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

Modelos lineares mistos em estudos toxicológicos longitudinais

Luzia Pedroso de Oliveira

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

Piracicaba
2015

Luzia Pedroso de Oliveira
Bacharel em Estatística

Modelos lineares mistos em estudos toxicológicos longitudinais
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 Ex-
perimentação Agronômica

Piracicaba
2015

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

Oliveira, Luzia Pedroso de
Modelos lineares mistos em estudos toxicológicos longitudinais / Luzia Pedroso de
Oliveira. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - -
Piracicaba, 2015.
159 p. : il.

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

1. Modelo linear hierárquico 2. Modelo multiníveis 3. Análise de dados longitudinais
4. Efeitos tóxicos 5. Modelos hierárquicos 6. Crescimento 7. Ganho de peso 8. Filhotes
de ratos I. Título

CDD 519.5
O48m

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

A Deus pela vida,

à minha irmã, Dra. Silvana (in memoriam), que
com seu exemplo, me inspirou a trabalhar com de-
dicação, responsabilidade e amor,

à minha filha Carla Vitória, minha
grande alegria e motivação

e aos meus pais Bruna e Walter
pelo apoio incondicional,

dedico.

AGRADECIMENTOS

Primeiramente a Deus, pela minha vida, pela alegria de aprender e ensinar, pela vontade e determinação para superar as dificuldades, pelas minhas conquistas e crescimento, pela paz e amor na realização do trabalho.

À CAPES pelo fundamental apoio financeiro concedido.

Ao professor Dr. Edwin Ortega, que me ajudou na elaboração do projeto de pesquisa, requisito para o ingresso no doutorado e foi meu orientador no período inicial.

Ao professor Dr. César Gonçalves de Lima, que desde o momento que passou a ser o meu orientador, a partir do 2º ano até o final, me motivou a crescer em áreas de meu interesse, tanto profissionais como pessoais. Com sua experiência de vida, como pesquisador, professor, orientador e amigo me proporcionou as melhores condições para eu desenvolver o meu trabalho. A ele os meus parabéns pela pessoa tão querida que é.

Aos professores da banca no exame de qualificação, Dra. Clarice Garcia Borges Demétrio, Dr. Euclides Braga Malheiros e Dr. Adriano Ferreti Borgatto, pelas sugestões, correções e críticas, contribuindo para o trabalho final. Aos professores suplentes da banca, Dr. Júlio Sílvio de Sousa Bueno Filho, em especial ao prof. Dr. Idemauro Antonio Rodrigues de Lara, por elogiar o meu trabalho dando-me estímulo e por sua amizade.

Aos professores do departamento de ciências exatas da ESALQ, pelos cursos de alta qualidade ministrados, pela amizade, por serem tão atenciosos para com os alunos compartilhando suas experiências de uma forma muito agradável, pela organização e realização dos vários *workshops* e cursos ministrados por pesquisadores renomados nacionais e internacionais. Em especial às professoras Dra. Roseli Aparecida Leandro e Dra. Sônia Maria de Stefano Piedade que me deram grande apoio antes e durante o doutorado, se preocupando comigo e com a minha filha e à professora Dra. Renata Alcarde, pela sua generosidade em todas as vezes que me ajudou.

Novamente às professoras Dra. Sônia Maria de Stefano Piedade e Dra. Renata Alcarde, pela valiosa experiência adquirida durante o estágio PAE - Programa de Aperfeiçoamento do Ensino Superior e por todos os momentos que compartilhamos.

Aos participantes do grupo de modelos mistos do departamento de ciências exatas da ESALQ, pela amizade e experiências compartilhadas.

À pesquisadora Dra. Aline de Holanda Nunes Maia, minha amiga desde a

graduação, por ter participado e me ajudado em alguns trabalhos de congressos, por ter me apresentado sua amiga, a pesquisadora Dra. Vera Lúcia Scherholz Salgado de Castro, também da EMBRAPA Meio Ambiente, em Jaguariúna e pelo conjunto de dados longitudinais bastante interessante que conseguiu com a mesma e que foi o meu marco inicial nessa área de estudo.

À pesquisadora Dra. Vera Lúcia Scherholz Salgado de Castro pela oportunidade de participar em alguns projetos, acompanhar alguns experimentos e pelos vários conjuntos de dados que me confiou, resultando em trabalhos apresentados em congressos de Estatística e em futuras co-autorias em artigos também na área de Toxicologia.

Aos técnicos do laboratório de Ecotoxicologia e Biossegurança da EMBRAPA Meio Ambiente, por serem tão atenciosos para comigo, mostrando os experimentos e respondendo às minhas dúvidas com entusiasmo.

Às recém-doutoras Aline Aparecida Becaro e Maria Célia Siqueira, integrantes de projetos da EMPRAPA que participei, pelo conhecimento compartilhado e por serem tão amáveis e pacientes com relação ao tempo necessário para as análises estatísticas.

Aos amigos e colegas de turma, pela amizade, pelos momentos de alegria, pela força nos momentos difíceis e pelos conhecimentos compartilhados. Em especial ao Everton Baptista da Rocha, meu grande amigo que me apoiou durante todo esse período e às ex-alunas, Dra. Simone Daniela Sartorio e Dra. Marina Rodrigues Maestre, por todas as vezes que me ajudaram.

Aos funcionários da ESALQ, em especial às secretárias Luciane Brajão e Solange de Assis Paes Sabadin, aos auxiliares Mayara Segatto e Rosni Honofre Aparecido Pinto e aos técnicos de laboratório Eduardo Bonilha e de redes de computadores Jorge Alexandre Wiendl, que sempre estiveram prontos para ajudar com muita boa vontade e amizade.

Aos funcionários das bibliotecas da ESALQ e também da UNICAMP pela atenção e bom atendimento. À USP pelo acesso aos artigos de periódicos assinados e à UNICAMP pelo acesso ao grande acervo de *e-books*, posteriormente também disponibilizado pela USP.

Aos amigos e colegas que encontro nos eventos científicos, pela amizade, atenção e troca de experiências.

Aos professores do Departamento de Estatística da UFSCar que contribuíram para a minha formação, em especial ao professor Dr. Carlos Alberto Ribeiro Diniz, do

qual nunca me esqueço, por tanto que fez por mim. Aos professores do Departamento de Estatística da UNICAMP, que também contribuíram para aumentar meu conhecimento na estatística.

À minha professora de matemática do ensino fundamental, de 5^a. a 7^a. série “dona” Irene, que me ajudou a despertar o gosto pela matemática. Por sua dedicação, por gostar muito de mim e valorizar o meu potencial.

À minha filha, Carla Vitória Pedroso Hernández, a minha grande alegria, amiga e companheira a 16 anos. Por ter contribuído com o silêncio em casa, tão importante para a minha maior concentração e por ser tão responsável e independente.

Ao amigo Pe. Sleiman Eid, que rezou muito por mim, pela minha família e amigos, para que eu pudesse ter paz e realizar um bom trabalho. Pelo seu coração tão generoso que proporcionou tantas alegrias para mim e para minha filha. Por ter me ensinado a ficar feliz diante de qualquer situação, para poder estar sempre inteira, dando o melhor de mim mesma.

A todos aqueles que não mencionei mas que contribuíram, direta ou indiretamente, para a realização desse trabalho e para o meu crescimento.

Agradeço de coração!

*“Não é o quanto fazemos, mas quanto
amor colocamos naquilo que fazemos.
Não é o quanto damos, mas quanto
amor colocamos em dar.”*

Madre Tereza de Calcutá

SUMÁRIO

RESUMO	13
ABSTRACT	15
LISTA DE FIGURAS	17
LISTA DE TABELAS	19
1 INTRODUÇÃO	21
2 DESENVOLVIMENTO	25
2.1 Dados longitudinais	25
2.2 Os estudos toxicológicos	26
2.2.1 As abordagens mais comuns de análise estatística de pesos diários de ratos em artigos científicos de diversas áreas	28
2.2.2 Considerações sobre as análises estatísticas de pesos diários de ratos dos artigos citados	33
2.3 Modelos lineares mistos para dados longitudinais	35
2.3.1 Especificação do modelo para dados longitudinais com dois níveis	36
2.3.1.1 O modelo linear misto na forma hierárquica	41
2.3.1.2 Tendências linear e quadrática	46
2.3.1.3 <i>Splines</i>	60
2.3.2 Especificação do modelo para dados longitudinais com três níveis	62
2.3.3 Os métodos de estimação e inferências	63
2.3.4 Etapas na seleção do modelo	65
2.3.5 O processo de validação do modelo	67
2.3.6 Os procedimentos computacionais	68
3 MATERIAL E MÉTODOS	71
3.1 Material	71
3.1.1 Experimento 1 - Avaliação da toxicidade por doses repetidas de um nanocomposto com relação aos perfis de pesos de ratos adultos machos e fêmeas	71
3.1.2 Experimento 2 - Avaliação dos efeitos de um herbicida na evolução de pesos de filhotes de ratas, sob dois aspectos: fase de gestação e dose	73
3.2 Métodos	82
3.2.1 Experimento 1	82
3.2.1.1 Implementação no SAS	85

3.2.2 Experimento 2	96
3.2.2.1 Implementação no SAS	103
4 RESULTADOS E DISCUSSÃO	117
4.1 Experimento 1	117
4.2 Experimento 2	125
5 CONSIDERAÇÕES FINAIS	133
REFERÊNCIAS	135
APÊNDICES	141

RESUMO

Modelos lineares mistos em estudos toxicológicos longitudinais

Os modelos mistos são apropriados na análise de dados longitudinais, agrupados e hierárquicos, permitindo descrever e comparar os perfis médios de respostas, levando em conta a variabilidade e a correlação entre as unidades experimentais de um mesmo grupo e entre os valores observados na mesma unidade experimental ao longo do tempo, assim como a heterogeneidade das variâncias. Esses modelos possibilitam a análise de dados desbalanceados, incompletos ou irregulares com relação ao tempo. Neste trabalho, buscou-se mostrar a flexibilidade dos modelos lineares mistos e a sua importância na análise de dados toxicológicos longitudinais. Os modelos lineares mistos foram utilizados para analisar os efeitos de dose no ganho de peso de ratos adultos machos e fêmeas, em teste de toxicidade por doses repetidas e também os efeitos de fase de gestação e dose nos perfis de pesos de filhotes de ratas tratadas. Foram comparados os modelos lineares mistos de regressão polinomial de grau 3, *spline* e de regressão por partes, ambos com um único ponto de mudança na idade média de abertura dos olhos dos filhotes, buscando o mais apropriado para descrever o crescimento dos mesmos ao longo do período de amamentação. São apresentados os códigos escritos no SAS/STAT para a análise exploratória dos dados, ajuste, comparação e validação dos modelos. Espera-se que o detalhamento da teoria e das aplicações apresentado contribua para a compreensão, interesse e uso desta metodologia por estatísticos e pesquisadores da área.

Palavras-chave: Modelo linear hierárquico; Modelo multiníveis; Análise de dados longitudinais; Efeitos tóxicos; Crescimento; Ganho de peso; Filhotes de ratos

ABSTRACT

Linear mixed models in longitudinal toxicological studies

Mixed models are appropriate in the analysis of longitudinal, grouped and hierarchical data, allowing describe and compare the average response profiles, taking into account the variability and correlation among the experimental units of the same group and among the values observed over the time in the same experimental unit, as well as the heterogeneity of variances. These models allow the analysis of unbalanced, incomplete or irregular data with respect to time. This work aimed to show the flexibility of linear mixed models and its importance in the analysis of longitudinal toxicological data. Linear mixed models were used to evaluate the effects of doses in the body weight gain of adult male and female Wistar rats, in repeated doses toxicity test and also the effects of pregnancy period and dose in the pups growth of treated dams. It were compared the linear mixed models of third degree polynomial regression, spline and piecewise regression, both with a single point of change in the average time of pups eyes opening, searching for the most appropriate one to describe their growth along the lactation period. The SAS/STAT codes used for exploratory data analysis, comparison and validation of fitted models are presented. It is expected that the detailing of the theory and of the applications presented contribute with the understanding, interest and use of this methodology by statisticians and researchers in the area.

Keywords: Hierarchical linear model; Multilevel model; Longitudinal data analysis; Toxic effects; Growth; Weight gain; Rat pups

LISTA DE FIGURAS

Figura 1 - Esquema da estrutura de dados longitudinais em 2 níveis	26
Figura 2 - Esquema da estrutura de dados longitudinais em 3 níveis	26
Figura 3 - Esquema das relações entre os coeficientes do modelo linear misto em 2 níveis	41
Figura 4 - Gráficos ilustrando alguns casos especiais de modelos mistos de tendência linear	48
Figura 5 - Representação gráfica de funções polinomiais com um único ponto de mu- dança com abcissa igual a 15	62
Figura 6 - Esquema representando o modelo misto em 3 níveis	63
Figura 7 - Foto do Experimento 1 - Pesagem rata	72
Figura 8 - Perfis individuais e médios de pesos nos níveis dos fatores sexo e dose . . .	73
Figura 9 - Fotos de um experimento conduzido na EMBRAPA em 2013	75
Figura 10 - Esquema do delineamento do Experimento 2	76
Figura 11 - Perfis de pesos individuais e médios dos filhotes ao longo de 23 dias	77
Figura 12 - Perfis de pesos individuais dos filhotes classificados por ninhada	78
Figura 13 - Perfis de pesos médios dos filhotes nas combinações dos níveis dos fatores <i>fase e dose</i>	79
Figura 14 - Diagramas de dispersão dos pesos dos filhotes em diferentes idades	80
Figura 15 - Gráficos de superfície dos pesos dos filhotes em função dos valores de cada covariável e das idades	81
Figura 16 - Estrutura hierárquica na unidade experimental (rata) em experimentos para avaliar os pesos longitudinais dos filhotes	97
Figura 17 - Pesos individuais, retas de regressão de mínimos quadrados ordinários e intervalos de 95 % de confiança para a média e para uma observação indi- vidual - Experimento 1	117
Figura 18 - Matriz de variâncias e covariâncias amostrais dos pesos de fêmeas (triân- gulo superior) e machos (triângulo inferior)	118
Figura 19 - Histograma, gráfico normal de probabilidades e de dispersão dos resíduos condicionais de Pearson - Experimento 1	122
Figura 19 - Histograma, gráfico normal de probabilidades e de dispersão dos resíduos condicionais de Pearson - Experimento 1 (cont.)	123

Figura 20 - Histograma, gráfico normal de probabilidades e de dispersão dos resíduos condicionais de Pearson de ajustes que convergiram para um máximo local	124
Figura 21 - Perfis individuais de pesos e retas individuais ajustadas a partir do modelo M1.12	124
Figura 22 - Gráficos mostrando a necessidade da covariável sazonal nos modelos lineares mistos polinomiais de grau 3 na idade - Experimento 2	128
Figura 23 - Gráfico normal de probabilidades, diagramas de caixas em função das idades e de dispersão em função dos preditos, dos resíduos condicionais de Pearson de modelos ajustados a partir das três abordagens	129
Figura 24 - Perfis médios de pesos observados e preditos de alguns modelos nas combinações de <i>fase</i> e <i>dose</i>	130
Figura 25 - Gráficos com os perfis individuais de pesos observados e preditos - Modelo M2.13	131

LISTA DE TABELAS

Tabela 1 - Casos particulares de modelos com tendência linear	47
Tabela 2 - Dados de entrada do Experimento 1 – Formato multivariado	87
Tabela 3 - Dados de entrada do Experimento 1 – Formato univariado	89
Tabela 4 - Modelos polinomiais de grau 3 selecionados para comparação - Experi- mento 2	101
Tabela 5 - Modelos das três abordagens selecionados para comparação - Experimento 2	102
Tabela 6 - Dados de entrada do Experimento 2 – Formato multivariado	103
Tabela 7 - Dados de entrada do Experimento 2 – Formato univariado	107
Tabela 8 - Características dos modelos ajustados e estatísticas de ajuste - Experi- mento 1	120
Tabela 9 - Estimativas e erros padrões dos parâmetros de efeitos fixos dos modelos ajustados	121
Tabela 10 - Níveis descritivos dos testes tipo III de efeitos fixos dos modelos ajustados	121
Tabela 11 - Estatísticas de ajuste dos modelos selecionados - Experimento 2	127
Tabela 12 - Comparação dos modelos de regressão polinomial por meio do teste da razão de verossimilhanças	127

1 INTRODUÇÃO

Na pesquisa, muitas vezes, temos como objetivo obter conclusões a respeito de características de uma ou mais populações em estudo, a partir de um conjunto de dados amostrados, o que faz da estatística uma ciência de grande importância prática. Para que seja possível a obtenção de resultados confiáveis, duas etapas são relevantes: um bom planejamento do experimento ou da amostragem e a formulação de um modelo correto que descreva de forma apropriada o processo que gerou os dados, ou seja, a variabilidade e a correlação entre os valores observados, assim como a relação entre a variável resposta e as variáveis independentes.

Os modelos estatísticos clássicos de regressão ou análise de variância (ANOVA - *analysis of variance*), também chamados de modelos de efeitos fixos, são apropriados nos casos em que os valores observados são provenientes de uma distribuição normal e não são correlacionados. Esses modelos agregam um único componente aleatório que representa a variação que é devida ao acaso (erro aleatório) e cuja variabilidade é assumida constante. Dessa forma, esses modelos não são apropriados na análise de dados agrupados, hierárquicos e longitudinais, dada a correlação entre as observações feitas nos indivíduos de um mesmo grupo ou em um mesmo indivíduo ao longo do tempo e as possíveis alterações nas variâncias das respostas ao longo do tempo.

Os modelos lineares mistos (*linear mixed models*), incluem além da variação aleatória do acaso, pelo menos mais uma fonte de variação aleatória (efeitos aleatórios não observáveis), ambas de forma linear. Esses modelos possibilitam uma abordagem mais realista das características de muitos conjuntos de dados, pela sua flexibilidade para modelar as estruturas de dependência e de variâncias não constantes. Esse modelo permite a análise de dados desbalanceados, incompletos ou irregulares com relação ao tempo. Dessa forma, o modelo referido como linear geral (GLM - *general linear model*) contendo somente efeitos fixos, variância homogênea e erros normais independentes, passou a ser um dos casos especiais mais restritos do modelo linear misto.

Algumas extensões dos modelos lineares mistos, formulados inicialmente assumindo distribuição normal multivariada para os erros e efeitos aleatórios, foram os modelos lineares generalizados mistos, para modelar variáveis respostas discretas, como contagens, dicotômicas ou com vários níveis categóricos (MCCULLOCH, 2003; JIANG, 2007; STROUP, 2012) e os modelos não lineares mistos, para descrever também as relações não lineares nos

parâmetros (DAVIDIAN; GILTINAN, 1995; VONESH; CHINCHILLI, 1996).

Os modelos lineares mistos de regressão polinomial de ordem baixa, ou seja, de grau 1 ou 2, são bastante úteis, pois possibilitam descrever as tendências lineares ou quadráticas das respostas, respetivamente, de uma forma mais simples. No caso de relações funcionais mais complexas pode haver necessidade de incluir termos cúbicos, quárticos, ou de ordem ainda maiores, o que compromete a interpretabilidade e a parcimônia desses modelos. Em situações em que as respostas apresentam um comportamento sigmoidal ou com assíntotas, por exemplo, os modelos não lineares nos parâmetros são preferíveis devido à melhor representação do processo e interpretabilidade dos parâmetros. Uma vez que os modelos não lineares mistos exigem métodos de estimação mais complexos, em alguns casos tenta-se primeiramente alguma transformação na variável resposta, de forma a buscar a sua linearização, possibilitando o uso de modelos lineares mistos.

Uma alternativa aos modelos lineares mistos paramétricos são os modelos mistos semiparamétricos (*semiparametric mixed models*), cujo desenvolvimento teórico na análise de dados longitudinais vem avançando rapidamente nos últimos 10 anos, motivado pela demanda das aplicações práticas em várias áreas (RUPPERT; WAND; CARROLL, 2003; WU; ZHANG 2006). Esses modelos são constituídos de componentes paramétricos e não paramétricos, sendo a parte não paramétrica usada para descrever as tendências sem nenhum padrão ao longo do intervalo de tempo estudado, por meio de formas funcionais flexíveis estimadas a partir dos dados. Os métodos para estimar as formas funcionais baseados em *splines*, ou seja, em polinômios ajustados por partes e unidos de forma contínua e suave (SMITH, 1979), tornaram-se preferidos, devido à maior simplicidade e conexão bastante atrativa com o método de estimação dos modelos lineares mistos (BRUMBACK; RICE, 1999; DURBÁN et al., 2005).

Em cenários mais simples, pode ser suficiente considerar apenas alguns pontos de ligação (mudança), ou até mesmo um único, para unir diferentes tendências ao longo do tempo (RAUDENBUSH; BRYK, 2002; FITZMAURICE; LAIRD; WARE, 2011). Esses modelos têm a vantagem de manter a interpretabilidade dos parâmetros, quando os polinômios em cada segmento forem de grau 1 ou 2. Há casos em que há interesse também na estimação do(s) ponto(s) de ligação que é(são) tratado(s) como parâmetro(s), sendo esses modelos conhecidos como modelos mistos de regressão segmentada (*segmented regression*) (GALLANT; FULLER, 1973; PIEPHO; OGUTU, 2003; ROBBINS; SAXTON; SOUTHERN, 2006).

Durante o processo de busca por um modelo estatístico que responda aos objetivos da pesquisa de forma correta, é preciso checar se as suposições do modelo proposto não estão sendo violadas, verificar a necessidade de refinar o ajuste incluindo ou excluindo covariáveis e também analisar a sensibilidade do modelo a observações ou conjunto de observações (perfis) atípicos, ou seja, fora do padrão da massa de dados. No caso dos modelos mistos também é necessário buscar estruturas apropriadas para representar as variâncias e as correlações entre as observações em tempos distintos e checar as suposições feitas sobre as demais fontes de variação aleatórias.

As análises de resíduo e de sensibilidade em modelos mistos são mais complexas quando comparadas com as dos modelos lineares clássicos, uma vez que há dois tipos de resíduos: condicionais e marginais e mais de uma fonte de variação aleatória.

As aplicações dos modelos mistos vêm sendo mais frequentes em diversas áreas, como por exemplo, ciências médicas, sociais, agrárias e biológicas, na psicologia e educação. Nos estudos toxicológicos envolvendo dados longitudinais para avaliar, por exemplo, os efeitos de produtos químicos nos perfis de pesos de ratos, as análises estatísticas apresentadas em trabalhos publicados em revistas da área costumam ser bastante simples, utilizando modelos que não admitem ou não descrevem de forma apropriada a estrutura de correlações entre os valores observados ao longo do tempo e a heterogeneidade de variâncias das respostas.

Neste trabalho, foram considerados os modelos lineares mistos assumindo distribuição normal multivariada para os erros e os efeitos aleatórios, sendo apropriados para variáveis respostas contínuas, cobrindo assim um grande número de aplicações práticas. O trabalho apresenta as vantagens desses modelos na avaliação da toxicidade de produtos químicos em estudos longitudinais, por meio das análises de dois experimentos, buscando fomentar o interesse e o uso da metodologia por pesquisadores da área.

De forma mais específica, o trabalho visa:

- apresentar uma revisão das técnicas estatísticas apresentadas em artigos de revistas de toxicologia, nutrição, áreas médicas, entre outras, na análise de dados longitudinais de pesos de ratos ao longo do tempo;
- apresentar uma revisão dos modelos lineares mistos para dados longitudinais com dois e três níveis de variação aleatória para modelar tendências linear, quadráticas e *splines*;
- propor o uso de modelos lineares mistos com dois e três níveis de variação aleatória para

avaliar os efeitos de produtos químicos na evolução de pesos de ratos, em testes de toxicidade sub-crônica e reprodutiva;

- desenvolver códigos no **SAS/STAT**, para obtenção de gráficos e ajuste dos modelos lineares mistos, visando simplificar o uso da metodologia.

Na seção 2 é apresentada uma revisão sobre os testes toxicológicos e principais técnicas estatísticas que vem sendo utilizadas na última década, na análise da evolução diária de pesos de ratos. As técnicas estatísticas empregadas nas análises longitudinais são discutidas levando em conta as suas aplicabilidades e restrições com relação ao uso. Ainda na seção 2 é apresentada uma revisão sobre os modelos lineares mistos, com relação a especificação, estimação, comparação e validação dos modelos. Na seção 3, são descritos os dois experimentos analisados, ilustrando o uso dos modelos lineares mistos nos estudos toxicológicos. Na seção 4, para cada experimento, são apresentadas as metodologias utilizadas para descrever as tendências dos pesos de ratos adultos e de filhotes e são comparados os modelos ajustados. Na seção 5 são apresentadas as considerações finais. No Apêndice A são detalhados os passos para se obter conjuntamente, por meio das equações de Henderson (1950), as estimativas dos parâmetros de efeitos fixos e as predições dos efeitos aleatórios. Nos Apêndices B e C são apresentadas as linhas de comandos escritas no **SAS/STAT**, versão 9.3, para o ajuste dos modelos.

2 DESENVOLVIMENTO

Os estudos longitudinais são indispensáveis para avaliar as mudanças de alguma característica que ocorrem ao longo do tempo em cada indivíduo, assim como os efeitos dos fatores que as influenciam.

De acordo com Singer e Willett (2003), somente a partir da década de 80, com a aplicação dos modelos mistos é que as mudanças nas respostas que ocorrem ao longo do tempo passaram a ser bem analisadas. Uma introdução e a perspectiva histórica dos avanços nas análises de dados longitudinais são apresentadas no livro de Fitzmaurice et al. (2009), assim como uma coletânea de artigos com alguns dos principais desenvolvimentos teóricos e aplicações de modelos mistos.

Neste capítulo inicialmente são apresentadas algumas estruturas de dados provenientes de estudos longitudinais (seção 2.1) e nas duas seções seguintes, os estudos toxicológicos e o modelos lineares mistos.

Com relação aos estudos toxicológicos (seção 2.2) é apresentada uma revisão das análises estatísticas de artigos científicos de toxicologia, nutrição e áreas médicas para avaliar a evolução dos pesos de ratos, em testes de toxicidade sub-crônica e reprodutiva ou em dietas (seção 2.2.1), seguida de algumas considerações sobre as análises que vêm sendo utilizadas (seção 2.2.2).

Com relação aos modelos lineares mistos para dados longitudinais (seção 2.3) é apresentada uma revisão sob vários aspectos: especificação para dados longitudinais com dois e três níveis de variação aleatória (seções 2.3.1 e 2.3.2), métodos de estimação e inferências (seção 2.3.3), etapas para seleção do modelo (seção 2.3.4), processo de validação do modelo (seção 2.3.5) e procedimentos computacionais (seção 2.3.6).

2.1 Dados longitudinais

Os dados longitudinais provêm de medidas repetidas em diferentes ocasiões, obtidas de uma maneira ordenada, por exemplo, ao longo do tempo, em cada indivíduo. Os dados longitudinais são casos particulares de dados multiníveis (hierárquicos ou aninhados) com dois níveis, sendo o nível 1 constituído pelas medidas repetidas nas diferentes ocasiões e o nível 2 pelos indivíduos, como ilustrado na Figura 1. Os dados longitudinais podem apresentar mais de um nível de agrupamento quando, por exemplo, as medidas são feitas

ao longo do tempo (nível 1) em dentes (nível 2) de pacientes (nível 3); medidas ao longo do tempo (nível 1) em alunos (nível 2) dentro de escolas (nível 3), como ilustrado na Figura 2.

Figura 1 – Esquema da estrutura de dados longitudinais em 2 níveis

Figura 2 – Esquema da estrutura de dados longitudinais em 3 níveis

Nos estudos longitudinais o planejamento pode ser balanceado ou não, com relação ao tempo. No caso balanceado, todas as unidades amostrais são observadas nos mesmos tempos, igualmente espaçados ou não, enquanto que no caso desbalanceado, conjuntos diferentes de unidades amostrais podem ser observados em conjuntos diferentes de tempo. De acordo com Singer, Nobre e Rocha (2011), se o planejamento foi balanceado mas ocorreu pelo menos uma observação perdida, os dados passam a ser desbalanceados com relação ao tempo.

A característica que distingue os dados em multiníveis é que as medidas nas unidades de um mesmo grupo são mais similares quando comparadas com as de outros grupos.

2.2 Os estudos toxicológicos

Para que as substâncias químicas usadas em produtos alimentares, cosméticos ou medicamentos sejam consideradas suficientemente seguras para o consumo ou exposição humana, elas devem ser testadas primeiramente em outros seres vivos.

Em modelo animal, os ratos são os mamíferos mais utilizados devido à sua pureza genética e semelhanças com a biologia humana. Devido à rápida reprodução e grandes ninhadas, possibilitam uma avaliação mais detalhada da segurança de um produto, por meio do estudo de várias gerações. O uso de animais de laboratório deve estar de acordo com os padrões éticos de cuidado e bem estar animal.

Os testes toxicológicos são também realizados com minhocas, peixes, crustáceos, hidra, algas, dentre outros organismos, sendo os organismos aquáticos usados para avaliar os efeitos sobre a biota e riscos ao ambiente.

A maioria dos produtos químicos está sujeita a rigorosas exigências de segurança impostas pelos governos para que possam ser comercializados. No Brasil, o órgão responsável pela segurança e regulamentação dos produtos é a ANVISA – Agência Nacional de Vigilância Sanitária.

Para avaliar a toxicidade de um produto são necessários vários tipos de testes, como, por exemplo, de toxicidade aguda, sub-crônica, crônica, reprodutiva e neurotoxicidade. Informações sobre os testes de toxicidade podem ser encontradas, por exemplo, na enciclopédia online “*of earth*” (2003).

Os protocolos para a realização de testes seguros fornecidos pela OECD (*Guidelines for Testing of Chemicals - Organization for Economic Cooperation and Development*), são aceitos internacionalmente e atualizados regularmente por milhares de especialistas dos seus países membros. Para cada tipo de teste há um protocolo específico. Por exemplo, o protocolo 422, (OECD, 1996), fornece as diretrizes para avaliar sinais de toxicidade em ratos com relação à reprodução e desenvolvimento, assim como os decorrentes da exposição continuada à substância química por um período limitado de tempo (sub-crônica). As instruções fornecidas vão desde a escolha da espécie do animal, condições de alojamento, alimentação, preparação e administração das doses, procedimentos de acasalamento, desenho experimental, pesagem, variáveis a serem coletadas, hematologia, patologia, características bioquímicas, aspectos comportamentais, autópsia, histopatologia até a obtenção e interpretação dos resultados. Para registro na ANVISA, os testes devem ser realizados de acordo com os Princípios das Boas Práticas de Laboratório (OECD, 1998) e seguir os protocolos de referência reconhecidos internacionalmente (sempre os mais atuais).

Nos ensaios para avaliar aspectos na reprodução e desenvolvimento, algumas das variáveis comumente observadas são: pesos diários das ratas durante a gestação, número

de dias de gestação, número de filhotes nascidos, número de filhotes natimortos, peso inicial dos filhotes, número de dias até o aparecimento de características físicas dos filhotes como aparecimento de penugens e pelos, descolamento das orelhas, abertura dos ouvidos e dos olhos, descida dos testículos dos filhotes machos, abertura das vaginas dos filhotes fêmeas, pesos dos filhotes ao longo da amamentação e pesos dos órgãos.

Os dados de peso corporal durante o período de exposição às substâncias são coletados em praticamente todos os ensaios com produtos químicos, uma vez que os animais devem ser pesados individualmente para a manipulação das dosagens e porque a evolução dos pesos é uma característica importante na avaliação de efeitos tóxicos. Nos ensaios reprodutivos são avaliados os pesos ao longo da gestação, assim como o efeito do produto no crescimento da prole durante a amamentação.

Com relação às características que são avaliadas ao longo do tempo, além dos pesos dos animais, consumo de ração/água, pode haver outras relacionadas a atividades motoras, como medições semanais dos tempos de sustentação dos filhotes em um fio de nylon ou dos números de passos em falso em uma malha etc. Dados desse tipo necessitam de uma análise que leve em conta a variabilidade biológica inter-indivíduos, devida à diferenças fisiológicas e genéticas, como a variabilidade intra-indivíduos, ou seja, ao longo do tempo de experimentação. Nos testes envolvendo várias gerações também é importante considerar a estrutura hierárquica dos dados, por exemplo, filhotes aninhados às progenitoras, uma vez que existe correlação nas respostas obtidas em filhotes de uma mesma rata.

2.2.1 As abordagens mais comuns de análise estatística de pesos diários de ratos em artigos científicos de diversas áreas

As técnicas estatísticas empregadas na análise de dados longitudinais em revistas de toxicologia, nutrição, medicina, entre outras, costumam ser bastante simples, deixando muitas vezes de explorar aspectos importantes dos dados coletados ou até mesmo exibindo resultados pouco precisos.

Analisar e comparar a evolução dos pesos de ratos durante a gestação e dos filhotes durante a lactação são aspectos importantes não somente na avaliação da toxicidade reprodutiva, mas também para avaliar efeitos de dietas, doenças, entre outros.

A seguir, são apresentadas algumas abordagens de análise estatística de pesos coletados diariamente, publicadas em revistas de diversas áreas, de modo a fornecer uma

revisão das técnicas que estão sendo utilizadas. Foram selecionados artigos em revistas de alto impacto nas respectivas áreas. Buscou-se as análises contendo o gráfico dos pesos ao longo do tempo. Foram relatadas as técnicas estatísticas e também as conclusões, visando apresentar as interpretações consideradas relevantes pelos autores. As publicações estão ordenadas a partir das mais recentes.

Vido et al. (2014) avaliaram o efeito da obesidade materna sobre o metabolismo da prole. Os filhotes das ratas foram pesados diariamente, sendo apresentados os gráficos com as médias e erros padrões diários por grupo, um deles durante o período de lactação que foi até 28 dias e o outro no período pós desmame até 90 dias. Os autores comentaram sobre as diferenças nas tendências de pesos sugeridas pelos gráficos, entretanto, as comparações são feitas apenas com relação aos ganhos de pesos durante os dois períodos analisados, a partir do teste t não pareado ou ANOVA seguida de Tukey.

Arzu et al. (2014) avaliaram os efeitos tóxicos do fungicida vinclozolin usado em frutas e vegetais. As ratas receberam as doses do fungicida por via oral sendo pesadas diariamente durante a gestação, mas somente foram apresentados, em tabelas para cada grupo, os pesos médios no início e no final da gestação e os respectivos desvios padrões, assim como os ganhos de peso médios e os desvios padrões no período (em g e %) juntamente com os resultados da ANOVA com um fator (dose) e dos testes de comparação de pares de médias de Games-Howell. Baseado nesses resultados os autores concluíram que não houve mudanças nos pesos e ganhos de pesos corporais das ratas.

Caporali et al. (2014) compararam o desenvolvimento de filhotes de ratas que desde o desmame até o acasalamento ficaram ou não em ambiente enriquecido, ou seja com interação social, atividade física e aumento das oportunidades de aprendizado. Os filhotes foram pesados diariamente e os autores apresentaram o gráfico com os pesos médios corporais diários dos grupos controle e tratado, ao longo de 25 dias de lactação. A análise estatística realizada foi ANOVA com um fator de tratamento (grupo) e um fator longitudinal (idade), mencionada como ANOVA com dois fatores para medidas repetidas, seguida do teste HSD de Tukey, sendo os dados primeiramente testados para normalidade (teste de Shapiro Wilk) e homocedasticidade (teste de Levene). Foram apresentados os níveis descritivos dos testes da interação e dos fatores principais e apenas o efeito de idade foi significativo.

