MODELOS LINEARES GENERALIZADOS MISTOS E EQUAÇÕES DE ESTIMAÇÃO GENERALIZADAS PARA DADOS BINÁRIOS APLICADOS EM ANESTESIOLOGIA VETERINÁRIA

Maicon Vinícius Galdino

Dissertação apresentada à Universidade
Estadual Paulista "Júlio de Mesquita Filho"
para a obtenção do título de
Mestre em Biometria

BOTUCATU
São Paulo - Brasil
Fevereiro - 2015

MODELOS LINEARES GENERALIZADOS MISTOS E EQUAÇÕES DE ESTIMAÇÃO GENERALIZADAS PARA DADOS BINÁRIOS APLICADOS EM ANESTESIOLOGIA VETERINÁRIA

Maicon Vinícius Galdino

Orientadora: Prof(a). Dr(a). Liciana Vaz de Arruda Silveira

Dissertação apresentada à Universidade Estadual Paulista "Júlio de Mesquita Filho" para a obtenção do título de Mestre em Biometria

BOTUCATU
São Paulo – Brasil
Fevereiro – 2015

Dedicatória

Para as minhas avós Sebastiana Martins Galdino (*in memoriam*) e Luiza Francisco da Silveira (*in memoriam*) – exemplos e lembranças deixados em minha vida.

Agradecimentos

Agradeço especialmente à Prof(a). Dr(a). Liciana Vaz de Arruda Silveira, pela paciência para orientar este trabalho, pelas dicas ao esclarecer algumas dúvidas, pelas risadas, pela amizade, ... enfim, por tudo.

Agradeço também à Prof(a). Dr(a). Maria del Pilar Diaz e ao médico veterinário e Prof. Dr. José Carlos de Figueiredo Pantoja por terem se disponibilizado a participar como membros do Exame de Qualificação, da Defesa da Dissertação e pelas sugestões feitas com o intuito de melhorar meu trabalho.

Agradeço aos professores do Curso de Mestrado em Biometria: Antônio Carlos Simões Pião, Carlos Roberto Padovani, José Silvio Govone, Lídia Raquel de Carvalho, Luzia Aparecida Trinca, Miriam Harumi Tsunemi, Paulo Fernando de Arruda Mancera e Paulo Milton Barbosa Landim por terem contribuído para o meu crescimento profissional.

Ao Prof. Dr. Luciano Barbosa pela disponibilidade em me aceitar como estagiário na disciplina de Bioestatística oferecida para o curso de Nutrição do IBB/Unesp.

Aos amigos e colegas Ariane Campolim, Clóvis Aparecido Caface Filho, Edimar Izidoro Novaes, Glauber Márcio Silveira Pereira, Jane Maiara Bertolla, Luiz Alberto Amaral Nardi, Maria Cecília Evangelista, Sophia Lanza de Andrade e Talita Tanaka Fernandes.

Ao Prof. Dr. Stelio Pacca Loureiro Luna e à Prof(a). Dr(a). Marilda Orghero Taffarel por ceder os dados que ilustrou as metodologias estatísticas estudadas.

A todos os meus alunos e a todos os meus colegas de profissão (professores) da Escola Estadual “João Queiroz Marques” (localizada em Rubião Júnior) que entenderam o motivo do meu afastamento no ano de 2014 para concluir esta dissertação.

À minha esposa Juliana Madureira do Santos Galdino, pelo companheirismo, amor e carinho.

À minha mãe Ivani Freire Galdino e aos meus pequenos sobrinhos Eduardo Galdino Alonso, Heitor de Oliveira Galdino e Flávia Galdino Alonso.

À Capes, pelo apoio financeiro.

SUMÁRIO

LISTA DE FIGURAS **4**

LISTA DE TABELAS **5**

RESUMO **6**

SUMMARY **8**

1) INTRODUÇÃO **10**

2) OBJETIVOS E JUSTIFICATIVA **14**

3) FUNDAMENTAÇÃO TEÓRICA **15**

3.1) Equações de Estimação Generalizadas **15**
3.1.1) Estrutura do modelo marginal 16
3.1.2) Estimação e inferência 21

3.2) Modelo Linear Generalizado Misto **23**
3.2.1) Estrutura do modelo linear generalizado misto 23
3.2.2) Estimação e inferência 26

3.3) Modelo Marginal versus Modelo com Efeitos Aleatórios **30**

4) ESTRUTURAS DE MATRIZES DE COVARIÂNCIAS **32**

5) RAZÃO DE CHANCES **36**

6) DESCRIÇÃO DOS DADOS **38**

7) ANÁLISE E RESULTADOS **41**

7.1) Análise Descritiva **42**

7.2) Modelagem Estatística **44**

8) CONCLUSÕES **55**

REFERÊNCIAS BIBLIOGRÁFICAS **58**

ANEXOS **61**

LISTA DE FIGURAS

Figura 1: Silhueta dos equinos (adaptado de Mair et al (2002))..... 39
Figura 2: Gráfico da proporção da variável resposta (valores “1”) por tratamento e por momento. 43
Figura 3: Histograma (a) e box-plot (b) das estimativas dos parâmetros de efeitos aleatórios do modelo 1..... 48
Figura 4: Gráficos dos resíduos condicionais estudentizados, (a) em função dos valores preditos, (b) quantil – quantil (c) histograma. 49
Figura 5: Gráficos de diagnósticos, (a) distância de Cook, (b) leverage. 52

LISTA DE TABELAS

Tabela 1: Frequência absoluta da variável resposta (valores “1”) por tratamento e por momento 42
Tabela 2: Frequência absoluta da variável resposta para valores “0” e para valores perdidos (entre parênteses) por tratamento e por momento..... 42
Tabela 3: Proporção da variável resposta (valores “1”) por tratamento e por momento 43
Tabela 4: Valores de AIC para diferentes estruturas de matrizes de covariâncias utilizando MLGMs. 44
Tabela 5: Estatísticas de ajuste do modelo 1 45
Tabela 6: Estatísticas e estimativas dos parâmetros de efeitos fixos do modelo 1 46
Tabela 7: Estatísticas e estimativas dos parâmetros de efeitos aleatórios do modelo 1..... 47
Tabela 8: Estatísticas descritivas das estimativas dos efeitos aleatórios do modelo 1..... 47
Tabela 9: Testes de normalidade para a distribuição das estimativas dos parâmetros de efeitos aleatórios do modelo 1..... 49
Tabela 10: Estimativas das razões de chances do modelo 1..... 50
Tabela 11: Valores de QIC para as estruturas de matrizes de covariâncias CS e AR(1). 51
Tabela 12: Valores de QIC para as estruturas de matrizes de covariâncias CS e AR(1) para o modelo sem a interação tratamento × momento..... 51
Tabela 13: Estatísticas de ajuste do modelo 2 51
Tabela 14: Estatísticas e estimativas dos parâmetros do modelo 2..... 52
Tabela 15: Estimativas das razões de chances do modelo 2..... 53
Tabela 16: Estatísticas e p-valores dos contrastes 54

