**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Modelos lineares mistos para análise de dados longitudinais
bivariados provenientes de ensaios agropecuários**

Simone Silmara Werner Gurgel do Amaral

Tese apresentada para obtenção do título de Doutora em
Ciências. Área de concentração: Estatística e Experi-
mentação Agronômica

**Piracicaba
2013**

Simone Silmara Werner Gurgel do Amaral
Bacharel e licenciada em Ciências Biológicas
Licenciada em Matemática

**Modelos lineares mistos para análise de dados longitudinais bivariados
provenientes de ensaios agropecuários**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **CÉSAR GONÇALVES DE LIMA**

Tese apresentada para obtenção do título de Doutora em
Ciências. Área de concentração: Estatística e Experi-
mentação Agronômica

**Piracicaba
2013**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Amaral, Simone Silmara Werner Gurgel do

Modelos lineares mistos para análise de dados longitudinais bivariados provenientes de ensaios agropecuários / Simone Silmara Werner Gurgel do Amaral. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2013.
78 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2013.

1. Armazenamento de produtos 2. Heterogeneidade de variâncias 3. Correlação
4. Modelos mistos bivariados 5. Dados longitudinais I. Título

CDD 519.50285
A485m

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

Aos meus pais,

*Onóbio Vicente Werner e
Delci Reis Monteiro Werner*

Às minhas irmãs,

*Suelem Samara Werner e
Milena Maiara Werner*

Ao meu companheiro,

*Marcus Vinicius Silva Gurgel do
Amaral*

Com amor, DEDICO.

AGRADECIMENTOS

Primeiramente agradeço a Deus, em quem acredito e confio, pela sua presença constante em minha vida, por ter me dado coragem e sabedoria para enfrentar os obstáculos.

Aos meus pais, Onóbio Vicente Werner e Delci Reis Monteiro Werner, por todo apoio e confiança, sem os quais a realização desse sonho não seria possível.

Ao Prof. Dr. César Gonçalves de Lima, pela orientação, amizade e confiança depositada. Sua paciência, humildade, respeito e carinho demonstrados desde a primeira conversa sempre serão um grande exemplo.

Aos professores do programa de Pós-graduação em Estatística e Experimentação Agronômica, Dr. Carlos Tadeu dos Santos Dias, Dra. Clarice Garcia Borges Demétrio, Dr. Décio Barbin, Dr. Gabriel Adrián Sarriés, Dr. Idemauro Antonio Rodrigues de Lara, Dr. Paulo Justiniano Ribeiro Jr., Dra. Roseli Aparecida Leandro, Dr. Sílvio Zocchi, Dra. Sônia Maria de Stefano Piedade, Dra. Taciana Villela Savian, por todos os ensinamentos e amizade.

Aos funcionários do Departamento, Luciane Brajão, Mayara Segatto, Solange Sabadin, Eduardo Bonilha, Jorge Wiendl e Rosni Honofre.

Ao CNPq, pelo fundamental suporte financeiro concedido.

A todos os alunos do curso de Pós-Graduação em Estatística e Experimentação Agronômica da ESALQ/USP, com os quais tive o prazer de conviver, pela amizade, pelo auxílio em momentos de dúvidas, a troca de conhecimentos e atenção recebida em todos os momentos.

Ao meu marido Marcus Vinicius, por colocar um sorriso no meu rosto em muitos momentos durante esses cinco anos, por toda dedicação e carinho, por ser capaz de confortar meu coração e dar a minha vida ainda mais sentido, pela decisão de compartilhar sua vida comigo para que possamos crescer juntos.

A todos que, de forma direta ou indireta, contribuíram para a realização deste trabalho.

*“Se a riqueza é um bem desejável na vida,
que há de mais rico que a sabedoria que tudo criou?”*

Sabedoria 8,5

*“Better to have estimated a correlation and found it to be zero than never
to have estimated a correlation at all.”*

Robert E. Weiss

SUMÁRIO

RESUMO	11
ABSTRACT	13
LISTA DE FIGURAS	15
LISTA DE TABELAS	17
1 INTRODUÇÃO	19
2 DESENVOLVIMENTO	21
2.1 Dados longitudinais univariados	21
2.2 Dados longitudinais multivariados	23
2.3 O modelo linear misto para dados longitudinais	24
2.3.1 Modelo linear misto em dois estágios	25
2.3.2 Exemplos de estruturas das matrizes de covariância	28
2.3.3 Estimação no modelo marginal	30
2.3.4 Seleção de modelos	32
2.3.4.1 Teste da razão de verossimilhanças	33
2.3.4.2 Critérios de informação	34
2.3.4.3 Teste de Wald aproximado	34
2.3.4.4 Teste F aproximado	35
2.3.5 Análise de Resíduos	35
2.4 O modelo linear misto para dados longitudinais bivariados	37
2.4.1 Diagnóstico para os modelos bivariados	41
2.5 O modelo linear misto para dados longitudinais multivariados	43
3 MATERIAL E MÉTODOS	45
3.1 Material	45
3.2 Métodos	46
3.2.1 Estratégia de análise	48
3.2.2 Recursos computacionais	52
3.2.3 Análise de resíduos	53
4 RESULTADOS E DISCUSSÃO	55
5 CONSIDERAÇÕES FINAIS	67
REFERÊNCIAS	69
ANEXOS	73

RESUMO

Modelos lineares mistos para análise de dados longitudinais bivariados provenientes de ensaios agropecuários

Em estudos longitudinais, repetidas observações de uma mesma variável resposta são coletadas na mesma unidade experimental, em diferentes ocasiões. Como diferentes observações são realizadas na mesma unidade, espera-se que estas sejam correlacionadas, e que exista uma heterogeneidade de variâncias nas diferentes ocasiões. Dados longitudinais multivariados são obtidos quando um conjunto de diferentes variáveis respostas são mensuradas na mesma unidade experimental repetidas vezes ao longo do tempo; nesse caso, além da correlação entre observações realizadas na mesma unidade experimental, deve-se considerar também a correlação entre diferentes variáveis respostas. Uma forma de analisar dados longitudinais bivariados é empregar um modelo misto para cada uma das variáveis respostas e uni-los em um modelo misto bivariado especificando a distribuição conjunta para os efeitos aleatórios. As estimativas dos parâmetros desta distribuição comum podem ser usadas para avaliar a relação entre as diferentes respostas. Para exemplificar a utilização da técnica, foram utilizados dados de armazenamento de leite UAT. Os modelos lineares mistos bivariados foram ajustados por meio do *software* SAS e a análise gráfica foi realizada por meio do *software* R. Para seleção dos modelos empregou-se os Critérios de Informação de Akaike (AIC) e Bayesiano (BIC), e o teste da razão de verossimilhanças para comparação de modelos encaixados. A utilização do modelo linear misto bivariado permitiu modelar a heterogeneidade de variâncias entre ocasiões e a correlação entre diferentes medidas na mesma unidade experimental, bem como a correlação entre as variáveis respostas.

Palavras-chave: Armazenamento de produtos; Heterogeneidade de variâncias; Correlação; Modelos mistos bivariados; Dados longitudinais

ABSTRACT

Linear mixed models in the analysis bivariate longitudinal data from agricultural essays

In longitudinal studies, repeated measurements of a response variable are taken in the same experimental unit over time. . Since different observations are measured on the same experimental unit, it is expected that there is correlation among the repeated measurements and heterogeneity of variances in different occasions. Multivariate Longitudinal Data are obtained when we measure a number of different response variables in the same experimental unit repeatedly over time; in this case, we should also observe a correlation between the different response variables. One way to analyze bivariate longitudinal data is to use a mixed model for each of the response variables, and unite them in bivariate mixed models specifying the joint distribution for random effects. Parameter estimates of this common distribution may be used to evaluate the relationship between different responses. As an example of the use of the technique, UHT milk storage data were used. Models were fitted using SAS software and the graphical analysis was done with software R. To model selection, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used, and maximum likelihood ratio test was used to compare nested models. The use of bivariate mixed linear model allowed to model the heteroscedasticity of the occasions, the correlation between the different measurements in the same experimental unit and also the correlation between the different response variables.

Keywords: Product storage; Heteroscedasticity; Correlation; Bivariate mixed models; Longitudinal data

LISTA DE FIGURAS

Figura 1 - Perfis individuais das concentrações de α_{s1} -caseína e β -caseína no leite UAT, por grupo	47
Figura 2 - (a) Boxplot bivariado para os resíduos da variável α_{s1} -caseína versus resíduos da variável β -caseína; (b) Gráfico quantil-quantil para a distância de Mahalanobis versus os quantis da quiquadrado	59
Figura 3 - Gráfico quantil-quantil para os resíduos condicionais padronizados para o modelo selecionado: (a) para a variável α -caseína; (b) para a variável β -caseína	59
Figura 4 - Perfis individuais das concentrações de α_{s1} -caseína e β -caseína no leite UAT, por grupo, retirando algumas das observações	60
Figura 5 - (a) Gráfico boxplot bivariado para os resíduos da variável α_{s1} -caseína versus resíduos da variável β -caseína; (b) Gráfico quantil-quantil para a distância de mahalanobis versus os quantis da qui quadrado	62
Figura 6 - Gráfico quantil-quantil para os resíduos condicionais padronizados para o modelo selecionado: (a) para a variável α -caseína; (b) para a variável β -caseína	63
Figura 7 - Gráfico de valores preditos versus resíduos condicionais padronizados para o modelo selecionado: (a) para a variável α -caseína; (b) para a variável β -caseína	63
Figura 8 - Gráfico de valores observados versus valores preditos para o modelo selecionado: (a) para a variável α -caseína; (b) para a variável α -caseína	64
Figura 9 - Gráfico dos EBLUPs (<i>empirical best linear unbiased predictor</i>) padronizados versus unidades experimentais	64
Figura 10 - Gráfico de perfis estimados segundo o modelo selecionado para as variáveis α -caseína e β -caseína	66

LISTA DE TABELAS

Tabela 1 - Estrutura de dados longitudinais balanceados e completos para o caso de g tratamentos, cinco repetições por tratamento e quatro medidas repetidas no tempo	22
Tabela 2 - Estrutura de dados longitudinais bivariados balanceados e completos para o caso de g tratamentos, cinco repetições por tratamento e quatro medidas repetidas no tempo	24
Tabela 3 - Utilização dos resíduos para fins de diagnóstico	37
Tabela 4 - Médias e erros padrões (e.p.) da concentração de α_{s1} -caseína e β -caseína em mg/mL do leite AUT nos cinco tempos de estocagem, por grupo	46
Tabela 5 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo método REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios	55
Tabela 6 - Número de parâmetros de efeito fixo (nf), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo método ML (LV) dos modelos ajustados, visando a escolher os efeitos fixos	57
Tabela 7 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo métodos REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios, após a seleção da parte fixa	58
Tabela 8 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo métodos REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios	61
Tabela 9 - Número de parâmetros de efeito fixo (nf), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo método ML (LV) dos modelos ajustados, visando a escolher os efeitos fixos	61

Tabela 10 -Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo métodos REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios, após a seleção da parte fixa . . . 62

1 INTRODUÇÃO

Os planejamentos longitudinais, em que a variável resposta é medida repetidas vezes na mesma unidade experimental, são uma prática comum em muitas áreas. Para análise deste tipo de dados é razoável considerar que existe uma correlação entre diferentes observações feitas em uma mesma unidade experimental, de tal forma que o uso do modelo linear clássico não é adequado neste caso, por pressupor independência entre as medidas repetidas. Além disso, é comum encontrar heterogeneidade de variâncias nas diversas ocasiões em que são feitas as medidas, fazendo com que a pressuposição de homocedasticidade do modelo linear clássico também não seja satisfeita. O modelo linear misto (ou modelo de efeitos aleatórios) é uma opção na modelagem deste tipo de dados, por possibilitar a utilização de diferentes estruturas para matrizes de covariância, que permitem modelar tanto a correlação entre as medidas repetidas quanto a heterocedasticidade.

Na maior parte dos estudos longitudinais muitas variáveis são coletadas na primeira ocasião, e muitas delas são reavaliadas nas ocasiões seguintes, o que produz um conjunto de dados longitudinais multivariado. Nesses casos, além da covariância existente entre observações referentes à mesma unidade experimental ao longo do tempo tem-se a covariância entre as diferentes variáveis respostas mensuradas (WEISS, 2005).

De acordo com Fieuws, Verbeke e Molenberghs (2007), quando mais de uma variável resposta está presente pode-se empregar um modelo misto para cada variável resposta e uni-los em uma classe de modelos mistos multivariados especificando a distribuição conjunta para os efeitos aleatórios. Tal abordagem é particularmente útil quando o interesse não está em descrever o comportamento individual de cada uma das variáveis respostas ou descrever o comportamento de uma das variáveis em função de outras (utilizando como variáveis explicativas), mas sim verificar a evolução conjunta das variáveis no decorrer do tempo.

Em trabalhos com armazenamento de produtos agropecuários é comum a mensuração de diversas variáveis nas mesmas unidades experimentais no decorrer do tempo, sendo de interesse do pesquisador conhecer a evolução conjunta dessas variáveis que irão determinar a qualidade do produto. Neste caso, conhecer o comportamento das variáveis respostas isoladamente não será suficiente, pois no processo de envelhecimento do produto as variáveis interagem.

A proposta deste trabalho é avaliar a aplicação da metodologia de análise

conjunta de dados longitudinais multivariados no contexto de um estudo que envolve variáveis respostas com distribuição normal inerentes ao armazenamento de produtos agropecuários. No ajuste dos modelos lineares mistos bivariados utilizou-se o procedimento mixed implementado no software SAS 9.2 (SAS INSTITUTE, 2008) e na análise de resíduos, algumas funções do *software* R (R DEVELOPMENT CORE TEAM, 2012).

2 DESENVOLVIMENTO

Nesta seção concentram-se os principais aspectos da teoria de modelos lineares mistos para análise de dados longitudinais bem como uma revisão dos principais artigos que contribuíram para o desenvolvimento dessa teoria; posteriormente faz-se uma revisão dos principais artigos que apresentam a utilização de modelos mistos na análise de dados longitudinais multivariados.

2.1 Dados longitudinais univariados

Em estudos longitudinais é comum coletar observações repetidas de uma variável resposta, em diferentes ocasiões, nas mesmas unidades experimentais. Essas unidades experimentais podem estar agrupadas ou recebendo níveis diferentes de um ou mais fatores de tratamento. Os dados provenientes de tais estudos são considerados: **regulares** (em relação ao tempo), quando é constante o intervalo de tempo entre as ocasiões subsequentes; **balanceados**, se as observações em todas as unidades experimentais forem feitas nos mesmos instantes de tempo e **completos**, quando o experimento não apresentar nenhuma observação perdida (LIMA, 1996). A Tabela 1 apresenta uma estrutura de dados longitudinais balanceados e completos, em que cada linha representa um perfil individual de respostas.

O modelo mais simples usado para analisar dados de ensaios com medidas repetidas é o modelo associado aos experimentos em parcela subdividida, dado por:

$$y_{ijk} = \mu + \alpha_k + \gamma_{ik} + \beta_j + (\alpha\beta)_{jk} + \varepsilon_{ijk} \quad (1)$$

na qual y_{ijk} representa o valor observado para a variável resposta no i -ésimo indivíduo, submetido ao k -ésimo tratamento e avaliado no j -ésimo tempo; μ é uma constante comum a todas as observações; α_k é o efeito do k -ésimo tratamento, γ_{ik} é o erro associado às parcelas com $\gamma_{ik} \sim N(0; \sigma_a^2)$; β_j é o efeito do j -ésimo tempo; $(\alpha\beta)_{jk}$ é o efeito da interação do k -ésimo tratamento com o j -ésimo tempo e ε_{ijk} é o erro associado às subparcelas, com $\varepsilon_{ijk} \sim N(0; \sigma^2)$. No caso geral: $i = 1, \dots, N_k$; $j = 1, \dots, n_i$ e $k = 1, \dots, g$, em que o total de indivíduos é dado por $N = \sum_{k=1}^g N_k$. Para experimentos balanceados e completos, assume-se $n_i = t$ ocasiões (LIMA, 1996).

Para o modelo (1), admite-se que os níveis do fator primário (tratamentos) foram aplicados nas parcelas, nas quais são tomadas medidas repetidas em ocasiões suces-

Tabela 1 - Estrutura de dados longitudinais balanceados e completos para o caso de g tratamentos, cinco repetições por tratamento e quatro medidas repetidas no tempo

Tratamentos	Unidade Experimental	Ocasões			
		1	2	3	4
1	1	y_{111}	y_{112}	y_{113}	y_{114}
1	2	y_{121}	y_{122}	y_{123}	y_{124}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	5	y_{151}	y_{152}	y_{153}	y_{154}
2	1	y_{211}	y_{212}	y_{213}	y_{214}
2	2	y_{221}	y_{222}	y_{223}	y_{224}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	5	y_{251}	y_{252}	y_{253}	y_{254}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
g	1	y_{g11}	y_{g12}	y_{g13}	y_{g14}
g	2	y_{g21}	y_{g22}	y_{g23}	y_{g24}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
g	5	y_{g51}	y_{g52}	y_{g53}	y_{g54}

sivas (o fator secundário é o tempo), e que essas medidas tomadas em ocasiões distintas têm variâncias homogêneas e são igualmente correlacionadas.

