
LCE 602 - Estatística Experimental (aulas práticas)

Escola Superior de Agricultura "Luiz de Queiroz"
Departamento de Ciências Exatas - ESALQ/USP
Março de 2002
(versão 3,14159265)

**Antonio Augusto Franco Garcia
Décio Barbin
Sônia Maria De Stefano Piedade**

LCE 602 - Estatística Experimental

Este texto destina-se aos alunos matriculados na referida disciplina, devendo servir como um guia para as aulas práticas, que são realizadas com auxílio do pacote estatístico SAS (Statistical Analysis System), em sua versão 6.12. Propositalmente, os resultados fornecidos pelo pacote não são apresentados, para que o aluno motive-se a executar e estudar os programas aqui apresentados.

- Recomenda-se que o aluno observe atentamente cada resultado, interpretando-o corretamente.
- A maioria dos programas é discutida nas aulas, e desse modo espera-se que esse material facilite a assimilação do conteúdo das aulas práticas.
- A resolução dos exercícios propostos é fundamental para entendimento dos assuntos aqui abordados. Os conjuntos de dados apresentados estão disponíveis no site do curso na Internet.

Críticas e sugestões, bem como correção de eventuais erros no material, serão bem recebidas.

1 Criação de data sets

A seguir, apresentamos diversos programas que ilustram a criação de data sets seguindo as regras do SAS. Esses arquivos têm extensão .SD2 e são apagados automaticamente no final de cada sessão do SAS, sendo gravados no diretório WORK definido pelo programa. Contudo, isso pode ser alterado pelo usuário, conforme será apresentado, resultado em data sets permanentes. Alguns deles mostram como a entrada incorreta dos dados pode afetar completamente as análises estatísticas a serem realizadas posteriormente. Lembre-se que cada data set criado deve ser visualizado com auxílio da procedure “proc print”, que não está incluída na maioria dos programas.

```
data teste1;
input trat $ bloco PESO;
cards;
A 1 100
A 2 130
B 1 99
B 2 145
;
```

```
data teste2;
input trat $ 1 bloco 3 PESO 5-7;
cards;
A 1 100
A 2 130
B 1 99
B 2 145
;
```

```
data;
input trat $ bloco PESO;
cards;
A 1 100 A 2 130
B 1 99 B 2 145
;
```

```
data;
input trat $ bloco PESO @@;
cards;
A 1 100 A 2 130
B 1 99 B 2 145
;
```

```
data;
do trat='A', 'B';
```

```
input bloco PES0;  
output;  
end;  
cards;  
1 100  
1 99  
2 130  
2 145  
;
```

```
data;  
do trat=1 to 2;  
do bloco=1 to 2;  
input PES0;  
output;  
end;  
end;  
cards;  
100  
130  
99  
145  
;
```

```
data;  
input trat $ bloco PES0;  
cards;  
A23456789 1 100  
A23456789 2 130  
B23456789 1 99  
B23456789 2 145  
;
```

```
data;  
length trat $ 9;  
input trat $ bloco PES0;  
cards;  
A23456789 1 100  
A23456789 2 130  
B23456789 1 99  
B23456789 2 145  
;
```

```
data;  
input trat $ bloco PES0;  
if bloco <=1 then CONJUNTO=100;  
else CONJUNTO=200;  
cards;  
A 1 100
```

```
A 2 130  
B 1 99  
B 2 145  
;
```

```
data;  
input trat $ bloco PES0;  
  if trat='A' then delete;  
cards;  
A 1 100  
A 2 130  
B 1 99  
B 2 145  
;
```

```
input trat $ bloco PES0;  
  if PES0=100 then PES0=333;  
cards;  
A 1 100  
A 2 130  
B 1 99  
B 2 145  
;
```

Atenção

Note que muitos data sets podem levar a resultados inadequados ou inesperados. Dessa forma, nessa parte do programa, em que estamos aprendendo como o SAS interpreta os dados, devemos observar atentamente como essas instruções devem ser escritas.

Exercício 1

1. Quantos data sets foram criados?
2. Quais seus nomes e conteúdos?
3. Indique, dentre os diversos programas apresentados, aqueles em que os data sets criados apresentaram problemas. Mostre como eles podem ser corrigidos.

2 Alterações no Data Step

Nos programas apresentados no item anterior, vimos algumas formas de efetuar operações e/ou alterar os dados no momento da criação dos data sets. Essa parte dos programas,

denominada Data Step, permite que o usuário defina novas variáveis, efetue operações com as já existentes, etc. A seguir, apresentamos alguns programas que exemplificam essas possibilidades. A primeira linha define opções para as saídas: não apresentar data, não numerar as páginas, apresentar 72 caracteres por linha (“linesize”).

```
options nodate nonumber ls=72;
data notas;
input aluno $ 1-14 sexo $ P1 P2 curso $;
media=(P1+2*P2)/3;
if media >=5 then result='aprov';
else result='reprov';
cards;
Joaqui m Si l va    M 4.9 5.1 AG
Pedro Al ei xo     M 3.8 2.5 AG
Ana Mari a        F 7.8 8.3 FL
Hel ena Souza     F 8.4 4.3 AG
Mari o Pi nhei ro M 6.4 7.3 FL
;
```

```
proc print;
title 'Dados para veri fi cação';
run;
```

```
proc sort;
by media;
run;
proc print;
title 'Dados ordenados pel a médi a fi nal';
run;
```

```
proc means;
title 'Médi a das notas de cada prova';
var P1 P2 medi a;
run;
```

Tabela 1. Operadores aritméticos

Operador	Operação
**	Exponenciação
*	Multiplicação
/	Divisão
+	Adição
-	Subtração

Tabela 2. Operadores para comparações

Operador	Significado
=	Igual
^=	Diferente
>	Maior
>=	Maior ou igual
<	Menor
<=	Menor ou igual

Exercício 2

1. Acrescente dois aos alunos ao conjunto já mencionado:
Carlos Alberto M 3.3 3.6 FL
Adriana Souza e Silva F 5.0 4.5 AG
2. Faça alterações no programa de forma que os resultados considerem o fato de que, se a média final for menor que 5.0, porém maior ou igual a 3.0, o aluno pode realizar uma prova de recuperação.

Dica

Procure entender muito bem a estrutura *if-then-else*, útil para programação.

3 Sintaxe geral das procedures do SAS

Algumas procedures (procedimentos) do SAS foram apresentadas nos dois itens anteriores. Todas elas têm a seguinte forma geral:

```
PROC NOME DA PROCEDURE <lista de opções>;  
    COMANDO / opções do comando;  
    COMANDO / opções do comando;
```

Podemos obter a sintaxe das procedures do SAS através dos menus. Por exemplo, utilizamos a procedure MEANS, cuja sintaxe é (usando *Help / SAS System Help: Main Menu / Modeling and Analysis Tools / Data Analysis / MEANS*):

```
PROC MEANS <option-list> <statistic-keyword-list>;
  VAR variable-list;
  BY variable-list;
  CLASS variable-list;
  FREQ variable;
  WEIGHT variable;
  ID variable-list;
  OUTPUT <OUT= SAS-data-set> <output-statistic-list>
    <MINID|MAXID <(var-1<(id-list-1)>
    <... var-n<(id-list-n)>>)>=name-list>;
```

Outras procedures muito usadas:

```
PROC SORT DATA= SAS-data-set
  OUT= SAS-data-set
  NODUPPLICATES|NODUPREC|NODUP
  NODUPKEY
  EQUALS|NOEQUALS
  REVERSE
  FORCE
  ASCII
  EBCDIC
  DANISH|NORWEGIAN
  FINNISH|SWEDISH
  NATIONAL
  SORTSEQ= collating-sequence
  SORTSIZ= memory-specification
  TAGSORT
  Host Options;
  BY <DESCENDING> variable ... ;                                /* required */
```

```
PROC PRINT DATA= SAS-data-set
  DOUBLE
  NOOBS
  UNIFORM
  LABEL
  SPLIT= 'split-character'
  N
  ROUND
  HEADING= direction
  ROWS= page-format
  WIDTH= column-width;
  VAR variable-list;
  ID variable-list;
  BY variable-list;
  PAGEBY BY-variable;
  SUMBY BY-variable;
```

```
SUM variable-list;
```

Exemplo 1 - Apresentamos a seguir dados de altura de plantas de milho obtidas em dois experimentos (ex1 e ex2, ver aula teórica). Calcular, para cada um deles, a variância, o desvio padrão, o erro-padrão da média, o coeficiente de variação, a média (com intervalo de confiança).

```
data teor2;
input ex1 ex2;
datalines;
203 203
208 198
198 199
200 200
202 201
192 202
197 197
;
proc print;
title 'Dados usados na aula teórica';
run;

proc means var std stderr cv mean clm;
title 'Estatísticas de interesse';
run;
```

4 Estatística descritiva

Uma forma de obter diversas estatísticas, incluindo testes para normalidade, gráficos ramo-e-folhas, etc, é com uso da proc univariate, cuja sintaxe é:

```
PROC UNIVARIATE DATA= SASdataset
                NOPRINT
                PLOT
                FREQ
                NORMAL
                PCTLDEF= value
                VARDEF= DF|WEIGHT|WGT|N|WDF
                ROUND= roundoff unit...;
VAR variables;
BY variables;
```

```
FREQ variabl e;  
WEIGHT variabl e;  
ID variabl es;  
OUTPUT OUT= SASdataset keyword= names...;
```

Exemplo 2 - Dados de pesos de colmos de uma certa variedade de cana-de-açúcar. Calcular a média e a variância, construir o gráfico ramo-e-folhas e um histograma, testar se os dados têm distribuição normal (dados disponíveis no arquivo EXEMP2.TXT). Interpretar os demais resultados.

```
data prati ca;  
input PESO@@;  
cards;  
16.89 11.84 18.69 16.51 18.97  
15.21 17.90 17.38 12.56 13.73  
14.60 18.94 16.72 14.41 13.23  
16.65 15.73 16.86 17.87 15.45  
18.19 16.04 15.72 16.68 17.76  
12.03 14.00 14.74 16.68 12.39  
15.30 15.49 11.53 19.13 15.01  
13.47 16.83 15.47 11.63 16.22  
12.14 17.69 12.54 18.73 17.02  
17.03 16.44 13.58 10.46 16.89  
;
```

```
proc univariate normal plot;  
title 'Estatística descrtiva';  
run;
```

```
proc chart;  
title 'Histograma';  
vbar PESO;  
run;
```

Nota:

A procedure chart permite a construção de gráficos de baixa resolução. No caso, utilizou-se o comando vbar (barras verticais), sendo também possível o emprego do comando hbar (barras horizontais). A sintaxe da PROC CHART é:

```
PROC CHART DATA= SAS-data-set  
          LPI = p  
          FORMCHAR (index list) = 'formchar-string';  
BY <variabl e-list>;
```

```
VBAR variable-list </ <option-list>>;
HVAR variable-list </ <option-list>>;
BLOCK variable-list </ <option-list>>;
PIE variable-list </ <option-list>>;
STAR variable-list </ <option-list>>;
```

Exercício 3

1. Alterar (com programação) um dos dados do peso de cana-de-açúcar para 5.0 kg, simulando um valor discrepante (ou “outlier”).
2. Fazer uma análise exploratória desses dados, interpretando os resultados.

5 Verificando as pressuposições do modelo

Apresentaremos a seguir programas que ilustram a possibilidade de verificar três das quatro pressuposições assumidas para realização da análise de variância (questões 3, 5 e 6 abaixo). O teste de não-aditividade do modelo será apresentado no capítulo 9.

Exemplo 3 - Um pesquisador pretende comparar 4 variedades de pêssego quanto ao enraizamento de estacas (ver aula teórica), instalado segundo o delineamento inteiramente ao acaso (arquivo EXEMP3.TXT).

1. Fazer uma análise exploratória dos dados de cada tratamento separadamente.
2. Construir um gráfico da dispersão dos dados para cada tratamento.
3. Verificar se há homogeneidade de variâncias.
4. Verificar se há presença de valores discrepantes através da análise de resíduos.
5. Testar se os resíduos têm distribuição normal.
6. Verificar se os resíduos são independentes.
7. Testar a hipótese de igualdade das médias dos tratamentos.

```
options nodate nonumber ps=60 ls=72;
data exemp3;
input trat ESTACAS;
cards;
1 2
1 2
1 1
1 1
1 0
2 1
2 0
```

```
2 0
2 1
2 1
3 12
3 10
3 14
3 17
3 11
4 7
4 9
4 15
4 8
4 10
;
```

```
proc print;
title 'Dados para verificação';
run;
```

```
proc sort;
  by trat;
run;
proc print;
title 'Dados ordenados por tratamento';
run;
```

```
title 'Análise exploratória com box-plot - questão 1';
proc univariate normal plot;
  by trat;
  var ESTACAS;
run;
title 'Dispersão dos dados para cada tratamento - questão 2';
proc plot;
  plot ESTACAS*trat;
run;
```

```
title 'Verificação de homog. de variâncias (Hartley) - questão 3';
proc means mean var CV;
  by trat;
  var ESTACAS;
run;
```

```
title 'ANOVA sem transformação, para obtenção dos res. padronizados';
proc glm;
  class trat;
  model ESTACAS = trat;
  output out=RESIDUOS PREDICTED=PREDTOS RESIDUAL=RESIDUOS
  STUDENT=RES_PAD;
```

```

run;

_____

title 'Exame dos resíduos - valores discrepantes';
proc print data=RESIDUOS;
run;
proc univariate data=RESIDUOS normal plot;
title 'Valores discrepantes, teste de normalidade';
var RESIDUOS;
run;

_____

title 'Gráfico dos resíduos padronizados - independência';
proc plot;
plot RES_PAD*PREDICTOS;
run;

_____

title 'Análise de variância sem transf. (iguald. trat.) - questão4';
proc anova;
class trat;
model ESTACAS = trat;
means trat; /* solicita o cálculo das médias por tratamento */
run;

_____

```

Comentários:

1. O data set EXEMP3.SD2 foi criado sem usar recursos de programação. Evidentemente, o leitor pode usar outras formas que lhe sejam convenientes, como por exemplo:

```

_____

data exemp3;
do trat=1 to 4;
input ESTACAS @;
output;
end;
cards;
2 1 12 7
2 0 10 9
1 0 14 15
1 1 17 8
0 1 11 10
;

```

```

_____

proc print;
title 'Outra forma de criar o data set';
run;

_____

```

É claro que os programas apresentados continuam válidos.

