]:
>             mataux[i,j] := mataux[i,j] + Telm[1,k]:
>         od:
>         matspl[i,j] := evalf(matspl[i,j] / ( mataux[i,j] - tgp ),6):
>         writeln(i,j,matspl[i,j]):
>     od:
> od:
> print(matspl);
> Yppk := array(1..tcond):
> for j from 1 to tcond do Yppk[j] := 0 :
>     for i from 1 to totelem do :
>         Yppk[j] := Yppk[j] + A[i,j] :
>     od:
> od:
> Yppk_Med := array(1..tcond):
> for j from 1 to tcond do Yppk_Med[j] := evalf(Yppk[j] / totelem,6) od:
> print (Yppk_Med);

```

```
#####
```

```

### Cálculo da matriz W
#####
> ng := totem - tgp:
> w := evalm(ng * matspl):
> print(w);
#####
##### Cálculo da Matriz B^0
#####
> z_aux := array(1..tcond, 1..tcond):
> aux_med := array(1..tcond, 1..tcond):
> aux_soma := array(1..tgp):
> for l from 1 to tgp do:
>   for c from 1 to tgp do:
>     aux := subvector(medcol, l, 1..tcond):
>   od:
>   aux_med := evalm(aux - Yppk_Med):
>   aux_soma[l] := evalf(multiply(aux_med, htranspose(aux_med)), 6):
> od:
> print(aux_soma);
> for z1 from 1 to tcond do for z2 from 1 to tcond do z_aux[z1, z2] := 0 od od:
> for z1 from 1 to tgp do z_aux := evalm(z_aux + aux_soma[z1]) od:
> print(z_aux);
>
#####
##### Preparando as Matrizes W e z_aux para cálculo das raízes
#####
> w_in := array(1..tcond, 1..tcond):
> w_bc := array(1..tcond, 1..tcond):
> for z1 from 1 to tcond do for z2 from 1 to tcond do w_in[z1, z2] := 0 od od:
> for z1 from 1 to tcond do for z2 from 1 to tcond do w_bc[z1, z2] := 0 od od:
> w_in := inverse(w):
> w_bc := evalf(multiply(w_in, z_aux), 6):
> print(w_bc);
>
>
#####
### Cálculo dos determinantes das matrizes e cálculo da constante M
#####
> m := 0:
> deslin := 0:
> matdet := array(1..tgp):
> for k from 1 to tgp do:
>   for i from 1 to tcond do:
>
>     for j from 1 to tcond do:
>       mataux[i, j] := matval[k, i, j]:

```

```

>      od:
>      od:
>      matdet[k] := det (mataux):
>      m := m + (Telm[1,k] - 1) * ln ( matdet [k] ):
>      deslin := deslin + Telm[1,k]:
> od:
> detpool := det (matspl):
> m := (deslin - tgp) * ln (detpool) - m:
#####
###Cálculo das constantes    C-1    e    Xo
#####
> deslin := 0:
> c := 0:
> for k from 1 to tgp    do:
>     c := c + 1 / ( Telm[1,k] - 1 ):
>     deslin := deslin + Telm[1,k]:
> od:
> totelem := deslin:
> c := c - 1 / ( deslin - tgp ):
> c := c * ( 2 * tcond^2 + 3 * tcond - 1 ) / ( 6 * ( tcond + 1 ) * ( tgp - 1 ) ):
> c := 1 - c:
> x := m * c:
> writeln(Nivel_Confianca):
> writeln(100- (Niv_Conf * 100)):
> writeln(Numero_Grupos):
> writeln(tgp):
> writeln(Numero_Condicoes):
> writeln(tcond):
> writeln(Qui_quadrado):
> writeln(x):
> ###   Cálculo dos Graus de Liberdade, do Q-Quadrado e do P-value
> gl := 1 / 2 * ( tgp - 1 ) * tcond * ( tcond + 1 ):
> Pvalue := 1 - statevalf[cdf, chisquare[gl]] (x):
> writeln(P_Value_1):
> writeln(Pvalue):
> writeln(GrauLiberdade_1):
> writeln(gl):
> writeln(Valor_X1):
> writeln(evalf(x)):
> writeln(Valor_C1):
> writeln(evalf(c)):
> statevalf[icdf, chisquare[gl]] (Niv_Conf ):

#####
### Determinando os 2 maiores auto-valores e seus vetores correspondentes
#####

```

```

> aux_valor:= array(1..tcond):
> aux_vetor:= array(1..tcond):
> pri_vetor := array(1..3):
> sec_vetor := array(1..3):
> At_valor:=evalf(eigenvals(w_bc),6):
> At_vetor := evalf(eigenvects(w_bc, 'radical'),6):
> for c from 1 to tcond do aux_valor[c] := 0 od:
> for v from 1 to tcond do aux_vetor[v] := 0 od:
> for z1 from 1 to 3 do pri_vetor[z1] := 0 od:
> for x1 from 1 to 3 do sec_vetor[x1] := 0 od:
> aux_valor:= sort([At_valor],(x,y)->evalb(x>y)):
> for b from 1 to tcond do:
>   aux_vetor := At_vetor[b]:
>   if (((aux_valor[1]+.000003) > aux_vetor[1]) and ((aux_valor[1] - .000003) <
aux_vetor[1]))then:
>     for c from 1 to 3 do:
>       pri_vetor[c] := aux_vetor[c]:
>     od:
>   fi:
>   if (((aux_valor[2]+.000003) > aux_vetor[1]) and ((aux_valor[2] - .000003) <
aux_vetor[1]))then:
>     for c from 1 to 3 do:
>       sec_vetor[c] := aux_vetor[c]:
>     od:
>   fi:
> od:
> print(aux_valor);
> print(aux_valor[1]);
> print(aux_valor[2]);
> print(pri_vetor);
> print(sec_vetor);
> #####
> #####Quanto explica cada autovetor (%)
> #####
> traco := 0:
> expec := 0:
> for l from 1 to tcond do:
>   for c from 1 to tcond do:
>     if l = c then:
>       traco := traco + w_bc[l,c]:
>     fi:
>
> od:
> od:
> exp_vec_maior := evalf((aux_valor[1] / traco) * 100,6):
> exp_vec_menor := evalf((aux_valor[2] / traco) * 100,6):

```

```
> expec := evalf(exp_vec_maior + exp_vec_menor,6):  
> print(exp_vec_maior);  
> print(exp_vec_menor);  
> print(expec);  
> close();
```

APÊNDICE 3

```

*****
*
* PROGRAMA : PROJETO - Menu Principal
*****
*
IF PCOUNT() > 0
    PARAMETER XYZPARM
ELSE
    XYZPARM="*****"
ENDIF
SET EXCLUSIVE OFF
SET TALK OFF
SET SCOREBOARD OFF
SET DATE BRITISH
SET DELETED ON
SET BELL OFF
SET COLOR TO W/N,W+/BG,,,W/N
SET KEY -3 TO CTRKEY
SET KEY -4 TO CALCULA
SET KEY -9 TO CTRKEY2
DECLARE
ZM[5],ZN[70],ZV1[4],ZV2[2],ZV3[1],ZV4[4],ZV5[1],LI[5],CI[5],LF[5],CF[5],VET[5],COR
ES[8]
BORDA="ÚÄ;³ÜÄÀ³ "
XYZHELP ="PROJETO.HLP"
XYZIMPRE="1"
XYZTEMPO=.03
XYZNOMSI='    Analise do Metodo Discriminante de Fisher
'+DTC(DATE())
XYZCABEC='Analise Discriminante'
XYZVER="Versao 1.0"
XYZESPER = 30
XYZINDEX = .T.
PUBLIC CAMPOAUX, RETORNA
RETORNA = .F.
CAMPOAUX = 0
PUBLIC X100
TIT = "01"
TIT = TIT + CHR(47)
TIT = TIT + "10"
TIT = TIT + "/"
TIT = TIT + "9"
TIT = TIT + "5"
DO WHILE DATE() < CTOD(TIT)
    CLEAR

```

```

! DATE
ENDDO
IF XYZPARM <> "*****"
    KEYBOARD(XYZPARM)
ENDIF
INICMENU()
SVTEL=GERCORES()
IF SVTEL="Z" .OR. XYZIMPRES="Z"
    CLEAR
    RETURN
ENDIF
OPC1=1
OPC2=1
OPC3=0
X100=1
VETOR="ZV1"
TECLA=0
TELA2=" "
DO WHILE TECLA != 27
    IF OPC3 = 0
        TECLA=PRILIN(OPC1)
    ENDIF
    IF TECLA != 13
        TECLA = 0
    LOOP
    ENDIF
SETCOLOR(STR(CORES[3],1)+"/"+STR(CORES[4],1)+", "+STR(CORES[4],1)+"/"+STR(CORES[3],1)+",,, "+STR(CORES[3],1)+"/"+STR(CORES[4],1))
@ 24,00 SAY SPACE(79)
@ 24,00 SAY "<F1> Auxilio    <F4> Tabelas    <F5> Calculadora    <Esc> Retorno ao Menu"
SETCOLOR(STR(CORES[5],1)+"/"+STR(CORES[6],1)+", "+STR(CORES[6],1)+"/"+STR(CORES[5],1)+",,, "+STR(CORES[5],1)+"/"+STR(CORES[6],1))
SET PRINTER TO PRN
IF OPC1 = 1
    IF OPC2 = 1
        DO ME010101    && Manutencao aos Dados Basicos da Pesquisa
    ELSEIF OPC2 = 2
        DO ME010401    && Importa dados do Excel
    ELSEIF OPC2 = 3