Assumir variâncias homogêneas em todas as ocasiões e covariâncias iguais entre quaisquer ocasiões é o mesmo que assumir que a matriz de covariâncias entre os tempos é do tipo simetria composta e envolve somente 2 parâmetros. No caso particular de 4 ocasiões, por exemplo, essa estrutura de covariâncias é dada por:

$$\begin{pmatrix} \sigma^2 + \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma^2 + \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma^2 + \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma^2 + \sigma_a^2 \end{pmatrix}$$

Uma condição necessária e suficiente para que possa ser empregada a análise de variância univariada no esquema de parcelas subdivididas para um experimento com medidas repetidas é que a matriz de covariâncias entre os tempos satisfaça à condição de esfericidade ou circularidade, o que é equivalente a especificar que as variâncias entre pares de erros associados a diferentes ocasiões sejam todas iguais (HUYNH; FELDT, 1970).

Para se testar a condição de esfericidade, pode-se utilizar o teste de Mauchly (1940). Caso o teste resulte em não significativo, conclui-se que a matriz de covariâncias

é do tipo esférica, podendo os dados do experimento serem analisados pelo esquema de parcelas subdivididas; caso contrário recomenda-se o uso da análise multivariada de perfis, ou a utilização da análise univariada adotando-se modelos mistos.

Na análise multivariada de perfis, as respostas obtidas em cada ocasião são consideradas observações de uma nova variável sendo realizada uma análise de variância multivariada (MANOVA) para comparar as médias dos tratamentos, dos tempos e da interação entre estes fatores. Nesse caso, a matriz utilizada para modelar a covariância entre os tempos é do tipo não estruturada e envolve $t(t+1)/2$ parâmetros. Para o caso em que $t = 4$, tem-se a seguinte matriz, que envolve 10 parâmetros:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

Uma grande limitação da análise multivariada de perfis é o fato de somente trabalhar com perfis individuais de respostas que sejam balanceados e completos. Uma vez que estudos longitudinais podem durar dias, semanas, meses ou até mesmo anos, é comum a ocorrência de dados irregulares no tempo, não balanceados ou incompletos, e por esse motivo, tem havido maior interesse no desenvolvimento de modelos estatísticos que possam ser utilizados em casos mais gerais. Os modelos (lineares ou não lineares) mistos, além de permitirem a utilização de diversos tipos de estruturas de covariâncias, permitem trabalhar com dados irregulares no tempo, não balanceados e incompletos (LIMA, 1996).

2.2 Dados longitudinais multivariados

Dados longitudinais multivariados são obtidos quando mais de uma variável resposta é mensurada na mesma unidade experimental (indivíduo), em repetidas ocasiões de tempos (FIEUWS, 2006). Nesse caso, além da covariância existente entre observações de uma variável resposta feitas na mesma unidade experimental ao longo do tempo, tem-se a correlação entre as diferentes variáveis respostas mensuradas. A Tabela 2 apresenta uma estrutura geral de dados longitudinais bivariados balanceados e completos, em que y_{rijl} representa a mensuração da r -ésima variável resposta ($r = 1, 2$) no i -ésimo indivíduo ($i = 1, 2, \dots, 5$), na j -ésima ocasião ($j = 1, 2, 3, 4$) para o k -ésimo tratamento ($k = 1, 2, \dots, g$).

Tabela 2 - Estrutura de dados longitudinais bivariados balanceados e completos para o caso de g tratamentos, cinco repetições por tratamento e quatro medidas repetidas no tempo

Tratamento	Unidade Experimental	Variável							
		1				2			
		Ocasão				Ocasão			
		1	2	3	4	1	2	3	4
1	1	y_{1111}	y_{1112}	y_{1113}	y_{1114}	y_{2111}	y_{2112}	y_{2113}	y_{2114}
1	2	y_{1121}	y_{1122}	y_{1123}	y_{1124}	y_{2121}	y_{2122}	y_{2123}	y_{2124}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	5	y_{1151}	y_{1152}	y_{1153}	y_{1154}	y_{2151}	y_{2152}	y_{2153}	y_{2154}
2	1	y_{1211}	y_{1212}	y_{1213}	y_{1214}	y_{2211}	y_{2212}	y_{2213}	y_{2214}
2	2	y_{1221}	y_{1222}	y_{1223}	y_{1224}	y_{2221}	y_{2222}	y_{2223}	y_{2224}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	5	y_{1251}	y_{1252}	y_{1253}	y_{1254}	y_{2251}	y_{2252}	y_{2253}	y_{2254}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
g	1	y_{1g11}	y_{1g12}	y_{1g13}	y_{1g14}	y_{2g11}	y_{2g12}	y_{2g13}	y_{2g14}
g	2	y_{1g21}	y_{1g22}	y_{1g23}	y_{1g24}	y_{2g21}	y_{2g22}	y_{2g23}	y_{2g24}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
g	5	y_{1g51}	y_{1g52}	y_{1g53}	y_{1g54}	y_{2g51}	y_{2g52}	y_{2g53}	y_{2g54}

Como no caso univariado, existe a possibilidade de se utilizar uma análise multivariada de perfis para dados longitudinais multivariados, por meio de um modelo “duplamente” multivariado - MDM. Admitindo, como exemplo, a estrutura de dados apresentada na Tabela 2, as 4 medidas repetidas de cada variável resposta formariam uma matriz de dados com 8 colunas de variáveis respostas. Neste caso, a matriz de variâncias e covariâncias do modelo duplamente multivariado teria dimensões 8×8 , não estruturada, com $(8)(8+1)/2 = 36$ parâmetros. No caso geral, a matriz de variâncias e covariâncias só pode ser estimada se o número de unidades experimentais for maior do que rt (ROY; KHATTREE, 2006). Mais uma vez, a grande restrição desse tipo de análise é o fato de apenas trabalhar com perfis de respostas balanceados e completos (ABDELMUR, 2009).

2.3 O modelo linear misto para dados longitudinais

O modelo linear misto inclui parâmetros de efeito fixo associados a uma ou mais covariáveis (contínuas ou categóricas) e de efeito aleatório associados a um ou mais fatores aleatórios (WEST; WELCH; GALECKI, 2007).

Os modelos mistos têm várias características desejáveis para a análise de dados longitudinais: podem ser utilizados para diferentes estruturas de dados, incluindo dados irregulares no tempo, não balanceados e incompletos; permitem a modelagem explícita da variação entre e intra indivíduos e, além disso, muitas vezes, os parâmetros individuais têm uma interpretação natural relevante para os objetivos do estudo (LAIRD; WARE, 1982).

2.3.1 Modelo linear misto em dois estágios

Seja a variável aleatória Y_{ij} , que denota a resposta para o i -ésimo indivíduo, mensurada no tempo t_{ij} , $i = 1, \dots, N$, $j = 1, \dots, n_i$, e $\mathbf{Y}_i$ o vetor aleatório de dimensão n_i de todas as medidas repetidas para o i -ésimo indivíduo, ou seja $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{in_i})'$.

O modelo linear misto de dois estágios foi definido por Laird e Ware (1982), assumindo no primeiro estágio que $\mathbf{Y}_i$ satisfaz o modelo de regressão dado por (VERBEKE; MOLENBERGHS, 2000):

$$\mathbf{Y}_i = \mathbf{Z}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i \quad (2)$$

em que $\mathbf{Z}_i$ é a matriz $n_i \times q$ de covariáveis conhecidas, que modela a evolução da variável resposta para o i -ésimo indivíduo; $\boldsymbol{\beta}_i$ é um vetor de dimensão q de coeficientes de regressão desconhecidos e específicos para cada indivíduo, e $\boldsymbol{\varepsilon}_i$ é o vetor de dimensão n_i formado pelos componentes de resíduo ε_{ij} , com $\boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{\Sigma}_i)$. Usualmente, assume-se que os $\boldsymbol{\varepsilon}_i$ são independentes e $\boldsymbol{\varepsilon}_i \sim N(0, \sigma^2 \mathbf{I})$.

Em um segundo estágio, assume-se que $\boldsymbol{\beta}_i$ pode ser escrito como:

$$\boldsymbol{\beta}_i = \mathbf{K}_i \boldsymbol{\beta} + \mathbf{b}_i \quad (3)$$

em que $\mathbf{K}_i$ é uma matriz de dimensões $q \times p$ de covariáveis conhecidas, e $\boldsymbol{\beta}$ é o vetor de dimensão p correspondentes aos parâmetros de regressão desconhecidos e $\mathbf{b}_i$ é um vetor de dimensão $q \times 1$. Assume-se que os $\mathbf{b}_i$ são independentes e $\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{D})$, o que pode explicar a variabilidade entre os indivíduos.

No intuito de combinar os modelos provenientes da formulação do modelo misto em dois estágios, substitui-se $\boldsymbol{\beta}_i$ definido em (3) na expressão (2) obtendo-se o modelo:

$$\begin{aligned}
\mathbf{Y}_i &= \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \\
\mathbf{b}_i &\sim N(\mathbf{0}, \mathbf{D}) \\
\boldsymbol{\varepsilon}_i &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_i) \\
\mathbf{b}_1, \dots, \mathbf{b}_N \text{ e } \boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_N, &\text{ independentes}
\end{aligned} \tag{4}$$

em que $\mathbf{X}_i = \mathbf{Z}_i\mathbf{K}_i$ representa a matrix $n_i \times p$ de covariáveis conhecidas. O modelo (4) é chamado de **modelo linear misto**, com parâmetros de efeito fixo $\boldsymbol{\beta}$ e parâmetros de efeito aleatório $\mathbf{b}_i$. Ou seja, assume-se que o vetor de medidas repetidas para cada indivíduo é representado por um modelo de regressão linear em que alguns parâmetros de regressão são específicos da população (ou seja, assumem os mesmos valores para todos os indivíduos), e alguns parâmetros são específicos a cada indivíduo (ou seja, assumem valores diferentes para os diferentes indivíduos).

O modelo linear misto (4) implica em:

$$\mathbf{Y}_i \sim N(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{V}_i)$$

em que $\mathbf{V}_i = \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \boldsymbol{\Sigma}_i$ representa a matriz de variâncias e covariâncias do modelo marginal.

Por exemplo: Seja um experimento em que as unidades experimentais foram avaliadas em quatro ocasiões (t_1, t_2, t_3, t_4) , sendo o comportamento dos dados em relação ao tempo descrito por uma reta, com efeitos aleatórios no intercepto e no coeficiente angular. Para cada uma das unidades experimentais será ajustada uma reta, e obtidos os coeficientes lineares (β_{0i}) e angulares (β_{1i}) . Utilizando esses coeficientes é obtido o vetor $\boldsymbol{\beta}$ e os componentes de variâncias associados ao vetor $\boldsymbol{\beta}$. Assim tem-se:

$$\begin{aligned}
\mathbf{Y}_i &= \begin{bmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \\ y_{4i} \end{bmatrix} \quad \mathbf{Z}_i = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{bmatrix} \quad \boldsymbol{\beta}_i = \begin{bmatrix} \beta_{0i} \\ \beta_{1i} \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} \sigma_{b_0}^2 & 0 \\ 0 & \sigma_{b_1}^2 \end{bmatrix} \\
\mathbf{X}_i &= \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad \boldsymbol{\Sigma}_i = \mathbf{I}_t\sigma^2
\end{aligned}$$

Dessa forma, admitindo-se uma variabilidade (entre indivíduos) associada aos parâmetros β_0 e β_1 (representada pela matriz $\mathbf{D}$), e que as medidas repetidas (intra-indivíduos) são condicionalmente independentes e apresentam variância constante (representada pela matriz Σ_i), a matriz de variâncias e covariâncias para o modelo marginal é dada por:

$$\begin{aligned}
 \mathbf{V}_i &= \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i' + \Sigma_i \\
 &= \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{bmatrix} \begin{bmatrix} \sigma_{b_0}^2 & 0 \\ 0 & \sigma_{b_1}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ t_1 & t_2 & t_3 & t_4 \end{bmatrix} + \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix} \\
 &= \begin{bmatrix} \sigma^2 + \sigma_{b_0}^2 + t_1^2 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_1 t_2 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_1 t_3 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_1 t_4 \sigma_{b_1}^2 \\ \sigma_{b_0}^2 + t_2 t_1 \sigma_{b_1}^2 & \sigma^2 + \sigma_{b_0}^2 + t_2^2 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_2 t_3 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_2 t_4 \sigma_{b_1}^2 \\ \sigma_{b_0}^2 + t_3 t_1 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_3 t_2 \sigma_{b_1}^2 & \sigma^2 + \sigma_{b_0}^2 + t_3^2 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_3 t_4 \sigma_{b_1}^2 \\ \sigma_{b_0}^2 + t_4 t_1 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_4 t_2 \sigma_{b_1}^2 & \sigma_{b_0}^2 + t_4 t_3 \sigma_{b_1}^2 & \sigma^2 + \sigma_{b_0}^2 + t_4^2 \sigma_{b_1}^2 \end{bmatrix}
 \end{aligned}$$

Essa estrutura de variâncias e covariâncias pode modelar a heterogeneidade de variâncias e a correlação positiva entre as medidas repetidas, pois a covariância entre os tempos é dada pela soma de duas variâncias.

Dado o efeito aleatório $\mathbf{b}_i$, tem-se que:

$$\begin{aligned}
 \mathbf{Y}_i | \mathbf{b}_i &\sim N(\mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i, \Sigma_i) \\
 \mathbf{b}_i &\sim N(\mathbf{0}, \mathbf{D})
 \end{aligned} \tag{5}$$

que é chamado de **modelo linear misto hierárquico** ou **condicional**. Exceto quando os dados são analisados no contexto bayesiano, a inferência baseia-se na distribuição marginal do vetor aleatório $\mathbf{Y}_i$, e a estrutura hierárquica do modelo original não é considerada no processo de estimação. O modelo hierárquico implica em um único modelo marginal, no entanto, diferentes modelos hierárquicos iniciais podem levar ao mesmo modelo marginal, de tal forma que a partir de um modelo marginal não se identifica o modelo hierárquico (VERBEKE; MOLENBERGHS, 2000).

Seja um experimento em que as unidades experimentais foram avaliadas em três ocasiões, t_1, t_2, t_3 , sendo o comportamento dos dados em relação ao tempo descrito por uma reta, com efeito aleatório no intercepto. A matriz de variâncias e covariâncias

para o modelo marginal é dada por:

$$\begin{aligned} \mathbf{V}_i &= \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i' + \boldsymbol{\Sigma}_i \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \sigma_{b_0}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} = \begin{bmatrix} \sigma^2 + \sigma_{b_0}^2 & \sigma_{b_0}^2 & \sigma_{b_0}^2 \\ \sigma_{b_0}^2 & \sigma^2 + \sigma_{b_0}^2 & \sigma_{b_0}^2 \\ \sigma_{b_0}^2 & \sigma_{b_0}^2 & \sigma^2 + \sigma_{b_0}^2 \end{bmatrix} \end{aligned}$$

Esta estrutura de variância pode ser utilizada em casos nos quais a variância e a correlação são constantes ao longo do tempo, e a correlação é positiva. Essa mesma estrutura pode ser obtida para o modelo marginal sem a utilização de efeitos aleatórios, admitindo para $\boldsymbol{\Sigma}_i$ uma estrutura uniforme ou de simetria composta.

Sejam $f_i(\mathbf{y}_i|\mathbf{b}_i)$ e $f_i(\mathbf{b}_i)$ as funções de densidade de probabilidade correspondentes a (5). A função de densidade de probabilidade marginal de $\mathbf{Y}_i$ é dada por:

$$f_i(\mathbf{y}_i) = \int f_i(\mathbf{y}_i|\mathbf{b}_i) f_i(\mathbf{b}_i) d\mathbf{b}_i$$

A distribuição conjunta de $\mathbf{Y}_i$ e $\mathbf{b}_i$ pode ser escrita como:

$$f_i(\mathbf{y}_i, \mathbf{b}_i) = f_i(\mathbf{y}_i) f_i(\mathbf{b}_i|\mathbf{y}_i) = f_i(\mathbf{y}_i|\mathbf{b}_i) f_i(\mathbf{b}_i) \quad (6)$$

Na expressão (6) observa-se que, ao considerar a distribuição marginal, a menos da utilização da distribuição a posteriori $f_i(\mathbf{b}_i|\mathbf{y}_i)$, perde-se a informação referente ao efeito aleatório $\mathbf{b}_i$.