Santos et al. (2013) avaliaram sinais de toxicidade do extrato de jacarandá, usado no tratamento de doenças inflamatórias e infecciosas, por meio de teste de toxicidade

aguda com ratos adultos machos que receberam a substância por via oral ao longo de 28 dias. Os animais foram pesados diariamente sendo apresentados e comparados somente os pesos médios finais e erros padrões por grupo, usando ANOVA e Tukey-Kramer (controle com os demais grupos). Os autores concluíram que a substância não causou efeitos adversos no ganho de peso corporal.

Andrade et al. (2012) avaliaram os efeitos da desnutrição materna nos seus filhotes. Foram considerados no estudo os filhotes machos de apenas duas ratas, sendo igualmente divididos em dois grupos, mães com restrição e sem restrição proteica durante a lactação. Os autores apresentaram o gráfico com os pesos médios diários dos filhotes dos grupos controle e tratado, ao longo de 21 dias de lactação mas comparam apenas os pesos aos 21 dias usando o teste de Mann-Whitney, concluindo que nesta idade os pesos corporais dos filhotes da rata com restrição proteica foram aproximadamente 61% menores que os dos filhotes da rata controle.

Dumortier et al. (2011) avaliaram os efeitos no desenvolvimento fetal de filhotes de ratas que receberam doses de dexametasona durante diferentes fases da gestação. O gráfico com os pesos médios e os erros padrões diários durante a gestação das ratas dos grupos tratados e do controle foi apresentado. Para comparar os pesos médios das ratas aos 21 dias de gestação e os pesos médios iniciais dos filhotes foi utilizada ANOVA com um fator e teste de Newman-Keuls comparando o controle com cada um dos grupos. Os autores concluíram que o uso de dexametasona prejudicou o ganho de peso corporal normal das ratas. Aos 21 dias os pesos das ratas tratadas apenas na última semana da gestação reduziram em 18,5% e das ratas tratadas durante toda a gestação reduziram em 25%, comparados com os do grupo controle. Os pesos corporais dos filhotes ao nascer também foram reduzidos, no primeiro caso em 18% e no segundo, em 30%, quando comparados com os pesos do grupo controle.

Brennan et al. (2008) avaliaram o efeito da restrição de nutrientes durante a gestação no desenvolvimento da prole. As ratas foram pesadas diariamente durante e após a gestação e os filhotes, pesados diariamente durante as 3 semanas de lactação e semanalmente após o desmame. Os efeitos dos fatores dieta e idade foram testados por meio de ANOVA e as médias foram comparadas duas a duas pelo teste de Holm-Sidak. Foram apresentados em gráficos os pesos médios diários e os erros padrões, para os dois grupos de ratas durante a gestação e de filhotes durante a lactação. Nos gráficos foram indicados os instantes em que

os pesos médios das mães que tiveram redução de nutrientes na alimentação foram diferentes dos pesos médios do grupo controle. Também foram indicados os instantes em que os pesos médios dessas mães foram maiores (menores) que o peso médio das mesmas no início do experimento e os instantes em que houve diferença entre os pesos médios dos filhotes. Os autores concluíram que as ratas do controle ganharam peso ao longo da gestação, enquanto que as ratas com restrição de nutrientes perderam peso de imediato e até 15-16 dias, sendo mais leves do que as do controle durante toda a gestação. Durante a lactação todas as ratas ganharam peso. Os filhotes de mães com restrição nasceram e permaneceram mais leves somente até o 3º dia. Aos 21 dias os filhotes do controle passaram a ser mais pesados.

Mazzanti et al. (2008) estudaram os efeitos do extrato de *cimicifuga racemosa*. Os ratos adultos machos usados no experimento foram pesados diariamente ao longo de 30 dias. Segundo os autores, os pesos corporais dos animais foram analisados usando ANOVA com dois fatores (tratamento $\times$ dia) para medidas repetidas, ou seja, ANOVA com um fator de tratamento (grupo) e um fator longitudinal (idade). Os autores concluíram que o extrato não provocou qualquer efeito no ganho de peso corporal dos ratos.

Silva et al. (2008), avaliaram os efeitos da administração neonatal de fluoxetina. Os filhotes de ratos foram pesados diariamente ao longo de 21 dias de lactação, sendo apresentado o gráfico dos pesos médios diários dos grupos comparados. Os autores descrevem que de acordo com as curvas de desenvolvimento ao longo de 21 dias após o nascimento, os animais dos grupos tratados apresentaram uma redução no peso corporal durante o período quando comparados com o grupo controle. Os pesos médios e os erros padrões aos 21 dias também foram apresentados em tabela, sendo comparados por meio de ANOVA e teste de Tukey.

Tousen et al. (2006) avaliaram o efeito da genisteína na gravidez e no desenvolvimento de filhotes por meio de dois experimentos. A progressão da prenhez foi observada medindo diariamente os ganhos de peso maternos e a quantidade de ração consumida. Foram apresentados, em tabela, os pesos médios das ratas no final da gestação e os respectivos desvios padrões, por tratamento, sendo utilizada ANOVA e teste de comparação de pares de médias de Duncan. Os autores concluíram que os pesos corporais das ratas na gestação não foram afetados pela substância. Também foram indicados os ganhos de pesos corporais diários dos filhotes (por ninhada) durante os 22 dias de lactação e semanais (por ninhada e sexo) a partir do desmame até o dia 75, sendo os respectivos valores médios (e intervalos de

variação) por tratamento, apresentados graficamente. As mudanças nos pesos médios dos filhotes ao longo do tempo foram analisadas a partir da ANOVA com medidas repetidas e graus de liberdade ajustados segundo Huynh e Feldt (1976). A substância também não provocou mudanças nos pesos corporais das ninhadas.

Šlamberová, Pometlová e Charousová (2006) avaliaram o desenvolvimento de filhotes de ratas expostas à droga ilícita metanfetamina no período pré-natal. Os filhotes foram pesados durante a lactação até 22 dias, sendo analisados somente o peso ao nascer e o ganho de peso no período por meio de ANOVA com dois fatores (dose e sexo) e do teste de comparação de pares de médias de Bonferroni. Com relação aos pesos ao nascer é apresentada uma tabela de dupla entrada com os valores médios e respectivos erros padrões para as combinações dos níveis de sexo e dose e também os valores da estatística F e os níveis descritivos associados a dose e sexo, não sendo comentada a interação dos fatores. Os autores concluíram que independente do sexo, o peso ao nascer de filhotes de mães que receberam a dose da droga foi menor comparado com os outros grupos (controle e solução salina) e que independente da dose, os machos foram mais pesados do que as fêmeas. Referente ao ganho de peso, não houve diferença significativa entre os grupos com relação às doses, entretanto, os filhotes machos ganharam mais peso durante a lactação do que as fêmeas.

Soubhia et al. (2005) avaliaram os efeitos nefrotóxicos da associação do tacrolimus e anti-inflamatórios não-hormonais em ratos. No experimento os animais foram pesados diariamente ao longo de 7 dias. Os ganhos de pesos médios no período, referentes aos grupos controle e tratados foram comparados usando ANOVA e teste de comparação de pares de médias de Student-Newman-Keuls.

Villarino et al. (2005) avaliaram alterações no crescimento de filhotes de ratas diabéticas. Os filhotes foram pesados diariamente ao longo de 21 dias de lactação, sendo apresentado o gráfico dos perfis médios de crescimento dos grupos controle e tratado. Os pesos médios corporais dos filhotes nos dois grupos foram comparados ao nascer, aos 7 e aos 21 dias usando ANOVA's independentes. Os autores concluíram que os pesos no parto nos dois grupos foram similares, enquanto que no desmame os filhotes de ratas diabéticas apresentaram pesos menores, sendo as diferenças altamente significativas também aos 7 dias e ainda mais pronunciadas aos 21 dias.

Vidinho et al. (2004) analisaram as consequências da ingestão perinatal de sal em ratos. Foram apresentados os gráficos de pesos semanais médios das ratas por

grupo, desde a 3^a. semana até o final da gestação e de pesos médios dos filhotes por grupo separados por sexo, ao longo de 21 dias de lactação (medidos 3 vezes em cada semana). Os efeitos do tempo e do sal sobre o peso foi avaliado usando ANOVA com dois fatores e teste de Bonferroni. Os autores concluíram que não houve diferenças entre os grupos maternos durante todo o período de acompanhamento. Nos gráficos de pesos dos filhotes foram indicados com asterisco os instantes em que houve diferença entre um dos grupos com relação aos demais, sendo as idades com diferenças significativas também citadas na discussão.

Abizaid et al. (2001) avaliaram os efeitos de 2-desoxi-D-glicose (2DG) isolado e combinado com mercaptoacetato no ciclo reprodutivo de ratas, por meio de 3 experimentos. As doses foram administradas ao longo de 2 dias, sendo usado como *baseline* o dia do início do tratamento. Os pesos diários das ratas e dos filhotes foram coletados nos períodos de gestação e lactação, sendo apresentados os gráficos com as alterações (ganho ou perda) diárias médias dos pesos. Os autores usam modelos mistos de análise de variância com dois fatores (tratamento e dia) na análise desses dados, mas não citam nenhuma referência com relação à abordagem da análise estatística realizada.

2.2.2 Considerações sobre as análises estatísticas de pesos diários de ratos dos artigos citados

Nesta seção, são discutidas algumas das vantagens e desvantagens das técnicas estatísticas, apresentadas na seção 2.2.1, usadas na análise dos pesos de ratos coletados ao longo do tempo. Vale destacar que a análise de dados longitudinais requer metodologias estatísticas que levem em conta as dependências entre as observações feitas no mesmo indivíduo ou em grupos de indivíduos e que possibilitem obter informações sobre as alterações na variável resposta ao longo do tempo e avaliar os efeitos de fatores de tratamento ou covariáveis sobre as respostas.

O uso da ANOVA com um único fator de tratamento (dose), seguido por algum teste de comparação de pares de médias de pesos em cada uma das idades selecionadas ou de ganhos médios de pesos em determinado período, descarta uma grande quantidade de informações coletadas e não leva em conta a estrutura de dependência entre os pesos no mesmo animal.

Alguns autores procuraram usar análises que agregam o tempo como fator,

ou seja, como variável qualitativa (com níveis ordenados), mas utilizando a ANOVA com dois fatores (dose e tempo) (*two way factorial*), seguida por algum teste de comparação de médias duas-a-duas. A ANOVA com dois fatores permite avaliar os efeitos principais de cada fator e também o efeito da interação dos níveis desses fatores sobre a variável resposta. Esse modelo é apropriado na análise de delineamentos completamente aleatorizados (ou aleatorizados em blocos ou quadrados latinos) com os tratamentos em esquema fatorial, ou seja, quando as unidades experimentais são atribuídas aleatoriamente às combinações dos níveis dos dois fatores. O modelo assume normalidade, independência e variância constante para os erros. Dessa forma, considerar o fator longitudinal como um fator de tratamento é inadequado, visto que existe uma dependência entre os pesos ao longo dos níveis do fator tempo e uma variabilidade não constante no caso de ratos em crescimento. A ANOVA com dois fatores pode ser adequada, por exemplo, para avaliar os efeitos de drogas e doses nos pesos das ratas ao final da gestação. Nesses casos, é indicado avaliar primeiramente o efeito da interação dos níveis dos dois fatores. No caso de interação significativa deve-se prosseguir com os desdobramentos, analisando os efeitos de um fator sobre os níveis do outro fator. A significância dos efeitos principais, ou seja, de cada fator isolado, deve ser analisada apenas quando a interação não resultar significativa. Quando o fator for qualitativo a análise deve prosseguir com as comparações de médias duas-a-duas e quando for quantitativo, com um estudo de regressão. A análise de um experimento fatorial é apresentada com detalhes em Steel e Torrie (1980), Pimentel-Gomes (2009), Barbin (2003), Montgomery (2008), dentre outros.

Alguns autores, por sua vez, consideraram o tempo como um fator longitudinal, utilizando a ANOVA com 2 fatores, um de tratamento (dose) e outro longitudinal (tempo). Essa técnica é conhecida como análise univariada de perfis (medidas repetidas), associados a experimentos em parcelas subdivididas no tempo (STEEL; TORRIE, 1980; PIMENTEL-GOMES, 2009; BARBIN, 2003; MONTGOMERY, 2008). Nesses experimentos a casualização é feita em dois níveis (parcelas e subparcelas) resultando em duas fontes de variação aleatória (erros associados às parcelas e às subparcelas) e na correlação suposta constante entre as medidas feitas nas unidades de uma mesma parcela. O modelo também assume normalidade e variâncias constantes dos erros nas parcelas e subparcelas e uma correlação constante entre os pares de observações nos diferentes tempos dentro da mesma parcela, não importando se os pesos foram observados em tempos próximos ou distantes,

nem tampouco em idades iniciais ou avançadas. No caso, por exemplo, dos pesos dos filhotes as correlações entre os pesos observados em tempos mais distantes são supostamente menores do que as obtidas em intervalos de tempo mais próximos e as variabilidades aumentam com o crescimento, sendo violadas as suposições do modelo. Esse modelo foi uma das poucas opções disponíveis para análise de dados longitudinais até a década de 70. Diante da violação das suposições do modelo, a análise univariada de perfis dos pesos de ratos pode levar a resultados enganosos.

Uma outra análise, apresentada por poucos autores, também considerando o tempo como um fator longitudinal é a análise multivariada de perfis, que utiliza o modelo associado à ANOVA multivariada (MANOVA - *multivariate analysis of variance*). Nesse caso, as medidas feitas em cada ocasião são entendidas como um vetor de respostas. A técnica se aplica a delineamentos balanceados com relação ao tempo, e no caso da ausência de alguma observação o vetor de respostas precisa ser excluído por inteiro. As suposições assumidas são independência entre as unidades experimentais, erros normais multivariados, e uma matriz de variâncias e covariâncias não estruturada, apresentada em (13), igual para todos os indivíduos. Dependendo da quantidade de ocasiões, o número de parâmetros pode ser tornar bastante grande. A análise não permite, por exemplo, o estudo das tendências nos perfis médios ao longo do tempo, assim como o uso de estruturas de covariâncias mais simples ou diferentes para grupos de indivíduos. Devido à suposição de independência entre as unidades experimentais, também não permite a análise de dados hierárquicos. Assim sendo, no caso de estudos longitudinais dos pesos de filhotes aninhados às ratas, a análise multivariada de perfis deve ser feita com os perfis médios de pesos das ninhadas e não com os perfis individuais de pesos dos filhotes.

Com relação ao resumo dos dados coletados e resultados obtidos, percebeu-se uma preferência pelo uso de tabelas, ainda que o gráfico com os valores médios e respectivos erros padrões seja mais informativo, pois mostra a evolução das respostas ao longo do tempo.

2.3 Modelos lineares mistos para dados longitudinais

Os modelos mistos (*mixed models*) ou modelo de efeitos mistos (*mixed effects models*) aparecem na literatura sob uma grande variedade de títulos. Por exemplo, nas ciências sociais e comportamentais, como educação e psicologia, eles são referidos como modelos lineares hierárquicos (*hierarchical linear models*) (RAUDENBUSH; BRYK, 2002, dentre ou-

tros) ou modelos multinível (*multilevel models*) (GOLDSTEIN, 2003, dentre outros) devido à natureza intrínseca dos níveis das unidades observadas: alunos dentro de classes, classes dentro de escolas, escolas dentro de municípios etc. Alguns autores utilizam também os termos: modelos de efeitos aleatórios (*random effects models*), modelos aleatórios (*random models*), ou modelos de componentes de variância (*variance components models*).

De acordo com Fitzmaurice et al (2009), as primeiras idéias com relação ao modelo linear misto foram propostas na astronomia, por Airy em 1861. Entretanto, o marco inicial das aplicações foi o artigo de Laird e Ware (1982).

A crescente popularidade dos modelos mistos é explicada, principalmente: pela flexibilidade que oferecem para analisar dados agrupados (dados longitudinais, medidas repetidas, delineamentos em blocos e dados em multiníveis); por apresentarem uma estrutura unificada para dados balanceados ou não-balanceados e pela disponibilidade de programas computacionais seguros e eficientes para ajustá-los (PINHEIRO; BATES, 2000).

2.3.1 Especificação do modelo para dados longitudinais com dois níveis

A principal condição dos modelos lineares de efeitos mistos é que algum subconjunto de parâmetros de regressão varie aleatoriamente de um indivíduo para outro, de modo que os indivíduos na população tenham seus próprios perfis de respostas individuais ao longo do tempo, o que faz com que as fontes naturais de heterogeneidade na população sejam levadas em conta. Assim sendo, a característica que distingue o modelo linear misto é que a resposta média é modelada como uma combinação de características da população compartilhadas com todos os indivíduos (efeitos fixos) e efeitos específicos que são próprios de cada indivíduo (efeitos aleatórios). O termo misto é usado para denotar que o modelo contém ambos efeitos fixos e aleatórios (FITZMAURICE; LAIRD; WARE, 2011).

A seguir o modelo linear misto é especificado para o vetor de respostas longitudinais de uma unidade amostral (indivíduo), de acordo com a formulação de Laird e Ware (1982), na qual os modelos condicional e marginal são definidos.

Modelo condicional

Seja $\mathbf{y}_i$, o vetor de dimensão $(n_i \times 1)$, contendo as n_i respostas ao longo do tempo, da unidade amostral i , que correspondem aos valores de um vetor aleatório observável $\mathbf{Y}_i$, $i = 1, \dots, m$, sendo m o número total de unidades e $n = \sum_{i=1}^m n_i$, o número

total de dados disponíveis no estudo, satisfazendo:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}_i \quad (1)$$

sendo $\boldsymbol{\beta}$, um vetor $(p \times 1)$ de efeitos fixos; $\mathbf{X}_i$ uma matriz $(n_i \times p)$ de delineamento específica para a unidade i ; $\mathbf{u}_i$ um vetor $(q \times 1)$ de efeitos aleatórios; $\mathbf{Z}_i$ uma matriz $(n_i \times q)$ de delineamento que vincula $\mathbf{y}_i$ aos efeitos aleatórios e $\mathbf{e}_i$ um vetor $(n_i \times 1)$ de erros aleatórios nas respostas de cada unidade i , sendo, portanto, os valores de um vetor aleatório não observável $\mathbf{E}_i$.

Para a estimação e testes de hipóteses, assume-se que o vetor $\mathbf{E}_i \sim \mathbf{N}(\mathbf{0}, \mathbf{R}_i)$, sendo $\mathbf{R}_i$, a matriz de covariâncias dos erros aleatórios na unidade i , geralmente, parametrizada em termos de um número pequeno de parâmetros de covariâncias. Dessa forma, o perfil médio de respostas ao longo do tempo de cada indivíduo é dado por:

$$\mathbb{E}(\mathbf{Y}_i|\mathbf{u}_i) = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i \quad (2)$$

sendo que $\mathbf{X}_i\boldsymbol{\beta}$ descreve as características que são comuns a população, ou seja, um comportamento médio e $\mathbf{Z}_i\mathbf{u}_i$ descreve as características que são particulares de cada indivíduo, ou seja, os desvios com relação ao comportamento médio.

No caso mais simples assume-se que os componentes do vetor de erros são não correlacionados e com mesma variância, σ^2 , e, portanto, $\mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i}$.

Desse modo,

$$\mathbf{Y}_i|\mathbf{u}_i \sim \mathbf{N}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i, \mathbf{R}_i) \quad (3)$$

uma vez que,

$$\text{Cov}(\mathbf{Y}_i|\mathbf{u}_i) = \text{Cov}(\mathbf{E}_i) = \mathbf{R}_i. \quad (4)$$

De acordo com Fitzmaurice, Laird e Ware (2011), ao assumir algum tipo de estrutura para $\mathbf{R}_i$ diferente de $\sigma^2 \mathbf{I}_{n_i}$, os erros nas medidas repetidas $\mathbf{e}_i$ passam a incluir também um componente devido à má especificação em $\mathbf{Z}_i\mathbf{u}_i$.

Algumas estruturas da matriz $\mathbf{R}_i$, são apresentadas a seguir, sendo a variância comum a todos os tempos denotada por σ^2 , a variância específica em cada tempo t , $t = 1, 2, \dots, n_i$, respectivamente, por σ_i^2 , $i = 1, 2, \dots, n_i$, a correlação constante entre os erros dois a dois, independente do intervalo de tempo entre as medidas repetidas por ρ e a

correlação específica que depende do intervalo de tempo entre as medidas repetidas por ρ_i , $i = 1, 2, \dots, n_i$. Detalhamentos sobre essas estruturas são encontrados, por exemplo, em Littell et al. (2006) e Fitzmaurice, Laird e Ware (2011).

Simetria composta (SC)

$$\mathbf{R}_i = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ & 1 & \rho & \cdots & \rho \\ & & 1 & \cdots & \rho \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & 1 \end{bmatrix}. \quad (5)$$

Auto-regressiva de ordem 1 (AR1)

$$\mathbf{R}_i = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n_i-1} \\ & 1 & \rho^3 & \cdots & \rho^{n_i-2} \\ & & 1 & \cdots & \rho^{n_i-3} \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & 1 \end{bmatrix}. \quad (6)$$

Toeplitz (Toep)

$$\mathbf{R}_i = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n_i-1} \\ & 1 & \rho_1 & \cdots & \rho_{n_i-2} \\ & & 1 & \cdots & \rho_{n_i-3} \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & 1 \end{bmatrix}. \quad (7)$$

Diagonal (D)

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ & \sigma_2^2 & 0 & \cdots & 0 \\ & & \sigma_3^2 & \cdots & 0 \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & \sigma_{n_i}^2 \end{bmatrix}. \quad (8)$$

Simetria composta heterogênea (SCH)

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \cdots & \sigma_1\sigma_{n_i}\rho \\ & \sigma_2^2 & \sigma_2\sigma_3\rho & \cdots & \sigma_2\sigma_{n_i}\rho \\ & & \sigma_3^2 & \cdots & \sigma_3\sigma_{n_i}\rho \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & \sigma_{n_i}^2 \end{bmatrix}. \quad (9)$$

Auto-regressiva de ordem 1, heterogênea (ARH1)

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \cdots & \sigma_1\sigma_{n_i}\rho^{n_i-1} \\ & \sigma_2^2 & \sigma_2\sigma_3\rho & \cdots & \sigma_2\sigma_{n_i}\rho^{n_i-2} \\ & & \sigma_3^2 & \cdots & \sigma_3\sigma_{n_i}\rho^{n_i-3} \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & \sigma_{n_i}^2 \end{bmatrix}. \quad (10)$$

Toeplitz heterogênea (Toeph)

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \cdots & \sigma_1\sigma_{n_i}\rho_{n_i-1} \\ & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \cdots & \sigma_2\sigma_{n_i}\rho_{n_i-2} \\ & & \sigma_3^2 & \cdots & \sigma_3\sigma_{n_i}\rho_{n_i-3} \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & \sigma_{n_i}^2 \end{bmatrix}. \quad (11)$$

Ante-dependência de ordem 1 (Ante1)

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \cdots & \sigma_1\sigma_{n_i}\rho_1\rho_2\rho_3\rho_4\rho_5\rho_6\cdots\rho_{n_i-1} \\ & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \cdots & \sigma_2\sigma_{n_i}\rho_2\rho_3\rho_4\rho_5\rho_6\cdots\rho_{n_i-1} \\ & & \sigma_3^2 & \cdots & \sigma_3\sigma_{n_i}\rho_3\rho_4\rho_5\rho_6\cdots\rho_{n_i-1} \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & \sigma_{n_i}^2 \end{bmatrix}. \quad (12)$$

Não estruturada (NE)

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n_i} \\ & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n_i} \\ & & \sigma_3^2 & \cdots & \sigma_{3n_i} \\ & & & \ddots & \vdots \\ \text{(simétrica)} & & & & \sigma_{n_i}^2 \end{bmatrix}. \quad (13)$$

Embora o modelo linear misto seja um modelo condicional, cujas respostas longitudinais dependem de uma combinação de parâmetros da população (efeitos fixos) e de efeitos específicos dos sujeitos (efeitos aleatórios não observáveis) é direta a obtenção de um modelo para a resposta média marginal (FITZMAURICE; LAIRD; WARE, 2011).

Modelo marginal

Assume-se nesse caso que os efeitos aleatórios $\mathbf{u}_i$ são provenientes de uma distribuição normal com vetor de médias $\mathbf{0}$ e matriz de covariâncias $\mathbf{G}$ ($q \times q$), independentes dos $\mathbf{e}_i$, ou seja,

$$\begin{bmatrix} \mathbf{U}_i \\ \mathbf{E}_i \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_i \end{bmatrix} \right) \quad (14)$$

e portanto, o perfil médio de respostas na população é dado por

$$E(\mathbf{Y}_i) = E[E(\mathbf{Y}_i|\mathbf{u}_i)] = \mathbf{X}_i\boldsymbol{\beta}. \quad (15)$$

Como consequência da inclusão dos efeitos aleatórios, tem-se:

$$\text{Cov}(\mathbf{Y}_i) = E[\text{Cov}(\mathbf{Y}_i|\mathbf{u}_i)] + \text{Cov}[E(\mathbf{Y}_i|\mathbf{u}_i)] = \mathbf{R}_i + \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i' = \mathbf{V}_i \quad (16)$$

e, portanto,

$$\mathbf{Y}_i \sim N(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{R}_i + \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i'). \quad (17)$$

Vale ressaltar que nos modelos marginais os erros representam desvios com relação às respostas médias na população, $\mathbf{X}_i\boldsymbol{\beta}$, enquanto que nos modelos condicionais os erros representam desvios com relação à resposta média específica do indivíduo, $\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i$.

Assim como no modelo de efeitos fixos, $E(\mathbf{Y}|\mathbf{X})$, as respostas médias nos modelos condicional e marginal são condicionais a $\mathbf{X}_i$, sendo essa matriz omitida da notação somente para simplificação (FITZMAURICE; LAIRD; WARE, 2011).

O modelo linear misto (modelo condicional) em (1), pode ser formulado em dois estágios, por meio de modelos mais simples especificados para cada nível de variação aleatória, o que possibilita um melhor entendimento das características principais dos modelos mistos. A especificação em estágios é usada principalmente na educação e ciências sociais (RAUDENBUSH; BRYK, 2002; GOLDESTINE, 2003).

De acordo com Fitzmaurice, Laird e Ware (2011), para propósitos pedagógicos a especificação em dois estágios é bastante útil, mas o leitor deve estar atento pois com essa formalização podem ser introduzidas algumas restrições não necessárias no modelo.

Nas próximas subseções, o modelo linear misto para dados longitudinais é especificado na forma hierárquica em dois estágios (níveis) (2.3.1.1), sendo apresentados alguns de seus casos especiais, em que as respostas médias são modeladas a partir de tendências lineares ou quadráticas (2.3.1.2) e de funções polinomiais por partes com um único ponto de mudança (2.3.1.3). Para simplificar a notação não será feita mais distinção entre uma variável aleatória e os seus possíveis valores, sendo ambos denotados por letras minúsculas.

2.3.1.1 O modelo linear misto na forma hierárquica

A Figura 3 apresenta um esquema para ilustrar a relação entre os coeficientes dos modelos nos dois estágios.

Figura 3 – Esquema das relações entre os coeficientes do modelo linear misto em 2 níveis

No primeiro estágio as mudanças nas respostas de cada indivíduo ao longo do tempo, podem ser descritas, por exemplo, por meio de modelos de regressão polinomial ou *splines*, específicos para cada indivíduo. O modelo no nível 1, pode conter além de variáveis que são função do tempo, como por exemplo, tempo elevado ao quadrado ou ao cubo, outras covariáveis que variam no tempo (*time-varying covariates*), que são incluídas para explicar a variabilidade das medidas feitas num mesmo indivíduo. Por exemplo, num estudo na área de educação, para acompanhar o desenvolvimento do aluno ao longo do ensino fundamental, uma possível covariável para capturar a variação temporal de uma medida de desempenho do aluno poderia ser o número total de faltas que ele teve em cada ano, uma vez que o desempenho do aluno (variável resposta) em um ano pode ser prejudicado em função de um grande número de faltas durante aquele ano. Num estudo toxicológico, por exemplo, poderia ser uma covariável definida pela interação entre o tamanho da ninhada e a idade

dos filhotes, e nesse caso, o efeito da covariável tamanho da ninhada nos pesos dos filhotes (variável resposta) seria diferente para cada idade.

O modelo no nível 1 é especificado a seguir.

Modelo nível 1

$$\begin{aligned}
 y_{ti} &= \beta_{0i} + \beta_{1i}x_{1ti} + \dots + \beta_{Qi}x_{Qti} + e_{ti} \\
 &= \beta_{0i} + \sum_{q=1}^Q \beta_{qi}x_{qti} + e_{ti} \\
 e_{ti} &\sim \mathbf{N}(0, \sigma^2) \text{ e } \text{Cov}(e_{ti}, e_{t'i}) = 0
 \end{aligned} \tag{18}$$

sendo, y_{ti} a resposta observada no indivíduo i , no tempo t , $t = 1, 2, \dots, n_i$ e n_i o número de ocasiões do indivíduo i . No caso de delineamento balanceado com relação ao tempo n_i é igual para todos os indivíduos. As covariáveis, denotadas por x_{qti} , $q = 1, \dots, Q$, com $x_{0ti} = 1$ para todos os tempos, são as mesmas para todos os indivíduos, mudando apenas os valores que elas assumem de acordo com a ocasião e indivíduo. Os coeficientes β_{qi} , $q = 1, \dots, Q$, associados às covariáveis e o intercepto β_{0i} podem ser interpretados como os coeficientes de regressão do indivíduo i que vão definir o perfil específico esperado de cada indivíduo, sendo portanto, as respostas longitudinais em cada indivíduo y_{ti} descritas em função do perfil individual esperado e de uma variação aleatória que corresponde ao erro amostral.

Usando notação matricial, o modelo em (18) é dado por:

$$\begin{aligned}
 \mathbf{y}_i &= \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{e}_i \\
 \mathbf{e}_i &\sim \mathbf{N}(\mathbf{0}, \mathbf{R}_i)
 \end{aligned} \tag{19}$$

$$\text{sendo, } \mathbf{y}_i = \begin{bmatrix} y_{1i} \\ y_{2i} \\ \vdots \\ y_{n_i i} \end{bmatrix} \quad \mathbf{X}_i = \begin{bmatrix} 1 & x_{11i} & x_{21i} & \cdots & x_{Q1i} \\ 1 & x_{12i} & x_{22i} & \cdots & x_{Q2i} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n_i i} & x_{2n_i i} & \cdots & x_{Qn_i i} \end{bmatrix} \quad \boldsymbol{\beta}_i = \begin{bmatrix} \beta_{0i} \\ \beta_{1i} \\ \vdots \\ \beta_{Qi} \end{bmatrix} \quad \mathbf{e}_i = \begin{bmatrix} e_{1i} \\ e_{2i} \\ \vdots \\ e_{n_i i} \end{bmatrix}$$

e $\mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i}$.

Modelo nível 2

No nível 2, assumir que os coeficientes $\beta_{0i}, \dots, \beta_{Qi}$, em (18) são variáveis alea-

tórias, significa dizer que os mesmos são funções dos erros aleatórios no nível 2, denotados por u_{qi} . Esses erros aleatórios correspondem à variação residual em β_{qi} não explicada pelas covariáveis do nível 2, denotadas por w_{si} , $s = 1, \dots, S_q$. Essas covariáveis correspondem às características dos indivíduos (unidades amostrais observadas nesse estágio) como, por exemplo, sexo, tratamento etc. e não precisam ser as mesmas para cada um dos β_q , por isso, o número de covariáveis S_q , varia em função de q .

Considerando todos os componentes do vetor β_i em (18) como aleatórios, o modelo no nível 2 é dado por

$$\begin{aligned}\beta_{qi} &= \gamma_{q0} + \gamma_{q1}w_{1i} + \dots + \gamma_{qS_q}w_{S_qi} + u_{qi} \\ &= \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs}w_{si} + u_{qi} \\ u_{qi} &\sim \mathbf{N}(0, \tau_{qq}) \text{ e } \text{Cov}(u_{qi}, u_{q'i}) = \tau_{qq'}.\end{aligned}\tag{20}$$

Nesse caso, como o nível 2 é o último nível, o intercepto γ_{q0} e os coeficientes das demais covariáveis $\gamma_{q1}, \dots, \gamma_{qS_q}$ são sempre fixos.

Na forma de vetores e matrizes, o modelo em (20) é dado por:

$$\begin{aligned}\beta_i &= \mathbf{W}_i\boldsymbol{\gamma} + \mathbf{u}_i \\ \mathbf{u}_i &\sim \mathbf{N}(\mathbf{0}, \mathbf{G})\end{aligned}\tag{21}$$

¹Usar $\beta_{0i}, \dots, \beta_{Q_i}$ para denotar variáveis aleatórias pode ser considerado um abuso de notação, pois usualmente as letras gregas são usadas para os efeitos fixos, entretanto, nessa especificação em estágios, em que um mesmo efeito é especificado como fixo em um estágio e como aleatório em um outro é usual usar essa notação.

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & \cdots & w_{S_0i} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & w_{1i} & \cdots & w_{S_1i} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & w_{1i} & \cdots & w_{S_Qi} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \vdots \\ \gamma_{0S_0} \\ \gamma_{10} \\ \gamma_{11} \\ \vdots \\ \gamma_{1S_1} \\ \vdots \\ \gamma_{Q0} \\ \gamma_{Q1} \\ \vdots \\ \gamma_{QS_Q} \end{bmatrix}$$

$$\mathbf{u}_i = \begin{bmatrix} u_{0i} \\ u_{1i} \\ \vdots \\ u_{Q_i} \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \sigma_{u_0}^2 & \sigma_{u_0u_1} & \cdots & \sigma_{u_0u_Q} \\ \sigma_{u_0u_1} & \sigma_{u_1}^2 & \cdots & \sigma_{u_1u_Q} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u_0u_Q} & \sigma_{u_0u_{Q-1}} & \cdots & \sigma_{u_Q}^2 \end{bmatrix}.$$

Nesse caso, como os coeficientes $\beta_{0i}, \dots, \beta_{Q_i}$, foram especificados como aleatórios, eles são interpretados como desvios em relação ao(s) valor(es) médio(s) e, portanto, a variação residual entre os indivíduos é expressa pela matriz $\mathbf{G}$, cujos elementos da diagonal representam as variâncias de cada um dos coeficientes especificados como aleatórios e os elementos fora da diagonal as covariâncias entre os coeficientes dois a dois.