2. O comando “by” requer que os dados estejam ordenados para a variável que será colocada após esse comando. Caso isso não tenha ocorrido, pode-se antes aplicar a PROC SORT para conseguir tal ordenação.
3. O comando OUTPUT OUT=RESIDUOS, incluído na proc glm, solicitou a criação de um data set com o nome RESIDUOS. Esse data set contém apenas as variáveis solicitadas: valores preditos (PREDICTED=PREDICTOS), resíduos associados a cada observação (RESIDUAL=RESIDUOS) e resíduos padronizados (STUDENT=RES_PAD). Os nomes colocados após o sinal de igual são arbitrários, tomados de forma a facilitar sua identificação.
4. Note que a PROC PRINT, a PROC MEANS e a PROC UNIVARIATE da última parte do programa continham o nome das variáveis para as quais esses procedimentos devem ser executados (após o comando VAR).

Algumas procedures usadas:

```
PROC ANOVA DATA= SAS-data-set
      MANOVA
      MULTIPASS
      OUTSTAT= SAS-data-set;
CLASS variable;                                /* required */
MODEL dependents=effects / options;           /* required */
ABSORB variable;
BY variable-list;
FREQ variable;
MANOVA H= effects E= effect M= equations...
      MNames= names PREFIX= name / options;
MEANS effects / options;
REPEATED factorname levels(level values)
      transformation<,...> / options;
TEST H= effects E= effect;
```

```
PROC GLM options ;
CLASS variable-list;
MODEL dependents= independents / options;     /* required */
ABSORB variable-list;
BY variable-list;          FREQ variable;
ID variable-list;          WEIGHT variable;
CONTRAST 'label' effect values... / options;
ESTIMATE 'label' effect values... / options;
LSMEANS effects / options;
MANOVA H= effects E= effect M= equations...
      MNames= names PREFIX= name / options;
MEANS effects / options;
OUTPUT OUT= SAS-data-set keywords= names... ;
RANDOM effects / options;
REPEATED factorname levels (level values)
```

```
transformation<...> / options;
TEST H= effects E= effect / options; .
```

Exercício 4

Um pesquisador deseja comparar 12 novos meios de cultura quanto ao desenvolvimento do fungo *Metharizium anisopliae*, usado no controle biológico da cigarrinha das pastagens. Para tanto, instalou um experimento segundo o delineamento inteiramente casualizado, com 4 repetições, e mediu o crescimento do fungo através do diâmetro das colônias. Baseando-se nos dados da Tabela 1 (arquivo EXERC4.TXT), pede-se:

1. Fazer uma análise exploratória dos dados para cada tratamento.
2. Testar a hipótese de igualdade das médias dos tratamentos com um nível de significância de 5% (análise de variância), sem transformação de dados.
3. Refazer o item anterior manualmente, indicando os cálculos envolvidos.
4. Verificar a pressuposição de homogeneidade de variâncias, através do teste do F máximo e do gráfico dos resíduos padronizados x valores estimados.
5. Verificar a pressuposição de normalidade dos resíduos, através de análises gráficas e do teste de Shapiro-Wilk (proc uni vari ate).

Tabela 3. Diâmetro das colônias.

Tratamentos	I	II	III	IV	Tratamentos	I	II	III	IV
Meio 1	0,31	0,45	0,46	0,43	Meio 7	0,44	0,35	0,31	0,40
Meio 2	0,82	1,10	0,88	0,72	Meio 8	0,56	1,02	0,71	0,38
Meio 3	0,43	0,45	0,63	0,76	Meio 9	0,22	0,21	0,18	0,23
Meio 4	0,45	0,71	0,66	0,62	Meio 10	0,30	0,37	0,38	0,29
Meio 5	0,29	0,40	0,23	0,36	Meio 11	0,23	0,25	0,24	0,22
Meio 6	0,92	0,61	0,49	1,24	Meio 12	0,30	0,36	0,31	0,33

Fonte: Modificado de Barbosa, J.C. 1984

6 Transformação de dados

Antes de apresentarmos exemplos sobre transformação de dados, vejamos como os comandos SET e MERGE podem auxiliar para realizarmos alterações em data sets existentes.

```
data exp1;
input local $ var $ alt_plt;
datalines;
```

```
Pi rac M 185
Pi rac N 196
Pi rac O 200
Pi rac P 189
;

data exp2;
input local $ var $ alt_plt;
datalines;
Araras P 180
Araras O 176
Araras N 177
Araras M 182
;

data tudo1;
set exp1 exp2;
run;
proc print data=tudo1;
title 'Comando SET';
run;

data exp3;
input var $ alt_plt;
datalines;
M 185
N 196
O 200
P 189
;
data exp4;
input var $ alt_esp;
datalines;
P 80
O 76
N 77
M 82
;

proc sort data=exp3;
by var;
run;
proc sort data=exp4;
by var;
run;

data tudo2;
merge exp3 exp4;
```



```
by var;
pos_rel =al t_esp/al t_pl t;
run;
```

```
proc print;
title 'Comando MERGE';
run;
```

Atenção:

1. Verifique atentamente a possibilidade de inserir instruções em data sets já criados, como foi feito na obtenção da variável pos_rel, criada a partir dos data sets exp3 e exp4.
2. Foi necessário ordenar os data sets exp3 e exp4 com o uso da proc sort, antes de executar o comando MERGE.

Veremos agora como alterar um data set criado com uso do comando INFILE, que permite a leitura de arquivos com extensão .PRN, gravados em algum diretório de interesse. Arquivos com essa extensão são do tipo “texto formatado (separado por espaços)”. O arquivo usado é denominado PRATICA4.PRN (disponível no site) e refere-se ao exemplo 3. Evidentemente, o diretório deve ser alterado para a situação específica. A forma de obter qual transformação pode ser aplicada será apresentado futuramente.

```
data aul a4;
infile 'O:\AUGUSTO\PRATI CA4. PRN' ;
input trat $ ESTACAS;
run;
```

A transformação desejada pode ser obtida da seguinte forma:

```
title 'Análise com transformação';
data aul a5;
set aul a4;
EST_TRAN=SQRT(ESTACAS+0.5);
run;
```

```
title 'Dados transformados - raiz quadrada (SQRT)';
proc print;
run;
```

```
title 'Análise de variância dos dados transformados';
proc glm data=aul a5;
class trat;
model EST_TRAN=trat;
```

```
output out=RES_TRAN PREDICTED=PREDTOS RESIDUAL=RESTRANS  
STUDENT=RESTOPAD;  
run;
```

```
title 'Resíduos dos dados transformados';  
proc print data=RES_TRAN;  
run;  
proc plot;  
plot RESTPAD*PREDTOS;  
run;  
proc univariate normal plot;  
var RESTPAD;  
run;
```

Nota:

Algumas funções do SAS muito úteis para efetuar transformação de dados são (x=argumento): SQRT(X) (raiz quadrada), LOG(X) (logaritmo neperiano), LOG10(X), ARSIN(X) (arco-seno).

Transformação estabilizadora da variância

Realizada com base na regressão linear $\log(\sigma^2) = \log \beta + \alpha \log \mu$, ou seja, entre o logaritmo da variância e o logaritmo da média. O valor do α (coeficiente angular) determinará a transformação a ser realizada (Tabela 4).

:

Tabela 4. Transformações estabilizadoras da variância

Relação entre a variância (σ_Y^2) e a média (μ)	α	$\lambda = 1 - \frac{\alpha}{2}$	Transformação $y^* = y^\lambda$	Observação
$\sigma_Y^2 \propto \text{constante}$	0	1	nenhuma	
$\sigma_Y^2 \propto \mu^1$	1	1/2	$y^* = y^{1/2} = \sqrt{y}$	Dados de contagem (Poisson)
$\sigma_Y^2 \propto \mu^2$	2	0	$y^* = \log y$	
$\sigma_Y^2 \propto \mu^3$	3	-1/2	$y^* = y^{-1/2} = \frac{1}{\sqrt{y}}$	
$\sigma_Y^2 \propto \mu^4$	4	-1	$y^* = y^{-1} = \frac{1}{y}$	
-	-	-	$y^* = \arcsen \sqrt{y}$	Dados com dist. binomial expressos como frações ($0 \leq y \leq 1$)

Fonte: Montgomery, D.C. (1991) Design and analysis of experiments

Exemplo 4 - Aplicando essa metodologia para estudar a transformação dos dados de enraizamento de estacas (exemplo 3), fazendo inicialmente o teste do F máximo, teremos:

```
options nodate nonumber ps=60 ls=72;
data exemp4;
input trat $ ESTACAS;
cards;
A 2
A 2
A 1
A 1
A 0
B 1
B 0
B 0
B 1
B 1
C 12
C 10
C 14
C 17
C 11
D 7
D 9
D 15
D 8
D 10
;
```

```
title 'Análise de variância e resíduos';
proc glm;
  class trat;
  model ESTACAS = trat;
  output out=resíduos PREDICTED=PREDITOS RESIDUAL=RESIDUOS
  STUDENT=RES_PAD;
run;

title 'Exame dos resíduos';
proc print data=resíduos;
run;

title 'Verificação da normalidade (Shapiro-Wilk), análise gráfica';
proc univariate normal plot;
  var RES_PAD;
run;

title 'Teste do F máximo - Hartley';
proc means noprint data=resíduos;
  var ESTACAS;
  by trat;
  output out=var_med MEAN=MEDIA VAR=VARIANC;
run;
data fmax;
set var_med;
run;
proc means noprint data=fmax;
  var VARIANC;
  output out=razao MIN=VMIN MAX=VMAX;
run;
data hartley;
set razao;
  FMAX=VMAX/VMIN;
run;
title 'Valor do F máximo';
proc print data=hartley;
  var VMIN VMAX FMAX;
run;

title 'Análise gráfica dos resíduos - independência';
proc plot data=resíduos;
  plot RES_PAD*PREDITOS;
run;

title 'Seleção empírica da transf. estabilizadora da variância';
data transf;
set var_med;
```

```

LMEDIA=LOG(MEDIA);
LVAR=LOG(VARIANC);
run;
title 'Obtenção dos valores de alfa e beta usando regressão';
proc reg data=transf;
  model LVAR=LMEDIA;
run;

_____

title 'Gráfico LOG(média) vs LOG(variância)';
proc plot;
  plot LVAR*LMEDIA;
run;

_____

title 'Anava com transformação';
data anava;
set exemp4;
  ESTAC_T=(ESTACAS)**(0.5);
run;
proc glm data=anava;
  class trat;
  model ESTAC_T=trat;
  output out=RESTRAN PREDICTED=PRED_T STUDENT=RES_T;
run;

_____

title 'Análise de resíduos dos dados transformados';
proc univariate normal plot data=RESTRAN;
  var RES_T;
run;
proc plot data=RESTRAN;
  plot RES_T*PRED_T;
run;

_____

```

Notas

1. A forma usada para cálculo do F máximo é mais trabalhosa que a apresentada no capítulo 5, mas permite a obtenção da razão diretamente, embora também seja necessário o uso da tabela.
2. Foram criados diversos data sets, com os seguintes nomes e conteúdos:
 - EXEMP4: dados originais
 - RESIDUOS: valores preditos, resíduos e resíduos padronizados
 - VAR_MED: médias e variâncias por tratamento
 - FMAX: o próprio data set var_med
 - HARTLEY: cálculo do F máximo, maior e menor variância
 - TRANSF: logaritmos da médias e das variâncias

- ANAVA: dados transformados
 - RESTRAN: valores preditos e resíduos padronizados dos dados transformados
3. Utilizou-se a PROC REG, cuja sintaxe é

```
PROC REG options;  
label: MODEL dependents= regressors / <options>;  
BY variable-list;  
FREQ variable;  
ID variable;  
VAR variable-list;  
ADD variable-list;  
DELETE variable-list;  
REWEIGHT <condition|ALLOBS> </options> | <STATUS|UNDO>;  
WEIGHT variable;  
label: MTEST <equation1, ... equationk / options>;  
OUTPUT OUT= SAS-data-set keyword= names ...;  
PAINT <condition|ALLOBS> </options> | <STATUS|UNDO>;  
PLOT <yvariable1*xvariable1> <=symbol 1>, ...  
      <yvariablek*xvariablek> <=symbol k> </options>;  
PRINT <options ANOVA MODELDATA>;  
REFIT;  
RESTRICT equation1, ... equationk;  
label: TEST equation1, ... equationk / option;
```

Exercício 5

1. Obtenha a transformação estabilizadora da variância adequada para os dados do exercício 4.
2. Refaça a análise de variância usando essa transformação.
3. Analise os resíduos dos dados transformados. As pressuposições de homogeneidade de variâncias, independência e distribuição normal dos resíduos estão satisfeitas?