Diferentes estruturas para as matrizes $\mathbf{D}$ e $\boldsymbol{\Sigma}_i$ podem ser utilizadas. Quando $\boldsymbol{\Sigma}_i$ é uma matriz diagonal, o modelo (4) é denominado **modelo de independência condicional** (NOBRE, 2004) .

2.3.2 Exemplos de estruturas das matrizes de covariância

A estrutura de covariâncias afeta as estimativas e os erros padrões de efeitos fixos, diagnósticos e inferência; sua escolha depende da estrutura dos dados, da informação disponibilizada pelo pesquisador e da disponibilidade computacional (BARBOSA, 2009).

A seguir, estão apresentadas algumas das estruturas mais utilizadas na prática para modelar as variâncias e covariâncias, considerando para efeito de ilustração somente $t = 4$ ocasiões.

- i) Componentes de variância - VC: admite variâncias iguais nas diferentes ocasiões, independência entre as medidas repetidas e envolve um único parâmetro.

$$\begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}$$

- ii) Componentes de variância heterogênea - UN(1): admite variâncias diferentes e independência entre as medidas repetidas. Envolve $t = 4$ parâmetros.

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}$$

- iii) Simetria Composta Homogênea - CS: admite variâncias e covariâncias iguais entre quaisquer medidas repetidas. Envolve dois parâmetros.

$$\begin{pmatrix} \sigma^2 + \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma^2 + \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma^2 + \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma^2 + \sigma_a^2 \end{pmatrix}$$

- iv) Simetria Composta Heterogênea - CSH: admite variâncias diferentes em cada ocasião e correlações constantes entre quaisquer dois instantes. Envolve $t+1 = 5$ parâmetros.

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \rho\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho\sigma_4\sigma_1 & \rho\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- v) Auto Regressiva de 1ª ordem - AR(1): admite variâncias iguais e correlação que decresce exponencialmente à medida que aumenta o intervalo de tempo entre as medidas repetidas. Pode ser utilizada apenas quando os intervalos de tempo forem igualmente espaçados. Envolve somente 2 parâmetros.

$$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

- vi) Auto Regressiva de 1ª ordem heterogênea - ARH(1): é uma generalização da matriz auto regressiva, impondo variâncias diferentes e correlação que decresce exponencialmente à medida que aumenta o intervalo de tempo entre as medidas. Envolve $t + 1 = 5$ parâmetros.

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & \rho^3\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho^2\sigma_2\sigma_4 \\ \rho^2\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho^3\sigma_4\sigma_1 & \rho^2\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- vii) Não Estruturada: admite que todas as variâncias e covariâncias podem ser desiguais. Envolve $t(t + 1)/2 = 10$ parâmetros.

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

2.3.3 Estimação no modelo marginal

Assumindo a distribuição normal e independência entre diferentes indivíduos, a função de distribuição marginal, para um determinado indivíduo i é uma distribuição normal com média $\mathbf{X}_i\boldsymbol{\beta}$ e matriz de variâncias e covariâncias $\mathbf{V}_i = \mathbf{V}_i(\boldsymbol{\alpha}) = \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \boldsymbol{\Sigma}_i$, em que $\boldsymbol{\alpha}$ representa o vetor dos parâmetros desconhecidos das matrizes de variâncias e covariâncias $\mathbf{D}$ e $\boldsymbol{\Sigma}_i$, denominados componentes de variância. A função de verossimilhança pode ser escrita da seguinte forma (MOLENBERGHS; VERBEKE, 2005):

$$L_{ML}(\boldsymbol{\theta}) = \prod_{i=1}^N \left\{ (2\pi)^{-\frac{n_i}{2}} |\mathbf{V}_i(\boldsymbol{\alpha})|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})' \mathbf{V}_i^{-1}(\boldsymbol{\alpha}) (\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta}) \right] \right\} \quad (7)$$

A estimação de $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha}')'$ requer a maximização conjunta da função de expressão (7) com relação aos elementos de $\boldsymbol{\theta}$.

Assumindo que os componentes de variância em $\boldsymbol{\alpha}$ são conhecidos, o estimador de máxima verossimilhança de $\boldsymbol{\beta}$, que coincide com o estimador obtido pelo método dos mínimos quadrados generalizados, é dado por (LAIRD; WARE, 1982):

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha}) = \left(\sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i^{-1}(\boldsymbol{\alpha}) \mathbf{X}_i \right)^{-1} \sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i^{-1}(\boldsymbol{\alpha}) \mathbf{Y}_i \quad (8)$$

Geralmente, tem-se que $\boldsymbol{\alpha}$ não é conhecido, porém desde que se tenha uma boa estimativa $\hat{\boldsymbol{\alpha}}$ pode-se substituir $\mathbf{V}_i(\boldsymbol{\alpha})$ por $\mathbf{V}_i(\hat{\boldsymbol{\alpha}})$ na expressão (8) para estimar $\boldsymbol{\beta}$.

Sob o modelo condicional a $\boldsymbol{\alpha}$ tem-se que $\hat{\boldsymbol{\beta}}$ segue uma distribuição normal multivariada.

Os métodos frequentemente utilizados para estimar os componentes de variância são: o método dos momentos ou método da análise de variância (ANOVA), método da máxima verossimilhança - ML (HARTLEY; RAO, 1967) e o método da máxima verossimilhança restrita - REML (PATTERSON; THOMPSON, 1971; HARVILLE, 1977). Existem ainda os métodos I, II e III de Henderson e aqueles derivados da estimação de funções quadráticas (MARCELINO; IEMMA, 2000).

O método da ANOVA consiste em obter os estimadores de $\boldsymbol{\alpha}$, que são as soluções do sistema obtido igualando-se os quadrados médios das diversas fontes de variação da análise de variância às suas respectivas esperanças matemáticas. Para dados balanceados os estimadores encontrados por esse método são não tendenciosos, eficientes e funções de estatísticas suficientes minimais, mas podem levar a estimativas negativas (SEARLE; CASELLA; McCULLOCH, 1992).

O método da máxima verossimilhança consiste em maximizar a função de verossimilhança das observações (7) em relação aos efeitos fixos e aos componentes de variância. As estimativas para $\boldsymbol{\alpha}$ assim obtidas são sempre não negativas, porém viesadas, devido ao fato de o método não considerar a perda de graus de liberdade resultante da estimação dos efeitos fixos do modelo (SEARLE; CASELLA; McCULLOCH, 1992).

No método da máxima verossimilhança restrita (REML), a função de verossimilhança $L(\boldsymbol{\theta})$ apresentada em (7) é dividida em duas partes, sendo uma delas totalmente livre dos efeitos fixos, obtida por meio do produto $\mathbf{A}'\mathbf{Y}_i$, em que $\mathbf{A}$ é uma matrix $n_i \times (n_i - p)$ de posto completo, que pode ser formada por quaisquer conjunto de

$n_i - p$ contrastes de erros independentes como, por exemplo, $\mathbf{A} = \mathbf{I} - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$, que tem colunas ortogonais às colunas da matriz $\mathbf{X}_i$ de tal forma que $E(\mathbf{A}'\mathbf{Y}_i) = \mathbf{0}$ e $Var(\mathbf{A}'\mathbf{Y}_i) = \mathbf{A}'\mathbf{V}_i\mathbf{A}$ (HARVILLE, 1977). A função de verossimilhança para $\mathbf{A}'\mathbf{Y}_i$ é dada por (HARVILLE, 1974):

$$\begin{aligned}
L_{REML}(\boldsymbol{\alpha}) = & (2\pi)^{-\frac{n_i-p}{2}} \left| \sum_{i=1}^N \mathbf{X}_i' \mathbf{X}_i \right|^{\frac{1}{2}} \\
& \times \left| \sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right|^{-\frac{1}{2}} \prod_{i=1}^N |\mathbf{V}_i|^{-\frac{1}{2}} \\
& \times \exp \left\{ -\frac{1}{2} \sum_{i=1}^N (\mathbf{Y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})' \mathbf{V}_i^{-1} (\mathbf{Y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}}) \right\}
\end{aligned} \tag{9}$$

em que $\hat{\boldsymbol{\beta}}$ é dado por (8). Assim, o estimador REML para $\boldsymbol{\alpha}$ é obtido maximizando-se a função dada em (9).

Quando os dados são balanceados, as estimativas obtidas para $\boldsymbol{\alpha}$ pelo método REML são idênticas àquelas obtidas pelo método da ANOVA, o que é desejável devido às propriedades ótimas deste último. O estimador de REML para $\boldsymbol{\alpha}$, bem como aquele obtido por ML, possuem normalidade assintótica, são consistentes e eficientes (VERBEKE; MOLENBERGHS, 2000).

2.3.4 Seleção de modelos

Para seleção dos modelos deve-se considerar o método de estimação adequado. Quando o interesse é testar hipóteses sobre os parâmetros da parte fixa do modelo, o método da máxima verossimilhança deve ser empregado, já quando o interesse está na parte aleatória tanto o método da máxima verossimilhança como o da máxima verossimilhança restrita podem ser utilizados. Vale ressaltar que, o método da máxima verossimilhança restrita (REML) apenas deve ser empregado para seleção da parte aleatória do modelo (VERBEKE; MOLENBERGHS, 2000; PINHEIRO; BATES, 2000).

A seguir, serão apresentados alguns resultados sobre o teste da razão de verossimilhanças e critérios de informação, bem como testes de Wald e F que podem ser empregados na seleção dos modelos.

2.3.4.1 Teste da razão de verossimilhanças

O teste da razão de verossimilhanças é um teste clássico para comparação de modelos encaixados, ou seja, quando o espaço gerado pelas colunas da matriz de especificação empregada no modelo sob a hipótese H_0 é um subespaço daquele gerado pelas colunas da matriz de especificação empregada no modelo sob a hipótese H_1 . A estatística para o teste da razão de verossimilhanças é definida por:

$$\Lambda = -2 \ln \left[\frac{L_{ML}(\hat{\boldsymbol{\theta}}_{ML,0})}{L_{ML}(\hat{\boldsymbol{\theta}}_{ML})} \right] \quad (10)$$

em que $\hat{\boldsymbol{\theta}}_{ML,0}$ é a estimativa de máxima verossimilhança obtida sob H_0 , e $\hat{\boldsymbol{\theta}}_{ML}$ é a estimativa de máxima verossimilhança obtida sob H_1 . Tem-se que, se o modelo reduzido é adequado, a estatística Λ sob H_0 segue, assintoticamente, uma distribuição qui-quadrado com número de graus de liberdade igual à diferença entre as dimensões dos espaços paramétricos dos modelos completo e reduzido (DIGGLE et al., 2002).

Considerando a interpretação hierárquica do modelo linear misto, pode existir o interesse científico em testar a necessidade da presença dos efeitos aleatórios. A menos da utilização da abordagem bayesiana, esse fato pode apenas ser testado indiretamente por meio do modelo marginal, testando a hipótese de que todos ou alguns dos componentes de variância presentes na matriz $\mathbf{D}$ são iguais a zero, de acordo com o interesse. Ao testar a hipótese de que um componente de variância, ou seja, um componente da diagonal da matriz $\mathbf{D}$, é igual a zero, pelo fato do teste ser realizado na borda do espaço paramétrico, a inferência clássica baseada na função de verossimilhança não pode ser aplicada diretamente. Sob a suposição de independência condicional, a distribuição nula assintótica da estatística da razão de verossimilhanças para testar hipóteses desse tipo é uma mistura de distribuições qui-quadrado. Tal problema não ocorre para o teste com componentes fora da diagonal da matriz $\mathbf{D}$, pois a covariância pode assumir valores positivos ou negativos (VERBEKE; MOLENBERGHS, 2003).

Para testar uma hipótese nula de que a matriz $\mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, em que $\mathbf{D}_{11}$ representa uma matriz positiva definida de dimensão $(q-1) \times (q-1)$, versus a hipótese alternativa de que a matriz $\mathbf{D}$ é uma matriz positiva semidefinida de dimensão $q \times q$, deve-se utilizar a estatística (10), que passa a ter como distribuição uma mistura de duas qui-quadrado $\frac{1}{2}\chi_{q-1}^2 + \frac{1}{2}\chi_q^2$ (STRAM; LEE, 1994).

Por exemplo, para verificar a necessidade de incluir um efeito aleatório ao coeficiente angular, quando o intercepto de efeito aleatório já está no modelo, tem-se:

$$H_0 : \mathbf{D} = \begin{pmatrix} d_{11} & 0 \\ 0 & 0 \end{pmatrix} \quad H_1 : \mathbf{D} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}$$

Nesse caso, deve-se utilizar uma mistura de $\frac{1}{2}\chi_1^2 + \frac{1}{2}\chi_2^2$ (STRAM; LEE, 1994).

2.3.4.2 Critérios de informação

Os principais critérios de informação usados na seleção de modelos são o critério de Akaike (*Akaike's Information Criterion* - AIC) e o critério bayesiano de Schwarz (*Bayesian Information Criterion* - BIC), que são baseados no valor da função de verossimilhança do modelo e dependem do número de observações e do número de parâmetros do modelo.

O Critério de informação de Akaike - AIC (AKAIKE, 1974), pode ser calculado por:

$$AIC = -2L(\hat{\boldsymbol{\theta}}) + 2d$$

em que $L(\hat{\boldsymbol{\theta}})$ é o valor da função de verossimilhança calculado em $\hat{\boldsymbol{\theta}}$, d representa o número total de parâmetros de efeito fixo e aleatório estimado no modelo.

O Critério de informação de Schwarz - BIC (SCHUWARZ, 1978) ou Schwarz *Information Criterion* - é calculado por:

$$BIC = -2L(\hat{\boldsymbol{\theta}}) + d \ln(N)$$

em que N é o número total de vetores $\mathbf{y}_i$.

Na seleção do modelo mais adequado, é necessário o cálculo do valor de AIC e BIC para cada modelo considerado, obtendo-se uma classificação dos modelos candidatos. O modelo com o menor valor de AIC ou de BIC é considerado o melhor modelo.

2.3.4.3 Teste de Wald aproximado

Seja $\mathbf{C}$ uma matriz qualquer conhecida tal que $\text{posto}(\mathbf{C}) = c$, e consideremos as hipóteses:

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0} \text{ versus } H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0}$$

Tem-se que a estatística de Wald é dada por:

$$(\hat{\beta} - \beta)' C' \left[C \left(\sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i(\hat{\alpha})^{-1} \mathbf{X}_i \right)^{-1} C' \right]^{-1} C(\hat{\beta} - \beta) \quad (11)$$

sob H_0 tem distribuição assintótica qui-quadrado com c graus de liberdade (VERBEKE; MOLENBERGHS, 2000).

2.3.4.4 Teste F aproximado

Para testar $H_0 : \mathbf{C}\beta = \mathbf{0}$ versus $H_1 : \mathbf{C}\beta \neq \mathbf{0}$, define-se a estatística:

$$F = \frac{(\hat{\beta} - \beta)' C' \left[C \left(\sum_{i=1}^N \mathbf{X}_i' \mathbf{V}_i(\hat{\alpha})^{-1} \mathbf{X}_i \right)^{-1} C' \right]^{-1} C(\hat{\beta} - \beta)}{\eta} \quad (12)$$

que sob H_0 segue assintoticamente uma distribuição F com q graus de liberdade do numerador (VERBEKE; MOLENBERGHS, 2000). O número η de graus de liberdade do denominador precisa ser estimado a partir dos dados utilizando, por exemplo, os métodos de aproximação como os propostos por Satterthwaite (1946), Kenward e Roger (1997), dentre outros.

2.3.5 Análise de Resíduos

Os resíduos são utilizados para avaliar a validade das suposições do modelo estatístico. Na regressão linear clássica, os resíduos podem ser utilizados para verificar homocedasticidade, existência de observações discrepantes, independência dos erros e normalidade (NOBRE, 2004). Como no modelo linear misto existe mais de uma fonte de erro, é possível definir diferentes “resíduos” que podem estar confundidos com outras fontes de variação no modelo (HILDEN-MILTON, 1995).

De acordo com Nobre e Singer (2007), podem-se definir três tipos de resíduos para o modelo (4):

(i) Resíduos marginais:

$$\hat{\mathbf{r}} = \mathbf{y} - \mathbf{X}\hat{\beta}$$

que são os valores preditos do erro marginal, $\mathbf{r} = \mathbf{y} - \mathbf{E}[\mathbf{Y}] = \mathbf{y} - \mathbf{X}\beta = \mathbf{Zb} + \varepsilon$;

(ii) Resíduos condicionais:

$$\hat{\varepsilon} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{Z}\hat{\mathbf{b}}$$

que são os valores preditos do erro condicional, $\boldsymbol{\varepsilon} = \mathbf{y} - E[\mathbf{Y}|\mathbf{b}] = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b}$;

(iii) O BLUP (Melhor preditor linear não-viesado):

$$\mathbf{Z}\hat{\mathbf{b}}$$

que consiste dos valores preditos para os efeitos aleatórios $\mathbf{Z}\mathbf{b} = E[\mathbf{Y}|\mathbf{b}] - E[\mathbf{Y}]$.

