

Análise de Experimentos em Genética e Melhoramento

Usando o Pacote Computacional SAS

Antonio Augusto Franco Garcia
Departamento de Genética, ESALQ/USP
Piracicaba, abril de 2004

Objetivos

Aprimorar os conhecimentos dos interessados na aplicação do SAS em situações encontradas em Genética e Melhoramento de Plantas. Pressupõe-se conhecimentos básicos de Biometria e Genética Quantitativa. Alguma familiaridade com o SAS, embora não exigida, facilitará o acompanhamento do curso.

Programa

i) Estrutura do SAS: noções gerais, criação e manipulação de arquivos; ii) Programação: linguagem de programação, variáveis e objetos, operadores lógicos, arrays; iii) Análise de dados de marcadores moleculares I (coeficientes de similaridade e mapas genéticos); iv) Bootstrap; v) Estatística descritiva; vi) Experimentos Inteiramente Casualizado, incluindo verificação das pressuposições do modelo; vii) Experimentos em Blocos Casualizados, incluindo análise conjunta e experimentos com parcelas perdidas; viii) Experimentos Fatoriais; ix) Experimentos Hierárquicos; x) Experimentos em Parcelas sub-divididas; xi) Experimentos em Látices; xii) Blocos Aumentados; xiii) Modelos Mistos e Componentes de Variância; xiv) Planejamento de Experimentos; xv) Análise de dados de marcadores moleculares II (mapeamento de QTL's: análise de marcas simples); xvi) Noções do SAS/IML para Álgebra de Matrizes, incluindo análise dialéctica.

Carga horária

10 horas

I. CRIAÇÃO DE DATA SETS

Neste item, serão apresentados alguns programas que ilustram a criação dos chamados SAS data sets. É importante que cada um deles seja estudado cuidadosamente. Note que em muitos casos os resultados obtidos podem ser diferentes dos esperados.

(prog1.sas)

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

```
B 2 145
;
proc print;
title 'Dados para verificação';
run;
```

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

```
data;
input trat $ bloco PESO;
cards;
A 1 100 A 2 130
B 1 99 B 2 145
;
proc print;
title 'Arquivo com problemas';
run;
```

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

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

```

run;

data;
input trat $ bloco PESO;
cards;
A23456789 1 100
A23456789 2 130
B23456789 1 99
B23456789 2 145
;
proc print;
run;

data;
length trat $ 9;
input trat $ bloco PESO;
cards;
A23456789 1 100
A23456789 2 130
B23456789 1 99
B23456789 2 145
;
proc print;
run;

data;
input trat $ bloco PESO;
  if bloco <=1 then CONJUNTO=100;
  else CONJUNTO=200;
cards;
A 1 100
A 2 130
B 1 99
B 2 145
;
proc print;
run;

data;
input trat $ bloco PESO;
  if trat='A' then delete;
cards;
A 1 100
A 2 130
B 1 99
B 2 145
;
proc print;
run;

data;
input trat $ bloco PESO;
  if PESO=100 then PESO=333;
cards;
A 1 100
A 2 130
B 1 99
B 2 145
;
proc print;
run;

```

```

data;
input trat $ bloco PESO;
PESO=SQRT(PESO+0.5);
cards;
A 1 100
A 2 130
B 1 99
B 2 145
;
proc print;
run;

proc print data=data2;
title 'data set data2';
run;

proc print data=data4;
title 'data set data4';
run;

```

- No DATA STEP do SAS são permitidos comandos de programação (Tabela 1), o que confere grande versatilidade ao pacote. Esses recursos podem ser utilizados para análise de dados, conforme será apresentado.
- Também é possível IMPORTAR/EXPORTAR arquivos de diversos formatos. Como exemplo, importe o arquivo do Excel chamado Exemplo 1.xls.

Tabela 1: Operadores aritméticos e de comparação

Operador	Operação	Operador	Operação
**	Exponenciação	=	Igual
*	Multiplicação	^=	Diferente
/	Divisão	>	Maior
+	Adição	>=	Maior ou igual
-	Subtração	<	Menor
		<=	Menor ou igual

II. MANIPULAÇÃO DE ARQUIVOS

Veremos agora como trabalhar com diversos arquivos criados separadamente. Os comandos SET e MERGE permitem combinar esses arquivos de duas formas diferentes, muito úteis na prática quando manipulamos grandes conjuntos de dados.

(prog2.sas)

```

data exp1;
input local $ var $ alt_plt;
datalines;
Pirac M 185
Pirac N 196
Pirac O 200
Pirac 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=alt_esp/alt_plt;
run;
proc print;
title 'Comando MERGE';
run;

```

III. VARIÁVEIS E INDIVÍDUOS

Exemplificaremos agora como diferenciar variáveis e indivíduos. Normalmente, as procedures (proc's) efetuam operações entre os indivíduos, enquanto que as funções (functions) realizam operações dentro dos indivíduos.

(prog3.sas)

```

options nodate nonumber ps=65 ls=72;
data media;
input alunos $ 1-14 sexo $ P1 P2 curso $;
medpond=(P1+2*P2)/3;
media=mean(of P1-P2);
if medpond >=5 then result='aprov';
else result='reprov';

```

```

cards;
Joaquim Silva M 4.9 5.1 AG
Pedro Aleixo M 3.8 2.5 AG
Ana Maria F 7.8 8.3 FL
Helena Souza F 8.4 4.3 AG
Mario Pinheiro M 6.4 7.3 FL
;
proc print;
title 'Dados para verificação';
run;

proc sort;
by medpond;
run;

proc print;
title 'Dados ordenados pela média ponderada';
run;

proc means;
title 'Estatísticas de interesse I';
run;

proc means N mean min max;
title 'Estatísticas de interesse II';
var P1 P2 media medpond;
run;

```

- Até o momento, foram apresentadas os seguintes procedimentos: PROC PRINT (visualização de resultados), PROC SORT (ordenação) e PROC MEANS (estatísticas univariadas). Muitas outras serão apresentadas oportunamente.
- A forma geral das funções do SAS tem a seguinte sintaxe:
funcname(arg1,arg2,...);
funcname(of arg1 arg2 ...);
funcname(of arg_st-arg_end);
- São possíveis combinações das funções, como por exemplo:

X=min(mean(of v1-v5), mean(of v6-v10));

Tabela 2: Algumas funções usadas para análise de dados (Estatística Descritiva).

Nome da função	Proposta
N([OF] argumentos)	Número de valores observados
NMISS([OF] argumentos)	Número de valores perdidos
CSS([OF] argumentos)	Soma de quadrados corrigida
CV([OF] argumentos)	Coefficiente de variação
KURTOSIS([OF] argumentos)	Coefficiente de curtose
SKEWNESS([OF] argumentos)	Coefficiente de assimetria
STD([OF] argumentos)	Desvio padrão
STDERR([OF] argumentos)	Desvio padrão da média
VAR([OF] argumentos)	Variância
USS([OF] argumentos)	Soma de quadrados não corrigida
SUM([OF] argumentos)	Soma dos valores
MEAN([OF] argumentos)	Média
MIN([OF] argumentos)	Menor valor
MAX([OF] argumentos)	Maior valor
RANGE([OF] argumentos)	Amplitude

V. USO DE ARRAYS

(prog4.sas)

```
data dentro;
input progenie bloco P1-P5;
MED_ALT=mean(of P1-P5);
VAR_ALT=var(of P1-P5);
cards;
 1 1 1.40 1.50 1.30 1.55 1.40
 2 1 1.30 1.50 1.55 1.70 1.40
 3 1 1.75 1.60 1.55 1.45 1.50
 4 1 1.30 1.30 1.30 1.50 1.40
 5 1 1.80 1.75 2.00 1.60 1.50
 1 2 1.35 1.90 1.65 1.45 1.80
 2 2 1.35 1.60 1.25 1.30 1.45
 3 2 1.55 1.60 1.50 1.60 1.85
 4 2 1.20 1.35 1.45 1.50 1.35
 5 2 1.55 1.50 1.45 1.50 1.45
;
proc print;
run;
proc print;
var progenie bloco MED_ALT VAR_ALT;
run;
```

IV. OUTROS TIPOS DE ARQUIVOS

Veremos agora como criar data sets permanentes e como manipular arquivos previamente gravados usando o formato PRN, ou “texto formatado (separado por espaços)”.

(prog5.sas)

```
libname local 'C:\MEUS DOCUMENTOS\';

data local.dados;
input trat$ bloco PESO;
cards;
A 1 1.26
A 2 0.99
B 1 2.37
B 2 1.17
;
proc print data=local.dados;
title 'SAS data set permanente';
run;

data arq_prn;
infile 'C:\MEUS DOCUMENTOS\PRAT1E.PRN';
input trat $ ESTACAS;
run;

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

Introduziremos agora o conceito de ARRAYS (arranjos), que facilita muito a programação no DATA STEP. Permitem a indexação de variáveis e, conseqüentemente, automatizam as operações a serem realizadas. Possui várias aplicações em situações comuns na Genética.

(prog6.sas)

```
data ntsys1;
input g1-g5;
if g1=999 then g1=.;
if g2=999 then g2=.;
if g3=999 then g3=.;
if g4=999 then g4=.;
if g5=999 then g5=.;
datalines;
999      0      0      1      0
999      1      1      0      1
999      0      0      1      0
 1      1      1      0      1
 0      1      1      0      1
 1      1      1      1      1
 1      0      0      0      0
 0      1      1      0      1
 1      0      0      0      0
 1      1      1      1      1
 0      0      999      0      0
;
proc print;
title 'Chato!';
run;

data ntsys2;
input g1-g5;
array g[5] g1-g5;
do i=1 to 5;
  if g[i]=999 then g[i]=.;
end;
drop i;
datalines;
999      0      0      1      0
999      1      1      0      1
999      0      0      1      0
 1      1      1      0      1
 0      1      1      0      1
 1      1      1      1      1
 1      0      0      0      0
 0      1      1      0      1
 1      0      0      0      0
 1      1      1      1      1
 0      0      999      0      0
;
proc print;
title 'Legal!';
run;
```

- O funcionamento das ARRAYS pode ser melhor visualizado

considerando o seguinte exemplo:

```
SUM=A+B+C;
é equivalente a
SUM=X[1]+X[2]+X[3];
```

No caso, empregou-se a operação

```
if g[i]=999 then g[i]=.;
```

sendo que o indexador i foi automaticamente incrementado com o comando “do”.

VI. CÁLCULO DE COEFICIENTES DE SIMILARIDADE

Uma vez introduzidos os conceitos de programação, ilustraremos agora como isso pode ser utilizado numa situação real, usando dados provenientes de marcadores moleculares dominantes. O programa a ser apresentado permite o cálculo de 7 diferentes coeficientes de similaridade, baseados em coincidências (ou não) de bandas. É importante notar que o mesmo pode ser facilmente alterado pelo usuário para calcular qualquer outro tipo de coeficiente de similaridade de interesse.

(prog7.sas)

```
options nodate nonumber ps=65 ls=78;
data rawdata;
%let m=40; /*número de marcadores*/
%let n=10; /*número de genótipos*/
%let x=100; /*square of n*/
%let c=1; /*1: Jaccard, 2: Sorensen-Dice,
3: Simple Matching, 4: Rogers
e Tanimoto, 5: Anderberg,
6: Russel e Rao, 7: Ochiai*/

input gen1-gen&n;
cards;
1 1 1 0 0 . 0 0 . 1
0 1 1 0 1 . 1 1 . 1
0 0 1 0 1 . 1 0 . 0
0 0 0 0 0 1 0 0 1 0
0 1 0 1 0 0 0 0 0 0
0 1 0 0 0 0 1 1 1 0
0 0 0 0 0 0 1 0 0 0
0 0 0 0 0 1 0 0 0 0
1 1 0 0 0 1 1 0 0 0
0 0 0 0 0 0 0 0 1 1
0 1 0 1 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 1 0 1 1
0 1 1 1 1 1 0 0 1 1
1 1 1 1 1 0 1 1 0 1
0 0 0 0 0 0 0 0 0 0
1 1 0 1 1 1 1 1 1 1
1 1 1 1 1 0 1 1 1 1
0 0 0 0 0 0 0 0 0 0
1 0 0 1 1 0 1 1 0 1
1 1 1 1 1 1 1 1 1 0
0 0 1 0 0 0 0 1 0 0
```

```
0 0 0 0 0 0 0 1 0 0
0 1 1 1 1 1 1 1 1 1
1 1 0 1 1 1 1 1 1 1
0 0 0 1 1 0 0 1 1 0
0 0 0 0 0 0 1 1 1 1
0 1 1 0 1 1 0 1 0 0
1 0 0 0 0 0 1 1 0 0
0 0 1 0 0 0 0 0 0 0
0 0 1 1 0 0 0 1 0 0
1 0 0 0 1 0 1 1 1 1
0 1 0 0 0 1 0 0 0 0
1 1 0 1 0 0 1 1 0 1
0 0 1 1 1 1 1 0 1 1
0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1
0 0 1 1 0 0 1 1 1 1
0 0 1 1 0 1 0 1 1 1
0 0 0 1 1 0 1 1 1 0
;
data simil;
set rawdata;
array gl[&n] gen1-gen&n;
array a[&n,&n] a1-a&x;
array b[&n,&n] b1-b&x;
array c[&n,&n] c1-c&x;
array d[&n,&n] d1-d&x;
do i=1 to &n;
do j=1 to &n;
if gl[i]=1 and gl[j]=1 then a[i,j]=1;
else a[i,j]=0;
if gl[i]=1 and gl[j]=0 then b[i,j]=1;
else b[i,j]=0;
if gl[i]=0 and gl[j]=1 then c[i,j]=1;
else c[i,j]=0;
if gl[i]=0 and gl[j]=0 then d[i,j]=1;
else d[i,j]=0;
end;
end;
drop i j;
run;
proc means data=simil noprint;
output out=simil1 sum=;
run;
data simil2;
set simil1;
array a[&n,&n] a1-a&x;
array b[&n,&n] b1-b&x;
array c[&n,&n] c1-c&x;
array d[&n,&n] d1-d&x;
array S[&n,&n] S1-S&x;
do i=1 to &n;
do j=1 to &n;
if &c=1 then
S[i,j]=(a[i,j])/(a[i,j]+b[i,j]+c[i,j]);
if &c=2 then
S[i,j]=(2*a[i,j])/
(2*a[i,j] + b[i,j] + c[i,j]);
if &c=3 then
S[i,j]=(a[i,j]+d[i,j])/
(a[i,j]+b[i,j]+c[i,j]+d[i,j]);
if &c=4 then
S[i,j]=(a[i,j]+d[i,j])/
```

```

(a[i,j]+(2*b[i,j]+2*c[i,j]+d[i,j]));
if &c=5 then
S[i,j]=(a[i,j])/(a[i,j]+2*(b[i,j]+c[i,j]));
if &c=6 then
S[i,j]=(a[i,j])/
(a[i,j]+b[i,j]+c[i,j]+d[i,j]);
if &c=7 then
S[i,j]=(a[i,j])/
(((a[i,j]+b[i,j])*(a[i,j]+c[i,j]))*0.5);
end;
end;
keep S1-S&x;
proc print;
title;
run;

proc iml;
reset autaname;
use simil2;
read all into FINAL;
FINAL2=shape(FINAL,&n);
print FINAL2 [format=6.4];
quit;

```

- Nota: esse programa inclui variáveis macro (criadas usando o comando %let), que podem ser usadas em qualquer parte do programa, sendo chamadas com o caráter &.
- No final, usou-se o PROC IML, que será abordado posteriormente, para formatação dos resultados. No caso, o data set simil2 foi convertido na matriz com nome FINAL, que foi posteriormente formatada com o comando shape.

VII. MAPAS GENÉTICOS

O programa a seguir exemplifica como o SAS pode ser usado para testar segregações de locos marcadores genéticos, bem como testar se há dependência entre eles (causada pela ligação genética), com empregado da PROC FREQ. Em seguida, foram estimadas as frequências de recombinação entre os locos, pelo método da máxima verossimilhança.

(prog8.sas)

```

options nodate nonumber ps=65 ls=78;
data mouse;
%let m=14; /*Número de Marcadores*/
%let n=196; /*Quadrado de m*/
%let g=103; /*Número de genótipos*/
input genot M1-M&m;
datalines;
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 1 1 1 1 1 1 1 1 1 1 1 1 1 0
3 0 1 1 1 1 1 1 1 1 1 1 1 1 1
4 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5 1 1 1 1 1 1 1 1 1 1 1 1 1 1
6 0 0 0 0 0 0 0 0 0 0 0 0 0 0
7 0 0 0 0 0 0 0 1 1 1 1 1 1 1
8 1 1 1 1 1 1 1 1 1 1 1 1 1 0

```

```

9 0 0 0 0 0 0 0 0 0 0 0 0 0 0
10 0 0 0 0 0 0 0 0 0 0 0 0 0 0
11 0 0 0 0 1 1 1 1 1 1 1 1 1 1
12 1 1 1 1 1 1 1 1 1 1 1 1 0 0
13 1 1 1 1 1 1 1 1 1 1 1 1 1 0
14 0 0 0 0 0 0 0 0 0 0 0 0 0 0
15 1 1 1 1 1 1 1 1 1 1 1 1 1 1
16 1 1 1 1 1 1 1 1 1 1 1 0 1 1
17 0 0 0 0 0 1 1 1 1 1 1 1 1 1
18 0 0 0 0 0 0 0 0 0 0 0 0 0 0
19 0 0 0 0 0 0 0 0 0 0 0 0 0 0
20 0 0 0 0 0 0 0 0 0 0 0 0 0 0
21 0 0 0 0 0 0 0 0 0 0 0 0 0 0
22 1 1 1 1 1 1 1 1 1 1 1 1 1 1
23 0 0 0 0 0 1 1 1 1 1 1 1 1 1
24 0 0 0 0 0 0 0 0 0 0 0 0 0 1
25 1 1 1 1 1 1 0 0 0 0 0 0 0 0
26 0 0 0 0 0 0 0 0 0 0 0 0 0 0
27 1 1 1 1 1 1 1 1 1 1 1 1 1 1
28 1 1 1 1 1 1 1 1 1 1 1 1 1 1
29 1 1 1 1 1 1 1 1 1 1 1 1 1 1
30 1 1 1 1 1 1 1 1 1 1 1 1 1 1
31 0 0 0 0 0 0 0 0 0 0 0 0 0 0
32 0 0 0 0 0 0 0 0 1 1 1 1 1 1
33 0 0 0 0 0 0 0 0 0 0 1 1 1 1
34 1 1 1 1 1 1 1 1 1 0 0 0 0 0
35 1 1 1 1 1 1 1 1 1 1 1 1 1 1
36 0 0 0 0 0 0 0 0 0 0 0 0 0 0
37 0 1 1 1 1 1 1 1 1 1 1 1 1 0
38 0 0 0 0 0 0 0 0 0 0 0 0 0 0
39 0 0 0 0 0 0 0 0 0 0 0 0 0 0
40 1 1 1 1 1 1 1 1 1 1 1 1 1 1
41 1 1 1 1 1 1 1 1 1 1 1 1 1 1
42 0 0 0 0 0 0 0 1 1 1 1 1 1 1
43 1 1 1 1 1 1 1 1 1 1 1 1 1 1
44 0 0 0 0 0 0 0 0 0 0 0 0 0 0
45 1 1 1 1 1 1 1 1 1 1 1 1 1 1
46 0 1 1 1 1 1 1 1 1 1 1 1 1 1
47 0 0 0 0 0 0 0 0 0 0 0 0 0 0
48 0 0 0 0 0 0 0 0 1 1 1 1 1 1
49 0 0 0 0 0 0 0 0 0 0 0 0 0 0
50 1 1 1 1 1 1 1 1 1 1 1 1 1 1
51 1 1 1 1 1 1 1 1 1 1 1 1 1 0
52 0 0 0 0 0 0 0 0 0 0 0 0 0 0
53 1 1 1 1 1 1 1 1 1 1 1 1 1 1
54 1 1 1 1 1 1 1 1 1 1 1 1 1 1
55 0 0 1 1 1 1 1 1 1 1 1 1 1 1
56 0 0 0 0 0 0 0 0 0 0 0 0 0 0
57 0 0 0 0 0 0 0 0 0 1 1 1 1 1
58 0 0 0 0 0 0 0 0 0 0 0 0 0 0
59 0 0 0 0 1 1 1 1 1 1 0 1 1 0
60 0 0 0 0 0 0 0 0 0 0 0 0 0 0
61 0 0 0 0 0 0 0 0 0 0 0 0 0 0
62 0 0 0 0 0 0 0 0 0 0 0 0 0 0
63 0 0 0 0 0 0 0 0 0 0 0 0 0 0
64 0 0 0 0 0 0 0 0 0 0 0 0 0 1
65 0 1 1 1 1 1 1 1 1 1 1 1 1 0
66 0 0 0 0 0 0 0 0 0 0 0 0 0 0
67 0 0 0 0 0 0 0 0 0 0 0 0 0 1
68 0 0 0 0 0 0 0 0 0 0 0 0 0 1
69 0 0 0 0 0 0 0 0 0 0 0 1 1 1
70 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