DO ME010301    && Listagem para Conferencia
    ELSEIF OPC2 = 4
        DO ME010201    && Manutencao aos Dados Coletados na Pesquisa
    ELSE
*      OPC2=0

```

```

ENDIF
ELSEIF OPC1 = 2
  IF OPC2 = 1
    ***** DO ME020101      && Exporta arquivo para o Maple
      DO KA020101      && Exporta arquivo para o Maple
    ELSEIF OPC2 = 2
    ***** DO ME020201      && Lista dados gerados pelo Maple
      DO KA020201      && Lista dados gerados pelo Maple
    ELSE
      OPC2=0
    ENDIF
  ELSEIF OPC1 = 3
    IF OPC2 = 1
      DO ME030101      && CONFIGURA DIRETORIO
    ELSE
      OPC2=0
    ENDIF
  ELSEIF OPC1 = 4
    IF OPC2 = 1
      DO HELP WITH "PROJETO",0,"PROJETO"
    ELSEIF OPC2 = 2
      MENSAGEM("CONFIRMA REORGANIZACAO DOS ARQUIVOS? (S/N) : ")
      IF LASTKEY() <> 83 .AND. LASTKEY() <> 115
        LOOP
      ENDIF
      MENSAGEM("DEVO COMPACTAR OS ARQUIVOS? (S/N) : ")
      IF LASTKEY() = 83 .OR. LASTKEY() = 115
        COMPACTA="S"
      ELSE
        COMPACTA="N"
      ENDIF
      RDX_ARQS()
      USE
      CLOSE ALL
      MENSAGEM("ARQUIVOS REORGANIZADOS... TECLE <Enter>")
    ELSEIF OPC2 = 3      && ATUALIZA DATA DO SISTEMA
      CLEAR
      ! DATE
      ! TIME

    ELSEIF OPC2 = 4      && ACERTA CORES
      ACERTACOR()
      if lastkey() <> 27
        KEYBOARD(CHR(13))
        SVTEL=GERCORES()
        IF SVTEL="Z"

```

```

        RETURN
    ENDIF
endif
opc2 = 0
else
    opc2 = 0
ENDIF
ELSE
    TECLA = 27
    SET COLOR TO
    CLEAR
    SELECT 1
    COMMIT
    CLOSE ALL
    RETURN
ENDIF
CLOSE ALL

SETCOLOR(STR(CORES[3],1)+"/"+STR(CORES[4],1)+",""+STR(CORES[4],1)+"/"+STR(C
ORES[3],1)+",,""+STR(CORES[3],1)+"/"+STR(CORES[4],1))
ENDDO
SET COLOR TO
CLEAR
COMMIT
CLOSE ALL
RETURN
*****
*
*           FUNCOES DO PROGRAMA
*
***  FUNCOES PARA REORGANIZAR ARQUIVOS
RETURN .T.
FUNCTION RDX_ARQS
    SET EXCLUSIVE ON
    XYZINDEX = .F.
    IF .NOT.ABRIR("/BAS/MOV/EST/FAL/CFG/",".T.")
        RETURN .F.
    ENDIF
    MOSTRA("REORGANIZANDO ARQUIVO DE INFORMACOES BASICAS DAS
PESQUISAS")
    SELECT MEBASE
    USE MEBASE
    IF COMPACTA = "S"
        PACK
    ENDIF
    INDEX ON MBCODPROJ
***

```

TO MEBASE

```

MOSTRA("REORGANIZANDO ARQUIVO COM OS DADOS LEVANTADOS")
SELECT MEMOVITO
USE MEMOVITO
IF COMPACTA = "S"
    PACK
ENDIF
INDEX ON MMCODPROJ+STR(MMGRUPO,3)+STR(MMELEMENTO,3)      TO
MEMOVITO
MOSTRA("REORGANIZANDO ARQUIVO COM OS DADOS ESTIMADOS")
SELECT MEESTIMA
USE MEESTIMA
IF COMPACTA = "S"
    PACK
ENDIF
INDEX ON MECODPROJ+STR(MEGRUPO,3)+STR(MEELEMENTO,3)      TO
MEESTIMA
***
MOSTRA("REORGANIZANDO ARQUIVO COM OS DADOS SEM FALHA")
SELECT MEFALHA
USE MEFALHA
IF COMPACTA = "S"
    PACK
ENDIF
INDEX ON MFCODPROJ+STR(MFGRUPO,3)+STR(MFELEMENTO,3)      TO
MEFALHA
***
***
MOSTRA("EFETIVANDO ATUALIZACOES EM DISCO")
COMMIT
SET EXCLUSIVE OFF
XYZINDEX = .T.
RETURN .T.

```

**** FUNCAO PARA ABERTURA DOS ARQUIVOS

*

FUNCTION ABRIR

* XYZARQ -> ARQUIVOS A SEREM ABERTOS SEPARADOS POR /

* XYZMOD0 -> MODO ABERTURA ARQUIVOS: .T.=EXCLUSIVO //

.F.=COMPARTILHADO

* XYZINDEX-> SE = .F. => NAO ABRIR INDICES

* SE = .T. => ABRIR INDICES (E' O DEFAULT)

PARAMETER XYZARQ,XYZMOD0

*

MOSTRA("VERIFICANDO DISPONIBILIDADE DOS ARQUIVOS")

IF AT("/BAS/",XYZARQ) > 0

IF NETUSE("MEBASE","MEBASE",XYZMOD0,XYZESPER)

IF XYZINDEX

SET INDEX TO MEBASE

ENDIF

ELSE

CLOSE ALL

RETURN .F.

ENDIF

ENDIF

IF AT("/MOV/",XYZARQ) > 0

IF NETUSE("MEMOVITO","MEMOVITO",XYZMOD0,XYZESPER)

IF XYZINDEX

SET INDEX TO MEMOVITO

ENDIF

ELSE

CLOSE ALL

RETURN .F.

ENDIF

ENDIF

IF AT("/EST/",XYZARQ) > 0

IF NETUSE("MEESTIMA","MEESTIMA",XYZMOD0,XYZESPER)

IF XYZINDEX

SET INDEX TO MEESTIMA

ENDIF

ELSE

CLOSE ALL

RETURN .F.

ENDIF

ENDIF

IF AT("/FAL/",XYZARQ) > 0

IF NETUSE("MEFALHA","MEFALHA",XYZMOD0,XYZESPER)

IF XYZINDEX

SET INDEX TO MEFALHA

```

    ENDIF
ELSE
    CLOSE ALL
    RETURN .F.
ENDIF
ENDIF
IF AT("/CFG/",XYZARQ) > 0
    IF NETUSE("MECONFIG","MECONFIG",XYZMODO,XYZESPER)
*   IF XYZINDEX
*       SET INDEX TO MECONFIG
*   ENDIF
    ELSE
        CLOSE ALL
        RETURN .F.
    ENDIF
ENDIF
XYZINDEX = .T.
MOSTRA(" ")
RETURN .T.
*****
*
*
* FUNCAO DE OPCA0 VIDEO/IMPRESSORA/ARQUIVO
*
FUNCTION IMPRVIDE
X = " "
DO WHILE (CONF = " ") .AND. (LASTKEY() # 27) .AND. (LASTKEY() # 5)
    MOSTRA("SAIDA EM V_ideo, I_mpressora ou A_rquivo? (V/I/A): V")
    @ 24,69 SAY " "
    INKEY(0)
    ZWPERIFE = "V"
    IF LASTKEY() = 27 .OR. LASTKEY() = 5
        X = " "
        LOOP
    ENDIF
    X = CHR(LASTKEY())
    NLIN = 22
    IF X = "v" .OR. LASTKEY() = 13
        X = "V"
    ELSEIF X = "i"
        X = "I"
    ELSEIF X = "a"
        X = "A"
    ENDIF
    IF X <> "V" .AND. X <> "I" .AND. X <> "A"
        MENSAGEM("OPCOES VALIDAS => V_ideo, I_mpressora ou A_rquivo")
        CONF = " "
    ENDIF

```