Modelo linear misto (combinação dos modelos nos níveis 1 e 2)

A combinação dos modelos definidos nos dois níveis é feita substituindo

$\beta_{0i}, \dots, \beta_{qi}$, definidos em (20) no modelo (18), resultando em:

$$\begin{aligned}
 y_{ti} &= \underbrace{\gamma_{00} + \sum_{s=1}^{S_0} \gamma_{0s} w_{si} + u_{0i}}_{\beta_{0i}} + \underbrace{\left(\gamma_{10} + \sum_{s=1}^{S_1} \gamma_{1s} w_{si} + u_{1i} \right)}_{\beta_{1i}} x_{1ti} + \dots + \underbrace{\left(\gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} w_{si} + u_{qi} \right)}_{\beta_{qi}} x_{qti} + e_{ti} \\
 &= \underbrace{\gamma_{00} + \gamma_{10} x_{1ti} + \dots + \gamma_{q0} x_{qti} + \sum_{s=1}^{S_0} \gamma_{0s} w_{si} + \sum_{s=1}^{S_1} \gamma_{1s} x_{1ti} w_{si} + \dots + \sum_{s=1}^{S_q} \gamma_{qs} x_{qti} w_{si}}_{\text{parte fixa}} + \underbrace{u_{0i} + u_{1i} x_{1ti} + \dots + u_{qi} x_{qti} + e_{ti}}_{\text{parte aleatória}}. \tag{22}
 \end{aligned}$$

Matricialmente, a combinação dos modelos consiste na substituição da expressão β_i em (21) em (19), resultando em

$$\begin{aligned}
 \mathbf{y}_i &= \mathbf{X}_i [\mathbf{W}_i \boldsymbol{\gamma} + \mathbf{u}_i] + \mathbf{e}_i \\
 &= \mathbf{X}_i \mathbf{W}_i \boldsymbol{\gamma} + \mathbf{X}_i \mathbf{u}_i + \mathbf{e}_i. \tag{23}
 \end{aligned}$$

Dependendo das características da população, pode ser mais apropriado reescrever o modelo no nível 2 de modo que alguns dos componentes do vetor β_i em (21) sejam considerados de efeitos fixos. Dessa forma, esses componentes podem assumir ou um valor médio comum a todos os indivíduos, denotado por γ_{q0} em (20), ou um valor médio comum a um subconjunto de indivíduos que apresentam as mesmas características, ou seja, os mesmos valores das covariáveis no nível 2. Se β_{qi} é definido como fixo assume-se que não há variação nesse coeficiente que não seja explicada pelas covariáveis no nível 2.

Ao considerar todos os coeficientes β_{qi} como aleatórios, as covariáveis no nível 1, x_{qti} , são especificadas tanto na parte fixa quanto na parte aleatória em (22). No caso em que somente alguns β_{qi} são considerados como aleatórios no nível 2, somente os x_{qti} associados aos coeficientes aleatórios é que estarão presentes na parte aleatória. Portanto, no caso em que alguns dos coeficientes são de efeitos fixos, devem ser consideradas duas matrizes, uma para a parte fixa, denotada por $\mathbf{X}_i$ e outra para a parte aleatória, denotada por $\mathbf{Z}_i$, formada por um subconjunto das colunas da matriz $\mathbf{X}_i$, sendo portanto, de dimensão $n_i \times p$, com $p < q$, uma vez que a dimensão do vetor $\mathbf{u}_i$ será reduzida $p \times 1$. As colunas da matriz $\mathbf{X}_i$ que não estão em $\mathbf{Z}_i$ correspondem aos componentes do vetor β_i que são fixos. Mais adiante, essas características serão discutidas com mais detalhes usando algumas situações hipotéticas.

Assim sendo, nos casos em que alguns dos coeficientes no nível 2 são especificados como fixos e outros como aleatórios, o modelo em (23) pode ser escrito como:

$$\mathbf{y}_i = \mathbf{X}_i \mathbf{W}_i \boldsymbol{\gamma} + \mathbf{Z}_i \mathbf{u}_i + \mathbf{e}_i. \quad (24)$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & \cdots & w_{S_0i} & x_{11i} & x_{11i}w_{1i} & \cdots & x_{11i}w_{S_1i} & \cdots & x_{Q1i} & x_{Q1i}w_{1i} & \cdots & x_{Q1i}w_{S_Qi} \\ 1 & w_{1i} & \cdots & w_{S_0i} & x_{12i} & x_{12i}w_{1i} & \cdots & x_{12i}w_{S_1i} & \cdots & x_{Q2i} & x_{Q2i}w_{1i} & \cdots & x_{Q2i}w_{S_Qi} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_{1i} & \cdots & w_{S_0i} & x_{1n_ii} & x_{1n_ii}w_{1i} & \cdots & x_{1n_ii}w_{S_1i} & \cdots & x_{Qn_ii} & x_{Qn_ii}w_{1i} & \cdots & x_{Qn_ii}w_{S_Qi} \end{bmatrix}.$$

A matriz $\mathbf{X}_i \mathbf{W}_i$ apresenta colunas que correspondem ao intercepto, às covariáveis nos níveis 1 e 2 e suas interações, sendo portanto, a matriz de delineamento dos efeitos fixos. Se denotarmos essa matriz por $\mathbf{X}_i$ e o vetor $\boldsymbol{\gamma}$ por $\boldsymbol{\beta}$, então o modelo em (24) corresponde ao modelo condicional em (1), exceto por uma restrição com relação à matriz $\mathbf{R}_i$. Na especificação em estágios a matriz $\mathbf{R}_i$ em (18) representa a variabilidade nas respostas ao longo do tempo que não é explicada pelas covariáveis no nível 1, ao passo que no modelo linear misto em (1), a matriz $\mathbf{R}_i$ inclui também a variabilidade não explicada pelas covariáveis no nível 2.

2.3.1.2 Tendências linear e quadrática

Alguns casos especiais de modelos lineares mistos para dados longitudinais são aqueles em que os comportamentos das respostas médias ao longo do tempo são modelados por meio de polinômios. Os modelos lineares mistos de regressão polinomial de ordem baixa, ou seja, de graus 1 ou 2, são bastante úteis quando os perfis populacionais seguem uma tendência linear ou quadrática, uma vez que os respectivos parâmetros apresentam interpretação prática e são facilmente estimados.

Nesta seção, são apresentados alguns exemplos hipotéticos representando situações que são comuns na prática. Os modelos são especificados na forma hierárquica em dois estágios, possibilitando uma apresentação mais detalhada das características dos modelos lineares mistos.

Nos exemplos, são considerados modelos de regressão linear admitindo-se que as mudanças nas respostas ao longo do tempo no nível 1 são modeladas por retas. A extensão para os modelos de regressão quadrática é direta. Foi suposta uma população com 3 grupos de indivíduos, classificados de acordo com os níveis de um tratamento: *controle*, *A* e *B*.

Os gráficos da Figura 4 apresentam as retas médias marginais (tracejadas) e as retas médias condicionais (demais retas), ilustrando algumas situações que resultam de diferentes especificações dos coeficientes no nível 2. Os casos abordados são apresentados na Tabela 1.

Tabela 1 – Casos particulares de modelos com tendência linear

Caso	Descrição
1	Retas com interceptos e inclinações diferentes nos três tratamentos e efeito aleatório nos dois parâmetros.
2	Retas com interceptos e inclinações diferentes nos três tratamentos e efeito aleatório somente no intercepto.
3	Retas com interceptos diferentes e mesma inclinação nos três tratamentos e efeito aleatório somente no intercepto.
4	Retas com interceptos e inclinações diferentes nos três tratamentos e efeito aleatório somente na inclinação.
5	Retas com mesmo intercepto e inclinações diferentes nos três tratamentos e efeito aleatório somente na inclinação.
6	Retas com interceptos e inclinações diferentes nos três tratamentos.
7	Retas com interceptos diferentes e mesma inclinação nos três tratamentos .
8	Retas com mesmo intercepto e inclinações diferentes nos três tratamentos.
9	Retas com mesmo intercepto e mesma inclinação para os três tratamentos.

Em todos os casos as tendências são lineares, sendo o tempo a única covariável (contínua) considerada para explicar a variabilidade das respostas ao longo do tempo. O modelo no nível 1 é apresentado a seguir e corresponde a um caso particular do modelo especificado em (18).

$$y_{ti} = \beta_{0i} + \beta_{1i}x_{ti} + e_{ti} \quad (25)$$

em que os valores de x_{ti} , t , $t = 1, \dots, n_i$, podem ser tempos decorridos após o início do experimento. O número total de medidas repetidas n_i , pode variar de indivíduo para indivíduo, ou seja, não é necessário que os indivíduos sejam observados nas mesmas ocasiões. O coeficiente β_{0i} corresponde ao intercepto da reta de crescimento do indivíduo i , e portanto,

Figura 4 – Gráficos ilustrando alguns casos especiais de modelos mistos de tendência linear

o valor y_{0i} , para $t=0$, é chamado *baseline*. O coeficiente β_{1i} denota a taxa de crescimento do indivíduo i , que é constante em relação ao tempo e corresponde à variação na resposta devido ao aumento de uma unidade no tempo.

A matriz $\mathbf{X}_i$ e o vetor $\boldsymbol{\beta}_i$ no modelo em (19) reduzem-se a

$$\mathbf{X}_i = \begin{bmatrix} 1 & x_{1i} \\ 1 & x_{2i} \\ \vdots & \vdots \\ 1 & x_{n_i i} \end{bmatrix} \quad \boldsymbol{\beta}_i = \begin{bmatrix} \beta_{0i} \\ \beta_{1i} \end{bmatrix}.$$

A seguir, são especificados os modelos no nível 2 para cada um dos casos descritos, sendo discutidas algumas características importantes, entre elas, a variação aleatória em cada um dos níveis. As possíveis covariáveis no nível 2 são as variáveis indicadoras w_{1i} e w_{2i} , que assumem valor igual a 1 no caso do indivíduo i ter sido aleatorizado, respectivamente, aos níveis A e B ou valor igual a 0, caso contrário, sendo a referência, $w_{1i}=w_{2i}=0$, o

grupo *controle*. Em todos os casos assumiu-se $\text{Var}(e_{ti}) = \sigma^2$ e $\text{Cov}(e_{ti}, e_{t'i}) = 0$.

No caso 1 (Tabela 1), o modelo de nível 2 é dado por:

$$\begin{aligned}\beta_{0i} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + u_{0i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i} + u_{1i}.\end{aligned}\quad (26)$$

O modelo linear misto obtido substituindo (26) em (25) é dado por:

$$\begin{aligned}y_{ti} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + u_{0i} + (\gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i} + u_{1i})x_{ti} + e_{ti} \\ &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + u_{0i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}x_{ti}w_{2i} + u_{1i}x_{ti} + e_{ti}.\end{aligned}\quad (27)$$

Da mesma forma as matrizes $\mathbf{W}_i$ e $\mathbf{Z}_i$ e os vetores $\boldsymbol{\gamma}$ e $\mathbf{u}_i$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & w_{1i} & w_{2i} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix} \quad \mathbf{Z}_i = \begin{bmatrix} 1 & x_{1i} \\ 1 & x_{2i} \\ \vdots & \vdots \\ 1 & x_{n_i i} \end{bmatrix} \quad \mathbf{u}_i = \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & x_{1i} & x_{1i}w_{1i} & x_{1i}w_{2i} \\ 1 & w_{1i} & w_{2i} & x_{2i} & x_{2i}w_{1i} & x_{2i}w_{2i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & w_{1i} & w_{2i} & x_{n_i i} & x_{n_i i}w_{1i} & x_{n_i i}w_{2i} \end{bmatrix}.$$

As interações entre tempo e tratamento são consequências do coeficiente β_{1i} ter sido definido no modelo de nível 2 em (26), em função dos níveis de tratamento, descrevendo portanto, as mudanças ao longo do tempo de acordo com o nível do tratamento.

O fato de especificar os coeficientes β_{0i} e β_{1i} no nível 2 como funções das variáveis indicadoras dos níveis do tratamento, faz com que os interceptos e as inclinações das retas médias sejam diferentes para cada nível de tratamento. Nesse caso, o intercepto da reta média do grupo *controle* é dado por γ_{00} , do grupo *A* por $(\gamma_{00} + \gamma_{01})$ e do grupo *B* por $(\gamma_{00} + \gamma_{02})$, sendo portanto, γ_{01} a diferença entre os interceptos das retas médias dos grupos

controle e A e γ_{02} a diferença entre os interceptos das retas médias dos grupos *controle* e B . Da mesma forma, γ_{10} , representa a inclinação da reta média do grupo *controle*, ou seja, a taxa de crescimento ao longo do tempo do grupo *controle*, $(\gamma_{10} + \gamma_{11})$, a inclinação da reta média do grupo A e $(\gamma_{10} + \gamma_{12})$ a inclinação da reta média do grupo B , sendo portanto, γ_{11} a diferença entre as inclinações das retas médias dos grupos *controle* e A e γ_{12} a diferença entre as inclinações das retas médias dos grupos *controle* e B .

Uma vez que os coeficientes β_{0i} e β_{1i} são aleatórios eles também são funções dos erros aleatórios no nível 2, respectivamente, u_{0i} e u_{1i} . As variáveis aleatórias u_{0i} e u_{1i} representam as diferenças entre os coeficientes de regressão (intercepto e inclinação) do perfil individual i com relação aos coeficientes da reta média obtida levando em conta o grupo, ou seja, o desvio específico do indivíduo i com relação ao perfil médio correspondente. Assim sendo, no modelo misto em (27) as retas individuais apresentam interceptos e inclinações diferentes e variam em torno da reta média referente ao seu respectivo grupo.

Assumindo que $\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{G})$, tem-se que:

$$\mathbf{G} = \begin{bmatrix} \sigma_{u_0}^2 & \sigma_{u_0 u_1} \\ \sigma_{u_0 u_1} & \sigma_{u_1}^2 \end{bmatrix},$$

sendo $\mathbf{G}$, a matriz de variâncias e covariâncias dos efeitos aleatórios u_{0i} e u_{1i} , que descreve a variação dos coeficientes das retas individuais residuais. Assim sendo, os elementos da diagonal $\sigma_{u_0}^2$, $\sigma_{u_1}^2$ representam, respectivamente, a variância dos interceptos e das inclinações das retas individuais residuais e $\sigma_{u_0 u_1}$ a covariância entre os interceptos e as inclinações.

$$\begin{aligned} \text{Var}(Y_{ti}) &= \text{Var}(u_{0i} + u_{1i}x_{1ti} + e_{ti}) = \\ &= \text{Var}(u_{0i}) + 2x_{1ti}\text{Cov}(u_{0i}, u_{1i}) + x_{1ti}^2\text{Var}(u_{1i}) + \text{Var}(e_{ti}) = \\ &= \sigma_{u_0}^2 + 2x_{1ti}\sigma_{u_0, u_1} + x_{1ti}^2\sigma_{u_1}^2 + \sigma^2. \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_{ti}, Y_{t'i}) &= \text{Cov}(u_{0i} + u_{1i}x_{1ti} + e_{ti}, u_{0i} + u_{1i}x_{1t'i} + e_{t'i}) = \\ &= \text{Var}(u_{0i}) + (x_{1ti} + x_{1t'i})\text{Cov}(u_{0i}, u_{1i}) + x_{1ti}x_{1t'i}\text{Var}(u_{1i}) = \\ &= \sigma_{u_0}^2 + (x_{1ti} + x_{1t'i})\sigma_{u_0, u_1} + x_{1ti}x_{1t'i}\sigma_{u_1}^2. \end{aligned}$$

Usando notação matricial,

$$\mathbf{V}_i = \text{Cov}(\mathbf{Y}_i) = \text{Var}(\mathbf{Z}_i\mathbf{u}_i + \mathbf{e}_i) = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i' + \mathbf{R}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i' + \sigma^2\mathbf{I}_{n_i}$$

sendo $\mathbf{V}_i$ a matriz de variâncias e covariâncias de $\mathbf{Y}_i$. Dessa forma, uma característica importante do modelo linear misto é que a matriz $\mathbf{V}_i$ é composta por parâmetros de covariâncias das matrizes $\mathbf{G}$ e $\mathbf{R}_i$, especificados de forma explícita, o que possibilita analisar separadamente as fontes de variação entre indivíduos ($\mathbf{G}$) e nas respostas dentro de um mesmo indivíduo ($\mathbf{R}_i$).

Uma vez que $\mathbf{V}_i$ é função de $\mathbf{G}$, os seus componentes dependem de como são especificados os coeficientes no nível 2. Neste caso, por exemplo, mesmo tendo escolhido no nível 1 a estrutura mais simples para a matriz de covariâncias dos erros, com e_{ti} independentes e variância constante igual a σ^2 em todos os tempos, a inclusão dos efeitos aleatórios no intercepto e inclinação, resultou em uma estrutura final com variâncias e correlações nas variáveis respostas que dependem do tempo.

A partir dos elementos da matriz $\mathbf{G}$, também é possível obter a correlação entre os interceptos e a inclinações, que é uma característica importante nos estudos de crescimento. Por exemplo, se a correlação for forte e positiva, implica que valores baixos no *baseline* estão associados a inclinações menores; já valores altos no *baseline* estão associados a inclinações maiores. Essa correlação é dada por:

$$\rho(\beta_{0i}, \beta_{1i}) = \frac{\sigma_{u_0 u_1}}{\sqrt{\sigma_{u_0}^2 \sigma_{u_1}^2}}.$$

Para o caso 2 (Tabela 1) o modelo de nível 2 é dado por:

$$\begin{aligned}\beta_{0i} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + u_{0i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i}.\end{aligned}\tag{28}$$

Substituindo (28) em (25) obtém-se o modelo linear misto dado por:

$$\begin{aligned}y_{ti} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + (\gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i})x_{ti} + u_{0i} + e_{ti} \\ &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}x_{ti}w_{2i} + u_{0i} + e_{ti}.\end{aligned}\tag{29}$$

Da mesma forma, a matriz $\mathbf{W}_i$ e os vetores $\mathbf{Z}_i$, $\boldsymbol{\gamma}$ e $\mathbf{u}_i$ do modelo em (21)

reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & w_{1i} & w_{2i} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix} \quad \mathbf{Z}_i = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{u}_i = \begin{bmatrix} u_{0i} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & x_{1i} & x_{1i}w_{1i} & x_{1i}w_{2i} \\ 1 & w_{1i} & w_{2i} & x_{2i} & x_{2i}w_{1i} & x_{2i}w_{2i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & w_{1i} & w_{2i} & x_{n_i i} & x_{n_i i}w_{1i} & x_{n_i i}w_{2i} \end{bmatrix}.$$

Assim como no caso 1, as retas médias referentes a cada grupo diferem entre si, apresentando tanto interceptos como inclinações diferentes, cujos valores são os mesmos detalhados no caso 1. O que muda é que foi associado um efeito aleatório apenas ao coeficiente β_{0i} , e portanto, apenas os interceptos individuais variam em relação aos interceptos das retas médias dos respectivos grupos. Nesse caso, os coeficientes β_{1i} são iguais para todos os indivíduos do mesmo grupo.

No modelo misto em (29), as retas individuais apresentam interceptos diferentes e as inclinações variam somente entre os grupos, sendo portanto, as retas individuais paralelas às retas médias dentro de cada grupo.

Assumindo que os erros aleatórios com relação ao intercepto, $u_{0i} \sim \mathbf{N}(0, \sigma_{u_0}^2)$, tem-se:

$$\begin{aligned} \text{Var}(Y_{ti}) &= \text{Var}(u_{0i} + e_{ti}) = \\ &= \text{Var}(u_{0i}) + \text{Var}(e_{ti}) = \\ &= \sigma_{u_0}^2 + \sigma^2. \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_{ti}, Y_{t'i'}) &= \text{Cov}(u_{0i} + e_{ti}, u_{0i} + e_{t'i'}) = \\ &= \text{Var}(u_{0i}) = \\ &= \sigma_{u_0}^2. \end{aligned}$$

Portanto, a matriz de variâncias e covariâncias de Y_i apresenta uma estrutura de simetria composta, sendo os elementos da diagonal iguais a $\sigma_{u_0}^2 + \sigma^2$ e os elementos fora da diagonal iguais a $\sigma_{u_0}^2$.

O modelo de nível 2 para o caso 3 (Tabela 1) é dado por:

$$\begin{aligned}\beta_{0i} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + u_{0i} \\ \beta_{1i} &= \gamma_{10}.\end{aligned}\tag{30}$$

Dessa forma, os interceptos individuais são funções de covariáveis que especificam os grupos e do erro aleatório específico de cada indivíduo, u_{0i} , assumindo, portanto, valores diferentes com relação aos indivíduos. As inclinações individuais são iguais a uma constante, dada por, γ_{10} , de modo que as mudanças ao longo do tempo são consideradas iguais para todos os indivíduos.

O modelo linear misto obtido substituindo (30) em (25) é dado por:

$$y_{ti} = \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + \gamma_{10}x_{ti} + u_{0i} + e_{ti}.\tag{31}$$

No modelo misto em (31), não há interação entre tratamento e tempo, uma vez que o coeficiente β_{1i} foi definido no modelo de nível 2 em (30), como sendo igual a uma constante.

Assim sendo, as retas que descrevem o crescimento médio com relação aos grupos (*controle*, *A* e *B*) têm interceptos diferentes entre si e iguais a γ_{00} , $(\gamma_{00} + \gamma_{01})$ e $(\gamma_{00} + \gamma_{02})$ e mesma inclinação γ_{10} , sendo as retas individuais paralelas às respectivas retas médias.

Da mesma forma, a matriz $\mathbf{W}_i$ e os vetores $\mathbf{Z}_i$, $\boldsymbol{\gamma}$ e $\mathbf{u}_i$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \end{bmatrix} \quad \mathbf{Z}_i = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{u}_i = \begin{bmatrix} u_{0i} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$, resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & x_{1i} \\ 1 & w_{1i} & w_{2i} & x_{2i} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & w_{1i} & w_{2i} & x_{n_i i} \end{bmatrix}.$$

Assumindo que os erros aleatórios com relação ao intercepto, $u_{0i} \sim \mathbf{N}(0, \sigma_{u_0}^2)$, em que $\sigma_{u_0}^2$ é a variância dos interceptos das retas individuais na população, a estrutura da matriz de variâncias e covariâncias de Y_i passa a ser de simetria composta, igual a do caso 2.

No caso 4 (Tabela 1), o modelo de nível 2 é dado por:

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i} + u_{1i}. \end{aligned} \quad (32)$$

O modelo linear misto obtido substituindo (32) em (25) é dado por:

$$y_{ti} = \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}x_{ti}w_{2i} + u_{1i} + e_{ti}. \quad (33)$$

Assim como nos casos 1 e 2, as retas médias referentes a cada grupo diferem entre si, apresentando tanto interceptos como inclinações diferentes, cujos valores são os mesmos detalhados no caso 1. O que muda com relação aos casos 1 e 2 é que se associou um efeito aleatório exclusivamente a β_{1i} , e portanto, apenas as inclinações individuais variam em relação às inclinações das retas médias dos respectivos grupos. Nesse caso, os coeficientes β_{0i} são iguais para todas as unidades experimentais de cada grupo.

No modelo linear misto em (33) as retas individuais apresentam inclinações diferentes mas com interceptos que variam somente com relação aos grupos.

A matriz $\mathbf{W}_i$ e os vetores $\mathbf{Z}_i$, $\boldsymbol{\gamma}$ e $\mathbf{u}_i$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & w_{1i} & w_{2i} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix} \quad \mathbf{Z}_i = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{n_i i} \end{bmatrix} \quad \mathbf{u}_i = \begin{bmatrix} u_{1i} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & x_{1i} & x_{1i}w_{1i} & x_{1i}w_{2i} \\ 1 & w_{1i} & w_{2i} & x_{2i} & x_{2i}w_{1i} & x_{2i}w_{2i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & w_{1i} & w_{2i} & x_{n_i i} & x_{n_i i}w_{1i} & x_{n_i i}w_{2i} \end{bmatrix}.$$

Assumindo que $u_{1i} \sim \mathbf{N}(0, \sigma_{u_1}^2)$, sendo $\sigma_{u_1}^2$, tem-se:

$$\begin{aligned} \text{Var}(Y_{ti}) &= \text{Var}(u_{1i} + e_{ti}) = \\ &= \text{Var}(u_{1i}) + \text{Var}(e_{ti}) = \\ &= \sigma_{u_1}^2 + \sigma^2. \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_{ti}, Y_{t'i'}) &= \text{Cov}(u_{1i} + e_{ti}, u_{1i} + e_{t'i'}) = \\ &= \text{Var}(u_{1i}) = \\ &= \sigma_{u_1}^2. \end{aligned}$$

Nesse caso, a matriz de variâncias e covariâncias de Y_i também apresenta uma estrutura de simetria composta, assim como nos casos 2 e 3, mas com elementos da diagonal iguais a $\sigma_{u_1}^2 + \sigma^2$ e elementos fora da diagonal iguais a $\sigma_{u_1}^2$.

Para o caso 5 (Tabela 1) o modelo de nível 2 é dado por:

$$\begin{aligned} \beta_{0i} &= \gamma_{00} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i} + u_{1i}. \end{aligned} \tag{34}$$

Dessa forma, os interceptos individuais são iguais a uma constante, dada por γ_{00} , sendo portanto, iguais para todos os indivíduos. As inclinações são funções das covariáveis que especificam os grupos e do erro aleatório específico de cada indivíduo, u_{1i} .

Substituindo (34) em (25) tem-se o modelo linear misto:

$$y_{ti} = \gamma_{00} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}x_{ti}w_{2i} + u_{1i} + e_{ti}. \tag{35}$$

No modelo 35, há interação entre tratamento e tempo, em consequência do coeficiente β_{1i} ter sido definido no modelo de nível 2 em (34), em função dos níveis de tratamento. Assim sendo, as retas individuais apresentam o mesmo intercepto, γ_{00} , e inclinações

que variam aleatoriamente em relação às inclinações médias de cada grupo, dadas por γ_{10} (*controle*), $\gamma_{10} + \gamma_{11}$ (*A*) e $\gamma_{10} + \gamma_{12}$ (*B*).

Da mesma forma, a matriz $\mathbf{W}_i$ e os vetores $\mathbf{Z}_i$, $\boldsymbol{\gamma}$ e $\mathbf{u}_i$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & w_{1i} & w_{2i} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix} \quad \mathbf{Z}_i = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{n_i i} \end{bmatrix} \quad \mathbf{u}_i = \begin{bmatrix} u_{1i} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & x_{1i} & x_{1i}w_{1i} & x_{1i}w_{2i} \\ 1 & x_{2i} & x_{2i}w_{1i} & x_{2i}w_{2i} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n_i i} & x_{n_i i}w_{1i} & x_{n_i i}w_{2i} \end{bmatrix}.$$

Assumindo que $u_{1i} \sim \mathbf{N}(0, \sigma_{u_1}^2)$, sendo $\sigma_{u_1}^2$ a variância das inclinações das retas individuais na população, a matriz de variâncias e covariâncias de $\mathbf{Z}_i$ apresenta uma estrutura de simetria composta igual a do caso 4.

Os modelos detalhados, a seguir, que correspondem aos casos 6 a 9, são de efeitos fixos e portanto, casos particulares dos modelos lineares mistos. De acordo com Fitzmaurice, Laird e Ware (2011) um modelo que apresenta somente coeficientes de efeitos fixos, mesmo assumindo uma estrutura da matriz $\mathbf{R}_i$ mais complexa, como as citadas na seção 2.3.1, pertence à classe de modelos de efeitos fixos, uma vez que a pressuposição básica do modelo linear misto é que algum subconjunto de parâmetros de regressão varie aleatoriamente de um indivíduo para outro, levando em conta as fontes de heterogeneidade natural na população.

O modelo de nível 2 para o caso 6 (Tabela 1), é dado por:

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i}. \end{aligned} \tag{36}$$

Substituindo (36) em (25) tem-se o seguinte modelo linear de efeitos fixos:

$$y_{ti} = \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{02}w_{2i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}x_{ti}w_{2i} + e_{ti}. \tag{37}$$

Da mesma forma, a matriz $\mathbf{W}_i$ e o vetor $\boldsymbol{\gamma}$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & w_{1i} & w_{2i} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & x_{1i} & x_{1i}w_{1i} & x_{1i}w_{2i} \\ 1 & w_{1i} & w_{2i} & x_{2i} & x_{2i}w_{1i} & x_{2i}w_{2i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & w_{1i} & w_{2i} & x_{n_i i} & x_{n_i i}w_{1i} & x_{n_i i}w_{2i} \end{bmatrix}.$$

O modelo (37) indica retas com interceptos e inclinações diferentes somente entre os grupos. Neste caso, como os dois coeficientes no nível 2 são considerados fixos, assume-se que não existem variações aleatórias significativas entre os indivíduos, resultando em uma matriz de variâncias e covariâncias de Y_i diagonal e com elementos iguais a σ^2 .

O modelo de nível 2 do caso 7 (Tabela 1) é dado por:

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01}w_{0i} + \gamma_{02}w_{2i} \\ \beta_{1i} &= \gamma_{10}. \end{aligned} \tag{38}$$

Substituindo (38) em (25) tem-se o modelo linear de efeitos fixos expresso como:

$$y_{ti} = \gamma_{00} + \gamma_{01}w_{0i} + \gamma_{02}w_{2i} + \gamma_{10}x_{ti} + e_{ti}. \tag{39}$$

Da mesma forma, a matriz $\mathbf{W}_i$ e o vetor $\boldsymbol{\gamma}$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & w_{1i} & w_{2i} & x_{1i} \\ 1 & w_{1i} & w_{2i} & x_{2i} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & w_{1i} & w_{2i} & x_{n_i i} \end{bmatrix}.$$

O modelo de efeitos fixos apresentado em (39) resulta em retas com interceptos diferentes para cada grupo, dados por γ_{00} (*controle*), $\gamma_{00} + \gamma_{01}$ (*A*) e $\gamma_{00} + \gamma_{02}$ (*B*) e uma única inclinação igual a γ_{10} , ou seja, retas médias paralelas para os três grupos.

O modelo de nível 2 para o caso 8 (Tabela 1) é expresso por:

$$\begin{aligned} \beta_{0i} &= \gamma_{00} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}w_{1i} + \gamma_{12}w_{2i}. \end{aligned} \quad (40)$$

Substituindo (40) em (25), obtém-se o modelo linear de efeitos fixos:

$$y_{ti} = \gamma_{00} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}x_{ti}w_{2i} + e_{ti}. \quad (41)$$

Da mesma forma, a matriz $\mathbf{W}_i$ e o vetor $\boldsymbol{\gamma}$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & w_{1i} & w_{2i} \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & x_{1i} & x_{1i}w_{1i} & x_{1i}w_{2i} \\ 1 & x_{2i} & x_{2i}w_{1i} & x_{2i}w_{2i} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n_i i} & x_{n_i i}w_{1i} & x_{n_i i}w_{2i} \end{bmatrix}.$$

O modelo de efeitos fixos apresentado em (41) resulta em retas com um intercepto comum igual a γ_{00} e com inclinações diferentes para cada grupo, dadas por γ_{10} (*controle*), $\gamma_{10} + \gamma_{11}$ (*A*) e $\gamma_{10} + \gamma_{12}$ (*B*).

Por último, o modelo de nível 2 para o caso 9 (Tabela 1) é dado por:

$$\begin{aligned}\beta_{0i} &= \gamma_{00} \\ \beta_{1i} &= \gamma_{10}.\end{aligned}\tag{42}$$

Substituindo (42) em (25), obtém-se o modelo linear de efeitos fixos mais simples:

$$y_{ti} = \gamma_{00} + \gamma_{10}x_{ti} + e_{ti}.\tag{43}$$

Da mesma forma, a matriz $\mathbf{W}_i$ e o vetor $\boldsymbol{\gamma}$ do modelo em (21) reduzem-se a

$$\mathbf{W}_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \end{bmatrix}.$$

O produto de $\mathbf{X}_i$ por $\mathbf{W}_i$ resulta em:

$$\mathbf{X}_i \mathbf{W}_i = \begin{bmatrix} 1 & x_{1i} \\ 1 & x_{2i} \\ \vdots & \vdots \\ 1 & x_{n_i i} \end{bmatrix}.$$

O modelo de efeitos fixos apresentado em (43) resulta em uma única reta, comum aos 3 grupos.

Com relação ao modelo linear misto de regressão quadrática, o modelo do nível 1 em (18) reduz-se à

$$y_{ti} = \beta_{0i} + \beta_{1i}x_{ti} + \beta_{2i}x_{ti}^2 + e_{ti}\tag{44}$$

sendo x_{ti} , $t = 1, \dots, n_i$, os tempos ou idades em que os indivíduos foram observados. Da mesma forma que no modelo misto de crescimento linear o número total de tempos n_i , pode variar de indivíduo para indivíduo. Neste caso, o intercepto β_{0i} , representa o valor da variável resposta do indivíduo i no tempo 0. O coeficiente β_{1i} , corresponde à inclinação (coeficiente angular) da reta tangente ao perfil do indivíduo i no tempo 0, podendo ser interpretado como a taxa de crescimento instantânea no tempo 0 e β_{2i} corresponde à curvatura da parábola ou à aceleração do processo de crescimento do indivíduo i .

Usando notação matricial, a matriz $\mathbf{X}_i$ e o vetor $\boldsymbol{\beta}_i$, descritos em (19), reduzem-se a

$$\mathbf{X}_i = \begin{bmatrix} 1 & x_{1i} & x_{1i}^2 \\ 1 & x_{2i} & x_{2i}^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{n_i i} & x_{n_i i}^2 \end{bmatrix} \quad \boldsymbol{\beta}_i = \begin{bmatrix} \beta_{0i} \\ \beta_{1i} \\ \beta_{2i} \end{bmatrix}.$$

Assim como no modelo misto de regressão linear, no modelo de regressão quadrática os coeficientes do nível 1, β_{0i} , β_{1i} e β_{2i} , são definidos no nível 2 como fixos ou aleatórios, resultando em vários casos particulares, que não serão detalhados neste trabalho.

2.3.1.3 *Splines*

Em algumas aplicações as tendências longitudinais das respostas médias podem ter comportamento não linear, não sendo possível representá-las por polinômios de 1º ou 2º grau ou até mesmo por polinômios de ordem maiores. Um exemplo é quando a resposta média aumenta (diminui) mais rapidamente em um sub-intervalo do que em outro (ou vice-versa). Nessas situações, pode ser mais apropriado modelar as tendências por meio de *splines*.

Um *spline* de grau k é definido com uma função por partes, constituída de polinômios de grau k que se unem em pontos de mudanças chamados nós. Tanto a função como as suas derivadas de ordem até $(k - 1)$ são contínuas nos nós, garantindo a suavidade da curva quando o grau do *spline* for maior do que 1. Os polinômios mais usados são os de 1º, 2º ou 3º grau.

A Figura 5 ilustra algumas situações em que as respostas médias ao longo do tempo, referentes a dois grupos, são modeladas por funções por partes com um único ponto de mudança com abcissa igual a 15.

O modelo *spline* linear, com um único ponto de mudança em x^* , ilustrado na Figura 5(a) é dado por:

$$\begin{aligned} E(\mathbf{Y}_{ti}) &= \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \\ &+ \gamma_{12}(x_{ti} - x^*)_+ + \gamma_{13}(x_{ti} - x^*)_+w_{1i} + e_{ti}, \end{aligned} \quad (45)$$

sendo x_{ti} o tempo, w_{1i} uma variável indicadora assumindo valor 1 se o indivíduo i não pertence ao grupo *de referência*, $(x_{ti} - x^*)_+$ uma função truncada, que recebe o valor de $(x_{ti} - x^*)$ se $x_{ti} > x^*$ e o valor 0 se $x_{ti} < x^*$.