71 1 1 1 1 1 1 1 1 1 1 1 1 1 1
72 0 0 0 0 0 0 0 0 0 0 0 0 0 0
73 1 0 0 0 0 0 0 0 0 0 0 0 0 0
74 1 1 1 1 1 1 1 1 1 1 1 1 1 1
75 1 1 1 1 1 1 1 1 1 1 1 1 1 0
76 1 1 1 1 1 1 1 1 1 1 1 1 1 1
77 1 1 1 1 1 1 1 1 1 1 1 1 1 1
78 1 1 1 1 1 1 1 1 1 1 1 1 1 1
79 1 1 1 1 1 1 0 0 0 0 0 0 0 0
80 1 1 1 1 1 1 1 1 1 1 1 1 1 1
81 1 0 0 0 0 0 0 0 0 0 0 0 1 1
82 0 0 0 0 0 0 0 0 0 0 0 0 0 1
83 1 1 1 1 1 1 1 1 1 1 1 1 1 1
84 0 0 0 0 0 0 0 0 1 1 1 1 1 1
85 1 1 1 1 1 1 1 1 1 1 1 1 1 1
86 0 0 0 0 0 0 0 0 0 0 0 0 0 0
87 0 0 0 0 0 0 0 0 0 0 0 0 0 0
88 0 0 0 0 0 0 0 0 0 0 0 0 0 0
89 0 0 0 0 0 0 0 0 0 0 0 0 0 0
90 0 0 0 0 0 0 0 0 0 0 0 0 0 1
91 0 0 0 0 0 0 0 0 0 0 0 0 0 0
92 0 0 0 0 0 0 0 0 0 0 0 0 0 0
93 1 1 1 1 1 0 0 0 0 0 0 0 0 0
94 0 0 0 0 0 0 0 0 0 0 0 0 0 0
95 0 0 0 0 0 0 0 0 1 1 1 1 1 0
96 0 0 0 0 0 0 0 1 1 1 1 1 1 1
97 0 0 0 0 0 0 0 0 0 0 0 0 0 0
98 0 0 0 0 0 0 0 0 0 0 0 0 0 0
99 0 0 0 0 0 0 0 0 0 0 0 0 0 0
100 1 1 1 1 0 0 0 0 0 0 0 0 0 0
101 1 1 1 0 0 0 0 0 0 0 0 0 0 0
102 0 0 0 0 0 0 0 0 0 0 0 0 0 1
103 1 0 0 0 0 0 0 0 0 0 0 0 0 0
;
proc print;
title 'Dados para verificação';
run;

proc freq;
title 'Frequências de cada genótipo -
Retrocruzamento';
table M1-M&m / TESTP=(0.50 0.50) NOCUM;
run;
proc freq;
title 'Tabelas de contingência';
table (M1-M&m)*(M1-M&m) / CHISQ;
run;

data ligacao;
set mouse;
array M[&m] M1-M&m;
array rec[&m,&m] rec1-rec&n;
do i=1 to &m;
do j=1 to &m;
if M[i]=1 and M[j]=1 then rec[i,j]=0;
if M[i]=0 and M[j]=0 then rec[i,j]=0;
if M[i]=1 and M[j]=0 then rec[i,j]=1;
if M[i]=0 and M[j]=1 then rec[i,j]=1;
end;
end;
run;
proc means noprint;

```

```

title;
var rec1-rec&n;
output out=recs sum=;
run;
data r;
set recs;
array rec[&m,&m] rec1-rec&n;
array r[&m,&m] r1-r&n;
do i=1 to &m;
do j=1 to &m;
r[i,j]=rec[i,j]/&g;
end;
end;
keep r1-r&n;
run;
proc iml;
use r;
read all into matrix;
recomb=shape(matrix,&m);
print 'Estimated pairwise recombination
frequency';
print recomb [format=4.2];
quit;

```

- Essas estimativas podem ser usadas para determinar a ordem dos marcadores dentro dos grupos de ligação.

VIII. BOOTSTRAP

O fato do DATA STEP do SAS ser também um linguagem de programação permite que o pacote seja usado para análises de interesse do usuário que não sejam contempladas diretamente em nenhuma proc. Como exemplo, apresentamos abaixo um programa para obtenção de amostras com reposição (bootstrap).

(prog9.sas)

```

/*****
/* Forneça os valores corretos de B e n */
*****/

/*INDIQUE O NÚMERO DE INDIVÍDUOS */
%let n=7;
/*INDIQUE O NÚMERO DE BOOTSTRAPS*/
%let B=100;

data dados;
input genotipo $ fenotipo;
cards;
aa 10
aa 12
aa 11
aa 15
aa 16
aa 10
aa 9
;
data boot;

```

```

do amostra=1 to &B;
  do i=1 to &n;
    select=ceil(ranuni(12345)*&n);
    set dados point=select nobs=n;
    position=select;
    output;
  end;
end;
stop;
run;
proc print data=boot;
run;

proc means mean;
  by amostra;
  var fenotipo;
  output out=medboot mean=media;
run;

proc univariate data=medboot;
var media;
run;

```

IX. ESTATÍSTICA DESCRITIVA

Veremos a seguir com obter estatísticas descritivas, usando como exemplo dados de altura de espigas de 100 plantas de milho (fonte: Geraldi, 2002).

(prog10.sas)

```

data altura;
input AE@@;
cards;
154 191 152 145 156
166 169 144 172 154
190 175 140 170 144
171 167 159 165 155
174 156 162 158 169
168 160 165 175 171
182 146 184 164 148
161 151 169 148 190
160 160 156 162 156
166 174 180 164 159
190 186 148 162 195
180 184 170 166 189
194 181 135 172 162
176 179 145 174 154
194 174 165 175 164
168 171 146 165 178
158 156 169 161 184
180 164 150 180 165
179 168 161 160 161
172 154 160 168 186
;
proc print;
title 'Dados para verificação';
run;
proc univariate normal plot;

```

```

title 'Estatística descritiva';
run;
proc chart;
title 'Histograma';
  vbar AE;
run;

```

- Note as diferenças da PROC UNIVARIATE em relação à PROC MEANS.

X. EXPERIMENTOS INTEIRAMENTE AO ACASO

Os dados abaixo referem-se a um experimento inteiramente ao acaso, feito com o objetivo de comparar 12 linhagens de fungo quanto ao diâmetro das colônias.

- Nota: os nomes escritos em letras maiúsculas foram criados pelo usuário e podem ser alterados.

A. Modelo matemático

$$y_{ij} = \mu + t_i + e_{ij} \quad (1)$$

B. Análise inicial

(prog11.sas)

```

options nodate nonumber ps=65 ls=78;
data FUNGOS;
input TRAT DIAM @@;
cards;
1 0.31 2 0.82 3 0.43 4 0.45 5 0.29 6 0.92
7 0.44 8 0.56 9 0.22 10 0.30 11 0.23 12 0.30
1 0.45 2 1.10 3 0.45 4 0.71 5 0.40 6 0.61
7 0.35 8 1.02 9 0.21 10 0.37 11 0.25 12 0.36
1 0.46 2 0.88 3 0.63 4 0.66 5 0.23 6 0.49
7 0.31 8 0.71 9 0.18 10 0.38 11 0.24 12 0.31
1 0.43 2 0.72 3 0.76 4 0.62 5 0.36 6 1.24
7 0.40 8 0.38 9 0.23 10 0.29 11 0.22 12 0.33
;
proc print;
title 'Dados para verificação';
run;
proc plot;
title 'Análise gráfica';
  plot DIAM*TRAT;
run;
proc glm;
title 'Análise de variância';
  class TRAT;
  model DIAM = TRAT;
run;

```

C. Verificação das pressuposições do modelo

```

proc glm noprint data=FUNGOS;
title 'Obtenção dos resíduos';

```



```

class TRAT;
model DIAM = TRAT;
output out=RESID residual=RESIDUO;
run;
proc print data=RESID;
title 'Resíduos';
run;

proc glm noprint data=FUNGOS;
title 'Obtenção dos resíduos padronizados';
class TRAT;
model DIAM = TRAT;
output out=RESPAD student=RES_PAD;
run;
proc print data=RESPAD;
title 'Resíduos Padronizados';
run;

proc glm noprint data=FUNGOS;
class TRAT;
model DIAM = TRAT;
output out=RESIDUOS predicted=PREDITOS
student=RES_PAD;
run;
proc print data=RESIDUOS;
title 'Res. pad. e ap. de mín. quad.';
run;

/*Análise dos resíduos - pressuposições*/

/*Teste do F máximo*/
proc sort data=RESIDUOS;
by trat;
run;
proc means noprint data=RESIDUOS;
var DIAM;
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;
proc print data=HARTLEY;
title 'Teste do F máximo';
var VMIN VMAX FMAX;
run;