```

    LOOP
ENDIF
CONF = X
IF X = "I"
    MENSAGEM("POSICIONE TOPO DA PAGINA / COLOQUE IMPRESSORA EM
LINHA / TECLE ALGO")
    NLIN = 60
    IF XYZIMPRE = "1"
        SET PRINTER TO LPT1
    ELSEIF XYZIMPRE = "2"
        SET PRINTER TO LPT2
    ELSEIF XYZIMPRE = "3"
        SET PRINTER TO LPT3
    ELSE
        SET PRINTER TO PRN
    ENDIF
    SET DEVICE TO PRINT
    DO WHILE .NOT. ISPRINTER()
SET DEVICE TO SCREEN
        MENSAGEM("COLOQUE A IMPRESSORA EM LINHA E TECLE ALGO PARA
INICIAR IMPRESSAO")
        IF LASTKEY() = 27
            RETURN " "
        ENDIF
        SET DEVICE TO PRINT
    ENDDO
SET DEVICE TO SCREEN
MOSTRA("IMPRESSAO EM ANDAMENTO... AGUARDE...")
    SET DEVICE TO PRINT
    PFKX=PCOL()
    SETPRC(0,PFKX)
ELSEIF X = "A"
    NLIN = 59
    SET PRINTER TO SPOOL.TXT
SET DEVICE TO SCREEN
    MOSTRA("GERACAO DO ARQUIVO SPOOL.TXT EM ANDAMENTO...
AGUARDE...")
    SET DEVICE TO PRINT
    ZWPERIFE = "A"
ELSE
    NLIN = 22
ENDIF
RETURN X
ENDDO
RETURN X
*****
FUNCTION ELICE

```

```

PARAMETER ELICE,ELICE1
SET DEVICE TO SCREEN
@ 21,75 SAY SUBSTR("/-\\",ELICE%4+1,1)
IF ELICE1 = "I"
    SET DEVICE TO PRINT
ENDIF
RETURN .T.
*
* FUNCAO PARA CONTROLAR SALTO DE PAGINA
*
FUNCTION SALTAPAG
PARAMETER WKDETREL,WKNOMREL
IF LIN > NLIN
    IF CONF = "V"
        IF PAG > 0
            MOSTRA("TECLE  =>  <Enter> P/ CONTINUAR   <Esc> P/ ENCERRAR")
            XYZABC = " "
            @ 24,78 GET XYZABC PICT "!"
            READ
            IF LASTKEY() = 27
                GOTO BOTTOM
            SKIP
            RETURN .T.
        ENDIF
    ENDIF
    CLEAR
    ELSE
        @ PROW(),PCOL() SAY CHR(18)
    ENDIF
    PAG = PAG + 1
    @ 00,00 SAY "Informiza    "+WKNOMREL
    @ 00,64 SAY "Emis.: " + DTOC(DATE())
    @ 01,00 SAY WKDETREL
    @ 01,64 SAY "Pag. : " + STR(PAG,8)
    @ 02,00 SAY REPLICATE("Ä",79)
ELSE
    RETURN .F.
ENDIF
LIN = 4
RETURN .T.
*
*
* FUNCAO PARA INICIALIZAR VARIABEIS DO MENU
*
*
FUNCTION INICMENU
*
```

```

ZM[1]=" EXPERIMENTOS "
ZM[2]=" MAPLE-V-3  "
ZM[3]=" CONFIGURA  "
ZM[4]=" APOIO  "
ZM[5]=" FINALIZA  "
*
ZV1[1]="DADOS BASICOS"
ZV1[2]=" PLANILHA  "
ZV1[3]=" RELATORIO  "
ZV1[4]=" VARIAVEIS  "
ZV2[1]=" EXPORTA  "
ZV2[2]=" RELATORIO  "
ZV3[1]=" DIRETORIO  "
ZV4[1]=" AUXILIO  "
ZV4[2]=" REORGANIZA  "
ZV4[3]=" DATA  "
ZV4[4]=" CORES  "
ZV5[1]=" FINALIZA  "
*
* MENSAGENS
*
AFILL(ZN,SPACE(1),1,70)
ZN[11]="Cadastramento Basicos do Experimento  "
ZN[12]="Importa Dados de Planilha (Formato TXT)  "
ZN[13]="Listagem para Conferencia  "
ZN[14]="Manutencao das Variaveis do Experimento  "
ZN[21]="Exporta Arquivo para Calculos no MAPLE-V-3"
ZN[22]="Relatorio com Dados Gerados no MAPLE-V-3  "
ZN[31]="Configura Diretorio p/ o MAPLE-V-3  "
ZN[41]="Acesso ao `HELP' do sistema  "
ZN[42]="Reorganiza os Arquivos do Sistema  "
ZN[43]="Altera data e hora da maquina  "
ZN[44]="Configuracao de cores da tela  "
*ZN[45]="Acesso ao DOS com retorno ao SISTEMA  "
*ZN[46]="Lista arquivo SPOOL.TXT na impressora  "
ZN[51]="Retorno ao DOS / Saida do Sistema  "
*
* POSICIONAMENTO DOS MENUS
*
AFILL(LI,6)
LI[5]=4
CI[1]=3
CI[2]=18
CI[3]=33
CI[4]=48
CI[5]=63
LF[1]=11

```

```

LF[2]=09
LF[3]=08
LF[4]=11
LF[5]=06
CF[1]=18
CF[2]=33
CF[3]=48
CF[4]=63
CF[5]=78
VET[1]=" INCLUSAO "
VET[2]=" EXCLUSAO "
VET[3]=" ALTERACAO "
VET[4]=" CONSULTA "
VET[5]=" RELACAO "
RETURN .T.
*
*
*****
*
* PROGRAMA : KA020201 - Lista Relatorio Gerado no Maple V.3 *
*****
*
IF .NOT. ABRIR("/BAS/MOV/EST/FAL/",F.)
RETURN
ENDIF
DIRETORIO = MEMOREAD("DIRETORI.DIR")
DIRETORIO = ALLTRIM(DIRETORIO)
CPO = 1
XMAXLIN=0
CONF="V"
DO WHILE CPO <> 0
DO CASE
CASE CPO = 1
CAM01 = SPACE(08)      && CODIGO DO EXPERIMENTO
CAM02 = SPACE(40)      && TITULO DA EXPERIMENTO
CAM03 = 0              && NUMERO DE MOMENTOS
CAM04 = 0              && NUMERO DE GRUPOS
DIA    = DTOC(DATE())
HORA   = SUBSTR(TIME(),1,5)
NIVELCONF = 0
VALLILL = 0
UMMENOSALFA = " "
TGP    = 0
TCOND  = 0
GLQQ   = 0
GLIH1  = 0
GLFH1  = 0

```

```

GLIH2 = 0
GLFH2 = 0
GLIH3 = 0
GLFH3 = 0
GLIH4 = 0
GLFH4 = 0
GLIH5 = 0
GLFH5 = 0
GLIH6 = 0
GLFH6 = 0
VALX1 = 0
VALC1 = 0
FTETA = 0
TETAMAX = 0
FH0 = 0
ESSE = 0
EME = 0
ENE = 0
PVALUE1 = 0
PVALUE2 = 0
PVALUE3 = 0
PVALUE4 = 0
PVALUE5 = 0
PVALUE6 = 0
CONF = "V"
XPG = 0
TIPOLIN = "NORMAL"
UNLOCK ALL
ME0202TELA()
SAVE SCREEN TO XTELA
@ 17,12 GET CAM01 PICT "!!!!!!!!" VALID ME0202CA01(CAM01)
@ 21,73 GET CONF PICT "!" VALID CONF="V" .OR. CONF="I" .OR. CONF="A"
.OR. LASTKEY() = 5
READ
IF LASTKEY() = 27
    CPO = 0
    LOOP
ELSEIF CONF = "V" .OR. CONF = "I" .OR. CONF = "A"
    ME0202REL()
    SET DEVICE TO SCREEN
    MENSAGEM("FIM DO RELATORIO. TECLE <Enter> ")
    RESTORE SCREEN FROM XTELA
ENDIF
CONF = " "
CPO = 1
ENDCASE
ENDDO

```

