

## ANÁLISE DE MEDIDAS REPETIDAS NO TEMPO USANDO O SAS

Euclides Braga MALHEIROS\*

**Medidas repetidas no tempo:** medidas tomadas em uma seqüência de tempos, em uma mesma unidade experimental.

Os experimentos com medidas repetidas no tempo envolvem geralmente 2 fatores: tratamentos e tempos, e são freqüentes em experimentos com animais, plantas, humanos, etc. O objetivo principal desse tipo de experimento é examinar e comparar as tendências dos tratamentos ao longo do tempo. Isto pode envolver comparações entre tratamentos dentro de cada tempo, ou comparações de tempos dentro de cada tratamento.

Considere o exemplo:

Cinco variedades de uma cultura (tratamentos) avaliadas ao longo do tempo (0, 30, 60, 90, 120 e 150 dias), em um experimento inteiramente casualizado, com 3 repetições.

Os dados observados são apresentados a seguir:

Tabela 1. Dados de porcentagem de açúcar na cana, em pol%, obtidos em um DIC com 3 repetições, envolvendo 5 variedades de cana, observados em 6 tempos de desenvolvimento da cultura.

Variedades	Rep.	Tempo em dias					
		0	30	60	90	120	150
V1	1	11,82	14,86	13,84	15,53	15,49	15,82
	2	12,07	14,44	13,92	15,47	16,34	18,64
	3	12,45	14,18	13,76	14,35	15,93	16,52
V2	1	12,47	15,19	15,02	15,54	18,53	15,76
	2	11,07	13,38	14,61	14,07	17,84	16,91
	3	10,66	14,22	13,54	15,93	15,94	16,81
V3	1	12,92	14,49	13,40	13,68	16,26	14,78
	2	10,29	14,42	14,62	15,84	16,29	15,62
	3	12,83	13,92	15,69	15,12	14,91	17,22
V4	1	11,96	14,71	14,98	15,25	16,21	15,53
	2	13,38	15,07	13,62	15,39	15,77	16,51
	3	10,37	15,78	13,33	14,50	16,66	16,34
V5	1	11,05	13,18	14,61	14,88	16,51	16,36
	2	10,63	13,14	14,53	14,21	16,57	15,24
	3	13,43	14,08	14,23	14,11	15,86	17,50

Fonte: Nogueira (1995)

Como se sabe, nas análises usuais do SAS (enfoque univariado) no SAS-DATA-SET as colunas são as variáveis e as linhas os registros, e assim sendo, os dados devem ser organizados na forma apresentada a seguir:

\* Departamento de Ciências Exatas – FCAV/UNESP, Campus de Jaboticabal. 14870-000 Jaboticabal SP

Dados na forma univariada:

Variedade	Tempo	Repetição	Y
1	0	1	11.82
1	0	2	12.07
1	0	3	12.45
1	30	1	14.86
...	...	...	...
5	3	150	17.50

Para um tipo de análise de medidas repetidas no tempo (enfoque multivariado) o SAS-DATA-SET deve conter uma variável (coluna) para cada tempo.

Dados na forma multivariada :

Variedade	Repetição	T1	T2	...	T6
1	1	11.82	14.86	...	15.82
1	2	12.07	14.44	...	18.64
1	3	12.45	14.18	...	16.52
2	1	12.47	15.19	...	15.76
...	...	...	...	...	...
5	150	13.43	14.08	...	17.50

### Passos para a análise desses dados:

**Passo 1:** Criar os SAS-DATA-SET multivariado quando os dados estão digitados da forma univariada (MRT1U.TXT), ou vice-versa.

```
/* CRIACAO DO SAS-DATA-SET MULTIVARIADO A PARTIR DO UNIVARIADO */
OPTIONS LS=78 PS=64;
LIBNAME PASTA "C:\MEUS DOCUMENTOS\MRT";
DATA PASTA.MRT1U;
INFILE "A:\MRT1U.TXT";
INPUT VR TP RP Y;
*PROC PRINT;
RUN;
PROC SORT DATA=PASTA.MRT1U; BY VR RP;
PROC TRANSPOSE OUT=PASTA.MRT1UM(RENAME=( _0=T1 _30=T2 _60=T3 _90=T4 _120=T5 _150=T6));
BY VR RP;
ID TP;
PROC PRINT;
RUN;
```