/*Normalidade dos resíduos*/
proc univariate data=RESIDUOS normal plot;
title 'Box-plot e normalidade dos residuos';
var RES_PAD;
run;

/*Independência dos resíduos*/

```

```

proc plot data=RESIDUOS;
title 'Gráfico: preditos vs res. pad.';
plot RES_PAD*PREDITOS;
run;

/*Escolha da transformação*/
data REGR;
set VAR_MED;
LMEDIA=log(MEDIA);
LVAR=log(VARIANC);
run;
proc reg data=REGR;
title 'Regressão log(media), log(var) - Escolha da transf.';
model LVAR=LMEDIA;
run;

D. Análise dos dados transformados

/*Transformação dos dados*/
data TRANSF;
set FUNGOS;
DIAM_T=(DIAM)**(-1);
run;

/*Verificação da validade da transformação*/
proc glm data=TRANSF noprint;
title 'Verificação da efetividade da transformação';
class TRAT;
model DIAM_T = TRAT;
output out=RESIDUO2 predicted=PREDITOS
student=RES_PAD;
run;
proc sort data=RESIDUO2;
by TRAT;
run;
proc means noprint data=RESIDUO2;
var DIAM_T;
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;
proc print data=hartley;
var VMIN VMAX FMAX;
run;
proc univariate data=RESIDUO2 normal plot;
var RES_PAD;
run;
proc plot data=RESIDUO2;
plot RES_PAD*PREDITOS;
run;

```

```
proc glm data=TRANSF;
title 'Análise de variância dos dados
transformados';
class TRAT;
model DIAM_T = TRAT;
means TRAT/tukey;
run;
```

Tabela 3: Transformação estabilizadora da variância (Fonte: Montgomery, 1991).

Relação entre médias e variâncias	α	$\lambda = 1 - \frac{\alpha}{2}$	Transformação $y^* = y^\lambda$
$\sigma_y^2 \propto cte$	0	1	nenhuma
$\sigma_y^2 \propto \mu^1$	1	1/2	$y^* = \sqrt{y}$
$\sigma_y^2 \propto \mu^2$	2	0	$y^* = \log y$
$\sigma_y^2 \propto \mu^3$	3	-1/2	$y^* = 1/(\sqrt{y})$
$\sigma_y^2 \propto \mu^4$	4	-1	$y^* = 1/y$
-	-	-	$y^* = \arcsen \sqrt{y}$

XI. EXPERIMENTOS EM BLOCOS AO ACASO

Exemplo: comparar 20 progenies (linhas puras) aleatórias de soja, avaliadas em 2 locais (fonte: Geraldi, 2002).

A. Modelo matemático

$$y_{ij} = \mu + p_i + b_j + e_{ij} \quad (2)$$

B. Programa

(prog12.sas)

```
data blocos;
input locais bloco progenie PRODUCAO;
cards;
1      1      1      251
1      1      2      198
1      1      3      186
1      1      4      121
1      1      5      159
1      1      6      141
1      1      7      232
1      1      8      161
1      1      9      277
1      1     10      144
1      1     11      153
1      1     12      195
```

1	1	13	159
1	1	14	203
1	1	15	204
1	1	16	166
1	1	17	185
1	1	18	183
1	1	19	214
1	1	20	141
1	2	1	192
1	2	2	200
1	2	3	156
1	2	4	147
1	2	5	173
1	2	6	167
1	2	7	242
1	2	8	168
1	2	9	253
1	2	10	142
1	2	11	142
1	2	12	176
1	2	13	155
1	2	14	198
1	2	15	187
1	2	16	142
1	2	17	199
1	2	18	149
1	2	19	198
1	2	20	141
1	3	1	262
1	3	2	153
1	3	3	182
1	3	4	181
1	3	5	142
1	3	6	169
1	3	7	160
1	3	8	155
1	3	9	289
1	3	10	137
1	3	11	149
1	3	12	277
1	3	13	157
1	3	14	183
1	3	15	199
1	3	16	163
1	3	17	183
1	3	18	151
1	3	19	186
1	3	20	176
2	1	1	218
2	1	2	150
2	1	3	206
2	1	4	162
2	1	5	176
2	1	6	166
2	1	7	180
2	1	8	204
2	1	9	332
2	1	10	270
2	1	11	158
2	1	12	221
2	1	13	224
2	1	14	282

2	1	15	248
2	1	16	203
2	1	17	147
2	1	18	211
2	1	19	220
2	1	20	136
2	2	1	227
2	2	2	153
2	2	3	226
2	2	4	172
2	2	5	263
2	2	6	147
2	2	7	194
2	2	8	229
2	2	9	243
2	2	10	207
2	2	11	163
2	2	12	204
2	2	13	259
2	2	14	279
2	2	15	224
2	2	16	297
2	2	17	223
2	2	18	226
2	2	19	245
2	2	20	155
2	3	1	316
2	3	2	172
2	3	3	234
2	3	4	183
2	3	5	182
2	3	6	204
2	3	7	220
2	3	8	232
2	3	9	273
2	3	10	278
2	3	11	189
2	3	12	205
2	3	13	225
2	3	14	361
2	3	15	184
2	3	16	293
2	3	17	154
2	3	18	258
2	3	19	218
2	3	20	218

```

;
proc print;
title 'Dados para verificação';
run;
proc sort; /*IMPORTANTE!*/
by locais;
run;
proc glm;
title 'Anavas individuais I';
by locais;
class bloco progenie;
model PRODUCAO = progenie bloco;
run;
proc glm;
title 'Identificando o resíduo';

```

```

by locais;
class bloco progenie;
model PRODUCAO = progenie bloco
progenie*bloco;
run;
proc glm;
title 'Anavas individuais II - modelo misto';
by locais;
class bloco progenie;
model PRODUCAO = progenie bloco;
random progenie bloco/test;
run;
proc mixed;
title 'Anavas com PROC MIXED';
by locais;
class bloco progenie;
model PRODUCAO =;
random progenie bloco;
run;

/*Análise conjunta - Modelo misto*/
proc glm;
title 'Análise conjunta';
class progenie bloco locais;
model PRODUCAO=locais bloco(locais) progenie
progenie*locais;
random bloco(locais) progenie
progenie*locais / test;
lsmeans progenie*locais / slice=locais;
lsmeans progenie; /*BLUE's*/
run;
proc mixed;
title 'Análise conjunta com PROC MIXED';
class progenie bloco locais;
model PRODUCAO=locais;
random bloco(locais) progenie
progenie*locais / solution; /*EBLUP's*/
run;

```

XII. EXPERIMENTOS EM BLOCOS AO ACASO COM DIFERENTES TIPOS DE TRATAMENTOS

No presente caso, temos 15 tratamentos (milho), sendo que 5 deles são variedades e 10 são híbridos. Deseja-se testar se há diferença dentro de cada tipo e se os tipos são diferentes entre si (Fonte: Geraldi, 2002).

A. Modelo matemático

$$y_{ij} = \mu + t_i + b_j + e_{ij} \quad (3)$$

B. Programa

(prog13.sas)

```
data blocos;
```

```

input trat bloco PG tipo $;
cards;
  1      1      6.76      V
  2      1      6.91      V
  3      1      6.41      V
  4      1      5.63      V
  5      1      5.79      V
  6      1      7.50      H
  7      1      8.37      H
  8      1      6.80      H
  9      1      8.54      H
10      1      7.59      H
11      1      7.12      H
12      1      8.22      H
13      1      7.61      H
14      1      7.72      H
15      1      6.87      H
  1      2      6.51      V
  2      2      5.71      V
  3      2      6.06      V
  4      2      7.06      V
  5      2      5.34      V
  6      2      7.03      H
  7      2      7.52      H
  8      2      9.05      H
  9      2      7.26      H
10      2      9.03      H
11      2      8.63      H
12      2      8.27      H
13      2      7.35      H
14      2      7.87      H
15      2      7.18      H
;

```

```

proc glm;
title 'Análise de variância e teste de Tukey';
  class trat bloco;
  model PG = trat bloco;
  means trat/tukey;
run;

```

```

proc glm;
title 'Anava com decomposição dos GLs';
  class trat bloco tipo;
  model PG = bloco tipo trat(tipo);
  lsmeans trat(tipo) / slice=tipo;
run;

```

```

proc glm;
title 'Anava com decomposição dos GLs -
  contrastes';
  class trat bloco;
  model PG = trat bloco;
  contrast 'Varied vs Hib'
    trat 2 2 2 2 2
      -1 -1 -1 -1 -1 -1 -1 -1 -1 -1;
  contrast 'Dentro Varied'
    trat -3 -3 2 2 2,
    trat 1 -1 0 0 0,
    trat 0 0 -2 1 1,
    trat 0 0 0 -1 1;
  contrast 'Dentro Hibrid'

```

```

  trat 0 0 0 0 0
    -1 -1 -1 -1 -1 1 1 1 1 1,
  trat 0 0 0 0 0
    -3 -3 2 2 2 0 0 0 0 0,
  trat 0 0 0 0 0
    -1 1 0 0 0 0 0 0 0 0,
  trat 0 0 0 0 0
    0 0 -2 1 1 0 0 0 0 0,
  trat 0 0 0 0 0
    0 0 0 -1 1 0 0 0 0 0,
  trat 0 0 0 0 0
    0 0 0 0 0 -3 -3 2 2 2,
  trat 0 0 0 0 0
    0 0 0 0 0 -1 1 0 0 0,
  trat 0 0 0 0 0
    0 0 0 0 0 0 0 -2 1 1,
  trat 0 0 0 0 0
    0 0 0 0 0 0 0 0 -1 1;
run;

```

XIII. EXPERIMENTOS EM BLOCOS AO ACASO COM OBSERVAÇÕES DENTRO DAS PARCELAS

Os dados seguintes referem-se à altura da espiga, anotados em 25 famílias de meios-irmãos de milho. O delineamento foi blocos casualizados. Fonte: Vencovsky e Barriga (1992).

A. Modelo matemático

$$y_{ij} = \mu + t_i + b_j + p_{k(ij)} + e_{ijk} \quad (4)$$

B. Programa

(prog14.sas)

