

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Cadernos de Matemática e Estatística
Série F: Trabalho de Divulgação

ANÁLISE DE OBSERVAÇÕES REPETIDAS ATRAVÉS DE
CONTRASTES NO TEMPO

João Riboldi
Dinara W. Xavier Fernandez

Série F, nº 4
Porto Alegre, julho de 1992

ANÁLISE DE OBSERVAÇÕES REPETIDAS ATRAVÉS DE
CONTRASTES NO TEMPO

João Riboldi

Dinara W. Xavier Fernandez

Departamento de Estatística

UFRGS

RESUMO

A análise de acordo com o modelo de parcelas subdivididas ou de experimentos em faixas é usualmente utilizada para analisar informações obtidas, de forma repetida, na mesma unidade experimental. Sabe-se que as suposições para tal análise são extremamente restritivas e que, frequentemente, não se verificam.

Procedimentos alternativos mais adequados são sugeridos na bibliografia, dentre os quais, a análise de variância multivariada e a utilização de contrastes no tempo. No presente trabalho, utiliza-se, quando verificada a inadequabilidade do modelo de parcelas subdivididas, o procedimento de contrastes no tempo, conforme sugerido por ROWELL e WALTERS (1976).

CONTRAST OVERTIME FOR ANALYSING DATA WITH
REPEATED OBSERVATIONS

ABSTRACT

Split-plot (or split-block) analyses are commonly applied to experimental results where several successive observations of the same variable have been recorded on each experimental unit. The assumptions required for such analyses receive scant attention and it often seems unlikely that these assumptions would be satisfied. Multivariate approach and method of analysing contrasts have been used.

In this paper, contrast overtime are analysed when the split-plot model is not valid, as suggested by ROWELL and WALTERS (1976).

1. INTRODUÇÃO.

Em experimentos com um conjunto de tratamentos sob investigação, instalados em determinado delineamento experimental, frequentemente várias observações de uma mesma variável são registradas em cada unidade experimental, observações estas coletadas em tempos diferentes. Este procedimento produz uma sequência de observações da mesma variável, com a característica de que as observações são correlacionadas ao longo do tempo.

Observações com a estrutura considerada são usualmente analisadas através da Análise de Variância para parcelas subdivididas (ou de experimentos em faixas), ou através da Análise de Variância Multivariada para medidas repetidas, ou pelo ajuste de equações polinomiais às sequências de observações e analisando-se seus coeficientes separadamente em uma série de Análises de Variância.

A Análise de Variância para parcelas subdivididas, considerando o tempo como tratamento secundário (tratamento da subparcela), pode não ser satisfatória, desde que certas suposições exigidas pela análise não são frequentemente satisfeitas. A Análise de Variância Multivariada, uma vez que o número de ocasiões (tempos) geralmente excede os graus de liberdade do erro, raramente pode ser utilizada.

Desta forma, dentre os procedimentos citados, o ajuste de equações polinomiais às sequências de observações é o de maior aplicabilidade. No presente trabalho, enfoca-se a análise de observações repetidas através do ajuste de equações polinomiais, que constituem contrastes no tempo.

2. MÉTODOS DE ANÁLISE.

Sejam experimentos organizados em blocos completos casualizados, com todos os t tratamentos ocorrendo em cada um dos r blocos. As observações para uma mesma variável são registradas em p tempos para cada uma das rt unidades experimentais.

Vários métodos de análise tem sido utilizados, os quais são caracterizados a seguir.

2.1. ANÁLISE DE VARIANCIA PARA CADA TEMPO SEPARADAMENTE

A Análise de Variância para cada tempo separadamente segue o seguinte esquema:

Causas de Variação	G.L.
Entre Blocos	$r - 1$
Entre Tratamentos	$t - 1$
Erro	$(r-1)(t-1)$

produzindo p análises nesta estrutura.