**Passo 2:** Representar graficamente os perfis médios 5 tratamentos (Variedades), ao longo do tempo.

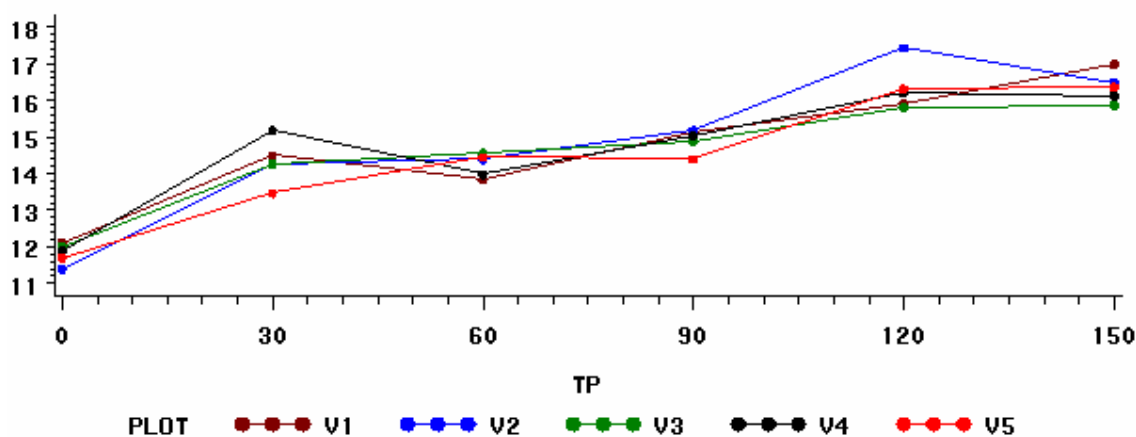
```
/* REPRESENTACAO GRAFICA DA TENDENCIA DOS TRATAMENTOS AO LONGO DO TEMPO */
OPTIONS LS=78 PS=64;
LIBNAME PASTA "C:\MEUS DOCUMENTOS\MRT";
DATA MRT1U; SET PASTA.MRT1U;
*PROC PRINT;
RUN;
PROC SORT; BY VR TP RP; RUN;
PROC MEANS MEAN NOPRINT;
OUTPUT OUT=MRT1G;
BY VR TP;
VAR Y;
```

```

*PROC PRINT;
RUN;
DATA MRT1G; SET MRT1G;
IF _STAT_ ^= "MEAN" THEN DELETE;
KEEP VR TP Y;
*PROC PRINT;
RUN;
PROC SORT DATA=MRT1G; BY TP;
PROC TRANSPOSE OUT=MRT1G(RENAME=(_1=V1 _2=V2 _3=V3 _4=V4 _5=V5));
BY TP;
ID VR;
PROC PRINT;
RUN;
PROC GPLOT;
PLOT V1*TP V2*TP V3*TP V4*TP V5*TP/OVERLAY LEGEND HAXIS=0 TO 150 BY 30;
SYMBOL1 COLOR=RED INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL2 COLOR=BLUE INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL3 COLOR=GREEN INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL4 COLOR=BLACK INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL5 COLOR=ORANGE INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
TITLE HEIGHT=1.4 "TENDENCIA DOS TRAT. AO LONGO DO TEMPO";
RUN;

```

## TENDENCIA DOS TRAT. AO LONGO DO TEMPO



## MÉTODOS PARA ANALISAR MEDIDAS REPETIDAS NO TEMPO

### 1. Análise como parcelas subdivididas.

Historicamente, o método mais utilizado para análise de medidas repetidas no tempo é a análise univariada num esquema em parcelas subdivididas, tendo o tempo como subparcelas. Referenciado na literatura como parcelas subdivididas no tempo.

O problema deste método de análise é que o delineamento em parcelas subdivididas a matriz de covariâncias é do tipo homogênea (condição suficiente para que os testes F sejam exatos), ou seja:

Se a matriz de variâncias e covariâncias é:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1t} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2t} \\ \dots & \dots & \dots & \dots \\ \sigma_{t1} & \sigma_{t2} & \dots & \sigma_t^2 \end{bmatrix},$$

onde  $\sigma_i^2$  é a variância no tempo  $i$ , e  $\sigma_{ij}$  é a covariância entre os tempos  $i$  e  $j$ , na estrutura homogênea  $\sigma_i^2 = \sigma^2, \forall i$  e  $\sigma_{ij} = \rho, \forall i \neq j$ .