```

UNLOCK ALL
CLOSE ALL
RETURN
* FUNCAO LISTA RELATORIO A PARTIR DO ARQUIVO DO MAPLE V-3
FUNCTION ME0202REL
  XX = ALLTRIM(CAM01)+"C"
  IF FILE(XX)
    ! DEL &XX
  ENDIF
  Z = DIRETORIO + "\BIN\" + ALLTRIM(CAM01) + "C > LI_XAO_X.LIX"
  @ 19,06 SAY ""
  ! COPY &Z
  XX = ALLTRIM(CAM01)+"E"
  IF FILE(XX)
    ! DEL &XX
  ENDIF
  Z = DIRETORIO + "\BIN\" + ALLTRIM(CAM01) + "E > LI_XAO_X.LIX"
  @ 19,06 SAY ""
  ! COPY &Z
  XX = ALLTRIM(CAM01)+"O"
  IF FILE(XX)
    ! DEL &XX
  ENDIF
  Z = DIRETORIO + "\BIN\" + ALLTRIM(CAM01) + "O > LI_XAO_X.LIX"
  @ 19,06 SAY ""
  ! COPY &Z
  @ 19,06 SAY " "
  DECLARE IDGRUPO[CAM04]
  FOR IX = 1 TO CAM04
    IDGRUPO[IX] = IX
  NEXT
  IF CONF = "I" .OR. CONF = "A"
    MOSTRA("IMPRIMINDO RELATORIO... AGUARDE...")
    SET DEVICE TO PRINT
    IF CONF="A"
      XREL = "REL02.TXT"
      SET PRINTER TO &XREL
    ENDIF
  ENDIF
  NATUREZA = " Original "
  XX = ALLTRIM(CAM01)+"O"
  IF FILE(XX)
    ME0202LISMAPLE()
  ENDIF
  IF CONF = "I" .OR. CONF = "A"
    EJECT
  ENDIF

```

```

IF CONF = "A"
  SET PRINTER TO PRN
ENDIF
RETURN .T.
FUNCTION ME0202LISMAPLE
  MAPLE = MEMOREAD(XX)
  XMAPLE = MAPLE
  FINAL = MLCOUNT(MAPLE,100)
  LINHA = " "
  TRATO = " "
  TRATO1 = 0
  IF CONF = "V"
    XMAXLIN = 24
    IF NATUREZA <> "Consistido"
      CLEAR
    ENDIF
  ELSE
    XMAXLIN = 60
  ENDIF
  XLIN = 100
  XCOL = 0
  FOR I = 1 TO FINAL
    XCOLFIM = AT(CHR(13)+CHR(10),XMAPLE)
    XLINMAPLE = SUBSTR(XMAPLE,1,XCOLFIM-1)
    XMAPLE = SUBSTR(XMAPLE,XCOLFIM+2,LEN(XMAPLE))
    IF LASTKEY() = 27 .OR. LEN(XLINMAPLE) < 1
      EXIT
    ENDIF
    LINHA = XLINMAPLE
    IF LINHA = "MediaColunas"
      ME0202MATRIZ()
      XLIN = XLIN + 2
      TRATO = "MatrizSpoolX"
      TIPOLIN = "CABEC"
      ME0102SALTAPG()
      @ XLIN,00 SAY "[ Matriz Amostral Comum de Variancia ]"
    ELSEIF LINHA = "Qui_quadrado"
      TRATO = "Qui_quadradoX"
    ELSEIF LINHA = "Nivel_Confianca"
      TRATO = "Nivel_ConfiancaX"
    ELSEIF LINHA = "Numero_Grupos"
      TRATO = "Numero_GruposX"
    ELSEIF LINHA = "Numero_Condicoes"
      TRATO = "Numero_CondicoesX"
    ELSEIF LINHA = "P_Value_1"
      TRATO = "P_Value_1X"
    ELSEIF LINHA = "GrauLiberdade_1"

```

```

        TRATO = "GrauLiberdade_1X"
ELSEIF LINHA = "MatrizW"
    XLIN = XLIN + 2
    TRATO = "MatrizWX"
    TIPOLIN = "CABEC"
    ME0102SALTAPG()
    @ XLIN,00 SAY "[ Matriz W ]"
ELSEIF LINHA = "MatrizB0"
    XLIN = XLIN + 2
    TRATO = "MatrizB0X"
    TIPOLIN = "CABEC"
    ME0102SALTAPG()
    @ XLIN,00 SAY "[ Matriz B0 ]"
ELSEIF LINHA = "MatrizXWB0"
    XLIN = XLIN + 2
    TRATO = "MatrizXWB0X"
    TIPOLIN = "CABEC"
    ME0102SALTAPG()
    @ XLIN,00 SAY "[ Matriz WB0 ]"
ELSEIF LINHA = "Autovalor_1"
    TRATO = "Autovalor_1X"
ELSEIF LINHA = "Autovalor_2"
    TRATO = "Autovalor_2X"
ELSEIF LINHA = "Autovetor_1"
    TRATO = "Autovetor_1X"
ELSEIF LINHA = "Autovetor_2"
    TRATO = "Autovetor_2X"
ELSEIF LINHA = "Vetormaior"
    RATO = "VetormaiorX"
ELSEIF LINHA = "Vetormenor"
    TRATO = "VetormenorX"
ELSEIF LINHA = "TotalFinal"
    TRATO = "TotalFinalX"
ELSEIF LINHA = "Error"
    TRATO = "Error"
ELSE
    IF TRATO = "MatrizSpoolX"
        CAM10 = ALLTRIM(SUBSTR(LINHA,1,4))
        CAM11 = VAL(ALLTRIM(SUBSTR(LINHA,5,4)))
        CAM20 = ME0202REAL(SUBSTR(LINHA,9,LEN(LINHA)))
        IF CAM11 = 1
            ME0102SALTAPG()
            @ XLIN,02 SAY "L=" + CAM10
            @ XLIN,09 SAY CAM20 PICT "9,999,999.999999"
        ELSE
            IF (CAM11 % 4) = 1
                ME0102SALTAPG()

```

```

ENDIF
XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 9
@ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
ENDIF
ELSEIF TRATO = "Numero_GruposX"
TGP = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
ELSEIF TRATO = "Numero_CondicoesX"
TCOND = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
ELSEIF TRATO = "Valor_X1X"
VALX1 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
ELSEIF TRATO = "Valor_C1X"
VALC1 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
ELSEIF TRATO = "Nivel_ConfiancaX"
NIVELCONF = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
UMMENOSALFA = SUBSTR(LINHA,1,LEN(LINHA))
ELSEIF TRATO = "Qui_quadradoX"
CAM20 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
XLIN = XLIN + 2
ME0102SALTAPG()
@ XLIN,00 SAY "<<< Teste de Homocedasticidade >>> "
XLIN = XLIN + 1
ME0102SALTAPG()
@ XLIN,00 SAY "[ Qui-Quadrado Calculado ] ="
@ XLIN,29 SAY CAM20 PICT "9,999,999.999999"
ELSEIF TRATO = "P_Value_1X"
PVALUE1 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
ELSEIF TRATO = "GrauLiberdade_1X"
GLQQ = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
XLIN = XLIN + 1
ME0102SALTAPG()
@ XLIN,00 SAY "[ P-Value ] ="
@ XLIN,14 SAY PVALUE1 PICT "9.999999"
@ XLIN,24 SAY "a"
@ XLIN,27 SAY UMMENOSALFA && PICT "99"
@ XLIN,32 SAY "% Com "
@ XLIN,40 SAY GLQQ PICT "99"
@ XLIN,42 SAY " G.L."
TRATO = " "
ELSEIF TRATO = "MatrizWX" .OR. TRATO = "MatrizB0X" .OR. TRATO =
"MatrizXWB0X"
CAM10 = ALLTRIM(SUBSTR(LINHA,1,4))
CAM11 = VAL(ALLTRIM(SUBSTR(LINHA,5,4)))
CAM20 = ME0202REAL(SUBSTR(LINHA,9,LEN(LINHA)))
IF CAM11 = 1
ME0102SALTAPG()
@ XLIN,02 SAY "L=" + CAM10
@ XLIN,09 SAY CAM20 PICT "9,999,999.999999"

```

```

ELSE
  IF (CAM11 % 4) = 1
    ME0102SALTAPG()
  ENDIF
  XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 9
  @ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
ENDIF
ELSEIF TRATO = "Autovalor_1X"
  CAM20 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
  XLIN = XLIN + 2
  ME0102SALTAPG()
  @ XLIN,00 SAY "<<< Determinando os dois maiores auto-valores e seus vetores
correspondentes >>> "
  XLIN = XLIN + 1
  ME0102SALTAPG()
  @ XLIN,00 SAY "[ Maior auto-valor      ] ="
  @ XLIN,32 SAY CAM20 PICT "9,999,999.999999"
ELSEIF TRATO = "Autovalor_2X"
  CAM20 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
  ME0102SALTAPG()
  @ XLIN,00 SAY "[ Segundo maior auto-valor ] ="
  @ XLIN,32 SAY CAM20 PICT "9,999,999.999999"
ELSEIF TRATO = "Autovetor_1X"
  XTRAT = SUBSTR(LINHA,1,2)
  IF XTRAT = "1 "
    CAM20 = ME0202REAL(SUBSTR(LINHA,5,LEN(LINHA)))
    XLIN = XLIN + 2
    ME0102SALTAPG()
    @ XLIN,00 SAY "<<< Auto-vetores associados a seus Auto-valores >>> "
    XLIN = XLIN + 1
    ME0102SALTAPG()
    @ XLIN,00 SAY "[ Primeiro auto-vetor ] ="
    @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
  ELSEIF XTRAT = "2 "
    CAM20 = ME0202REAL(SUBSTR(LINHA,5,LEN(LINHA)))
    ME0102SALTAPG()
    @ XLIN,00 SAY "[ Multiplicidade      ] ="
    @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
    ME0102SALTAPG()
    @ XLIN,00 SAY " Vetor: "
    XLIN = XLIN - 1
  ELSEIF XTRAT = "3 "
    CAM10 = ALLTRIM(SUBSTR(LINHA,5,4))
    CAM11 = VAL(ALLTRIM(SUBSTR(LINHA,5,4)))
    CAM20 = ME0202REAL(SUBSTR(LINHA,9,LEN(LINHA)))
    IF CAM11 = 1
      ME0102SALTAPG()

```