```

data dentro;
input progenie bloco@;
do plantas=1 to 5;
  input PRODUCAO@;
  output;
end;
cards;
  1 1 1.40 1.50 1.30 1.55 1.40
  2 1 1.30 1.50 1.55 1.70 1.40
  3 1 1.75 1.60 1.55 1.45 1.50
  4 1 1.30 1.30 1.30 1.50 1.40
  5 1 1.80 1.75 2.00 1.60 1.50
  6 1 1.40 1.50 1.75 1.60 1.55
  7 1 1.20 1.25 1.45 1.45 1.60
  8 1 1.45 1.60 1.60 1.60 1.70
  9 1 1.40 1.35 1.45 1.40 1.25
10 1 1.55 1.50 1.65 1.55 1.70
11 1 1.30 1.15 1.30 1.60 1.60
12 1 1.70 1.30 1.80 1.60 1.70
13 1 1.50 1.40 1.40 1.35 1.20
14 1 1.80 1.80 1.55 1.50 1.40

```

```

15 1 1.45 1.50 1.45 1.30 1.40
16 1 1.60 1.45 1.35 1.40 1.45
17 1 1.55 1.30 1.60 1.60 1.40
18 1 1.70 1.25 1.30 1.50 1.45
19 1 1.35 1.45 1.50 1.50 1.30
20 1 1.60 1.25 1.40 1.50 1.45
21 1 1.70 1.70 1.35 1.45 1.25
22 1 1.55 1.30 1.15 1.35 1.30
23 1 1.60 1.45 1.30 1.65 1.60
24 1 1.40 1.70 1.50 1.15 1.30
25 1 1.55 1.25 1.45 1.45 1.20
  1 2 1.35 1.90 1.65 1.45 1.80
  2 2 1.35 1.60 1.25 1.30 1.45
  3 2 1.55 1.60 1.50 1.60 1.85
  4 2 1.20 1.35 1.45 1.50 1.35
  5 2 1.55 1.50 1.45 1.50 1.45
  6 2 1.85 1.55 1.75 1.50 1.65
  7 2 1.25 1.50 1.50 1.65 1.30
  8 2 1.75 1.45 1.55 1.70 1.30
  9 2 1.35 1.55 1.20 1.30 1.20
10 2 1.60 1.50 1.60 1.65 1.95
11 2 1.60 1.90 1.55 1.90 1.80
12 2 1.30 1.30 1.55 1.40 1.45
13 2 1.65 1.40 1.35 1.60 1.20
14 2 1.60 1.60 1.45 1.55 1.75
15 2 1.35 1.20 1.50 1.75 1.00
16 2 1.50 1.30 1.45 1.50 1.60
17 2 1.50 1.35 1.35 1.45 1.25
18 2 1.60 1.75 1.75 1.50 1.80
19 2 1.45 1.65 1.55 1.45 1.40
20 2 1.65 1.35 1.25 1.45 1.65
21 2 1.50 1.70 1.60 1.55 1.60
22 2 1.40 1.30 1.10 1.40 1.40
23 2 1.50 1.65 1.40 1.80 1.15
24 2 1.75 1.40 1.50 1.50 1.45
25 2 1.70 1.55 1.30 1.30 1.60
;

proc glm;
title 'Análise de variância I';
  class bloco progenie;
  model PRODUCAO = progenie bloco
    progenie*bloco;
  random progenie bloco progenie*bloco/test;
run;
proc mixed;
title 'Análise de variância II';
  class bloco progenie;
  model PRODUCAO =;
  random progenie bloco progenie*bloco;
run;

C. Componentes de variância

proc varcomp method=type1;
title 'Método dos momentos';
  class bloco progenie;
  model PRODUCAO = progenie bloco
    progenie*bloco;
run;
proc varcomp method=mivque0;
title 'Estimador não-viesado de variância

```

```

mínima';
class bloco progenie;
model PRODUCAO = progenie bloco
  progenie*bloco;
run;
proc varcomp method=ml;
title 'Estimador de máxima verossimilhança';
  class bloco progenie;
  model PRODUCAO = progenie bloco
    progenie*bloco;
run;
proc varcomp method=reml;
title 'Estimador de máx. veros. restrita';
  class bloco progenie;
  model PRODUCAO = progenie bloco
    progenie*bloco;
run;

```

XIV. EXPERIMENTOS EM BLOCOS AO ACASO COM PARCELAS PERDIDAS

Estudaremos agora um exemplo de experimento em Blocos ao Acaso com perda de uma parcela. Esse experimento é discutido em detalhes em Barbin (1994) e foi conduzido com o objetivo de comparar 9 porta-enxertos para a laranjeira Valência. Foi simulada a perda de uma parcela, sendo que seu valor foi estimado como sendo 193, de tal forma a minimizar o resíduo.

A. Modelo matemático

$$y_{ij} = \mu + t_i + b_j + e_{ij} \quad (5)$$

B. Programa

```

(prog15.sas)

options nodate nonumber ps=62 ls=78;
data citros1;
do trat=1 to 9;
  do bloco=1 to 3;
    input PRODUCAO@;
    output;
  end;
end;
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=citros1;
title 'Dados para verificação';
run;

title1 'Análise de variância - teste de
Tukey';

```

```

title2 '(atenção para a ordem dos efeitos no
modelo)';
proc glm data=citros1;
  class trat bloco;
  model PRODUCAO=trat bloco;
  means trat/tukey lines;
run;

title1 'Análise de variância - teste de
Tukey';
title2 '(atenção para a ordem dos efeitos no
modelo)';
proc glm data=citros1;
  class trat bloco;
  model PRODUCAO=bloco trat;
  means trat/tukey lines;
  lsmeans trat;
run;

title1 'Análise de variância com valor da
parc. perd.';
title2 '(cuidado: SQtrat não está corrigida)';
data citros2;
set citros1;
if PRODUCAO=. then PRODUCAO=193;
run;

proc glm data=citros2;
  class trat bloco;
  model PRODUCAO=bloco trat;
  means trat/tukey;
run;

```

XV. ENSAIOS FATORIAIS

Exemplo: Produção de matéria seca (em miligramas) de 10 linhagens de *A.nidulans*, em meio sem e com ácido nicotínico. Delineamento em blocos ao acaso com duas repetições. Fonte: Geraldi, 2002.

(prog16.sas)

```

options nodate nonumber ps=62 ls=78;
data aula6;
length linhagem $ 14 acido $ 9;
input linhagem $ acido $ bloco MS;
cards;
MSE_(Original)      Sem_acido      1      47.2
MSE_(Original)      Sem_acido      2      38.2
MSE_(Original)      Com_acido      1     140.0
MSE_(Original)      Com_acido      2     152.0
MSE_su_1_nic_8      Sem_acido      1      37.6
MSE_su_1_nic_8      Sem_acido      2      44.4
MSE_su_1_nic_8      Com_acido      1      79.6
MSE_su_1_nic_8      Com_acido      2      74.0
MSE_su_2_nic_8      Sem_acido      1     103.2
MSE_su_2_nic_8      Sem_acido      2     106.0
MSE_su_2_nic_8      Com_acido      1     102.6

```

```

MSE_su_2_nic_8      Com_acido      2     112.0
MSE_su_3_nic_8      Sem_acido      1     163.6
MSE_su_3_nic_8      Sem_acido      2     176.0
MSE_su_3_nic_8      Com_acido      1     120.8
MSE_su_3_nic_8      Com_acido      2     128.2
MSE_su_4_nic_8      Sem_acido      1     133.0
MSE_su_4_nic_8      Sem_acido      2     125.0
MSE_su_4_nic_8      Com_acido      1     147.3
MSE_su_4_nic_8      Com_acido      2     161.9
MSE_su_5_nic_8      Sem_acido      1     103.6
MSE_su_5_nic_8      Sem_acido      2      98.4
MSE_su_5_nic_8      Com_acido      1     168.2
MSE_su_5_nic_8      Com_acido      2     160.2
MSE_su_6_nic_8      Sem_acido      1      85.3
MSE_su_6_nic_8      Sem_acido      2      92.7
MSE_su_6_nic_8      Com_acido      1      93.5
MSE_su_6_nic_8      Com_acido      2      84.9
MSE_su_7_nic_8      Sem_acido      1      56.8
MSE_su_7_nic_8      Sem_acido      2      46.8
MSE_su_7_nic_8      Com_acido      1      87.3
MSE_su_7_nic_8      Com_acido      2      93.1
MSE_su_8_nic_8      Sem_acido      1     110.6
MSE_su_8_nic_8      Sem_acido      2     119.2
MSE_su_8_nic_8      Com_acido      1     162.6
MSE_su_8_nic_8      Com_acido      2     177.8
MSE_su_9_nic_8      Sem_acido      1      36.2
MSE_su_9_nic_8      Sem_acido      2      37.8
MSE_su_9_nic_8      Com_acido      1      73.0
MSE_su_9_nic_8      Com_acido      2      74.0
;

```

```

proc glm;
title 'Análise de variância';
class linhagem acido bloco;
model MS = bloco linhagem acido
linhagem*acido;
* model MS = bloco linhagem|acido;
run;

```

```

proc glm;
title 'Análise de variância na presença de
interação significativa I';
class linhagem acido bloco;
model MS = bloco acido linhagem(acido);
run;

```

```

proc glm;
title 'Análise de variância na presença de
interação significativa II';
class linhagem acido bloco;
model MS = bloco linhagem acido(linhagem);
run;

```

```

proc glm;
title 'Análise de variância na presença de
interação significativa III';
class linhagem acido bloco;
model MS = bloco linhagem|acido;
lsmeans acido*linhagem / slice=linhagem;
run;

```

```

proc glm;

```

```

title 'Análise de variância na presença de
interação significativa IV';
class linhagem acido bloco;
model MS = bloco linhagem|acido;
lsmeans acido*linhagem / slice=acido
adjust=tukey PDIFF;
run;

```

XVI. MODELOS COM CLASSIFICAÇÃO HIERÁRQUICA

Exemplo: Peso de galinhas (em gramas) com 20 semanas de idade. Fonte: Geraldi (2002).

(prog17.sas)