Essa é a considerada na análise em parcelas subdivididas.

O que se encontra na literatura é que as medidas repetidas em uma mesma unidade experimental (animal, plantas ou humanos) são correlacionadas e que medidas em tempos mais próximos apresentam correlações mais altas que em tempos mais distantes. Uma estrutura que tem sido estudada é a autoregressiva.

Segundo Huynh & Feldt (1970) uma condição necessária e suficiente para que os testes F sejam exatos, mais geral da forma de  $\Sigma$ , é que:

$$\sigma_{kk'} = \frac{(\sigma_k^2 + \sigma_{k'}^2)}{2} - \lambda, \quad k \neq k'.$$

onde  $\lambda$  é a diferença entre as médias das variâncias e as medias das covariâncias.

Esta condição, denominada condição H-F ou condição de esfericidade (ou circularidade) da matriz  $\Sigma$ , equívale a especificar que as variâncias das diferenças entre pares de tempos sejam todas iguais, ou seja:

$$\sigma_{Y_k - Y_{k'}}^2 = c, \quad \forall k \neq k'$$

Para exemplificar, verifique se a matriz abaixo satisfaz a condição de esfericidade:

$$\Sigma = \begin{bmatrix} 5,0 & 2,5 & 5,0 & 7,5 \\ 2,5 & 10,0 & 7,5 & 10,0 \\ 5,0 & 7,5 & 15,0 & 12,5 \\ 7,5 & 10,0 & 12,5 & 20,0 \end{bmatrix}.$$

Observe que

$$\lambda = \frac{(5,0 + 10,0 + 15,0 + 20,0)}{4} - \frac{(2,5 + 5,0 + 7,5 + 7,5 + 10,0 + 12,5)}{6} = 5,0$$

$$\text{e } \sigma_{12} = \frac{(5,0 + 10,0)}{2} - 5,0 = 2,5, \quad \sigma_{13} = \frac{(5,0 + 15,0)}{2} - 5,0 = 5,0; \quad \text{e assim por diante.}$$

Ainda mais:

$$\sigma_{Y_1 - Y_2}^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} = 5,0 + 10,0 - 2(2,5) = 10,0;$$

$$\sigma_{Y_1 - Y_3}^2 = \sigma_1^2 + \sigma_3^2 - 2\sigma_{13} = 5,0 + 15,0 - 2(5,0) = 10,0; \quad \text{e assim por diante.}$$

Logo a matriz satisfaz a condição de esfericidade.

**Passo 3 (opcional):** Fazer um programa SAS para obter a matriz de covariâncias entre os tempos dos dados da Tabela 1.

```
/* OBTENÇÃO DA MATRIZ DE COVARIÂNCIAS ENTRE TEMPOS */
OPTIONS LS=78 PS=64;
LIBNAME PASTA "C:\MEUS DOCUMENTOS\MRT";
DATA MRT1M; SET PASTA.MRT1M;
*PROC PRINT;
RUN;
PROC CORR COV NOPRINT OUTP=MVC;
VAR T1-T6;
RUN;
DATA MVC; SET MVC;
IF _TYPE_ ^= "COV" THEN DELETE;
RUN;
PROC PRINT DATA=MVC;
RUN;
```

Observando a matriz de covariâncias resultante, observa-se que não é fácil visualizar se a condição de esfericidade é satisfeita.

Murchly (1940) apresenta um teste para a condição de esfericidade.

Segundo Box (1954), Geisser & Greenhouse (1958), Greenhouse & Geisser (1959) e Huynh & Feldt (1976), ainda que a matriz  $\Sigma$  não satisfaça a condição de esfericidade, a distribuição F central poderá ser usada, de maneira aproximada, para sub-parcelas ( $S$ ) e para a interação entre parcelas e sub-parcelas ( $P*S$ ), se forem efetuadas correções nos graus de liberdade dessas fontes de variação (multiplicando os originais por um fator  $\epsilon$ ,  $(t-1)^{-1} \leq \epsilon \leq 1$ , onde  $t$  é o número de tempos).

Em 1984 o comando REPEATED foi incluído no PROC GLM do SAS. Este comando executa análises uni e multivariadas, incluindo o teste de esfericidade de  $\Sigma$  e as correções de graus de liberdade sugeridas na literatura.