7 Testes de comparações múltiplas

Ainda considerando o mesmo exemplo:

```
options nodate nonumber ps=60 ls=72;  
data exemp4b;  
input trat $ ESTACAS@@;  
ESTAC_T=(ESTACAS)**(0.5);  
cards;  
A 2 A 2 A 1 A 1 A 0
```

```
B 1 B 0 B 0 B 1 B 1
C 12 C 10 C 14 C 17 C 11
D 7 D 9 D 15 D 8 D 10
;
```

```
proc glm;
title 'Análise de variância dos dados transformados, com testes';
class trat;
model ESTAC_T=trat;
means trat / tukey;
means trat / tukey alpha=0.01;
means trat / duncan;
means trat / duncan alpha=0.01;
contrast 'A_B vs C_D' trat 1 1 -1 -1;
contrast 'A vs B' trat 1 -1 0 0;
contrast 'C vs D' trat 0 0 1 -1;
run;
```

Nota:

1. Note que usamos o mesmo data set, criado aqui de outra forma.
2. O SAS possui diversas opções que podem ser incluídas após o comando means (no caso, usamos Tukey e Duncan, com níveis $\alpha=0.05$ - default - e 0.01 , além dos contrastes):

ALPHA=	BON	CLDIFF	CLM	DEPNLY
DUNCAN	DUNNETT	DUNNETTL	DUNNETTU	E=
ETYP=	GABRIEL	HOVTEST=	HTYPE=	KRATIO=
LINE	NOSORT	REGWF	REGWQ	SCHEFFE
SIDAK	SMM GT2	SNK	T LSD	TUKEY
WALLER	WELCH			

8 Ensaios inteiramente ao acaso com número diferente de repetições por tratamento

Seguindo o que foi apresentado no texto básico desse curso, consideraremos que houve perda de 3 parcelas. Devemos codificar esses valores no data set através de pontos (.), um para cada valor perdido. O data step pode ser:

```
options nodate nonumber ps=60 ls=72;
data parcer;
input trat $ ESTACAS;
```

```
data lines;
```

```
A          2  
A          2  
A          1  
A          .  
A          0  
B          1  
B          0  
B          0  
B          1  
B          1  
C          12  
C          10  
C          14  
C          .  
C          .  
D          7  
D          9  
D          15  
D          8  
D          10  
;
```

Podemos proceder a análise normalmente, como feito anteriormente. Aplicaremos (pelas razões já discutidas) a transformação $\sqrt{y+0,5}$. A análise dos resíduos pode ser feita normalmente, da mesma maneira. Vale lembrar que essa notação para observações perdidas também pode ser usada para eliminar dados discrepantes, detectados pela análise dos resíduos.

```
title 'Análise de variância e resíduos';  
proc glm;  
  class trat;  
  model ESTACAS = trat;  
  output out=resíduos PREDICTED=PREDICTOS RESIDUAL=RESIDUOS  
         STUDENT=RES_PAD;  
run;
```

```
title 'Exame dos resíduos';  
proc print data=resíduos;  
run;
```

```
title 'Verificação da normalidade (Shapiro-Wilk), análise gráfica';  
proc univariate normal plot;  
  var RES_PAD;  
run;
```

```
title 'Análise gráfica dos resíduos';
```



```

proc plot;
  plot RES_PAD*PREDI TOS;
run;

-----

title 'Teste do F máximo - Hartley';
proc means noprint data=resi duos;
  var ESTACAS;
  by trat;
  output out=var_med MEAN=MEDI A VAR=VARI ANC;
run;
proc print data=var_med;
run;
data fmax;
set var_med;
run;
proc means noprint data=fmax;
  var VARI ANC;
  output out=razao MI N=VMI N MAX=VMAX;
run;
data hartley;
set razao;
  FMAX=VMAX/VMI N;
run;
title 'Valor do F máximo';
proc print data=hartley;
  var VMI N VMAX FMAX;
run;

-----

title 'Seleção empírica da transf. estabilizadora da variância';
data transf;
set var_med;
  LMEDI A=LOG(MEDI A);
  LVAR=LOG(VARI ANC);
run;
title 'Obtenção do valor de alfa';
proc reg data=transf;
  model LVAR=LMEDI A;
run;

-----

title 'Gráfico LOG(medi a) vs LOG(vari ânci a)';
proc plot;
  plot LVAR*LMEDI A;
run;

-----

title 'Anava com transformação';
data anava;
set pratica7;
  ESTAC_T=(ESTACAS+0.5)**(0.5);
run;

```

```
proc glm data=anava;
  class trat;
  model ESTAC_T=trat;
  output out=RESTRAN PREDICTED=PRED_T STUDENT=RES_T;
run;

_____

title 'Análise de resíduos dos dados transformados';
proc univariate normal plot data=RESTRAN;
  var RES_T;
run;
proc plot data=RESTRAN;
  plot RES_T*PRED_T;
run;

_____

title 'Testes de Tukey - desbalanceado';
proc glm data=anava;
  class trat;
  model ESTAC_T=trat;
  means trat / tukey;
  means trat / tukey lines; /*Opções para Tukey, escolher uma*/
  lsmeans trat / adjust=tukey PDIF=all;
run;

_____
```

Notas:

1. A última linha de comando permite a obtenção de médias de mínimos quadrados para o efeito 'trat'. É possível associá-la à opção ADJUST, em que são feitos ajustes para os níveis de significância das comparações múltiplas. PDIF solicita a impressão dos níveis de significância (p-values).
2. Quando há número diferente de repetições por tratamento, o SAS usa a média harmônica do número de repetições de cada tratamento para determinar a diferença mínima significativa, se for usada a opção LINES (ver final do programa).

Exercício 6

Um ensaio visando verificar a micropropagação do capim limão sob diferentes níveis de BAP foi conduzido, segundo o delineamento inteiramente casualizado (dados no arquivo EXERC6.SAS).

Tabela 5. Número de brotos por perfilho, em meios com diferentes níveis do regulador de crescimento BAP, 30 dias após a introdução do perfilho no meio.

Níveis de BAP	Número de brotos				
I: 0,5 mg	3	4	1	2	4
II: 1,0 mg	4	-	2	7	6
III: 2,0 mg	2	2	3	2	7
IV: 3,0 mg	9	4	2	5	10

Fonte: Adaptado de Latado, R.R. (1999) - CENA/USP

Pede-se:

1. Fazer a análise de variância para a variável número de brotos. Apresentar o quadro da forma usual.
2. Testar se os resíduos têm distribuição normal.
3. Representar os resíduos num gráfico de caixa (box-plot) e num diagrama ramo-e-folhas.
4. Verificar se há dados discrepantes, pela análise dos resíduos.
5. Fazer um gráfico dos valores preditos vs resíduos padronizados. Verificar se há indícios da necessidade de transformação dos dados.
6. Testar se há homogeneidade de variâncias pelo teste do F máximo.
7. Fazer uma regressão linear entre o logaritmo das médias e o logaritmo das variâncias. Há relação linear entre elas? Em caso afirmativo, qual o valor de λ deve ser usado para aplicar a transformação estabilizadora da variância? Qual transformação é sugerida?
8. Refazer a análise de variância (com teste de Tukey) e dos resíduos usando a transformação proposta (se necessário). Verificar se ela foi efetiva.

9 Ensaio em blocos casualizados

Veremos a seguir um exemplo de análise de um experimento instalado no delineamento blocos ao acaso. Os programas permitem também a verificação de algumas pressuposições da análise de variância.

Exemplo 5 - Um pesquisador deseja estudar o comportamento de nove porta-enxertos para a laranjeira Valência: Tangerina Sunki (1), Limão rugoso nacional (2), Limão rugoso da Flórida (3), Tangerina Cleópatra (4), Citrange-troyer (5), Trifoliata (6), Tangerina cravo (7), Laranja caipira (8) e Limão cravo (9). Foram medidos dados de produção aos 12 anos de idade (detalhes na aula teórica). O data set pode ser criado da seguinte maneira (arquivo EXEMP5.SAS):

```
options nodate nonumber ps=60 ls=78;
proc format;
  value trat 1='Tangerina sunki'
            2='Limão rugoso nacional'
            3='Limão rugoso da Flórida'
            4='Tangerina Cleópatra'
            5='Citrange-troyer'
            6='Trifoliata'
            7='Tangerina cravo'
            8='Laranja caiçara'
            9='Limão cravo';
run;

data exemp5;
do trat=1 to 9;
  do bloco=1 to 3;
    input PRODUCAO@;
    output;
  end;
end;
format trat trat.;
datalines;
145 155 166    200 190 190    183 186 208
190 175 186    180 160 156    130 160 130
206 165 170    250 271 230    164 190 193
;

proc print;
title1 'Dados para verificação';
title2 '(uso da proc format)';
run;
```

Nota:

Aproveitamos o exemplo para introduzir o uso da PROC FORMAT, que pode ser muito útil para a identificação dos tratamentos, dentre outras aplicações. Sua sintaxe é:

```
PROC FORMAT CNTLIN= SAS-data-set
            CNTLOUT= SAS-data-set
            FMTLIB
            LIBRARY= libref<. catalog-name>
            MAXLABEL= n
            MAXSELEN= n
            NOREPLACE
            NOTSORTED
```

```

PAGE;
VALUE name <(format-option-list)> range='formatted-value-1'
<... range-n='formatted-value-n'>;
PICTURE name <(format-option-list)> range-1='picture-1'
<(picture-option-list)> <... range-n='picture-n'
<(picture-option-list-n)>;
INVALUE name <(informat-option-list)> <'>range1<'>
=i nformatted-
value-1 <... <'>range-n<'>=i nformatted-value-n>;
SELECT entry_list;
EXCLUDE entry_list;

```

A seguir, da mesma forma que para exemplos anteriores, verificaremos se há homogeneidade de variâncias pelo teste do F máximo de Hartley. Para o exemplo em questão, o valor tabelado com $k = 9$ e $v = 2$ (5%) é $F = 475$.

```

title1 'Teste do F máximo - Hartley';
title2;
proc glm noprint;
class trat bloco;
model PRODUCAO = trat bloco;
output out=RESIDUOS residual=residual student=res_pad
predicted=predictos;
run;
proc means noprint data=RESIDUOS;
var residual;
by trat;
output out=VARIANC var=varian;
run;
proc means noprint data=VARIANC;
var varian;
output out=RAZAO MIN=VMIN MAX=VMAX;
run;
data hartley;
set razao;
FMAX=VMAX/VMIN;
run;
proc print;
run;

```

Note que foi usada a opção NOPRINT na PROC GLM e na PROC MEANS, com o objetivo de evitar saídas desnecessárias nesse ponto da análise.

O teste de Shapiro-Wilk para a normalidade dos resíduos pode ser feito da seguinte maneira:

```

title 'Teste de normalidade dos resíduos';
proc univariate normal plot data=residuals;

```

```
var resíduo;  
run;
```

A independência dos resíduos pode ser verificada graficamente com o seguinte programa:

```
title 'Independência dos resíduos';  
proc plot data=resíduos;  
plot res_pad*preditos;  
run;
```

Teste de não-aditividade de Tukey (1 gl)

Agora, apresentaremos um programa que permite fazer o teste de não-aditividade do modelo de Tukey, com 1 grau de liberdade.

```
title 'Aditividade do modelo';  
proc means data=exemp5 nway noprint;  
class trat;  
var PRODUCAO;  
output out=trat_m mean=medtrat;  
run;
```

```
proc means data=exemp5 nway noprint;  
class bloco;  
var PRODUCAO;  
output out=bloco_m mean=medbloco;  
run;
```

```
proc means data=exemp5 nway noprint;  
var PRODUCAO;  
output out=geral_m mean=medgeral;  
run;
```

```
proc sort data=exemp5;  
by trat;  
run;  
data temp1;  
merge exemp5 trat_m;  
by trat;  
run;
```

```
proc sort data=exemp5;  
by bloco;
```

```
run;
data temp2;
merge exemp5 bloco_m;
  by bloco;
run;

-----

data temp3;
set geral_m;
do trat=1 to 9; /*CUI DADO*/
  do bloco=1 to 3; /*CUI DADO*/
    output;
  end;
end;
run;

-----

proc sort data=temp1;
  by trat bloco;
proc sort data=temp2;
  by trat bloco;
proc sort data=temp3;
  by trat bloco;
run;
data tudo;
merge temp1 temp2 temp3;
  by trat bloco;
keep trat bloco PRODUCAO medtrat medbloco medgeral;
run;

-----

data tudo2;
set tudo;
tratest=medtrat-medgeral;
blocoest=medbloco-medgeral;
keep trat bloco PRODUCAO tratest blocoest;
run;
proc print;
title1 'Valores para Tukey 1 gl não-adi ti vi dade';
run;

-----

proc glm data=tudo2;
title1 'Análise de variância';
title2 'com Tukey 1 gl para não-adi ti vi dade';
title3 'OBS: considerar apenas tratest*blocoest';
class trat bloco;
model PRODUCAO = trat bloco tratest*blocoest/SS3;
output out=tukey_p predicted=tukeypre;
run;
proc plot;
title1;
title2;
```

```
title3;  
plot tukeypre*tratest/vpos=40;  
run;
```

Finalmente, após verificar que as pressuposições da análise foram atendidas, podemos realizar a análise de variância com o teste de Tukey.

```
title 'Análise de variância - teste de Tukey';  
title2;  
title3;  
proc glm data=exemp5;  
  class trat bloco;  
  model PRODUCAO=trat bloco;  
  means trat/tukey;  
run;
```

Atenção

1. Percebe-se que o modelo matemático considerado foi $y_{ij} = m + t_i + b_j + e_{ij}$, diferente dos exemplos referentes ao delineamento inteiramente ao acaso. Isso faz que os comandos CLASS e MODEL, na PROC GLM, sejam acrescidos desse novo efeito considerado no modelo. É muito importante que isso seja observado, uma vez que o SAS realizaria a análise (erroneamente) mesmo que indicássemos incorretamente o modelo. Desse modo, é muito importante que verifiquemos atentamente qual foi o delineamento empregado, para indicar ao SAS o modelo adequado.
2. A aditividade do modelo foi testada, o que não havia sido feito nos exemplos anteriores. A elaboração do programa está além dos objetivos desse curso, mas ele pode ser usado em situações reais, apenas alterando o número de tratamentos e blocos.