MODELOS LINEARES GENERALIZADOS MISTOS E EQUAÇÕES DE ESTIMAÇÃO GENERALIZADAS PARA DADOS BINÁRIOS APLICADOS EM ANESTESIOLOGIA VETERINÁRIA

Autor: Maicon Vinícius Galdino
Orientadora: Prof(a). Dr(a). Liciana Vaz de Arruda Silveira

RESUMO

O objetivo deste trabalho é propor, comparar e selecionar modelos estatísticos baseados na metodologia dos MLGMs e das EEGs. Analisar em que situações cada modelo pode ser melhor utilizado. Discorrer sobre as técnicas de estimação dos parâmetros, escolha da estrutura da matriz de covariância, definição e utilidade das razões de chances e a aplicação em um conjunto de dados da área de anestesiologia veterinária envolvendo equinos.

Os dados utilizados neste trabalho foram obtidos por Taffarel (2013). Os vinte e quatro equinos utilizados no experimento foram agrupados da seguinte maneira: (1) animais anestesiados (tratamento 1), (2) animais anestesiados e que receberam analgesia prévia (tratamento 2), (3) animais anestesiados e submetidos à orquiectomia com administração de analgésico após o processo operatório (tratamento 3) e (4) animais anestesiados e submetidos à orquiectomia com administração de analgésico antes de serem operados (tratamento 4).

Os animais foram observados por cinco avaliadores em diferentes momentos: uma hora antes do procedimento cirúrgico ou anestésico (M1), quatro horas após a recuperação anestésica e antes da aplicação de analgésicos nos animais do tratamento 3 (M2), duas horas após M2 (M3) e

vinte e quatro horas após a cirurgia (M4). A variável resposta (com distribuição Bernoulli) era “olhar o flanco”, um comportamento comum em equinos com dores abdominais.

O tratamento 3 e o momento 2 foram os efeitos que apresentaram as maiores razões de chances de ocorrência da variável resposta (independente de qual metodologia estatística foi empregada na análise) quando comparados com os demais tratamentos e momentos.

Ambos os modelos estatísticos ajustados aos dados (modelos 1 e 2) obtidos através das EEGs e dos MLGMs mostraram-se ser uma poderosa ferramenta para explicar a variável resposta “olhar o flanco” principalmente em relação a seleção das covariáveis estatisticamente significativas.

Nas EEGs, como não se exige a especificação da distribuição dos vetores-resposta, esta tornou-se uma metodologia estatística bastante vantajosa, principalmente para dados binários como os analisados neste trabalho.

Os MLGMs, por terem como uma de suas suposição a heterogeneidade de indivíduo para indivíduo, estes se mostraram adequados quando o objetivo de inferência é sobre um indivíduo específico.

GENERALIZED LINEAR MIXED MODELS AND GENERALIZED ESTIMATING EQUATIONS FOR BINARY DATA APPLIED IN VETERINARY ANESTHESIOLOGY

Author: Maicon Vinícius Galdino
Adviser: Prof(a). Dr(a). Liciana Vaz de Arruda Silveira

SUMMARY

The objective of this work is to propose, compare and select statistical models based on the methodology of GLMMs and GEEs. Examine in which situations each model can be better used. Discuss about the techniques of parameter estimation, choice of the covariance matrix structure, definition and utility of the odds ratios and application in a set of data in the area of veterinary anesthesiology involving equines.

The data used in this work were obtained by Taffarel (2013). The twenty-four equines used in the experiment were grouped as follows: (1) anesthetized animals (treatment 1), (2) animals anesthetized and that received prior analgesia (treatment 2), (3) animals anesthetized and underwent orchiectomy with administration of analgesics after surgery process (treatment 3) and (4) animals anesthetized and underwent orchiectomy with administration of analgesics before being operated (treatment 4).

The animals were observed by five evaluators at different times: one hour before surgery or anesthetic (M1), four hours after recovery from anesthesia and before the application of analgesics in animals of treatment 3 (M2), two hours after M2 (M3) and twenty-four hours after surgery (M4). The response variable (with Bernoulli distribution) was "look the flank", a common behavior in equines with abdominal pain.

The treatment 3 and the moment 2 were the effects with the highest occurrence of odds ratios of the response variable (regardless of what statistical methodology used in the analysis) compared to other treatments and moments.

Both statistical models fitted to the data (models 1 and 2) obtained through the EEGs and MLGMs shown to be a powerful tool to explain the response variable "look flank" especially regarding the selection of statistically significant covariates.

In the EEG, is not required as the specification of the distribution of vectors response, it became quite advantageous statistical methods, particularly for binary data as analyzed in this work.

The MLGMs for having as one of its assumption heterogeneity among individuals, these were suitable when the goal of inference is about a specific individual.

1) INTRODUÇÃO

É comum em pesquisas, experimentos ou ensaios clínicos, o pesquisador estar interessado se certo evento ocorreu ou não (como por exemplo, se ocorreu ou não óbito de um paciente, se houve ou não recidiva de certa doença em mamíferos e se certa praga em uma lavoura criou ou não resistência a um determinado pesticida, dentre outros) gerando assim dados binários. Além disso, é importante estudar quais variáveis contribuíram para que o evento de interesse ocorresse.

Neste sentido, a estatística é utilizada para se estudar a relação entre as variáveis. Por exemplo, analisar a influência de variáveis explicativas sobre uma variável resposta de interesse ou sobre o seu valor médio esperado. De um modo geral, os estatísticos utilizam um modelo de regressão múltipla para relacionar a variável resposta com as variáveis explicativas (regressoras). Esta metodologia é uma das mais amplas e eficientes para modelar dados.

Para se analisar dados binários, os métodos estatísticos tradicionais, em especial os modelos lineares gerais (Gaussianos) e/ou a análise de variância (ANOVA), são inadequados, visto que estes supõem que a variável resposta (contínua) tenha distribuição Normal como estrutura probabilística, homocedasticidade (dispersão em torno do valor médio constante) e independência entre as observações, o que não ocorre em situações descritas acima, pois a variável resposta de interesse (denotada por Y) recebe o valor “1” se ocorreu o evento de interesse (sucesso) e “0” caso contrário (fracasso). Portanto, Y tem distribuição Bernoulli (variável discreta binária).

Até o início da década de 70 faziam-se transformações nesta variável (na proporção de ocorrência do evento), que produziam alterações na relação entre a variável resposta e as variáveis explicativas o que dificultava a interpretação do modelo, mostrando-se pouco eficiente.