Como a obtenção do BLUP depende dos componentes de variância e covariância desconhecidos, é razoável obter o BLUP utilizando um estimador dos componentes de variância e covariância, obtendo o BLUP empírico - EBLUP (NOBRE, 2004).

De acordo com Hilden-Milton (1995), idealmente um resíduo deve ser uma função que depende apenas dos componentes de efeitos fixos do modelo e do respectivo erro do qual ele é preditor. Quando isto ocorre ele é chamado de **resíduo puro**, mas quando o resíduo depende de dois ou mais erros ele é chamado de **resíduo confundido**.

Os resíduos condicionais ($\hat{\varepsilon}$) e os efeitos aleatórios $\mathbf{Z}\hat{\mathbf{b}}$ são resíduos confundidos pela presença de $\mathbf{b}$ e $\boldsymbol{\varepsilon}$; no entanto, se $\mathbf{Z} \in C(\mathbf{X})$, em que $C(\mathbf{X})$ representa o subespaço gerado pelas colunas da matriz $\mathbf{X}$, tais resíduos serão considerados puros. Mais informações podem ser obtidas em Nobre (2004).

Como os elementos de $\hat{\varepsilon}$ podem ter variâncias distintas, West, Welch e Galecki (2007) sugerem a utilização dos resíduos padronizados, obtidos por meio da divisão dos resíduos pelos seus respectivos desvios padrões verdadeiros. No entanto, como esses desvios são raramente conhecidos na prática, os autores sugerem que sejam utilizadas as estimativas dos desvios padrões de forma a obter o **resíduo padronizado**. Esses resíduos podem ser utilizados para verificação da normalidade e homocedasticidade dos dados, desde que tais resíduos não estejam confundidos (NOBRE, 2004).

Os efeitos aleatórios refletem a diferença entre o valor predito e o valor médio populacional da i -ésima unidade experimental, dessa forma os BLUPs podem ser utilizados para verificar a presença de indivíduos discrepantes, ou ainda pode-se utilizar a distância de Mahalanobis para os efeitos aleatórios $\mathbf{b}_i$ ($\hat{\zeta}_i = \hat{\mathbf{b}}_i' \hat{Var}[\hat{\mathbf{b}}_i]^{-1} \hat{\mathbf{b}}_i$) conforme sugere Waternaux, Laird e Ware (1989). A Tabela 3 indica a utilização de cada tipo de resíduo, segundo Nobre e Singer (2007).

Tabela 3 - Utilização dos resíduos para fins de diagnóstico

Diagnóstico	Tipo de resíduo	Gráfico
Linearidade dos efeitos ($E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$)	Marginal	$\hat{\mathbf{r}}$ versus variáveis explanatórias
Presença de observações discrepantes	Condicional	$\hat{\boldsymbol{\varepsilon}}$ versus índices
Homocedasticidade dos erros condicionais	Condicional	$\hat{\boldsymbol{\varepsilon}}$ versus valores ajustados
Normalidade dos erros condicionais	Condicional	quantil-quantil para os resíduos condicionais com confundimento mínimo*
Presença de perfis individuais discrepantes	EBLUP	$\hat{\boldsymbol{\zeta}}$ versus índices dos indivíduos
Normalidade dos efeitos aleatórios ($\mathbf{b}_i$)	EBLUP	quantil-quantil ponderado para $\hat{\mathbf{b}}_i$

*Para maiores detalhes consultar Nobre e Singer (2007).

2.4 O modelo linear misto para dados longitudinais bivariados

Segundo Weiss (2005), dados longitudinais bivariados são comuns e a abordagem que trata uma das variáveis resposta como uma covariável que varia com o tempo, na maioria dos casos, não é a mais adequada.

Existem diferentes abordagens para modelagem de dados longitudinais com múltiplas respostas, no entanto, quando o objetivo é estudar como a evolução de uma resposta está relacionada com a evolução de outra resposta, a modelagem utilizando efeitos aleatórios é uma alternativa muito flexível. Essa abordagem tem muitas vantagens, é aplicável a uma ampla variedade de situações e permite vincular modelos a variáveis respostas com mesma distribuição, ou com distribuições diferentes (FIEUWS, 2006).

Seja Y_{rij} a mensuração obtida no i -ésimo indivíduo para a r -ésima variável resposta na j -ésima ocasião (tempo), com intervalos entre ocasiões possivelmente diferentes, em que $i = 1, \dots, N$, $r = 1, 2$ e $j = 1, \dots, n_{ri}$. Assume-se que não é necessário que o

mesmo número de medidas seja feito em todos os indivíduos, nem para todas as variáveis respostas. Seja $\mathbf{Y}_{ri}$ o vetor das n_{ri} medidas tomadas no indivíduo i , para a variável r . Assume-se que cada $\mathbf{Y}_{ri}$ satisfaz um modelo misto. No caso de duas variáveis respostas tem-se:

$$\begin{aligned}\mathbf{Y}_{1i} &= \mathbf{X}_{1i}\boldsymbol{\beta}_1 + \mathbf{Z}_{1i}\mathbf{b}_{1i} + \boldsymbol{\varepsilon}_{1i} \\ \mathbf{Y}_{2i} &= \mathbf{X}_{2i}\boldsymbol{\beta}_2 + \mathbf{Z}_{2i}\mathbf{b}_{2i} + \boldsymbol{\varepsilon}_{2i}\end{aligned}$$

em que $\boldsymbol{\beta}_r$, $r = 1, 2$, representa o vetor de parâmetros de efeito fixo para cada uma das variáveis; $\mathbf{X}_{ri}$ e $\mathbf{Z}_{ri}$ são, respectivamente, as matrizes de especificação dos efeitos fixos e dos efeitos aleatórios para cada uma das variáveis. Ambas as trajetórias de resposta são ligadas por meio da especificação da distribuição conjunta dos efeitos aleatórios $\mathbf{b}_{ri}$. Assume-se que o vetor $\mathbf{b}_i$ de todos os efeitos aleatórios para o i -ésimo indivíduo é proveniente de uma distribuição normal multivariada, com vetor de médias nulo e matriz de covariâncias $\mathbf{D}$, ou seja:

$$\mathbf{b}_i = \begin{pmatrix} \mathbf{b}_{1i} \\ \mathbf{b}_{2i} \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix} \right]$$

em que $\mathbf{D}_{12}$ representa a covariância entre $\mathbf{b}_{1i}$ e $\mathbf{b}_{2i}$. A matriz de covariância $\mathbf{D}_{12} = \mathbf{D}'_{21}$ permite que os níveis médios dos dois efeitos sejam correlacionados; caso eles não sejam correlacionados $\mathbf{D}_{12}$ será uma matriz nula e, neste caso, espera-se que as estimativas produzidas pelo método REML para os componentes de $\mathbf{D}_{12}$ estejam próximas de zero quando comparadas com o erro padrão (WEISS, 2005).

Assume-se ainda que os vetores de erros ($\boldsymbol{\varepsilon}_{1i}$ e $\boldsymbol{\varepsilon}_{2i}$) são não correlacionados, independentes dos efeitos aleatórios e que:

$$\boldsymbol{\varepsilon}_i = \begin{pmatrix} \boldsymbol{\varepsilon}_{1i} \\ \boldsymbol{\varepsilon}_{2i} \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{1i} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{2i} \end{pmatrix} \right].$$

O fato de admitir que os vetores de erros $\boldsymbol{\varepsilon}_{1i}$ e $\boldsymbol{\varepsilon}_{2i}$ são não correlacionados implica que, condicional aos efeitos aleatórios, as trajetórias das respostas são independentes (FIEUWS; VERBEKE, 2004).

Admitindo, por exemplo, que em um experimento as unidades experimentais são avaliadas em relação a duas variáveis, $r = 1, 2$, nas mesmas quatro ocasiões, t_1, t_2, t_3, t_4 , sendo o comportamento dos dados em relação ao tempo, para cada uma das variáveis,

descrito por uma reta, com efeito aleatório no intercepto da reta de cada uma das variáveis.

Assim tem-se:

$$\begin{aligned}
 \mathbf{Y}_i &= \begin{bmatrix} y_{11i} \\ y_{12i} \\ y_{13i} \\ y_{14i} \\ y_{21i} \\ y_{22i} \\ y_{23i} \\ y_{24i} \end{bmatrix} \quad \mathbf{X}_i = \begin{bmatrix} 1 & t_1 & 0 & 0 \\ 1 & t_2 & 0 & 0 \\ 1 & t_3 & 0 & 0 \\ 1 & t_4 & 0 & 0 \\ 0 & 0 & 1 & t_1 \\ 0 & 0 & 1 & t_2 \\ 0 & 0 & 1 & t_3 \\ 0 & 0 & 1 & t_4 \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{20} \\ \beta_{21} \end{bmatrix} \\
 \mathbf{Z}_i &= \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} \sigma_{b_{10}}^2 & \sigma_{b_{10}b_{20}} \\ \sigma_{b_{20}b_{10}} & \sigma_{b_{20}}^2 \end{bmatrix} \quad \boldsymbol{\Sigma}_i = \begin{bmatrix} \boldsymbol{\Sigma}_{1i} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{2i} \end{bmatrix} = \mathbf{I}_4 \otimes \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}
 \end{aligned}$$

Para facilitar a notação, no caso de apenas um efeito aleatório para cada variável, será utilizado $b_{10} = b_1$ e $b_{20} = b_2$. Dessa forma, a matriz de variâncias e co-variâncias para o modelo marginal é dada por:

$$\begin{aligned}
 \mathbf{V}_i &= \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i' + \boldsymbol{\Sigma}_i \\
 &= \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_1 b_2} \\ \sigma_{b_2 b_1} & \sigma_{b_2}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} + \mathbf{I}_4 \otimes \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}
 \end{aligned}$$

E então:

$$\mathbf{V}_i = \left[\begin{array}{cccc|cccc} \sigma_{b_1}^2 + \sigma_1^2 & \sigma_{b_1}^2 & \sigma_{b_1}^2 & \sigma_{b_1}^2 & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} \\ \sigma_{b_1}^2 & \sigma_{b_1}^2 + \sigma_1^2 & \sigma_{b_1}^2 & \sigma_{b_1}^2 & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} \\ \sigma_{b_1}^2 & \sigma_{b_1}^2 & \sigma_{b_1}^2 + \sigma_1^2 & \sigma_{b_1}^2 & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} \\ \sigma_{b_1}^2 & \sigma_{b_1}^2 & \sigma_{b_1}^2 & \sigma_{b_1}^2 + \sigma_1^2 & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} \\ \hline \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_2}^2 + \sigma_2^2 & \sigma_{b_2}^2 & \sigma_{b_2}^2 & \sigma_{b_2}^2 \\ \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_2}^2 & \sigma_{b_2}^2 + \sigma_2^2 & \sigma_{b_2}^2 & \sigma_{b_2}^2 \\ \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_2}^2 & \sigma_{b_2}^2 & \sigma_{b_2}^2 + \sigma_2^2 & \sigma_{b_2}^2 \\ \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_1 b_2} & \sigma_{b_2}^2 & \sigma_{b_2}^2 & \sigma_{b_2}^2 & \sigma_{b_2}^2 + \sigma_2^2 \end{array} \right]$$

Essa estrutura de variâncias e covariâncias pode modelar tanto a correlação (positiva) entre as medidas feitas em ocasiões diferentes como a correlação entre as duas variáveis respostas. No entanto, é uma estrutura bastante restritiva, pois assume que as variâncias das medidas é a mesma em todas as ocasiões e para as duas variáveis respostas, que a covariância entre observações da mesma variável realizadas em diferentes ocasiões é constante e positiva, e que a covariância entre as duas variáveis é a mesma em todas as ocasiões.

Segundo Fieuws e Verbeke (2004), a suposição de independência condicional pode ser relaxada, admitindo-se que os erros não são independentes. Neste caso, assim como no modelo para dados longitudinais univariados, a escolha entre modelar a correlação entre as variáveis por meio de efeitos aleatórios ou diretamente na matriz Σ_i depende do interesse do pesquisador, lembrando que, da mesma forma que acontece nos modelos univariados, diferentes modelos hierárquicos podem levar ao mesmo modelo marginal.

Muitos trabalhos têm sido desenvolvidos em que a estrutura da matriz Σ_i é especificada por um produto Kronecker entre uma matriz $\mathbf{R}_i$ de dimensões $t \times t$, utilizada para modelar a estrutura de variâncias e covariâncias entre os tempos, e uma matriz $\mathbf{G}_i$, que serve para modelar as variâncias e covariâncias entre as diferentes variáveis e que no caso bivariado terá dimensões 2×2 (ROY; KHATTREE, 2006; THIÈBAUT et al., 2002). Dessa forma tem-se:

$$\Sigma_i = \mathbf{G}_i \otimes \mathbf{R}_i \quad (13)$$

A estrutura da matriz $\mathbf{R}_i$ é admitida ser a mesma para as duas variáveis.

Utilizar a forma (13) para representar a matriz $\mathbf{V}_i$ sem a necessidade de inclusão de efeitos aleatórios tem algumas vantagens, pois facilita a implementação de algumas estruturas especiais. Por exemplo, ao assumir que $\mathbf{G}_i$ é do tipo não estruturada e $\mathbf{R}_i$ é do tipo simetria composta, obtém-se a seguinte estrutura para matriz $\mathbf{\Sigma}_i$, que para 4 ocasiões é dada por:

$$\mathbf{\Sigma}_i = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \otimes \begin{bmatrix} \sigma^2 & \sigma_a & \sigma_a & \sigma_a \\ \sigma_a & \sigma^2 & \sigma_a & \sigma_a \\ \sigma_a & \sigma_a & \sigma^2 & \sigma_a \\ \sigma_a & \sigma_a & \sigma_a & \sigma^2 \end{bmatrix}$$

$$\mathbf{\Sigma}_i = \left[\begin{array}{cccc|cccc} \sigma_1^2 \sigma^2 & \sigma_1^2 \sigma_a & \sigma_1^2 \sigma_a & \sigma_1^2 \sigma_a & \sigma_{12} \sigma^2 & \sigma_{12} \sigma_a & \sigma_{12} \sigma_a & \sigma_{12} \sigma_a \\ \sigma_1^2 \sigma_a & \sigma_1^2 \sigma^2 & \sigma_1^2 \sigma_a & \sigma_1^2 \sigma_a & \sigma_{12} \sigma_a & \sigma_{12} \sigma^2 & \sigma_{12} \sigma_a & \sigma_{12} \sigma_a \\ \sigma_1^2 \sigma_a & \sigma_1^2 \sigma_a & \sigma_1^2 \sigma^2 & \sigma_1^2 \sigma_a & \sigma_{12} \sigma_a & \sigma_{12} \sigma_a & \sigma_{12} \sigma^2 & \sigma_{12} \sigma_a \\ \sigma_1^2 \sigma_a & \sigma_1^2 \sigma_a & \sigma_1^2 \sigma_a & \sigma_1^2 \sigma^2 & \sigma_{12} \sigma_a & \sigma_{12} \sigma_a & \sigma_{12} \sigma_a & \sigma_{12} \sigma^2 \\ \hline \sigma_{21} \sigma^2 & \sigma_{21} \sigma_a & \sigma_{21} \sigma_a & \sigma_{21} \sigma_a & \sigma_2^2 \sigma^2 & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma_a \\ \sigma_{21} \sigma_a & \sigma_{21} \sigma^2 & \sigma_{21} \sigma_a & \sigma_{21} \sigma_a & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma^2 & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma_a \\ \sigma_{21} \sigma_a & \sigma_{21} \sigma_a & \sigma_{21} \sigma^2 & \sigma_{21} \sigma_a & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma^2 & \sigma_2^2 \sigma_a \\ \sigma_{21} \sigma_a & \sigma_{21} \sigma_a & \sigma_{21} \sigma_a & \sigma_{21} \sigma^2 & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma_a & \sigma_2^2 \sigma^2 \end{array} \right]$$

Observe que a matriz $\mathbf{\Sigma}_i$, neste caso $\mathbf{\Sigma}_i = \mathbf{V}_i$, também assume que a variância das medidas feitas em todas as ocasiões é constante para cada um das variáveis ($\sigma_1^2 \sigma^2$ e $\sigma_2^2 \sigma^2$, respectivamente), no entanto ela difere da estrutura obtida com a utilização de efeito aleatório de intercepto para as duas variáveis apresentada anteriormente. A covariância entre as medidas das duas variáveis em ocasiões distintas ($\sigma_{12} \sigma_a$) também é constante, mas é menor que a covariância entre as medidas feitas nas mesmas ocasiões ($\sigma_{12} \sigma^2$).

Esta característica é comum em situações práticas: uma covariância maior entre as medidas das duas variáveis feitas na mesma ocasião, do que feitas em ocasiões distintas.