Exercício 7

Um experimento de competição de variedades de cana-de-açúcar foi conduzido em 1999, com a finalidade de verificar qual(is) delas pode(m) ser recomendadas para plantio (arquivo: EXERC7.XLS). Para tanto, foi conduzido um experimento em blocos ao acaso. Pede-se:

1. Verificar se os resíduos têm distribuição normal. Representá-los num diagrama ramo-e-folhas e num gráfico de caixas.
2. Verificar graficamente se há independência dos resíduos.
3. Testar se há homogeneidade de variâncias pelo teste do F máximo.
4. Fazer o teste de não-aditividade do modelo.

5. Observar os resíduos padronizados e verifique se há algum(ns) valores discrepantes. Em caso afirmativo, eliminá-lo(s) e refazer os itens 1 a 4.
6. Há necessidade de realizar transformação dos dados? Em caso afirmativo, qual seria a transformação indicada?
7. Fazer a análise de variância, transformando os dados se necessário. Recomendar a(s) variedade(s) mais produtivas para plantio, com base no teste de Tukey.

Tabela 6. Produção de genótipos de cana-de-açúcar em ton/ha num ensaio em blocos ao acaso.

Genótipo	bloco I	bloco II	bloco III	bloco IV
RB72454	147,857	141,429	141,429	143,571
RB835054	138,571	134,286	137,143	130,714
RB835486	121,429	128,571	127,857	114,286
RB845210	150,000	128,571	140,000	138,571
RB845257	159,286	149,286	156,429	141,429
RB855002	127,143	137,857	137,143	127,143
RB855036	172,143	157,857	161,429	145,714
RB855113	161,429	150,714	150,714	156,429
RB855453	136,429	132,143	134,286	125,000
RB855536	148,571	142,857	137,857	142,857
RB855546	159,286	144,286	158,571	156,429
RB855584	160,714	162,143	165,714	160,714
RB865230	177,143	172,143	176,429	169,286

Fonte: Programa de Melhoramento Genético da Cana-de-açúcar (PMGCA), UFSCar, Araras. 2000

10 Ensaio em blocos casualizados com parcelas perdidas

Antes de iniciarmos a análise desse tipo de experimento, veremos como criar data sets permanentes. Vale lembrar que, até o momento, criamos data sets que são automaticamente apagados pelo SAS sempre que uma sessão se encerra. Dessa forma, caso desejemos mantê-los, devemos torná-los permanentes. Essa abordagem é muito comum em publicações sobre o SAS, sendo também muito usada por usuários avançados para a troca de arquivos. A obtenção desses arquivos é possível com o uso do comando LIBNAME, seguido de um nome que identificará o diretório que conterá o data set permanente, e depois do caminho indicando onde o arquivo será gravado. Por exemplo,

```
libname local 'O:\augusto\';
```

cria a libname com nome local, especificando o diretório O:\augusto\, no qual gravamos diversos arquivos úteis nesse curso. É evidente que o leitor pode alterar esse diretório da forma que lhe for mais conveniente, como por exemplo, 'C:\Meus Documentos\Estatística Experimental\Programas no SAS\ Datasets permanentes\'. Agora, basta especificar os data sets que serão permanentes, incluindo o nome da libname antes do nome do data set:

```
data local.pratica9;
```

Este comando criará um data set de nome pratica9, gravado no diretório O:\AUGUSTO\ e que não será apagado no final da sessão. O nome do data set permanente é definido após o ponto, sendo a parte anterior (local, no caso) o nome da libname. Os data sets permanentes têm extensão .SD2 e só podem ser lidos no próprio SAS (não são arquivos do tipo texto). Uma vez que esses arquivos estão prontamente disponíveis, podemos iniciar a próxima sessão do SAS com o seguinte procedimento (sem ser necessário criar novamente os data sets):

```
libname local 'O:\augusto\';  
proc print data=local.pratica9;  
title 'Dados para verificação';  
run;
```

Notas:

1. Esse tipo de arquivo permanente é muito útil quando a criação do data set correto é muito laboriosa, sendo possível retomar a análise em sessões posteriores com os arquivos já criados.
2. Lembre-se que os usuários da rede do Departamento de Ciências Exatas não têm permissão para gravar arquivos no drive O:\AUGUSTO\ e, portanto, os programas aqui apresentados devem ser adaptados à situação específica.

Voltemos agora ao exemplo 5. Imaginemos que a parcela do tratamento 7 no bloco 2 foi perdida. O data set (permanente) pode ser criado da seguinte maneira:

```
options nodate nonumber ps=62 ls=78;  
libname local 'R:\';  
proc format;  
  value trat 1=' (1) Tangerina sunki '  
             2=' (2) Limão rugoso nacional '  
             3=' (3) Limão rugoso da Flórida '  
             4=' (4) Tangerina Cleópatra '
```

```

5=' (5) Citrange-troyer'
6=' (6) Trifoliata'
7=' (7) Tangerina cravo'
8=' (8) Laranja caiçara'
9=' (9) Limão cravo';

run;

data local.pratic9a;
do trat=1 to 9;
  do bloco=1 to 3;
    input PRODUCAO@;
    output;
  end;
end;
format trat trat.;
datalines;
145 155 166    200 190 190    183 186 208
190 175 186    180 160 156    130 160 130
206 .    170    250 271 230    164 190 193
;

proc print data=local.pratic9a;
title 'Dados para verificação';
run;

```

Conforme discutido em aula, em função dessa perda de parcelas, a ordem de entrada dos efeitos do modelo altera os resultados, já que as SQ's do tipo I e III são diferentes. A análise de variância pode ser feita da seguinte maneira:

```

title1 'Análise de variância - teste de Tukey';
title2 '(atenção para a ordem dos efeitos no modelo)';
proc glm data=local.pratic9a;
  class trat bloco;
  model PRODUCAO=trat bloco;
  means trat/tukey;
  means trat/tukey lines; /*escolha uma das opções para means*/
run;

title1 'Análise de variância - teste de Tukey';
title2 '(atenção para a ordem dos efeitos no modelo)';
proc glm data=local.pratic9a;
  class trat bloco;
  model PRODUCAO=bloco trat;
  means trat/tukey lines;
run;

```

Atenção:

Compare as saídas com as análises feitas na aula teórica. São muito comuns erros nesse tipo de análise. Foram empregadas aqui duas ordens para os efeitos no modelo

Exercício 8

1. Apresente o quadro da maneira usual (aula teórica) para os programas acima, selecionando os resultados adequados.
2. Examine os graus de liberdade fornecidos pelo SAS. Eles estão levando em conta a perda de uma parcela? Justifique.
3. Como a ordem dos efeitos no modelo afeta as análises de variância feitas pelo SAS?

Com base no exercício 8, conclui-se que a análise de variância correta pode ser feita da seguinte maneira:

```
title1 'Análise de variância, Tukey (med. ajust.)';
title2;
proc glm data=local.pratic9a;
  class trat bloco;
  model PRODUCAO=bloco trat;
  lsmeans trat/adjust=tukey PDIFF=all;
run;
```

Na aula teórica foi apresentada a fórmula para cálculo da parcela perdida, de forma a minimizar a SQresíduo: $y = 193$. Substituindo esse valor no data set, teremos

```
title1 'Análise de variância com valor da parc. perd.';
data local.parcper1;
set local.pratic9a;
  if PRODUCAO=. then PRODUCAO=193;
run;
proc print;
run;
title2 '(cuidado: SQtrat não está corrigida)';
proc glm data=local.parcper1;
  class trat bloco;
  model PRODUCAO=bloco trat;
run;
```

Cuidado

A soma de quadrados de tratamentos não está corrigida, e também não foi subtraído um grau de liberdade do resíduo. Essa forma de análise, substituindo a parcela perdida, não é

necessária quando usamos o SAS para as análises, sendo aqui apresentada apenas com finalidade didática.

Para o caso de duas parcelas perdidas (ver aula teórica):

```
data local.pratic9b;
do trat=1 to 9;
  do bloco=1 to 3;
    input PRODUCAO@;
    output;
  end;
end;
format trat trat.;
datalines;
145 155 166    200 190 190    183 186 208
.    175 186    180 160 156    130 160 130
206 .    170    250 271 230    164 190 193
;

proc print data=local.pratic9b;
title 'Dados para verificação';
run;

title 'Análise de variância, Tukey (med. ajust.)';
proc glm data=local.pratic9b;
  class trat bloco;
  model PRODUCAO=bloco trat;
  lsmeans trat/adjust=tukey PDIFF=all;
run;
```

Exercício 9

Os resultados apresentados a seguir (Tabela 5) referem-se a um experimento em blocos ao acaso para comparar progênies de milho quanto à produtividade de grãos ($\text{kg}/10 \text{ m}^2$) (arquivo: EXERC9.SD2, data set permanente). Pede-se:

1. Testar se os resíduos têm distribuição normal
2. Represente os resíduos num diagrama ramo-e-folhas
3. Represente os resíduos num gráfico de caixas.
4. Testar se há homogeneidade de variâncias pelo teste do F máximo.
5. Observe os resíduos padronizados e verifique se há algum(ns) valor(es) discrepante(s).
6. Caso houver valor(es) discrepante(s), eliminá-lo usando programação no Data Step, e refazer os itens 1, 2 e 3.

7. Qual a importância em fazer essas análises propostas nos itens 1 a 5? Justifique.
8. Há necessidade de realizar transformação dos dados? Por que?
9. Caso haja essa necessidade, qual seria a transformação a ser realizada?
10. Que tipo de soma de quadrados (I ou III) é mais indicada para analisar estatisticamente os dados desse experimento? Justifique.
11. Fazer a análise de variância, transformando os dados se necessário. Comparar as progênies com base no teste de Tukey.
12. Compare as progênies também com o teste de Duncan.

Tabela 7. Produtividade de grãos de milho em kg/10 m².

Progenie	I	II	III	Progenie	I	II	III
1	16,1	17,1	15,8	11	17,1	-	15,0
2	17,6	17,0	14,5	12	28,9	23,4	22,4
3	20,5	-	16,2	13	22,5	20,0	20,4
4	22,5	16,5	15,4	14	-	16,2	15,2
5	18,5	19,8	16,0	15	20,7	18,9	19,3
6	18,8	18,4	17,6	16	16,3	17,4	15,8
7	17,4	20,5	-	17	18,8	16,9	15,1
8	23,7	20,1	19,1	18	20,3	15,5	-
9	19,8	17,1	16,8	19	22,1	20,8	16,2
10	21,9	15,5	15,7	20	20,4	23,0	18,6

Fonte: Adaptado de Vencovsky e Barriga, 1992

11 Ensaio em quadrados latinos

Os ensaios em quadrados latinos levam em conta o controle local, aplicado em dois sentidos. Como exemplo, vejamos o experimento apresentado na aula teórica.

Exemplo 6 - Um ensaio para comparação de 5 variedades de cana-de-açúcar, A, B, C, D e E, foi conduzido em um quadrado latino. Foi avaliada a produção, em kg/parcela. Os valores obtidos podem ser usados para a criação do arquivo da seguinte forma (arquivo EXEMP6.SAS):

```
options nodate nonumber ps=65 ls=78;
data latino;
input linha coluna trat $ PROD;
datalines;
1      1      D      432
1      2      A      518
1      3      B      458
```

1	4	C	583
1	5	E	331
2	1	C	724
2	2	E	478
2	3	A	524
2	4	B	550
2	5	D	400
3	1	E	489
3	2	B	384
3	3	C	556
3	4	D	297
3	5	A	420
4	1	B	494
4	2	D	500
4	3	E	313
4	4	A	486
4	5	C	501
5	1	A	515
5	2	C	660
5	3	D	438
5	4	E	394
5	5	B	318

```

;
proc print;
title 'Dados para verificação';
run;

```

A análise de algumas pressuposições da análise de variância é feita de forma análoga ao que já foi apresentado.

```

title 'Análise de variância e resíduos';
proc glm noprint;
class trat;
model PROD = linha coluna trat;
output out=resíduos PREDICTED=PREDTOS RESIDUAL=RESIDUOS
STUDENT=RES_PAD;
run;

```

```

title 'Exame dos resíduos';
proc print data=resíduos;
run;

```

```

title 'Verificação da normalidade (teste Shapiro-Wilk), análise
gráfica';
proc univariate normal plot;
var RES_PAD;
run;

```

```
title 'Análise gráfica dos resíduos';  
proc plot;  
  plot RES_PAD*PREDICTOS;  
run;  
  
title 'Teste do F máximo - Hartley';  
proc sort;  
  by trat;  
run;  
proc means noprint data=resíduos;  
  var PROD;  
  by trat;  
  output out=var_med MEAN=MEDIA VAR=VARIANC;  
run;  
proc print data=var_med;  
run;  
data fmax;  
set var_med;  
run;  
proc means noprint data=fmax;  
  var VARIANC;  
  output out=razao MIN=VMIN MAX=VMAX;  
run;  
data hartley;  
set razao;  
  FMAX=VMAX/VMIN;  
run;  
title 'Valor do F máximo';  
proc print data=hartley;  
  var VMIN VMAX FMAX;  
run;
```

A análise de variância (com Tukey) pode ser feita da seguinte forma:

```
proc glm data=latino;  
title 'Análise de variância';  
class linha coluna trat;  
model PROD = linha coluna trat;  
means trat/tukey;  
run;
```


Nota:

Veja como é simples analisar experimentos usando a PROC GLM. Porém, deve-se estar atento quanto ao modelo matemático que deve ser empregado.