No modelo em (45), o intercepto da reta média do grupo *de referência* é dado por γ_{00} , sendo portanto, γ_{01} a diferença entre os interceptos das retas médias dos dois grupos. Com relação às inclinações, no intervalo $x_{ti} < x^*$, γ_{10} corresponde à inclinação do grupo *de referência* e $(\gamma_{10} + \gamma_{11})$ à inclinação do outro grupo. As inclinações das retas no intervalo $x_{ti} > x^*$, correspondem à $(\gamma_{10} + \gamma_{12})$ para o grupo *de referência* e $(\gamma_{10} + \gamma_{11} + \gamma_{12} + \gamma_{13})$ para o outro grupo.

O modelo *spline* quadrático, ilustrado na Figura 5(b), é dado por:

$$\begin{aligned} E(\mathbf{Y}_{ti}) = & \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{20}x_{ti}^2 + \gamma_{21}x_{ti}^2w_{1i} + \\ & + \gamma_{22}(x_{ti} - x^*)_+^2 + \gamma_{23}(x_{ti} - x^*)_+^2w_{1i} + e_{ti}. \end{aligned} \quad (46)$$

sendo $(x_{ti} - x^*)_+^2$ uma função truncada, que recebe o valor de $(x_{ti} - x^*)^2$ se $x_{ti} > x^*$ e o valor 0 caso contrário.

No modelo em (46), os interceptos e os coeficientes dos termos lineares permanecem iguais nos intervalos anterior e posterior ao ponto de mudança, nos dois grupos. Com relação aos coeficientes associados ao termo quadrático, no intervalo anterior ao ponto de mudança corresponde à γ_{20} se for do grupo *de referência* e à $(\gamma_{20} + \gamma_{21})$ se for do outro grupo, enquanto que no intervalo após o ponto de mudança corresponde à $(\gamma_{20} + \gamma_{22})$ e $(\gamma_{20} + \gamma_{21} + \gamma_{22} + \gamma_{23})$, respectivamente.

O modelo quadrático por partes, ilustrado na Figura 5(c), é dado por:

$$\begin{aligned} E(\mathbf{Y}_{ti}) = & \gamma_{00} + \gamma_{01}w_{1i} + \gamma_{10}x_{ti} + \gamma_{11}x_{ti}w_{1i} + \gamma_{12}(x_{ti} - x^*)_+ + \gamma_{13}(x_{ti} - x^*)_+w_{1i} \\ & + \gamma_{20}x_{ti}^2 + \gamma_{21}x_{ti}^2w_{1i} + \gamma_{22}(x_{ti} - x^*)_+^2 + \gamma_{23}(x_{ti} - x^*)_+^2w_{1i} + e_{ti}. \end{aligned} \quad (47)$$

No modelo em (47), além dos coeficientes dos termos quadráticos, os referentes aos lineares também são diferentes nos intervalos anterior e posterior ao ponto de mudança. O coeficiente associado ao termo linear no intervalo após o ponto de mudança corresponde à $(\gamma_{10} + \gamma_{12})$, para o grupo *de referência* e à $(\gamma_{10} + \gamma_{11} + \gamma_{12} + \gamma_{13})$ para o outro grupo. Uma vez que somente os interceptos permanecem iguais nos dois intervalos apenas a continuidade das funções é garantida e, portanto, a curva pode deixar de ser suave.

A inclusão de efeitos aleatórios nos modelos envolvendo funções por partes é realizada da mesma forma que nos modelos de tendência linear e quadrática discutidos, por se tratar de um modelo linear nos parâmetros.

De acordo com Fitzmaurice, Laird e Ware (2011), em muitos estudos longitudinais os perfis de respostas podem ser bem representados por meio de *splines* com até dois

Figura 5 – Representação gráfica de funções polinomiais com um único ponto de mudança com abcissa igual a 15

nós localizados criteriosamente no intervalo de tempo.

No caso de mudanças não lineares mais complexas ao longo do tempo, os modelos semi-paramétricos de suavização são apropriados. Entre os métodos de suavização, os *splines* tornaram-se preferidos nos últimos anos, devido à simplicidade de ajuste por meio de ferramentas desenvolvidas para os modelos lineares mistos. Ruppert, Wand e Carroll (2003) mostraram a relação entre os *splines* penalizados (*penalized splines*) e os *splines* suavizados (*smoothing splines*) com os modelos mistos. Wu e Zhang (2006) estenderam a ideia para os polinômios locais (*local polynomial*) e regressão *splines* (*regression splines*).

2.3.2 Especificação do modelo para dados longitudinais com três níveis

Os modelos lineares mistos com 3 níveis de variação aleatória são extensões dos modelos lineares mistos com 2 níveis introduzidos na seção 2.3.1. Esses modelos são também conhecidos como modelos com dois níveis aninhados ou hierárquicos de variação aleatória (WEST; WELCH; GALECKI, 2007; PINHEIRO; BATES, 2000), uma vez que os efeitos aleatórios nas unidades do nível 2 estão aninhados nas unidades do nível 3. A Figura 6 apresenta um esquema representando os coeficientes de um modelo linear misto com 3 níveis.

O modelo linear misto para dados longitudinais com dois níveis hierárquicos

Figura 6 – Esquema representando o modelo misto em 3 níveis

de variação aleatória, apresentado em Pinheiro e Bates (2000) é expresso por:

$$\begin{aligned} \mathbf{y}_{ij} &= \mathbf{X}_{ij}\boldsymbol{\beta} + \mathbf{Z}_{i,j}\mathbf{u}_i + \mathbf{Z}_{ij}\mathbf{u}_{ij} + \mathbf{e}_{ij}, \quad i = 1, \dots, M \quad j = 1, \dots, M_i \\ \mathbf{u}_i &\sim N(\mathbf{0}, \mathbf{G}_1), \quad \mathbf{u}_{ij} \sim N(\mathbf{0}, \mathbf{G}_2), \quad \mathbf{e}_{ij} \sim N(\mathbf{0}, \mathbf{R}_{ij}) \end{aligned} \quad (48)$$

em que $\mathbf{y}_{ij}$ de dimensão $(n_{ij} \times 1)$ é o perfil de respostas da j -ésima unidade de nível 2 aninhada a i -ésima unidade de nível 3, M é número de unidades de nível 3, M_i é número de unidades de nível 2 aninhadas a unidade i de nível 3, $\mathbf{X}_{ij}$ $(n_{ij} \times p)$ é a matriz conhecida de covariáveis (variáveis contínuas ou indicadoras) associadas aos efeitos fixos $\boldsymbol{\beta}$ $(p \times 1)$, $\mathbf{Z}_{i,j}$ $(n_{ij} \times q_1)$ e $\mathbf{Z}_{ij}$ $(n_{ij} \times q_2)$, são matrizes conhecidas de covariáveis associadas, respectivamente, aos efeitos aleatórios de nível 1 e de nível 2, dados por $\mathbf{u}_i$ $(q_1 \times 1)$ e $\mathbf{u}_{ij}$ $(q_2 \times 1)$; $\mathbf{e}_{ij}$ $(n_{ij} \times 1)$ é o vetor de erros aleatórios associados às medidas repetidas na unidade j aninhada na unidade i , $\mathbf{G}_1$ $(q_1 \times q_1)$ e $\mathbf{G}_2$ $(q_2 \times q_2)$, respectivamente, são as matrizes de covariâncias de $\mathbf{u}_i$ e $\mathbf{u}_{ij}$ e $\mathbf{R}_{ij}$ $(n_{ij} \times n_{ij})$ é a matriz de covariâncias dos erros.

Das suposições de normalidade em (48) segue que

$$\mathbf{y}_{ij} \sim N(\mathbf{X}_{ij}\boldsymbol{\beta}, \mathbf{Z}_{i,j}\mathbf{G}_1\mathbf{Z}'_{i,j} + \mathbf{Z}_{ij}\mathbf{G}_2\mathbf{Z}'_{ij} + \mathbf{R}_{ij}). \quad (49)$$

2.3.3 Os métodos de estimação e inferências

Nos modelos mistos é necessário estimar os parâmetros de efeito fixo e os de covariâncias e prever os efeitos aleatórios. Os efeitos aleatórios são preditos e não estimados porque eles não são parâmetros (valores fixos e desconhecidos da população), mas sim, realizações de variáveis aleatórias.

Assumindo que $\mathbf{G}$ e $\mathbf{R}$ são matrizes conhecidas, as estimativas dos parâmetros dos efeitos fixos e as predições dos efeitos aleatórios, podem ser obtidas conjuntamente, a

partir do sistema de equações de modelos mistos- EMM, proposto por Henderson (1950):

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

em que $\hat{\boldsymbol{\beta}}$ corresponde ao BLUE (*Best Linear Unbiased Estimator*), ou seja, o estimador não viesado de $\boldsymbol{\beta}$ que é função linear de $\mathbf{y}$ e tem o menor erro quadrático médio e $\hat{\mathbf{u}}$ corresponde ao BLUP (*Best Linear Unbiased Predictor*) de $\mathbf{u}$ (LITTELL et al., 2006). O detalhamento dos cálculos matriciais necessários para a obtenção do BLUE e BLUP são apresentados no Apêndice A.

Na maioria das situações práticas, $\mathbf{G}$ e $\mathbf{R}$ não são conhecidas, sendo necessário estimá-las. Substituindo nas equações de Henderson seus valores estimados obtêm-se as soluções apresentadas a seguir, que correspondem aos estimadores empíricos do BLUE e BLUP, sendo denotados por EBLUE e EBLUP, respectivamente.

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= \left[\mathbf{X}' \left(\hat{\mathbf{R}}^{-1} - \hat{\mathbf{R}}^{-1} \mathbf{Z} \left(\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1} \right)^{-1} \mathbf{Z}' \hat{\mathbf{R}}^{-1} \right) \mathbf{X} \right]^{-1} \\ &\quad \mathbf{X}' \left(\hat{\mathbf{R}}^{-1} - \hat{\mathbf{R}}^{-1} \mathbf{Z} \left(\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1} \right)^{-1} \mathbf{Z}' \hat{\mathbf{R}}^{-1} \right) \mathbf{y} \\ \hat{\mathbf{u}} &= \left(\mathbf{Z}' \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1} \right)^{-1} \mathbf{Z}' \hat{\mathbf{R}}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}). \end{aligned}$$

De acordo com Littell et al. (2006) é mais apropriado obter as soluções por meio de uma forma de varredura, proposta por Goodnight (1979), que é relacionada à eliminação de Gauss-Jordan e ao procedimento Doolittle.

Dentre os vários métodos para estimar os componentes de variância nos modelos mistos, com efeitos aleatórios normais, o método de máxima verossimilhança restrita (*Restricted Maximum Likelihood-REML*) ou máxima verossimilhança residual (MVR), tem sido preferido, uma vez que o método de máxima verossimilhança (*Maximum Likelihood-ML*) produz estimativas dos componentes de variância e de seus erros padrões viesadas para baixo. Todavia, em alguns casos de delineamentos mais simples e balanceados, os dois métodos produzem o mesmo resultado, que também confere com o mais antigo método de estimação que é baseado nos quadrados médios esperados da ANOVA (método de ANOVA) (LITTELL et al., 2006). A evolução histórica e uma descrição dos vários métodos disponíveis para estimar os componentes de variância, está presente em Searle, Casella e McCulloch, 2006.

A estimação usando MVR envolve métodos numéricos e para grande conjuntos de dados pode tornar-se mais complicada. Vários algoritmos têm sido apresentados para resolver este problema como: o algoritmo EM (*Expectation Maximum*) (DEMPSTER; LAIRD; RUBIN, 1977), o algoritmo *ridge-stabilized* de Newton-Rapson com derivadas analíticas, implementado no procedimento MIXED do SAS (LITTELL et al., 2006) e o *AI-Average Information* (GILMOUR; THOMPSON; CULLIS, 1995), implementado no programa ASRem1. Informações sobre o ASRem1 podem ser obtidas em Butler et al. (2009).

2.3.4 Etapas na seleção do modelo

Na etapa de seleção de um modelo é importante buscar por um modelo simples que se ajuste bem aos dados observados, ou seja, um modelo que seja parcimonioso com relação ao número de parâmetros e que ao mesmo tempo explique bem a variação da variável resposta. No processo de busca é preciso levar em conta os objetivos da pesquisa, o tipo de amostragem ou delineamento experimental utilizado, o conhecimento prévio sobre variáveis preditoras importantes e sobre a variabilidade das unidades amostrais. Nesse sentido, a análise descritiva, os testes de hipóteses e os critérios de informação são ferramentas importantes.

O teste da razão de verossimilhanças é utilizado para comparar dois modelos encaixados, que diferem com relação aos parâmetros de covariâncias (estimativas por MVR) ou de efeitos fixos (estimativas por MVR). Nesse teste utiliza-se a distribuição χ^2 padrão com graus de liberdade igual à diferença entre os números de parâmetros dos modelos comparados. No caso em que os parâmetros de covariâncias encontram-se no limite do espaço paramétrico, assume-se uma mistura de distribuições χ^2 para a estatística do teste (PINHEIRO; BATES, 2000). Entretanto, para se testar a significância de parâmetros de efeitos fixos, esses autores recomendam o uso dos testes F e t , condicionais às estimativas dos parâmetros de covariâncias dos efeitos aleatórios, alertando que a distribuição χ^2 usada no TRV, nesse caso, pode fornecer nível descritivo não conservativo, ou seja, menor que o verdadeiro. Os critérios de informação de Akaike (AIC), de Hurvich and Tsai (AICC), bayesiano de Schwarz (BIC), dentre outros, são úteis no caso de modelos não encaixados. Esses critérios são baseados nos valores da verossimilhança do modelo ajustado e ponderados pela quantidade de observações e pelo número de parâmetros no modelo (WEST; WELCH; GALECKI, 2007, dentre outros). Alguns métodos e gráficos para escolha dos parâmetros

de covariâncias são apresentados em Singer, Nobre, Rocha (2011).

O processo de seleção do modelo misto torna-se mais complexo comparado com o modelo de efeitos fixos uma vez que uma série de escolhas deve ser considerada: definir os efeitos como fixos ou aleatórios, verificar quais efeitos devem ser mantidos ou retirados do modelo e escolher dentre as muitas possibilidades as estruturas das matrizes de covariâncias $\mathbf{G}$ e $\mathbf{R}$, buscando representar as médias e as variabilidades dos dados observados.

Não há uma estratégia única que possa ser utilizada no processo de obtenção e seleção de um modelo, sendo necessário o ajuste e investigação de uma série deles. Entretanto, diante de tantos modelos a serem testados, alguns autores propuseram estratégias que auxiliam na seleção. São apresentadas a seguir, duas delas, uma sugerida por Verbeke e Molenberghs (2000) e a outra por Raudenbush e Bryk (2002), sendo esta última comum em trabalhos nos quais os modelos mistos são especificados da forma hierárquica.

Etapas de seleção sugeridas por Verbeke e Molenberghs

1. Iniciar o processo com um modelo formulado com o maior número possível de efeitos fixos de fatores e covariáveis, incluindo suas interações, buscando melhor explicar o comportamento médio das respostas, antes de investigar a variação aleatória dos dados.
2. Selecionar os efeitos aleatórios a serem incluídos no modelo. Para testar quais deles são significativos pode ser utilizado o teste de razão de verossimilhanças, considerando nulo os respectivos parâmetros de covariâncias.
3. Selecionar a estrutura de covariância para os resíduos no modelo. Uma vez que os efeitos aleatórios e fixos foram adicionados ao modelo, a variância restante na resposta é atribuída ao erro residual, e uma estrutura de covariâncias para os resíduos deve ser investigada.
4. Verificar se há fatores ou covariáveis desnecessários na parte fixa do modelo.

Etapas sugeridas por Raudenbush

No caso da especificação hierárquica do modelo misto, de acordo com Raudenbush e Bryk (2002), é típico começar a análise com o ajuste de um modelo sem covariáveis,

pois esse modelo fornece evidência empírica útil para definir uma especificação apropriada da equação de crescimento individual e várias estatísticas iniciais que servem de base para a avaliação de modelos subsequentes no nível 2. Métodos empíricos para auxiliar na especificação do modelo são apresentados pelos autores, que também discutem as vantagens ao centrar os preditores nos modelos em cada nível, possibilitando interpretações mais interessantes e evitando viés dos estimadores, problemas de multicolinearidade e dificuldades de convergência.

2.3.5 O processo de validação do modelo

A validade das inferências estatísticas depende, sobretudo, da escolha adequada do modelo. Nesse sentido, é importante checar se o modelo está bem ajustado, se as suposições assumidas foram satisfeitas e se existem observações ou perfis atípicos.

Nos modelos lineares tradicionais existem vários tipos de resíduos e de estatísticas para diagnóstico que são bem estabelecidos, como por exemplo, os resíduos ordinários, obtidos pelas diferenças entre os valores observados na amostra e os valores ajustados (preditos sob o modelo ajustado) e variações dos mesmos, como os resíduos estudentizados internamente e os resíduos estudentizados externamente, as medidas de *leverage*, distâncias de Cook, DFBETAS, DFFITS e COVRATIO (COOK; WEISBERG, 1982; BELSLEY; KUH; WELSCH, 2004; DEMÉTRIO; ZOCCHI, 2010, dentre outros). A análise gráfica é bastante utilizada, sendo considerados importantes, o diagrama de dispersão dos resíduos versus valores ajustados, o gráfico da variável adicionada ou da regressão parcial (*added variable plot*), o gráfico de resíduos parciais ou de resíduos mais componente (*partial residual plot*) e os gráficos de probabilidades normal (*normal plot*) e semi-normal (*half normal plot*) com envelope de simulação (ATKINSON, 1987).

Os resíduos marginais têm comportamentos diferentes dos resíduos do modelo clássico, devido à matriz de covariâncias $\mathbf{V}$ não ser diagonal. As estimativas dos efeitos fixos no modelo misto dependem da estimativa desta matriz (LITTELL et al., 2006). Os resíduos marginais são utilizados para diagnósticos sobre os componentes de efeitos fixos (SCHABENBERGER, 2004).

Os resíduos condicionais, obtidos considerando as previsões dos efeitos aleatórios, são específicos do indivíduo e úteis para diagnosticar se os componentes de efeitos aleatórios estão especificados corretamente no modelo (SCHABENBERGER, 2004).

Outro agravante no diagnóstico do modelo misto está na detecção de pontos influentes pois, a adição ou remoção de um ponto não altera somente as estimativas dos efeitos fixos, mas também as estimativas dos componentes de variância (LITTELL et al., 2006). Uma diferença também importante é que nos estudos longitudinais o interesse maior está em detectar perfis individuais influentes ao invés de pontos isolados.

Nos modelos mistos um ponto influente pode afetar as estimativas dos efeitos fixos, dos parâmetros de covariâncias, dos valores preditos dos efeitos aleatórios e dos valores ajustados, assim como suas respectivas precisões. De acordo com Schabenberger (2004), o simples fato de detectar um ponto influente não justifica sua remoção da análise sendo necessário verificar em que aspecto o modelo está sendo afetado.

No caso dos modelos lineares mistos, os resíduos e as medidas de diagnóstico são encontrados em diversas publicações, sendo que alguns deles já se encontram implementados em programas computacionais importantes (CHRISTENSEN; PEARSON; JOHNSON, 1992; SCHABENBERGER, 2004, dentre outros).

Zewotir e Galpin (2006) apresentam estudos de simulação Monte Carlo para avaliar a adequação de várias estatísticas de diagnósticos propostas na literatura para modelos lineares mistos, examinando a sensibilidade a pontos atípicos e *leverage*.

2.3.6 Os procedimentos computacionais

Existem vários softwares estatísticos que podem ser utilizados para ajustar modelos mistos, como o SAS, R, ASReml, SPSS, S-Plus, Mplus, Stata, MLwiN e HLM.

West, Welsh e Galecki (2007) apresentaram uma evolução histórica desses *softwares* salientando, por exemplo, que: Bryk e Raudenbush publicaram em 1982 o primeiro programa, HLM (*hierarchical linear model*), para análise de modelos lineares hierárquicos; Schluchter e Jennrich introduziram em 1988 a rotina BMDP5-V para ajustar modelos com medidas repetidas e desbalanceados; SAS introduziu em 1992 o `proc mixed` como parte do pacote SAS/STAT; Stata versão 5 lançou em 1995 o procedimento `xtreg` para análise de modelos com um único fator aleatório e o procedimento `xtgee` para análise de modelos de dados em painel; Bates e Pinheiro introduziram em 1988 e 2001 no *software* R as funções `lme` e `lmer`, respectivamente, que são usadas no ajuste de modelos lineares mistos. Apresentaram detalhes da sintaxe dos comandos, interpretação e comparação dos resultados de vários tipos de modelos lineares mistos ajustados para conjuntos de dados reais, usando

os softwares estatísticos: SAS (`proc mixed`), SPSS (procedimentos `mixed` e `genlinmixed`), Stata (`mixed`/`xtmixed`), R (funções `lme` e `lmer`) e HLM.

O SAS é um produto altamente reconhecido no mercado e com uma grande confiabilidade de seus resultados, sendo utilizado em vários centros de pesquisas e universidades. Além do `proc mixed`, disponibiliza vários procedimentos para a análise de dados com medidas repetidas, como: o `glimmix`, que é usado para ajustar modelos lineares mistos generalizados; o `hpmixed`, para ajustar modelos lineares mistos complexos com estruturas simples de componentes de variância utilizando técnicas de matrizes esparsas e o `nlmixed`, para ajustar modelos mistos em que os efeitos fixos e aleatórios são incorporados de forma não linear.

O ajuste de modelos com o `proc mixed` pressupõe que os dados sejam normalmente distribuídos, que a média dos dados seja uma função linear de um conjunto de parâmetros de efeitos fixos e que as variâncias e covariâncias apresentem uma das estruturas disponíveis no programa. No procedimento `mixed`, estão implementados diversos tipos de resíduos e estatísticas de diagnóstico que podem ser apresentados em gráficos de alta qualidade utilizando o sistema `ods graphics`.

O R é um ambiente de *software* livre para cálculos estatísticos e construção de gráficos. É um *software* de código aberto em que é possível examinar e alterar os códigos fontes de suas funções e conhecer como os cálculos são realizados. As principais funções usadas no ajuste de modelos lineares mistos podem ser encontradas no pacote `nlme` e `lme4`. Informações adicionais sobre esse e outros pacotes podem ser encontrados no sítio <<http://www.r-project.org>>.

O HLM foi projetado para o ajuste de modelos lineares hierárquicos e oferece várias ferramentas para a descrição, construção de gráficos e análise de dados longitudinais. Nesse programa o usuário especifica os modelos em cada um dos níveis, de forma hierárquica, e a partir deles o programa gera o modelo linear misto (combinado), que pode ser visualizado (RAUDENBUSH et al., 2004). A versão atual ajusta uma larga variedade de modelos lineares hierárquicos, incluindo os modelos com variáveis respostas sem distribuição normal.

3 MATERIAL E MÉTODOS

Nesta seção, são apresentadas algumas aplicações do modelo linear misto para avaliar a evolução de pesos de ratos em testes de toxicidade de produtos químicos. Os dados longitudinais utilizados são provenientes de experimentos conduzidos no laboratório de Ecotoxicologia e Biossegurança da EMBRAPA Meio Ambiente.

As análises estatísticas realizadas são referentes a dois experimentos, sendo um deles para avaliar a toxicidade por doses repetidas (sub-crônica) de um nanocomposto em ratos adultos machos e fêmeas (Experimento 1) e o outro para avaliar a toxicidade reprodutiva, ou seja, os efeitos tóxicos de um herbicida (Experimento 2) no crescimento de filhotes de ratas Wistar tratadas durante a gestação. Os experimentos são descritos na subseção 3.1 e os métodos estatísticos incluindo a implementação no SAS/STAT, versão 9.3 são apresentados na subseção 3.2. Os resultados são discutidos na seção 4.

3.1 Material

3.1.1 Experimento 1 - Avaliação da toxicidade por doses repetidas de um nanocomposto com relação aos perfis de pesos de ratos adultos machos e fêmeas

Com o uso crescente de nanocompostos com características ainda não completamente elucidadas, torna-se imprescindível avaliar o risco que eles podem representar à saúde humana. As nanopartículas avaliadas possuem propriedades bactericidas e estão sendo produzidas para serem incluídas em filmes comestíveis para revestimento de frutas e legumes, visando sua conservação.

Para avaliar a toxicidade do nanocomposto de interesse realizou-se em 2013 um teste com doses repetidas em ratos Wistar adultos, machos e fêmeas. Nesse experimento é estudada a variação de peso corporal durante o período de exposição à substância, pois em caso de efeito tóxico mais evidente pode ocorrer alguma diminuição no ganho natural de peso.

Os ratos usados no experimento foram de colônias mantidas no CEMIB (Centro Multidisciplinar para Investigação Biológica na área da ciência em animais de laboratório - UNICAMP), sendo o estado geral de saúde das colônias controlado por meio de padrões reconhecidos (ISO 9001:2000). Cinco machos e cinco fêmeas foram distribuídos aleatoria-

mente a cada uma das quatro doses do nanocomposto testadas (0, 1, 3 e 5 $\mu\text{g/kg}$ de peso corporal), totalizando 40 ratos adultos. As doses foram administradas diariamente, via oral, ao longo de 28 dias e os animais foram acompanhados e pesados nesses dias, incluindo o *baseline*. Os ratos do grupo controle (dose 0) foram tratados com solução salina e tiveram o mesmo manejo dos ratos tratados com as outras doses. O teste foi realizado de acordo com os Princípios das Boas Práticas de Laboratório (OECD, 1997) e protocolo 422 (OECD, 1996), sendo as doses do nanocomposto preparadas por pesquisadores da EMBRAPA Instrumentação.

A foto na Figura 7 é de uma das ratas deste experimento sendo pesada. A Figura 8(a) apresenta o gráfico com os pesos individuais, ao longo do tempo, de machos (azul) e fêmeas (vermelho) e com os pesos médios por grupo nas combinações dos níveis de sexo e dose (linhas tracejadas). Em ambos os casos os pesos sucessivos foram unidos por segmentos de retas facilitando a visualização dos perfis. Por meio deste gráfico foi possível visualizar a variabilidade que existe entre os perfis de animais do mesmo grupo e a grande discrepância entre os pesos de machos e fêmeas. Na Figura 8(b) foram apresentados os oito perfis médios de pesos nas combinações dos níveis de sexo e dose, possibilitando visualizar as semelhanças e as diferenças entre eles. Os perfis referentes aos machos e às fêmeas que receberam a dose 0 (solução salina) estão em cinza e preto, respectivamente. Para os demais perfis de machos (fêmeas) foram usadas as tonalidades de azul (vermelho), sendo as mais fortes atribuídas as doses maiores.

Figura 7 – Foto do Experimento 1 - Pesagem rata

(a) Pesos individuais das fêmeas (vermelho), dos machos (azul) e médios (tracejados)

(b) Pesos médios

Figura 8 – Perfis individuais e médios de pesos nos níveis dos fatores sexo e dose

3.1.2 Experimento 2 - Avaliação dos efeitos de um herbicida na evolução de pesos de filhotes de ratas, sob dois aspectos: fase de gestação e dose

O herbicida analisado é um agrotóxico utilizado nas principais culturas do estado de São Paulo, para combater ervas daninhas. O experimento foi conduzido em 2006 e dentre os vários aspectos reprodutivos analisados, avaliou-se o efeito da exposição materna durante a gestação no crescimento e desenvolvimento dos filhotes no período de amamentação.

Foram usados ratos Wistar machos e fêmeas nulíparas, com idades entre 75 e 90 dias, de colônias mantidas no CEMIB. Os ratos foram alojados individualmente (exceto durante o acasalamento) em gaiolas de policarbonato com lascas de madeira (maravalha) e dado livre acesso a ração (Purina Lab Chow) e água da torneira. Os animais foram mantidos em ambiente com temperatura controlada ($22 \pm 2^\circ\text{C}$) em 70% de umidade e em um ciclo de 12 horas luz/escuro. Cada macho previamente determinado a ser fértil foi alojado em gaiola com a fêmea no período escuro. O dia em que se encontrou ou um rolinha vaginal ou espermatozoides no esfregaço vaginal foi considerado o início da gestação (0 dias).

As ratas possivelmente prenhas foram alocadas aleatoriamente aos grupos

formados pelas combinações dos níveis de dois fatores: fase da gestação e doses do herbicida. O fator fase da gestação apresenta 2 níveis: inicial, sendo a dose aplicada do 1º ao 6º dia e organogênese do 6º ao 15º dia. O fator dose apresenta 3 níveis: 0, 25, 50 mg/kg de peso corporal. Utilizou-se um delineamento inteiramente casualizado com os tratamentos (combinação dos níveis de fase e dose) em esquema fatorial 2×3 . As ratas dos dois grupos controle, um para cada fase de gestação, foram tratadas com solução salina sob o mesmo manejo das ratas tratadas com as doses da substância.

Após o parto as ratas foram alojadas com as suas ninhadas em gaiolas separadas, continuando com livre acesso a ração e água. No total foram analisadas 60 ratas, sendo 10 em cada grupo. Uma amostra de 5 filhotes foi selecionada aleatoriamente de cada rata, a menos de três delas que tiveram 4 filhotes apenas, sendo todos analisados. Os filhotes foram pesados individualmente e diariamente ao longo de 23 dias. Os filhotes de mesma rata permaneceram na gaiola com a mãe alimentando-se somente do leite materno. Três das ratas tratadas tiveram menos de 4 filhotes. Outras características com relação às ratas e às ninhadas também foram registradas: pesos diários das ratas durante a gestação, tempo de gestação, número de filhotes nascidos (tamanho da ninhada pós parto), número de filhotes natimortos e número de filhotes que morreram durante a amamentação. A Figura 9 apresenta fotos de um experimento que foi conduzido em 2013 nas mesmas condições do Experimento 2 e a Figura 10 o esquema resumindo o delineamento do experimento.

A Figura 11(a) apresenta os perfis de pesos dos filhotes ao longo do período de lactação (23 dias), classificados de acordo com as combinações dos níveis dos fatores *fase* e *dose*, mostrando que há uma grande variabilidade entre os animais e ainda um aumento nas variâncias dos pesos a medida que os filhotes ganham peso. A Figura 11(b) apresenta os perfis de pesos médios dos filhotes por ninhada e também nas combinações dos níveis dos fatores *fase* e *dose* e possibilitou visualizar a variabilidade dos perfis médios de ninhadas com relação ao perfil médio na respectiva combinação dos níveis dos fatores *fase* e *dose*.

A Figura 12 possibilitou visualizar uma variabilidade entre os filhotes de mesma rata. Observou-se também que quanto menor o tamanho da ninhada maior o ganho de peso dos filhotes.

A Figura 13 apresenta os perfis de pesos médios nas combinações dos níveis de fase e dose. O gráfico do item (a) possibilitou comparar os perfis de pesos médios com relação às doses, fixada a fase e o gráfico do item (b) os perfis de pesos médios nas duas

(a) Pesagens dos filhotes

(b) Numeração dos filhotes

(c) Ninhadas

(d) Ração, maravalha e laboratório

Figura 9 – Fotos de um experimento conduzido na EMBRAPA em 2013

Figura 10 – Esquema do delineamento do Experimento 2

fases, fixada a dose. Esses gráficos auxiliam na detecção de interação entre os dois fatores.

Os diagramas de dispersão dos pesos dos filhotes nos diferentes pares de idades são apresentados na Figura 14. Para facilitar a visualização dos diagramas foi necessário considerar subintervalos de idades. Dessa forma, foram apresentados três gráficos contendo os seguintes pares de idades: até 12 dias (a), após 12 dias (b) e somente as idades ímpares (c). A partir dos gráficos, ao traçar retas paralelas à diagonal principal, que contém os histogramas dos pesos em cada idade, observou-se que considerando o mesmo intervalo entre as pesagens as correlações aumentam a medida que os pesos aumentam; ao traçar retas verticais ou horizontais observou-se que a medida que aumenta o intervalo entre as pesagens diminui a correlação entre os pesos e ao traçar uma reta vertical ou horizontal ao longo de p1 (peso no dia 1) observou-se que as correlações com o peso inicial são bem fracas. Também pode ser observado que as correlações entre os pesos em idades menores foram mais fracas e que elas aumentaram a medida que os pesos aumentam.

A Figura 15 apresenta os gráficos de superfície dos pesos dos filhotes em função da idade e das covariáveis: número de filhotes nascidos (a), número de dias de gestação (b) e peso da rata no início da gestação (c). Observou-se, por exemplo, que a relação entre o peso do filhote e o tamanho da ninhada depende da idade, de modo que, a diferença entre os pesos de filhotes de ninhadas menores e maiores aumenta com o aumento da idade.

(a) Pesos individuais

(b) Pesos médios por ninhada e nas combinações dos níveis de fase e dose

Figura 11 – Perfis de pesos individuais e médios dos filhotes ao longo de 23 dias

Figura 12 – Perfis de pesos individuais dos filhotes classificados por ninhada

(a) Classificados por dose

(b) Classificados por fase

Figura 13 – Perfis de pesos médios dos filhotes nas combinações dos níveis dos fatores *fase* e *dose*

(a) Até 12 dias.

(b) A partir de 12 dias.

(c) Idades ímpares.

Figura 14 – Diagramas de dispersão dos pesos dos filhotes em diferentes idades

(a) Número de filhotes nascidos

(b) Número de dias de gestação

(c) Peso da rata no início da gestação

Figura 15 – Gráficos de superfície dos pesos dos filhotes em função dos valores de cada covariável e das idades

3.2 Métodos

Nesta seção são descritas as metodologias usadas nas análises dos experimentos apresentados na seção 3.1.

3.2.1 Experimento 1

A análise exploratória dos dados possibilita detectar características importantes que contribuem para o processo de seleção do modelo. Por exemplo, o gráfico de perfis individuais e médio para cada combinação dos níveis dos fatores *fase* e *dose*, fornece indicativo da variabilidade dos perfis individuais com relação ao respectivo perfil médio, auxiliando na escolha dos efeitos aleatórios. O gráfico contendo uma reta de regressão ajustada para o perfil de pesos de cada animal, obtida por mínimos quadrados e os correspondentes intervalos de confiança para as médias e para as observações individuais, fornecem indicativo da variabilidade das medidas repetidas com relação à reta individual. No caso dos intervalos de confiança de machos e fêmeas apresentarem amplitudes diferentes, ou ainda, quando as variâncias e as covariâncias amostrais forem diferentes para os grupos, há necessidade de testar estruturas diferentes para a matriz de covariâncias dos erros com relação ao sexo.

Utilizou-se o modelo linear misto com dois fatores: *dose* e *sexo*, para avaliar o efeito das doses do nanocomposto no ganho de peso de ratos adultos machos e fêmeas.

No processo de seleção de modelos, seguiu-se as etapas sugeridas por Verbeke e Molenberghs (2000), apresentadas na seção 2.3.4. A primeira etapa foi modelar as respostas médias por meio de um polinômio de ordem 1, sendo portanto, as tendências dos perfis médios, representadas por retas. Dessa forma, considerou-se o *tempo* como covariável contínua, os fatores *sexo* e *dose*, a interação $sexo \times dose$ e as interações dos mesmos com o *tempo*, para descrever a parte fixa. Na segunda etapa foram ajustados e comparados modelos com diferentes estruturas de covariâncias para os efeitos aleatórios e erros com o intuito de explicar as variabilidades devida aos animais e ao longo do tempo, assim como as correlações entre as respostas obtidas repetidamente nas mesmas ratas.