Apesar da validade de cada uma das p análises, desde que as observações sejam independentes, normalmente distribuídas com variância constante, os p testes dos efeitos de tratamentos nitidamente não são independentes, pois tem origem na mesma unidade experimental. Além disso, procedendo-se às p análises separadamente, torna-se inviável verificar a mudança do efeito de tratamentos com o tempo, ou seja, é inviável verificar existência da interação Tratamento x Tempo.

2.2. ANÁLISE DE VARIANÇIA CONSIDERANDO OS TEMPOS CONJUNTAMENTE

A Análise de Variância considerando-se os tempos conjuntamente segue o modelo de parcelas subdivididas (split-plot), cujo esquema é dado por:

Causas de Variação	G.L.
Entre Blocos	$r - 1$
Entre Tratamentos	$t - 1$
Erro(a)	$(r-1)(t-1)$
Entre Tempos	$p - 1$
Tempo x Tratamento	$(p-1)(t-1)$
Erro(b)	$t(r-1)(p-1)$

A Análise de Variância tem a familiar estrutura hierárquica com dois termos de erro. O primeiro (Erro(a)) representando a variação entre unidades experimentais e o segundo (Erro(b)) representando a variação dentro das unidades experimentais. A parte da parcela principal da análise é semelhante a análise de cada tempo separadamente; a variável, neste caso, é a soma das p respostas separadas.

A análise em parcelas subdivididas tem sido amplamente defendida por diversos autores, dentre os quais PEARCE (1953) que sugeriu sua adequabilidade para a análise de dados de plantas perenes. PEARCE também sugere que, desde que o tempo pode ser considerado como um fator sistemático, pode ser desejável particionar o Erro(b) como segue:

Causas de Variação	G.L.
Tempo x Bloco	$(p-1)(r-1)$
Tempo x Tratamento x Bloco	$(p-1)(t-1)(r-1)$

utilizando-se a interação de três fatores como Erro(b) propriamente dito. Este procedimento às vezes referido como

análise de experimentos em faixas (split-block) tem sido defendido por muitos autores, incluindo STEEL e TORRIE (1980). De uma forma ou de outra, para se testar o efeito de tratamentos, é correto utilizar o quadrado médio do Erro(a). Agora, a fim de justificar o uso do quadrado médio do Erro(b) em combinação com a distribuição F, para a parte da subparcela da análise, algumas pressuposições adicionais são necessárias.

A diferença fundamental entre delineamentos com observações repetidas e o verdadeiro delineamento em parcelas subdivididas é que a classificação da subparcela (tempo) não é casualizada para formar os arranjos. A casualização, além de produzir estimativas não viesadas para o erro experimental, também produz igualdade de covariâncias das classes de subparcelas. A condição de homogeneidade de covariâncias não é fácil de ser satisfeita em dados com observações repetidas, uma vez que observações mais próximas numa escala de tempo tendem a ser mais intimamente correlacionadas do que observações mais espaçadas no tempo.

Para justificar o uso da análise em parcelas subdivididas, dever-se-ia utilizar um teste para a homogeneidade de variâncias e covariâncias, às vezes também chamado de uniformidade. Um teste deste tipo foi proposto por WILKS (1946). A não uniformidade da matriz de variâncias e covariâncias provoca problemas na parte da subparcela da análise, onde os testes de Tempo e da interação Tempo x Tratamento não são válidos. Para contornar esses problemas nos testes, frequentemente utiliza-se as correções propostas por GREENHOUSE e GEISSER (1959) e HUYNH e FELDT (1970).

A alta frequência com que a falta de uniformidade tem sido detectada em dados experimentais, conduzem a que métodos alternativos de análise sejam utilizados.

2.3. ANÁLISE DE VARIÂNCIA MULTIVARIADA PARA MEDIDAS REPETIDAS

A análise de observações repetidas pode ser feita utilizando-se procedimentos multivariados e vários autores, dentre outros STEEL (1955), DEMPSTER (1963), DANFORD, HUGHES e McNEE (1960), tem adotado esses procedimentos. A validade dos métodos multivariados não é questionável pois supõem estrutura geral da matriz de variâncias e covariâncias e são úteis quando a condição de uniformidade de variâncias e covariâncias não se verifica. Esses métodos, no entanto, muitas vezes introduzem dificuldades de interpretação, ausentes em procedimentos univariados. Além disso, o número de ocasiões (tempos) pode exceder aos graus de liberdade do erro, o que inviabiliza o uso do procedimento.