No caso: $y_{ijk} = m + I_i + C_j + t_{k(i,j)} + e_{ijk}$.

Exercício 10

Os dados abaixo referem-se a um experimento para avaliação de variedades de trigo, plantadas num quadrado latino (arquivo EXERC10.XLS).

Pede-se:

1. Fazer a análise de variância.
2. Indicar a(s) melhor(es) variedade(s), usando o teste de Duncan.
3. Caso utilizássemos outro critério para as comparações múltiplas, como por exemplo o de Tukey, as conclusões seriam diferentes? Qual o motivo?
4. Cite vantagens e desvantagens desse delineamento experimental.

Tabela 8. Produção de trigo de 4 variedades (kg/parcela).

Linhas	Colunas			
	1	2	3	4
1	C= 10,5	D= 7,7	B= 12,0	A= 13,2
2	B= 11,1	A= 12,0	C= 10,3	D= 7,5
3	D= 5,8	C= 12,2	A= 11,2	B= 13,7
4	A= 11,6	B= 12,3	D= 5,9	C= 10,2

Fonte: Steel e Torrie, 1960

12 Ensaio fatorial

Neste tipo de ensaio, são avaliadas combinações entre os níveis de diferentes fatores. Se todas as combinações forem consideradas, temos os fatoriais completos.

Essas combinações são então avaliadas segundo delineamentos adequados, como por exemplo, blocos casualizados. Dessa forma, o modelo matemático empregado deve ser adequado à situação em questão, devendo ser cuidadosamente definido pelo pesquisador.

Por exemplo, no caso de termos dois fatores (A e B) com dois níveis cada, teremos $2 \times 2 = 4$ tratamentos, que, se forem avaliados segundo delineamento inteiramente ao acaso, implicarão no seguinte modelo matemático: $y_{ijk} = m + a_i + b_j + ab_{ij} + e_{ijk}$. Por outro lado, se

o delineamento for blocos ao acaso, o modelo será $y_{ijk} = m + a_i + b_j + ab_{ij} + r_k + e_{ijk}$ (em que r_k é o efeito de blocos, não incluído no modelo anterior).

Exemplo 7 - Um experimento foi instalado com o objetivo de estudar a influência do aparelho e do operador dele na medição de alturas de árvores de *Eucaliptus saligna* aos 7 anos de idade. Foram avaliados 5 diferentes aparelhos, por 4 operadores (fatorial 5 x 4), em 10 árvores (cada árvore foi considerada como um bloco) (maiores detalhes: ver aula teórica).

No SAS, a criação do data set fica (arquivo EXEMP7.SAS):

```
options nodate nonumber ps=65 ls=78;
data fatorial;
input trat apar oper bloco ALTURA@@;
datalines;
1 1 1 1 22.40 1 1 1 2 20.85
1 1 1 3 23.60 1 1 1 4 21.00
1 1 1 5 19.10 1 1 1 6 19.80
1 1 1 7 16.55 1 1 1 8 14.75
1 1 1 9 21.10 1 1 1 10 14.30
2 2 1 1 22.90 2 2 1 2 21.40
2 2 1 3 23.95 2 2 1 4 22.25
2 2 1 5 21.40 2 2 1 6 21.00
2 2 1 7 16.90 2 2 1 8 14.85
2 2 1 9 22.00 2 2 1 10 15.00
3 3 1 1 23.50 3 3 1 2 21.00
3 3 1 3 23.75 3 3 1 4 20.75
3 3 1 5 19.50 3 3 1 6 19.50
3 3 1 7 17.50 3 3 1 8 14.50
3 3 1 9 20.00 3 3 1 10 14.00
4 4 1 1 22.50 4 4 1 2 20.50
4 4 1 3 23.20 4 4 1 4 21.00
4 4 1 5 21.00 4 4 1 6 18.90
4 4 1 7 17.80 4 4 1 8 14.30
4 4 1 9 20.60 4 4 1 10 14.20
5 5 1 1 21.45 5 5 1 2 19.20
5 5 1 3 23.35 5 5 1 4 20.35
5 5 1 5 19.95 5 5 1 6 19.35
5 5 1 7 17.45 5 5 1 8 14.45
5 5 1 9 22.00 5 5 1 10 14.75
6 1 2 1 22.65 6 1 2 2 20.65
6 1 2 3 23.00 6 1 2 4 20.75
6 1 2 5 20.25 6 1 2 6 19.80
6 1 2 7 17.25 6 1 2 8 15.00
6 1 2 9 19.75 6 1 2 10 14.25
7 2 2 1 23.00 7 2 2 2 20.70
7 2 2 3 22.50 7 2 2 4 20.95
7 2 2 5 22.25 7 2 2 6 20.75
7 2 2 7 18.00 7 2 2 8 14.75
7 2 2 9 20.50 7 2 2 10 15.25
8 3 2 1 22.00 8 3 2 2 19.50
```

8	3	2	3	23.25	8	3	2	4	20.50
8	3	2	5	21.25	8	3	2	6	19.75
8	3	2	7	17.75	8	3	2	8	14.75
8	3	2	9	20.50	8	3	2	10	14.25
9	4	2	1	22.90	9	4	2	2	21.20
9	4	2	3	24.60	9	4	2	4	21.50
9	4	2	5	21.20	9	4	2	6	20.00
9	4	2	7	18.70	9	4	2	8	15.00
9	4	2	9	21.50	9	4	2	10	14.20
10	5	2	1	21.45	10	5	2	2	18.90
10	5	2	3	23.20	10	5	2	4	20.25
10	5	2	5	19.95	10	5	2	6	19.20
10	5	2	7	17.35	10	5	2	8	14.35
10	5	2	9	21.80	10	5	2	10	14.65
11	1	3	1	22.50	11	1	3	2	21.25
11	1	3	3	23.10	11	1	3	4	20.60
11	1	3	5	21.00	11	1	3	6	19.50
11	1	3	7	16.60	11	1	3	8	14.35
11	1	3	9	20.75	11	1	3	10	14.10
12	2	3	1	22.50	12	2	3	2	21.00
12	2	3	3	23.00	12	2	3	4	21.75
12	2	3	5	22.75	12	2	3	6	20.35
12	2	3	7	17.20	12	2	3	8	14.85
12	2	3	9	22.35	12	2	3	10	16.00
13	3	3	1	22.75	13	3	3	2	20.50
13	3	3	3	22.75	13	3	3	4	19.50
13	3	3	5	20.50	13	3	3	6	19.75
13	3	3	7	17.25	13	3	3	8	14.25
13	3	3	9	21.50	13	3	3	10	14.25
14	4	3	1	21.75	14	4	3	2	19.35
14	4	3	3	21.75	14	4	3	4	19.50
14	4	3	5	20.50	14	4	3	6	19.00
14	4	3	7	16.35	14	4	3	8	14.10
14	4	3	9	20.85	14	4	3	10	13.85
15	5	3	1	21.35	15	5	3	2	19.20
15	5	3	3	23.20	15	5	3	4	20.30
15	5	3	5	20.00	15	5	3	6	19.30
15	5	3	7	17.50	15	5	3	8	14.40
15	5	3	9	21.90	15	5	3	10	14.80
16	1	4	1	21.25	16	1	4	2	21.25
16	1	4	3	22.25	16	1	4	4	21.25
16	1	4	5	18.00	16	1	4	6	20.00
16	1	4	7	17.25	16	1	4	8	14.65
16	1	4	9	21.00	16	1	4	10	14.25
17	2	4	1	22.10	17	2	4	2	21.60
17	2	4	3	22.35	17	2	4	4	21.75
17	2	4	5	19.75	17	2	4	6	20.65
17	2	4	7	16.70	17	2	4	8	15.75
17	2	4	9	20.85	17	2	4	10	15.40
18	3	4	1	21.25	18	3	4	2	21.50
18	3	4	3	22.10	18	3	4	4	21.70

18	3	4	5	19.75	18	3	4	6	19.75
18	3	4	7	18.20	18	3	4	8	14.60
18	3	4	9	21.25	18	3	4	10	14.75
19	4	4	1	21.90	19	4	4	2	21.00
19	4	4	3	22.75	19	4	4	4	20.75
19	4	4	5	19.70	19	4	4	6	20.00
19	4	4	7	18.45	19	4	4	8	14.30
19	4	4	9	20.75	19	4	4	10	15.10
20	5	4	1	21.20	20	5	4	2	18.90
20	5	4	3	23.30	20	5	4	4	20.30
20	5	4	5	19.90	20	5	4	6	19.30
20	5	4	7	17.40	20	5	4	8	14.50
20	5	4	9	22.00	20	5	4	10	14.40

;

```
proc print;
title 'Dados para verificação';
run;
```

Uma análise inicial (opcional) para verificar a existência de diferenças entre os tratamentos, sem considerar o esquema fatorial (modelo matemático $y_{ij} = m + t_i + b_j + e_{ij}$), é:

```
proc glm;
title 'Análise de variância sem discriminar os fatores';
class trat bloco;
model ALTURA = trat bloco;
run;
```

Atenção:

Usualmente, a coluna referente aos tratamentos (de 1 a 20, no caso) não é incluída no arquivo. Isso foi feito aqui com apenas com finalidades didáticas. Evidentemente, se não houver a variável classificatória tratamentos, essa análise apresentada acima não é possível.

Já a análise considerando os fatores é ($y_{ijk} = m + a_i + b_j + ab_{ij} + r_k + e_{ijk}$):

```
proc glm;
title 'Análise de variância discriminando os fatores';
class apar oper bloco;
model ALTURA = apar oper apar*oper bloco;
run;
```

Uma vez que nessa análise a interação entre operadores e aparelhos foi significativa, podemos estudar se há diferença entre os operadores para cada aparelho (com Tukey):

```
proc glm;  
title 'Análise de variância na presença de interação I';  
class apar oper bloco;  
model ALTURA = apar oper apar*oper bloco;  
lsmeans apar*oper/slice=apar adjust=tukey PDIFF=all;  
run;
```

Caso desejemos estudar se há diferença entre os aparelhos para cada operador (com Tukey):

```
proc glm;  
title 'Análise de variância na presença de interação II';  
class apar oper bloco;  
model ALTURA = apar oper apar*oper bloco;  
lsmeans apar*oper/slice=oper adjust=tukey PDIFF=all;  
run;
```

Nota:

Na presença de interação significativa, o pesquisador deve definir qual a melhor forma de complementar o estudo, já que diferentes decomposições são possíveis (deve-se tomar cuidado com o nível conjunto de significância).

Exercício 11

Os dados apresentados na tabela a seguir referem-se a um experimento fatorial 2 x 3 inteiramente casualizado (9 repetições). Faça a análise de variância correta e interprete os resultados (dados no arquivo EXERC11.SAS).

Tabela 9. Crescimento de ramos de agrião (cm) após uma semana de cultivo em solução nutritiva.

Horas de luz/dia	Temperatura noturna					
	baixa			alta		
8	3,5	2,5	3,0	8,5	6,5	7,0
	4,0	4,5	3,0	6,0	7,0	7,0
	3,0	5,5	2,5	9,0	8,0	7,0
12	5,0	3,5	4,5	5,5	6,0	6,5
	5,5	3,5	4,0	6,0	8,5	6,5
	4,0	3,0	4,0	3,5	4,5	8,5
16	4,5	5,0	5,5	8,5	9,0	9,5
	5,0	6,0	4,5	9,0	7,0	7,0
	5,0	5,0	6,5	8,5	7,0	9,0

Fonte: Dados fictícios

13 Ensaio fatoriais das séries 2^n e 3^n

Conforme apresentado nas aulas teóricas, esses ensaios têm grande aplicação na área agrônômica, principalmente em ensaios de adubação N-P-K. Caso não seja empregada a técnica de confundimento, a análise desses ensaios é feita da mesma maneira que apresentada no item anterior. Tomemos como exemplo um ensaio fatorial 2^3 .