### 3. Análise usando o comando REPEATED do PROC GLM.

Para este tipo de análise os dados devem estar na forma multivariada (uma coluna para cada tempo).

A sintaxe para o uso desse procedimento é:

#### Sintaxe:

```
PROC GLM <opções 1>;
CLASS <fator Tra.>;
MODEL <Lista Var.Tempo>=<fator Trat.>/<opções 2>;
REPEATED <fator tempo> <nº níveis fator tempo> [<(valores níveis fator tempo)>]
        <tipo de contrastes>/<opções 3>;
RUN;
```

Uma das possíveis <opções 1> é:

- **DATA=<SDS>** - especifica o SAS-DATA-SET a ser usado.

Uma das possíveis <opções 2> é:

- **NOUNI** – não executa as análises unidimensionais, dentro de cada tempo.

Alguns dos possíveis <tipos de contrastes> são:

- **CONTRAST [(nível referencial)]** – gera contrastes entre cada nível do fator tempo com o nível referencial. Quando o nível referencial não for especificado, considera o último.
- **POLYNOMIAL** – gera contrastes de polinômios ortogonais para os níveis do fator tempo.
- **HELMERT** – gera contrastes entre cada nível do fator tempo com a média dos subsequentes.
- **MEAN (nível referencial)** – gera contrastes entre o cada nível (exceto o referencial) com a média dos outros.
- **PROFILE** – gera contrastes entre níveis adjacentes do fator tempo.

Algumas das <opções 3> são:

- **CANONICAL** – executa uma análise canônica das matrizes H e E.
- **HTYPE=n** – especifica o tipo da soma de quadrados a ser usado.
- **NOM** – Não executa a análise multivariada.
- **NOUNI** – Não executa as análises univariadas.
- **PRINT|H|M|V** – Imprime a matriz especificada:  
**E** - matriz da soma de quadrados dos produtos cruzados (SS&CP) para erros. Com a opção PRINTE o SAS apresenta o teste de esfericidade da matriz de covariâncias.  
**H** - matriz da soma de quadrados dos produtos cruzados (SS&CP) para a Hipótese.  
**M** - matriz dos contrastes.  
**V** - matriz dos auto-valores e auto-vetores associada ao teste.
- **SUMMARY** – apresenta a tabela da análise da variância dos contrastes para o fator tempo.

**Passo 4:** Fazer um programa SAS obter a análise usando o comando REPEATED do PROC GLM, para os dados da Tabela 1.

```
/* Análise usando o comando REPEATED do PROC GLM */
OPTIONS LS=78 PS=64;
LIBNAME PASTA "C:\MEUS DOCUMENTOS\MRT";
DATA M2; SET PASTA.MRT1U;
PROC SORT; BY VR RP;
PROC TRANSPOSE OUT=MRT3(RENAME=( _0=T1 _30=T2 _60=T3 _90=T4 _120=T5 _150=T6 ));
BY VR RP;
ID TP;
*PROC PRINT;
RUN;
PROC GLM DATA=MRT3;
CLASS VR;
MODEL T1-T6=VR/NOUNI;
REPEATED TP 6 POLYNOMIAL/PRINTE SUMMARY;
RUN;
```

## INTERPRETAÇÃO DOS RESULTADOS:

### a) Teste de esfericidade.

Test for Sphericity: Mauchly's Criterion = 0.1956602
Chisquare Approximation = 13.214144 with 14 df    Prob > Chisquare = 0.5097
Applied to Orthogonal Components:
Test for Sphericity: Mauchly's Criterion = 0.1956602
Chisquare Approximation = 13.214144 with 14 df    Prob > Chisquare = 0.5097

O SAS apresenta 2 testes de esfericidade. O primeiro depende do tipo de contrastes solicitado e o segundo é válido para qualquer conjunto de contrastes ortogonais.

Observa-se com esses resultados que a condição de esfericidade da matriz não deve ser rejeitada. Assim sendo os testes envolvidos na análise anterior (Parcelas subdivididas) são exatos.

**Passo 5:** Fazer a análise como medidas repetidas no tempo.

/\* Análise univariada como parcelas subdivididas \*/

```

OPTIONS LS=76 PS=64;
LIBNAME PASTA "C:\MEUS DOCUMENTOS\MRT";
DATA M; SET PASTA.MRT1U;
PROC GLM;
CLASS VR TP RP;
MODEL Y= VR RP(VR) TP VR*TP/SS3;
TEST H=VR E=RP(VR);
MEANS VR/TUKEY E=RP(VR);
MEANS TP/TUKEY;
RUN;
```

Resultado da análise:

Tabela 3: Valores de F, com respectivas probabilidades, Coeficiente de Variação (CV), e médias obtidas na análise da variância.