```

options nodate nonumber;
data hierarq1;
input machos femeas descend PESO;
cards;
1      1      1      1550
1      1      2      1500
1      1      3      1550
1      1      4      1600
1      2      1      1700
1      2      2      1600
1      2      3      1300
1      2      4      1500
1      3      1      1600
1      3      2      1500
1      3      3      1550
1      3      4      1650
1      4      1      1350
1      4      2      1700
1      4      3      1500
1      4      4      2050
2      1      1      1550
2      1      2      1650
2      1      3      1500
2      1      4      1550
2      2      1      1400
2      2      2      1400
2      2      3      1350
2      2      4      1300
2      3      1      1550
2      3      2      1450
2      3      3      1650
2      3      4      1650
2      4      1      1300
2      4      2      1350
2      4      3      1500
2      4      4      1500
3      1      1      1650
3      1      2      1600
3      1      3      1600
3      1      4      1550
3      2      1      1300

```

```

3      2      2      1600
3      2      3      1250
3      2      4      1500
3      3      1      1250
3      3      2      1300
3      3      3      1350
3      3      4      1400
3      4      1      1700
3      4      2      1500
3      4      3      1700
3      4      4      1350
4      1      1      1650
4      1      2      1350
4      1      3      1250
4      1      4      1150
4      2      1      1350
4      2      2      1150
4      2      3      1250
4      2      4      1350
4      3      1      1300
4      3      2      1200
4      3      3      1300
4      3      4      1250
4      4      1      1250
4      4      2      1450
4      4      3      1450
4      4      4      1400
;

```

```

proc glm;
title 'Sem residuo';
class machos femeas descend;
model PESO = machos femeas(machos)
descend(machos*femeas);
test h=machos e=femeas(machos);
test h=femeas(machos)
e=descend(machos*femeas);
run;

```

```

proc glm;
title 'Com residuo';
class machos femeas descend;
model PESO = machos femeas(machos);
test h=machos e=femeas(machos);
run;

```

XVII. ENSAIOS EM PARCELAS SUBDIVIDIDAS

Apresentamos a seguir os dados referentes à porcentagem de BRIX no caldo de cana-de-açúcar, num experimento de aplicação de maturadores. Foram coletadas amostras após diferentes semanas de aplicação dos maturadores. Fonte: Barbin, 1994.

A. Modelo matemático

$$y_{ijk} = \mu + t_i + b_j + e_{ij} + s_k + ts_{ik} + e_{ijk} \quad (6)$$

B. Programa

(prog18.sas)

```
data subdiv;
input trat $ epoca bloco BRIX@@;

IF trat='T1' then EpDtr1=epoca;
ELSE EpDtr1=100;
IF trat='T2' then EpDtr2=epoca;
ELSE EpDtr2=100;
IF trat='T3' then EpDtr3=epoca;
ELSE EpDtr3=100;

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

datalines;

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;

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; /*indicação*/
```

```
* means epoca/tukey; /*indicação*/
run;
```

```
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;

proc glm;
title 'Anava - trat. dentro de época';
class trat epoca bloco;
model BRIX = bloco trat epoca epoca*trat;
/*Note a ausência do Erro(a)*/
lsmeans epoca*trat/slice=epoca adjust=tukey
PDIF=all;
run;
```

```
/*Regressões*/
proc glm;
title 'Anava - época dentro de trat. com
regressão 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;
```

```
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;
```


XVIII. EXPERIMENTOS EM LÁTICE

Os dados a seguir referem-se a um experimento em látice 4 x 4, com 3 repetições. Fonte: Gomes, 1990. O modelo foi considerado fixo, e o objetivo é verificar se os tratamentos diferem entre si.

A. Modelo matemático

$$y_{ijk} = \mu + r_i + b_{j(i)} + t_k + e_{ijk} \quad (7)$$

B. Programa

(prog19.sas)

```
data latice;
input rep bloco trat PRODUCAO@@;
cards;
1 1 1 2.0 1 1 2 2.9
1 1 3 2.2 1 1 4 3.9
1 2 5 2.3 1 2 6 2.5
1 2 7 1.4 1 2 8 1.7
1 3 9 1.6 1 3 10 3.0
1 3 11 1.5 1 3 12 2.1
1 4 13 2.3 1 4 14 3.4
1 4 15 2.0 1 4 16 2.8
2 1 1 2.2 2 1 5 2.3
2 1 9 2.7 2 1 13 1.4
2 2 2 3.1 2 2 6 2.8
2 2 10 2.6 2 2 14 2.8
2 3 3 3.1 2 3 7 2.9
2 3 11 2.5 2 3 15 2.4
2 4 4 4.0 2 4 8 2.8
2 4 12 2.7 2 4 16 1.5
3 1 1 3.0 3 1 6 2.9
3 1 11 2.6 3 1 16 3.1
3 2 2 1.8 3 2 5 1.9
3 2 12 2.9 3 2 15 2.5
3 3 3 1.7 3 3 8 2.0
3 3 9 1.4 3 3 14 2.3
3 4 4 4.4 3 4 7 3.7
3 4 10 3.3 3 4 13 2.2
;
proc glm;
title 'Análise de variância, com médias
ajustadas';
title2 'Modelo fixo';
class rep bloco trat;
model PRODUCAO= rep bloco(rep) trat;
means trat;
lsmeans trat/adjust=tukey PDIF;
run;
data latice2;
set latice;
GROUP=rep;
BLOCK=bloco;
TREATMNT=trat;
DROP rep bloco trat;
run;
```

```
proc lattice;
title 'Análise com PROC LATTICE - CUIDADO';
title2 'OBS: notar que as med.aj. estão
diferentes';
run;
```

C. Análise conjunta

Para ilustrar, serão apresentados dados (fictícios) de um látice 4 x 4, com 2 repetições, em 2 locais. Apenas o efeito de locais foi considerado fixo.

Modelo matemático

$$y_{ijk} = \mu + l_l + r_{i(l)} + b_{j(il)} + t_{kl} + t_k + e_{ijk} \quad (8)$$

(prog20.sas)

```
data latice;
input local rep bloco trat PRODUCAO@@;
cards;
1 1 1 1 8.3 1 1 1 2 8.9
1 1 1 3 4.5 1 1 1 4 5.9
1 1 2 5 9.4 1 1 2 6 1.2
1 1 2 7 0.1 1 1 2 8 0.2
1 1 3 9 7.7 1 1 3 10 2.0
1 1 3 11 9.5 1 1 3 12 9.1
1 1 4 13 8.5 1 1 4 14 8.0
1 1 4 15 3.8 1 1 4 16 1.8
1 2 1 1 8.1 1 2 1 5 10.5
1 2 1 9 2.8 1 2 1 13 3.4
1 2 2 2 2.9 1 2 2 6 2.6
1 2 2 10 2.5 1 2 2 14 1.9
1 2 3 3 3.0 1 2 3 7 0.3
1 2 3 11 1.5 1 2 3 15 2.4
1 2 4 4 3.7 1 2 4 8 2.7
1 2 4 12 8.1 1 2 4 16 1.4

2 1 1 1 9.0 2 1 1 2 9.9
2 1 1 3 8.1 2 1 1 4 8.9
2 1 2 5 9.9 2 1 2 6 1.4
2 1 2 7 0.2 2 1 2 8 0.7
2 1 3 9 1.6 2 1 3 10 3.0
2 1 3 11 8.1 2 1 3 12 1.0
2 1 4 13 9.3 2 1 4 14 9.4
2 1 4 15 0.8 2 1 4 16 8.7
2 2 1 1 2.6 2 2 1 5 10.9
2 2 1 9 2.7 2 2 1 13 1.5
2 2 2 2 3.3 2 2 2 6 1.9
2 2 2 10 4.6 2 2 2 14 3.8
2 2 3 3 2.2 2 2 3 7 0.6
2 2 3 11 2.5 2 2 3 15 2.4
2 2 4 4 1.5 2 2 4 8 2.7
2 2 4 12 1.9 2 2 4 16 9.5
;

proc sort;
by local;
run;
proc glm;
title 'Anavas individuais';
```

```

by local;
  class rep bloco trat;
  model PRODUCAO = rep bloco(rep) trat;
  random rep bloco(rep) trat/test;
run;

proc glm;
title 'Análise conjunta';
  class local rep bloco trat;
  model PRODUCAO = local rep(local)
    bloco(rep*local) trat*local trat;
  random rep(local) bloco(rep*local)
    trat*local trat/test;
  lsmeans trat;
run;
proc mixed;
title 'Análise conjunta';
  class local rep bloco trat;
  model PRODUCAO = local;
  random rep(local) bloco(rep*local)
    trat*local trat/solution;
run;

```

XIX. EXPERIMENTOS EM BLOCOS AUMENTADOS

Consideraremos agora um experimento para avaliar 10 linhas puras de soja, com uma única repetição, divididas ao acaso em três blocos. Cada bloco possuiu 4 testemunhas comuns, que são genótipos comerciais. Fonte: Scott e Milliken, 1993.

A. Modelo matemático

$$y_{ij} = \mu + b_j + t_k + e_{ij} \quad (9)$$

B. Programa

(prog21.sas)

```

data;
input entryno entry$ block YIELD;
X = entry; if entryno < 5 then X = 0;
C = entry; if entryno > 4 then C = 0;
CARDS;
1 A 1 83
2 B 1 77
3 C 1 78
4 D 1 78
7 g 1 70
11 k 1 75
12 l 1 74
1 A 2 79
2 B 2 81
3 C 2 81
4 D 2 91
5 e 2 79
9 i 2 78

```