Como solução para esse problema Nelder e Wedderburn (1972) propuseram os modelos lineares generalizados (MLGs), que unificou os

procedimentos de inferência tanto para dados contínuos como discretos, eliminando a necessidade de transformação na variável resposta.

Além disso, em experimentos, algo comumente feito são medidas repetidas em um mesmo indivíduo. Muitas vezes por interesse científico (quando um médico deseja estudar o efeito ao longo do tempo de um antinflamatório aplicado em pacientes, por exemplo) e outras por custos ou escassez de unidades amostrais (no caso em que se deseja estudar uma doença rara ou uma pesquisa com animais silvestres). Nestes casos deve-se levar em consideração uma possível correlação entre as diferentes medidas realizadas em um mesmo indivíduo. Assim a suposição de independência intra-indivíduo pode não ser válida.

“Um perigo em potencial que pode ocorrer ao desconsiderar a correlação entre as observações é que, ao ajustar um modelo de regressão para dados correlacionados, uma covariável pode não apresentar significância estatística, mas se ajustarmos um modelo de regressão para dados independentes, esta mesma covariável apresente significância estatística”. (HOSMER, LEMESHOW e STURDIVANT, 2013).

“A vantagem deste tipo de estudo longitudinal (em que as variáveis respostas estão sendo medidas repetidamente sobre o mesmo indivíduo ao longo do tempo) é que se pode distinguir as mudanças ocorridas dentro de cada indivíduo em diferentes momentos fixos.” (MYERS et al, 2010).

Para variáveis contínuas, uma opção para se analisar tais dados seria a utilização da análise de variância para medidas repetidas (ANOVA-MR), mas para dados binários a ANOVA-MR é inapropriada, pois esta supõe distribuição Normal multivariada, variâncias iguais em todos os momentos e correlação constante entre quaisquer dois momentos (algo difícil de obter em experimentos da área de medicina veterinária, pois é esperado, por exemplo, que o paciente mude sua situação clínica ao longo do tempo).

Dentre as diferentes maneiras de se analisar dados binários correlacionados (sob o enfoque da estatística clássica), duas delas se mostram

importantes: os modelos lineares generalizados mistos (MLGMs) e as equações de estimação generalizadas (EEGs).

Os MLGMs (também chamados de modelos sujeitos-específicos) têm por objetivo estimar os parâmetros associados aos efeitos fixos e aos efeitos aleatórios. Entende-se por efeitos fixos aqueles em que os níveis do fator são controlados (escolhidos) pelo pesquisador, como por exemplo, quantidade de fertilizante que será aplicado em uma determinada região (o fator é a quantidade de fertilizante e os níveis são fixados, tais como: 0, 15, 30, 45 e 60 kg/ha); e efeitos aleatórios aqueles em que os níveis do fator são uma amostra aleatória da população dos possíveis níveis. Por exemplo, suponhamos que certa indústria de derivados de leite queira saber se as fazendas fornecedoras interferem na qualidade de seus produtos. Como o número de fazendas (N) é grande torna-se inviável fazer o experimento com todas, então o experimentador sorteia e retira uma amostra de fazendas para o experimento. Assim o estudo trará informações sobre a população de fazendas, não apenas para as fazendas amostradas.

Um das suposições básicas dos MLGMs é o pressuposto de heterogeneidade das unidades amostrais do experimento. Eles procuram explicar a variabilidade natural intrínseca das unidades amostrais por adicionar no modelo efeitos aleatórios não observáveis nos quais as covariáveis (inseridas no modelo) foram incapazes de capturá-la. Por exemplo, por mais homogênea que seja a amostra, o tempo de reação de determinado princípio ativo em pacientes submetidos a experimentos farmacocinéticos difere uns dos outros por diversos fatores não observados. Recorrendo aos MLGMs, o pesquisador poderá explicar esta variabilidade extra.

As EEGs (também chamadas de modelos população-média), desenvolvidas por Liang e Zeger (1986) propõem analisar dados com medidas repetidas utilizando MLGs. Eles obtêm estimativas consistentes para os parâmetros de regressão, utilizando as EEGs e tratam os parâmetros de correlação (isto é a dependência nas respostas) como sendo de perturbação. Além disso, as EEGs são robustas para uma má especificação da matriz de correlação (para um tamanho amostral n suficientemente grande).

Prentice e Zhao (1991) utilizam EEGs para obter estimativas consistentes dos parâmetros de regressão e de correlação e, neste caso, é necessário que tanto o modelo de regressão como a estrutura da matriz de correlação estejam corretamente especificados. Métodos para análise de dados que ignoram essa dependência entre as observações tendem a subestimar os erros-padrões (ZEGER, KARIM, 1991; DIGGLE, LIANG e ZEGER, 1994).

Fitzmaurice, Laird e Ware (2011) citam uma situação que ilustra bem a diferença entre os objetivos práticos dos MLGMs e das EEGs: “se um médico está interessado no efeito (ao longo do tempo) de um determinado fármaco em um paciente específico, então o recomendado é utilizar MLGMs. Por outro lado, se uma empresa farmacêutica está interessada no efeito médio do mesmo fármaco na população, então as EEGs se adaptam melhor a situação”. Portanto, as duas metodologias são recomendadas quando se trabalha com dados correlacionados.

Entre as metodologias estatísticas descritas (MLGMs e EEGs), não há uma que seja “melhor” ou “ideal”. Cada uma tem seus objetivos específicos, vantagens e desvantagens (principalmente quanto a diferença na interpretação dos parâmetros de seus modelos).

Diante do exposto acima, este trabalho ficou estruturado da seguinte maneira: o capítulo 2 mostrará os objetivos, justificativa e relevância deste trabalho. O capítulo 3 discorrerá sobre as metodologias estatísticas (MLGMs e EEGs) e, para cada uma delas abordará a estrutura dos modelos, métodos de estimação e inferência dos parâmetros. Os capítulos 4 e 5 descreverão brevemente algumas estruturas de matrizes de covariâncias utilizadas na modelagem envolvendo MLGMs e EEGs e razões de chances, respectivamente. O capítulo 6 fará a descrição dos dados, a metodologia aplicada no experimento (em anestesiologia veterinária envolvendo equinos) e quais são as covariáveis que serão incorporadas na análise dos dados e na construção dos modelos. O capítulo 7 ficou reservado a análise e aos resultados obtidos e o capítulo 8 destinado as conclusões e considerações finais.

2) OBJETIVOS E JUSTIFICATIVA

O objetivo deste trabalho é propor, comparar e selecionar modelos estatísticos no enfoque dos MLGMs e das EEGs. Analisar em que situações cada metodologia pode ser melhor utilizada. Discorrer sobre as metodologias de estimação dos parâmetros, escolha da estrutura da matriz de covariância, definição e utilidade das razões de chances e a aplicação em um conjunto de dados da área de anestesiologia veterinária.