```

        @ XLIN,09 SAY CAM20 PICT "9,999,999.999999"
    ELSE
        IF (CAM11 % 4) = 1
            ME0102SALTAPG()
        ENDIF
        XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 9
        @ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
    ENDIF
ENDIF
ELSEIF TRATO = "Autovetor_2X"
    XTRAT = SUBSTR(LINHA,1,2)
    IF XTRAT = "1 "
        XLIN = XLIN + 1
        CAM20 = ME0202REAL(SUBSTR(LINHA,5,LEN(LINHA)))
        ME0102SALTAPG()
        @ XLIN,00 SAY "[ Segundo auto-vetor ] ="
        @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
    ELSEIF XTRAT = "2 "
        CAM20 = ME0202REAL(SUBSTR(LINHA,5,LEN(LINHA)))
        ME0102SALTAPG()
        @ XLIN,00 SAY "[ Multiplicidade   ] ="
        @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
        ME0102SALTAPG()
        @ XLIN,00 SAY " Vetor:"
        XLIN = XLIN - 1
    ELSEIF XTRAT = "3 "
        CAM10 = ALLTRIM(SUBSTR(LINHA,5,4))
        CAM11 = VAL(ALLTRIM(SUBSTR(LINHA,5,4)))
        CAM20 = ME0202REAL(SUBSTR(LINHA,9,LEN(LINHA)))
        IF CAM11 = 1
            ME0102SALTAPG()
            @ XLIN,09 SAY CAM20 PICT "9,999,999.999999"
        ELSE
            IF (CAM11 % 4) = 1
                ME0102SALTAPG()
            ENDIF
            XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 9
            @ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
        ENDIF
    ENDIF
ELSEIF TRATO = "VetormaiorX"
    CAM20 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
    XLIN = XLIN + 2
    ME0102SALTAPG()
    @ XLIN,00 SAY "<<< Retencao de informacao nos auto-vetores >>> "
    XLIN = XLIN + 1
    ME0102SALTAPG()

```

```

      @ XLIN,00 SAY "[ Primeiro auto-vetor ] ="
      @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
ELSEIF TRATO = "VetormenorX"
  CAM20 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
  ME0102SALTAPG()
  @ XLIN,00 SAY "[ Segundo auto-vetor ] ="
  @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
  ELSEIF TRATO = "TotalFinalX"
    CAM20 = ME0202REAL(SUBSTR(LINHA,1,LEN(LINHA)))
    XLIN = XLIN + 1
    ME0102SALTAPG()
    @ XLIN,00 SAY "[ T o t a l      ] ="
    @ XLIN,26 SAY CAM20 PICT "9,999,999.999999"
  ENDIF
ENDIF
NEXT
RETURN .T.
* FUNCAO PARA CONVERSAO DE NOTACAO CIENTIFICA PARA REAL
FUNCTION ME0202REAL
PARAMETER X
IF AT("e",X) = 0
  CAMY = VAL(X) * 1
ELSE
  CAMY1 = VAL(SUBSTR(X,1,(AT("e",X)-1))) * 1
  CAMY2 = VAL(SUBSTR(X,(AT("e",X)+1),LEN(X))) * 1
  CAMY = CAMY1*(10^CAMY2)
ENDIF
RETURN CAMY

```

```

*#####
*  FUNCOES DEFINIDAS PARA CONSISTENCIA DOS CAMPOS SOLOCITADOS  #
*#####
*  FUNCAO ME0202CA01 - CAM01 - CODIGO DO EXPERIMENTO
FUNCTION ME0202CA01
  PARAMETER X
  IF X = "      " .OR. EMPTY(X)
    MENSAGEM("INFORME O CODIGO DO EXPERIMENTO")
    RETURN .F.
  ENDIF
  SELECT MEBASE
  GOTO TOP
  SEEK X
  IF EOF()
    MENSAGEM("EXPERIMENTO NAO CADASTRADO")
    RETURN .F.
  ELSE
    CAM02 = MBTITPROJ
    CAM03 = MBNUMEXP
    CAM04 = MBNUMGP
    @ 17,23 SAY CAM02 PICT
    "XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX"
    @ 18,24 SAY CAM03 PICT "99"
    @ 18,48 SAY CAM04 PICT "999"
  ENDIF
  Z = DIRETORIO + "\BIN"
  SET PATH TO &Z
  XXX = ALLTRIM(X)
  IF .NOT. FILE(XXX+"O") .AND. .NOT. FILE(XXX+"C") .AND. .NOT. FILE(XXX+"C")
    SET PATH TO
    MENSAGEM("ARQUIVO GERADO PELO MAPLE NAO DISPONIVEL [NA PASTA
DO MAPLE]")
    RETURN .F.
  ENDIF
  SET PATH TO
  RETURN .T.
*
*  FUNCAO ME0202TELA - TELA DO PROGRAMA
*
FUNCTION ME0202TELA
  TELABASE(3)
  @ 17,04 SAY "Codigo:      - "
  @ 18,04 SAY "Numero de Variaveis:      Numero de Grupos: "
  @ 21,04 SAY "Listar em Video ou Impressora? (<V>ideo /<I>mpressora /<Esc>Aborta): "
  RETURN .T.
*
*  FUNCAO ME0102SALTAPG      -  VERIFICA SALTO DE PAGINA

```

```

*
FUNCTION ME0102SALTAPG()
XLIN = XLIN + 1
IF XLIN < XMAXLIN
  IF TIPOLIN <> "NORMAL"
    IF (XLIN + 4) < XMAXLIN
      TIPOLIN = "NORMAL"
      RETURN .F.
    ENDIF
  ELSE
    TIPOLIN = "NORMAL"
    RETURN .F.
  ENDIF
ENDIF
TIPOLIN = "NORMAL"
XPG = XPG + 1
IF CONF = "V" .AND. XPG > 1
  XLIN = XLIN - 1
  @ XLIN,79 SAY ""
  INKEY(0)
  CLEAR
ENDIF
@ 0,00 SAY "Experimentos de Fisher                                PG:" +
STRZERO(XPG,3,0)
@ 1,00 SAY "-----"
@ 2,00 SAY "Experimento: " + CAM01 + " - " + SUBSTR(CAM02,1,55)
IF NATUREZA = " Original "
  @ 3,00 SAY "[" + NATUREZA + "]"
ELSEIF NATUREZA = " Estimado "
  @ 3,00 SAY "[ Valores Estimados ]"
ELSEIF NATUREZA = "Consistido"
  @ 3,00 SAY "[ Falhas Desprezadas ]"
ENDIF
@ 3,27 SAY "Grupos: " + ALLTRIM(STR(CAM04)) + " / Momentos: " +
ALLTRIM(STR(CAM03))
@ 3,65 SAY DIA + "-" + HORA
@ 4,00 SAY "-----"
XLIN = 5
RETURN .T.
*
*   FUNCAO ME0202MATRIZ      -   LISTA MATRIZ ORIGINAL, MEDIA E
COVARIANCIA
*
FUNCTION ME0202MATRIZ()
XMAPLE = XLINMAPLE + CHR(13) + CHR(10) + XMAPLE
X = CAM03 * CAM04
Y = CAM03 * CAM03 * CAM04

```