2.4. ANÁLISE DE CONTRASTES ATRAVÉS DO TEMPO

O método proposto por ROWELL e WALTERS (1976) consiste no ajuste de equações polinomiais às sequências de dados e na análise separadamente de seus coeficientes através de uma série de Análises de Variância. Ou seja, o método consiste na avaliação e análise de funções lineares de variáveis sobre o tempo, sendo a escolha da função governada principalmente pelo interesse do pesquisador.

Num experimento com observações para p ocasiões (tempos), as funções consistem de $(p-1)$ contrastes ortogonais, então o procedimento é equivalente a particionar a seção da subparcela da análise na forma apresentada na Tabela 1.

A heterocedasticidade dos componentes da interação Tempo x Tratamento x Bloco, que é de fato o Erro(b) na análise de experimentos em faixas, invalida a

parte referente a subparcela da análise. Desta forma, cada componente da interação Tempo x Tratamento necessita ser testado através de um componente apropriado do Erro(b), não propriamente o Erro(b).

Tabela 1: Análise de variância em parcelas subdivididas com partição do efeito de Tempo, Tempo x Tratamento e Tempo x Tratamento x Bloco

Causas de Variação	G.L.
Entre Tratamentos	$t - 1$
Entre Blocos	$r - 1$
Erro(a)	$(t-1)(r-1)$
Entre Tempos	$p - 1$
Contraste (1)	1
⋮	⋮
Contraste (p-1)	1
Tempo x Tratamento	$(p-1)(t-1)$
Contraste (1) x Tratamento	$t-1$
⋮	⋮
Contraste (p-1) x Tratamento	$t-1$
Tempo x Bloco	$(p-1)(r-1)$
Tempo x Tratamento x Bloco	$(p-1)(t-1)(r-1)$
Contraste (1) x Tratamento x Bloco	$(t-1)(r-1)$
⋮	⋮
Contraste (p-1) x Tratamento x Bloco	$(t-1)(r-1)$

A análise particionada é obtida através da avaliação do contraste considerado para cada unidade experimental e procedendo-se uma análise de variância para os valores dos contrastes de acordo com o esquema:

Causas de Variação	G.L.
Fator de correção (Contraste)	1
Tratamento (x Contraste)	$t-1$
Bloco (x Contraste)	$r-1$
Tratamento x Bloco x Contraste	$(t-1)(r-1)$
Total	rt

O fator de correção, obtido da forma usual, consistiria num teste de significância para o efeito médio

do contraste em questão. Além disso, as somas de quadrados são divididas pela soma dos quadrados dos coeficientes dos contrastes.

Os contrastes adequados podem ser sugeridos por gráficos plotados para indicar a natureza da relação com o Tempo. Em curvas de crescimento, por exemplo, o uso de polinômios ortogonais pode ser apropriado. WISHART (1938) investigou os resultados de um experimento de nutrição em suínos, pela análise de coeficientes de regressão individuais para animais. STEVENS (1949), ao investigar os resultados de um ensaio de variedades de café, mediu a capacidade bianual, ajustando polinômios de somas e diferenças de resultados em anos sucessivos.

O procedimento de utilizar contrastes que são importantes ao experimento é mais prontamente interpretável que um procedimento multivariado. STELL (1955), para a análise de um ensaio com variedades, repetido num segundo ano nas mesmas parcelas, utilizou uma Análise de Variáveis Canônicas resultando no uso de funções lineares. Essas funções foram incapazes de interpretações práticas, como bem observou FINNEY (1956), segundo o qual o uso da soma e da diferença da produção de dois anos, conduziriam a interpretações mais claras.