```

1 A 3 92
2 B 3 79
3 C 3 87
4 D 3 81
6 f 3 89
8 h 3 96
10 j 3 82
;
proc print;
title;
run;

proc glm;
  class block entry;
  model YIELD = block entry;
  *random block/test;
  *random entry/test;
  *random block entry/test;
  lsmeans entry;
run;

proc glm;
  class block X C;
  model YIELD = block C X(C);
  *random block/test;
  *random C/test;
  *random X(C)/test;
  *random C X(C)/test;
  *random block C X(C)/test;
  lsmeans C X(C);
run;

proc mixed;
  class block X C;
  model YIELD =;
  random block C X(C)/solution;
run;

proc varcomp method=TYPE1;
  class block X C;
  model YIELD = block C X(C);
run;

```

XX. PLANEJAMENTO DE EXPERIMENTOS

O SAS permite planejar experimentos, incluindo desde delineamentos simples até os mais sofisticados. Apresentamos a seguir alguns exemplos de casualização de experimentos inteiramente casualizados, experimentos em parcelas subdivididas e em látices 3 x 3 e 5 x 5. No caso dos látices, é possível a partir desses exemplos construir os demais casos (ver Cochran e Cox 1950, página 304).

(prog22.sas)

```

/* Inteiramente ao acaso */
data orig;
input parcela trat;

```

```

datalines;
  1 1
  2 1
  3 1
  4 1
  5 1
  6 2
  7 2
  8 2
  9 2
 10 2
;
proc plan seed=123;
  factors parcela=10;
  output data=orig out=casual;
run;
proc sort;
by parcela;
run;
proc print;
run;

/* Blocos ao Acaso */
proc plan;
  factors blocos=3 ordered t=4;
  output out=rcbd;
run;
proc print;
run;

/* Parcela Sub-dividida*/
proc plan seed=123;
  factors blocos=3 ordered parc=2 random
subparc=3 random;
  /*Obs: random É opcional */
  output out=subdiv;
run;
proc print;
run;

/* LÁTICE 3 x 3, COM 4 REPETIÇÕES */
/* SE NECESSÁRIO, FIXE A SEMENTE */
proc plan;
  treatments trat=3 of 9 cyclic (1 2 3) 3;
  factors bloco=3 parcela=3;
  output out=rep1;
run;
proc plan;
  treatments trat=3 of 9 cyclic (1 4 7) 1;
  factors bloco=3 parcela=3;
  output out=rep2;
run;
proc plan;
  treatments trat=3 of 9 cyclic (1 5 9) 6;
  factors bloco=3 parcela=3;
  output out=rep3;
run;
proc plan;
  treatments trat=3 of 9 cyclic (1 8 6) 3;
  factors bloco=3 parcela=3;
  output out=rep4;
run;

data repet1;
rep = 1; set rep1;
data repet2;
rep = 2; set rep2;
data repet3;
rep = 3; set rep3;
data repet4;
rep = 4; set rep4;
data latices;
set repet1 repet2 repet3 repet4;
run;
proc print;
run;

/* LÁTICE 5 x 5, COM 6 REPETIÇÕES */
/* SE NECESSÁRIO, FIXE A SEMENTE */
proc plan;
  treatments trat=5 of 25 cyclic (1 2 3 4 5) 5;
  factors bloco=5 parcela=5;
  output out=rep1;
run;
proc plan;
  treatments trat=5 of 25 cyclic (1 6 11 16 21)
1;
  factors bloco=5 parcela=5;
  output out=rep2;
run;
proc plan;
  treatments trat=5 of 25 cyclic (1 7 13 19 25)
20;
  factors bloco=5 parcela=5;
  output out=rep3;
run;
proc plan;
  treatments trat=5 of 25 cyclic (1 12 23 9 20)
15;
  factors bloco=5 parcela=5;
  output out=rep4;
run;
proc plan;
  treatments trat=5 of 25 cyclic (1 17 8 24 15)
10;
  factors bloco=5 parcela=5;
  output out=rep5;
run;
proc plan;
  treatments trat=5 of 25 cyclic (1 22 18 14 10)
5;
  factors bloco=5 parcela=5;
  output out=rep6;
run;
data repet1;
rep = 1; set rep1;
data repet2;
rep = 2; set rep2;
data repet3;
rep = 3; set rep3;
data repet4;
rep = 4; set rep4;
data repet5;
rep = 5; set rep5;
data repet6;

```

```

rep = 6; set rep6;
data latice;
set repet1 repet2 repet3 repet4 repet5 repet6;
run;
proc print;
run;

```

XXI. MAPEAMENTO DE QTL's - ANÁLISE DE MARCAS SIMPLES

O SAS também pode ser usado para mapeamento de QTL's, mesmo com modelos mais elaborados. No exemplo a seguir, apresentamos como testar se o marcador M1 está ligado a algum(ns) QTL(s), usando a análise de variância de marcas simples, normalmente empregada para uma abordagem inicial do assunto. Para tanto, retomamos os dados já apresentados anteriormente, acrescentando porém dados referentes ao peso corporal (Fonte: Zeng, 2001). O procedimento pode ser repetido para todas as marcas. Além disso, mostramos também como selecionar um modelo de regressão múltipla via “stepwise”.

A. Modelo matemático

$$y_{ij} = \mu + m_i + g_{j(i)} \quad (10)$$

B. Programa

(prog23.sas)

```

data mouse;
input genot$ BW M1-M14;
datalines;
1 50 1 1 1 1 1 1 1 1 1 1 1 1 1
2 54 1 1 1 1 1 1 1 1 1 1 1 1 0
3 49 0 1 1 1 1 1 1 1 1 1 1 1 1
4 41 0 0 0 0 0 0 0 0 0 0 0 0 0
5 36 1 1 1 1 1 1 1 1 1 1 1 1 1
6 48 0 0 0 0 0 0 0 0 0 0 0 0 0
7 37 0 0 0 0 0 0 1 1 1 1 1 1 1
8 55 1 1 1 1 1 1 1 1 1 1 1 1 0
9 42 0 0 0 0 0 0 0 0 0 0 0 0 0
10 46 0 0 0 0 0 0 0 0 0 0 0 0 0
11 39 0 0 0 0 1 1 1 1 1 1 1 1 1
12 58 1 1 1 1 1 1 1 1 1 1 1 1 0
13 56 1 1 1 1 1 1 1 1 1 1 1 1 0
14 44 0 0 0 0 0 0 0 0 0 0 0 0 0
15 60 1 1 1 1 1 1 1 1 1 1 1 1 1
16 70 1 1 1 1 1 1 1 1 1 1 0 1 1
17 62 0 0 0 0 0 1 1 1 1 1 1 1 1
18 48 0 0 0 0 0 0 0 0 0 0 0 0 0
19 51 0 0 0 0 0 0 0 0 0 0 0 0 0
20 48 0 0 0 0 0 0 0 0 0 0 0 0 0
21 44 0 0 0 0 0 0 0 0 0 0 0 0 0
22 71 1 1 1 1 1 1 1 1 1 1 1 1 1
23 49 0 0 0 0 0 1 1 1 1 1 1 1 1
24 40 0 0 0 0 0 0 0 0 0 0 0 1 1

```

```

25 51 1 1 1 1 1 1 0 0 0 0 0 0 0
26 43 0 0 0 0 0 0 0 0 0 0 0 0 0
27 63 1 1 1 1 1 1 1 1 1 1 1 1 1
28 62 1 1 1 1 1 1 1 1 1 1 1 1 1
29 71 1 1 1 1 1 1 1 1 1 1 1 1 1
30 52 1 1 1 1 1 1 1 1 1 1 1 1 1
31 47 0 0 0 0 0 0 0 0 0 0 0 0 0
32 56 0 0 0 0 0 0 0 1 1 1 1 1 1
33 53 0 0 0 0 0 0 0 0 0 1 1 1 1
34 50 1 1 1 1 1 1 1 1 1 0 0 0 0
35 71 1 1 1 1 1 1 1 1 1 1 1 1 1
36 53 0 0 0 0 0 0 0 0 0 0 0 0 0
37 57 0 1 1 1 1 1 1 1 1 1 1 1 0
38 41 0 0 0 0 0 0 0 0 0 0 0 0 0
39 30 0 0 0 0 0 0 0 0 0 0 0 0 0
40 56 1 1 1 1 1 1 1 1 1 1 1 1 1
41 59 1 1 1 1 1 1 1 1 1 1 1 1 1
42 60 0 0 0 0 0 0 1 1 1 1 1 1 1
43 46 1 1 1 1 1 1 1 1 1 1 1 1 1
44 50 0 0 0 0 0 0 0 0 0 0 0 0 0
45 49 1 1 1 1 1 1 1 1 1 1 1 1 1
46 71 0 1 1 1 1 1 1 1 1 1 1 1 1
47 55 0 0 0 0 0 0 0 0 0 0 0 0 0
48 52 0 0 0 0 0 0 0 1 1 1 1 1 1
49 44 0 0 0 0 0 0 0 0 0 0 0 0 0
50 35 1 1 1 1 1 1 1 1 1 1 1 1 1
51 57 1 1 1 1 1 1 1 1 1 1 1 1 0
52 54 0 0 0 0 0 0 0 0 0 0 0 0 0
53 62 1 1 1 1 1 1 1 1 1 1 1 1 1
54 49 1 1 1 1 1 1 1 1 1 1 1 1 1
55 59 0 0 1 1 1 1 1 1 1 1 1 1 1
56 35 0 0 0 0 0 0 0 0 0 0 0 0 0
57 43 0 0 0 0 0 0 0 0 1 1 1 1 1
58 45 0 0 0 0 0 0 0 0 0 0 0 0 0
59 44 0 0 0 0 1 1 1 1 1 1 0 1 0
60 47 0 0 0 0 0 0 0 0 0 0 0 0 0
61 51 0 0 0 0 0 0 0 0 0 0 0 0 0
62 50 0 0 0 0 0 0 0 0 0 0 0 0 0
63 44 0 0 0 0 0 0 0 0 0 0 0 0 0
64 44 0 0 0 0 0 0 0 0 0 0 0 0 1
65 49 0 1 1 1 1 1 1 1 1 1 1 1 0
66 43 0 0 0 0 0 0 0 0 0 0 0 0 0
67 45 0 0 0 0 0 0 0 0 0 0 0 0 1
68 53 0 0 0 0 0 0 0 0 0 0 0 0 1
69 42 0 0 0 0 0 0 0 0 0 0 1 1 1
70 43 1 1 1 1 1 1 1 1 1 1 1 1 1
71 58 1 1 1 1 1 1 1 1 1 1 1 1 1
72 36 0 0 0 0 0 0 0 0 0 0 0 0 0
73 51 1 0 0 0 0 0 0 0 0 0 0 0 0
74 47 1 1 1 1 1 1 1 1 1 1 1 1 1
75 68 1 1 1 1 1 1 1 1 1 1 1 1 0
76 56 1 1 1 1 1 1 1 1 1 1 1 1 1
77 56 1 1 1 1 1 1 1 1 1 1 1 1 1
78 69 1 1 1 1 1 1 1 1 1 1 1 1 1
79 46 1 1 1 1 1 1 0 0 0 0 0 0 0
80 52 1 1 1 1 1 1 1 1 1 1 1 1 1
81 43 1 0 0 0 0 0 0 0 0 0 0 1 1
82 40 0 0 0 0 0 0 0 0 0 0 0 0 1
83 79 1 1 1 1 1 1 1 1 1 1 1 1 1
84 57 0 0 0 0 0 0 0 0 1 1 1 1 1
85 43 1 1 1 1 1 1 1 1 1 1 1 1 1
86 56 0 0 0 0 0 0 0 0 0 0 0 0 0

```