2.4.1 Diagnóstico para os modelos bivariados

Duas variáveis aleatórias com distribuição normal podem ter distribuição conjunta normal bivariada, porém o fato de duas variáveis aleatórias terem distribuição

normal univariada não implica em uma distribuição conjunta normal bivariada. Para verificação da pressuposição de normalidade bivariada pode-se utilizar um gráfico de dispersão das duas variáveis com as curvas de níveis. Quando a normalidade bivariada for satisfeita espera-se que os pontos concentrem-se no interior de uma elipse. Uma série de linhas de contorno é utilizada para demonstrar uma série de elipses com quantidades variáveis de distância a partir do centro comum (centróide) (BURDENSKI, 2000).

De acordo com Everitt (2005), o boxplot bivariado pode ser útil para indicar as propriedades de distribuição dos dados e para identificar possíveis *outliers*. O boxplot bivariado baseia-se no cálculo robusto de medidas de locação, escala e correlação. É constituído essencialmente por um par de elipses concêntricas, uma das quais inclui 50% dos dados, e outra que delinea potenciais *outliers*. Além disso, o boxplot bivariado apresenta as linhas de regressão de cada variável em relação à outra, cuja interseção é a estimativa do parâmetro de locação. O ângulo agudo entre as linhas de regressão será pequeno se o valor absoluto da correlação entre as variáveis for grande, e será grande para um valor absoluto de correlações pequeno.

De acordo com Timm (2002), se $\mathbf{Y}_i \sim N_p(\boldsymbol{\mu}, \mathbf{V})$, então a forma quadrática

$$\Delta_i^2 = (\mathbf{Y}_i - \boldsymbol{\mu})' \mathbf{V}^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}) \sim \chi_p^2$$

pode ser utilizada como uma medida de distância entre o vetor de valores observados e o vetor de valores esperados, e leva em consideração as diferentes variâncias das variáveis e as covariâncias entre duas variáveis (EVERITT, 2005). A distância de Mahalanobis (d_i^2), utilizada para estimar Δ_i^2 a partir de uma amostra, é dada por:

$$d_i^2 = (\mathbf{y}_i - \hat{\mathbf{y}})' \hat{\mathbf{V}}^{-1} (\mathbf{y}_i - \hat{\mathbf{y}})$$

Se as observações são provenientes de uma distribuição normal bivariada, essas distâncias têm, aproximadamente, uma distribuição qui-quadrado com 2 graus de liberdade. Dessa forma, o gráfico das distâncias ordenadas versus os quantis da distribuição qui-quadrado (q_i) pode ser utilizado para verificar a normalidade bivariada. Se os dados são provenientes de uma distribuição normal bivariada, o gráfico com os pontos (d_i^2, q_i) devem estar próximos de uma reta e os pontos distantes dela são considerados potenciais *outliers* (TIMM, 2002).

2.5 O modelo linear misto para dados longitudinais multivariados

A mesma abordagem apresentada para dados longitudinais bivariados pode ser empregada no caso de dados longitudinais multivariados, utilizando um modelo misto para cada uma das respostas e unindo esses modelos por meio da especificação da distribuição conjunta para os efeitos aleatórios (FIEUWS; VERBEKE; MOLENBERGHS, 2007). As estimativas dos parâmetros desta distribuição comum podem ser usadas para avaliar a relação entre as diferentes variáveis respostas (FIEUWS, 2006).

Devido à sua flexibilidade, a abordagem utilizando efeitos aleatórios para a modelagem conjunta de perfis longitudinais multivariados tem recebido especial atenção em diversas publicações, como Shah, Laird e Schoenfeld (1997); Chakraborty et al. (2003); Fieuws e Verbeke (2004), dentre outros.

Seja m o número de variáveis respostas a serem modeladas conjuntamente, com $r = 1, 2, \dots, m$. Assume-se:

$$\mathbf{Y}_{ri} | \mathbf{b}_{ri} \sim N_{n_{ri}}(\mathbf{X}_{ri}\boldsymbol{\beta}_r + \mathbf{Z}_{ri}\mathbf{b}_{ri}, \boldsymbol{\Sigma}_{ri}) \quad (14)$$

em que $N_{n_{ri}}$ representa a distribuição normal n_{ri} -dimensional, $\mathbf{X}_{ri}$ e $\mathbf{Z}_{ri}$ são as matrizes do delineamento conhecidas para os efeitos fixos e aleatórios, respectivamente, e $\boldsymbol{\beta}_r$ é o vetor de parâmetros que inclui todos os efeitos fixos necessários para modelar a média da variável r . Não é necessário assumir o mesmo modelo para todas as variáveis respostas, sendo possível assumir um modelo misto linear, linear generalizado, ou até mesmo um modelo misto não-linear. Também não é necessária a utilização do mesmo número (q_r) de efeitos aleatórios para todas as m variáveis respostas.

Os modelos lineares mistos específicos para cada variável resposta definidos em (14) são combinados por meio da especificação de uma distribuição conjunta (normal multivariada) para $\mathbf{b}_i = (\mathbf{b}_{i1}, \mathbf{b}_{i2}, \dots, \mathbf{b}_{im})'$ de todos os efeitos aleatórios para o i -ésimo indivíduo, com média zero e matriz de covariâncias $\mathbf{D}$, ou seja:

$$\mathbf{b}_i = \begin{pmatrix} \mathbf{b}_{1i} \\ \mathbf{b}_{2i} \\ \vdots \\ \mathbf{b}_{mi} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} & \dots & \mathbf{D}_{1m} \\ \mathbf{D}_{21} & \mathbf{D}_{22} & \dots & \mathbf{D}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{D}_{m1} & \mathbf{D}_{m2} & \dots & \mathbf{D}_{mm} \end{pmatrix} \right]$$

em que $\mathbf{D}_{rs}$ representa a covariância entre $\mathbf{b}_{ri}$ e $\mathbf{b}_{si}$, com $r, s = 1, 2, \dots, m$.

Um caso especial do modelo (14) é o chamado modelo de parâmetros compartilhados (*shared-parameter model*), que assume o mesmo conjunto de efeitos aleatórios para todas as variáveis respostas (VERBEKE; MOLENBERGHS; RIZOPOULOS, 2010).

Os planejamentos longitudinais uni ou multivariados não incluem os estudos denominados **análise de sobrevivência**, em que são avaliados os tempos antes da ocorrência de um determinado evento de interesse, porém muitos estudos vem sendo desenvolvidos para modelagem conjunta de dados longitudinais e dados de tempos para ocorrência do evento (HEMDERSON; DIGGLE; DOBSON, 2000; ALBERT; SHIH, 2010).

É importante perceber que o modelo (14) ainda é um modelo linear misto, de tal forma que podem ser empregadas as mesmas técnicas de estimação, inferência e seleção adotadas para o modelo linear misto univariado.

3 MATERIAL E MÉTODOS

3.1 Material

O estudo dos procedimentos de investigação utilizou os dados da tese de Fernandes (2007) que avaliou os efeitos da contagem de células somáticas (CCS) nas frações de caseína do leite UAT (Ultra Alta Temperatura), ao longo do tempo de armazenamento.

Dentre as variáveis respostas estudadas por Fernandes (2007), somente foram utilizadas no presente trabalho as concentrações em (mg/mL) de α_{S1} -caseína e beta-caseína avaliadas utilizando o método por cromatografia líquida de alta eficiência (CLAE).

As amostras foram coletadas em usina de beneficiamento localizada no município de Casa Branca-SP. A amostragem foi realizada no período de março de 2005 a março de 2006, seguindo-se a mesma rotina de procedimentos adotados pela usina no recebimento de leite cru.

Os leites dos lotes experimentais foram adicionados de citrato de sódio (0,05%, p/v), transferidos para silo isotérmico específico e monitorados durante todo o processo.

Quinze lotes de leite UAT foram categorizados em grupos de acordo com a CCS, que variou de 197.000 a 800.000 CS/mL ou 5,29 a 5,90 log CS/mL. O grupo com baixa CCS (Grupo 1) constituiu-se de lotes com contagens entre 197.000 e 316.000 CS/mL; o grupo com CCS intermediária (Grupo 2), por lotes com contagens entre 379.000 e 560.000 CS/mL e o grupo com alta CCS (Grupo 3) constituiu-se de lotes com contagem entre 600.000 e 800.000 CS/mL.

O produto foi envasado automaticamente em embalagens assépticas cartonadas de 1 litro, em sistema de torre construído pela mesma empresa. Todas as unidades produzidas em cada lote ficaram retidas na usina até o término do período de observação (sete dias) de amostras do lote mantidas em estufa a 35°C. Após o processamento dos lotes, e de acordo com os procedimentos de rotina da empresa, algumas unidades do produto foram submetidas ao teste de condutividade elétrica, para a verificação da existência de microfuros na embalagem, os quais poderiam interferir negativamente na qualidade do leite processado.

Cinco amostras por lote do produto (ou seja, cinco embalagens originais de

leite UAT contendo, cada uma, 1 litro do produto) foram avaliadas 8, 30, 60, 90 e 120 dias após a fabricação, quanto à quantidade de α_{s1} -caseína e β -caseína (mg/mL) (FERNANDES, 2007).

Os valores médios e os respectivos erros padrões para as variáveis consideradas para os diferentes tempos e grupos estão apresentados na Tabela 4.

Tabela 4 - Médias e erros padrões (e.p.) da concentração de α_{s1} -caseína e β -caseína em mg/mL do leite AUT nos cinco tempos de estocagem, por grupo

α_{s1} – caseína						
Dia	Grupo 1		Grupo 2		Grupo 3	
	Média	e.p.	Média	e.p.	Média	e.p.
8	11,30	2,18	10,23	2,60	8,70	4,28
30	8,94	4,52	9,76	3,33	9,32	3,36
60	9,21	2,07	9,07	2,83	8,37	2,22
90	7,10	6,02	7,50	4,45	7,16	3,82
120	6,99	3,62	6,95	4,14	6,65	3,94
β – caseína						
Dia	Grupo 1		Grupo 2		Grupo 3	
	Média	e.p.	Média	e.p.	Média	e.p.
8	11,54	2,47	11,32	2,47	9,58	3,57
30	11,24	3,37	10,53	3,37	8,27	4,60
60	10,56	0,55	9,88	0,55	8,38	3,88
90	9,43	1,74	8,47	1,74	6,97	4,76
120	7,45	3,00	7,81	3,00	6,39	4,71

A Figura 1 apresenta os perfis individuais para as duas variáveis respostas. As variáveis apresentam comportamento decrescente no decorrer do tempo, sendo a semelhança no comportamento das duas variáveis mais acentuada para o grupo 2. Pode-se observar também uma heterogeneidade de variâncias nos diferentes tempos, sendo mais evidente para a variável β -caseína nos grupos 1 e 2, e para α_{s1} -caseína no grupo 1.

3.2 Métodos

Para analisar os dados apresentados na seção 3.1.1, empregou-se o modelo de efeitos aleatórios bivariado, para o qual considerou-se y_{rij} a j -ésima mensuração obtida na amostra de leite proveniente do i -ésimo lote para as variáveis respostas concentração de α_{s1} -caseína e β -caseína e $\mathbf{y}_{ri}$, o vetor das medidas da variável r observadas no indivíduo i . Assume-se que cada y_{rij} satisfaz um modelo misto:

Figura 1 - Perfis individuais das concentrações de α_{s1} -caseína e β -caseína no leite UAT, por grupo

$$\begin{aligned}
 y_{rij} = & \overbrace{\beta_{0r}G1_i + \beta_{1r}G1_it_j + \beta_{2r}G2_i + \beta_{4r}G2_it_j + \beta_{3r}G3_i + \beta_5G3_it_j}^{\text{parte fixa}} \\
 & + \underbrace{b_{ri0} + b_{ri1}t_j + \epsilon_{rij}}_{\text{parte aleatória}}
 \end{aligned}$$

para $i = 1, \dots, 15$, $j = 1, \dots, 5$ e $t_j = 8, 30, 60, 90$ e 120 , com a variável indicadora, Gk_i , assumindo os valores:

$$Gk_i = \begin{cases} 1 & \text{se } y_{rij} \in \text{Grupo } k, k=1,2,3 \\ 0 & \text{caso contrário.} \end{cases}$$

Matricialmente tem-se:

$$\mathbf{Y}_{ri} | \mathbf{b}_{ri} \sim N_{ri}(\mathbf{X}_{ri}\boldsymbol{\beta}_r + \mathbf{Z}_{ri}\mathbf{b}_{ri}, \boldsymbol{\Sigma}_{ri})$$

em que N_{ri} representa a distribuição normal r_i -dimensional, $\mathbf{X}_{ri}$ e $\mathbf{Z}_{ri}$ são as matrizes do delineamento, conhecidas, para os efeitos fixos e aleatórios respectivamente, e $\boldsymbol{\beta}_r$ é o vetor de parâmetros fixos desconhecidos associado à variável r , $r = 1, 2$.

Assume-se que o vetor $\mathbf{b}_i$ de todos os efeitos aleatórios para o i -ésimo indivíduo é proveniente de uma distribuição normal multivariada de dimensão $q_1 + q_2$ (em que q_r representa o número de efeitos aleatórios associados à r -ésima variável), com média zero e matriz de covariâncias $\mathbf{D}$, ou seja:

$$\mathbf{b}_i = \begin{pmatrix} \mathbf{b}_{1i} \\ \mathbf{b}_{2i} \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix} \right]$$

em que $\mathbf{D}_{11}$ representa a matriz de variâncias e covariâncias associadas aos efeitos aleatórios da variável *alpha*_{s1}—caseína, $\mathbf{D}_{22}$ representa a matriz de variâncias e covariâncias associadas aos efeitos aleatórios da variável *beta*—caseína e $\mathbf{D}_{12}$ representa a matriz de covariâncias entre $\mathbf{b}_{1i}$ e $\mathbf{b}_{2i}$. Esta matriz permite que os níveis médios dos dois efeitos sejam correlacionados; caso eles não sejam correlacionados, $\mathbf{D}_{12}$ será uma matriz nula (WEISS, 2005).

3.2.1 Estratégia de análise

Para ajuste do modelo, realizou-se primeiramente a seleção dos efeitos aleatórios utilizando o modelo maximal para a parte fixa (considerando todos os efeitos de grupo e as possíveis interações e o tempo como fator qualitativo) e o método REML de estimação. Em seguida, realizou-se a seleção da parte fixa (método da máxima verossimilhança - ML) sendo posteriormente realizada uma nova seleção da parte aleatória, empregando o modelo selecionado para a parte fixa e o método REML. Na seleção dos modelos não encaixados empregou-se os Critérios de Informação de Akaike (AIC) e Bayesiano (BIC), e o teste da razão de verossimilhanças para comparação de modelos encaixados (FIEUWS; VERBEKE, 2004).

Considerando o modelo fixo maximal, foram testadas as seguintes estruturas para matriz $\mathbf{D}$, sugeridas pelas características observadas no gráfico de perfis individuais (Figura 1).

M1. Efeito aleatório no intercepto e coeficiente angular para ambas as variáveis, com correlação entre intercepto e coeficiente angular, e correlação entre os efeitos das duas variáveis. A estrutura fica:

$$\mathbf{D} = \left(\begin{array}{cc|cc} \sigma_{b_{10}}^2 & \sigma_{b_{10}b_{11}} & \sigma_{b_{10}b_{20}} & \sigma_{b_{10}b_{21}} \\ \sigma_{b_{10}b_{11}} & \sigma_{b_{11}}^2 & \sigma_{b_{11}b_{20}} & \sigma_{b_{11}b_{21}} \\ \hline \sigma_{b_{20}b_{10}} & \sigma_{b_{20}b_{11}} & \sigma_{b_{20}}^2 & \sigma_{b_{20}b_{21}} \\ \sigma_{b_{21}b_{10}} & \sigma_{b_{21}b_{11}} & \sigma_{b_{20}b_{21}} & \sigma_{b_{21}}^2 \end{array} \right)$$

em que o índice b_{rs} representa a variável (r) e o coeficiente (s), assim $\sigma_{b_{10}b_{11}}$ representa a covariância entre os efeitos aleatórios de intercepto ($s = 0$) e coeficiente

angular ($s = 1$) associados a $alfa_{s1}$ –caseína ($r = 1$), enquanto $\sigma_{b_{10}b_{20}}$ representa a covariância entre os efeitos aleatórios de intercepto nas diferentes variáveis.