```

DECLARE MEDCOL[X],COVAR[Y]
XMEDCOL = 0
XCOVAR = 0
FOR J = I TO FINAL
  XCOLFIM = AT(CHR(13)+CHR(10),XMAPLE)
  XLINMAPLE = SUBSTR(XMAPLE,1,XCOLFIM-1)
  XMAPLE = SUBSTR(XMAPLE,XCOLFIM+2,LEN(XMAPLE))
  IF LASTKEY() = 27 .OR. LEN(XLINMAPLE) < 1 .OR. XLINMAPLE = "MatrizSpool"
    EXIT
  ENDIF
  LINHA = XLINMAPLE
  IF LINHA = "MediaColunas"
    TRATO = "MediaColunasX"
  ELSEIF LINHA = "MatrizCovariancia"
    TRATO = "MatrizCovarianciaX"
  ELSE
    IF TRATO = "MediaColunasX"
      XMEDCOL = XMEDCOL + 1
      MEDCOL[XMEDCOL] = LINHA
    ELSEIF TRATO = "MatrizCovarianciaX"
      XCOVAR = XCOVAR + 1
      COVAR[XCOVAR] = LINHA
    ENDIF
  ENDIF
NEXT
IF NATUREZA = " Original "
  SELECT MEMOVITO
  ZZARQ = "MM"
ELSEIF NATUREZA = " Estimado "
  SELECT MEESTIMA
  ZZARQ = "ME"
ELSE
  SELECT MEFALHA
  ZZARQ = "MF"
ENDIF
Y = CAM01
GOTO TOP
SET SOFTSEEK ON
SEEK Y
SET SOFTSEEK OFF
XGRUPO = 0
XNUMGP = 0
XLIN = 100
DO WHILE .NOT. EOF()
  IF NATUREZA = " Original "
    ZZCODPROJ = MMCODPROJ
    ZZGRUPO = MMGRUPO

```

```

    ZZELEMENTO = MMELEMENTO
ELSEIF NATUREZA = " Estimado "
    ZZCODPROJ = MECODPROJ
    ZZGRUPO  = MEGRUPO
    ZZELEMENTO = MEELEMENTO
    XV = ZZCODPROJ + STR(ZZGRUPO,3) + STR(ZZELEMENTO,3)
    SELECT MEMOVITO
    GOTO TOP
    SEEK XV
    SELECT MEESTIMA
ELSE
    ZZCODPROJ = MFCODPROJ
    ZZGRUPO  = MFGRUPO
    ZZELEMENTO = MFELEMENTO
ENDIF
IF CAM01 <> ZZCODPROJ
    GO BOTTOM
    SKIP
    LOOP
ELSE
    IF XGRUPO <> ZZGRUPO
        ME0202FIMGP()
        XGRUPO = ZZGRUPO
        XNUMGP = XNUMGP + 1
        IDGRUPO[XNUMGP] = XGRUPO
    ENDIF
    CAM10 = STRZERO(ZZELEMENTO,3)
    FOR I = 1 TO CAM03
        XVAR = ZZARQ + "VAL" + STRZERO(I,2,0)
        CAM11 = I
        CAM20 = &XVAR
        IF CAM11 = 1
            ME0102SALTAPG()
            @ XLIN,02 SAY "UE=" + CAM10
            @ XLIN,08 SAY CAM20 PICT "9,999,999.999999"
        ELSE
            IF (CAM11 % 4) = 1
                ME0102SALTAPG()
            ENDIF
            XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 8
            @ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
        ENDIF
    ENDIF
    IF NATUREZA = " Estimado "
        SELECT MEMOVITO
        IF .NOT. EOF()
            XVAR = "MMVAL" + STRZERO(I,2,0)
            CAM20 = &XVAR

```

```

        IF CAM20 = 0
            IF CONF = "V"
                ?? "*"
            ELSE
                @ PROW(),PCOL() SAY "*"
            ENDIF
        ENDIF

    ENDIF
        SELECT MEESTIMA
    ENDIF
    NEXT
    SKIP
    ENDIF
ENDDO
ME0202FIMGP()
RETURN .T.
*
*  LISTA DADOS DOS GRUPOS
*
FUNCTION ME0202FIMGP
ME0102SALTAPG()
IF NATUREZA = " Original "
    ZZGRUPO = MMGRUPO
ELSEIF NATUREZA = " Estimado "
    ZZGRUPO = MEGRUPO
ELSE
    ZZGRUPO = MFGRUPO
ENDIF
IF XGRUPO = 0
    @ XLIN,00 SAY "[ Unidades Experimentais do GRUPO " +
ALLTRIM(STR(ZZGRUPO,3)) + " ]"
    RETURN .T.
ENDIF
ME0102SALTAPG()
@ XLIN,00 SAY "[ Vetor de Media do GRUPO " + ALLTRIM(STR(XGRUPO,3)) + " ]"
FOR I = 1 TO CAM03
    ***** XMEDCOL = (XGRUPO - 1) * CAM03 + I
    XMEDCOL = (XNUMGP - 1) * CAM03 + I
    LINHA = MEDCOL[XMEDCOL]
    CAM10 = ALLTRIM(SUBSTR(LINHA,1,4))
    CAM11 = VAL(ALLTRIM(SUBSTR(LINHA,5,4)))
    CAM20 = ME0202REAL(SUBSTR(LINHA,9,LEN(LINHA)))
    IF CAM11 = 1
        ME0102SALTAPG()
        @ XLIN,02 SAY "MEDIA >"
        @ XLIN,09 SAY CAM20 PICT "9,999,999.999999"
    
```

```

ELSE
  IF (CAM11 % 4) = 1
    ME0102SALTAPG()
  ENDIF
XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 9
  @ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
ENDIF
NEXT
XLIN = XLIN + 1
ME0102SALTAPG()
@ XLIN,00 SAY "[ Matriz de Variancia do GRUPO " + ALLTRIM(STR(XGRUPO,3)) + "
]"
Z = CAM03 * CAM03
FOR I = 1 TO Z
  ***** XCOVAR = (XGRUPO - 1) * (CAM03 * CAM03) + I
  XCOVAR = (XNUMGP - 1) * (CAM03 * CAM03) + I
  LINHA = COVAR[XCOVAR]
  CAM10 = ALLTRIM(SUBSTR(LINHA,5,4))
  CAM11 = VAL(ALLTRIM(SUBSTR(LINHA,9,4)))
  CAM20 = ME0202REAL(SUBSTR(LINHA,13,LEN(LINHA)))
  IF CAM11 = 1
    ME0102SALTAPG()
    @ XLIN,02 SAY "L=" + CAM10
    @ XLIN,09 SAY CAM20 PICT "9,999,999.999999"
  ELSE
    IF (CAM11 % 4) = 1
      ME0102SALTAPG()
    ENDIF
    XCOL = ( ( CAM11 - 1 ) % 4 ) * 18 + 9
    @ XLIN,XCOL SAY CAM20 PICT "9,999,999.999999"
  ENDIF
NEXT
IF .NOT. EOF()
  XLIN = XLIN + 1
  ME0102SALTAPG()
  @ XLIN,00 SAY "[ Elementos do GRUPO " + ALLTRIM(STR(ZZGRUPO,3)) + " ]"
ENDIF
RETURN .T.

```

```

*****
*
* PROGRAMA : ME020101 - Gera Base Para o Maple V.3
*
*****
*
IF .NOT. ABRIR("/BAS/MOV/EST/FAL/",.F.)
    RETURN
ENDIF
DIRETORIO = MEMOREAD("DIRETORI.DIR")
DIRETORIO = ALLTRIM(DIRETORIO)
CPO = 1
XLIN = 0
DECLARE TELM[10],MAT[500]
DO WHILE CPO <> 0
    DO CASE
    CASE CPO = 1
        CAM01 = SPACE(08)    && CODIGO DO EXPERIMENTO
        CAM02 = SPACE(40)    && TITULO DA EXPERIMENTO
        CAM03 = 0            && NUMERO DE MOMENTOS
        CAM04 = 0            && NUMERO DE GRUPOS
        CAM05 = 0            && NIVEL DE SIGNIFICANCIA
        NIVCONF = 0
        CONF = " "
        UNLOCK ALL
        ME0201TELA()
        @ 17,12 GET CAM01 PICT "!!!!!!!"    VALID ME0201CA01(CAM01)
        @ 21,31 GET CONF PICT "!" VALID CONF="S" .OR. CONF="N" .OR. LASTKEY() =
5
        READ
        IF LASTKEY() = 27
            CPO = 0
            LOOP
        ELSEIF CONF = "S"
            *
            ORIGINAL = .T.    && INDICADOR P/ LISTAR ARQUIVO ORIGINAL
            CONSISTIDO = .F.  && INDICADOR P/ LISTAR ARQUIVO CONSISTIDO
            ESTIMADO = .F.    && INDICADOR P/ LISTAR ARQUIVO ESTIMADO
            *
            DO ME020102
            *
            ME0201GERA()
            SET DEVICE TO PRINT
            SET PRINTER TO PRN
            SET DEVICE TO SCREEN
            MENSAGEM("ARQUIVO GERADO COM SUCESSO. TECLE <Enter> ")

```