Embora o uso do conjunto completo de $(p-1)$ contrastes ortogonais seja necessário para fornecer a decomposição completa da análise da subparcela, conforme consta na Tabela 1, o exame de algumas, não necessariamente ortogonais, funções de variáveis sobre o tempo pode fornecer, muitas vezes, informações suficientes na prática.

3. EXEMPLO.

O exemplo considerado no presente trabalho foi extraído de COLE e GRIZZLE (1966), cujos dados encontram-se na Tabela 2.

3.1. DADOS.

Dezesseis cães mestiços foram divididos em quatro grupos de tratamentos de quatro cães cada, com o objetivo de investigar o efeito das drogas morfina e trimetafam na liberação de histamina no sangue.

Os cães dos primeiros dois grupos receberam sulfato de morfina intravenosa e os cães restantes receberam trimetafam intravenoso. Além disso, os cães do segundo e quarto grupos foram tratados de modo que suprimentos de histamina disponível foram esvaziados no momento da inoculação com as drogas (tratamentos).

Os níveis de histamina no sangue foram observados nos dezesseis cães em 1, 3 e 5 minutos após a administração da droga.

Tabela 2: Níveis de histamina no sangue observados para diferentes tempos em cães submetidos a diferentes tratamentos

Tratamentos	Repetições (cães)	Tempos (min)		
		1	3	5
Grupo 1	1	0,20	0,10	0,08
	2	0,06	0,02	0,02
	3	1,40	0,48	0,24
	4	0,57	0,35	0,24
	média	0,86	0,24	0,15
Grupo 2	1	0,09	0,13	0,14
	2	0,11	0,10	0,14
	3	0,07	0,07	0,07
	4	0,07	0,06	0,07
	média	0,09	0,09	0,09
Grupo 3	1	0,62	0,31	0,22
	2	1,05	0,73	0,60
	3	0,83	1,07	0,80
	4	3,13	2,06	1,23
	média	1,41	1,04	0,71
Grupo 4	1	0,09	0,09	0,08
	2	0,09	0,09	0,10
	3	0,10	0,12	0,12
	4	0,05	0,05	0,05
	média	0,08	0,09	0,09

Utilizou-se a transformação logaritmo para proceder a análise dos dados.

3.2. ANÁLISE DOS DADOS

(a) ANÁLISE DE VARIÂNCIA NO MODELO DE PARCELA SUBDIVIDIDA

Causas de Variação	G.L.	Q.M.
Tratamento	3	2,6330
Erro(a)	12	0,3469
Tempo	2	0,1320
Tempo x Tratamento	6	0,0730
Erro(b)	24	0,0068

Como era de se esperar, o quadrado médio do Erro(a) excede ao quadrado médio do Erro(b).

(b) ANÁLISE DE VARIÂNCIA PARA CADA TEMPO

Tempo (min)	G.L.	Q.M. Erro
1	12	0,117
3	12	0,140
5	12	0,103

As variâncias (QM Erro) para cada tempo não aparentam serem heterogêneas.

(c) CORRELAÇÕES ENTRE OS TEMPOS

As correlações residuais (GL = 11) entre os tempos foram as seguintes:

	1 min	3 min	5 min
1 min	1,00	0,95	0,91
3 min		1,00	0,99
5 min		(simétrica)	1,00

A correlação é mais forte entre observações separadas por 2 minutos do que por 4 minutos. Esse fato evidencia a inadequabilidade do modelo de parcela subdividida (ou de experimentos em faixas) que supõe correlação constante.

(d) TESTE DA VALIDADE DO MODELO DE PARCELA SUBDIVIDIDA

Utilizando-se o teste de WILKS (1946) obteve-se $\chi^2 = 20$ com G.L. = 4, significativo ao nível de significância de 5%, pois $\chi^2_{.05(4)} = 9,49$.

O teste mostra desvios significativos de uniformidade da matriz de variâncias e covariâncias, isto é, as variâncias para cada tempo foram heterogêneas e/ou as

correlações foram diferentes. No presente caso, a diferença entre as correlações parece ser o fator determinante desta não uniformidade.