- M2. Efeito aleatório no intercepto para $alpha_{s1}$ –caseína, efeito aleatório no intercepto e coeficiente angular, com correlação entre intercepto e no coeficiente angular para a variável $beta$ –caseína, e correlação entre os efeitos aleatórios nas duas variáveis. A estrutura fica:

$$D = \left(\begin{array}{c|cc} \sigma_{b_{10}}^2 & \sigma_{b_{10}b_{20}} & \sigma_{b_{10}b_{21}} \\ \hline \sigma_{b_{20}b_{10}} & \sigma_{b_{20}}^2 & \sigma_{b_{20}b_{21}} \\ \sigma_{b_{21}b_{10}} & \sigma_{b_{21}b_{20}} & \sigma_{b_{21}}^2 \end{array} \right)$$

- M3. Efeito aleatório no intercepto em ambas as variáveis e correlação entre os efeitos nas duas variáveis. A estrutura fica:

$$D = \left(\begin{array}{c|c} \sigma_{b_{10}}^2 & \sigma_{b_{10}b_{20}} \\ \hline \sigma_{b_{20}b_{10}} & \sigma_{b_{20}}^2 \end{array} \right)$$

- M4. Efeito aleatório no intercepto e no coeficiente angular em ambas as variáveis, com correlação entre intercepto e coeficiente angular para cada variável, e sem correlação entre os efeitos nas duas variáveis. A estrutura fica:

$$D = \left(\begin{array}{cc|cc} \sigma_{b_{10}}^2 & \sigma_{b_{10}b_{11}} & 0 & 0 \\ \sigma_{b_{10}b_{11}} & \sigma_{b_{11}}^2 & 0 & 0 \\ \hline 0 & 0 & \sigma_{b_{20}}^2 & \sigma_{b_{20}b_{21}} \\ 0 & 0 & \sigma_{b_{20}b_{21}} & \sigma_{b_{21}}^2 \end{array} \right)$$

- M5. Efeito aleatório no intercepto para $alpha_{s1}$ –caseína, efeito aleatório no intercepto e coeficiente angular, com correlação entre intercepto e coeficiente angular para a variável $beta$ –caseína, e correlação entre os efeitos de intercepto nas duas variáveis. A estrutura fica:

$$D = \left(\begin{array}{c|cc} \sigma_{b_{10}}^2 & \sigma_{b_{10}b_{20}} & 0 \\ \hline \sigma_{b_{20}b_{10}} & \sigma_{b_{20}}^2 & \sigma_{b_{20}b_{21}} \\ 0 & \sigma_{b_{21}b_{20}} & \sigma_{b_{21}}^2 \end{array} \right)$$

- M6. Efeito aleatório no intercepto e no coeficiente angular em ambas as variáveis, sem

correlação entre os efeitos. A estrutura fica:

$$\mathbf{D} = \left(\begin{array}{cc|cc} \sigma_{b_{10}}^2 & 0 & 0 & 0 \\ 0 & \sigma_{b_{11}}^2 & 0 & 0 \\ \hline 0 & 0 & \sigma_{b_{20}}^2 & 0 \\ 0 & 0 & 0 & \sigma_{b_{21}}^2 \end{array} \right)$$

M7. Efeito aleatório no intercepto para α_{s1} —caseína, efeito aleatório no intercepto e coeficiente angular para a variável β —caseína, sem correlação entre os efeitos. A estrutura fica:

$$\mathbf{D} = \left(\begin{array}{c|cc} \sigma_{b_{10}}^2 & 0 & 0 \\ \hline 0 & \sigma_{b_{20}}^2 & 0 \\ 0 & 0 & \sigma_{b_{21}}^2 \end{array} \right)$$

M8. Efeito aleatório no intercepto em ambas as variáveis, sem correlação entre os efeitos nas duas variáveis. A estrutura fica:

$$\mathbf{D} = \left(\begin{array}{c|c} \sigma_{b_{10}}^2 & 0 \\ \hline 0 & \sigma_{b_{20}}^2 \end{array} \right)$$

M9. Efeito aleatório no intercepto compartilhado pelas variáveis. A estrutura fica:

$$\mathbf{D} = \left(\sigma_{b_0}^2 \right)$$

M10. Efeito aleatório no intercepto compartilhado pelas variáveis, e efeito de coeficiente angular apenas para β —caseína. A estrutura fica:

$$\mathbf{D} = \left(\begin{array}{cc} \sigma_{b_0}^2 & 0 \\ 0 & \sigma_{b_{21}}^2 \end{array} \right)$$

M11. Efeito aleatório no intercepto compartilhado pelas variáveis e efeito de coeficiente angular semelhante em cada uma das variáveis, sem correlação. A estrutura fica:

$$\mathbf{D} = \left(\begin{array}{ccc} \sigma_{b_0}^2 & 0 & 0 \\ 0 & \sigma_{b_{11}}^2 & 0 \\ 0 & 0 & \sigma_{b_{21}}^2 \end{array} \right)$$

M12. Sem utilização de efeitos aleatórios, matriz Σ_i modelada pelo produto Kroneker entre uma matriz não estruturada 2×2 e uma matriz do tipo simetria composta 5×5 .

Avaliando-se as médias apresentadas na Tabela 4 e os perfis individuais da Figura 1, pode-se admitir que as respostas de ambas as variáveis decrescem na mesma velocidade em todos os grupos. Observa-se também que as concentrações iniciais de α_s1 -caseína e de β -caseína são semelhantes nos grupos 1 e 2. Dessa forma, com relação à parte fixa foram comparados os seguintes modelos:

1. Interceptos e coeficientes angulares distintos para todos os grupos para cada variável.
2. Interceptos distintos para os grupos em cada variável e coeficientes angulares distintos para os grupos e iguais para as duas variáveis.
3. Interceptos iguais para os grupos 1 e 2 para α_s1 -caseína e interceptos distintos para β -caseína, interceptos iguais para o grupo 3 em relação às duas variáveis; coeficientes angulares distintos para os grupos e iguais para as duas variáveis.
4. Interceptos iguais para os grupos 1 e 2 e distintos para as duas variáveis; mesmo intercepto para o grupo 3 nas duas variáveis; coeficientes angulares distintos para os grupos e iguais para as duas variáveis.
5. Interceptos iguais para os grupos 1 e 2 e distintos para as duas variáveis; mesmo intercepto para o grupo 3 nas duas variáveis; coeficientes angulares iguais para os grupos 1 e 2 e para as duas variáveis; mesmo coeficiente angular para o grupo 3 nas duas variáveis.
6. Interceptos iguais para os grupos 1 e 2 para as duas variáveis; mesmo intercepto para o grupo 3 nas duas variáveis; coeficientes angulares iguais para todos os grupos e para as duas variáveis.
7. Interceptos iguais para os grupos 1 e 2 nas duas variáveis; mesmo intercepto para o grupo 3 nas duas variáveis; coeficientes angulares iguais para todos os grupos e para as duas variáveis.
8. Interceptos iguais para todos os grupos para α_s1 -caseína e o grupo 3 para β -caseína; interceptos iguais para os grupos 1 e 2 para β -caseína; coeficientes angulares iguais para todos os grupos e para as duas variáveis.

3.2.2 Recursos computacionais

Para ajuste dos modelos aos dados utilizou-se o procedimento *mixed* implementado no *software* estatístico SAS (SAS INSTITUTE, 2008) e a análise de diagnóstico foi realizada no software R (R DEVELOPMENT CORE TEAM, 2012), cujos comandos estão apresentados detalhadamente no Anexo A. Para todos os testes empregou-se o nível de 5% de significância.

Para possibilitar que ambas as variáveis fossem incluídas em um mesmo modelo foram utilizadas variáveis indicadoras para α_{s1} e β -caseína, que foram combinadas com os demais fatores de interesse (tempo e grupo) para descrever os modelos de interesse, tanto na parte fixa como na parte aleatória.

O *proc mixed* permite que os componentes de variância e covariância sejam especificados por meio do comando *random* e *repeated*. O comando *random* é utilizado para descrever os efeitos aleatórios a serem utilizados e especificar a estrutura da matriz D , já o comando *repeated* é utilizado para especificar a estrutura da matriz Σ . Ambas as linhas de comandos, *random* e *repeated*, permitem a utilização do comando *group*, que pode ser utilizado para que a estimação seja feita de forma independente de acordo com os grupos especificados.

No caso do modelo bivariado, utilizou-se o comando *group* (conforme apresentado no Anexo A) na parte relacionada à estimação da matriz Σ para ajuste do modelo de independência condicional, sendo calculadas duas submatrizes para compor a matriz Σ : uma relacionada à variável α_{s1} -caseína e outra relacionada à variável β -caseína. Na grande maioria dos casos (exceto para o modelo 12), especificou-se $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \otimes \mathbf{I}_t$, em que a primeira matriz resulta da indicação das variáveis no comando *group* e a segunda resulta da indicação da estrutura *type = vc*. Sem a utilização do comando *group*, teríamos uma estrutura mais simples: $\Sigma = \sigma^2 \mathbf{I}_{2t}$.

Na análise dos resíduos, foram utilizados gráficos construídos por funções do software R (R DEVELOPMENT CORE TEAM, 2012): os gráficos *boxplot* bivariado, por meio da função *bv.boxplot* disponível no pacote *asbio* (AHO, 2011), com algumas modificações; os gráficos quantil-quantil da distribuição normal com envelope simulado por meio da função *qqPlot* disponível no pacote *car* (FOX; WEISBERG, 2011); o gráfico de quantil-quantil da distribuição qui-quadrado disponível no pacote *mvoutlier* (FILZMO-

SER; GSCHWANDTNER, 2012), com algumas modificações para construção do envelope simulado. Os gráficos de perfis individuais foram obtidos utilizando o pacote *lattice* (SARKAR, 2008). Os comandos utilizados estão apresentados no Anexo B.

3.2.3 Análise de resíduos

Utilizou-se os resíduos condicionais padronizados de cada variável, obtidos conforme descrito na seção 2.3.5, para verificar a pressuposição de normalidade univariada por meio do quantil-quantil da distribuição normal, bem como a homocedasticidade por meio do gráfico de resíduos versus valores preditos.

O gráfico *boxplot* bivariado dos resíduos condicionais padronizados obtidos para a variável α_{s1} —caseína versus resíduos condicionais padronizados da variável β —caseína foi empregado para verificar a presença de *outliers* no modelo bivariado. Para verificar a suposição de normalidade bivariada, utilizou-se o gráfico das distâncias de Mahalanobis estimadas versus os quantis da distribuição qui-quadrado, conforme indicado por Timm (2002) e Everitt (2005).

Para verificação da presença de indivíduos discrepantes utilizou-se o gráfico dos EBLUPs padronizados versus as unidades experimentais. Os EBLUPs padronizados foram obtidos pelo quociente entre os valores preditos e o desvio padrão estimado correspondente.

4 RESULTADOS E DISCUSSÃO

As informações necessárias para a comparação dos modelos M1-M12 estão apresentadas na Tabela 5. Vale notar que, iniciando pelo modelo M1, que apresenta o maior número de parâmetros possíveis, pode-se utilizar diferentes sequências de modelos encaixados. A Tabela 5 apresenta apenas uma das possíveis sequências de comparações utilizando o teste da razão de verossimilhanças. Os critérios de informação (AIC e BIC) podem ser empregados mesmo que os modelos não sejam encaixados.

O valor do logarítmo da função de verossimilhança, que é utilizado tanto no cálculo dos valores dos critérios de informação quanto na realização do teste da razão de verossimilhanças, também pode ser utilizado na comparação de dois modelos, desde que eles tenham o mesmo número de parâmetros. Neste caso a escolha feita coincidirá com a dos critérios de informação.

Tabela 5 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo método REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios

Modelo	nc	AIC	BIC	-2LVR	Comparação	Distribuição utilizada	Λ	p -valor
M1	12	464,8	473,3	440,8	-	-	-	-
M2	8	459,9	465,5	443,9	M1×M2	$\frac{1}{2}\chi_3^2 + \frac{1}{2}\chi_4^2$	3,1	0,4588
M3	5	462,7	466,2	452,7	M2×M3	$\frac{1}{2}\chi_2^2 + \frac{1}{2}\chi_3^2$	8,8	0,0443
M4	8	464,4	470,0	448,4	-	-	-	-
M5	7	460,7	465,7	446,7	M2×M5	χ_1^2	2,8	0,0943
M6	6	461,3	465,6	449,3	-	-	-	-
M7	5	460,5	464,0	450,5	M5×M7	χ_2^2	3,8	0,1496
M8	4	466,8	469,7	458,8	M7×M8	$\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$	8,3	0,0039
M9	3	463,8	465,9	457,8	-	-	-	-
M10	4	462,2	465,0	454,2	M9×M10	$\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$	3,6	0,0578
M11	4	456,9	459,7	448,9	-	-	-	-
M12	4	458,0	460,9	450,0	-	-	-	-

Na comparação do modelo 1 com o modelo 2, testa-se a inclusão de efeito aleatório no coeficiente angular para $alpha_{s1}$ —caseína, sendo a hipótese nula dada por:

$$H_0 : \begin{cases} \sigma_{b_{10}b_{11}} = \sigma_{b_{11}b_{20}} = \sigma_{b_{11}b_{21}} = 0 \\ \sigma_{b_{11}}^2 = 0 \end{cases}$$

Neste caso testa-se a inclusão de 4 parâmetros, sendo que um deles ($\sigma_{b_{11}}$) está na borda do espaço paramétrico, logo, a distribuição a ser utilizada no teste será $\frac{1}{2}\chi_3^2 + \frac{1}{2}\chi_4^2$. O teste resultou não significativo (p -valor = 0,4588), dessa forma o modelo selecionado é o modelo 2, que indica que o efeito aleatório do coeficiente angular para a variável $alpha_{s1}$ —caseína pode ser retirado do modelo.

Para verificar a necessidade de inclusão de efeito aleatório para o coeficiente angular para a variável $beta$ —caseína, deve-se comparar o modelo 2 com o modelo 3, sendo a hipótese nula dada por:

$$H_0 : \begin{cases} \sigma_{b_{21}b_{10}} = \sigma_{b_{21}b_{20}} = 0 \\ \sigma_{b_{21}}^2 = 0 \end{cases}$$

Testa-se a inclusão de 3 parâmetros, sendo que um deles ($\sigma_{b_{21}}$) está na borda do espaço paramétrico, logo, a distribuição a ser utilizada no teste será $\frac{1}{2}\chi_2^2 + \frac{1}{2}\chi_3^2$. A hipótese H_0 deve ser rejeitada (p -valor = 0,0443) e o efeito aleatório no coeficiente angular para a variável $beta$ —caseína deve ser mantido.

Na comparação dos modelos 4 e 2, que apresentam o mesmo número de parâmetros, verifica-se que o valor do AIC para o modelo 4 (464,4) é superior ao valor do AIC obtido para o modelo 2 (459,9), dessa forma conclui-se que o modelo 2 é mais adequado.

Na comparação do modelo 2 com o modelo 5 testa-se a necessidade de incluir um componente para correlação entre o intercepto da variável $alpha_{s1}$ —caseína e o coeficiente angular da $beta$ —caseína, ou seja, testa-se a hipótese $H_0 : \sigma_{b_{10}b_{21}} = 0$. Neste caso, testa-se a inclusão de apenas um parâmetro, utilizando uma χ_1^2 com 1 g.l. Como o teste resultou em não significativo (p -valor=0,0943), conclui-se que não há necessidade de inclusão deste parâmetro e o modelo 5 é considerado mais adequado.

O modelo 5 não pode ser comparado com o modelo 6 utilizando o teste da razão de verossimilhanças pois esses modelos não são encaixados. Como o valor do AIC é menor para o modelo 5 e os valores do BIC estão próximos, considerou-se o modelo 5 como mais adequado do que o modelo 6.

Ao realizar a comparação do modelo 5 com o modelo 7, testa-se a necessidade de se admitir uma correlação entre os interceptos das variáveis, bem como entre o

intercepto e coeficiente angular para a variável *beta*–caseína, $H_0 : \sigma_{b_{10}b_{20}} = \sigma_{b_{20}b_{21}} = 0$. Neste caso, testa-se a inclusão de dois parâmetros e a distribuição a ser utilizada no teste será a χ^2_2 . Ao comparar a diferença entre os valores de $-2LV$ dos modelos envolvidos no teste com o quantil da distribuição de referência, conclui-se que a hipótese H_0 não deve ser rejeitada (valor $p = 0,1496$), logo opta-se pelo modelo 7.

Na comparação dos modelos 7 e 8, testa-se a hipótese $H_0 : \sigma_{b_{21}}^2 = 0$. A estatística tem somente 1 grau de liberdade e como o parâmetro está na borda do espaço paramétrico o teste deve ser realizado utilizando $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$ como referência. Observa-se que este teste resultou em significativo ($p = 0,0039$), indicando que o modelo 7 continua sendo a melhor escolha.

O modelo 7 e os modelos 9, 10 e 11 não são encaixados, dessa forma será empregado um critério de informação para seleção do modelo mais apropriado. Observa-se que os valores referentes ao modelo 11 são inferiores, tanto para o AIC como para o BIC, logo opta-se pelo modelo 11. Para seleção entre o modelo 11 e o modelo 12, novamente utiliza-se um critério de informação, sendo selecionado o modelo 11.