A justificativa e relevância desse trabalho baseiam-se na importância do tema apresentado, o qual está engajado com a linha de pesquisa (bioestatística) do programa de pós-graduação em Biometria onde tem por finalidade o uso de metodologias estatísticas no estudo de problemas biológicos. Além disso, ter contato com pesquisadores da área de medicina veterinária propiciou ao aluno habilidades e competências essenciais que se deve ter ao trabalhar com profissionais e/ou equipes multidisciplinares.

3) FUNDAMENTAÇÃO TEÓRICA

A necessidade de distinguir os modelos para analisar dados longitudinais discretos de acordo com a interpretação dos coeficientes leva ao uso de termos como modelos marginais (no contexto das EEGs) e modelos com efeitos aleatórios (MLGMs). Estes modelos, devido ao fato da interpretação dos coeficientes serem diferentes, têm objetivos de inferência distintos. Para os modelos marginais o objetivo da inferência é a população enquanto que para os modelos com efeitos aleatórios o objetivo da inferência é o indivíduo. A escolha do modelo para analisar os dados longitudinais discretos tem então de ser feita tendo em conta a questão a que se quer dar resposta.

Para o estudo de análise de dados com medidas repetidas através das EEGs e dos MLGMs utilizamos alguns modelos estatísticos (para resposta binária) que serão descritos conforme se segue.

3.1) Equações de Estimação Generalizadas

Como não é uma tarefa simples a obtenção de um correspondente multivariado para distribuições de probabilidades discretas para vetor(es) resposta para modelos marginais, precisamos de opções que não levam em consideração a função de verossimilhança. Uma delas foi proposta por Liang e Zeger (1986) e tal método é baseado na ideia de “equações de estimação”.

Em modelos marginais, este método baseia-se na estimação por quase-verossimilhança e como tal não necessita que a distribuição conjunta de $\boldsymbol{Y}_i$ seja especificada. A sua popularidade deve-se ao fato de combinar, de forma simples e flexível, a modelagem do valor médio da variável resposta em função das covariáveis, isto é, a componente de maior interesse, com a acomodação da dependência entre as observações sobre o mesmo indivíduo sem necessidade de especificar a estrutura de dependência.

3.1.1) Estrutura do modelo marginal

Os modelos marginais são uma maneira natural de estender os modelos lineares generalizados a dados longitudinais (discretos ou contínuos). Num modelo marginal o valor esperado marginal $E(Y_{it})$ é modelado como função das covariáveis. “O termo valor esperado marginal é usado para realçar o fato de que a resposta média não está condicionada a outras variáveis resposta ou a efeitos aleatórios não observáveis” (FITZMAURICE, LAIRD E WARE, 2011).

Como as medições repetidas, em cada indivíduo, não têm a tendência de ser independente, a análise marginal tem que incluir pressupostos em relação à correlação.

Seja y_{it} a resposta observada do i-ésimo indivíduo no tempo t, em que $t = 1, \dots, T_i$ e $i = 1, \dots, n$, realização da variável resposta Y_{it}. O vetor das variáveis resposta para o i-ésimo indivíduo é dado por $\boldsymbol{Y}_i = (Y_{i1}, \dots, Y_{iT_i})^{\mathrm{T}}$ e $\mathbf{x}_{it}^{\mathrm{T}} = (1, x_{it1}, \dots, x_{it(p-1)})$ representa o vetor $p \times 1$ das covariáveis associado ao ponto (i, t) do experimento. Para cada componente de $\boldsymbol{Y}_i$, Y_{it}, tem-se $E\left(Y_{it}\right) = \mu_{\mathrm{it}}$ e $var\left(Y_{it}\right) = v_{\mathrm{it}}$.

O modelo linear clássico pode ser escrito na forma:

$$Y_{it} = \beta_0 + \beta_1 x_{it1} + \beta_2 x_{it2} + \cdots + \beta_{(p-1)} x_{it(p-1)} + \epsilon_{it} = \boldsymbol{x}_{it}^{T}\boldsymbol{\beta} + \epsilon_{it}$$

em que $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1})^T$ é o vetor $p \times 1$ dos parâmetros de regressão desconhecidos e ϵ_{it} é uma variável aleatória com valor esperado zero.

Considera-se que Y_{it} tem uma distribuição pertencente à família exponencial cuja densidade marginal é definida por:

$$f(y_{it}) = \exp\left[\frac{\omega_{it}}{\phi}\left(y_{it}\theta_{it} - c(\theta_{it})\right) + d(y_{it}, \phi)\right].$$

Segundo Fitzmaurice, Laird e Ware (2011) um modelo marginal para dados longitudinais tem as seguintes características:

1) A esperança marginal para cada variável resposta,

$$E(Y_{it}) = \mu_{it} = \frac{\partial c(\theta_{it})}{\partial \theta_{it}},$$

depende das covariáveis, x_{it} através de uma função de ligação,

$$g(\mu_{it}) = \eta_{it} = \boldsymbol{x}_{it}^{T}\boldsymbol{\beta}\ ,$$

tal como a identidade para respostas contínuas normalmente distribuídas, o $logit$ para respostas binárias e o log para contagens;

2) A variância marginal depende do valor médio marginal de acordo com

$$var(Y_{it}) = v_{it} = V(\mu_{it})\frac{\phi}{\omega_{it}} = \frac{\partial^2 c(\theta_{it})}{\partial \theta_{it}^2}\frac{\phi}{\omega_{it}},$$

em que $V(.)$ é uma função de variância conhecida, ϕ é um parâmetro de dispersão ou de escala (que pode ser conhecido ou estimado) e ω_{it} é uma constante conhecida;

3) A correlação entre Y_{it} e $Y_{it'}$ (em que t' denota o tempo em que a variável resposta Y_i foi observada, sendo $t \neq t'$) é uma função do valor médio marginal e, possivelmente, de parâmetros adicionais, α; e escreve-se $corr(Y_{it}, Y_{it'}) = \rho(\mu_{it}, \mu_{it'})$ em que $\rho(.)$ é uma função conhecida

As duas primeiras hipóteses do modelo marginal descrevem o valor médio e a variância de Y_{it}, e correspondem ao MLG inicialmente proposto por McCullagh e Nelder (1989). A terceira hipótese, com a incorporação da correlação entre as respostas referentes ao mesmo indivíduo, que representa a principal extensão dos MLGs para dados longitudinais (FITZMAURICE, LAIRD e WARE, 2011).