(e) ANÁLISE DE CONTRASTES SOBRE O TEMPO

Em muitas situações, polinômios ortogonais (linear, quadrático, etc) podem ser adequados para estabelecer contrastes sobre o tempo. No presente caso, desde que as observações foram tomadas em três tempos igualmente espaçados (1, 3 e 5 min), os dois contrastes escolhidos foram:

Contrastes	Tempo (min)			S.Q. dos coeficientes
	1	3	5	
Linear	-1	0	1	2
Quadrático	1	-2	1	6
Total	1	1	1	3

A análise das três funções fornece a análise de parcela subdividida por inteiro, sendo que o total fornece a parte da parcela principal.

A soma dos produtos de cada par dos coeficientes dos contrastes é zero, indicando ortogonalidade.

(f) ANÁLISES DE VARIANCIA SUCESSIVAS

As três funções lineares (contrastos) foram calculadas para cada uma das 16 unidades experimentais (cães) e três Análises de Variância foram realizadas dividindo-se as Somas de Quadrados por 2, 6 e 3, que representam as somas de quadrados dos coeficientes dos

contrastes, obtendo-se os seguintes resultados:

Causas de Variação	G.L.	Total	Q.M.	
			Linear	Quadrático
Fator de correção	1	-	0,2566	0,0074
Entre Tratamentos	3	2,6330	0,1375	0,0089
Dentro de Tratamentos (Erro)	12	0,3469	0,0104	0,0032

(g) ANÁLISE COMPLETA

A combinação das análises do item (f) fornece a análise em parcela subdividida que foi apresentada em (a), agora com as devidas partições. As Soma de Quadrados de Tempo, da interação Tempo x Tratamento e do Erro(b) foram particionadas em componentes Linear e Quadrático, como segue:

Causas de Variação	G.L.	S.Q.	Q.M.
Tratamento	3		2,6330
Erro(a)	12		0,3469
Tempo	2	0,2640	0,1320
Linear	1	0,2566	0,2566
Quadrático	1	0,0074	0,0074
Tempo x Tratamento	6	0,4392	0,0731
Linear x Tratamento	3	0,4125	0,1375
Quadrático x Tratamento	3	0,0267	0,0089
Erro(b)	24	0,1632	0,0068
Componente Linear	12	0,1248	0,0104
Componente Quadrático	12	0,0384	0,0032

Os dois componentes do Erro(b) (0,0104 e 0,0032) mostram sinais evidentes de heterocedasticidade, enfatizando a prudência de particionar o Erro(b).

As partições das Somas de Quadrados procedidas nas análises conduzem a resultados idênticos àqueles obtidos por COLE e GRIZZLE (1966), utilizando análise de variância multivariada.

4. DISCUSSÃO

A simplicidade de cálculo e de interpretação do método de análise de contrastes sobre o tempo, sugerem a utilidade do método para dados com observações repetidas. Para experimentos pequenos, entretanto, os quadrados médios dos erros para os contrastes serão baseados em poucos graus de liberdade. Em tal caso, seria importante investigar a validade da análise em parcelas subdivididas (ou de experimentos em faixas), na qual uma estimativa total do erro baseia-se substancialmente em maior número de graus de liberdade.

Os contrastes adequados a serem utilizados podem ser sugeridos por gráficos construídos para verificar a natureza da relação da resposta com o tempo. Em muitos casos, polinômios ortogonais podem ser recomendados e sua eficácia dependerá da porção da variância do erro que foi decomposta. Apesar de, em muitos casos, polinômios de baixa ordem serem suficientes, não é pouco frequente, para dados com medidas repetidas na agricultura, a exigência de coeficientes de ordem elevada, que são de difícil interpretação. Daí propostas que objetivam eliminar esta dificuldade, como a apresentada por EVANS e ROBERTS (1979).