```
ENDIF  
CONF = " "  
CPO = 1  
ENDCASE  
ENDDO  
UNLOCK ALL  
CLOSE ALL  
RETURN  
*  
* FUNCAO GERA ARQUIVO PARA O MAPLE V  
*  
FUNCTION ME0201GERA  
* <<<<<<<<<<<<<<<<<< PREPARA TELM E A MATRIZ >>>>>>>>>>>  
MOSTRA("EXPORTANDO ARQUIVOS PARA O MAPLE... AGUARDE...")  
SET DEVICE TO PRINT  
XARQUIVO = DIRETORIO + "\BIN\" + CAM01 + ".MS"  
SET PRINTER TO &XARQUIVO  
MAPLE = MEMOREAD("PROJETO.MS")  
FINAL = MLCOUNT(MAPLE,80)  
XLIN = 0  
PRIMEIRO = .T.  
IF ORIGINAL  
    ME0201ORIG()  
ENDIF  
IF CONSISTIDO  
    ME0201CONSIG()  
ENDIF  
IF ESTIMADO  
    ME0201EST()  
ENDIF  
XLIN = XLIN + 1  
@ XLIN,0 SAY "{END}"  
RETURN .T.  
*  
* PREPARA EXPORTACAO DO ARQUIVO ORIGINAL  
*  
FUNCTION ME0201ORIG  
Y = CAM01  
SELECT MEMOVITO  
GOTO TOP  
SET SOFTSEEK ON  
SEEK Y  
SET SOFTSEEK OFF  
XGRUPO = MMGRUPO  
XREGTOT = 0  
XREGGP = 0  
XNUMGP = 1
```

```

XJ = 2
MAT[01]= CHR(34) + "A:=matrix("
MAT[02]=" , " + STRZERO(CAM03,2,0) + ",["
XT = 2
TELM[01]= CHR(34) + " Telm:=matrix(1,"
TELM[02]=" ,["
DO WHILE .NOT. EOF()
  IF CAM01 <> MMCODPROJ
    GO BOTTOM
    SKIP
    LOOP
  ELSE
    FOR I = 1 TO CAM03
      XVAR = "MMVAL" + STRZERO(I,2,0)
      XTEMP = &XVAR
      MAT[XJ] = MAT[XJ] + ALLTRIM(STR(XTEMP,13,6)) + ","
      IF LEN(MAT[XJ]) > 60
        MAT[XJ] = SUBSTR(MAT[XJ],2,LEN(MAT[XJ]))
        XJ = XJ + 1
        MAT[XJ] = " "
      ENDIF
    NEXT
    IF XGRUPO <> MMGRUPO
      XGRUPO = MMGRUPO
      TELM[XT] = TELM[XT] + STRZERO(XREGGP,3,0) + ","
      XREGGP = 0
      XNUMGP = XNUMGP + 1
      IF LEN(TELM[XT]) > 60
        TELM[XT] = SUBSTR(TELM[XT],2,LEN(MAT[XT]))
        XT = XT + 1
        TELM[XT] = " "
      ENDIF
    ENDIF
    XREGTOT = XREGTOT + 1
    XREGGP = XREGGP + 1
    SKIP
  ENDIF
ENDDO
IF LEN(MAT[XJ]) = 1
  XJ = XJ -1
ENDIF
MAT[XJ] = SUBSTR(MAT[XJ],2,LEN(MAT[XJ])-2) + "):" + CHR(34)
MAT[01] = CHR(34) + "A:=matrix(" + STRZERO(XREGTOT,5,0)
XTAMMAT = 0
FOR KK = 1 TO XJ
  XTAMMAT = XTAMMAT + LEN(MAT[KK])
NEXT

```

```

MAT[01] = STRZERO((XTAMMAT-2),5,0) + " " + MAT[01]
TELM[XT] = SUBSTR(TELM[XT],2,LEN(MAT[XT])) + STRZERO(XREGGP,3,0)
TELM[XT] = TELM[XT] + "]:" + CHR(34)
TELM[01] = CHR(34) + "Telm:=matrix(1," + STRZERO(XNUMGP,3,0)
XTAMMAT = 0
FOR KK = 1 TO XT
    XTAMMAT = XTAMMAT + LEN(TELM[KK])
NEXT
TELM[01] = STRZERO((XTAMMAT-2),4,0) + " " + TELM[01]
SUFIXO = "O"
ME0201MAPLE()
RETURN .T.
*
*   PREPARA EXPORTACAO DO ARQUIVO CONSISTIDO
*
FUNCTION ME0201CONSIS
Y = CAM01
SELECT MEFALHA
GOTO TOP
SET SOFTSEEK ON
SEEK Y
SET SOFTSEEK OFF
XGRUPO = MFGRUPO
XREGTOT = 0
XREGGP = 0
XNUMGP = 1
XJ = 2
MAT[01]= CHR(34) + "A:=matrix("
MAT[02]=" ," + STRZERO(CAM03,2,0) + ",["
XT = 2
TELM[01]= CHR(34) + " Telm:=matrix(1,"
TELM[02]=" ,["
DO WHILE .NOT. EOF()
IF CAM01 <> MFCODPROJ
    GO BOTTOM
    SKIP
    LOOP
ELSE
    FOR I = 1 TO CAM03
        XVAR = "MFVAL" + STRZERO(I,2,0)
        XTEMP = &XVAR
        MAT[XJ] = MAT[XJ] + ALLTRIM(STR(XTEMP,13,6)) + ","
        IF LEN(MAT[XJ]) > 60
            MAT[XJ] = SUBSTR(MAT[XJ],2,LEN(MAT[XJ]))
            XJ = XJ + 1
            MAT[XJ] = " "
        ENDIF
    
```

```

NEXT
IF XGRUPO <> MFGRUPO
  XGRUPO = MFGRUPO
  TELM[XT] = TELM[XT] + STRZERO(XREGGP,3,0) + ","
  XREGGP = 0
  XNUMGP = XNUMGP + 1
  IF LEN(TELM[XT]) > 60
    TELM[XT] = SUBSTR(TELM[XT],2,LEN(MAT[XT]))
    XT = XT + 1
    TELM[XT] = " "
  ENDIF
ENDIF
ENDIF
XREGTOT = XREGTOT + 1
XREGGP = XREGGP + 1
SKIP
ENDIF
ENDDO
IF LEN(MAT[XJ]) = 1
  XJ = XJ -1
ENDIF
MAT[XJ] = SUBSTR(MAT[XJ],2,LEN(MAT[XJ])-2) + "):" + CHR(34)
MAT[01] = CHR(34) + "A:=matrix(" + STRZERO(XREGTOT,5,0)
XTAMMAT = 0
FOR KK = 1 TO XJ
  XTAMMAT = XTAMMAT + LEN(MAT[KK])
NEXT
MAT[01] = STRZERO((XTAMMAT-2),5,0) + " " + MAT[01]
TELM[XT] = SUBSTR(TELM[XT],2,LEN(MAT[XT])) + STRZERO(XREGGP,3,0)
TELM[XT] = TELM[XT] + "):" + CHR(34)
TELM[01] = CHR(34) + "Telm:=matrix(1," + STRZERO(XNUMGP,3,0)
XTAMMAT = 0
FOR KK = 1 TO XT
  XTAMMAT = XTAMMAT + LEN(TELM[KK])
NEXT
TELM[01] = STRZERO((XTAMMAT-2),4,0) + " " + TELM[01]
SUFIXO = "C"
ME0201MAPLE()
RETURN .T.
*
*   PREPARA EXPORTACAO DO ARQUIVO ESTIMADO
*
FUNCTION ME0201EST
Y = CAM01
SELECT MEESTIMA
GOTO TOP
SET SOFTSEEK ON
SEEK Y

```

```

SET SOFTSEEK OFF
XGRUPO = MEGRUPO
XREGTOT = 0
XREGGP = 0
XNUMGP = 1
XJ = 2
MAT[01]= CHR(34) + "A:=matrix("
MAT[02]=" , " + STRZERO(CAM03,2,0) + ",["
XT = 2
TELM[01]= CHR(34) + " Telm:=matrix(1,"
TELM[02]=" ,["
DO WHILE .NOT. EOF()
  IF CAM01 <> MECODPROJ
    GO BOTTOM
    SKIP
    LOOP
  ELSE
    FOR I = 1 TO CAM03
      XVAR = "MEVAL" + STRZERO(I,2,0)
      XTEMP = &XVAR
      MAT[XJ] = MAT[XJ] + ALLTRIM(STR(XTEMP,13,6)) + ","
      IF LEN(MAT[XJ]) > 60
        MAT[XJ] = SUBSTR(MAT[XJ],2,LEN(MAT[XJ]))
        XJ = XJ + 1
        MAT[XJ] = " "
      ENDIF
    NEXT
    IF XGRUPO <> MEGRUPO
      XGRUPO = MEGRUPO
      TELM[XT] = TELM[XT] + STRZERO(XREGGP,3,0) + ","
      XREGGP = 0
    XNUMGP = XNUMGP + 1
    IF LEN(TELM[XT]) > 60
      TELM[XT] = SUBSTR(TELM[XT],2,LEN(TELM[XT]))
      XT = XT + 1
      TELM[XT] = " "
    ENDIF
  ENDIF
  XREGTOT = XREGTOT + 1
  XREGGP = XREGGP + 1
  SKIP
ENDIF
ENDDO
IF LEN(MAT[XJ]) = 1
  XJ = XJ -1
ENDIF
MAT[XJ] = SUBSTR(MAT[XJ],2,LEN(MAT[XJ])-2) + "]:" + CHR(34)

```