A princípio foram considerados modelos com estruturas de covariâncias mais simples, sendo a complexidade das mesmas aumentada gradativamente buscando melhorar a qualidade do ajuste sem comprometer a parcimônia com relação ao número de parâmetros do modelo. Em todos os casos assumiu-se distribuição normal para os erros e efeitos aleatórios.

A parte fixa, comum a todos os modelos, é dada por:

$$\begin{aligned}
 y_{ti} = & \gamma_{00} + \gamma_{01}M_i + \gamma_{02}D_{1i} + \gamma_{03}D_{3i} + \gamma_{04}D_{5i} + \\
 & + \gamma_{05}M_iD_{1i} + \gamma_{06}M_iD_{3i} + \gamma_{07}M_iD_{5i} + \\
 & + \gamma_{10}d_t + \gamma_{11}M_id_t + \gamma_{12}D_{1i}d_t + \gamma_{13}D_{3i}d_t + \\
 & + \gamma_{14}D_{5i}d_t + \gamma_{15}M_iD_{1i}d_t + \gamma_{16}M_iD_{3i}d_t + \gamma_{17}M_iD_{5i}d_t.
 \end{aligned} \tag{50}$$

sendo y_{ti} , a resposta (peso) observada no rato i no tempo t , $t = 0, 1, 2, \dots, 28$, y_{0i} o peso do rato i antes de receber a dose correspondente (*baseline*); d_t o número de dias após o início do experimento; M_i uma variável indicadora que recebe o valor 1 se o animal i for macho e 0 caso contrário; D_{1i} , D_{3i} e D_{5i} variáveis indicadoras das doses 1, 3 e 5 $\mu\text{g/kg}$ de peso corpóreo, sendo portanto, o grupo de referência formado por fêmeas que receberam a solução salina e os coeficientes γ_{0j} e γ_{1j} , $j = 1, 2, \dots, 7$, respectivamente, os efeitos dos níveis de sexo, dose e da interação $\text{sexo} \times \text{dose}$ nos interceptos e nas inclinações das retas médias.

Assim sendo, os coeficientes das retas médias dos 6 grupos formados pelas combinações dos níveis dos fatores *sexo* e *dose* foram obtidos substituindo em (50) os valores das variáveis indicadoras correspondentes. Por exemplo, no grupo de referência o intercepto e a inclinação da reta média correspondem a γ_{00} e γ_{10} , respectivamente. De forma análoga, tem-se que o grupo formado por ratos machos que receberam a dose de 5 $\mu\text{g/kg}$ apresenta reta média com intercepto e inclinação iguais a $\gamma_{00} + \gamma_{01} + \gamma_{04}$ e $\gamma_{10} + \gamma_{11} + \gamma_{14}$, respectivamente.

Com relação à parte aleatória dos modelos foram consideradas várias estruturas para as matrizes de covariâncias $\mathbf{V}_i$ do vetor de respostas do i -ésimo indivíduo, $\mathbf{Y}_i$, a partir da definição dos efeitos aleatórios e matriz $\mathbf{R}_i$. Em alguns modelos assumiu-se uma matriz $\mathbf{V}_i$ para cada sexo. A matriz de variâncias e covariâncias $\mathbf{V}$ das respostas dos n animais corresponde a uma matriz bloco diagonal, sendo cada bloco igual a $\mathbf{V}_i$.

A seguir são descritos alguns dos modelos que foram ajustados durante o processo de seleção, ressaltando as suas principais características.

M1.1 Modelo assumindo uma estrutura de simetria composta para a matriz de variâncias e covariâncias, $\mathbf{V}_i$, ou seja, considerou-se uma variabilidade constante em todos os tempos e uma covariância (correlação) também constante, de modo que, a semelhança entre os pares de pesos é a mesma independente do intervalo de tempo entre os mesmos. Para obtenção dessa estrutura considerou-se efeito aleatório no intercepto e erros aleatórios independentes com mesma variância.

M1.2 Modelo assumindo estruturas de simetria composta, V_i , diferentes para machos e fêmeas. Essas estruturas foram obtidas considerando variâncias diferentes para os interceptos das retas individuais de machos e fêmeas e erros aleatórios independentes com mesma variância.

M1.3 Modelo com efeito aleatório no intercepto e erros independentes com variâncias homogêneas com relação ao tempo mas diferentes para machos e fêmeas.

M1.4 Modelo com efeitos aleatórios no intercepto e na inclinação e erros aleatórios independentes com mesma variância.

M1.5 Modelo com efeitos aleatórios no intercepto e na inclinação, sendo as variâncias das inclinações das retas individuais de machos e fêmeas diferentes e erros aleatórios independentes com mesma variância.

M1.6 Modelo com efeitos aleatórios no intercepto e na inclinação e erros aleatórios independentes com variâncias homogêneas com relação ao tempo mas diferentes para machos e fêmeas.

M1.7 Modelo com efeitos aleatórios no intercepto e na inclinação e erros aleatórios independentes com variâncias heterogêneas com relação ao tempo.

M1.8 Modelo com efeitos aleatórios no intercepto e na inclinação e erros aleatórios independentes com variâncias heterogêneas com relação ao tempo e sexo.

M1.9 Modelo com efeitos aleatórios no intercepto e na inclinação e uma estrutura autoregressiva de ordem 1 (AR1), descrita em (6), para as variâncias e covariâncias dos erros aleatórios ao longo do tempo.

M1.10 Modelo com efeitos aleatórios no intercepto e na inclinação e estruturas AR1 para as variâncias e covariâncias dos erros aleatórios ao longo do tempo diferentes para machos e fêmeas.

M1.11 Modelo com efeitos aleatórios no intercepto e na inclinação e uma estrutura autoregressiva de ordem 1 heterogênea (ARH1), descrita em (10), para as variâncias e covariâncias dos erros aleatórios ao longo do tempo.

M1.12 Modelo com efeitos aleatórios no intercepto e na inclinação e estruturas ARH1 para as variâncias e covariâncias dos erros aleatórios ao longo do tempo diferentes para machos e fêmeas.

M1.13 Modelo com efeitos aleatórios no intercepto e na inclinação e uma estrutura de ante-dependência de ordem 1 (Ante1), descrita em (12), para as variâncias e

covariâncias dos erros aleatórios ao longo do tempo.

M1.14 Modelo com efeitos aleatórios no intercepto e na inclinação e estruturas Ante1 diferentes para machos e fêmeas.

Como a parte fixa é comum nos modelos M1.1 a M1.14, os mesmos foram comparados a partir do teste da razão de verossimilhanças e dos critérios de informação AIC, AICC e BIC, obtidos usando o método de máxima verossimilhança restrita. Nas análises gráficas dos resíduos condicionais de Pearson, realizadas para checar as suposições do modelo, também utilizou-se o método de máxima verossimilhança restrita.

Os efeitos das interações e os efeitos principais dos fatores foram testados por meio do teste tipo III ¹ de efeitos fixos, considerando o método de estimação de máxima verossimilhança e a correção de Kenward-Roger para o número de graus de liberdade (SPEED, 2013). Os efeitos não significativos foram retirados um a um, ou seja, o modelo foi reajustado após a retirada de cada efeito não significativo.

3.2.1.1 Implementação no SAS

O sistema SAS disponibiliza a ferramenta Enterprise Guide, que possibilita entrar diretamente com as linhas de comandos na linguagem SAS ou ainda definir as tarefas por meio de *menus*. Uma das vantagens dessa ferramenta é a facilidade de armazenar os resultados das análises em um único projeto salvo na extensão **egp**. Ao executar as linhas de comandos na aba **Program** a saída correspondente é apresentada na aba **Results**. Nessa última ao clicar na opção **Create > Report Snapshot** cria-se na janela **Project Tree** um novo item, contendo a respectiva saída, podendo ser renomeado para facilitar a identificação. Dessa forma, os resultados das análises gráficas e de ajuste dos modelos podem ser acessados a partir dos itens da árvore do projeto.

São apresentados nesta seção as linhas de comandos escritas para as análises do Experimento 1, de modo a simplificar o uso da metodologia. São descritos os procedimentos para a entrada dos dados, análise descritiva, ajuste e validação dos modelos lineares mistos.

A Tabela 2 apresenta a estrutura original dos dados de entrada. As informações referentes às unidades experimentais encontram-se nas linhas, sendo que a coluna *ue* identifica a unidade experimental de 1 a 40. A coluna *sexo* indica o sexo do animal:

¹Maiores detalhes em SAS Institute Inc. SAS/STAT 9.3 User's Guide: The MIXED Procedure, 2011, p. 4824.

fêmea (F) ou macho (M), sendo que as 20 primeiras unidades experimentais correspondem as fêmeas e as 20 restantes aos machos. Na coluna *dose* os níveis de dose foram denotados por: 0, 1, 3 e 5 correspondendo, respectivamente, à solução salina, 1, 3 e 5 $\mu\text{g/kg}$ de peso corpóreo. As colunas *peso0*, *peso1*, ..., *peso28* contém os pesos observados nos dias 0, 1,..., 28, respectivamente e dessa forma, os valores da variável resposta nos diferentes tempos são apresentados em colunas separadas, e portanto, na forma multivariada. A coluna *peso0* corresponde aos pesos dos animais no início do experimento, antes de receber a dose correspondente.

O arquivo com os dados foi salvo no formato *xls* e lido a partir do procedimento *import*, como segue:

```
*cria uma planilha no SAS nomeada dados_multivar no formato sas7bdat
a partir da planilha no formato xls importada;
proc import
    datafile="c:\experimento1\dados.xls"
    out=dados_multivar dbms=excel replace;
    getnames=yes;
run;
*salva a planilha dados_multivar na pasta experimento1 em c:,
cujo acesso é feito pelo atalho exper1;
libname exper1 "c:\experimento1"
data exper1.dados_multivar;
    set dados_multivar;
run;
*obtem os valores médios dos pesos nas combinações
dos níveis dos fatores sexo e dose;
proc means data=exper1.dados_multivar noprint;
    by dose sexo;
    var peso0-peso28;
    output out=medias mean=pesomediadosesexo0-pesomediadosesexo28;
run;
*inclui os valores médios obtidos no arquivo dados_multivar;
data exper1.dados_multivar;
    merge exper1.dados_multivar medias;
    by sexo dose;
run;
```

Apenas alguns procedimentos do SAS/STAT requerem o formato de entrada

Tabela 2 – Dados de entrada do Experimento 1 – Formato multivariado

u_e	sexo	dose	peso0	peso1	peso2	...	peso26	peso27	peso28
1	F	0	260	260	268	...	256	261	266
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
5	F	0	290	290	294	...	290	296	292
6	F	1	214	213	220	...	220	214	212
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
10	F	1	272	270	279	...	279	278	279
11	F	3	230	230	232	...	237	238	240
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
15	F	3	279	277	280	...	288	288	287
16	F	5	225	226	230	...	245	241	243
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
20	F	5	290	290	293	...	288	288	294
21	M	0	438	450	465	...	461	468	467
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
25	M	0	465	480	494	...	497	493	493
26	M	1	454	470	486	...	504	497	502
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
30	M	1	390	400	400	...	414	413	410
31	M	3	460	458	456	...	450	464	462
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
35	M	3	490	500	500	...	503	496	502
36	M	5	490	500	500	...	516	517	512
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
40	M	5	435	440	438	...	429	444	445

multivariado, sendo que, na maioria dos procedimentos os valores da variável resposta devem estar em uma única coluna, ou seja, no formato univariado. Nesse caso é necessário uma coluna adicional indicando os tempos correspondentes. A Tabela 3 apresenta os dados do Experimento 1 no formato univariado, obtido a partir dos seguintes comandos:

```
*coloca os dados de entrada no formato univariado;
data exper1.dados_univar;
  set exper1.dados_multivar;
  keep u_e sexo dose peso pesomediodosessexo dia;
```



```

    peso=peso0; pesomediodosesexo=pesomediodosesexo0; dia=0; output exper1.dados_univar;
    peso=peso1; pesomediodosesexo=pesomediodosesexo1; dia=1; output exper1.dados_univar;
    peso=peso2; pesomediodosesexo=pesomediodosesexo2; dia=2; output exper1.dados_univar;
    ...
    peso=peso28; pesomediodosesexo=pesomediodosesexo28; dia=28; output exper1.dados_univar;
run;

```

Dentre as análises do Experimento 1 que foram apresentadas, o formato de entrada multivariado foi necessário apenas para o procedimento `corr` que possibilitou atribuir gradientes de cores às matrizes de variâncias e covariâncias amostrais (Figura 18), obtidas separadamente para machos e fêmeas, a partir dos seguintes comandos:

```

*matriz de covariâncias com cores;
*para que as alterações na saída padrão sejam provisórias;
ods path work.templat(update) sasuser.templat(read) sashelp.tmplmst(read);
*define o gradiente de cores;
%paint(values=600 to 1200 by 1, colors=cyan blue magenta red, macro=setstylecorr)
*faz as alterações na saída padrão;
proc template;
    edit base.corr.matrix;
        column (rowname rowlabel)(matrix)*(matrix2);
    edit matrix;
        %setstylecorr(backgroundcolor)
    end;
end;
run;
*obtem a matriz de variâncias e covariâncias amostral;
proc corr data=exper1.dados_multivar cov noprob nocorr nosimple nomiss;
    where sexo="F"; *ou sexo="M";
    var peso0-peso28;
    ods select all;
run;
*desconsidera as alterações na saída padrão sem precisar fechar o programa;
proc template;
    delete Base.Corr.Matrix;
run;

```

Os gráficos da Figura 8 foram obtidos a partir dos seguintes comandos:

```

*acesso aos arquivos da pasta experimento1 em c;;

```

Tabela 3 – Dados de entrada do Experimento 1 – Formato univariado

u_e	dose	sexo	dia	peso
1	0	F	0	260
1	0	F	1	260
⋮	⋮	⋮	⋮	⋮
1	0	F	27	261
1	0	F	28	266
2	0	F	0	223
2	0	F	1	222
⋮	⋮	⋮	⋮	⋮
2	0	F	27	233
2	0	F	28	235
⋮	⋮	⋮	⋮	⋮
20	5	F	0	290
20	5	F	1	290
⋮	⋮	⋮	⋮	⋮
20	5	F	27	288
20	5	F	28	294
21	0	M	0	438
21	0	M	1	450
⋮	⋮	⋮	⋮	⋮
21	0	M	27	468
21	0	M	28	467
22	0	M	0	430
22	0	M	1	440
⋮	⋮	⋮	⋮	⋮
22	0	M	27	464
22	0	M	28	466
⋮	⋮	⋮	⋮	⋮
40	5	M	0	435
40	5	M	1	440
⋮	⋮	⋮	⋮	⋮
40	5	M	27	444
40	5	M	28	445

```

libname exper1 "c:\experimento1"
*obtem as variáveis pesoF e pesoM;
data exper1.dados_univar;
    set exper1.dados_univar;
    if sexo = "F" then pesoF = peso;
    if sexo = "M" then pesoM = peso;
run;
*item (a)
*defini os nomes dos níveis de dose;
proc format;
    value dosefmt 0="solução salina" 1= "1ug/kg" 3= "3ug/kg" 5= "5ug/kg";
run;
*obtem o gráfico com os perfis de pesos individuais e médios;
ods graphics on / reset=all antialiasmax=1200 imagename='perfis' imagefmt=pdf
    width=4.5in height=4.5in;
ods listing gpath="c:\experimento1";
proc sgpanel data=exper1.dados_univar noautolegend;
    label pesoF="pesos (g) individuais e médios" dia="dias";
    panelby dose / columns=2 rows=2;
    series x=dia y=pesoF / group=u_e lineattrs=(color=red pattern=1 thickness=1);
    series x=dia y=pesoM / group=u_e lineattrs=(color=blue pattern=1 thickness=1);
    series x=dia y=pesomediadosesexos / group=u_e
        lineattrs=(color=black pattern=2 thickness=3);
    colaxis values=(0, 4, 8, 12, 16, 20, 24, 28);
    rowaxis values=(210, 260, 310, 360, 410, 460, 510);
    format dose dosefmt.; *para usar os nomes atribuídos às doses;
run;
*item (b);
goptions reset=all htext=1.5 ftext="arial";
symbol1 c=black      v=none h=.5 i=j r=5 line=1;
symbol2 c=pappk      v=none h=.5 i=j r=5 line=2;
symbol3 c=deeppink   v=none h=.5 i=j r=5 line=3;
symbol4 c=darkred     v=none h=.5 i=j r=5 line=4;
symbol5 c=gray       v=none h=.5 i=j r=5 line=1;
symbol6 c=lightblue  v=none h=.5 i=j r=5 line=2;
symbol7 c=blue       v=none h=.5 i=j r=5 line=3;
symbol8 c=darkblue   v=none h=.5 i=j r=5 line=4;
axis1 order=(220 to 500 by 20) label=(a=90 "pesos médios (g)") length=30;
axis2 order=(0 to 28 by 4) label=('dia') minor=none length=60;

```

```
proc gplot data = exper1.dados_univar;
  plot pesomediadosesexo*dia=u_e / vaxis=axis1 haxis=axis2 nolegend;
run;
```

O procedimento específico para o ajuste de modelos lineares mistos, nos casos em que a variável resposta segue uma distribuição normal é o MIXED, entretanto, os procedimentos GLIMMIX e NLMIXED também podem ser usados para esse fim, ainda que sejam mais específicos para o ajuste de modelos lineares generalizados mistos e para o ajuste de modelos não lineares mistos, respectivamente.

No procedimento MIXED, assim como em alguns outros, ao definir uma variável como fator (classificatória), a partir do comando **class**, são criadas automaticamente as variáveis indicadoras correspondentes. Por exemplo, ao especificar **class dose**, sendo os níveis do fator *dose*: 0, 1, 3 e 5, são criadas automaticamente 3 (número de níveis menos 1) variáveis indicadoras, cada uma assumindo valor 1 se a unidade experimental pertencer à categoria correspondente e 0 caso contrário. Na forma padrão, o último nível ordenado (neste caso, a dose 5) corresponde a referência. Da mesma forma, ao especificar **class sexo**, sendo os níveis do fator *sexo*: *F* e *M* é criada 1 (2-1) variável indicadora *F* e, portanto, o último nível ordenado *M* corresponde a referência. Todavia, em algumas situações é mais interessante considerar algum outro nível como referência ao invés do último, como é o caso do *controle* (*dose* = 0), sendo, portanto, necessário criar uma nova variável com os níveis ordenados de modo que o *controle* seja o último, identificando cada um dos níveis. Por exemplo, para criar as variáveis indicadoras dos níveis de dose D_{1i} , D_{3i} , D_{5i} e de sexo M_i no modelo em (50), foram criadas duas novas variáveis *dose_ord* e *sexo_ord*, de modo que o *controle* e *fêmea* passaram a ser os últimos níveis, respectivamente. Os níveis das novas variáveis foram associados às categorias correspondentes por meio do procedimento **format**. Dessa forma, o grupo formado pelas fêmeas do grupo controle passou a ser o grupo de referência. O código utilizado é apresentado a seguir:

```
data exper1.dados_univar;
  set exper1.dados_univar;
  dia_d = dia;
  if dose = 0 then dose_ord = 4;
  if dose = 1 then dose_ord = 1;
  if dose = 3 then dose_ord = 2;
  if dose = 5 then dose_ord = 3;
```

```

if sexo = "M" then sexo_ord = 1;
if sexo = "F" then sexo_ord = 2;
run;

proc format;
  value dose_ordfmt 1 = "dose1"
                    2 = "dose3"
                    3 = "dose5"
                    4 = "controle";
  value sexo_ordfmt 1 = "M"
                    2 = "F";
run;

```

No procedimento **MIXED** a parte fixa do modelo é especificada por meio do comando **model** e as estimativas e testes de hipóteses dos efeitos fixos são fornecidas acrescentando a opção **solution** após a especificação do modelo. Quando o interesse são os resultados dos testes de hipóteses é necessário especificar **method=ml** e **ddfm=kr**. Os resíduos marginais e condicionais dos modelos lineares mistos são armazenados em planilhas por meio dos comandos **outp** e **outpm**, respectivamente, desde que a opção **residual** também seja usada. Para a verificação das pressuposições do modelo não é necessário definir **method=reml**, pois esse é o método padrão do **SAS**. O procedimento **MIXED** disponibiliza um conjunto de gráficos para análise dos resíduos condicionais e marginais e análise de pontos influentes, quando especificadas as opções **residual** e **influence**, respectivamente e conjuntamente o comando **ods graphics on**.

Para definir os efeitos aleatórios é conveniente usar a opção **subject** do comando **random**, pois dessa forma o processamento é realizado por sujeitos, acarretando uma economia de tempo e memória. A opção **covtest** fornece estimativas dos erros padrões das variâncias e covariâncias dos efeitos aleatórios e também os resultados do teste Z de Wald que servem apenas como um indicativo de quais parâmetros contribuem para melhorar o ajuste. No procedimento **MIXED** estão disponíveis várias estruturas de covariâncias para as matrizes $\mathbf{G}$ e $\mathbf{R}_i$, a partir do comando **type**. Para visualizar essas matrizes devem ser especificadas, respectivamente, as opções **g** e **r** do comando **random**. Uma outra opção desse comando é a matriz $\mathbf{V}_i$ obtida a partir de **v**. Por meio da opção **group** disponível nos comandos **repeated** ou **random** é possível considerar matrizes $\mathbf{R}_i$ e $\mathbf{V}_i$ diferentes para indivíduos de grupos distintos. Para visualizá-las é necessário indicar os indivíduos corres-

pondentes, por exemplo, ao especificar $\mathbf{v}=(1 \ 21)$, são exibidas as matrizes dos indivíduos 1 e 21, ou seja, $\mathbf{V}_1$ e $\mathbf{V}_{21}$. O MIXED fornece de forma automática o resultado do teste da razão de verossimilhanças que compara o modelo que está sendo ajustado, com mais de um componente de variância, com o modelo reduzido com uma única variância residual igual para todos os níveis do fator longitudinal. É importante destacar que o nível descritivo desse teste não é calculado levando em conta as misturas de χ^2 .

Nos casos em que houve problemas de convergência seguiu-se as dicas e estratégias apresentadas por Kiernan, Tao e Gibbs (2012), para problemas com os ajustes de modelos mistos no *SAS*.

As linhas de comandos usadas para o ajuste do modelo M1.1, a partir do procedimento MIXED, são apresentadas a seguir e às referentes aos modelos M1.2 a M1.14 encontram-se no Apêndice B.

```
*M1.1;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal; *method=ml;
  class u_e dose_ord sexo_ord dia_d ;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
            dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia/
            residual solution outp=exper1.res_cond_m1 outpm=exper1.res_marg_m1;
            *ddfm=kr;
  random intercept / subject=u_e solution g v vcorr;
  repeated dia_d/subject=u_e r rcorr;
  format dose_ord dose_ordfmt.;
  format sexo_ord sexo_ordfmt.;
run; *obs: incluir os comandos comentados para obter os resultados dos testes de
      hipóteses dos efeitos fixos;
ods graphics off;
```

Para obter os gráficos da Figura 17 utilizou-se os comandos:

```
*retas de regressão e IC;
*item (a);
ods graphics on / reset=all width=5in height=5in;
proc sgpanel data=exper1.dados_univar noautolegend;
  where (sexo="F");
  panelby u_e / columns=5 rows=4 ;
```

```

title;
label dia="dia" peso="pesos, retas de regressão individuais (MQ) e ICs" u_e="rata";
reg x=dia y=peso / group=u_e clm cli;
colaxis values=(0, 4, 8, 12, 16, 20, 24, 28);
rowaxis values=(200, 250, 300, 350);
run;
ods graphics off;
*item (b);
ods graphics on / reset=all width=5in height=5in;
proc sgpanel data=exper1.dados_univar noautolegend;
  where (sexo="M");
  panelby u_e / columns=5 rows=4 ;
  title;
  label dia="dia" peso="pesos, retas de regressão individuais (MQ) e ICs" u_e="rato";
  reg x=dia y=peso / group=u_e clm cli;
  colaxis values=(0, 4, 8, 12, 16, 20, 24, 28);
  rowaxis values=(400, 450, 500, 550);
run;
ods graphics off;

```

Para obter os gráficos da Figuras 19 e 20, referentes à análise de resíduos dos modelos ajustados e o gráfico com os perfis individuais de pesos observados e respectivas retas ajustadas (Figura 21) utilizou-se as seguintes linhas de comandos:

```

*histograma dos resíduos;
goptions reset=all device=jpeg ftext="arial" htext=14pt vsize=4 hsize=5;
axis1 order=(0 to 30 by 5) label=(a=90"%");
proc capability data=exper1.res_cond_m12 noprint;
*exper1.res_cond_m12 resíduos condicionais do modelo M1.12;
var PearsonResid;
label PearsonResid = "resíduos condicionais de Pearson";
histogram /normal (mu=0 sigma=1 w=2 color=bib) midpoints=(-6 to 6 by 0.5)
  nolegend cfill=brpk vaxis=axis1;
inset mean="média" (5.2) std="d.p." (4.2) min="min." (5.2) max="máx." (4.2) /
  pos=ne;
run;

*gráfico normal de probabilidades;
goptions reset=all ftext="arial" htext=14pt vsize=4 hsize=5;

```

```

symbol v=circle w=2 h=2 color=brgr;
axis1 order=(-6.5 to 6.5 by 1) label=(a=90 "resíduos condicionais de Pearson");
axis2 order=(-4 to 4 by 1) label=("quantis");
proc capability data=exper1.res_cond_m12 noprint;
    qqplot PearsonResid / normal(mu=0 sigma=1 color=black w=2)
        nolegend vaxis=axis1 haxis=axis2 ;
run;

```

*gráfico dos resíduos vs preditos;

```

goptions reset=all device=jpeg ftext="arial" htext=14pt vsize=4 hsize=5;
symbol1 c=brgr v=circle h=.8;
axis1 order=(-6.5 to 6.5 by 1) label=(a=90 "resíduos condicionais de Pearson")
    minor=none;
axis2 order=(200 to 520 by 40) label=("preditos (g)") minor=none;
proc gplot data=exper1.res_cond_m12;
    plot PearsonResid*pred/ vaxis=axis1 haxis=axis2 vref=(-3 0 3);
run;

```

*gráfico dos pesos observados e retas ajustadas para cada animal (mesmas cores para ambos);

*para definir os nomes dos níveis de dose;

```

proc format;
    value dosefmt 0="solução salina" 1="1ug/kg" 3="3ug/kg" 5="5ug/kg";
run;

data myattrmap;
    input id $ value linecolor $ @@;
    datalines;
u_e 1 bippk      u_e 2 red      u_e 3 cxe9967a u_e 4 darkred
u_e 5 deeppink  u_e 6 gray     u_e 7 blue    u_e 8 bib
u_e 9 cx6495ed u_e 10 cx00bfff u_e 11 bippk   u_e 12 red
u_e 13 cxe9967a u_e 14 darkred u_e 15 deeppink u_e 16 gray
u_e 17 blue     u_e 18 bib     u_e 19 cx6495ed u_e 20 cx00bfff
u_e 21 bippk    u_e 22 red     u_e 23 cxe9967a u_e 24 darkred
u_e 25 deeppink u_e 26 gray    u_e 27 blue     u_e 28 bib
u_e 29 cx6495ed u_e 30 cx00bfff u_e 31 bippk    u_e 32 red
u_e 33 cxe9967a u_e 34 darkred u_e 35 deeppink u_e 36 gray
u_e 37 blue     u_e 38 bib     u_e 39 cx6495ed u_e 40 cx00bfff
;

run;

ods graphics on / reset=all antialiasmax=1200 imagename="predeobs" imagefmt=pdf

```



```

width=4.5in height=4.5in;
ods listing gpath="c:\experimento1";
proc sgpanel data=exper1.res_cond_m12 noautolegend dattrmap=myattrmap;
  label dia="dias" pred="pesos (g) observados e preditos";
  panelby dose / columns=2 rows=2;
  series x=dia y=pred / group=u_e attrid=u_e lineattrs=(pattern=2);
  series x=dia y=peso / group=u_e attrid=u_e lineattrs=(pattern=1);
  colaxis values=(0,2,4,6,8,10,12,14,16,18,20,22,24,26,28);
  rowaxis values=(200,220,240,260,280,300,320,340,360,380,400,420,440,460,480,500,520);
  format dose dosefmt.;
run;
ods graphics off;

```

3.2.2 Experimento 2

Uma das características dos experimentos realizados para avaliar os efeitos de produtos químicos aplicados nas mães, nas gerações futuras, durante alguma fase da vida dos animais é a estrutura hierárquica das unidades. Por exemplo, nesse experimento conduzido para estudar a influência das doses aplicadas nas mães durante a gestação, sobre os pesos de seus filhotes ao longo do período de amamentação é natural assumir: que os pesos observados ao longo do tempo no mesmo filhote sejam dependentes; que os pesos de filhotes de mesma rata sejam dependentes, uma vez que há dependência entre filhotes de mesma rata e que os pesos de filhotes de ratas diferentes sejam independentes, uma vez que as ratas são independentes e provenientes de uma amostra aleatória de uma população de ratas. Essa estrutura de dependência entre os pesos observados nos filhotes está esquematizada na Figura 16, sendo que os tempos, $tempos_{tij}$ (tempo t do filhote j da rata i), correspondem as idades em que foram feitas as pesagens.

As ratas e os filhotes foram considerados fatores de efeito aleatório, uma vez que representam amostras aleatórias das respectivas populações. Como as ratas e os filhotes são fatores aleatórios e os filhotes são aninhados às ratas, tem-se dois níveis hierárquicos de variação aleatória.

Os fatores *fase* e *dose* foram considerados de efeito fixo, uma vez que há interesse em comparar os perfis de pesos dos filhotes nas combinações dos níveis desses dois fatores.

A idade (em dias) dos filhotes foi considerada uma covariável contínua, uma

Figura 16 – Estrutura hierárquica na unidade experimental (rata) em experimentos para avaliar os pesos longitudinais dos filhotes

vez que o interesse foi comparar os perfis médios de pesos ao longo do tempo e não os pesos médios nas idades. A idade foi centrada em 1, ou seja, utilizou-se $(idade-1)$, de modo que o intercepto passou a ser o 1o. dia de vida (valor que foi observado).

Buscando explicar parte da variabilidade das respostas, foram testadas as covariáveis número de dias de gestação G_i , número de filhotes nascidos T_i (tamanho da ninhada) e o peso inicial das ratas no início do experimento P_i . Essas covariáveis foram centradas nas respectivas médias para facilitar a interpretação dos parâmetros, de modo que seus valores iguais a zero correspondem aos valores médios.

Para modelar a relação entre os pesos dos filhotes e as idades nas pesagens foram comparadas três tipos de abordagens: o modelo de regressão polinomial de grau 3 (abordagem 1), o modelo *spline* quadrático com um único ponto de mudança fixado aos 15 dias (abordagem 2) e uma modificação deste último relaxando a restrição de continuidade das derivadas de primeira ordem (abordagem 3). Todos os modelos incluíram as covariáveis de ratas centradas nas respectivas médias.

O modelo de regressão polinomial de grau 3 apresenta uma desvantagem comparado com os modelos *spline* quadrático e de regressão quadrática por partes, em relação à interpretação dos parâmetros, já que nesse caso não há uma correspondência direta entre cada parâmetro e a taxa ou aceleração de crescimento. Além disso, a estimação dos parâmetros no modelo de regressão polinomial pode ser mais instável. Nesse modelo, uma forma para comparar as curvas médias de crescimento seria comparando conjuntamente os coeficientes dos termos de ordem 1, 2 e 3. No caso dos modelos *spline* quadrático e de regressão quadrática por partes (abordagens 2 e 3), o crescimento dos filhotes é modelado por duas parábolas, uma delas no intervalo anterior e a outra no intervalo posterior ao ponto

de mudança, que corresponde a idade média de abertura dos olhos dos filhotes. Os ganhos de peso nos intervalos anterior e posterior a abertura dos olhos são obtidos derivando as funções quadráticas correspondentes. Uma vez que a derivada de uma função quadrática resulta numa função linear, tem-se que o ganho de peso diário aumenta de forma linear com o aumento da idade em cada um dos intervalos. As derivadas de ordem 2 das funções quadráticas anterior e posterior ao ponto de mudança correspondem ao aumento no ganho diário de peso, constante e diferente em cada um dos intervalos.

As partes fixas dos modelos lineares mistos nas três abordagens comparadas são apresentadas a seguir:

Modelo polinomial de grau 3

$$\begin{aligned}
 y_{tij} = & (\gamma_{00} + \gamma_{01}F_{2i} + \gamma_{02}D_{2i} + \gamma_{03}D_{3i} + \gamma_{04}F_{2i}D_{2i} + \gamma_{05}F_{2i}D_{3i} + \gamma_{06}G_i + \gamma_{07}T_i + \gamma_{08}P_i) + \\
 & + (\gamma_{10} + \gamma_{11}F_{2i} + \gamma_{12}D_{2i} + \gamma_{13}D_{3i} + \gamma_{14}F_{2i}D_{2i} + \gamma_{15}F_{2i}D_{3i})d_{tij} + \\
 & + (\gamma_{20} + \gamma_{21}F_{2i} + \gamma_{22}D_{2i} + \gamma_{23}D_{3i} + \gamma_{24}F_{2i}D_{2i} + \gamma_{25}F_{2i}D_{3i})d_{tij}^2 + \\
 & + (\gamma_{30} + \gamma_{31}F_{2i} + \gamma_{32}D_{2i} + \gamma_{33}D_{3i} + \gamma_{34}F_{2i}D_{2i} + \gamma_{35}F_{2i}D_{3i})d_{tij}^3 + e_{tij} \quad (51)
 \end{aligned}$$

sendo y_{tij} e d_{tij} , respectivamente, o peso e a idade (em dias) centrada em 1, ou seja, (idade-1), do filhote j da rata i no tempo t , $t = 1, \dots, 23$; d_{tij}^2 a $(idade - 1)^2$ e d_{tij}^3 a $(idade - 1)^3$; F_{2i} uma variável indicadora que recebe o valor 1 se a rata i for da fase 2, D_{2i} e D_{3i} , variáveis indicadoras das doses 25 e 50 mg/kg, respectivamente e G_i , T_i e P_i variáveis contínuas centradas nas médias, sendo respectivamente, número de dias de gestação, tamanho da ninhada e peso inicial da rata i .

Modelo *spline* quadrático com ponto de mudança em $d_{tij} = 14$

$$\begin{aligned}
 y_{tij} = & (\gamma_{00} + \gamma_{01}F_{2i} + \gamma_{02}D_{2i} + \gamma_{03}D_{3i} + \gamma_{04}F_{2i}D_{2i} + \gamma_{05}F_{2i}D_{3i} + \gamma_{06}G_i + \gamma_{07}T_i + \gamma_{08}P_i) + \\
 & + (\gamma_{10} + \gamma_{11}F_{2i} + \gamma_{12}D_{2i} + \gamma_{13}D_{3i} + \gamma_{14}F_{2i}D_{2i} + \gamma_{15}F_{2i}D_{3i})d_{tij} + \\
 & + (\gamma_{20} + \gamma_{21}F_{2i} + \gamma_{22}D_{2i} + \gamma_{23}D_{3i} + \gamma_{24}F_{2i}D_{2i} + \gamma_{25}F_{2i}D_{3i})d_{tij}^2 + \\
 & + (\gamma_{30} + \gamma_{31}F_{2i} + \gamma_{32}D_{2i} + \gamma_{33}D_{3i} + \gamma_{34}F_{2i}D_{2i} + \gamma_{35}F_{2i}D_{3i})(d_{tij} - 14)_+^2 + e_{tij} \quad (52)
 \end{aligned}$$

sendo $(d_{tij} - 14)_+^2$ o quadrado de uma função truncada, que recebe o valor de $(d_{tij} - 14)$ se $d_{tij} > 14$ e o valor 0 caso contrário. As demais variáveis correspondem às descritas no modelo em (51).