Após a seleção da estrutura de covariâncias utilizando o modelo fixo maximal, realizou-se a seleção da parte fixa do modelo. As informações necessárias para a comparação dessas estruturas estão apresentadas na Tabela 6. A numeração dos modelos utilizada na tabela indica a estrutura de covariâncias selecionada (modelo 11) e o modelo de efeitos fixos após o ponto.

Tabela 6 - Número de parâmetros de efeito fixo (nf), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo método ML (LV) dos modelos ajustados, visando a escolher os efeitos fixos

Modelo	nf	AIC	BIC	-2LV	Comparação	Distribuição utilizada	Λ	p -valor
M11.1	12	520,8	532,1	488,8	-	-	-	-
M11.2	9	515,0	524,2	489,0	M11.1×M11.2	3	0,2	0,9776
M11.3	7	511,2	519,0	489,2	M11.2×M11.3	2	0,2	0,9048
M11.4	6	510,2	517,3	490,2	M11.3×M11.4	1	1,0	0,3173
M11.5	5	508,4	514,7	490,4	M11.4×M11.5	1	0,2	0,6547
M11.6	4	508,8	514,5	492,8	M11.5×M11.6	1	2,4	0,1213
M11.7	3	520,8	525,8	506,8	M11.6×M11.7	1	14,0	0,0002
M11.8	3	509,9	514,9	495,9	M11.6×M11.8	1	3,1	0,0783

Os testes da razão de verossimilhanças foram favoráveis ao modelo M11.8, que envolve 3 parâmetros de efeito fixo. No entanto o valor da estatística Λ para comparação desses modelos está próximo do valor crítico do teste e os critérios de informação foram favoráveis ao modelo M11.6, que envolve 4 parâmetros. Optou-se por este modelo que envolve: interceptos semelhantes entre os grupos 1 e 2 para $alpha_{s1}$ -caseína e para $beta$ -caseína, interceptos semelhantes para o grupo 3 em relação a $alpha_{s1}$ e $beta$ -caseína e coeficientes angulares semelhantes para as duas variáveis e todos os grupos.

Após a seleção da parte fixa do modelo, estudou-se novamente a seleção da parte aleatória do modelo misto. A partir dos resultados apresentados na Tabela 7, concluiu-se novamente pela escolha do modelo 11.

Tabela 7 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo métodos REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios, após a seleção da parte fixa

Modelo	nc	AIC	BIC	-2LVR	Comparação	Distribuição utilizada	Λ	p -valor
M1	12	519,0	527,5	495,0	-	-	-	-
M2	8	515,8	521,4	499,8	M1×M2	$\frac{1}{2}\chi_3^2 + \frac{1}{2}\chi_4^2$	4,8	0,4588
M3	5	518,4	521,9	508,4	M2×M3	$\frac{1}{2}\chi_2^2 + \frac{1}{2}\chi_3^2$	9,5	0,0160
M4	8	521,3	527,0	505,3	-	-	-	-
M5	7	517,1	522,0	503,1	M2×M5	χ_1^2	3,3	0,0693
M6	6	518,4	522,7	506,4	-	-	-	-
M7	5	517,7	521,2	507,7	M5×M7	χ_2^2	4,6	0,1003
M8	4	523,7	526,5	515,7	M7×M8	$\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$	8,0	0,0023
M9	3	518,7	520,8	512,7	-	-	-	-
M10	4	517,4	520,2	509,4	M9×M10	$\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$	3,3	0,0346
M11	4	512,1	514,9	504,1	-	-	-	-
M12	4	515,5	518,3	507,5	-	-	-	-

No *boxplot* bivariado dos resíduos condicionais padronizados da Figura 2 (a), observa-se a presença de alguns valores fora da elipse tracejada. Esses valores são candidatos a *outliers*.

No gráfico quantil-quantil para as distâncias de Mahalanobis (Figura 2 (b)) observa-se a presença de diversos pontos fora do envelope de simulação, o que indica falta de ajuste do modelo escolhido.

Os gráficos quantil-quantil dos resíduos de cada uma das variáveis versus

Figura 2 - (a) Boxplot bivariado para os resíduos da variável α_{s1} -caseína versus resíduos da variável β -caseína; (b) Gráfico quantil-quantil para a distância de Mahalanobis versus os quantis da quiquadrado

os quantis da distribuição normal (Figura 3) mostram alguns pontos fora do envelope de simulação.

Figura 3 - Gráfico quantil-quantil para os resíduos condicionais padronizados para o modelo selecionado: (a) para a variável α -caseína; (b) para a variável β -caseína

De acordo com os gráficos do resíduo e segundo indicações da responsável pelos dados, optou-se pela exclusão de cinco observações da variável α_{s1} -caseína (três

do grupo 1 e duas do grupo 3) e uma observação da variável *beta*-caseína (grupo 3) do conjunto de dados. A Figura 4 apresenta os perfis individuais por grupo após a retirada das observações consideradas fora do padrão.

Após a retirada das seis observações, foram realizados novamente os ajustes dos modelos visando à seleção dos componentes da parte aleatória, utilizando o modelo maximal para a parte fixa. Em seguida, procedeu-se à seleção da parte fixa e uma posterior seleção de componentes da parte aleatória, utilizando o modelo com a parte fixa selecionada no passo anterior.

Figura 4 - Perfis individuais das concentrações de α_{s1} -caseína e *beta*-caseína no leite UAT, por grupo, retirando algumas das observações

As Tabelas 8, 9 e 10 apresentam as informações necessárias para a seleção dos modelos (parte fixa e parte aleatória), recalculadas após a retirada dos pontos. O modelo 11 foi selecionado para a parte aleatória, enquanto para parte fixa foi selecionado o modelo M11.6, que inclui interceptos semelhantes entre os grupos 1 e 2 para α_{s1} -caseína e para *beta*-caseína, interceptos semelhantes para o grupo 3 em relação a α_{s1} e *beta*-caseína e coeficientes angulares semelhantes para as duas variáveis e todos os grupos.

Tabela 8 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo métodos REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios

Modelo	nc	AIC	BIC	-2LVR	Comparação	Distribuição utilizada	Λ	p -valor
M1	12	395,6	404,1	371,6	-	-	-	-
M2	8	398,2	403,9	382,2	M1×M2	$\frac{1}{2}\chi_3^2 + \frac{1}{2}\chi_4^2$	10,6	0,0455
M3	5	403,7	407,3	393,7	-	-	-	-
M4	8	396,1	401,8	380,1	M1×M4	χ_4^2	8,5	0,0748
M5	7	399,4	404,4	385,4	M1×M5	$\frac{1}{2}\chi_4^2 + \frac{1}{2}\chi_5^2$	13,8	0,0249
M6	6	392,2	396,4	380,2	-	-	-	-
M7	5	400,6	404,1	390,6	M1×M7	$\frac{1}{2}\chi_6^2 + \frac{1}{2}\chi_7^2$	19,0	0,0062
M8	4	409,7	412,5	401,7	-	-	-	-
M9	3	404,2	406,3	398,2	-	-	-	-
M10	4	399,6	402,4	391,6	-	-	-	-
M11	4	388,6	391,4	380,6	M1×M11	χ_8^2	9,0	0,3423
M12	4	395,8	398,7	387,8	-	-	-	-

Tabela 9 - Número de parâmetros de efeito fixo (nf), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo método ML (LV) dos modelos ajustados, visando a escolher os efeitos fixos

Modelo	nf	AIC	BIC	-2LV	Comparação	Distribuição utilizada	Λ	p -valor
M11.1	12	435,3	446,6	403,3	-	-	-	-
M11.2	9	429,7	438,9	403,7	M11.1×M11.2	χ_3^2	0,4	0,9402
M11.3	7	427,8	435,6	405,8	M11.2×M11.3	χ_2^2	2,1	0,3499
M11.4	6	426,0	433,1	406,0	M11.3×M11.4	χ_1^2	0,2	0,6547
M11.5	5	424,1	430,4	406,1	M11.4×M11.5	χ_1^2	0,1	0,7518
M11.6	4	423,7	429,3	407,7	M11.5×M11.6	χ_1^2	1,6	0,2059
M11.7	3	434,8	439,8	420,8	M11.6×M11.7	χ_1^2	13,1	0,0003
M11.8	3	426,8	431,7	412,8	M11.6×M11.8	χ_1^2	5,1	0,0239

As Figuras de 5 a 8 apresentam os gráficos de diagnóstico para o modelo selecionado. Pode-se visualizar apenas um ponto fora da elipse no *boxplot* bivariado, porém esses pontos não estão fora do envelope simulado no gráfico quantil-quantil para a distância de Mahalanobis versus os quantis da quiquadrado com 2 graus de liberdade (Figura 5). Observa-se ainda que o ângulo entre os segmentos de reta no *boxplot* bivariado está próximo de 90°, indicando que a maior parte da correlação entre as variáveis foi explicada pelo modelo utilizado. Nos gráficos quantil-quantil dos resíduos condicio-

nais padronizados das duas variáveis (Figura 6) observa-se que a maior parte dos pontos encontra-se dentro do envelope de simulação da distribuição normal, com exceção de três observações da variável *beta*-caseína que se encontram na borda do envelope.

Tabela 10 - Número de componentes de variância e covariância (nc), critérios de informação de Akaike (AIC) e Bayesiano (BIC) e o logaritmo da função de verossimilhança pelo métodos REML (LVR) dos modelos ajustados, visando à escolha dos efeitos aleatórios, após a seleção da parte fixa

Modelo	nc	AIC	BIC	-2LV	Comparação	Distribuição utilizada	Λ	p -valor
M1	12	434,2	442,7	410,2	-	-	-	-
M2	8	439,6	445,2	423,6	M1×M2	$\frac{1}{2}\chi_3^2 + \frac{1}{2}\chi_4^2$	13,4	0,0067
M3	5	444,5	448,0	434,5	-	-	-	-
M4	8	437,1	442,8	421,1	-	-	-	-
M5	7	440,9	445,8	426,9	-	-	-	-
M6	6	433,1	437,3	421,1	-	-	-	-
M7	5	443,9	447,4	433,9	-	-	-	-
M8	4	452,5	455,3	444,5	-	-	-	-
M9	3	444,7	446,9	438,7	-	-	-	-
M10	4	440,5	443,3	432,5	-	-	-	-
M11	4	428,0	430,9	420,0	M1×M11	χ_8^2	9,8	0,2793
M12	4	434,9	437,7	426,9	-	-	-	-

(a)

(b)

Figura 5 - (a) Gráfico boxplot bivariado para os resíduos da variável α_{s1} -caseína versus resíduos da variável β -caseína; (b) Gráfico quantil-quantil para a distância de mahalanobis versus os quantis da qui quadrado

Figura 6 - Gráfico quantil-quantil para os resíduos condicionais padronizados para o modelo selecionado: (a) para a variável alpha-caseína; (b) para a variável beta-caseína

Na Figura 7 observa-se que todos os resíduos padronizados encontram-se no intervalo $(-3, 3)$, sem apresentar qualquer tendência que possa indicar a falta de algum termo na parte fixa do modelo, confirmando que o modelo selecionado é adequado para descrever o comportamento dos dados. O gráfico dos EBLUPs (*empirical best linear unbiased predictor*) versus unidades experimentais está apresentado na Figura 9.

Figura 7 - Gráfico de valores preditos versus resíduos condicionais padronizados para o modelo selecionado: (a) para a variável alpha-caseína; (b) para a variável beta-caseína

(a)

(b)

Figura 8 - Gráfico de valores observados versus valores preditos para o modelo selecionado: (a) para a variável alpha-caseína; (b) para a variável alpha-caseína

Figura 9 - Gráfico dos EBLUPs (*empirical best linear unbiased predictor*) padronizados versus unidades experimentais

As estimativas dos parâmetros de efeito fixo para o modelo selecionado estão apresentadas na expressão 15.

$$\hat{y}_{rij} = 10,94G12_{1i} + 11,83G12_{2i} + 9,86G3_i - 0,0314t_j \quad (15)$$

Observa-se na expressão (15) que os grupos 1 e 2 apresentam concentrações iniciais de *beta*–caseína superiores as concentrações de *alpha*–caseína, o que não ocorre no grupo 3, para o qual as concentrações iniciais de *alpha*_{s1}–caseína e *beta*–caseína foram consideradas iguais, estimadas em 9,8 mg/mL. O grupo 3 possui a menor concentração das caseínas em estudo, no entanto a redução da concentração no decorrer do armazenamento ocorre de forma semelhante para os três grupos e para as duas caseínas estudadas, a uma taxa estimada em 0,0314 mg/mL por dia de armazenamento. A Figura 10 apresenta o gráfico dos perfis individuais e perfil médio estimados para cada um dos grupos.

A matriz $\hat{\mathbf{V}}_i$ apresenta as estimativas dos componentes de variâncias e covariâncias para o modelo marginal, na qual observa-se que tanto a covariância entre as duas variáveis respostas como a covariância entre medidas da mesma variável reposta e a heterocedasticidade foram modeladas marginalmente por meio dos efeitos aleatórios considerados.

$$\hat{\mathbf{V}}_i = \left[\begin{array}{ccccc|ccccc} 1,037 & 0,545 & 0,570 & 0,594 & 0,618 & 0,521 & 0,521 & 0,521 & 0,521 & 0,521 \\ 0,546 & 1,121 & 0,703 & 0,793 & 0,884 & 0,521 & 0,521 & 0,521 & 0,521 & 0,521 \\ 0,570 & 0,703 & 1,393 & 1,065 & 1,246 & 0,521 & 0,521 & 0,521 & 0,521 & 0,521 \\ 0,594 & 0,793 & 1,065 & 1,846 & 1,609 & 0,521 & 0,521 & 0,521 & 0,521 & 0,521 \\ 0,618 & 0,884 & 1,246 & 1,609 & 2,480 & 0,521 & 0,521 & 0,521 & 0,521 & 0,521 \\ \hline 0,521 & 0,521 & 0,521 & 0,521 & 0,521 & 1,266 & 0,545 & 0,570 & 0,594 & 0,618 \\ 0,521 & 0,521 & 0,521 & 0,521 & 0,521 & 0,545 & 1,351 & 0,703 & 0,793 & 0,884 \\ 0,521 & 0,521 & 0,521 & 0,521 & 0,521 & 0,570 & 0,703 & 1,622 & 1,065 & 1,246 \\ 0,521 & 0,521 & 0,521 & 0,521 & 0,521 & 0,594 & 0,793 & 1,065 & 2,075 & 1,609 \\ 0,521 & 0,521 & 0,521 & 0,521 & 0,521 & 0,618 & 0,884 & 1,246 & 1,609 & 2,710 \end{array} \right]$$

A correlação marginal entre *alpha*_{s1}–caseína e *beta*–caseína, considerando $t = 0$, obtida por meio dos componentes de variância e covariância estimados é dada por:

$$r_M = \frac{\hat{\sigma}_{b_{01}b_{02}}}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_{b_{01}}^2} \sqrt{\hat{\sigma}_2^2 + \hat{\sigma}_{b_{02}}^2}}$$

$$r_M = \frac{\sqrt{0,5218}}{\sqrt{0,5091 + 0,5218} \sqrt{0,7387 + 0,5218}}$$

$$r_M = 0,6337$$

No modelo considerado, a covariância entre as variáveis respostas é constante, no entanto, a correlação entre elas diminui com o passar do tempo devido ao aumento da variância com o passar do tempo, conforme observado na matriz $\hat{V}_i$.

Em sua tese, Fernandes (2007) concluiu que houve atividade proteolítica no leite UAT durante o armazenamento, com consequente diminuição de *beta*-caseína e de *alpha*_{s1}-caseína, similar à ordenação da taxa de proteólise encontrada para plasmina descrita por Grieve e Kitchen (1985). No entanto não foram realizados testes para verificar se a taxa de degradação da *beta*-caseína era superior à taxa de degradação da *alpha*_{s1}-caseína.

Figura 10 - Gráfico de perfis estimados segundo o modelo selecionado para as variáveis *alpha*-caseína e *beta*-caseína

Com a utilização do modelo bivariado, a realização de testes para as taxas de redução das caseínas possibilitou verificar que as concentrações iniciais de *alpha*_{s1} e *beta*-caseína foram diferentes apenas para os grupos com CCS baixa ou intermediária, sendo que as taxas de redução foram consideradas semelhantes nos três grupos.

5 CONSIDERAÇÕES FINAIS

A utilização de modelos mistos bivariados para dados longitudinais é bastante atraente porque permite estudar a inclusão de componentes de variância e covariância, que servem para explicar a variabilidade das respostas ao longo do tempo e também a correlação entre as variáveis respostas.

Permite a realização de testes estatísticos visando à escolha de parâmetros comuns às duas variáveis em estudo, o que é particularmente interessante quando as variáveis têm a mesma unidade de medida. Neste caso, as conclusões tiradas com o uso dos modelos mistos bivariados não seriam possíveis com a utilização de modelos mistos univariados.