Modelo marginal para dados binários

Seja $\boldsymbol{x}_{it} = (1, t_1, \dots, t_{p-1})^T$ e $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$ e considere-se Y_{it} uma resposta binária tomando os valores 0 e 1 e a função de ligação logística $(logit)$. Neste caso as hipóteses do modelo marginal assumem a forma:

1) $logit(\mu_{it}) = log\left(\frac{\mu_{it}}{1-\mu_{it}}\right) = log\left(\frac{\Pr(Y_{it}=1)}{\Pr(Y_{it}=0)}\right) = \beta_0 + \beta_1 t + \cdots + \beta_{p-1} t_{p-1}$;

2) $var(Y_{it}) = \mu_{it}(1 - \mu_{it})$;

3) $corr(Y_{it}, Y_{it'}) = \alpha$.

Ao modelo marginal acima indicado, dá-se o nome de modelo de regressão logística, o qual assume o pressuposto de variância de Bernoulli, $var(Y_{it}) = \mu_{it}(1 - \mu_{it})$.

Para descrever a terceira hipótese, é necessário ser criterioso ao utilizar o termo correlação. Nas variáveis resposta discretas, as correlações estão condicionadas pelo valor médio da variável resposta, e vice-versa, como é o caso em que a variável resposta é binária.

A correlação entre duas respostas binárias é dada por (DIGGLE et al, 2002):

$$corr(Y_{it-1}, Y_{it}) = \frac{\Pr(Y_{it-1}=1, Y_{it}=1) - \mu_{it-1}\mu_{it}}{[\mu_{it-1}(1-\mu_{it-1})\mu_{it}(1-\mu_{it})]^{1/2}}$$

em que $\Pr(Y_{it-1} = 1, Y_{it} = 1)$ satisfaz a desigualdade

$$\max(0, \mu_{it-1} + \mu_{it} - 1) < \Pr(Y_{it-1} = 1, Y_{it} = 1) < \min(\mu_{it-1}, \mu_{it}).$$

Por exemplo, se $\mu_{it-1} = E(Y_{it-1}) = \Pr(\mathrm{Y_{it-1}} = 1) = 0{,}3$ e $\mu_{it} = E(Y_{it}) = \Pr(\mathrm{Y_{it}} = 1) = 0{,}7$ então $corr(Y_{it-1}, Y_{it}) \leq 0{,}49$. O que significa que a correlação não pode exceder 0,49 quando as probabilidades de sucesso têm os valores 0,3 e 0,7. (ZUUR et al, 2009).

Com este resultado, verifica-se que, quando a variável resposta é binária, a correlação não é a medida mais natural de associação entre as respostas para um mesmo indivíduo. Neste caso, muitos autores utilizam a razão de chances que será descrita com mais detalhes no capítulo 5.

O vetor $\boldsymbol{\beta}$ representa o efeito das covariáveis no valor médio populacional e não como a resposta do indivíduo depende da covariável.

Ao analisarmos as três hipóteses do modelo marginal, verifica-se que não é necessário qualquer pressuposto em relação à distribuição das observações; apenas é necessário um modelo de regressão para o valor médio da variável resposta.

Esta omissão de um pressuposto em relação à distribuição é vantajosa, sobretudo no caso em que a variável resposta é discreta, em que especificar a distribuição conjunta multivariada de $\boldsymbol{Y}_i$ não é trivial. Por outro lado, uma desvantagem ao omitir-se a distribuição de $\boldsymbol{Y}_i$ é que os métodos de estimação baseados na máxima verossimilhança não podem ser utilizados e, por isso, recorre-se ao método das EEGs.

As EEGs, propostas por Liang e Zeger (1986), é a extensão multivariada do método de estimação por quase-verossimilhança. O método das EEGs para $\boldsymbol{\beta}$ tem a mesma forma das equações *score* para os modelos univariados; o que difere é a forma de escolher a matriz de covariância.

Assim, tendo em atenção que as componentes do vetor $\boldsymbol{Y}_i = (Y_{i1}, \dots, Y_{iT_i})^{\mathrm{T}}$ para cada indivíduo não são independentes, o método das EEGs para $\boldsymbol{\beta}$ é dado por:

$$\boldsymbol{U}_{\boldsymbol{\beta}}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^{n} \boldsymbol{D}_i^T \boldsymbol{V}_i[(\boldsymbol{\alpha})]^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0} \ (3.1)$$

em que $\boldsymbol{D}_i = \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}$, é uma matriz $T_i \times p$, $\boldsymbol{Y}_i = \left(Y_{i1}, \dots, Y_{iT_i}\right)$, $\boldsymbol{\mu}_i$ o vetor do valor esperado de $\boldsymbol{Y}_i$ e $\boldsymbol{V}_i(\boldsymbol{\alpha})$ é uma matriz diagonal $T_i \times T_i$ que para o i-ésimo indivíduo é dada por:

$$\boldsymbol{V}_i(\boldsymbol{\alpha}) = \phi(\boldsymbol{A}_i^{1/2} \boldsymbol{R}_i(\boldsymbol{\alpha}) \boldsymbol{A}_i^{1/2}) \ (3.2)$$

Em (3.2) $\boldsymbol{A}_i$ é uma matriz $T_i \times T_i$ cujo elemento na diagonal t é $V(\mu_{it})$, função variância ($var(Y_{it}) = \phi V(\mu_{it})$) e $\boldsymbol{R}_i(\boldsymbol{\alpha}) = corr(\boldsymbol{Y}_i)$ é uma matriz diagonal $T_i \times T_i$ que recebe o nome de "matriz de correlação de trabalho".

Se $\boldsymbol{R}_i(\boldsymbol{\alpha}) = \mathbf{I}$, a matriz identidade $T_i \times T_i$, então o método das EEGs reduz-se às equações de quase-verossimilhança para um modelo linear generalizado que assume que as medidas repetidas são independentes.

Quando $\boldsymbol{R}_i(\boldsymbol{\alpha})$ está corretamente especificada então $\boldsymbol{V}_i(\boldsymbol{\alpha}) = var(\boldsymbol{Y}_i)$.

O sistema de equações (3.1), concebido para estimar $\boldsymbol{\beta}$, necessita de $\boldsymbol{\alpha}$ e ϕ mas a estimação destes depende de $\boldsymbol{\beta}$**.** Este processo de dependência é quebrado através da utilização de um processo iterativo.