Modelo de regressão quadrática por partes com ponto de mudança em $d_{tij} = 14$

$$\begin{aligned}
y_{tij} = & (\gamma_{00} + \gamma_{01}F_{2i} + \gamma_{02}D_{2i} + \gamma_{03}D_{3i} + \gamma_{04}F_{2i}D_{2i} + \gamma_{05}F_{2i}D_{3i} + \gamma_{06}G_i + \gamma_{07}T_i + \gamma_{08}P_i) + \\
& + (\gamma_{10} + \gamma_{11}F_{2i} + \gamma_{12}D_{2i} + \gamma_{13}D_{3i} + \gamma_{14}F_{2i}D_{2i} + \gamma_{15}F_{2i}D_{3i})d_{tij} + \\
& + (\gamma_{20} + \gamma_{21}F_{2i} + \gamma_{22}D_{2i} + \gamma_{23}D_{3i} + \gamma_{24}F_{2i}D_{2i} + \gamma_{25}F_{2i}D_{3i})d_{tij}^2 + \\
& + (\gamma_{30} + \gamma_{31}F_{2i} + \gamma_{32}D_{2i} + \gamma_{33}D_{3i} + \gamma_{34}F_{2i}D_{2i} + \gamma_{35}F_{2i}D_{3i})(d_{tij} - 14)_+ + \\
& + (\gamma_{40} + \gamma_{41}F_{2i} + \gamma_{42}D_{2i} + \gamma_{43}D_{3i} + \gamma_{44}F_{2i}D_{2i} + \gamma_{45}F_{2i}D_{3i})(d_{tij} - 14)_+^2 + e_{tij}. \quad (53)
\end{aligned}$$

As variáveis do modelo 53 correspondem às descritas no modelo em (52).

Considerando cada uma das abordagens foram ajustados vários modelos, alterando a parte aleatória, a matriz $\mathbf{R}_i$ e em alguns casos também a parte fixa. Assumiu-se distribuição normal para os erros e efeitos aleatórios e utilizou-se matrizes de covariâncias não estruturadas para descrever as variâncias e covariâncias dos efeitos aleatórios de rata e/ou filhote, ou seja, variâncias e covariâncias diferentes, buscando reduzir as dificuldades computacionais, como sugere Fitzmaurice, Laird e Ware (2011). A Tabela 4 apresenta alguns modelos de regressão polinomial de grau 3 selecionados para comparação.

De acordo a Tabela 4 apenas os modelos M2.1 e M2.10 não apresentam a parte fixa adicional, que corresponde a covariável $\cos(a \times t)$ que foi incluída em alguns modelos buscando melhorar a qualidade dos ajustes. Essa covariável foi criada a partir do ajuste de um modelo não linear, $Y = w \times \cos(a \times d_t)$ sendo a variável resposta Y os resíduos do modelo M2.1., $d_t = 0, 1, \dots, 22$ a (idade-1) no tempo $t = 1, 2, \dots, 23$ e w e a parâmetros que foram estimados conjuntamente. Somente a estimativa de a foi usada no cálculo dos valores dessa covariável, sendo fixado $w=1$. O valor da constante usado foi $a = 5,753054$. Esse artifício foi apresentado por Pinheiro e Bates (2000) e também por Goldstein (2003) e visa modelar o comportamento periódico da variável resposta com relação ao tempo não explicado pelas covariáveis inclusas no modelo. O fato dos resíduos do modelo M2.1 apresentar um comportamento periódico centrado em zero é que possibilitou representá-lo por uma oscilação tipo cosseno.

Na Tabela 4, $u_{0i}, u_{1i}, u_{2i}, u_{3i}$ são os efeitos aleatórios de rata no intercepto e nos coeficientes: linear, quadrático e de $\cos(a \times d_t)$, respectivamente. Os efeitos aleatórios de filhote aninhado a rata, no intercepto e nos coeficientes linear e quadrático são denotados por $u_{1ij}, u_{2ij}, u_{3ij}$, respectivamente. O erro aleatório na resposta no tempo t do filhote j da rata i é dado por e_{tij} , sendo independente dos efeitos aleatórios de rata e filhote. As variâncias populacionais dos efeitos aleatórios de rata no intercepto, coeficiente linear, quadrático e de

$\cos(a \times d_t)$ são denotadas, respetivamente, por σ_{r0}^2 , σ_{r1}^2 e σ_{r2}^2 , σ_{r3}^2 e as covariâncias entre os efeitos aleatórios por σ_{r01} , σ_{r02} , σ_{r03} , σ_{r12} , σ_{r13} , σ_{r23} . Por exemplo, σ_{r01} denota a covariância entre os efeitos aleatórios de rata no intercepto e no coeficiente linear. Para denotar as variâncias e covariâncias dos efeitos aleatórios de filhote, substitui-se σ_r por σ_f .

Foram também ajustados modelos incluindo a interação entre a covariável número de filhotes nascidos e o fator longitudinal, $T_i \times d_{tij}$ nas três abordagens consideradas. A Tabela 5 apresenta o modelo de melhor ajuste em cada uma das abordagens, dentre vários modelos comparados considerando efeitos aleatórios de filhote em alguns dos coeficientes lineares e quadráticos das curvas antes e após do ponto de mudança e também diferentes estruturas para a matriz $\mathbf{R}_i$.

No modelo polinomial M2.11 apresentado na Tabela 5 utilizou-se as mesmas notações dos modelos polinomiais apresentados na Tabela 4. Os modelos M2.12 e M2.13, também apresentados na Tabela 5, foram ajustados a partir das abordagens 2 e 3, sendo os perfis individuais dos filhotes especificados em (54) e (55), respetivamente.

$$\begin{aligned}
& y_{tij}|u_{0i}, u_{1i}, u_{1ij}, u_{2i}, u_{2ij}, u_{3i} = \\
& = (\gamma_{00} + u_{0i} + \gamma_{01}F_{2i} + \gamma_{02}D_{2i} + \gamma_{03}D_{3i} + \gamma_{04}F_{2i}D_{2i} + \gamma_{05}F_{2i}D_{3i} + \gamma_{06}G_i + \gamma_{07}T_i + \gamma_{08}P_i) + \\
& \quad + (\gamma_{10} + u_{1i} + u_{1ij} + \gamma_{11}F_{2i} + \gamma_{12}D_{2i} + \gamma_{13}D_{3i} + \gamma_{14}F_{2i}D_{2i} + \gamma_{15}F_{2i}D_{3i})d_{tij} + \\
& \quad + (\gamma_{20} + u_{2i} + u_{2ij} + \gamma_{21}F_{2i} + \gamma_{22}D_{2i} + \gamma_{23}D_{3i} + \gamma_{24}F_{2i}D_{2i} + \gamma_{25}F_{2i}D_{3i})d_{tij}^2 + \\
& \quad + (\gamma_{30} + u_{3i} + \gamma_{31}F_{2i} + \gamma_{32}D_{2i} + \gamma_{33}D_{3i} + \gamma_{34}F_{2i}D_{2i} + \gamma_{35}F_{2i}D_{3i})(d_{tij} - 14)_+^2 + e_{tij} \quad (54)
\end{aligned}$$

$$\begin{aligned}
& y_{tij}|u_{0i}, u_{1i}, u_{1ij}, u_{2i}, u_{3i}, u_{4i} = \\
& = (\gamma_{00} + u_{0i} + \gamma_{01}F_{2i} + \gamma_{02}D_{2i} + \gamma_{03}D_{3i} + \gamma_{04}F_{2i}D_{2i} + \gamma_{05}F_{2i}D_{3i} + \gamma_{06}G_i + \gamma_{07}T_i + \gamma_{08}P_i) + \\
& \quad + (\gamma_{10} + u_{1i} + u_{1ij} + \gamma_{11}F_{2i} + \gamma_{12}D_{2i} + \gamma_{13}D_{3i} + \gamma_{14}F_{2i}D_{2i} + \gamma_{15}F_{2i}D_{3i})d_{tij} + \\
& \quad + (\gamma_{20} + u_{2i} + \gamma_{21}F_{2i} + \gamma_{22}D_{2i} + \gamma_{23}D_{3i} + \gamma_{24}F_{2i}D_{2i} + \gamma_{25}F_{2i}D_{3i})d_{tij}^2 + \\
& \quad + (\gamma_{30} + u_{3i} + u_{3i} + \gamma_{31}F_{2i} + \gamma_{32}D_{2i} + \gamma_{33}D_{3i} + \gamma_{34}F_{2i}D_{2i} + \gamma_{35}F_{2i}D_{3i})(d_{tij} - 14)_+ + \\
& \quad + (\gamma_{40} + u_{4i} + \gamma_{41}F_{2i} + \gamma_{42}D_{2i} + \gamma_{43}D_{3i} + \gamma_{44}F_{2i}D_{2i} + \gamma_{45}F_{2i}D_{3i})(d_{tij} - 14)_+^2 + e_{tij} \quad (55)
\end{aligned}$$

Tabela 4 – Modelos polinomiais de grau 3 selecionados para comparação - Experimento 2

Modelo	parte fixa adicional	ef. aleatórios nível 1	matriz covariâncias nível 1	ef. aleatórios nível 2	matriz covar. nível 2	matriz covar. dos erros
M2.1	-	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} \\ & \sigma_{r1}^2 & \sigma_{r12} \\ & & \sigma_{r2}^2 \end{bmatrix}$	$\begin{bmatrix} u_{1ij} \\ u_{2ij} \end{bmatrix}$	$\begin{bmatrix} \sigma_{f1}^2 & \sigma_{f12} \\ & \sigma_{f2}^2 \end{bmatrix}$	$\sigma^2 I$
M2.2	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} \\ & \sigma_{r1}^2 & \sigma_{r12} \\ & & \sigma_{r2}^2 \end{bmatrix}$	$\begin{bmatrix} u_{1ij} \\ u_{2ij} \end{bmatrix}$	$\begin{bmatrix} \sigma_{f1}^2 & \sigma_{f12} \\ & \sigma_{f2}^2 \end{bmatrix}$	$\sigma^2 I$
M2.3	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	$\begin{bmatrix} u_{1ij} \\ u_{2ij} \end{bmatrix}$	$\begin{bmatrix} \sigma_{f1}^2 & \sigma_{f12} \\ & \sigma_{f2}^2 \end{bmatrix}$	$\sigma^2 I$
M2.4	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	$\begin{bmatrix} u_{1ij} \\ u_{2ij} \end{bmatrix}$	$\begin{bmatrix} \sigma_{f1}^2 & \sigma_{f12} \\ & \sigma_{f2}^2 \end{bmatrix}$	AR1
M2.5	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	-	-	ARH1
M2.6	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	-	-	Ante1
M2.7	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	-	-	Toeph
M2.8	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	-	-	Ne
M2.9	$\cos(a \times t)$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} \\ & \sigma_{r1}^2 & \sigma_{r12} \\ & & \sigma_{r2}^2 \end{bmatrix}$	-	-	Ne
M2.10	-	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} \\ & \sigma_{r1}^2 & \sigma_{r12} \\ & & \sigma_{r2}^2 \end{bmatrix}$	-	-	Ne

a=5,753054.

Tabela 5 – Modelos das três abordagens selecionados para comparação - Experimento 2

Modelo	parte fixa adicional	ef. aleatórios nível 1	matriz covariâncias nível 1	ef. aleatórios nível 2	matriz covar. nível 2	matriz covar. dos erros
M2.11	$\cos(a \times t)$ $T_i \times d_{tij}$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	-	-	Toeplitz
M2.12	$T_i \times d_{tij}$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} \\ & & \sigma_{r2}^2 & \sigma_{r23} \\ & & & \sigma_{r3}^2 \end{bmatrix}$	$\begin{bmatrix} u_{1ij} \\ u_{2ij} \end{bmatrix}$	$\begin{bmatrix} \sigma_{f1}^2 & \sigma_{f12} \\ & \sigma_{f2}^2 \end{bmatrix}$	Ante1
M2.13	$T_i \times d_{tij}$	$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \\ u_{4j} \end{bmatrix}$	$\begin{bmatrix} \sigma_{r0}^2 & \sigma_{r01} & \sigma_{r02} & \sigma_{r03} & \sigma_{r04} \\ & \sigma_{r1}^2 & \sigma_{r12} & \sigma_{r13} & \sigma_{r14} \\ & & \sigma_{r2}^2 & \sigma_{r23} & \sigma_{r24} \\ & & & \sigma_{r3}^2 & \sigma_{r34} \\ & & & & \sigma_{r5}^2 \end{bmatrix}$	$\begin{bmatrix} u_{1ij} \\ u_{2ij} \end{bmatrix}$	$\begin{bmatrix} \sigma_{f1}^2 & \sigma_{f12} \\ & \sigma_{f2}^2 \end{bmatrix}$	Ante1

$a=5,753054$.

3.2.2.1 Implementação no SAS

Nesta seção são apresentadas as linhas de comandos escritas no SAS/STAT para a entrada dos dados, análise descritiva, ajuste e validação dos modelos lineares mistos do Experimento 2.

A Tabela 6 apresenta a estrutura original dos dados de entrada. As informações referentes aos filhotes das unidades experimentais (progenitoras) encontram-se nas linhas, sendo que as colunas *u_e* e *filhote* identificam, respectivamente, a unidade experimental de 1 a 60 e seus filhotes. No total foram analisados 295 filhotes identificados na coluna *obs*. As colunas *fase* e *dose* indicam, respectivamente, os níveis correspondentes aos fatores *dose* (solução salina, 25 ou 50mg/kg) e *fase* (1 ou 2). As colunas *peso1*, *peso2*, ..., *peso23* contêm os pesos observados nos dias 1, ..., 23, respectivamente apresentados em colunas separadas, e portanto, na forma multivariada. As colunas *no. filh. nasc.*, *no. filh. n. vivos*, *no. dias gest.* e *peso inic. rata* indicam, respectivamente, os valores das covariáveis: número de filhotes nascidos, número de filhotes nascidos vivos, número de dias de gestação e peso inicial da progenitora no início do experimento, referentes à cada *u_e*.

Tabela 6 – Dados de entrada do Experimento 2 – Formato multivariado

obs	u_e	filhote	fase	dose	peso1	peso2	...	peso23	no. filh. nasc.	no. filh. n. vivos	no. dias gest.	peso inic. rata
1	1	1	1	0	5	8	...	60	6	6	23	268
2	1	2	1	0	5	7	...	61	6	6	23	268
⋮	⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮	⋮
5	1	5	1	0	5	8	...	60	6	6	23	268
6	2	1	1	0	5	7	...	39	10	10	21	239
7	2	2	1	0	7	8	...	38	10	10	21	239
⋮	⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮	⋮
10	2	5	1	0	6	7	...	38	10	10	21	239
⋮	⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮	⋮
291	60	1	2	50	7	8	...	58	6	6	23	235
292	60	2	2	50	8	8	...	55	6	6	23	235
⋮	⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮	⋮
295	60	5	2	50	7	9	...	55	6	6	23	235

A leitura dos dados e a obtenção da planilha *dados_multivar* no formato

`sas7bdat` foi realizada da mesma forma que no Experimento 1, usando as linhas de comandos apresentadas na seção 3.2.1.1.

Para obter os valores das covariáveis centradas nas respectivas médias, foi necessário primeiramente criar uma nova planilha na qual foram mantidas para cada *u_e* uma única linha, lembrando que na planilha *dados_multivar* como as linhas contém as informações individuais de cada filhote, os valores das covariáveis referentes às ratas se repetem. As covariáveis centradas foram obtidas subtraindo dos valores originais das covariáveis, os valores médios correspondentes. Também foram criadas novas variáveis: *dose_ord* usada no ajuste dos modelos de modo que o controle passou a ser a dose referência; *fase_dose* para obter os níveis formados pelas combinações dos níveis de *fase* e *dose*; *pesomedioninhada1*, ..., *pesomedioninhada23* com os pesos médios dos filhotes por ninhada e dia e *pesomediofase-dose1*, ..., *pesomediofasedose23* com os pesos médios dos filhotes por combinação dos níveis de fase e dose e dia. A partir do arquivo no formato multivariado obteve-se o arquivo no formato univariado. Os códigos utilizados são apresentados a seguir:

```
*nova planilha para cálculo dos valores médios das covariáveis;
data exper2.medias_covar;
    set exper2.dados_multivar;
    if filhote ne 1 then delete;
run;

*calcula os valores médios das covariáveis;
proc means data=exper2.medias_covar noprint;
    var no_filhotes_nasc no_filhotes_nasc_vivos no_dias_gestacao pesorata1 u_e;
    output out=medias_covar mean=mno_filhotes_nasc mno_filhotes_nasc_vivos
            mno_dias_gestacao mpesorata1 u_e;
run;

*atualiza a planilha acrescentando as covariáveis centradas e
as novas variáveis dose_ord e fase_dose.
data exper2.dados_multivar;
    set exper2.dados_multivar;
    pesorata1_cent=pesorata1-231.0667;
    no_filhotes_cent=no_filhotes_nasc-9.95;
    no_dias_gest_cent=no_dias_gestacao-22.0333;
    if dose = 0 then dose_ord = 3;
    if dose = 25 then dose_ord = 1;
    if dose = 50 then dose_ord = 2;
```

```

    if (fase = 1 and dose = 0) then fase_dose = 1;
    if (fase = 1 and dose = 25) then fase_dose = 2;
    if (fase = 1 and dose = 50) then fase_dose = 3;
    if (fase = 2 and dose = 0) then fase_dose = 4;
    if (fase = 2 and dose = 25) then fase_dose = 5;
    if (fase = 2 and dose = 50) then fase_dose = 6;
run;
*obtem a variavel pesomedioninhada por dia;
proc means data=exper2.dados_multivar noprint;
    by u_e;
    var peso1-peso23;
    output out=ninhadas mean=pesomedioninhada1-pesomedioninhada23;
run;
*inclui a variavel pesomedioninhada no arquivo de dados;
data exper2.dados_multivar;
    merge exper2.dados_multivar ninhadas;
    by u_e;
run;
*ordena os dados de acordo com os niveis de fase e dose;
proc sort data=exper2.dados_multivar;
    by fase dose;
run;
*obtem a variavel pesomediofasedose por dia;
proc means data=exper2.dados_multivar noprint;
    by fase dose;
    var peso1-peso23;
    output out=mfase_dose mean=pesomediofasedose1-pesomediofasedose23;
run;
*inclui a variavel pesomediofasedose no arquivo de dados;
data exper2.dados_multivar;
    merge exper2.dados_multivar mfase_dose;
    drop _type_ _freq_;
    by fase dose;
run;
*coloca os dados de entrada no formato univariado;
data exper2.dados_univar;
    set exper2.dados_multivar;
    drop peso1-peso23 pesomedioninhada1-pesomedioninhada23
        pesomediofasedose1-pesomediofasedose23;

```

```
pesofilhote=peso1; idade=1; pesomedioninhada=pesomedioninhada1;
pesomediofasedose=pesomediofasedose1; output exper2.dados_univar;
pesofilhote=peso2; idade=2; pesomedioninhada=pesomedioninhada2;
pesomediofasedose=pesomediofasedose2; output exper2.dados_univar;
...
pesofilhote=peso23; idade=23; pesomedioninhada=pesomedioninhada23;
pesomediofasedose=pesomediofasedose23; output exper2.dados_univar;
run;
```

Tabela 7 – Dados de entrada do Experimento 2 – Formato univariado

obs	u_e	filhote	idade	fase	dose	peso	no. filh.	no. filh.	no. dias	peso inic.
						filhote	nasc.	n. vivos	gest.	rata
1	1	1	1	1	0	5	6	6	23	268
2	1	1	2	1	0	8	6	6	23	268
3	1	1	3	1	0	9	6	6	23	268
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
23	1	1	23	1	0	60	6	6	23	268
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
93	1	5	1	1	0	5	6	6	23	268
94	1	5	2	1	0	8	6	6	23	268
95	1	5	3	1	0	8	6	6	23	268
96	1	5	4	1	0	11	6	6	23	268
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
115	1	5	23	1	0	60	6	6	23	268
116	2	1	1	1	0	5	10	10	21	239
117	2	1	2	1	0	7	10	10	21	239
118	2	1	3	1	0	8	10	10	21	239
119	2	1	4	1	0	9	10	10	21	239
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
138	2	1	23	1	0	39	10	10	21	239
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6671	60	1	1	2	50	7	6	6	23	235
6672	60	1	2	2	50	8	6	6	23	235
6673	60	1	3	2	50	11	6	6	23	235
6674	60	1	4	2	50	11	6	6	23	235
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6693	60	1	23	2	50	58	6	6	23	235
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6763	60	5	1	2	50	7	6	6	23	235
6764	60	5	2	2	50	9	6	6	23	235
6765	60	5	3	2	50	10	6	6	23	235
6766	60	5	4	2	50	11	6	6	23	235
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6785	60	5	23	2	50	55	6	6	23	235

Como descrito no Experimento 1, o procedimento MIXED requer o formato univariado de entrada de dados. Para o ajuste dos modelos referentes ao Experimento 2 foi necessário criar as outras covariáveis relacionadas à idade: d_{tij} , d_{tij}^2 , d_{tij}^3 , $(d_{tij} - 14)_+$, $(d_{tij} - 14)_+^2$ e $\cos(5,753054 \times d_{tij})$. O código utilizado é apresentado a seguir:

```
* novas variáveis função da idade;
data exper2.dados_univar;
  set exper2.dados_univar;
  idadm1=idade-1;
  idadm1_2=idadm1*idadm1;
  idadm1_3=idadm1_2*idadm1;
  idadm1m14=max(idadm1-14,0);
  idadm1m14_2=idadm1m14*idadm1m14;
  cos_idadm1=cos(5.753054*idadm1);
run;
```

Para obter os gráficos das Figuras (11), (12), (13), (14) e (15) foram utilizadas as seguintes linhas de comandos:

```
*gráficos dos perfis de pesos individuais de filhotes e médios de ninhadas;
*atribui nomes às combinações de fase e dose;
proc format;
  value fasedosefmt 1 = "fase 1, solução salina"
    2 = "fase 1, dose=25mg/kg"
    3 = "fase 1, dose=50mg/kg"
    4 = "fase 2, solução salina"
    5 = "fase 2, dose=25mg/kg"
    6 = "fase 2, dose=50mg/kg";
run;
*perfis filhotes;
ods graphics on/ reset=all imagename="exp2_perfindfil" imagefmt=pdf
  width=6in height=5in;
ods listing gpath="c:\experimento2";
proc sgpanel data=exper2.dados_univar noautolegend;
  label pesofilhote="peso (g)";
  label dia='idade (dias)';
  panelby fase_dose /columns=3 rows=2 novarname;
  series x=dia y=pesofilhote / group=u_o lineattrs=(color=brgr pattern=1);
  colaxis values=(1,3,5,7,9,11,13,15,17,19,21,23);
```

```

    format fase_dose fasedosefmt.;
run;

*perfis médios ninhadas;

ods graphics on/ reset=all imagename="exp2_perfmninh" imagefmt=pdf
    width=6in height=5in;
ods listing gpath="c:\experimento2";
proc sgpanel data=exper2.dados_univar noautolegend;
    label pesomedioninhada="peso médio (g)";
    label dia="idade (dias)";
    panelby fase_dose /columns=3 rows= 2 novarname;
    series x=dia y=pesomedioninhada/ group=u_o lineattrs=(color=brgr pattern=1);
    series x=dia y=pesomediofasedose/ group=u_o
        lineattrs=(color=red pattern=2 thickness=3);
    colaxis values=(1,3,5,7,9,11,13,15,17,19,21,23);
    format fase_dose fasedosefmt.;
run;

*gráficos dos perfis médios nas combinações de fase e dose;

*atribui nomes às doses;

proc format;
    value dosefmt 0 = "solução salina"
        25 = "dose=25mg/kg"
        50 = "dose=50mg/kg";

run;

*atribui nomes às fases;

proc format;
    value fasefmt 1 = "1"
        2 = "2";

run;

*ordena os dados;

proc sort data=exper2.univar;
    by fase dose dia;

run;

*calcula as médias de pesos por fase dose e dia;

proc means data=exper2.dados_univar noprint;
    var pesofilhote;
    by fase dose dia;
    output out=media mean= media_peso;

run;

```

```

*exibe os perfis médios de pesos (painéis=fase);
ods graphics on/ reset=all imagename="exp2_fase" imagefmt=pdf
                    width=3.5in height=2.5in;
ods listing gpath="c:\experimento2";
proc sgpanel data=media;
    panelby fase / columns=2 rows=1 ;
    label media_peso="peso médio";
    scatter x=dia y=media_peso / group=dose;
    series x=dia y=media_peso / group=dose lineattrs=(pattern=1);
    format dose dosefmt.;
    format fase fasefmt.;
run;
ods graphics off;

*gráfico dos perfis médios de pesos por fase e dose (painéis = dose);
ods graphics on/ reset=all imagename = "exp2_dose" imagefmt=pdf
                    width=5in height=2.5in;
ods listing gpath="c:\experimento2";
proc sgpanel data=media;
    panelby dose / columns=3 rows=1 novarname;
    label media_peso="peso médio";
    scatter x=dia y=media_peso / group=fase;
    series x=dia y=media_peso / group=fase lineattrs=(pattern=1);
    format dose dosefmt.;
    format fase fasefmt.;
run;
ods graphics off;

*gráfico dos perfis de filhotes classificados por ninhada;
*ordena planilha de dados com relação a u_e;
proc sort data=exper2.dados_univar;
    by u_e;
run;

*cria arquivo de dados para localização das informações;
data exper2.dados_univar2;
set exper2.dados_univar;
by u_e;
if first.u_e then do;
    yloc=65; yloc2=55; xloc_pct=10; xloc_title=5;

```

```

    text="f.n."; text2="f.n.v.";
end;
else do;
    yloc=.; yloc2=.; xloc_pct=.;
    xloc_title=.; text="";
end;
run;

*exibe o gráfico;

ods graphics on/ reset=all imagename="exp2_perfindporninh" imagefmt=pdf
                width=8in height=8in;

ods listing gpath="c:\experimento2";

proc sgpanel data=exper2.dados_univar2 noautolegend;
    label yloc="peso (g)";
    label xloc_title="idade (dias)" u_e="rata";
    panelby u_e / columns=10 rows= 6;
    series x = dia y = pesofilhote / group=u_o lineattrs=(color=brgr pattern=1);
    scatter y=yloc x=xloc_title / markerchar=text;
    scatter y=yloc2 x=xloc_title / markerchar=text2;
    scatter y=yloc x=xloc_pct / markerchar=no_filhotes_nasc;
    scatter y=yloc2 x=xloc_pct / markerchar=no_filhotes_nasc_vivos;
    colaxis values=(1,12,23) offsetmin=0.1 offsetmax=0.1;
    rowaxis values=(0 to 65 by 5);
run;

*diagramas de dispersão dos pesos dos filhotes nos pares de idades;

*cria uma variável com nome reduzido;

data exper2.dados_multivar;
    set exper2.dados_multivar;
    p1=peso1;    p2=peso2;    p3=peso3;    p4=peso4;    p5=peso5;    p6=peso6;
    p7=peso7;    p8=peso8;    p9=peso9;    p10=peso10; p11=peso11; p12=peso12;
    p13=peso13;  p14=peso14; p15=peso15; p16=peso16; p17=peso17; p18=peso18;
    p19=peso19;  p20=peso20; p21=peso21; p22=peso22; p23=peso23;
run;

*obtem o painel com os gráficos;

ods graphics on/ reset=all imagename="exp2_cor" imagefmt=png
                width=6in height=6in;

ods listing gpath="c:\experimento2";

proc sgscatter data=exper2.dados_multivar;
    *idades até 12 dias;

```



```

matrix p1 p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12 /
      markerattrs=(color=brgr) diagonal=(histogram normal);
*idades após 12 dias;
*matrix p12 p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23/
      markerattrs=(color=brgr) diagonal=(histogram normal);
*idades ímpares;
*matrix p1 p3 p5 p7 p9 p11 p13 p15 p17 p19 p21 p23 /
      markerattrs=(color=brgr) diagonal=(histogram normal);

run;

ods graphics off;

*gráfico de superfície - exemplo covariável número de filhotes nascidos;
goptions reset=all border cback=white htext=11pt htitle=14pt device=png;
proc g3d data=exper2.dados_univar;
  label no_filhotes_nasc="no filhotes" peso_filhote="peso (g)" idade="idade (dias)";
  plot no_filhotes_nasc*idade=peso_filhote / grid yticknum=5 zticknum=5 ;
run;

```

As linhas de comandos usadas para o ajuste do modelo M2.1 são apresentadas a seguir e às referentes aos modelos M2.2 a M2.13 encontram-se no Apêndice C.

```

data exper2.dados_univar;
  set exper2.dados_univar;
  if fase = 1 then fase_ord = 2;
  if fase = 2 then fase_ord = 1;
  if dose = 0 then dose_ord = 3;
  if dose = 25 then dose_ord = 1;
  if dose = 50 then dose_ord = 2;
run;

proc format;
  value fase_ordfmt 1 = "fase2"
                  2 = "fase1";
run;

proc format;
  value dose_ordfmt 1 = "dose25"
                  2 = "dose50"
                  3 = "controle";

```

```

run;

*M2.1;
proc mixed data=exper2.dados_univar covtest order=internal; *method=ml;
  class u_e filhote fase_ord dose_ord dia_d;
  model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
         fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
         idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
         idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
         no_filhotes_cent pesoratal_cent no_dias_gest_cent /
         solution residual outpm=exper2.res_marg_m1 outp=exper2.res_cond_m1;
  random intercept idadm1 idadm1_2 / subject=u_e type=un solution v vcorr g;
  random idadm1 idadm1_2 / subject=filhote(u_e) type=un;
  repeated dia_d/subject=filhote(u_e) r rcorr;
  format fase_ord fase_ordfmt.;
  format dose_ord dose_ordfmt.;
run;

```

Os gráficos referente a análise dos resíduos apresentada nas Figuras 22 e 23 e os gráficos com os perfis médios e individuais observados e estimados (Figuras 24 e 25), foram obtidos a partir das seguintes linhas de comandos.

```

*boxplot dos resíduos em função da idade;
ods graphics on / reset=all antialiasmax=1200 imagename="boxresm7" imagefmt=jpeg
               width=3.5in height=3in;
ods listing gpath="c:\experimento2";
proc sgplot data=exper2.res_cond_m13;
  label pearsonresid="resíduos condicionais de Pearson" idadm1="idade-1 (dias)";
  vbox pearsonresid / category=idadm1;
  reflate 0 / axis=y lineattrs=(color=black);
  yaxis values=(-5.5 -4.5 -3.5 -2.5 -1.5 -0.5 0.5 1.5 2.5 3.5 4.5 5.5);
run;

*resíduos vs preditos incluindo curva suavizada;
ods graphics on / reset=all antialiasmax=6800 imagename="resvspredm7" imagefmt=jpeg
               width=3.5in height=3in;
ods listing gpath="c:\experimento2";
proc sgplot data=Luzia.res_cond_m7;
  label pearsonresid="resíduos condicionais de Pearson" pred="preditos M2.7";

```

```

scatter x=pred y=pearsonresid;
loess x=pred y=pearsonresid / lineattrs=(color=lightblue);
refline 0 / axis=y lineattrs=(color=black thickness=1);
yaxis values=(-5.5 -4.5 -3.5 -2.5 -1.5 -0.5 0.5 1.5 2.5 3.5 4.5 5.5);
run;

*gráficos para comparar os perfis de pesos médios observados e preditos;
*ordena de acordo com os níveis de fase e dose;
proc sort data=exper2.res_marg_m13; *planilha corresponde a saída de outpm do Mixed;
  by fase_dose dia;
run;
*calcula os pesos médios observados e preditos nas combinações dos níveis de fase e dose;
proc means data=exper2.res_marg_m13 noprint;
  var pred pesofilhote fase dose;
  by fase_dose dia;
  output out=media mean=media_pred media_peso fase dose;
  format fase_dose fasedosefmt.;
run;
*exibe o gráfico;
ods graphics on / reset=all imagename="exp2_perfmpredm7" imagefmt=pdf
  width=5in height=4.4in;
ods listing gpath="c:\experimento2";
proc sgpanel data=media;
  panelby fase_dose / columns=3 rows=2 novarname;
  label media_peso="peso médio" media_pred="predito";
  scatter x=dia y=media_peso / markerattrs=(symbol=circlefilled color=red)
    legendlabel="observado";
  series x=dia y=media_pred / lineattrs=(pattern=1 thickness=1 color=blue)
    legendlabel="previsto";
  format fase_dose fasedosefmt.;
run;
ods graphics off;

*gráfico de perfis de pesos observados e preditos dos filhotes
(apenas o filhote 1 de cada ninhada);
ods graphics on / reset=all imagename="exp2_perfipredm13" imagefmt=pdf
  width=10in height=9in;
ods listing gpath="c:\experimento2";
proc sgpanel data=exper2.res_cond_m13 noautolegend;

```

```
label u_o="filhote" pesofilhote="peso (g)" dia="idade (dias)";
where filhote=1;
panelby u_o / columns=10 rows=6;
scatter x=dia y=pesofilhote / group=u_o
               markerattrs=(size=4 symbol=circlefilled color=red);
series x=dia y=pred / group=u_o lineattrs=(pattern=1 thickness=1 color=bib);
colaxis values=(1,12,23);
rowaxis values=(0 to 60 by 20);
run;
ods graphics off;
```


4 RESULTADOS E DISCUSSÃO

4.1 Experimento 1

O gráfico da Figura 11 forneceu indicativo de variabilidade entre os animais do mesmo grupo, em relação ao intercepto e a inclinação. A Figura 17 apresenta os pesos individuais ao longo do tempo das fêmeas (a) e dos machos (b) e a reta de regressão de mínimos quadrados ajustada para cada animal. Os gráficos contêm também as bandas de confiança com os valores dos intervalos de 95% de confiança para a média e para uma observação futura. Observou-se que a dispersão dos pesos de cada animal com relação à reta de regressão individual é maior nos machos do que nas fêmeas, dando indicativo da necessidade de considerar matrizes de covariâncias dos erros diferentes para machos e fêmeas.

Figura 17 – Pesos individuais, retas de regressão de mínimos quadrados ordinários e intervalos de 95 % de confiança para a média e para uma observação individual - Experimento 1

A Figura 18 apresenta a matriz com os valores das variâncias e covariâncias amostrais dos pesos nos pares de idades. No triângulo superior estão as informações das fêmeas e no triângulo inferior as informações dos machos. Foi utilizado o gradiente de cores: ciano (600), azul (800), magenta (1000) e vermelho (1200), evidenciando que o grupo de machos apresentou variâncias e covariâncias amostrais maiores comparados com o grupo de

fêmeas.