5. BIBLIOGRAFIA

- COLE, J.W. & GRIZZLE, J.E. Applications of multivariate analysis of variance to repeated measurements experiments. *Biometrics*, 22: 810-28, 1966.
- DANFORD, M.B., HUGHES, H.M. & McNEE, R.C. On analysis of repeated-measurements experiments. *Biometrics*, 16: 547-65, 1960.
- DEMPSTER, A.P. Stepwise multivariate analysis of variance based on principal variables. *Biometrics*, 19: 478-90, 1963.
- EVANS, J.C. & ROBERTS, E.A. Analysis of sequential observations with applications to experiments on grazing animals and perennial plants. *Biometrics*, 35: 687-93, 1979.
- FINNEY, D.J. Multivariate analysis and agricultural experiments. *Biometrics*, 12: 67-71, 1956.
- GREENHOUSE, S.W. & GEISSER, S. On methods in analysis of profile data. *Psychometrika*, 24: 95-112, 1959.
- HUYNH, H. & FELDT, L.S. Conditions under which mean square ratios in repeated measurement designs have exact F-distributions. *Journal of the American Statistical Association*, 65: 1582-89, 1970.

- PEARCE, S.C. Field experimentation with fruit trees and perennial plants. Commonwealth Bureau of Horticulture and Plantation Crops, Kent, 1953.
- ROWELL, J.G. & WALTERS, D.E. Analysing data with repeated observations on each experimental unit. *Journal of Agricultural Science*, 87; 423-32, 1976.
- STEEL, R.G.D. An analysis of perennial crop data. *Biometrics*, 11: 201-12, 1955.
- STEEL, R.G.D. & TORRIE, J.H. Principles and procedures of statistics. 2. ed. New York, McGraw-Hill, 1980. 633 p.
- STEVENS, W.L. Análise estatística do ensaio de variedades de café. *Bragantia*, 9: 103-23, 1949.
- WILKS, S. S. Sample criterion for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Annals of Mathematical Statistics*, 17: 257-81, 1946.
- WISHART, J. Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika*, 30: 16-28, 1938.

Publicações do Instituto de Matemática da UFRGS
Cadernos de Matemática e Estatística

Série F: Trabalho de Divulgação

1. Nubem A. C. Medeiros e Jaime B. Ripoll - Superfícies Invariantes - MAI/90.
2. Sídia Maria Callegari-Jacques - Análise de Correspondência: Aplicações em Genética - DEZ/91
3. Jandira M.G.Fachel et al - Correspondence Analysis Applied to Ethnographic Data : Case Examples - JAN/92
4. João Riboldi e Dinara W. Xavier Fernandez - Análise de Observações Repetidas Através de contrastes no Tempo - JUL/92
5. Maria Teresa Albanese e Martin Knott - Twomiss: a Computer Program For Fitting a one or Two-Factor Logit-Probit Latent Variable Model to Binary Data When Observations May be Missing - JUL/92
6. Vera Clotilde Carneiro - Retrato Atual do Curso de Licenciatura em Matemática da UFRGS e Plano de Ações Renovadoras - JUL/92

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
NÚCLEO DE ATIVIDADES EXTRA CURRICULARES

Os Cadernos de Matemática e Estatística publicam as seguintes séries:

Série A: Trabalho de Pesquisa

Série B: Trabalho de Apoio Didático

Série C: Colóquio de Matemática SBM/UFRGS

Série D: Trabalho de Graduação

Série F: Trabalho de Divulgação

Série G: Textos para Discussão

Toda correspondência com solicitação de números publicados e demais informações deverá ser enviada para:

NAEC - NÚCLEO DE ATIVIDADES EXTRA CURRICULARES
INSTITUTO DE MATEMÁTICA - UFRGS
AV. BENTO GONÇALVES, 9500 - PRÉDIO 43111
CEP 91509 - 900 AGRONOMIA - POA/RS
FONE: 336 92 22 OU 339 13 55 OU 228 16 33
RAMAL 6197
FAX: 336 15 12