Para que o uso do modelo misto bivariado seja proveitoso, é necessário que a variabilidade conjunta dos dados seja explicada tanto pela correlação entre as medidas repetidas ao longo do tempo, quanto pela correlação entre as variáveis respostas.

O uso de uma estrutura muito complexa para explicar as variâncias e covariâncias entre as medidas repetidas pode tornar desnecessária a inclusão de parâmetros visando a explicar a correlação entre as variáveis respostas.

A utilização do box-plot bivariado e do gráfico de quantil-quantil da distribuição quiquadrado na análise de resíduos facilitou bastante a identificação de observações discrepantes, pelo fato de permitir a visualização dos resultados para as duas variáveis, simultaneamente.

O proc mixed do SAS e as funções gráficas do R facilitaram bastante o ajuste dos modelos e a análise de resíduos. Entretanto, mostraram limitações por não permitir a especificação de estruturas diferentes para modelar as variâncias e covariâncias de cada uma das variáveis respostas. Em todos os casos estudados, os parâmetros podem ser diferentes, mas a estrutura é sempre a mesma.

Trabalhos futuros

Muito trabalho ainda precisa ser feito para consolidar a metodologia de análise de dados longitudinais bivariados, entre eles:

- (i) Implementar um pacote estatístico que torne mais flexível a especificação das estruturas de variâncias e covariâncias;

- (ii) Estudar a utilização de modelos mistos bivariados para variáveis que tenham um comportamento não linear em relação ao tempo;
- (iii) Estudar a utilização de modelos mistos bivariados quando uma variável tem comportamento linear em relação ao tempo e a outra variável apresentar um comportamento não linear;
- (iv) Estudar e consolidar um metodologia para análise de resíduo para modelos mistos bivariados;
- (v) Utilizar a metodologia bayesiana para o ajuste de modelos mistos bivariados para dados longitudinais.

REFERÊNCIAS

- ABDELMUR, A.S. **Análise de dados multivariados com medidas repetidas**. 2009. 89 p. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística, Universidade de São Paulo, Piracicaba, 2009.
- AHO, K. **Asbio**: a collection of statistical tools for biologists. R package version 0.3-40. 2011. Disponível em: <<http://CRAN.R-project.org/package=asbio>>. Acesso em: 18 out. 2012.
- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, New York, v. 19, n. 6, p. 716-723, Dec. 1974.
- ALBERT, P.S.; SHIH, J.H. An approach for jointly modeling multivariate longitudinal measurements and discrete time-to-event data. **The Annals of Applied Statistics**, Beachwood, v. 4, n. 3, p. 1517-1532, Feb. 2010.
- BARBOSA, M. **Uma abordagem para análise de dados com medidas repetidas utilizando modelos lineares mistos**. 2009. 134 p. Dissertação (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2009.
- BURDENSKI, T. Evaluating univariate, bivariate, and multivariate normality using graphical and statistical procedures. **Multiple Linear Regression Viewpoints**, Alabama, v. 26, n. 2, p. 15-28, 2000.
- CHAKRABORTY, H.; HELMS, R.W.; SEN, P.K.; COHEN, M.S. Estimating correlation by using a general linear mixed model: evaluation of the relationship between the concentration of HIV-1 RNA in blood and semen. **Statistics in Medicine**, North Carolina, v. 22, p. 1457-1464, May 2003.
- DIGGLE, P.J.; HEAGERTY, P.T.; KUNG-LEE, L.; ZEGER, S.L. **Analysis of longitudinal data**. New York: Oxford University Press, 2002. 369 p.
- EVERITT, B. **An R and S-PLUS R companion to multivariate analysis**. London: Springer-Verlag, 2005. 221 p.
- FERNANDES, A.M. **Efeitos dos níveis de células somáticas sobre a qualidade do leite integral obtido por processo UAT direto**. 2007. 109 p. Tese (Doutorado em Zootecnia) - Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, 2007.
- FIEUWS, S. **Mixed models for multivariate longitudinal data**. 2006. 126 p. Thesis (PhD in Medical Sciences) - Katholieke Universiteit Leuven, Leuven, 2006.
- FIEUWS, S.; VERBEKE, G. Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach. **Statistics in Medicine**, Diepenbeek, v. 23, p. 3093-3104, 2004.
- FIEUWS, S.; VERBEKE, G.; MOLENBERGHS, G. Random-effects models for multivariate repeated measures. **Statistical Methods in Medical Research**, London, v. 16, n. 5, p. 387-397, 2007.

FILZMOSER, P.; GSCHWANDTNER, M. Mvoutlier: multivariate outlier detection based on robust methods. 2012. R package version 1.9.7. Disponível em: <<http://CRAN.R-project.org/package=mvoutlier>>. Acesso em: 15 out. 2012

FITZMAURICE, G.M.; LAIRD, N.M.; ROTNITZKY, A.G. Regression models for discrete longitudinal responses. **Statistical Science**, Maryland, v. 8, n.3, p. 284-299, 1993.

FOX, J.; WEISBERG, S. **An R companion to applied regression**. 2nded. Thousand Oaks: Sage, 2011. Disponível em: <<http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>>. Acesso em: 27 out. 2011.

GRIEVE, P.A.; KITCHEN, B.J. Proteolysis in milk: the significance of proteinases originating from milk leucocytes and a comparison of the action of leucocyte, bacterial and natural milk proteinases on casein. **Journal of Dairy Research**, Cambridge, v. 52, p. 101-12, 1985.

HARVILLE, D.A. Bayesian inference for variance components using only error contrasts. **Biometrika**, Oxford, v. 61, n. 2, p. 383-385, 1974.

----- Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, Alexandria, VA, v. 72, n. 358, p. 320-338, 1977.

HARTLEY, H.O.; RAO, J.N.K. Maximum-likelihood estimation for the mixed analysis of variance model. **Biometrika**, Oxford, v. 54, n. 1/2, p. 93-108, 1967.

HEMDERSON, R.; DIGGLE, P.; DOBSON, A. Joint modelling of longitudinal measurements and event time data. **Biostatistics**, Oxford, v. 1, n. 4, p. 465-480, 2000.

HILDEN-MILTON, J.A. **Multilevel diagnostics for mixed and hierarchical linear models**. 1995. 110 p. Theis (PhD in Mathematics) - University of California, Los Angeles, 1995.

HUYNH, H.; FELDT, L.S. Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. **Journal of the American Statistical Association**, Boston, v. 65, n. 322, p. 1582-1589, 1970.

KENWARD, M.G.; ROGER, J.H. Small sample inference for fixed effects from restricted maximum likelihood. **Biometrics**, Arlington, v. 53, p. 983-997, 1997.

KITCHEN, B.J. Proteolysis in milk: the significance of proteinases originating from milk leucocytes and a comparison of the action of leucocyte, bacterial and natural milk proteinases on casein. **Journal of Dairy Research**, Cambridge, v. 52, p. 101-12, 1985.

LAIRD N.M.; WARE J.H. Random-effects models for longitudinal data. **Biometrics**, Arlington, v. 38, n. 4, p. 963-974, 1982.

LIMA, C.G. **Análise de dados longitudinais provenientes de experimentos em blocos casualizados**. 1996. 119 p. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1996.

MARCELINO, S.D.R.; IEMMA, A.F. Métodos de estimação de componentes de variância em modelos mistos desbalanceados. **Scientia Agricola**, Piracicaba, v. 57, n. 4, p. 643-652, 2000.

MAUCHLY, J.W. Significance test for sphericity of a normal n-variate distribution. **The Annals of Mathematical Statistics**, Beachwood, v. 11, n. 2, p. 204-209, June 1940.

MOLENBERGHS, G.; VERBEKE, G. **Models for discrete longitudinal data**. New York: Springer-Verlag, 2005. 683 p.

NOBRE, J.S.; SINGER, J.M. Residual analysis for linear mixed models. **Biometrical Journal**, Weinheim, v. 49, n. 6, p. 863-875, 2007.

NOBRE, J.S.N. **Métodos de diagnóstico para modelos lineares mistos**. 2004. 85 p. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2004.

PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, Oxford, v. 58, n. 3, p. 545-554, 1971.

PINHEIRO, J.C.; BATES, D.M. **Mixed-effects models in S and S-PLUS**. New York: Springer- Verlag, 2000. 528 p.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2012. Disponível em: <<http://www.R-project.org/>>. Acesso em: 16 out. 2012.

ROY, A.; KHATTREE, R. Testing the hypothesis of a Kronecker product covariance matrix in multivariate repeated measures data. Statistics and data analysis. In: SAS Users Group International. **Proceedings of the Statistics and Data Analysis Section**, n.199-130, p.1-11, 2006. Disponível em: <<http://www2.sas.com/proceedings/sugi30/toc.html>>. Acesso em: 14 out. 2012.

SARKAR, D. **Lattice: multivariate data visualization with R**. New York: Springer, 2008. 259 p.

SAS INSTITUTE. **SAS=STAT 9.2 user's guide**. Cary, 2008. Disponível em: <<http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#titlepage.htm>> . Acesso em: 27 out. 2012.

SATTERTHWAITE, F.E. An approximate distribution of estimates of variance components. **Biometrics Bulletin**, New Jersey, v. 2, n. 6, p. 110-114, 1946.

SCHUWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, Philadelphia, v. 6, n. 2, p. 461-464, 1978.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. **Variance components**. New York: John Wiley, 1992. 509 p.

SHAH, A.; LAIRD, N.; SCHOENFELD, D. A random-effects model for multiple characteristics with possibly missing data. **Journal of the American Statistical Association**, Alexandria, v. 92, n. 438, p. 775-779, 1997.

STRAM, D.O.; LEE, G.W. Variance components testing in the longitudinal mixed effects model. **Biometrics**, Arlington, v. 50, n. 4, p. 1171-1177, 1994.

THIÉBAUT, R.; JACQMIN-GADDA, H., CHÊNE, G.; LEPORT, C.; COMMENGES, D. Bivariate linear mixed models using SAS proc MIXED. **Computer Methods and Programs in Biomedicine**, Philadelphia, v. 69, p. 249-256, 2002.

TIMM, N.H. **Applied multivariate analysis**. New-York, Springer, 2002. 693 p.

VERBEKE, G.; MOLENBERGHS, G. **Linear mixed models for longitudinal data**. New York: Springer-Verlag, 2000. 569 p.

----- The use of score tests for inference on variance components. **Biometrics**, Leuven, v. 59, n. 2, p. 254-262, June 2003.

VERBEKE, G.; MOLENBERGHS, G.; RIZOPOULOS, D. Random effects models for longitudinal data. In: van MONTFORT, K.; OUD, J.H.L.; SATORRA, A. **Longitudinal research with latent variables**. Berlin: Springer, 2010. p. 37-96.

WATERNAUX, C.; LAIRD, N.M.; WARE, J.H. Methods for analysis of longitudinal data: blood-lead concentrations and cognitive development. **Journal of the American Statistical Association**, v. 84, n. 405, p. 33-41, Mar. 1989.

WEISS, R. **Modeling longitudinal data**. New York: Springer-Verlag, 2005. 432 p.

WEST, B.T.; WELCH, K.B.; GALECKI, A.T. **Linear mixed models: a practical guide using statistical software**. Washington: Taylor Francis Group; LLC Chapman Hall; CRC, 2007. 339 p.

ANEXOS

ANEXO A - Códigos para ajuste do modelo para os dados de armazenamento de leite UAT

```
proc mixed data=UATM covtest method=REML;
Title 'Modelo Beta e alpha - M1';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Ind_beta Ind_beta*Dia /type=UN subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo) R;
Run;
```

```
proc mixed data=UATM covtest method=REML;
Title 'Modelo Beta e alpha - M2';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Ind_beta Ind_t_beta /type=UN subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;
```

```
proc mixed data=UATM covtest method=REML;
Title 'Modelo Beta e alpha - M3';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Ind_beta /type=UN subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;
```

```
proc mixed data=UATM covtest method=REML;
Title 'Modelo Beta e alpha - M4';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Int Dia /type=UN subject=lote(Grupo) group=Ind_beta g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;
```

```
proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M5';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Ind_beta Ind_t_beta/type=UN(2) subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;
```

```
proc mixed data=UATM covtest method=REML;
Title 'Modelo Beta e alpha - M6';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Int Dia /type=VC group=Ind_beta subject=lote(Grupo)g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;
```

```

proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M7';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Ind_beta Ind_t_beta/type=UNR(1) subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;

```

```

proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M8';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Int /type=VC group=Ind_beta subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;

```

```

proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M9';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Int /type=VC subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;

```

```

proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M10';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
random Int Ind_t_beta /type=VC subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;

```

```

proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M11';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint
outp=Saida_comp residual;
random Int Ind_beta*Dia /type=VC subject=lote(Grupo) g v solution;
repeated /type=VC group=Ind_beta sub=lote(Grupo);
Run;

```

```

proc mixed data=UATM covtest method=REML maxiter=1000;
Title 'Modelo Beta e alpha - M12';
class lote Grupo Grupo4 Ind_beta Ind_alpha DiaC;
model Variavel = Grupo4 Dia / solution ddfm=sat noint residual;
repeated Ind_beta DiaC /type=UN@CS sub=lote(Grupo) r;
Run;

```

ANEXO B - Códigos para construção dos gráficos para os dados de armazenamento de leite UAT

```

require(car)
require(asbio)
require(mvoutlier)

RES=(read.table("dados_sem_obs.txt", header=T))
RES1=subset(RES,RES$Ind_alpha==1)
RES2=subset(RES,RES$Ind_alpha==0)

x=na.omit(cbind(RES1$Resid,RES2$Resid))
  covr <- covMcd(x, alpha = 0.5)
  dist <- mahalanobis(x, center = covr$center, cov = covr$cov)
  s <- sort(dist, index = TRUE)
  res <-s$x
a=(seq(1:nrow(x))-0.5)/75
env<-quantile(res,c(.025,0.5,.975))
for (i in 1:999){
  res<-cbind(res,sort(rchisq(nrow(x),2))))}
for (i in 1:nrow(res)){
  env<-cbind(env,quantile(res[i,],c(.025,0.5,.975))))}
matplot(qchisq(a,2),cbind(t(env[,-1]),res[,1]),
  col=c(2,2,2,1),pch=c(NULL,NULL,1),cex=0.8,type=c("l","l","l","p"),
  lty=c(2,1,2,1), lwd=c(2,2,2,1),
  xlab="quantis teóricos (Quiquadrado)", ylab="distâncias ordenadas")

# Adaptação da função bv.boxplot
biv.boxplot(x[,1],x[,2], robust=T, a=3.5, xlab="Resíduos condicionais
padronizados - variável alpha-caseína",
ylab="Resíduos condicionais padronizados - variável beta-caseína")

qqPlot(RES1$Resid, xlab="Quantis teóricos", ylab="Resíduos condicionais
padronizados",id.method="identify",id.cex=0.8)
qqPlot(RES2$Resid, xlab="Quantis teóricos", ylab="Resíduos condicionais
padronizados",id.method="identify",id.cex=0.8)

plot(RES1$Pred, RES1$Resid, xlab="Valores preditos",ylab="Resíduos condicionais
padronizados",ylim=c(-4,4))
abline(h=0, col="red")
abline(h=c(-3,3), lty=2)
identify(RES1$Pred, RES1$Resid,cex=0.8)
plot(RES2$Pred, RES2$Resid, xlab="Valores preditos",ylab="Resíduos condicionais
padronizados",ylim=c(-4,4))
abline(h=0, col="red")
abline(h=c(-3,3), lty=2)
identify(RES2$Pred, RES2$Resid,cex=0.8)

plot(RES1$Variavel,RES1$Pred,xlab="Valores observados",ylab="Valores preditos",
  xlim=c(4,12), ylim=c(4,12))
abline(a=0,b=1, col="red")

```

```

identify(RES1$Variavel,RES1$Pred,cex=0.8)
plot(RES2$Variavel,RES2$Pred,xlab="Valores observados",ylab="Valores preditos")
abline(a=0,b=1, col="red")
identify(RES2$Variavel,RES2$Pred,cex=0.8)

plot(RES1$lote, RES1$Resid, xlab="Unidades Experimentais (lotes)",
ylab="Resíduos condicionais padronizados - alpha-caseína",
xlim=c(0,16), ylim=c(-4.5,4.5),xaxt="n")
axis(1,1:15,1:15)
abline(h=c(-3,3), col="red",lty=2)
identify(RES1$lote, RES1$Resid,cex=0.8)
X11()
plot(RES2$lote, RES2$Resid, xlab="Unidades Experimentais (lotes)",
ylab="Resíduos condicionais padronizados - beta-caseína",
xlim=c(0,16), ylim=c(-4.5,4.5),xaxt="n")
axis(1,1:15,1:15)
abline(h=c(-3,3), col="red",lty=2)
identify(RES2$lote, RES2$Resid,cex=0.8)

```