Exemplo 8 - Um experimento em blocos casualizados no esquema fatorial 2^3 foi instalado para estudar efeitos da adubação na produção da cultura do cafeeiro. O programa com os dados para a análise fica

```
options nodate nonumber ps=65 ls=72;
data exemp8;
input N P K bloco PROD@@;
cards;
0 0 0 1 31.8 0 0 0 2 40.5 0 0 0 3 25.7 0 0 0 4 25.7 0 0 0 5 37.2 0 0 0 6 45.3
0 0 1 1 25.6 0 0 1 2 32.4 0 0 1 3 39.6 0 0 1 4 48.9 0 0 1 5 20.6 0 0 1 6 33.7
0 1 0 1 36.2 0 1 0 2 37.8 0 1 0 3 40.9 0 1 0 4 44.8 0 1 0 5 32.4 0 1 0 6 38.4
0 1 1 1 37.1 0 1 1 2 53.0 0 1 1 3 36.4 0 1 1 4 43.0 0 1 1 5 19.7 0 1 1 6 30.4
1 0 0 1 35.3 1 0 0 2 39.0 1 0 0 3 36.0 1 0 0 4 33.5 1 0 0 5 28.2 1 0 0 6 42.4
1 0 1 1 51.5 1 0 1 2 66.1 1 0 1 3 51.7 1 0 1 4 52.0 1 0 1 5 56.5 1 0 1 6 58.2
1 1 0 1 43.8 1 1 0 2 32.7 1 1 0 3 43.3 1 1 0 4 41.8 1 1 0 5 31.9 1 1 0 6 37.7
1 1 1 1 47.0 1 1 1 2 49.9 1 1 1 3 50.9 1 1 1 4 49.1 1 1 1 5 71.7 1 1 1 6 39.6
;

proc print;
```

```

title 'Dados para verificação';
run;

proc glm;
title 'Análise de variância';
class N P K bloco;
model PROD = N|P|K bloco;
* model PROD = N P K N*P N*K P*K N*P*K bloco;
run;

```

Note:

O modelo matemático adotado é $y_{ijkl} = m + n_i + p_j + k_k + np_{ij} + nk_{ik} + pk_{jk} + npk_{ijk} + b_l + e_{ijkl}$.

```

proc glm;
title 'Análise com interação I';
class N P K bloco;
model PROD = N|P|K bloco;
lsmeans N*K/slice=K ADJUST=TUKEY PDIFF=ALL;
run;

```

```

proc glm;
title 'Análise com interação II';
class N P K bloco;
model PROD = N|P|K bloco;
lsmeans N*K/slice=N ADJUST=TUKEY PDIFF=ALL;
run;

```

Comentários:

Veja que nesses modelos em que há combinações de níveis de fatores é possível simplificar sua indicação no SAS. Assim, são equivalentes

```

model PROD = N|P|K bloco;

```

e

```

model PROD = N P K N*P N*K P*K N*P*K bloco;

```

sendo necessário obviamente especificar somente um deles. Nesse exemplo, a interação N x K foi a única significativa, e desse modo efetuou-se a decomposição apenas dentro de N e de K. Em outras situações, o leitor pode facilmente escolher a análise de interesse.

Exemplo 9 - Vejamos agora um outro exemplo (hipotético) para um fatorial 2^3 (sem confundimento) (arquivo EXEMP9.SAS).

```
data exemp9;
input N P K bl oco bl oco2 PROD@@;
cards;
1 0 0 1 1 34.5 1 0 0 2 2 39.3 1 0 0 3 3 36.7 1 0 0 4 4 34.3
0 0 1 1 1 23.8 0 0 1 2 2 31.4 0 0 1 3 3 39.0 0 0 1 4 4 49.8
1 1 1 1 1 46.5 1 1 1 2 2 50.2 1 1 1 3 3 51.9 1 1 1 4 4 50.1
0 1 0 1 1 35.9 0 1 0 2 2 36.5 0 1 0 3 3 42.8 0 1 0 4 4 45.0

1 1 0 1 5 44.0 1 1 0 2 6 32.8 1 1 0 3 7 43.2 1 1 0 4 8 42.9
0 0 0 1 5 32.8 0 0 0 2 6 41.5 0 0 0 3 7 26.7 0 0 0 4 8 26.5
1 0 1 1 5 52.9 1 0 1 2 6 65.3 1 0 1 3 7 52.7 1 0 1 4 8 52.1
0 1 1 1 5 36.0 0 1 1 2 6 53.4 0 1 1 3 7 35.8 0 1 1 4 8 43.1
;

proc print;
title 'Dados para veri fi cação';
run;

proc glm;
title 'Anál ise de vari ânci a usual';
class N P K bl oco;
model PROD = N|P|K bl oco;
run;

proc glm;
title 'Anál ise de vari ânci a com i nteração';
class N P K bl oco;
model PROD = N|P|K bl oco;
lsmeans N*K/sl i ce=K;
lsmeans N*K/sl i ce=N;
run;
```

Atenção:

Nesse exemplo, há no arquivo uma variável denominada bloco2, que será usada a seguir. No momento, deve ser ignorada. Fica claro que esse tipo de análise é muito simples. A decomposição da(s) interação(ões) significativa(s) fica(m) como exercício para o leitor.

Confundimento

Vejamos agora como realizar essa análise caso tenha sido empregada a técnica de confundimento. Para tanto, usaremos a variável `bloco2`, correspondente aos blocos usando confundimento, e não mais `bloco`, como na análise acima. Recomenda-se que o leitor compare as diferenças entre os resultados das duas análises, prestando também muita atenção para as diferenças que existem na restrição à casualização. É evidente que, em situações reais, o experimento usará ou não a técnica de confundimento; aqui usamos `bloco` e `bloco2` com o intuito de fazer que o leitor consiga diferenciar essas situações.

```
proc glm;
title 'Análise de variância com confundimento';
class N P K bloco2;
model PROD = bloco2 N|P|K/SS1;
run;
```

Esse comando `SS1` colocado no modelo indica que estamos pedindo apenas as somas de quadrado do tipo 1, adequadas a esse caso. É essencial que `bloco2` (blocos com confundimento) seja colocado antes dos demais efeitos do modelo, para que os ajustes sejam feitos para os blocos. Verifique que o SAS colocará 0 graus de liberdade para a interação tripla, já que ela está confundida com o efeito de blocos e não pode ser estudada separadamente.

Na presença de interação dupla, teremos:

```
proc glm;
title 'Análise com interação e confundimento I';
class N P K bloco2;
model PROD = bloco2 N|P|K/SS1;
lsmeans N*K/slice=K ADJUST=TUKEY PDIFF=ALL;
run;
```

```
proc glm;
title 'Análise com interação e confundimento II';
class N P K bloco2;
model PROD = bloco2 N|P|K/SS1;
lsmeans N*K/slice=N ADJUST=TUKEY PDIFF=ALL;
run;
```

Exercício 12

1. Faça croquis mostrando a disposição do experimento apresentado no exemplo 9, sem aplicar o confundimento (usando a variável `bloco`) e com a aplicação do confundimento (usando `bloco2`).

Exemplo 10 - Os dados a seguir referem-se à produção de algodão herbáceo, em kg/ha, num ensaio fatorial 3^3 adubação N-P-K, em blocos ao acaso, com confundimento (arquivo EXEMP10.SAS).

```
options nodate nonumber ps=65 ls=72;
data prat11;
input N P K bloco PROD@@;
cards;
0 0 0 1 868 0 0 0 2 319
0 1 2 1 951 0 1 2 2 521
0 2 1 1 694 0 2 1 2 868
1 0 1 1 972 1 0 1 2 486
1 1 0 1 1319 1 1 0 2 1139
1 2 2 1 812 1 2 2 2 868
2 0 2 1 951 2 0 2 2 1076
2 1 1 1 1493 2 1 1 2 1146
2 2 0 1 1076 2 2 0 2 1215

0 0 1 3 625 0 0 1 4 764
0 1 0 3 1319 0 1 0 4 1042
0 2 2 3 1042 0 2 2 4 729
1 0 2 3 729 1 0 2 4 833
1 1 1 3 764 1 1 1 4 1215
1 2 0 3 806 1 2 0 4 625
2 0 0 3 1285 2 0 0 4 604
2 1 2 3 972 2 1 2 4 1153
2 2 1 3 1042 2 2 1 4 729

0 0 2 5 465 0 0 2 6 1181
0 1 1 5 833 0 1 1 6 729
0 2 0 5 1069 0 2 0 6 660
1 0 0 5 1215 1 0 0 6 937
1 1 2 5 729 1 1 2 6 1083
1 2 1 5 660 1 2 1 6 625
2 0 1 5 1076 2 0 1 6 972
2 1 0 5 1250 2 1 0 6 1139
2 2 2 5 1076 2 2 2 6 1264
;
```

```
proc print;
title 'Dados para verificação';
run;
```

O programa para a análise é

```
proc glm;
title 'Análise de variância com confundimento';
class N P K bloco;
```

```
model PROD = bloco N P K N*P N*K P*K N*P*K/SS1;
run;
```

Exercício 13

A tabela abaixo apresenta as produções, em kg/ha, de parcelas de um experimento fatorial 2^3 de adubação, em blocos casualizados. Os fatores são N, P e K (presentes ou ausentes). O número 1 representa a testemunha (sem adubo), N indica presença apenas de nitrogênio e ausência dos demais elementos, e assim por diante (arquivo EXERC13.SAS).

Tabela 10. Produção (t/ha) de um experimento de adubação de milho (fatorial 2^3).

	1	N	P	NP	K	NK	PK	NPK
Bloco 1	1,32	1,80	1,66	1,72	2,58	2,72	2,26	2,95
Bloco 2	2,12	2,20	2,66	3,85	3,56	3,20	2,08	3,28
Bloco 3	1,75	2,95	1,73	2,62	2,86	2,25	1,95	2,40
Bloco 4	2,35	2,96	2,58	3,00	2,75	2,75	2,70	3,35

Fonte: Pimentel Gomes, 1987

Pede-se:

1. Criar o arquivo SAS adequado para analisar esses dados.
2. Obtenha também um data set permanente.
3. Fazer a análise de variância
4. Por que a presença de interação significativa requer que sejam feitos estudos de um fator dentro dos níveis do outro com o qual interagiu?
5. Efetue as recomendações de adubação para N, P e K (se necessário, especificando em que níveis dos fatores a recomendação é válida).

Exercício 14

A seguir são apresentados resultados de um experimento de adubação de cana-de-açúcar, com confundimento de 2 graus de liberdade da interação tripla (arquivo EXERC14.TXT).

Tabela 11. Produção de cana-de-açúcar (t/ha) de um experimento fatorial (blocos casualizados) com confundimento.

Trat.	Rep. 1	Rep. 2	Trat.	Rep. 1	Rep. 2	Trat	Rep. 1	Rep. 2
000	37,0	29,3	001	60,2	39,1	002	28,3	24,3
012	42,6	37,0	010	57,6	19,5	011	66,6	41,8
021	68,4	64,4	022	76,0	66,0	020	66,8	55,1
101	33,5	28,6	102	42,8	27,7	100	32,6	37,0
110	45,7	61,3	111	63,4	63,0	112	63,4	50,8
122	49,7	52,8	120	58,2	53,4	121	71,1	56,8
202	36,2	35,2	200	56,6	30,5	201	41,3	36,9
211	47,4	55,4	212	69,1	49,5	210	50,2	67,2
200	59,0	56,2	221	59,4	43,2	222	55,4	69,5

Fonte: Pimentel Gomes, 1987

1. Criar o arquivo SAS com esses dados e fazer a análise de variância
2. Fazer os desdobramentos necessários, com teste de Tukey, interpretando os resultados das análises.

14 Uso da regressão na análise de variância

Em diversas situações práticas temos que os níveis dos fatores são quantitativos. Nesse caso, diversos autores recomendam o estudo de regressões, normalmente mais informativo que os testes de comparações múltiplas (ver aula teórica).

No caso do exemplo 10, temos que os níveis de N, P e K são quantitativos, sendo recomendável efetuar esse estudo. Embora K não tenha apresentado F significativo, apresentaremos aqui a regressão apenas por razões didáticas.

Nesse caso, são criadas variáveis auxiliares indicando o grau do polinômio a ser considerado. Note que agora devem ser especificados os níveis de cada fator (por exemplo, N com níveis 0-40-80), o que não foi feito no exemplo anterior (arquivo: EXEMP10b.SAS).

```
options nodate nonumber ps=65 ls=72;
data;
input N P K bloco PROD@@;

LI NEAR_N=N; QUAD_N=N*N; LI NEAR_P=P; QUAD_P=P*P; LI NEAR_K=K; QUAD_K=K*K;

cards;
0 0 0 1 868 0 0 0 2 319
```

```

0 60 120 1 951      0 60 120 2 521
0 120 60 1 694      0 120 60 2 868
40 0 60 1 972       40 0 60 2 486
40 60 0 1 1319      40 60 0 2 1139
40 120 120 1 812    40 120 120 2 868
80 0 120 1 951      80 0 120 2 1076
80 60 60 1 1493     80 60 60 2 1146
80 120 0 1 1076     80 120 0 2 1215

0 0 60 3 625        0 0 60 4 764
0 60 0 3 1319       0 60 0 4 1042
0 120 120 3 1042    0 120 120 4 729
40 0 120 3 729      40 0 120 4 833
40 60 60 3 764      40 60 60 4 1215
40 120 0 3 806      40 120 0 4 625
80 0 0 3 1285       80 0 0 4 604
80 60 120 3 972     80 60 120 4 1153
80 120 60 3 1042    80 120 60 4 729

0 0 120 5 465       0 0 120 6 1181
0 60 60 5 833       0 60 60 6 729
0 120 0 5 1069      0 120 0 6 660
40 0 0 5 1215       40 0 0 6 937
40 60 120 5 729     40 60 120 6 1083
40 120 60 5 660     40 120 60 6 625
80 0 60 5 1076      80 0 60 6 972
80 60 0 5 1250      80 60 0 6 1139
80 120 120 5 1076   80 120 120 6 1264
;

```

```

proc print;
title 'Dados para verificação';
run;

```

Atenção:

As variáveis auxiliares criadas foram LINEAR_N, QUAD_N, LINEAR_P, QUAD_P, LINEAR_K, QUAD_K.

O programa para a análise de variância já foi apresentado:

```

proc glm;
title 'Análise de variância com confundimento';
class N P K bloco;
model PROD = bloco N P K N*P N*K P*K N*P*K/SS1;
run;

```

Vejam agora os programas que possibilitam testar os componentes lineares e quadráticos para N, P e K:

```
proc glm;
title 'Anava tipo I com confundimento e regressões';
class N P K bloco;
model PROD = bloco LINEAR_N QUAD_N LINEAR_P QUAD_P LINEAR_K QUAD_K
          N*P N*K P*K N*P*K/SS1;
run;
```

Atenção:

Devemos, necessariamente, usar as somas de quadrado do tipo I. O efeito de blocos e os demais no modelo devem estar na ordem indicada acima.