O processo iterativo divide-se em quatro passos (DAVIS, 2002; FITZMAURICE, LAIRD e WARE, 2011; MYERS et al, 2010; MOREL e NEERCHAL, 2012):

1) Determinar-se as estimativas iniciais de $\boldsymbol{\beta}$, $\widehat{\boldsymbol{\beta}}^{(0)}$, através do ajustamento de um MLG assumindo que as observações são independentes;

2) (i) Calcular os resíduos de Pearson e_{it};

(ii) Calcular $\boldsymbol{\alpha}$;

(iii) Calcular $\boldsymbol{R}_i\ (\boldsymbol{\alpha})$;

(iv) Calcular ϕ;

(v) Calcular $\boldsymbol{V}_i(\boldsymbol{\alpha}) = \phi(\boldsymbol{A}_i^{1/2}\boldsymbol{R}_i\ (\boldsymbol{\alpha})\ \boldsymbol{A}_i^{1/2})$;

3) Atualizar a estimativa de $\boldsymbol{\beta}$:

$$\widehat{\boldsymbol{\beta}}^{(t+1)} = \widehat{\boldsymbol{\beta}}^{(t)} - \left[\sum_{i=1}^{n} \widehat{\boldsymbol{D}}_i^T\ [\boldsymbol{V}_i(\widehat{\boldsymbol{\alpha}})]^{-1}\widehat{\boldsymbol{D}}_i\right]^{-1} \times \left[\sum_{i=1}^{n} \widehat{\boldsymbol{D}}_i^T\ [\boldsymbol{V}_i(\widehat{\boldsymbol{\alpha}})]^{-1}(y_i - \widehat{\boldsymbol{\mu}}_i)\right].$$

4) Repetir os passos 2-3 até o processo convergir.

As estimativas de $\boldsymbol{\alpha}$ e ϕ são facilmente obtidas a partir dos resíduos de Pearson, e_{it}, definidos por

$$e_{it} = \frac{(y_{it} - \hat{\mu}_{it})}{\sqrt{var(\hat{\mu}_{it})}}$$

em que o parâmetro de dispersão ou de escala, ϕ, é estimado através de

$$\hat{\phi} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T_i}\sum_{i=1}^{T_i} e_{it}^2$$

e a estimação de $\boldsymbol{\alpha}$, pode ser obtida de forma semelhante, dependendo da estrutura de correlação definida. Os pormenores das estruturas de matrizes de correlação de trabalho serão descritos no capítulo 4.

3.1.2) Estimação e inferência

Liang e Zeger (1986) mostraram que, sob certas condições de regularidade, o estimador de $\boldsymbol{\beta}$, $\widehat{\boldsymbol{\beta}}$, solução das EEGs, possui as seguintes propriedades:

1. É um estimador consistente de $\boldsymbol{\beta}$.

Esta propriedade é válida quer se tenha modelado corretamente a estrutura de correlação ou não (MYERS et al, 2010; FITZMAURICE, LAIRD e WARE, 2011; STOKES, DAVIS e KOCK, 2000). Desta forma, para que $\widehat{\boldsymbol{\beta}}$ proporcione estimativas válidas de $\boldsymbol{\beta}$ necessita-se apenas que o modelo para o valor médio da variável resposta tenha sido corretamente especificado.

2. Para grandes amostras, a distribuição amostral de $\widehat{\boldsymbol{\beta}}$ é Normal multivariada, com valor médio $\boldsymbol{\beta}$ e matriz de covariância

$$I_0^{-1} I_1 I_0^{-1} \ (3.3)$$

sendo que

$$I_0 = \sum_{i=1}^{n} \mathbf{D}_i^{\mathrm{T}} [\mathbf{V}_i(\boldsymbol{\alpha})]^{-1} \mathbf{D}_i$$

e

$$I_1 = \sum_{i=1}^{n} \boldsymbol{D}_i^{\mathrm{T}} \, [\mathbf{V}_i(\boldsymbol{\alpha})]^{-1} var(Y_i) [\mathbf{V}_i(\boldsymbol{\alpha})]^{-1} \mathbf{D}_i$$

calculada em $(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\alpha}})$. Este estimador da matriz de covariância é consistente à medida que o número de indivíduos que contribuem para cada elemento da

matriz tende para infinito (DIGGLE et al, 2002) e é conhecido como estimador "*sandwich*" de $var(\widehat{\boldsymbol{\beta}})$.

Embora com a metodologia das EEGs se obtenham estimadores consistentes de $\boldsymbol{\beta}$, mesmo quando a matriz de correlação de trabalho não está corretamente especificada, os erros-padrão usuais dos estimadores não são consistentes nesta situação. Felizmente, a utilização do estimador "*sandwich*" de $var(\widehat{\boldsymbol{\beta}})$ permite obter, em muitas situações, erros padrão válidos para $\widehat{\boldsymbol{\beta}}$. Baseados no estimador "*sandwich*" obtêm-se versões robustas do teste Wald.

Apesar da importante propriedade de somente ser necessário que o modelo para o valor médio da variável resposta seja corretamente especificado para obter estimadores consistentes de $\boldsymbol{\beta}$, a escolha da matriz de covariância é bastante importante. São duas as razões para esse fato. Primeiro, porque quanto mais próxima a matriz de covariância, $\boldsymbol{V}_i$, estiver da verdadeira matriz de covariância, maior é a eficiência ou precisão com a qual $\boldsymbol{\beta}$ pode ser estimado. Segundo, porque a propriedade de robustez é uma propriedade assintótica, que é válida apenas para grandes amostras. (FITZMAURICE, LAIRD e WARE, 2011).

Quando se utiliza o método das EEGs o teste de razão de verossimilhanças não pode ser usado em virtude da estimação ter se baseado na quase-verossimilhança. No entanto, tendo em atenção à distribuição assintótica dos estimadores das EEGs, o teste de Wald pode ser usado.

Por exemplo, quando o objetivo é testar

$$H_0: \boldsymbol{\beta}_j = 0 \; vs \; H_0: \boldsymbol{\beta}_j \neq 0$$

para $j = 1, \dots, p$ a estatística de teste de Wald é dada por

$$W = \frac{\widehat{\boldsymbol{\beta}}_j}{\sqrt{var(\widehat{\boldsymbol{\beta}}_j)}}$$

que, sob H_0, tem distribuição assintótica Normal e $var(\widehat{\boldsymbol{\beta}}_j)$ é o j-ésimo elemento da diagonal da matriz (3.3).

Para testar a hipótese mais geral

$$H_0: \boldsymbol{L\beta} = 0 \; vs \; H_0: \boldsymbol{L\beta} \neq 0$$

em que $\boldsymbol{L}$ é uma matriz $r \times p$ de posto completo r, a estatística de teste de Wald é dada por:

$$W = \left(\boldsymbol{L}\widehat{\boldsymbol{\beta}}\right)^T \left[\boldsymbol{L}\, var\left(\widehat{\boldsymbol{\beta}}\right)\boldsymbol{L}^T\right]^{-1} \left(\boldsymbol{L}\widehat{\boldsymbol{\beta}}\right)$$

que, sob H_0, tem uma distribuição assintótica qui-quadrado com r graus de liberdade (Davis, 2002).