Os principais resultados dos modelos ajustados estão apresentados nas Tabelas 8, 9 e 10 e correspondem, respectivamente, às estatísticas de ajuste, às estimativas de máxima verossimilhança restrita e respectivos erros padrões dos parâmetros de efeitos fixos e os níveis descritivos (p-valores) dos testes tipo III de efeitos fixos.

A necessidade de considerar variâncias residuais diferentes ao longo do tempo ficou evidente, visto que os valores das estatísticas de ajuste dos modelos com variâncias constantes ou diferentes somente em relação ao sexo, foram maiores do que 7200 unidades. Ao incluir variâncias diferentes ao longo do tempo esses valores reduziram consideravelmente (modelos M1.7, M1.8, M1.11, M1.12, M1.13). Observou-se também que ao considerar estruturas de covariâncias residuais diferentes para as fêmeas e os machos, os ajustes melhoraram.

como pode ser notado, por exemplo, comparando os modelos M1.7 com M1.8 e M1.11 com M1.12. Da mesma forma, aumentando a complexidade dessas estruturas os resultados ficaram ainda melhores, como pode ser observado comparando, por exemplo, os modelos M1.7, M1.11 e M1.13 e por essa razão testou-se a estrutura Ante1 diferente para machos e fêmeas, contendo 114 parâmetros, cuja convergência não foi alcançada. Assim sendo, o modelo que apresentou os menores valores das estatísticas de ajuste foi o M1.12.

De acordo com os resultados da Tabela 9 observou-se uma diferença de até 7,37g entre as estimativas dos parâmetros, obtida comparando os coeficientes de $d3 \times m$ nos modelos M1.10 e M1.13. Com relação aos erros padrões, a diferença máxima foi de 0,586g, obtida comparando os erros padrões dos estimadores dos coeficientes $d1 \times m$, $d3 \times m$ e $d5 \times m$ nos modelos (M1.1, M1.2) com M1.12.

Os níveis descritivos dos testes tipo III de efeitos fixos, apresentados na Tabela 10, foram obtidos pelo método de máxima verossimilhança e aproximação de Kenword Roger para o número de graus de liberdade do denominador. De acordo com os resultados, observou-se que os modelos assumindo uma estrutura de simetria composta para $\mathbf{R}_i$, levaram a conclusões distintas com relação aos efeitos fixos, comparadas com a do modelo de melhor ajuste. Nos primeiros modelos, por exemplo, o efeito da interação dia $\times$ dose $\times$ sexo ao longo do tempo foram altamente significativos. Também pode ser observado que todos os modelos assumindo variâncias diferentes com relação ao tempo apresentaram as mesmas conclusões com relação aos efeitos fixos.

A Figura 19 apresenta o histograma, o gráfico normal de probabilidades e o diagrama de dispersão dos resíduos condicionais de Pearson vs preditos dos modelos M1.1, M1.3, M1.4, M1.6, M1.7, M1.8, M1.10, M1.11, M1.12 e M1.13. Com relação às médias dos resíduos condicionais dos modelos M1.7, M1.9, M1.10, M1.11, M1.12 e M1.13, observou-se uma pequena diferença com relação ao valor 0, sendo, respectivamente, iguais a -0,01; -0,01; 0,02; 0,01; 0,04 e -0,01.

Alguns dos modelos testados apresentaram problemas de convergência ou impossibilidade de obtenção de uma matriz hessiana positiva definida, mas na maioria dos casos a simples alteração dos valores iniciais dos parâmetros resolveu o problema. Isso ocorreu com os modelos M1.5, M1.8, M1.11, M1.12 e M1.13. Houve também alguns casos de ajustes que apresentaram anomalias nos resíduos, devido à provável convergência para algum máximo local. A Figura 20 apresenta os resíduos condicionais de dois modelos ilustrando o

problema. Os gráficos correspondem aos modelos com estrutura ARH1 comum para machos e fêmeas e com estrutura Ante1 diferente para cada sexo. Os valores de $-2 \times \log(\text{REML})$, AIC, AICC e BIC dos ajustes foram respectivamente: (7005,0; 7071,0; 7073,1; 7126,8) e (6618,2; 6850,2; 6876,7; 7046,1). No caso da estrutura ARH1, foi possível resolver o problema alterando os valores iniciais, que corresponde ao modelo M1.11, entretanto, o mesmo não ocorreu no caso da estrutura Ante1.

A Figura 21 apresenta o gráfico com os perfis individuais observados e preditos a partir de M1.12.

No modelo M1.2 (melhor ajuste), de acordo com os resultados da Tabela 10, foram significativos os efeitos de sexo no intercepto e o ganho de peso diário (dia), havendo, portanto, um ganho de peso comum dos animais ao longo do experimento e diferença entre os pesos médios iniciais das fêmeas e dos machos. Desta forma, concluiu-se que o nanocomposto não provocou alterações diferentes nos ganhos de pesos de ratos machos e fêmeas.

Tabela 8 – Características dos modelos ajustados e estatísticas de ajuste - Experimento 1

Modelo	nº parâm.	covar. ef. aleat.	estrutura matriz $\mathbf{R}_i$	$-2\log \text{REML}$	AIC	AICC	BIC
M1.1	2	σ_{u0}^2	$\sigma^2 \mathbf{I}_{n_i}$	7562,4	7566,4	7566,4	7569,8
M1.2	3	$\sigma_{u0_F}^2, \sigma_{u0_M}^2$	$\sigma^2 \mathbf{I}_{n_i}$	7562,3	7568,3	7568,3	7573,4
M1.3	3	σ_{u0}^2	$\sigma_F^2 \mathbf{I}_{n_i}, \sigma_M^2 \mathbf{I}_{n_i}$	7408,5	7414,5	7414,5	7419,5
M1.4	3	$\sigma_{u0}^2, \sigma_{u1}^2$	$\sigma^2 \mathbf{I}_{n_i}$	7457,0	7463,0	7463,0	7468,1
M1.5	4	$\sigma_{u0}^2, \sigma_{u1_F}^2, \sigma_{u1_M}^2$	$\sigma^2 \mathbf{I}_{n_i}$	7454,7	7462,7	7462,8	7469,5
M1.6	4	$\sigma_{u0}^2, \sigma_{u1}^2$	$\sigma_F^2 \mathbf{I}_{n_i}, \sigma_M^2 \mathbf{I}_{n_i}$	7287,2	7295,2	7295,3	7302,0
M1.7	31	$\sigma_{u0}^2, \sigma_{u1}^2$	D	7126,3	7188,3	7190,1	7240,7
M1.8	60	$\sigma_{u0}^2, \sigma_{u1}^2$	D _F , D _M	6971,5	7091,5	7098,3	7192,8
M1.9	4	$\sigma_{u0}^2, \sigma_{u1}^2$	AR1	7363,0	7371,0	7371,0	7377,8
M1.10	6	$\sigma_{u0}^2, \sigma_{u1}^2$	AR1 _F , AR1 _M	7194,6	7206,6	7206,7	7216,7
M1.11	32	$\sigma_{u0}^2, \sigma_{u1}^2$	ARH1	7024,4	7088,4	7090,3	7142,4
M1.12	62	$\sigma_{u0}^2, \sigma_{u1}^2$	ARH1 _F , ARH1 _M	6859,4	6983,4	6990,6	7088,1
M1.13	59	$\sigma_{u0}^2, \sigma_{u1}^2$	Ante1	6882,5	7000,5	7007,0	7100,1

Tabela 9 – Estimativas e erros padrões dos parâmetros de efeitos fixos dos modelos ajustados

Efeitos	M1.1	M1.3	M1.4	M1.6	M1.7	M1.8	M1.10	M1.11	M1.12	M1.13
intercepto	254,84 (-11,49)	254,84 (11,47)	254,84 (11,31)	254,84 (11,33)	254,16 (11,25)	254,23 (11,28)	254,80 (11,37)	253,75 (11,29)	254,11 (11,34)	252,20 (11,20)
dose1 (d1)	0,31 (16,25)	0,31 (16,22)	0,31 (15,99)	0,31 (16,03)	-0,37 (15,91)	-0,67 (15,96)	0,46 (16,07)	0,07 (15,96)	-1,00 (16,03)	0,36 (15,83)
dose3 (d3)	-9,74 (16,25)	-9,74 (16,22)	-9,74 (15,99)	-9,74 (16,03)	-9,55 (15,91)	-9,62 (15,96)	-9,66 (16,07)	-9,50 (15,96)	-9,96 (16,03)	-10,10 (15,83)
dose5 (d5)	-7,63 (16,25)	-7,63 (16,22)	-7,63 (15,99)	-7,63 (16,03)	-8,32 (15,91)	-8,14 (15,96)	-7,53 (16,07)	-7,91 (15,96)	-8,15 (16,03)	-9,12 (15,83)
macho (m)	203,37 (16,25)	203,37 (16,25)	203,37 (15,99)	203,37 (16,05)	208,12 (15,91)	207,91 (15,97)	202,08 (16,13)	208,27 (15,96)	207,82 (16,05)	211,79 (15,83)
d1 × m	-24,00 (22,98)	-24,00 (22,98)	-24,00 (22,62)	-24,00 (22,70)	-22,04 (22,49)	-21,92 (22,58)	-24,26 (22,81)	-22,23 (22,58)	-21,01 (22,70)	-20,80 (22,39)
d3 × m	5,42 (22,98)	5,42 (22,98)	5,42 (22,62)	5,42 (22,70)	1,16 (22,49)	1,24 (22,58)	6,63 (22,81)	1,08 (22,58)	1,43 (22,70)	-0,74 (22,39)
d5 × m	23,36 (22,98)	23,36 (22,98)	23,36 (22,62)	23,36 (22,70)	21,10 (22,49)	20,75 (22,58)	24,22 (22,81)	21,46 (22,58)	21,93 (22,70)	21,25 (22,39)
dia	0,14 (0,06)	0,14 (0,04)	0,14 (0,12)	0,14 (0,11)	0,19 (0,12)	0,19 (0,12)	0,14 (0,10)	0,23 (0,12)	0,20 (0,11)	0,28 (0,13)
dia × d1	-0,11 (0,08)	-0,11 (0,06)	-0,11 (0,17)	-0,11 (0,16)	-0,12 (0,17)	-0,09 (0,17)	-0,12 (0,14)	-0,18 (0,17)	-0,11 (0,16)	-0,13 (0,18)
dia × d3	0,15 (0,08)	0,15 (0,06)	0,15 (0,17)	0,15 (0,16)	0,12 (0,17)	0,12 (0,17)	0,15 (0,14)	0,09 (0,17)	0,12 (0,16)	0,15 (0,18)
dia × d5	0,01 (0,08)	0,01 (0,06)	0,01 (0,17)	0,01 (0,16)	0,01 (0,17)	0,01 (0,17)	0,00 (0,14)	-0,03 (0,17)	0,00 (0,16)	0,07 (0,18)
dia × m	0,38 (0,08)	0,38 (0,08)	0,38 (0,17)	0,38 (0,17)	0,14 (0,17)	0,15 (0,17)	0,44 (0,16)	0,11 (0,17)	0,14 (0,16)	-0,03 (0,18)
dia × d1 × m	-0,04 (0,11)	-0,04 (0,11)	-0,04 (0,24)	-0,04 (0,24)	-0,11 (0,24)	-0,14 (0,24)	-0,02 (0,23)	-0,09 (0,24)	-0,18 (0,23)	-0,18 (0,26)
dia × d3 × m	-0,48 (0,11)	-0,48 (0,11)	-0,48 (0,24)	-0,48 (0,24)	-0,25 (0,24)	-0,25 (0,24)	-0,53 (0,23)	-0,23 (0,24)	-0,26 (0,23)	-0,15 (0,26)
dia × d5 × m	-0,22 (0,11)	-0,22 (0,11)	-0,22 (0,24)	-0,22 (0,24)	-0,07 (0,24)	-0,07 (0,24)	-0,27 (0,23)	-0,10 (0,24)	-0,14 (0,23)	-0,09 (0,26)

Tabela 10 – Níveis descritivos dos testes tipo III de efeitos fixos dos modelos ajustados

Efeitos	M1.1	M1.3	M1.4	M1.6	M1.7	M1.8	M1.10	M1.11	M1.12	M1.13
dose	0,5281	0,5280	0,5147	0,5177	0,5572	0,5490	0,5117	0,5495	0,5275	0,5870
sexo	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001
dose × sexo	0,2435	0,2434	0,2310	0,2338	0,3121	0,3248	0,2169	0,3044	0,3243	0,3307
dia	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001	<,0001
dia× dose	0,1434	0,1436	0,7558	0,7421	0,4397	0,5031	0,6151	0,2642	0,3084	0,1310
dia× sexo	<,0001	<,0001	0,0281	0,0248	0,7392	0,6844	0,0069	0,9238	0,9629	0,1556
dia× dose × sexo	<,0001	<,0001	0,2011	0,1842	0,7856	0,7631	0,0874	0,8033	0,7211	0,9036

Figura 19 – Histograma, gráfico normal de probabilidades e de dispersão dos resíduos condicionais de Pearson - Experimento 1

Figura 19 – Histograma, gráfico normal de probabilidades e de dispersão dos resíduos condicionais de Pearson - Experimento 1 (cont.)

Figura 20 – Histograma, gráfico normal de probabilidades e de dispersão dos resíduos condicionais de Pearson de ajustes que convergiram para um máximo local

Figura 21 – Perfis individuais de pesos e retas individuais ajustadas a partir do modelo M1.12

4.2 Experimento 2

Nos ajustes dos modelos do Experimento 2 foram considerados os valores iniciais e critérios de convergência padrões do procedimento MIXED. Os modelos polinomiais de grau 3 com efeitos aleatórios de rata no coeficiente de grau 3 não convergiram mesmo considerando erros independentes. Também não foi possível ajustar modelos polinomiais de grau 3 contendo conjuntamente efeito aleatório de filhote e estruturas mais complexas para a matriz $\mathbf{R}_i$, pois ou não convergiram ou não foi possível obter uma matriz hessiana positiva definida. De acordo com o gráfico apresentado na Figura 12, optou-se por excluir o efeito aleatório no intercepto de filhote, dada a pequena variabilidade dos interceptos de filhotes com relação aos interceptos das curvas médias por ninhada e a dificuldade de convergência.

A partir da análise de resíduos do modelo M2.1 (gráficos (a) e (b) da Figura 22) observou-se um comportamento periódico dos mesmos, violando a suposição de aleatoriedade e simetria ao redor de zero. A sazonalidade dos mesmos ficou bem evidente no diagrama de dispersão dos resíduos condicionais em função da idade (Figura 22(b)), sendo semelhante em todas as combinações dos níveis dos fatores *fase* e *dose*. Essa sazonalidade pôde ser modelada por meio da covariável $\cos(5,753054 \times idade)$, sendo o valor da constante a obtido como descrito na seção 3.2.2. Os valores dessa covariável são apresentados em azul claro na Figura 22(b). Dessa forma, ao incluir essa covariável, assim como o efeito aleatório de rata no seu coeficiente (M2.3), os resíduos condicionais obtidos não mais violaram a suposição de aleatoriedade como pode ser observado na 22(c).

A melhora do ajuste com a inclusão da covariável $\cos(5,753054 \times idade)$, considerando erros independentes, também pôde ser observada por meio das estatísticas de ajuste dos modelos, apresentadas na Tabela 11. Verificou-se, por exemplo, uma redução de aproximadamente 1250 unidades nos valores de AIC e de BIC com a inclusão dessa covariável na parte fixa do modelo (M2.2). Ao incluir também o efeito aleatório no coeficiente dessa covariável (M2.3) a redução foi em torno de 210 unidades. Até mesmo no caso da estrutura mais geral (Ne) para a matriz de covariâncias dos erros (M2.10), a inclusão dessa covariável (M2.9) reduziu em aproximadamente 190 unidades os valores de AIC e BIC, sendo a redução em torno de 100 unidades após a inclusão também do efeito aleatório no coeficiente dessa covariável (M2.8).

O ajuste do modelo considerando a matriz NE, foi bastante demorado, sendo superior a 30 horas de processamento. Pelo fato de cada filhote ser observado em 23 tempos

torna-se necessário, nesse caso, estimar $(23 \times 24)/2 = 276$ parâmetros de covariâncias.

Comparando os modelos polinomiais de grau 3, M2.1 à M2.10, por meio do teste da razão de verossimilhanças (Tabela 12) observou-se que o modelo de melhor ajuste foi o M2.8, que apresenta estrutura NE, a covariável $\cos(5,753054 \times idade)$ e o efeito aleatório de rata no coeficiente dessa covariável. Entretanto, considerando o critério BIC, o melhor ajuste foi do modelo M2.7, uma vez que nesse critério, considera-se uma penalização maior para o aumento do número de parâmetros no modelo. Pelo critério AIC e teste da razão de verossimilhanças o modelo M2.8 foi a segunda melhor opção. Portanto, pelo princípio da parcimônia o modelo M2.7 foi escolhido dentre os modelos M2.1 à M2.10.

Na Tabela 11 também são apresentadas as estatísticas de ajuste dos modelos *spline* (M2.12) e quadrático por partes (M2.13), ambos com ponto de mudança na idade média de abertura dos olhos dos filhotes. O modelo M2.13 apresentou melhor ajuste de acordo com os critérios apresentados. A análise de resíduos, apresentada na Figura 23, evidenciou a qualidade do ajuste do modelo M2.13, uma vez que os resíduos condicionais apresentaram comportamento aleatório em torno de 0.

A Figura 24 possibilitou visualizar a qualidade dos ajustes com relação aos perfis médios nas combinações de *fase* e *dose* e a Figura 25 mostrou que os perfis individuais (um de cada ninhada), ainda que bem diferentes entre si, dependendo das características das ninhadas foram bem descritos pelo modelo M2.13.

De acordo com o modelo M2.13, fixado o nível de significância em 5%, foram significativos os efeitos da interação *fase* $\times$ *dose* ($p=0,0003$), d_{tij} ($p<0,0001$), d_{tij}^2 ($p<0,0001$), $(fase \times dosed_{tij} - 14)_+$ ($p=0,0116$), $(d_{tij} - 14)_+^2$ ($p<0,0001$), T_i ($p=0,0332$), $T_i \times d_{tij}$ ($p<0,0001$) e P_i ($p=0,0038$).

Tabela 11 – Estatísticas de ajuste dos modelos selecionados - Experimento 2

	nº parâm.	nº parâm.	-2log veros.	AIC	BIC
modelo	ef. fixos	covar.	MVR (MV)	MVR (MV)	MVR (MV)
M2.1	27	10	(23363)	(23437)	(23515)
M2.2	28	10	22309 (22108)	22329 (22184)	22350 (22263)
M2.3	28	14	22082	22110	22139
M2.4	28	15	21632	21662	21693
M2.5	28	34	21103	21171	21242
M2.6	28	55	20944	21054	21169
M2.7	28	55	20768	20878	20993
M2.8	28	286	20013	20585	21184
M2.9	28	282	20128 (19926)	20692 (20546)	21283 (21195)
M2.10	27	282	(20121)	(20739)	(21387)
M2.11	29	55	20732 (20520)	20842 (20688)	20957 (20864)
M2.12	28	58	20638 (20483)	20754 (20655)	20876 (20835)
M2.13	34	63	20153 (19987)	20279 (20180)	20411 (20384)

Os valores entre parênteses foram obtidos considerando a estimação por MV.

Tabela 12 – Comparação dos modelos de regressão polinomial por meio do teste da razão de verossimilhanças

modelos		método	diferença	diferença	dist. assint.	nível
aninhado	referência	estimação	nº parâm.	-2log veros.	estat. teste	descritivo
M2.1	M2.2	MV	1	1.255	χ_1^2	<,0001
M2.2	M2.3	MVR	4	227	$0,5\chi_3^2 + 0,5\chi_4^2$	<,0001
M2.3	M2.4	MVR	1	450	χ_1^2	<,0001
M2.4	M2.5	MVR	19	529	$0,5\chi_{17}^2 + 0,5\chi_{19}^2$	<,0001
M2.5	M2.6	MVR	21	159	χ_{21}^2	<,0001
M2.6	M2.8	MVR	231	931	χ_{231}^2	<,0001
M2.7	M2.8	MVR	231	755	χ_{231}^2	<,0001
M2.8	M2.9	MVR	4	115	χ_4^2	<,0001
M2.9	M2.10	MV	1	195	χ_1^2	<,0001

(a) Resíduos vs preditos do modelo M2.1

(b) Resíduos vs idade do modelo M2.1 e em azul claro os valores da covariável sazonal

(c) Resíduos vs idade do modelo M2.3 que inclui a covariável sazonal na parte fixa e aleatória

Figura 22 – Gráficos mostrando a necessidade da covariável sazonal nos modelos lineares mistos polinomiais de grau 3 na idade - Experimento 2

Figura 23 – Gráfico normal de probabilidades, diagramas de caixas em função das idades e de dispersão em função dos preditos, dos resíduos condicionais de Pearson de modelos ajustados a partir das três abordagens

(a) Modelo M2.7

(b) Modelo M2.11

(c) Modelo M2.12

(d) Modelo M2.13

Figura 24 – Perfis médios de pesos observados e preditos de alguns modelos nas combinações de *fase* e *dose*

Figura 25 – Gráficos com os perfis individuais de pesos observados e preditos - Modelo M2.13

5 CONSIDERAÇÕES FINAIS

A revisão bibliográfica das técnicas usadas na análise de dados longitudinais em estudos toxicológicos revelou que a maioria não leva em conta a correlação que existe entre os valores observados ao longo do tempo na mesma unidade experimental. Por meio da comparação de alguns modelos pôde ser verificado que a especificação incorreta da matriz de covariâncias leva a resultados pouco confiáveis.

O modelo misto de regressão linear com matrizes de covariâncias dos erros diferentes para machos e fêmeas, mostrou-se apropriado para comparar o ganho de peso de ratos adultos, nos testes de toxicidade por doses repetidas e levou em conta a variabilidade biológica dos animais e a correlação entre os pesos observados no mesmo animal ao longo do tempo.

No experimento de toxicidade reprodutiva, o modelo linear misto possibilitou modelar as correlações entre os pesos de filhotes de mesma rata e entre os pesos ao longo do tempo no mesmo filhote. O modelo também levou em conta o aumento da variabilidade ao longo do tempo e possibilitou especificar explicitamente os perfis de pesos nos níveis da estrutura hierárquica.

O crescimento dos filhotes depende do tamanho da ninhada, pois quanto menor o tamanho da ninhada mais rápido é o crescimento dos filhotes. A covariável número de filhotes nascidos mostrou-se altamente significativa para explicar parte da variação dos pesos, devendo ser incluída no modelo.

Observou-se que o crescimento dos filhotes torna-se mais acelerado após a abertura dos olhos, que ocorre em média aos 15 dias de idade, possivelmente devido ao fato de que nos primeiros dias de vida ocorrem também outras mudanças físicas, como crescimento de penugens e pelos, descolamento das orelhas e abertura dos ouvidos, acarretando num gasto maior de energia. Dessa forma, fixar o ponto de mudança aos 15 dias, nos modelos quadrático por partes e spline, fez sentido também do ponto de vista prático.

Modelar o crescimento dos filhotes usando curvas com mudanças suaves, como é o caso do polinômio de grau 3 e do *splines* com um único ponto de mudança na idade média de abertura dos olhos, não é o mais apropriado pois acarreta em resíduos condicionais com comportamento cíclico, violando a suposição de aleatoriedade em torno de 0.

No caso do modelo misto polinomial de grau 3, a inclusão de uma covariável com comportamento sazonal, obtida com base no comportamento cíclico dos resíduos, me-

lhorou a qualidade dos ajustes. Entretanto, a não praticidade na obtenção dos valores dessa covariável e a falta de interpretação prática dos parâmetros do modelo desfavoreceram a escolha por essa metodologia.

Os modelos com um ponto de mudança na idade média de abertura dos olhos, além de fornecer melhor ajuste, comparado com o modelo polinomial de grau 3, facilitam a comparação entre os perfis médios dos filhotes nas diferentes doses, tornando possível comparar o ganho de peso e o aumento no ganho diário de peso (aceleração no crescimento) em cada uma das fases: anterior e posterior ao ponto de mudança, que são as informações de interesse dos pesquisadores da área.

O modelo linear misto de regressão quadrática por partes, com ponto de mudança aos 15 dias, foi o que melhor descreveu o crescimento dos filhotes (perfis de pesos) durante o período de lactação, dentre as abordagens comparadas. Nesse modelo a restrição de suavidade da curva é relaxada.

REFERÊNCIAS

- ABIZAID, A. et al. Effect of metabolic fuel availability on fertility varies with reproductive state. **Physiology & Behavior**, v. 74, n. 1-2, p. 77-83, 2001. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/11564455>>. Acesso em: 1 jun. 2014.
- ANDRADE, F. M. ; JUDICE, L. F.; CARDOSO, G. P.; CISNE, R.; RAMOS, C. F.; BABINSKI, M. A. Maternal malnutrition during lactation in Wistar rats: effects on elastic fibers of the extracellular matrix in the trachea of offspring. **Jornal Brasileiro de Pneumologia**, Distrito Federal, v. 38, n. 5, p. 588-594, 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23147051>>. Acesso em: 1 jun. 2014.
- ARZU, E.; SÜLOGLU, A. K.; SELMANOGLU, G. Vinclozolin exposure throughout pregnancy and its developmental toxicity. **Toxicology Research**, Lucknow, n. 1, p. 375-383, 2014. Disponível em: <<http://pubs.rsc.org/en/content/articlepdf/2014/tx/c4tx00037d>>. Acesso em: 1 set. 2014.
- ATKINSON, A. C. **Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis**. New York: Oxford University Press, 1987. 282p. (Oxford Statistical Science, 1).
- BARBIN, D. **Planejamento e análise estatística de experimentos agrônomicos**. Arapongas: Midas, 2003. 194 p.
- BATES, D. M. **Examples from multilevel software comparative reviews**. R Development Core Team. Disponível em: <<http://cran.r-project.org/web/packages/mlmRev/vignettes/MlmSoftRev.pdf>>. Acesso em: 1 nov. 2011.
- BATES, D. M. **lme4: mixed-effects modeling with R**. New York: Springer, 2010. 131p. Disponível em: <<http://lme4.r-forge.r-project.org/book/>>. Acesso em: 1 nov. 2011.
- BELSLEY, D. A.; KUH, E.; WELSCH, R. E. **Regression diagnostics: identifying influential data and sources of collinearity**. New Jersey: John Wiley & Sons, 2004. 292 p.
- BRENNAN, K. A.; KAUFMAN, S; REYNOLDS, S. W.; MCCOOK, B. T.; KAN, G.; CHRISTIAENS, I.; SYMONDS, M. E.; OLSON, D. M. Differential effects of maternal nutrient restriction through pregnancy on kidney development and later blood pressure control in the resulting offspring. **American Journal of Physiology. Regulatory, Integrative and Comparative Physiology**, Bethesda, v. 295, n. 1, p. R197-205, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18480243>>. Acesso em: 1 jun. 2014.
- BRUMBACK, B. A.; RICE, J. A. Smoothing Spline Models for the Analysis of Nested and Crossed Samples of Curves. **Journal of the American Statistical Association**, Alexandria, v. 93, n. 443 , p. 961-976, 1998 . Disponível em <<http://www.jstor.org/stable/2669837>>. Acesso em: 1 jun. 2013.
- CAPORALI, P. et al. Pre-reproductive maternal enrichment influences offspring developmental trajectories: motor behavior and neurotrophin expression. **Frontiers in Behavioral Neuroscience**, Lausanne, v. 8, art. 195 p. 1-14, 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24910599>>. Acesso em: 1 ago. 2014.

CHRISTENSEN, R.; PEARSON, L. M.; JOHNSON, W. Case-deletion diagnostics for mixed models. **Technometrics**, Alexandria, v. 34, n. 1, p. 38-45, 1992. Disponível em <<http://www.jstor.org/stable/1269550>>. Acesso em: 1 dez. 2011.

COMBINED repeated dose toxicity study with the reproduction/developmental toxicity screening test. Paris: OECD Publishing, n. 422, 1996. 14p. (OECD Guidelines for the testing of chemicals, 4). Disponível em: <http://www.oecd-ilibrary.org/environment/test-no-422-combined-repeated-dose-toxicity-study-with-the-reproduction-developmental-toxicity-screening-test_9789264070981-em>. Acesso em: 1 jun. 2013.

COOK, D. R.; WEISBERG, S. **Residuals and influence in regression**. Great Britain: Chapman & Hall, 1982. 230 p. (Monographs on Statistics and Applied Probability).

DAVIDIAN, M.; GILTINAN, D. M. **Nonlinear models for repeated measurement data**. New York: Chapman & Hall, 1995. 359 p. (Monographs on Statistics and Applied Probability, 62).

DEMÉTRIO, C. G. B. **An introduction to mixed models in agriculture**. Arquivo: MixedModelsGembloux.pdf, Disponível em: <http://www.lce.esalq.usp.br/arquivos/arquivos.php?dir=aulas/2011/LCE5805>. Acesso em: 20 maio 2011.

DEMÉTRIO, C. G. B.; ZOCCHI, S. S. **Modelos de regressão**. 2010. 220 p.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society**, v. 39, n. 1, p. 1-38, 1977. (Methodological, B). Disponível em <<http://links.jstor.org/sici?sici=0035-9246%281977%2939%3A1%3C1%3AMLFIDV%3E2.0.CO%3B2-Z>>. Acesso em: 1 dez. 2011.

DUMORTIER, O.; THEYS, N.; AHN, M. T.; REMACLE, C.; REUSENS, B. Impairment of rat fetal beta-cell development by maternal exposure to dexamethasone during different time-windows. **PLoS One**, San Francisco, v. 6, n. 10, p.1-10, 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21991320>>. Acesso em: 1 jun. 2014.

DURBÁN, M.; Harezlak, J.; Wand, M. P.; Carroll, R. J. Simple fitting of subject-specific curves for longitudinal data. **Statistics in Medicine**, Hoboken, v. 24, n. 8, p. 1153-1167, 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15568201>>. Acesso em: 1 jun. 2013.

FITZMAURICE, G.; DAVIDIAN, M.; VERBEKE, G.; MOLENBERGHS, G. **Longitudinal data analysis**. Boca Raton: Chapman & Hall, 2009. 619 p. (Handbooks of Modern Statistical Methods).

FITZMAURICE, G. M.; LAIRD, N. M.; WARE, J. H. **Applied longitudinal analysis**. 2nd ed. Hoboken: John Wiley & Sons, 2011. 701 p.

GALLANT, A. R.; FULLER, W. A. Fitting segmented polynomial regression models whose join points have to be estimated. **Journal of the American Statistical Association**, Alexandria, v. 68, n. 341, p. 144-147, 1973. Disponível em <<http://www.jstor.org/stable/2284158>>. Acesso em: 1 jun. 2013.

- GILMOUR, A. R.; THOMPSON, R.; CULLIS, B. R. Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. **Biometrics**, Washington, v. 51, n. 4, p. 1440-1450, 1995. Disponível em: < <http://www.jstor.org/>>. Acesso em: 15 dez. 2011.
- GOLDSTEIN, H. **Multilevel statistical models**. 3rd ed. London: Arnold, v. 3. 2003. 253 p. (Kendall's Advanced Theory of Statistics).
- GOODNIGHT, J. H. A Tutorial on the sweep operator. **American Statistical Association**, Alexandria, v. 33, n. 3, p.149-158, 1979. Disponível em: <<http://www.jstor.org/stable/2683825>>. Acesso em: 1 dez. 2011.
- HENDERSON, C. R. Estimation of genetic parameters. **The Annals of Mathematical Statistics**, Beachwood, v. 21, p. 309-310, 1950.
- JIANG, J. **Linear and generalized linear mixed and their applications**. New York: Springer, 2007. 257 p.
- LAIRD, N.M.; WARE, J.H. Random-effects models for longitudinal data. **Biometrics**, Washington, v. 38, n. 4, p. 963-974, 1982. Disponível em: <<http://www.jstor.org/stable/2529876>>. Acesso em: 15 jan. 2010.
- LITTELL, R.C.; MILLIKEN, G. A.; STROUP, W. W.; WOLFINGER, R. D. **SAS for Mixed Models**, 2nd ed., North Carolina: SAS Institute, 2006. p. 814.
- MCCULLOCH, C. E. **Generalized linear mixed models**. Beachwood: Institute of Mathematical Statistics, 2003. 84 p.
- MAZZANTI, G.; SOTTO, A. D.; FRANCHITTO, A.; MASTRANGELO, S.; PEZZELLAC, M.; VITALONE, A.; MAMMOLA, C. L. Effects of Cimicifuga racemosa extract on liver morphology and hepatic function indices. **Phytomedicine**, Sidney, v. 15, n. 11, p. 1021-1024, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18434119>>. Acesso em: 1 jun. 2014.
- MONTGOMERY, D. C. **Design and analysis of experiments**. 7th ed. New York: John Wiley & Sons, 2008. 680 p.
- PETERSEN, K. B.; PETERSEN, M. S. **The Matrix Cookbook**. 2006. 56p. Disponível em: <http://www.mit.edu/~wingated/stuff_i_use/matrix_cookbook.pdf>. Acesso em 1 dez. 2013.
- PIEPHO, H.P.; OGUTU, J.O. Inference for the break point in segmented regression with application to longitudinal data. **Biometrical Journal**, Weinheim, v. 45, n. 5, p. 591-601, 2003. Disponível em <<http://onlinelibrary.wiley.com/doi/10.1002/bimj.200390035/>>. Acesso em: 1 jun. 2013.
- PIMENTEL-GOMES, F. **Curso de estatística experimental**. 15 ed. Piracicaba: Fealq, 2009. 451 p.
- PRINCIPLES on good laboratory practice. Paris: OECD Publishing, 1998. 41 p. (OECD Principles of good laboratory practice and compliance monitoring, 1). Disponível em: <http://www.oecd-ilibrary.org/environment/oecd-principles-on-good-laboratory-practice_9789264078536-en>. Acesso em: 1 jun. 2013.
- PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer, 2000. 528 p. (Statistics and Computing).

RAUDENBUSH, S. W.; BRYK, A. S. **Hierarchical linear models: applications and data analysis methods**. 2nd ed. London: Sage Publications, 2002. 485 p. (Advanced Quantitative Techniques in the Social Sciences Series, 1).

RAUDENBUSH, S. W.; BRYK, A. S.; CHEONG, Y. F.; CONGDON, R. T. J. **HLM 6: hierarchical linear and nonlinear modeling**. Lincolnwood: Scientific Software International, 2004. 297 p.

ROBBINS, K. R.; SAXTON, A. M.; SOUTHERN, L. L. Estimation of nutrient requirements using broken-line regression analysis. **Jornal of Animal Science**, Champaign, v. 84, suppl., p. E155-E165, 2006. Disponível em <<http://www.ncbi.nlm.nih.gov/pubmed/16582088>>. Acesso em: 1 jun. 2013.

RUPPERT, D; WAND, M. P. ; CARROLL, R. J. **Semiparametric regression**. New York: Cambridge University Press, 2003. 386 p.

SANTOS, J. A.; ARRUDA, A.; CARDOSO, C. A. L.; VIEIRA, M. C.; PICCINELLI, A. C.; AQUINO, D. F. S. KASSUYA, C. A. L.; ARENA, A. C. Subacute and reproductive oral toxicity assessment of the hydroethanolic extract of jacaranda decurrens roots in adult male rats. **Evidence-Based Complementary and Alternative Medicine**, Nasr City, v. 2013, p. 1-6, 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24348699>>. Acesso em: 1 jun. 2014.