```
MAT[01] = CHR(34) + "A:=matrix(" + STRZERO(XREGTOT,5,0)
XTAMMAT = 0
FOR KK = 1 TO XJ
    XTAMMAT = XTAMMAT + LEN(MAT[KK])
NEXT
MAT[01] = STRZERO((XTAMMAT-2),5,0) + " " + MAT[01]
TELM[XI] = SUBSTR(TELM[XI],2,LEN(MAT[XI])) + STRZERO(XREGGP,3,0)
TELM[XI] = TELM[XI] + "):" + CHR(34)
TELM[01] = CHR(34) + "Telm:=matrix(1," + STRZERO(XNUMGP,3,0)
XTAMMAT = 0
FOR KK = 1 TO XI
    XTAMMAT = XTAMMAT + LEN(TELM[KK])
NEXT
TELM[01] = STRZERO((XTAMMAT-2),4,0) + " " + TELM[01]
SUFIXO = "E"
ME0201MAPLE()
RETURN .T.
*
* <<<<<<<<<<<<<<<<<<<<<< GRAVA O ARQUIVO MAPLE >>>>>>>>>>>>>>>>
*
FUNCTION ME0201MAPLE
XMAPLE = MAPLE
J = 0
FOR I = 1 TO FINAL
    XCOLFIM = AT(CHR(13)+CHR(10),XMAPLE)
    XLINMAPLE = SUBSTR(XMAPLE,1,XCOLFIM-1)
    XMAPLE = SUBSTR(XMAPLE,XCOLFIM+2,LEN(XMAPLE))
    IF LASTKEY() = 27
        EXIT
    ENDIF
IF I = 18
    IF SUBSTR(XLINMAPLE,4,9) <> "tcond:=3:"
        SET DEVICE TO SCREEN
        @ 10,00 SAY XLINMAPLE
        MENSAGEM("ARQUIVO BASE PARA MAPLE FOI ALTERADO. PROCESSO
ABORTADO")
        RETURN .F.
    ENDIF
    XLIN = XLIN + 1
    @ XLIN,0 SAY "10 " + CHR(34) + "tcond:=" + STRZERO(CAM03,2,0) + ":" +
CHR(34)
ELSEIF I = 21
    IF SUBSTR(XLINMAPLE,4,7) <> "tgp:=2:"
        SET DEVICE TO SCREEN
        @ 10,00 SAY XLINMAPLE
        MENSAGEM("ARQUIVO BASE PARA MAPLE FOI ALTERADO. PROCESSO
ABORTADO")
```

```

    RETURN .F.
ENDIF
XLIN = XLIN + 1
@ XLIN,0 SAY "9 " + CHR(34) + "tgp:=" + STRZERO(XNUMGP,3,0) + ":" + CHR(34)
ELSEIF I = 23
    IF SUBSTR(XLINMAPLE,5,26) <> "Telm := matrix(1,2,[5,5]):"
        SET DEVICE TO SCREEN
        @ 10,00 SAY XLINMAPLE
        MENSAGEM("ARQUIVO BASE PARA MAPLE FOI ALTERADO. PROCESSO
ABORTADO")
    RETURN .F.
ENDIF
FOR KK = 1 TO XT
    XLIN = XLIN + 1
    @ XLIN,0 SAY TELM[KK]
NEXT
ELSEIF I = 26
    IF SUBSTR(XLINMAPLE,5,20) <> "Ret_Clip := RETORNO:"
        SET DEVICE TO SCREEN
        @ 10,00 SAY XLINMAPLE
        MENSAGEM("ARQUIVO BASE PARA MAPLE FOI ALTERADO. PROCESSO
ABORTADO")
    RETURN .F.
ENDIF
XLIN = XLIN + 1
@ XLIN,0 SAY "21 " + CHR(34) + "Ret_Clip := " +
SUBSTR(ALLTRIM(CAM01)+SUFIXO+"      ",1,8) + ":" + CHR(34)
ELSEIF I = 29
    IF SUBSTR(XLINMAPLE,5,17) <> "Niv_Conf := 0.95:"
        SET DEVICE TO SCREEN
        @ 10,00 SAY XLINMAPLE
        MENSAGEM("ARQUIVO BASE PARA MAPLE FOI ALTERADO. PROCESSO
ABORTADO")
    RETURN .F.
ENDIF
XLIN = XLIN + 1
@ XLIN,0 SAY "17 " + CHR(34) + "Niv_Conf := " + STRZERO(NIVCONF,4,2) + ":" +
CHR(34)
ELSEIF I = 32
    IF SUBSTR(XLINMAPLE,6,11) <> "A:=matrix(1"
        SET DEVICE TO SCREEN
        @ 10,00 SAY XLINMAPLE
        MENSAGEM("ARQUIVO BASE PARA MAPLE FOI ALTERADO. PROCESSO
ABORTADO")
    RETURN .F.
ENDIF

```

```

    FOR KK = 33 TO 36 && P/ DESPRESAR DE MAIS LINHAS DA MATRIZ NO
    ARQ.ORIG.
        XCOLFIM = AT(CHR(13)+CHR(10),XMAPLE)
        XLINMAPLE = SUBSTR(XMAPLE,1,XCOLFIM-1)
    XMAPLE = SUBSTR(XMAPLE,XCOLFIM+2,LEN(XMAPLE))
    NEXT
    FOR KK = 1 TO XJ
        XLIN = XLIN + 1
        @ XLIN,0 SAY MAT[KK]
    NEXT
    ELSEIF I < 7
        IF PRIMEIRO
            IF I = 6
                IF .NOT. ORIGINAL
                    TOTCOMAND = VAL(SUBSTR(XLINMAPLE,12,5)) * 2
                    XLINMAPLE = "{SCP_R 0 0 " + STRZERO(TOTCOMAND,5)
                ENDIF
                PRIMEIRO = .F.
            ENDIF
            XLIN = XLIN + 1
            @ XLIN,0 SAY XLINMAPLE
        ENDIF
    ELSE
        IF XLINMAPLE <> "{END}"
            XLIN = XLIN + 1
            @ XLIN,0 SAY XLINMAPLE
        ENDIF
    ENDIF
NEXT
RETURN .T.
*#####
*  FUNCOES DEFINIDAS PARA CONSISTENCIA DOS CAMPOS SOLOCITADOS  #
*#####
*
*  FUNCAO ME0201CA01 - CAM01 - CODIGO DO EXPERIMENTO
*
FUNCTION ME0201CA01
PARAMETER X
IF X = " " .OR. EMPTY(X)
    MENSAGEM("INFORME O CODIGO DO EXPERIMENTO")
    RETURN .F.
ENDIF
SELECT MEBASE
GOTO TOP
SEEK X
IF EOF()
    MENSAGEM("EXPERIMENTO NAO CADASTRADO")

```

```

    RETURN .F.
ELSE
    CAM02 = MBTITPROJ
    CAM03 = MBNUMEXP
    CAM04 = MBNUMGP
    CAM05 = MBNIVSIG
    NIVCONF = (100 - CAM05) / 100
    @ 17,23 SAY CAM02 PICT "XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX"
    @ 18,24 SAY CAM03 PICT "99"
    @ 18,48 SAY CAM04 PICT "999"
ENDIF
Y = CAM01
SELECT MEMOVITO
GOTO TOP
SET SOFTSEEK ON
SEEK Y
SET SOFTSEEK OFF
IF EOF()
    MENSAGEM("MOMENTOS AINDA NAO INFORMADOS PARA ESSE
EXPERIMENTO")
    RETURN .F.
ENDIF
IF CAM01 <> MMCODPROJ
    MENSAGEM("MOMENTOS AINDA NAO INFORMADOS PARA ESSE
EXPERIMENTO")
    RETURN .F.
ENDIF
RETURN .T.
* FUNCAO ME0201TELA - TELA DO PROGRAMA
FUNCTION ME0201TELA
TELABASE(3)
@ 17,04 SAY "Codigo:      - "
@ 18,04 SAY "Numero de Condicao:      Numero de Grupos: "
@ 21,04 SAY "Confirma Execucao? (S/N) : "
RETURN .T.
* FUNCAO VEARQUIVO
FUNCTION VEARQUIVO
PARAMETER Y
IF FILE(Y)
    @ 23,10 SAY SPACE(65)
    X=" "
    @ 23,10 SAY "ARQUIVO `"+Y+"` JA EXISTE! DELETA? (S/N) : " GET X PICT "!"
VALID X="S" .OR. X="N"
    READ
    IF X <> "S"
        RETURN .F.
    ENDIF

```

```
! DEL &Y  
ENDIF  
RETURN .T.
```