```

87 40 0 0 0 0 0 0 0 0 0 0 0 0 0 0
88 44 0 0 0 0 0 0 0 0 0 0 0 0 0 0
89 40 0 0 0 0 0 0 0 0 0 0 0 0 0 0
90 66 0 0 0 0 0 0 0 0 0 0 0 0 1 1
91 32 0 0 0 0 0 0 0 0 0 0 0 0 0 0
92 45 0 0 0 0 0 0 0 0 0 0 0 0 0 0
93 45 1 1 1 1 1 0 0 0 0 0 0 0 0 0
94 53 0 0 0 0 0 0 0 0 0 0 0 0 0 0
95 48 0 0 0 0 0 0 0 1 1 1 1 1 1 0
96 36 0 0 0 0 0 0 1 1 1 1 1 1 1 1
97 50 0 0 0 0 0 0 0 0 0 0 0 0 0 0
98 56 0 0 0 0 0 0 0 0 0 0 0 0 0 0
99 50 0 0 0 0 0 0 0 0 0 0 0 0 0 0
100 45 1 1 1 1 0 0 0 0 0 0 0 0 0 0
101 43 1 1 1 0 0 0 0 0 0 0 0 0 0 0
102 37 0 0 0 0 0 0 0 0 0 0 0 0 0 1
103 35 1 0 0 0 0 0 0 0 0 0 0 0 0 0
;

proc glm;
title 'Análise da Marca M1';
class M1;
model BW = M1;
run;

proc reg;
model BW = M1-M14 / selection=stepwise;
run;

```

XXII. ÁLGEBRA DE MATRIZES (INTRODUÇÃO AO SAS/IML)

Apresentamos abaixo alguns exemplos de como o PROC IML do SAS pode ser usado para realizar operações com matrizes.

(prog24.sas)

```

proc iml;
a={3 -1 2,
   2 -2 3,
   4 1 -4};
print a;

inv_a=inv(a);
print inv_a;

A={2 2, 3 4};
B={4 5, 1 0};
SOMA=A+B;
print A B soma;

DIF=A-B;
print dif;

AxB=A*B;
print AxB;

```

```

x={1 2 3,
   4 5 6,
   6 2 4};
traco=trace(x);
print traco;

x={ 1.9 2 3,
   40 5.2 6,
   60 20 4.3};
transp1=t(x);
transp2=x`;
print transp1 transp2;

A={4 3,
   -2 5};
b={10,
   8};
C={7 -10,
   4 13};
A_b=A||b;
A_C=A//C;
print A_b A_C;

A={4 2 2,
   2 2 0,
   2 0 2};
RANK_A=round(trace(ginv(A)*A));
print RANK_A;

a={1 1 1,
   1 2 4,
   1 3 9};
determ=det(a);
print determ;

A={4 3, -2 5};
b={10, 8};
x=solve(A,b);
print x;

X={1 1 0,
   1 1 0,
   1 0 1,
   1 0 1};
MP=ginv(X);
print MP;

quit;

```

Nota

Recomenda-se que o leitor verifique com muito cuidado os resultados obtidos em cada operação.

A. Exemplo: análise dialélica

Tomaremos os dados apresentados por Vencovsky e Barriga (1992) para um dialelo (cruzamentos interraciais de milho) analisado segundo o método 4 de Griffing (presença somente dos híbridos, sem os recíprocos). O quadrado médio do resíduo (dividido por r) foi 0,1080. O objetivo é estimar os efeitos das

capacidades geral e específica de combinação, usando o método dos mínimos quadrados.

Tabela 4: Médias dos cruzamentos para a análise dialélica.

	1	2	3	4	5
1	...	3,65	2,64	2,44	2,11
2		...	4,23	3,29	3,91
3			...	2,85	3,18
4				...	2,47
5					...

Modelo

$$y_{ij} = \mu + g_i + g_j + s_{ij} + \bar{e}_{ij} \quad (11)$$

Programa

(prog25.sas)

```
proc iml;
reset print;
y={3.65,
    2.64,
    2.44,
    2.11,
    4.23,
    3.29,
    3.91,
    2.85,
    3.18,
    2.47,
    0,
    0,
    0,
    0,
    0,
    0,
    0};
X={1 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0,
    1 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0,
    1 1 0 0 1 0 0 0 1 0 0 0 0 0 0 0,
    1 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0,
    1 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0,
    1 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0,
    1 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0,
    1 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0,
    1 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0,
    1 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1,
    0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0,
    0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0,
    0 0 0 0 0 0 1 0 0 0 1 1 1 0 0 0,
    0 0 0 0 0 0 0 1 0 0 1 0 0 1 1 0,
    0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 1,
    0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 1};
XtX=t(X)*X;
Xty=t(X)*y;
teta=inv(XtX)*(Xty);

quit;
```

XXIII. BIBLIOGRAFIA RECOMENDADA

- [1] Alves, M. C. SAS Básico. Série Didática Ciagri nº 9. ESALQ/USP, Piracicaba, SP. 1990.
- [2] Barbin, D. Planejamento e Análise Estatística de Experimentos Agronômicos. Apostila. Departamento de Matemática e Estatística, ESALQ/USP, Piracicaba, SP. 1994.
Discute detalhes teóricos de alguns dos modelos e métodos aqui apresentados.
- [3] Cochran, W.G.; Cox, G.M. Experimental Designs. John Wiley and Sons, Inc. New York, 1950.
- [4] Cody, R.P.; Smith, J.K. Applied Statistics and the SAS Programming Language. 3rd ed. Elsevier Science Publishing Co., New York. 1991. 403p.
Excelente didática e ótimo para uma introdução ao SAS como pacote. Não apresenta detalhes dos métodos estatísticos.
- [5] Diório, F.C. SAS Applications Programming: A Gentle Introduction. Duxbury Press, California. 1997. 684 p.
Leitura recomendada para os que se interessam em usar o SAS para escrever seus próprios programas.
- [6] 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. Disponível em <http://www.ciagri.usp.br/~aafgarci/estaplic.html>
Aborda a maioria dos exemplos aqui apresentados.
- [7] Garcia, A.A.F.; Barbin, D.; Piedade, S.M.S. LCE 602 - Estatística Experimental (aulas práticas). Publicação didática. Departamento de Ciências Exatas, Escola Superior de Agricultura "Luiz de Queiroz". Março de 2002. 63p. Disponível em <http://www.ciagri.usp.br/~aafgarci/estexp.html>
Vários exemplos com os respectivos programas e arquivos de dados. Adotado como texto da disciplina LCE 602 - Estatística Experimental, oferecida como curso optativo aos alunos de graduação da ESALQ/USP.
- [8] Garcia, A.A.F. Uso do pacote computacional SAS para análises estatístico-genéticas. 2002. Disponível em <http://www.ciagri.usp.br/~aafgarci/estaplic.html>
Versão anterior desse curso, que foi também ministrado em 2002.
- [9] Latour, D.; Littell. Advanced General Linear Models with an Emphasis on Mixed Models - Course Notes. SAS Institute Inc., Cary, NC. 1996.
Técnicas avançadas, recomendado aos leitores com forte base teórica.
- [10] Littell, R. C.; Milliken, G. A.; Stroup, W. W.; Wolfinger, R. D. SAS System for Mixed Models. SAS Institute Inc., Cary, NC. 1996.
- [11] 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.
- [12] Piedade, S. M. S. Aulas práticas de Estatística Experimental, 1º semestre de 2000. Publicação didática.

Departamento de Ciências Exatas, ESALQ/USP, Piracicaba, SP.

- [13] Reyes, A. E. L. SAS - Estatística Básica. Série Didática Ciagri nº 17. ESALQ/USP, Piracicaba, SP. 1990.
- [14] SAS/STAT User's Guide, Version 6, Fourth Edition, vol. 1 e 2. SAS Institute Inc., Cary, NC. 1994.
Referência completa sobre todas as proc's do SAS.
- [15] Scott, R.A.; Milliken, G.A. A SAS program for analysing augmented randomized complete block design. Crop Science 33: 865-867, 1993.
- [16] Vencovsky, R.; Barriga, P. Genética Biométrica no Fitomelhoramento. Sociedade Brasileira de Genética, Ribeirão Preto, 1992. 486 p.
- [17] Zocchi, S.S.; Barbin, D. LCE 602 - Estatística Experimental, 2º semestre de 1999. Publicação didática. Departamento de Ciências Exatas, ESALQ/USP, Piracicaba, SP.

PÁGINAS NA INTERNET:

- [18] <http://www.sas.com>
Página do SAS Institute.
- [19] <http://smoo.ciagri.usp.br/sasdoc/sasdoc/sashtml/onldoc.htm>
Manual do SAS na internet, que pode ser consultado em computadores da USP.
- [20] <http://www.r-project.org>
Não deixe de consultar caso deseje conhecer o R, que é um excelente pacote estatístico (na verdade, é um ambiente - linguagem de programação). É gratuito e multiplataforma; permite fazer todas as análises aqui apresentadas (e muito mais).