3.2) Modelo Linear Generalizado Misto

Os MLGMs, propostos por Breslow e Clayton (1993), são uma extensão dos MLGs que permitem a adição de componentes de variabilidade devida a efeitos não observados. Os MLGMs são, portanto, MLGs que incluem efeitos aleatórios no preditor linear, além dos efeitos fixos.

A introdução destes efeitos aleatórios no preditor linear permite modelar a estrutura de correlação entre observações que pertencem ao mesmo indivíduo. A correlação entre observações de um mesmo indivíduo surge assim do fato destas partilharem os mesmos efeitos aleatórios não observáveis.

Os MLGMs têm como objetivo descrever as alterações da resposta média de cada indivíduo e a relação destas alterações com as covariáveis. Assim sendo, estes modelos têm como finalidade realizar inferências para o indivíduo e não para a população.

3.2.1) Estrutura do modelo linear generalizado misto

Tal como na seção 3.1.1, seja y_{it} a resposta observada no indivíduo i para o tempo t, realização da variável resposta Y_{it} em que $t = 1, \dots, T_i$ e $i = 1, \dots, n$ sendo $\boldsymbol{x}_{it}^T = (1, x_{it1}, \dots, x_{it(p-1)})$ o vetor $p \times 1$ das covariáveis associadas

a cada vetor resposta Y_{it} e $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1})^T$o vetor $p \times 1$ dos parâmetros de regressão desconhecidos. Considere-se agora $\boldsymbol{z}_{it}$ um vetor $q \times 1$ de efeitos aleatórios $\boldsymbol{b}_i$.

Os MLGMs podem ser especificados do seguinte modo (DIGGLE et al, 2002):

1) Dado $\boldsymbol{b}_i$, as variáveis resposta $(Y_{i1}, \dots, Y_{iTi})$, são mutuamente independentes e seguem um MLG com densidade

$$f(Y_{it}|\boldsymbol{b}_i) = exp\left[\frac{\omega_{it}}{\phi}\big(y_{it}\theta_{it} - c(\theta_{it})\big) + d(y_{it}, \phi)\right].$$

O valor médio e a variância condicionais são dados, respectivamente, por:

$$E(Y_{it}|\boldsymbol{b}_i) = \mu_{it}^b = \frac{\partial c(\theta_{it})}{\partial \theta_{it}}$$

e

$$var(Y_{it}|\boldsymbol{b}_i) = v_{it}^b = \frac{\partial^2 c(\theta_{it})}{\partial \theta_{it}^2}\frac{\phi}{\omega_{it}},$$

que se assume que satisfaçam,

$$g\big(\mu_{it}^b\big) = \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \mathbf{z}_{it}^{\mathrm{T}}\mathbf{b}_i$$

e

$$v_{it}^b = V(\mu_{it}^b)\frac{\phi}{\omega_{it}}$$

em que $g(.)$ é uma função de ligação e $V(\mu_{it}^b)$ uma função de variância, ambas conhecidas, ϕ é um parâmetro de dispersão ou de escala e ω_{it} é uma constante conhecida;

2) Os efeitos aleatórios, $\boldsymbol{b}_i$, $i = 1, \dots, n$ são independentes entre si com uma distribuição multivariada comum F.

De uma maneira geral a equação dos MLGMs podem ser escrita na forma:

$$g\{E(Y_{it}|\boldsymbol{b}_i)\} = \eta_{it}^b = \boldsymbol{x}_{it}^T\boldsymbol{\beta} + \boldsymbol{z}_{it}^T\boldsymbol{b}_i$$

com as especificações dadas em 1 e 2, sendo a distribuição comum aos efeitos aleatórios $\boldsymbol{b}_i$, Normal multivariada com valor esperado zero e matriz de covariância $\boldsymbol{D}$.

Modelo linear generalizado misto para dados binários

Seja $\boldsymbol{z}_{it} = \boldsymbol{x}_{it} = (1, t_1, \dots, t_{p-1})^T$ e $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$ e $\boldsymbol{b}_i = (b_{0i}, b_{1i}, \dots, b_{(p-1)i})^T$, em que se assume que $\boldsymbol{b}_i$ é um vetor aleatório independente com uma distribuição Normal multivariada, com valor médio zero e matriz de covariância, $\boldsymbol{D}$.

Y_{it} é uma resposta bináhria que assume os valores 0 e 1 e a função de ligação logística, $logit$, é a escolha natural para a função de ligação. Neste caso as hipóteses do MLGM são:

1) Condicional ao vetor de efeitos aleatórios $\boldsymbol{b}_i$, Y_{it} são independentes e seguem uma distribuição Bernoulli, com $E(Y_{it}|\boldsymbol{b}_i) = \mu_{it}^b$ e $var(Y_{it}|\boldsymbol{b}_i) = E(Y_{it}|\boldsymbol{b}_i)\{1 - E(Y_{it}|\boldsymbol{b}_i)\} = \mu_{it}^b(1 - \mu_{it}^b)$, $(\phi = 1)$;

2) $logit(\mu_{it}^b) = \boldsymbol{x}_{it}^T\boldsymbol{\beta} + \boldsymbol{z}_{it}^T\boldsymbol{b}_i = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_1 + \cdots + (\beta_{p-1} + b_{(p-1)i})t_{p-1}$.

Interpretação dos coeficientes do modelo

A introdução de efeitos aleatórios nos MLGs, juntamente com covariáveis associadas ao experimento, permite modelar a estrutura de correlação entre observações que pertencem ao mesmo indivíduo. No entanto, no modelo linear geral (Gaussiano) esses efeitos não afetam a interpretação dos parâmetros fixos do modelo; já nos MLGs a interpretação desses efeitos é afetada.

A esperança de Y_{it}, $E(Y_{it})$, é dada por

$$E(Y_{it}) = E\{E(Y_{it}|\boldsymbol{b}_i)\} = E[g^{-1}(\boldsymbol{x}_{it}^T\boldsymbol{\beta} + \boldsymbol{z}_{it}\boldsymbol{b}_i)].$$

Em geral, não é possível simplificar esta expressão devido a não linearidade da função $g^{-1}(.)$ e, como tal, De Boeck, Naberezny e Tavares (2011) mencionam que:

$$E[g(Y_{it}|\boldsymbol{b}_i)] \neq g[E(Y_{it}|\boldsymbol{b}_i)].$$

Assim, nos MLGMs têm-se as seguintes interpretações para os parâmetros fixos e para a componente aleatória do modelo:

(i) Parte fixa: As componentes de $\boldsymbol{\beta}, \beta_k$, têm uma interpretação em termos específicos do indivíduo. Isto é, representam a influência das covariáveis na alteração da resposta média esperada de um indivíduo específico.

(ii) Parte aleatória: As componentes de $\boldsymbol{b}_i$ expressam a heterogeneidade natural de cada indivíduo específico.