A obtenção da função de primeiro grau associando os níveis de N e a produção pode ser obtida com o seguinte programa:

```
proc reg;
title 'Equação de primeiro grau para N';
model PROD = LINEAR_N;
run;
```

Já a função de segundo grau para P é obtida com

```
proc reg;
title 'Equação de segundo grau para P';
model PROD = LINEAR_P QUAD_P;
run;
```

Nota:

Optamos por essas funções com base no resultado da análise de variância, que deve ser cuidadosamente interpretado pelo leitor. Note que não foi feito nenhum teste de comparações múltiplas. É importante ressaltar que as interações não foram significativas. Caso isso não tivesse ocorrido, seriam necessários desdobramentos.

Exercício 15

Os dados abaixo referem-se a um experimento (hipotético) estudando a influência do número de dias e quantidade de um dado fungicida na quantidade de matéria seca

produzida por plantas de uma horta. O delineamento foi blocos casualizados (arquivo não fornecido).

Tabela 12. Dados resultantes do experimento

Dias	Quantidade de fungicida	Bloco 1	Bloco 2	Bloco 3	Bloco 4
3	0	15,7	14,6	16,5	14,7
	4	9,8	14,6	11,9	12,4
	8	7,9	10,3	9,7	9,6
10	0	18,0	17,4	15,1	14,4
	4	13,6	10,6	11,8	13,3
	8	8,8	8,2	11,3	11,2

1. Crie o arquivo SAS adequado.
2. Faça a análise de variância, desdobrando a interação, se necessário.
3. Faça um estudo usando regressão para os fatores onde há diferença entre os níveis. Apresente a(s) função(ões) ajustadas.
4. Discuta os resultados detalhadamente, explicando o significado das funções encontradas.

Exercício 16

Os dados abaixo referem-se a um experimento de adubação de milho em blocos ao acaso. Os tratamentos constaram de adubação com 0, 25, 50, 75 e 100 kg/ha de P_2O_5 (arquivo EXERC16.XLS).

Tabela 13. Produção de milho, em kg/parcela, de um experimento de adubação de milho.

0	25	50	75	100
3,38	7,15	10,07	9,55	9,14
5,77	9,78	9,73	8,95	10,17
4,90	9,99	7,92	10,24	9,75
4,54	10,10	9,48	8,66	9,50

Fonte: Pimentel Gomes, 1987

1. Faça a análise de variância.
2. Faça um estudo usando regressão para os fatores onde há diferença entre os níveis. Apresente a(s) função(ões) ajustadas.
3. Discuta os resultados, explicando o significado das funções encontradas

15 Ensaio em parcelas subdivididas

Exemplo 11 - Apresentamos a seguir os dados referentes a % de brix no caldo de um experimento de aplicação de maturadores de cana-de-açúcar (aula teórica). O delineamento empregado foi blocos ao acaso com 3 parcelas. De cada parcela, foram coletadas amostras após 0, 2, 4, 6, 8 e 10 semanas de aplicação dos maturadores (arquivo EXEMP11.SAS).

```
data;
input trat $ epoca bloco BRI X@@;
cards;
T1 0 1 17.70 T1 0 2 17.50 T1 0 3 17.95 T1 0 4 17.90
T1 0 5 17.30 T1 0 6 17.60 T1 0 7 17.40 T1 0 8 17.75
T1 2 1 16.95 T1 2 2 18.60 T1 2 3 17.05 T1 2 4 18.00
T1 2 5 19.40 T1 2 6 18.55 T1 2 7 17.80 T1 2 8 18.05
T1 4 1 18.95 T1 4 2 18.10 T1 4 3 17.60 T1 4 4 18.35
T1 4 5 19.20 T1 4 6 18.90 T1 4 7 18.25 T1 4 8 18.70
T1 6 1 18.65 T1 6 2 18.50 T1 6 3 19.65 T1 6 4 18.30
T1 6 5 20.10 T1 6 6 18.75 T1 6 7 19.25 T1 6 8 17.40
T1 8 1 19.70 T1 8 2 20.20 T1 8 3 19.70 T1 8 4 19.70
T1 8 5 20.55 T1 8 6 20.25 T1 8 7 19.80 T1 8 8 18.70
T1 10 1 20.75 T1 10 2 19.00 T1 10 3 19.85 T1 10 4 20.80
T1 10 5 19.45 T1 10 6 21.70 T1 10 7 19.10 T1 10 8 18.25

T2 0 1 16.83 T2 0 2 16.96 T2 0 3 17.25 T2 0 4 18.22
T2 0 5 17.75 T2 0 6 16.90 T2 0 7 16.70 T2 0 8 17.19
T2 2 1 17.70 T2 2 2 17.56 T2 2 3 18.17 T2 2 4 17.72
T2 2 5 19.48 T2 2 6 18.20 T2 2 7 18.58 T2 2 8 18.14
T2 4 1 18.00 T2 4 2 18.05 T2 4 3 18.55 T2 4 4 19.52
T2 4 5 20.70 T2 4 6 18.85 T2 4 7 18.41 T2 4 8 19.20
T2 6 1 18.68 T2 6 2 18.83 T2 6 3 17.65 T2 6 4 20.40
T2 6 5 20.03 T2 6 6 19.63 T2 6 7 18.70 T2 6 8 19.71
T2 8 1 19.55 T2 8 2 19.38 T2 8 3 18.10 T2 8 4 20.85
T2 8 5 20.80 T2 8 6 20.42 T2 8 7 19.33 T2 8 8 21.10
T2 10 1 15.56 T2 10 2 17.18 T2 10 3 19.00 T2 10 4 19.40
T2 10 5 20.45 T2 10 6 19.92 T2 10 7 20.00 T2 10 8 19.65

T3 0 1 16.63 T3 0 2 17.70 T3 0 3 17.52 T3 0 4 17.52
T3 0 5 17.26 T3 0 6 17.18 T3 0 7 17.43 T3 0 8 16.52
T3 2 1 17.18 T3 2 2 17.75 T3 2 3 17.65 T3 2 4 17.55
T3 2 5 18.61 T3 2 6 18.26 T3 2 7 18.00 T3 2 8 16.56
T3 4 1 18.05 T3 4 2 18.20 T3 4 3 18.57 T3 4 4 19.57
T3 4 5 19.03 T3 4 6 18.91 T3 4 7 18.48 T3 4 8 17.73
T3 6 1 18.50 T3 6 2 19.53 T3 6 3 19.08 T3 6 4 19.03
T3 6 5 19.48 T3 6 6 18.95 T3 6 7 18.77 T3 6 8 16.54
T3 8 1 20.65 T3 8 2 20.90 T3 8 3 18.68 T3 8 4 18.55
T3 8 5 20.28 T3 8 6 20.25 T3 8 7 19.27 T3 8 8 15.60
T3 10 1 19.90 T3 10 2 23.96 T3 10 3 19.08 T3 10 4 19.03
T3 10 5 20.62 T3 10 6 20.33 T3 10 7 20.06 T3 10 8 18.85
;
```

```
proc print;
title 'Dados para verificação';
run;
```

O modelo matemático é $y_{ijk} = m + a_i + b_j + e_{ij} + s_k + as_{ik} + e_{ijk}$ (ver aula teórica). O programa para realizar a análise de variância pode ser:

```
proc glm;
title 'Análise de variância - parc. subd.';
class trat epoca bloco;
model BRIX = bloco trat trat*bloco epoca epoca*trat;
test h=trat e=trat*bloco;
* means trat/tukey e=trat*bloco; /*apenas para indicação*/
* means epoca/tukey; /*apenas para indicação*/
run;
```

Note que, como agora temos dois resíduos envolvidos, é necessário usar um artifício na indicação do modelo: como a soma de quadrados da interação entre tratamentos e blocos coincide nesse caso com a soma de quadrados do resíduo (a), especificamos esse efeito no modelo, devendo contudo indicar que a fonte de variação tratamentos deve ser testada com esse resíduo (comando `test h=trat e=trat*bloco`). O SAS testa todas as fontes de variação automaticamente com o resíduo do modelo (no caso, resíduo b), o que é incorreto para tratamentos. Note que blocos foram testados com o resíduo b, uma vez que não especificamos outro testador. Caso tivéssemos interesse nesse teste, deveríamos especificar como testador o resíduo a.

Cuidado

O programa apresenta o teste de Tukey para os tratamentos das parcelas, especificando que o resíduo correto é o Resíduo (a), além do teste para as subparcelas (Resíduo b, feito automaticamente). Contudo, na presença de interação significativa, como é o caso, devemos fazer o desdobramento da interação, e não esse teste, aqui mostrado apenas para exemplificar com deve ser especificado o resíduo na ausência de interação significativa.

Consideremos o estudo de épocas dentro de tratamentos (efeito das subparcelas dentro dos efeitos das parcelas):

```
proc glm;
title 'Anava - época dentro de trat.';
class trat epoca bloco;
model BRIX = bloco trat trat*bloco epoca epoca*trat;
lsmeans epoca*trat/slice=trat;
* test h=trat e=trat*bloco;
```

```
run;
```

Note que o teste F para tratamentos está incorreto, pois está usando o Resíduo (b) como testador. A linha com o asterisco mostra como corrigir o problema. Lembre-se que o asterisco “desliga” o processamento da linha em questão, devendo ser retirado caso haja interesse em processá-la.

Para estudar tratamentos dentro de épocas (efeito da parcela dentro dos efeitos das subparcelas):

```
proc glm;  
title 'Anava - trat. dentro de época';  
class trat época bloco;  
model BRIX = bloco trat época época*trat;  
lsmeans época*trat/slice=época adjust=tukey PDIFF=all;  
run;
```

Atenção

Há, nesse segundo desdobramento, a necessidade de realizar uma combinação dos resíduos (ver aula teórica). Optamos aqui por adotar o procedimento recomendado por Perecin et al. (2000) para essa análise no SAS. Para tanto, suprimimos o Resíduo (a) (tratamentos x blocos) do modelo.

Vejamos agora como fazer um estudo usando regressão sobre os níveis de época, dentro de cada tratamento, uma vez que a interação foi significativa.

Novamente, faremos uso de variáveis auxiliares:

- EpDtr1 significa épocas dentro do tratamento 1;
- X1EpDtr1, será a variável auxiliar para ajuste linear (x^1) em função das épocas dentro do tratamento 1;
- X2EpDtr1, será a variável auxiliar para ajuste quadrático (x^2) para épocas dentro do tratamento 1;
- As demais variáveis auxiliares têm significado análogo;
- Note que época é uma variável já presente no data set.

Exceto pela presença dessas variáveis, o data set ficará idêntico ao já criado no exemplo 11.

```
data subdiv;  
input trat $ época bloco BRIX@@;  
  
if trat='T1' then EpDtr1=época; ELSE EpDtr1=100;  
if trat='T2' then EpDtr2=época; ELSE EpDtr2=100;  
if trat='T3' then EpDtr3=época; ELSE EpDtr3=100;
```

```

X1EpDtr1=EpDtr1; X2EpDtr1=EpDtr1*EpDtr1;
X1EpDtr2=EpDtr2; X2EpDtr2=EpDtr2*EpDtr2;
X1EpDtr3=EpDtr3; X2EpDtr3=EpDtr3*EpDtr3;

```

```
cards;
```

```

T1  0 1 17.70   T1  0 2 17.50   T1  0 3 17.95   T1  0 4 17.90
T1  0 5 17.30   T1  0 6 17.60   T1  0 7 17.40   T1  0 8 17.75
T1  2 1 16.95   T1  2 2 18.60   T1  2 3 17.05   T1  2 4 18.00
T1  2 5 19.40   T1  2 6 18.55   T1  2 7 17.80   T1  2 8 18.05
T1  4 1 18.95   T1  4 2 18.10   T1  4 3 17.60   T1  4 4 18.35
T1  4 5 19.20   T1  4 6 18.90   T1  4 7 18.25   T1  4 8 18.70
T1  6 1 18.65   T1  6 2 18.50   T1  6 3 19.65   T1  6 4 18.30
T1  6 5 20.10   T1  6 6 18.75   T1  6 7 19.25   T1  6 8 17.40
T1  8 1 19.70   T1  8 2 20.20   T1  8 3 19.70   T1  8 4 19.70
T1  8 5 20.55   T1  8 6 20.25   T1  8 7 19.80   T1  8 8 18.70
T1 10 1 20.75   T1 10 2 19.00   T1 10 3 19.85   T1 10 4 20.80
T1 10 5 19.45   T1 10 6 21.70   T1 10 7 19.10   T1 10 8 18.25

```

```

T2  0 1 16.83   T2  0 2 16.96   T2  0 3 17.25   T2  0 4 18.22
T2  0 5 17.75   T2  0 6 16.90   T2  0 7 16.70   T2  0 8 17.19
T2  2 1 17.70   T2  2 2 17.56   T2  2 3 18.17   T2  2 4 17.72
T2  2 5 19.48   T2  2 6 18.20   T2  2 7 18.58   T2  2 8 18.14
T2  4 1 18.00   T2  4 2 18.05   T2  4 3 18.55   T2  4 4 19.52
T2  4 5 20.70   T2  4 6 18.85   T2  4 7 18.41   T2  4 8 19.20
T2  6 1 18.68   T2  6 2 18.83   T2  6 3 17.65   T2  6 4 20.40
T2  6 5 20.03   T2  6 6 19.63   T2  6 7 18.70   T2  6 8 19.71
T2  8 1 19.55   T2  8 2 19.38   T2  8 3 18.10   T2  8 4 20.85
T2  8 5 20.80   T2  8 6 20.42   T2  8 7 19.33   T2  8 8 21.10
T2 10 1 15.56   T2 10 2 17.18   T2 10 3 19.00   T2 10 4 19.40
T2 10 5 20.45   T2 10 6 19.92   T2 10 7 20.00   T2 10 8 19.65