SAS Institute Inc. **SAS OnlineDoc®9.2**. 2009. Disponível em: <<http://support.sas.com/index.html>>. Acesso em: 1 nov. 2011.

SCHABENBERGER, O. Mixed model influence diagnostics. **Statistics and data analysis**, North Carolina, SAS Institute, n. 189-29, p. 1-17, 2004. (SUGI, 29). Disponível em: <<http://www2.sas.com/proceedings/sugi29/18929.pdf>>. Acesso em 1 dez. 2011.

SEARLE S. R.; CASELLA, G.; McCULLOCH, C. E. **Variance components**. New Jersey: John Wiley & Sons, 2006. 501 p.

SILVA, V. A. J.; AMORIM, M. J. A. A. L.; AMORIM, A. A. J.; PINTO, C. F.; DEIRÓ, T. B. J.; OLIVEIRA, J. R. M.; PEIXOTO, C. A.; CASTRO, R. M. Neonatal administration of fluoxetine decreased final sertoli cell number in Wistar rats. **International Journal of Morphology**, Temuco, v. 26, n. 1, p. 51-62, 2008. Disponível em: <<http://www.scielo.cl/pdf/ijmorphol/v26n1/art09.pdf>>. Acesso em 1 mar. 2014.

SINGER, J. M.; NOBRE, J. S.; ROCHA, F. M. M. **Análise de dados longitudinais**. São Paulo: IME-USP, 2011. 221 p. Disponível em: <<http://people.ufpr.br/~lucambio/CE075/1S2011/Singer&Nobre&Rocha2011mar.pdf>>. Acesso em: 1 set. 2011.

SINGER, J.D.; WILLETT, J.B. **Applied longitudinal data analysis: modelling change and event occurrence**. New York: Oxford University Press, 2003. 644 p.

SLAMBEROVÁ, R.; POMETLOVÁ, M.; CHAROUSOVÁ, P. Postnatal development of rat pups is altered by prenatal methamphetamine exposure. **Progress in Neuro-Psychopharmacology & Biological Psychiatry**, Sidney, v. 30, n. 1, p. 82-88, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16046043>>. Acesso em: 1 jun. 2014.

SMITH, P. L. Splines as a Useful and Convenient Statistical Tool. **The American Statistician**, Alexandria, v. 33, n. 2, p. 57-62, 1979. Disponível em: <<http://www.tandfonline.com/doi/pdf/10.1080/00031305.1979.10482661>>. Acesso em: 1 jun. 2013.

SOUBHIA, R. M. C.; MENDES, G. E. F.; MENDONÇA, F. Z.; BAPTISTA, M. A. S.; CIPULLO, J. P.; BURDMANN, E. A. Tacrolimus and nonsteroidal anti-inflammatory drugs: an association to be avoided. **American Journal of Nephrology**, Chicago, v. 25, n. 4, p. 327-334, 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15976495>>. Acesso em: 1 jun. 2014.

STEEL, R.G.D.; TORRIE, J.H. **Principles and procedures of statistics : a biometrical approach**. 2 ed., Singapura: McGraw-Hill, 1980. 633 p.

STROUP, W. W. **Generalized linear mixed models: modern concepts, methods and applications**. Boca Raton: Taylor & Francis, 2012. 555 p.

TOUSEN, Y; UMEKI, M.; NAKASHIMA, Y.; ISHIMI, Y.; IKEGAMI, S. Effects of genistein, an isoflavone, on pregnancy outcome and organ weights of pregnant and lactating rats and development of their suckling pups. **Journal of Nutritional Science and Vitaminology**, Tokyo, v. 52, n. 3, p. 174-182, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16967761>>. Acesso em: 1 jun. 2014.

TOXICITY testing methods. **Encyclopedia of earth**, 2013. Disponível em: <<http://www.eoearth.org/view/article/51cbef177896bb431f69c4fc>>. Acesso em: 1 set. 2014.

VERBEKE, G.; MOLENBERGHS, G. **Linear mixed models for longitudinal data**. New York: Springer-Verlag, 2000. 568 p.

VIDO, D. S. et al. Maternal obesity and late effects on offspring metabolism. **Arquivos Brasileiros de Endocrinologia & Metabologia**, São Paulo, v. 58, n. 3, p. 301-307, 2014. ISSN 1677-9487. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24863094>>. Acesso em: 1 jul. 2014.

VIDONHO, A. F. J; SILVA, A. A. ; CATANOZI, S.; ROCHA, J. C.; BEUTEL, A.; CARILLO, B. A.; FURUKAWA, L. N. S.; CAMPOS, R. R.; BERGAMASCHI, C. M. T.; CARPINELLI, A. R.; QUINTÃO, E. C. R.; DOLNIKOFF, M. S.; HEIMANN, J. C. Perinatal salt restriction: a new pathway to programming insulin resistance and dyslipidemia in adult wistar rats. **Pediatric Research**, New York, v. 56, n. 6, p. 842-848, 2004. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15470199>>. Acesso em: 1 mar. 2014.

VILLARINO, M. E.; GOYA, J. A.; DE LUCCA, R. C.; UBIOS, A. M. Alterations of tooth eruption and growth in pups suckling from diabetic dams. **Pediatric Research**, New York, v. 58, n. 4, p. 695-699, 2005. Disponível em: <<http://www.nature.com/pr/journal/v58/n4/pdf/pr2005726a.pdf>>. Acesso em: 1 mar. 2014.

VONESH, E. F.; CHINCHILLI, V. M. **Linear and Nonlinear Models for the Analysis of Repeated Measurements**. New York: Marcel Dekker, 1997. 560 p. (Statistics: textbooks and monographs, 154).

WEST, B. T.; WELCH, K. B.; GALECKI, A. T. **Linear mixed models: a practical guide using statistical software**. New York, Chapman & Hall, 2007. 339 p.

WU, H.; ZHANG, J. T. **Nonparametric regression methods for longitudinal data analysis**: mixed-effects modeling approaches. Hoboken: John Wiley & Sons, 2006. 384 p. (Series in Probability and Statistics, 515).

ZEWOTIR, T.; GALPIN, J. S. Evaluation of linear mixed model case deletion diagnostic tools by Monte Carlo simulation. **Communications in Statistics - Simulation and Computation**, Alexandria, v. 35, n.3, p. 645-682, 2006. Disponível em: <<http://dx.doi.org/10.1080/03610910600716795>>. Acesso em: 1 dez. 2011.

APÊNDICES

APÊNDICE A - Obtenção do BLUE e BLUP de β e u

Nesta seção são apresentados os cálculos matriciais necessários para a obtenção conjunta do BLUE de β e do BLUP de u por meio das Equações de Modelos Mistos de Henderson, obtidas pela maximização da função de densidade conjunta $f(\mathbf{y}, \mathbf{u})$.

Seja o modelo linear misto

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}u + \mathbf{e}. \quad (56)$$

Assume-se,

$$\begin{bmatrix} u \\ e \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{G} & 0 \\ 0 & \mathbf{R} \end{bmatrix} \right). \quad (57)$$

Por definição de função de densidade condicional,

$$f(\mathbf{y}|\mathbf{u}) = \frac{f(\mathbf{y}, \mathbf{u})}{f(\mathbf{u})},$$

assim sendo,

$$f(\mathbf{y}, \mathbf{u}) = f(\mathbf{y}|\mathbf{u})f(\mathbf{u}). \quad (58)$$

De (56) e (57) tem-se que,

$$f(\mathbf{y}|\mathbf{u}) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} e^{-1/2((\mathbf{y}-\mathbf{X}\beta-\mathbf{Z}u)' \mathbf{R}^{-1}(\mathbf{y}-\mathbf{X}\beta-\mathbf{Z}u))}$$

$$f(\mathbf{u}) = \frac{1}{(2\pi)^{n/2}|\mathbf{G}|^{1/2}} e^{-1/2((\mathbf{u}-0)' \mathbf{G}^{-1}(\mathbf{u}-0))}.$$

Portanto, de (58) e assumindo $\mathbf{G}$ e $\mathbf{R}$ conhecidas, tem-se:

$$f(\mathbf{y}, \mathbf{u}) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} e^{-1/2((\mathbf{y}-\mathbf{X}\beta-\mathbf{Z}u)' \mathbf{R}^{-1}(\mathbf{y}-\mathbf{X}\beta-\mathbf{Z}u))} \frac{1}{(2\pi)^{n/2}|\mathbf{G}|^{1/2}} e^{-1/2(\mathbf{u}' \mathbf{G}^{-1} \mathbf{u})}.$$

Aplicando o logaritmo na função $f(\mathbf{y}, \mathbf{u})$ e desconsiderando os termos constantes, obtém-se a função

$$g(\beta, \mathbf{u}) = -\frac{1}{2} ((\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}u)' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}u)) - \frac{1}{2} (\mathbf{u}' \mathbf{G}^{-1} \mathbf{u}) =$$

$$\begin{aligned}
&= -\frac{1}{2} \left((\mathbf{y}' - \boldsymbol{\beta}' \mathbf{X}' - \mathbf{u}' \mathbf{Z}')' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) + \mathbf{u}' \mathbf{G}^{-1} \mathbf{u} \right) = \\
&= -\frac{1}{2} \left(\mathbf{y}' \mathbf{R}^{-1} \mathbf{y} - \underline{\mathbf{y}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta}} - \underline{\mathbf{y}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u}} - \underline{\boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{y}} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta} + \underline{\boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u}} - \underline{\mathbf{u}' \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y}} \right) + \\
&\quad -\frac{1}{2} \left(\underline{\mathbf{u}' \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta}} + \mathbf{u}' \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u} + \mathbf{u}' \mathbf{G}^{-1} \mathbf{u} \right) = \\
&= -\frac{1}{2} \left(\mathbf{y}' \mathbf{R}^{-1} \mathbf{y} - 2\boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} - 2\mathbf{u}' \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta} + 2\boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u} + \mathbf{u}' \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u} + \mathbf{u}' \mathbf{G}^{-1} \mathbf{u} \right).
\end{aligned}$$

As derivadas parciais de $g(\boldsymbol{\beta}, \mathbf{u})$ com relação a $\boldsymbol{\beta}$ e $\mathbf{u}$ são dadas por:

$$\begin{bmatrix} \frac{\partial g(\boldsymbol{\beta}, \mathbf{u})}{\partial \boldsymbol{\beta}} \\ \frac{\partial g(\boldsymbol{\beta}, \mathbf{u})}{\partial \mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} - \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta} - \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} - \underbrace{\boldsymbol{\beta}' \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z}}_{\mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta}} - \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} \mathbf{u} - \mathbf{G}^{-1} \mathbf{u} \end{bmatrix}. \quad (59)$$

Igualando-se a zero as derivadas parciais em (59), obtêm-se as Equações de Modelos Mistos de Henderson, dadas por:

$$\begin{aligned}
&\begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} - \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} - \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \hat{\mathbf{u}} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} - \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} - (\mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \\
&\Leftrightarrow \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \hat{\mathbf{u}} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + (\mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix} \Leftrightarrow \\
&\Leftrightarrow \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} & (\mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}. \quad (60)
\end{aligned}$$

As soluções das equações em 60 são pontos de máximo da função $g(\boldsymbol{\beta}, \mathbf{u})$, uma vez que o determinante da matriz hessiana em 61 é maior que zero e $\frac{\partial^2 g(\boldsymbol{\beta}, \mathbf{u})}{\partial \boldsymbol{\beta}^2} < 0$.

$$H = \begin{vmatrix} \frac{\partial^2 g(\boldsymbol{\beta}, \mathbf{u})}{\partial \boldsymbol{\beta}^2} & \frac{\partial^2 g(\boldsymbol{\beta}, \mathbf{u})}{\partial \boldsymbol{\beta} \partial \mathbf{u}} \\ \frac{\partial^2 g(\boldsymbol{\beta}, \mathbf{u})}{\partial \mathbf{u} \partial \boldsymbol{\beta}} & \frac{\partial^2 g(\boldsymbol{\beta}, \mathbf{u})}{\partial \mathbf{u}^2} \end{vmatrix} = \begin{vmatrix} -\mathbf{X}' \mathbf{R}^{-1} \mathbf{X} & -\mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \\ -\mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} & -(\mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \end{vmatrix}. \quad (61)$$

Portanto,

$$\begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ \hline Z'R^{-1}X & (Z'R^{-1}Z + G^{-1}) \end{bmatrix}^{-1} \begin{bmatrix} X'R^{-1}y \\ \hline Z'R^{-1}y \end{bmatrix}. \quad (62)$$

A matriz inversa em (62) foi obtida utilizando o resultado R1¹.

$$\begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} X'R^{-1}y \\ \hline Z'R^{-1}y \end{bmatrix}. \quad (63)$$

$$\begin{aligned} B_{11} &= (X'R^{-1}X)^{-1} + (X'R^{-1}X)^{-1} (X'R^{-1}Z) \\ &\quad \left[(Z'R^{-1}Z + G^{-1}) - (Z'R^{-1}X) (X'R^{-1}X)^{-1} (X'R^{-1}Z) \right]^{-1} \\ &\quad (Z'R^{-1}X) (X'R^{-1}X)^{-1} =^2 \\ &= \left[(X'R^{-1}X) - (X'R^{-1}Z) [Z'R^{-1}Z + G^{-1}]^{-1} (X'R^{-1}Z)' \right]^{-1} = \\ &= \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1}. \\ B_{12} &= - \left[(X'R^{-1}X) - (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \right]^{-1} \\ &\quad (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1} = \\ &= - \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} \\ &\quad (X'R^{-1}Z) (Z'R^{-1}Z + D^{-1})^{-1}. \end{aligned}$$

$$\begin{aligned} B_{21} &= - (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \\ &\quad \left[(X'R^{-1}X) - (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \right]^{-1} = \\ &= - (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \\ &\quad \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1}. \end{aligned}$$

1

$$R1 : \begin{bmatrix} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}C_2^{-1}A_{21}A_{11}^{-1} & -C_1^{-1}A_{12}A_{22}^{-1} \\ \hline -A_{22}^{-1}A_{21}C_1^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}C_1^{-1}A_{12}A_{22}^{-1} \end{bmatrix}$$

$$C_1 = A_{11} - A_{12}A_{22}^{-1}A_{21} \text{ e } C_2 = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

$$^2\text{identidade de Woodbury: } (A - CBC')^{-1} = \left(A^{-1} + A^{-1}C (B^{-1} - C'A^{-1}C)^{-1} C'A^{-1} \right).$$

$$\begin{aligned}
B_{22} &= (Z'R^{-1}Z + G^{-1})^{-1} + (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \\
&\left[(X'R^{-1}X) - (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \right]^{-1} \\
&\quad (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1} = \\
&= (Z'R^{-1}Z + G^{-1})^{-1} + (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \\
&\quad \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} \\
&\quad (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1}.
\end{aligned}$$

Portanto, de (63),

$$\hat{\beta} = B_{11} (X'R^{-1}y) + B_{12} (Z'R^{-1}y) \quad (64)$$

$$\hat{u} = B_{21} (X'R^{-1}y) + B_{22} (Z'R^{-1}y). \quad (65)$$

Substituindo B_{11} e B_{12} em (64) tem-se,

$$\begin{aligned}
\hat{\beta} &= \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} (X'R^{-1}y) - \\
&\quad - \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} (X'R^{-1}Z) \\
&\quad \quad (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}y) = \\
&= \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} \\
&\quad (X'\Sigma^{-1}y) - (X'R^{-1}Z) (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}y).
\end{aligned}$$

$$\begin{aligned}
\therefore \hat{\beta} &= \left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} \\
&\quad X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) y.
\end{aligned}$$

Substituindo B_{21} e B_{22} em (65) tem-se,

$$\begin{aligned}
\hat{u} &= - (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \\
&\left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} (X'R^{-1}y) + \\
&\quad + (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}y) + \\
&\quad + (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}X) \\
&\left[X' \left(R^{-1} - R^{-1}Z (Z'R^{-1}Z + G^{-1})^{-1} Z'R^{-1} \right) X \right]^{-1} (X'R^{-1}Z) \\
&\quad (Z'R^{-1}Z + G^{-1})^{-1} (Z'R^{-1}y) =
\end{aligned}$$

$$\begin{aligned}
&= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \\
&\left\{ -\mathbf{X} \left[\mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{X} \right]^{-1} \right. \\
&(\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}) + \mathbf{y} + \mathbf{X} \left[\mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{X} \right]^{-1} \\
&\quad \left. (\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}) (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}) \right\} = \\
&= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \\
&\left\{ \mathbf{y} - \mathbf{X} \left[\mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{X} \right]^{-1} \right. \\
&\quad \left. \left[(\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}) + (\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}) (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}) \right] \right\} = \\
&= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \\
&\left\{ \mathbf{y} - \mathbf{X} \left[\mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{X} \right]^{-1} \right. \\
&\quad \left. \left[\mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{y} \right] \right\}.
\end{aligned}$$

$$\therefore \hat{\mathbf{u}} = (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}).$$

Assim sendo, os vetores $\hat{\boldsymbol{\beta}}$ e $\hat{\mathbf{u}}$, que correspondem, respectivamente, ao BLUE e BLUP nos modelos lineares mistos, são dados por:

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= \left[\mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{X} \right]^{-1} \\
&\quad \mathbf{X}' \left(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z} (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \right) \mathbf{y}
\end{aligned}$$

e

$$\hat{\mathbf{u}} = (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}).$$

Apêndice B - Códigos escritos no SAS para ajuste dos modelos M1.2 a M1.14

*M1.2;

```
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
  class u_e dose_ord sexo_ord dia_d;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
            dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
            residual solution outp=exper1.res_cond_m2 outpm=exper1.res_marg_m2;
  random intercept / subject=u_e group=sexo_ord solution g v=(1 21) vcorr=(1 21);
  repeated dia_d/subject=u_e r=(1 21) rcorr=(1 21);
  format dose_ord dose_ordfmt.;
  format sexo_ord sexo_ordfmt.;
run;
ods graphics off;
```

*M1.3;

```
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
  class u_e dose_ord sexo_ord dia_d;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
            dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
            residual solution outp=exper1.res_cond_m3 outpm=exper1.res_marg_m3;
  random intercept / subject=u_e solution g v=(1 21) vcorr=(1 21);
  repeated dia_d/subject=u_e group=sexo_ord r=(1 21) rcorr=(1 21);
  format dose_ord dose_ordfmt.;
  format sexo_ord sexo_ordfmt.;
run;
ods graphics off;
```

*M1.4;

```
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
  class u_e dose_ord sexo_ord dia_d;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
            dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
            residual solution outp=exper1.res_cond_m4 outpm=exper1.res_marg_m4;
  random intercept dia / subject=u_e solution g v vcorr;
  repeated dia_d/subject=u_e r rcorr;
```

```

format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.5;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
  class u_e dose_ord sexo_ord dia_d;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
              dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
              residual solution outp=exper1.res_cond_m5 outpm=exper1.res_marg_m5;
  random intercept / subject=u_e;
  random dia / subject=u_e group=sexo_ord solution g v=(1 21) vcorr=(1 21);
  repeated dia_d/subject=u_e r rcorr;
  format dose_ord dose_ordfmt.;
  format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.6;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
  class u_e dose_ord sexo_ord dia_d;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
              dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
              residual solution outp=exper1.res_cond_m6 outpm=exper1.res_marg_m6;
  random intercept dia / subject=u_e solution g v=(1 21) vcorr=(1 21);
  repeated dia_d/subject=u_e group=sexo_ord r=(1 21) rcorr=(1 21);
  format dose_ord dose_ordfmt.;
  format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.7;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
  class u_e dose_ord sexo_ord dia_d;
  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia

```

```

        dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
        residual solution outp=exper1.res_cond_m7 outpm=exper1.res_marg_m7;
random intercept dia / subject=u_e solution g v vcorr;
repeated dia_d/subject=u_e type=toeph(1) r rcorr;
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.8;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
    class u_e dose_ord sexo_ord dia_d;
    model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
        dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
        residual solution outp=exper1.res_cond_m8 outpm=exper1.res_marg_m8;
parms
786.65  0.08 189.07 61.51 49.83 47.58 23.07 13.37 14.01 13.61  6.03
    9.34 11.87  36.26 20.20 40.24 27.02 19.39 13.22 16.26 13.25  8.12
    18.11 80.63  11.22 10.68  7.55 28.10 53.54 11.13 13.66 189.0 61.51
    49.83 47.58  23.07 13.37 14.01 13.61  6.03  9.30 11.87 36.26 20.20
    40.24 27.02  19.39 13.22 16.26 13.25  8.12 18.11 80.63 11.22 10.68
    7.55 28.10  53.54 11.13 13.66
; *os valores iniciais correspondem as estimativas dos parâmetros do modelo m1.7
    que foram considerados iguais para machos e fêmeas;
random intercept dia / subject=u_e solution g v=(1 21) vcorr=(1 21);
repeated dia_d/subject=u_e type=toeph(1) group=sexo_ord r=(1 21) rcorr=(1 21);
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.9;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
    class u_e dose_ord sexo_ord dia_d;
    model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
        dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
        residual solution outp=exper1.res_cond_m9 outpm=exper1.res_marg_m9;

```

```

random intercept dia / subject=u_e solution g v vcorr;
repeated dia_d/subject=u_e type=ar(1) r rcorr;
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.10;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
class u_e dose_ord sexo_ord dia_d;
model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
            dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
            residual solution outp=exper1.res_cond_m10 outpm=exper1.res_marg_m10;
random intercept dia / subject=u_e solution g v=(1 21) vcorr=(1 21);
repeated dia_d/subject=u_e type=ar(1) group=sexo_ord r=(1 21) rcorr=(1 21);
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.11 ajuste que convergiu para um máximo local;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
class u_e dose_ord sexo_ord dia_d;
model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
            dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
            residual solution outp=exper1.res_cond_m11 outpm=exper1.res_marg_m11;
random intercept dia / subject=u_e solution g v vcorr;
repeated dia_d/subject=u_e type=arh(1) r rcorr;
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;

*M1.11 ajuste que convergiu para o máximo global;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
class u_e dose_ord sexo_ord dia_d;

```



```

model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
              dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
              residual solution outp=exper1.res_cond_m11 outpm=exper1.res_marg_m11;

parms
786.65  0.08 189.07 61.51 49.83 47.58 23.07 13.37 14.01 13.61 6.03
    9.34 11.87 36.26 20.20 40.24 27.02 19.39 13.22 16.26 13.25 8.12
    18.11 80.63 11.22 10.68 7.55 28.10 53.54 11.13 13.66 0.06
; *os valores iniciais correspondem as estimativas dos parâmetros do modelo M1.7
    incluindo a estimativa inicial do parâmetro autoregressivo;
random intercept dia / subject=u_e solution g v vcorr;
repeated dia_d/subject=u_e type=arh(1) r rcorr;
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;

run;
ods graphics off;

*M1.12;
ods graphics on;
proc mixed data=exper1.dados_univar covtest order=internal;
    class u_e sexo_ord dose_ord dia_d;
    model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
                  dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia/
                  residual solution outp=exper1.res_cond_m12 outpm=exper1.res_marg_m12;

    parms
792.8  0.1 353.1 98.1 44.7 74.2 17.7 16.2 14.6 15.4 8.0 15.9 8.9
    53.3 35.6 65.4 42.7 21.3 18.8 23.5 16.3 10.0 19.2 140.0 9.7 12.4
    6.1 44.4 92.7 14.4 19.0 0.1 16.2 20.3 53.8 16.7 26.2 9.3 12.3
    12.7 4.9 4.0 14.7 21.5 6.4 16.5 10.9 17.3 8.4 8.7 9.8 5.5
    17.8 18.9 12.4 9.8 9.7 11.5 14.0 7.4 8.8 0.1
; *os valores iniciais correspondem as estimativas dos parâmetros do modelo M1.8
    incluindo as estimativas iniciais dos parâmetros autoregressivos;
random intercept dia / subject=u_e solution g v=(1 21) vcorr=(1 21);
repeated dia_d/subject=u_e type=arh(1) group=sexo_ord r=(1 21) rcorr=(1 21);
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;

run;
ods graphics off;

```

```

*M1.13;

ods graphics on;

proc mixed data=exper1.dados_univar covtest order=internal;

  class u_e dose_ord sexo_ord dia_d;

  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
              dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
              residual solution outp=exper1.res_cond_m13 outpm=exper1.res_marg_m13;

  parms
786.65  0.08 189.07 61.51 49.83 47.58 23.07 13.37 14.01 13.61  6.03
    9.34 11.87  36.26 20.20 40.24 27.02 19.39 13.22 16.26 13.25  8.12
18.11 80.63  11.22 10.68  7.55 28.10 53.54 11.13 13.66  0.01  0.01
    0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01
    0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01
    0.01  0.01  0.01  0.01

; *os valores iniciais correspondem as estimativas dos parâmetros do modelo M1.7
  incluindo as correlações iguais a 0,01;

  random intercept dia / subject=u_e solution g v vcorr;

  repeated dia_d/subject=u_e type=ante(1) r rcorr;

  format dose_ord dose_ordfmt.;

  format sexo_ord sexo_ordfmt.;

run;

ods graphics off;

*M1.14;

ods graphics on;

proc mixed data=exper1.dados_univar covtest order=internal;

  class u_e dose_ord sexo_ord dia_d ;

  model peso = dose_ord sexo_ord dose_ord*sexo_ord dia
              dose_ord*dia sexo_ord*dia dose_ord*sexo_ord*dia /
              residual solution outp=exper1.res_cond_m14 outpm=exper1.res_marg_m14;

  parms
797.26  0.06 288.79  78.27  45.03 73.08 16.21 18.82  16.95  18.75   7.93 16.16  11.34
52.24 27.01  70.27  49.63  20.38 21.65 28.55 22.16  11.07  30.52 194.96 14.88  17.40
 5.86 78.93 110.94  11.15  14.25  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01
 0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01  0.01
 0.01  0.01  0.01  0.01  0.01  0.01  0.01 16.51 18.64 57.64 19.56 26.12  9.40
17.37 14.85  4.54  4.43 17.29 28.53  7.64 21.63  9.45 15.61 10.07  8.90 11.14
 6.84 20.50 21.79 14.21  9.44  8.16 12.25 17.52  8.02  8.38  0.01  0.01  0.01

```

```
0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01
0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01
; *os valores iniciais correspondem as estimativas dos parâmetros do modelo M1.16
    incluindo as correlações iguais a 0,01;
random intercept dia / subject=u_e solution g v=(1 21) vcorr=(1 21);
repeated dia_d/subject=u_e type=ante(1) group=sexo_ord r=(1 21) rcorr=(1 21);
format dose_ord dose_ordfmt.;
format sexo_ord sexo_ordfmt.;
run;
ods graphics off;
```

Apêndice C - Códigos escritos no SAS para ajuste dos modelos M2.2 a M2.13

```

data exper2.dados_univar;
  set exper2.dados_univar;
  if fase = 1 then fase_ord = 2;
  if fase = 2 then fase_ord = 1;
  if dose = 0 then dose_ord = 3;
  if dose = 25 then dose_ord = 1;
  if dose = 50 then dose_ord = 2;
run;

proc format;
  value fase_ordfmt 1 = "fase2"
                  2 = "fase1";
run;

proc format;
  value dose_ordfmt 1 = "dose25"
                  2 = "dose50"
                  3 = "controle";
run;

*M2.1;
proc mixed data=exper2.dados_univar covtest order=internal;
  class u_e filhote fase_ord dose_ord dia_d;
  model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
          fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
          idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
          idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
          no_filhotes_cent pesorata1_cent no_dias_gest_cent /
          solution residual outpm=exper2.res_marg_m1 outp=exper2.res_cond_m1;
  random intercept idadm1 idadm1_2 / subject=u_e type=un solution v vcorr g;
  random idadm1 idadm1_2 / subject=filhote(u_e) type=un;
  repeated dia_d/subject= filhote(u_e) r rcorr;
  format fase_ord fase_ordfmt.;
  format dose_ord dose_ordfmt.;
run;

*M2.2;

```

```

proc mixed data=exper2.dados_univar covtest order=internal;
  class u_e filhote fase_ord dose_ord dia_d;
  model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
         fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
         idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
         idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
         no_filhotes_cent pesorata1_cent no_dias_gest_cent cos_idadm1 /
         solution residual outpm=exper2.res_marg_m2 outp=exper2.res_cond_m2;
  random intercept idadm1 idadm1_2 / subject=u_e type=un solution v vcorr g;
  random idadm1 idadm1_2/ subject=filhote(u_e) type=un;
  repeated dia_d/subject= filhote(u_e) r rcorr;
  format fase_ord fase_ordfmt.;
  format dose_ord dose_ordfmt.;
run;

```

*M2.3 e M2.4;

```

proc mixed data=exper2.dados_univar covtest order=internal;
  class u_e filhote fase_ord dose_ord dia_d;
  model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
         fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
         idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
         idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
         no_filhotes_cent pesorata1_cent no_dias_gest_cent cos_idadm1 /
         solution residual outpm=exper2.res_marg_m3 outp=exper2.res_cond_m3; *se M2.3;
         *solution residual outpm=exper2.res_marg_m4 outp=exper2.res_cond_m4; *se M2.4;
  random intercept idadm1 idadm1_2 cos_idadm1 / subject=u_e type=un solution v vcorr g;
  random idadm1 idadm1_2/ subject=filhote(u_e) type=un;
  repeated dia_d/subject= filhote(u_e) r rcorr; *se M2.3;
  *repeated dia/subject= filhote(u_e) r rcorr type=ar(1); *se M2.4;
  format fase_ord fase_ordfmt.;
  format dose_ord dose_ordfmt.;
run;

```

*M2.5 a M2.8;

```

proc mixed data=exper2.dados_univar covtest order=internal;
  class u_e filhote fase_ord dose_ord dia_d;
  model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
         fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
         idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2

```

```

        idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
        no_filhotes_cent pesorata1_cent no_dias_gest_cent cos_idadm1 /
        solution residual outpm=exper2.res_marg_m5 outp=exper2.res_cond_m5; *se M2.5
        *solution residual outpm=exper2.res_marg_m6 outp=exper2.res_cond_m6; *se M2.6;
        *solution residual outpm=exper2.res_marg_m7 outp=exper2.res_cond_m7; *se M2.7;
        *solution residual outpm=exper2.res_marg_m8 outp=exper2.res_cond_m8; *se M2.8;
random intercept idadm1 idadm1_2 cos_idadm1 / subject=u_e type=un solution v vcorr g;
repeated dia_d/subject= filhote(u_e) type=arh(1) r rcorr; *se M2.5;
*repeated dia_d/subject= filhote(u_e) type=ante(1) r rcorr; *se M2.6;
*repeated dia_d/subject= filhote(u_e) type=toeph r rcorr; *se M2.7; *cpu time 17h;
*repeated dia_d/subject= filhote(u_e) type=un r rcorr; *se M2.8;
format fase_ord fase_ordfmt.;
format dose_ord dose_ordfmt.;
run;

*M2.9;
proc mixed data=exper2.dados_univar covtest order=internal;
class u_e filhote fase_ord dose_ord dia_d;
model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
        fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
        idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
        idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
        no_filhotes_cent pesorata1_cent no_dias_gest_cent cos_idadm1 /
        solution residual outpm=exper2.res_marg_m9 outp=exper2.res_cond_m9;
random intercept idadm1 idadm1_2 / subject=u_e type=un solution v vcorr g;
repeated dia_d / subject= filhote(u_e) type=un r rcorr;
format fase_ord fase_ordfmt.;
format dose_ord dose_ordfmt.;
run;

*M2.10;
proc mixed data=exper2.dados_univar covtest order=internal;
class u_e filhote fase_ord dose_ord dia_d;
model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
        fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
        idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
        idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
        no_filhotes_cent pesorata1_cent no_dias_gest_cent /
        solution residual outpm=exper2.res_marg_m10 outp=exper2.res_cond_m10;

```

```

random intercept idadm1 idadm1_2 / subject=u_e type=un solution v vcorr g;
repeated dia_d / subject= filhote(u_e) type=un r rcorr;
format fase_ord fase_ordfmt.;
format dose_ord dose_ordfmt.;
run;

*M2.11;
proc mixed data=exper2.dados_univar covtest order=internal;
class u_e filhote fase_ord dose_ord dia_d;
model peso=fase_ord dose_ord fase_ord*dose_ord idadm1
        fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
        idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
        idadm1_3 fase_ord*idadm1_3 dose_ord*idadm1_3 fase_ord*dose_ord*idadm1_3
        no_filhotes_cent no_filhotes_cent*idadm1 pesorata1_cent no_dias_gest_cent /
        solution residual outpm=exper2.res_marg_m11 outp=exper2.res_cond_m11;
random intercept idadm1 idadm1_2 / subject=u_e type=un solution v vcorr g;
repeated dia_d / subject= filhote(u_e) type=toeph r rcorr;
format fase_ord fase_ordfmt.;
format dose_ord dose_ordfmt.;
run;

*M2.12;
ods graphics on;
proc mixed data=exper2.dados_univar covtest order=internal;
class u_e filhote fase_ord dose_ord dia_d;
model peso=fase_ord dose_ord fase_ord*dose_ord
        idadm1 fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1
        idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2
        idadm1m14_2 fase_ord*idadm1m14_2 dose_ord*idadm1m14_2
        fase_ord*dose_ord*idadm1m14_2
        no_filhotes_cent no_filhotes_cent*idadm1 pesorata1_cent no_dias_gest_cent /
        solution residual outpm=exper2.res_marg_m12 outp=exper2.res_cond_m12;
random intercept idadm1 idadm1_2 idadm1m14_2 / subject=u_e type=un solution v vcorr g;
random idadm1 idadm1_2 / subject=filhote(u_e) type=un;
repeated dia_d / subject= filhote(u_e) type=ante(1) r rcorr;
format fase_ord fase_ordfmt.;
format dose_ord dose_ordfmt.;
run;
ods graphics off;

```

```

*M2.13;

ods graphics on;

proc mixed data=exper2.dados_univar covtest order=internal;

  class u_e filhote fase_ord dose_ord dia_d;

  model peso=fase_ord dose_ord fase_ord*dose_ord

          idadm1 fase_ord*idadm1 dose_ord*idadm1 fase_ord*dose_ord*idadm1

          idadm1_2 fase_ord*idadm1_2 dose_ord*idadm1_2 fase_ord*dose_ord*idadm1_2

          idadm1m14 fase_ord*idadm1m14 dose_ord*idadm1m14 fase_ord*dose_ord*idadm1m14

          idadm1m14_2 fase_ord*idadm1m14_2 dose_ord*idadm1m14_2

          fase_ord*dose_ord*idadm1m14_2

          no_filhotes_cent no_filhotes_cent*idadm1 pesorata1_cent no_dias_gest_cent /

          solution residual outpm=exper2.res_marg_m13 outp=exper2.res_cond_m13;

  random intercept idadm1 idadm1_2 idadm1m14 idadm1m14_2 /

                                     subject=u_e type=un solution v vcorr g;

  random idadm1 idadm1m14 / subject=filhote(u_e) type=un;

  repeated dia_d / subject= filhote(u_e) type=ante(1) r rcorr;

  format fase_ord fase_ordfmt.;

  format dose_ord dose_ordfmt.;

run;

ods graphics off;

```