ESTATÍSTICAS	Variável Y
F p/ Tratam.(Tr)	0,56(0,6969)
F p/ Tempo (Tp)	54,24(0,0001)
F p/ int. Tr x Tp	0,84(0,6546)
CV% - Parcelas	6,25%
CV% - Subparcelas	6,02%

O problema deste método de análise é que o delineamento em parcelas subdivididas a matriz de covariâncias é do tipo homogênea (condição suficiente para que os testes F sejam exatos), ou seja:

**Exemplo 2:** Fazer os passos descritos anteriormente para os dados da Tabela 4.

Tabela 4. Dados de IAF de cinco genótipos de leguminosas, avaliados em 7 épocas (dias), obtidos em um DBC.

Genótipo	Bloco	Número de dias						
		88	104	120	137,5	153,5	181,5	209,5
1	1	0,13	0,39	0,46	0,52	1,18	0,87	0,51
1	2	0,31	0,25	1,07	0,44	1,11	1,41	1,08
1	3	0,22	0,40	0,53	3,61	1,11	0,98	0,78
1	4	0,08	0,17	0,97	1,11	1,70	0,92	0,74
2	1	0,13	0,36	0,89	0,62	1,64	1,61	1,42
2	2	0,27	0,10	0,53	0,60	1,52	1,01	1,09
2	3	0,13	0,41	1,03	3,60	1,92	0,56	1,09
2	4	0,08	0,55	0,62	1,04	2,47	1,45	1,12
3	1	0,84	1,44	2,31	6,07	3,90	3,42	2,00
3	2	0,45	1,18	2,66	3,88	4,03	3,09	0,99
3	3	0,67	2,39	4,25	6,33	4,13	3,46	0,96
3	4	1,28	3,45	5,04	5,57	3,87	0,36	0,77
4	1	0,42	0,80	0,72	1,07	1,20	1,08	1,26
4	2	0,15	0,40	0,42	0,85	0,66	0,85	0,71
4	3	0,22	0,30	0,77	0,94	1,44	1,49	0,62
4	4	0,28	0,36	0,74	0,73	1,62	1,84	1,36
5	1	0,67	1,77	2,09	3,27	3,92	2,36	2,72
5	2	0,66	1,07	2,39	4,19	4,89	1,86	1,61
5	3	1,41	2,55	3,87	4,62	3,62	3,87	0,02
5	4	1,30	2,16	5,78	8,62	7,92	0,26	0,26

Fonte: Castro (1999)

### Os dados estão digitados na forma multivariada (MRT2M.txt)

#### Entrada de dados e criação de um SDS na forma univariada

```

OPTIONS LS=78 PS=64;
LIBNAME PASTA "C:\MEUS DOCUMENTOS\MRT";
DATA PASTA.MRT2M (KEEP=GN BL T1-T7)
    PASTA.MRT2U (KEEP=GN BL TP Y);
INFILE "A:\MRT2M.TXT";
INPUT GN BL T1-T7;
OUTPUT PASTA.MRT2M;
TP=88; Y=T1; OUTPUT PASTA.MRT2U;
TP=104; Y=T2; OUTPUT PASTA.MRT2U;
TP=120; Y=T3; OUTPUT PASTA.MRT2U;
TP=137.5; Y=T4; OUTPUT PASTA.MRT2U;
TP=153.5; Y=T5; OUTPUT PASTA.MRT2U;
TP=181.5; Y=T6; OUTPUT PASTA.MRT2U;
TP=209.5; Y=T7; OUTPUT PASTA.MRT2U;
RUN;
PROC PRINT DATA=PASTA.MRT2M;
PROC PRINT DATA=PASTA.MRT2U;
RUN;