```

```

T3  0 1 16.63   T3  0 2 17.70   T3  0 3 17.52   T3  0 4 17.52
T3  0 5 17.26   T3  0 6 17.18   T3  0 7 17.43   T3  0 8 16.52
T3  2 1 17.18   T3  2 2 17.75   T3  2 3 17.65   T3  2 4 17.55
T3  2 5 18.61   T3  2 6 18.26   T3  2 7 18.00   T3  2 8 16.56
T3  4 1 18.05   T3  4 2 18.20   T3  4 3 18.57   T3  4 4 19.57
T3  4 5 19.03   T3  4 6 18.91   T3  4 7 18.48   T3  4 8 17.73
T3  6 1 18.50   T3  6 2 19.53   T3  6 3 19.08   T3  6 4 19.03
T3  6 5 19.48   T3  6 6 18.95   T3  6 7 18.77   T3  6 8 16.54
T3  8 1 20.65   T3  8 2 20.90   T3  8 3 18.68   T3  8 4 18.55
T3  8 5 20.28   T3  8 6 20.25   T3  8 7 19.27   T3  8 8 15.60
T3 10 1 19.90   T3 10 2 23.96   T3 10 3 19.08   T3 10 4 19.03
T3 10 5 20.62   T3 10 6 20.33   T3 10 7 20.06   T3 10 8 18.85

```

```
;
```

```

proc print;
title 'Dados para verificação';
run;

```

O estudo das regressões pode ser feito com o seguinte programa:

```
proc glm;
title 'Anava - época dentro de trat. com reg. linear e quadrática';
class trat epoca bloco EpDtr1 EpDtr2 EpDtr3;
model BRIX = bloco trat trat*bloco
          X1EpDtr1(trat) X2EpDtr1(trat) EpDtr1(trat)
          X1EpDtr2(trat) X2EpDtr2(trat) EpDtr2(trat)
          X1EpDtr3(trat) X2EpDtr3(trat) EpDtr3(trat)/SS1;
run;
```

Fato:

Esse programa não é simples, e seu entendimento está além dos objetivos desse curso. Porém, pode ser usado em situações reais, com as adaptações necessárias. O programa está no arquivo EXEMP11b.SAS.

A obtenção das funções lineares e quadráticas é mais simples. Note que é necessário criar novos data sets selecionando apenas os dados de cada tratamento. A escolha do grau do polinômio baseou-se na análise de variância realizada com o programa apresentado acima.

```
data trat1;
set subdiv;
if trat='T1';
run;
proc reg data=trat1;
title 'Equação de primeiro grau para trat1';
model BRIX = X1EpDtr1;
run;
```

```
data trat2;
set subdiv;
if trat='T2';
run;
proc reg data=trat2;
title 'Equação de segundo grau para trat2';
model BRIX = X1EpDtr2 X2EpDtr2;
run;
```

```
data trat3;
set subdiv;
if trat='T3';
run;
proc reg data=trat3;
title 'Equação de primeiro grau para trat3';
model BRIX = X1EpDtr3;
```

run;

Exercício 17

Os dados a seguir referem-se a um experimento com parcelas subdivididas, sendo as parcelas referentes a métodos de irrigação e as subparcelas a espaçamentos. O delineamento foi blocos ao acaso (arquivo EXERC17.XLS).

Tabela 14. Dados (altura de plantas) de um experimento em blocos ao acaso com parcelas subdivididas.

Bloco	Irrigação A		Irrigação B		Irrigação C	
	esp.1	esp. 2	esp. 1	esp. 2	esp. 1	esp. 2
I	58	44	85	59	66	54
II	77	59	90	68	93	75
III	38	30	73	45	67	53
IV	52	34	77	55	64	48

Fonte: Vieira, 1999.

1. Fazer a análise de variância adequada, com teste de Tukey.

16 Análise de grupos de experimentos

No caso de diversos experimentos semelhantes serem repetidos em diversas condições, como, por exemplo, diversos locais, anos, épocas, etc, podemos usar essa técnica de análise.

No SAS, essas análises são relativamente simples se os experimentos repetidos forem iguais. Para tanto, basta especificar corretamente qual é o modelo matemático a ser usado.

Exemplo 12 - Três experimentos para comparação de plantas de *Eucaliptus grandis* quanto à altura (metros) aos 7 anos de idade foram realizados nas localidades de Araraquara, Bento Quirino e Mogi-Guaçu (Fonte: Eng. Agrôn, L. C. Costa Coelho, Tupi, SP). O delineamento empregado foi blocos ao acaso (dados no arquivo EXEMP12.SAS).

Os resultados podem ser colocados num único arquivo:

```
options nodate nonumber ps=65 ls=72;
data conjunta;
length local $ 12;
input local $ trat $ bloco ALTURA@@;
datalines;
```

```

ARARAQUARA      T1  1  20. 3  ARARAQUARA      T1  2  19. 6
ARARAQUARA      T1  3  23. 5  ARARAQUARA      T1  4  19. 1
ARARAQUARA      T2  1  21. 7  ARARAQUARA      T2  2  19. 3
ARARAQUARA      T2  3  16. 7  ARARAQUARA      T2  4  18. 5
ARARAQUARA      T3  1  22. 0  ARARAQUARA      T3  2  24. 9
ARARAQUARA      T3  3  24. 4  ARARAQUARA      T3  4  20. 8
ARARAQUARA      T4  1  20. 8  ARARAQUARA      T4  2  23. 0
ARARAQUARA      T4  3  21. 3  ARARAQUARA      T4  4  24. 9
ARARAQUARA      T5  1  21. 5  ARARAQUARA      T5  2  22. 3
ARARAQUARA      T5  3  22. 1  ARARAQUARA      T5  4  21. 9
ARARAQUARA      T6  1  19. 6  ARARAQUARA      T6  2  17. 7
ARARAQUARA      T6  3  18. 7  ARARAQUARA      T6  4  22. 0
BENTOQUI RI NO  T1  1  10. 2  BENTOQUI RI NO  T1  2  11. 7
BENTOQUI RI NO  T1  3   9. 1  BENTOQUI RI NO  T1  4   8. 1
BENTOQUI RI NO  T2  1  16. 1  BENTOQUI RI NO  T2  2  10. 8
BENTOQUI RI NO  T2  3  10. 9  BENTOQUI RI NO  T2  4  10. 3
BENTOQUI RI NO  T3  1  17. 7  BENTOQUI RI NO  T3  2  13. 1
BENTOQUI RI NO  T3  3  14. 2  BENTOQUI RI NO  T3  4  11. 0
BENTOQUI RI NO  T4  1  13. 5  BENTOQUI RI NO  T4  2  14. 4
BENTOQUI RI NO  T4  3  11. 2  BENTOQUI RI NO  T4  4  12. 8
BENTOQUI RI NO  T5  1  20. 5  BENTOQUI RI NO  T5  2  12. 5
BENTOQUI RI NO  T5  3  11. 3  BENTOQUI RI NO  T5  4  12. 2
BENTOQUI RI NO  T6  1  12. 0  BENTOQUI RI NO  T6  2  13. 0
BENTOQUI RI NO  T6  3  12. 3  BENTOQUI RI NO  T6  4  10. 6
MOGI _GUACU     T1  1  22. 7  MOGI _GUACU     T1  2  21. 4
MOGI _GUACU     T1  3  22. 9  MOGI _GUACU     T1  4  22. 0
MOGI _GUACU     T2  1  22. 6  MOGI _GUACU     T2  2  21. 4
MOGI _GUACU     T2  3  20. 7  MOGI _GUACU     T2  4  20. 8
MOGI _GUACU     T3  1  21. 4  MOGI _GUACU     T3  2  21. 7
MOGI _GUACU     T3  3  22. 5  MOGI _GUACU     T3  4  19. 4
MOGI _GUACU     T4  1  25. 0  MOGI _GUACU     T4  2  23. 6
MOGI _GUACU     T4  3  23. 3  MOGI _GUACU     T4  4  24. 8
MOGI _GUACU     T5  1  26. 4  MOGI _GUACU     T5  2  26. 4
MOGI _GUACU     T5  3  28. 0  MOGI _GUACU     T5  4  27. 3
MOGI _GUACU     T6  1  20. 6  MOGI _GUACU     T6  2  23. 5
MOGI _GUACU     T6  3  19. 4  MOGI _GUACU     T6  4  21. 9
;

```

```

proc print;
title 'Dados para verificação';
run;

```

E a análise para cada local pode ser

```

proc sort;
  by local;
run;

```

```
proc glm;
title 'Análises individuais';
by local;
class trat bloco;
model ALTURA = bloco trat;
run;
```

Finalmente, podemos fazer a análise de variância conjunta:

```
proc glm;
title 'Análise conjunta I';
class local trat bloco;
model ALTURA = local bloco(local) trat trat*local;
run;
```

```
proc glm;
title 'Análise conjunta II';
class local trat bloco;
model ALTURA = local bloco(local) trat trat*local;
lsmeans trat*local / slice=local;
run;
```

Notas:

1. O segundo programa deve ser usado quando a interação for significativa.
2. No primeiro programa, indicou-se bloco(local), que significa “blocos dentro de locais”, pois a numeração em cada local é arbitrária (I, II, etc), não tendo nenhuma relação com a numeração em outro local.

Exercício 18

Os dados contidos no arquivo EXERC18.SAS e apresentados na Tabela 15, referem-se a experimentos em blocos ao acaso para avaliação de 20 progênies de milho em dois locais (Fonte: Geraldi, I.O.; 2000 - Departamento de Genética, ESALQ/USP). Pede-se:

1. Fazer as análises de variância em cada local.
2. Fazer a análise de variância conjunta.
3. Recomendar as progênies superiores.
4. Qual foi o modelo matemático empregado? (se necessário, pesquise na bibliografia)

Tabela 15. Dados de avaliação de 20 progênies de milho em blocos ao acaso, em dois locais. Produção em g/parcela.

	LOCAL 1			LOCAL 2		
	I	II	III	I	II	III
1	251	192	262	218	227	316
2	198	200	153	150	153	172
3	186	156	182	206	226	234
4	121	147	181	162	172	183
5	159	173	142	176	263	182
6	141	167	169	166	147	204
7	232	242	160	180	194	220
8	161	168	155	204	229	232
9	277	253	289	332	243	273
10	144	142	137	270	207	278
11	153	142	149	158	163	189
12	195	176	277	221	204	205
13	159	155	157	224	259	225
14	203	198	183	282	279	361
15	204	187	199	248	224	184
16	166	142	163	203	297	293
17	185	199	183	147	223	154
18	183	149	151	211	226	258
19	214	198	186	220	245	218
20	141	141	176	136	155	218

Fonte: Geraldi, 2000

17 Bibliografia recomendada

- Alves, M. C. SAS Básico. Série Didática Ciagri nº 9. ESALQ/USP, Piracicaba, SP. 1990.
- Barbin, D. Planejamento e Análise Estatística de Experimentos Agronômicos. Apostila. Departamento de Matemática e Estatística, ESALQ/USP, Piracicaba, SP. 1994.
- Cody, R.P.; Smith, J.K. Applied Statistics and the SAS Programming Language. 3rd ed. Elsevier Science Publishing Co., New York. 1991. 403p.
- Diiorio, F.C. SAS Applications Programming: A Gentle Introduction. Duxbury Press, California. 1997. 684 p.

- Garcia, A.A.F.; Geraldi, I.O. Alguns Programas no SAS. Publicação didática. Departamento de Ciências Exatas, ESALQ/USP, Piracicaba, SP. Julho de 2000. 6p.
- Garcia, A.A.F.; Iemma, A.F. Introdução ao SAS/IML. Publicação didática. Departamento de Ciências Exatas, Escola Superior de Agricultura "Luiz de Queiroz". Março de 2000. 20p.
- Iemma, A. F. Que hipóteses estatísticas testamos através do SAS em presença de caselas vazias. Sci. Agric. 52(2): 210-220, 1995.
- Latour, D.; Littell. Advanced General Linear Models with an Emphasis on Mixed Models - Course Notes. SAS Institute Inc., Cary, NC. 1996.
- Littell, R. C.; Milliken, G. A.; Stroup, W. W.; Wolfinger, R. D. SAS System for Mixed Models. SAS Institute Inc., Cary, NC. 1996.
- Perecin, D.; Malheiros, E. B.; Pereira, G. T. Variáveis auxiliares para expressar desdobramentos de graus de liberdade e contrastes com o programa SAS. Anais da 45ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria. Universidade Federal de São Carlos, São Carlos, SP. 2000. p.137-140.
- Piedade, S. M. S. Aulas práticas de Estatística Experimental, 1º semestre de 2000. Publicação didática. Departamento de Ciências Exatas, ESALQ/USP, Piracicaba, SP.
- Reyes, A. E. L. SAS - Estatística Básica. Série Didática Ciagri nº 17. ESALQ/USP, Piracicaba, SP. 1990.
- SAS/STAT User's Guide, Version 6, Fourth Edition, vol. 1 e 2. SAS Institute Inc., Cary, NC. 1994.
- Scott, R.A.; Milliken, G.A. A SAS program for analysing augmented randomized complete block design. Crop Science 33: 865-867, 1993.
- Zocchi, S.S.; Barbin, D. LCE 602 - Estatística Experimental, 2º semestre de 1999. Publicação didática. Departamento de Ciências Exatas, ESALQ/USP, Piracicaba, SP.