```

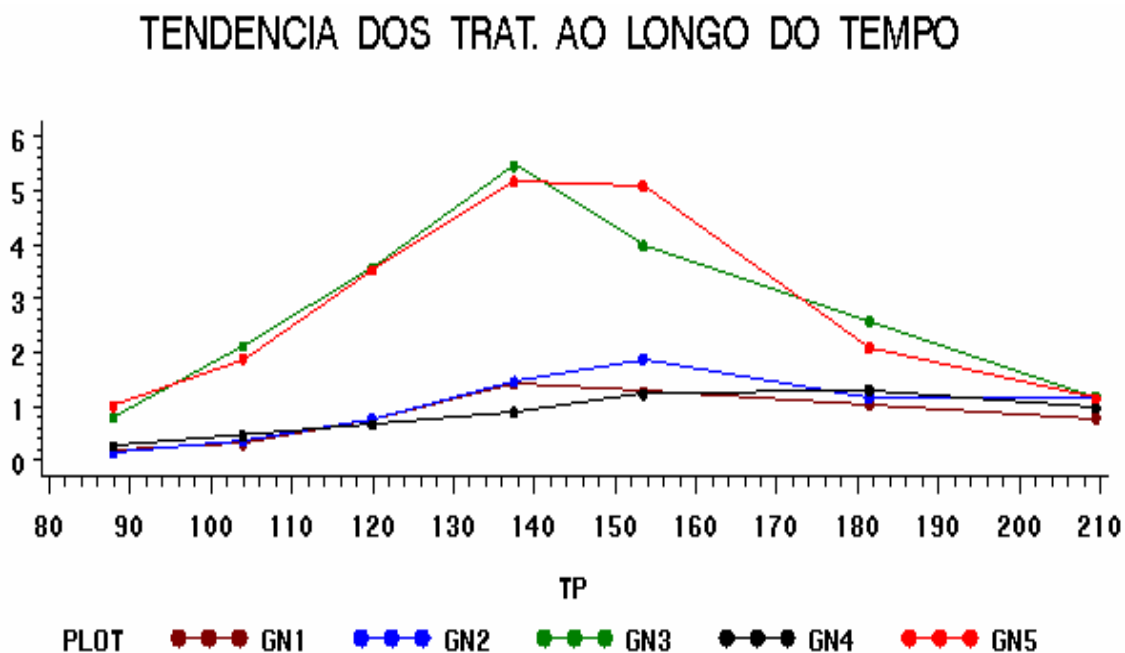


```

/* REPRESENTAÇÃO GRÁFICA */
DATA UNI ; SET PASTA.MRT2U;
RUN;
PROC SORT; BY GN TP BL; RUN;
PROC MEANS MEAN NOPRINT;
OUTPUT OUT=UNIM MEAN=YM;
BY GN TP;
VAR Y;
*PROC PRINT;
RUN;
PROC SORT DATA=UNIM; BY TP;
PROC TRANSPOSE OUT=UNIMT(RENAME=(_1=GN1 _2=GN2 _3=GN3 _4=GN4 _5=GN5));
BY TP;
ID GN;
VAR=YM;
*PROC PRINT;
RUN;
PROC GPLOT;
PLOT GN1*TP GN2*TP GN3*TP GN4*TP GN5*TP/OVERLAY LEGEND;
SYMBOL1 COLOR=RED INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL2 COLOR=BLUE INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL3 COLOR=GREEN INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL4 COLOR=BLACK INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
SYMBOL5 COLOR=ORANGE INTERPOL=JOIN VALUE=DOT HEIGHT=0.5;
TITLE HEIGHT=1.4 "TENDENCIA DOS TRAT. AO LONGO DO TEMPO";
RUN;

```

### Representação gráfica:



### Teste de esfericidade da matrix

```
/* Análise usando o comando REPEATED do PROC GLM */  
PROC GLM DATA=PASTA.MRT2M;  
CLASS BL GN;  
MODEL T1-T7=GN/NOUNI;  
REPEATED TP 6 POLYNOMIAL/PRINTE SUMMARY;  
RUN;
```

### Interpretação dos resultados:

#### a) Resultados do teste de esfericidade.

<p>Test for Sphericity: Mauchly's Criterion = 0.0027667 Chisquare Approximation = 75.262197 with 20 df    Prob &gt; Chisquare = 0.0000</p> <p>Applied to Orthogonal Components: Test for Sphericity: Mauchly's Criterion = 0.0027667 Chisquare Approximation = 75.262197 with 20 df    Prob &gt; Chisquare = 0.0000</p>
---

Observa-se com esses resultados que a condição de esfericidade da matriz deve ser rejeitada. Assim sendo os testes da análise em parcelas subdivididas não são exatos.