3.2.2) Estimação e inferência

Os MLGMs permitem a acomodação de variáveis resposta com distribuições não-Normais, da especificação de funções não lineares entre o valor esperado da resposta e dos preditores, e podem modelar a correlação através de efeitos aleatórios. No sentido de serem feitas inferências corretas, esta correlação tem de ser considerada e a distribuição conjunta, quer do vetor das respostas, quer do vetor de efeitos aleatórios, tem de ser completamente especificada.

Como resultado a função de verossimilhança é usada e, consequentemente, a estimação e a inferência baseiam-se no método da máxima verossimilhança.

No MLGM a contribuição do i-ésimo indivíduo para a função de verossimilhança é:

$$L_i^A(\boldsymbol{\beta}, \phi, \boldsymbol{D}) = \int \prod_{t=1}^{T_i} f(y_{it}|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi) f(\boldsymbol{b}_i|\boldsymbol{D}) d\boldsymbol{b}_i.$$

O sobrescrito A em $L_i^A(\boldsymbol{\beta}, \phi, \boldsymbol{D})$ é uma notação utilizada para indicar que se trata da função de verossimilhança do modelo com efeitos aleatórios. Para os n indivíduos a função de verossimilhança é apresentada na forma:

$$L^A(\boldsymbol{\beta}, \phi, \boldsymbol{D}) = \prod_{i=1}^{n} L_i^A(\boldsymbol{\beta}, \phi, \boldsymbol{D}) = \prod_{i=1}^{n} \int L_i^F(\boldsymbol{\beta}, \phi|\boldsymbol{D})\, f(\boldsymbol{b}_i|\boldsymbol{D}) d\boldsymbol{b}_i,$$

com $L_i^F(\boldsymbol{\beta}, \phi|\boldsymbol{D}) = \prod_{t=1}^{T_i} f(y_{it}|b_i, \beta, \phi)$, em que o sobrescrito F refere-se ao modelo de efeitos fixos.

O logaritmo da função de verossimilhança é dado por:

$$l^A(\beta, \phi, \boldsymbol{D}) = \sum_{i=1}^{n} \log L_i^A(\boldsymbol{\beta}, \phi, \boldsymbol{D})$$

que assume a expressão

$$l^A(\boldsymbol{\beta}, \phi, \boldsymbol{D}) = \sum_{i=1}^{n} \log \int L_i^F(\boldsymbol{\beta}, \phi|\boldsymbol{D})|\boldsymbol{D}|^{-1/2} \times \frac{1}{(2\pi)^{q/2}} \exp\left(-\frac{1}{2}\mathbf{b}_\mathrm{i}^\mathrm{T}\mathbf{D}^{-1}\boldsymbol{b}_i\right) d\boldsymbol{b}_i.$$

Em geral, o problema não tem uma solução analítica. Desta forma faz-se necessário recorrer a métodos de cálculo numérico para realizar inferência com base na máxima verossimilhança. Os métodos mais conhecidos são: a aproximação dos dados, do integrando (aproximação de Laplace) e da integral. Descreveremos brevemente somente o método baseado na aproximação da integral conforme se segue:

Método de aproximação da integral

Dos vários métodos de integração que podem ser usados, muitos dos quais se encontram implementados em vários programas estatísticos, destaca-se o método numérico da quadratura de Gauss-Hermite e o método numérico adaptativo da quadratura de Gauss-Hermite. Estes métodos baseiam-se na aproximação de integrais na forma:

$$\int f(z)\phi(z)dz \quad (3.4)$$

em que $f(z)$ é uma função densidade conhecida e $\phi(z)$ é a função densidade de uma distribuição Normal multivariada. Começa-se por padronizar os efeitos aleatórios de forma a obter a matriz de covariância. Seja $\boldsymbol{\delta}_i$ com $\boldsymbol{\delta}_i = \mathbf{D}^{-1/2}\mathbf{b}_i$ em que $\boldsymbol{\delta}_i$ segue uma distribuição Normal com valor médio e matriz de covariância dada por $\mathbf{I}$. Desta forma o preditor linear assume a forma:

$$\eta_{it} = \boldsymbol{x}_{it}^{\mathrm{T}}\beta + \boldsymbol{z}_{it}^{\mathrm{T}}\mathbf{D}^{-1/2}\boldsymbol{\delta}_i$$

em que $\mathbf{b}_i$ foi substituído por $\mathbf{D}^{-1/2}\boldsymbol{\delta}_i$.

A contribuição do i-ésimo indivíduo para a função de verossimilhança é, assim, dada por

$$\begin{aligned} L_i^A(\boldsymbol{\beta}, \phi, \boldsymbol{D}) &= \int \prod_{t=1}^{T_i} f(y_{it}|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi) f(\boldsymbol{b}_i|\boldsymbol{D}) d\boldsymbol{b}_i \\ &= \int \prod_{t=1}^{T_i} f(y_{it}|\boldsymbol{\delta}_i, \boldsymbol{\beta}, \mathbf{D}, \phi) f(\boldsymbol{\delta}_i) d\boldsymbol{\delta}_i \,. \quad (3.5) \end{aligned}$$

A expressão (3.5) está, portanto, na forma da expressão (3.4), a qual é necessária para aplicar o método da quadratura de Gauss-Hermite (ou o método adaptativo da quadratura de Gauss-Hermite).

Método da quadratura de Gauss-Hermite

Segundo McCulloch, Searle e Neuhaus (2008) o método da quadratura de Gauss-Hermite, $\int f(z)\phi(z)dz$ é aproximada por uma soma ponderada,

$$\int f(z)\phi(z)dz \approx \sum_{q=1}^{Q} w_q f\left(z_q\right),$$

em que Q é a ordem da aproximação. Quanto maior for Q mais precisa é a aproximação. Os pontos de quadratura, z_q, são as soluções de um polinômio de Hermite de ordem Q e w_q são os pesos escolhidos. Os pontos de quadratura, z_q, e os pesos w_q encontram-se tabelados.

Método adaptativo da quadratura de Gauss-Hermite

No método adaptativo da quadratura de Gauss-Hermite, os pontos de quadratura são centrados e dimensionados para que $f(z)\phi(z)$ assuma uma distribuição Gaussiana. O valor médio desta distribuição será igual à estimativa, $\hat{z}$, de $\ln[f(z)\phi(z)]$ e a variância igual a

$$\left[-\frac{\partial^2}{\partial z^2}\ln[f(z)\phi(z)]|_{z=\hat{z}}\right]^{-1}.$$

Assim sendo, a quadratura de pontos adaptativa assume a forma,

$$z_q^+ = \hat{z} + \left[-\frac{\partial^2}{\partial z^2}\ln[f(z)\phi(z)]|_{z=\hat{z}}\right]^{-1/2} z_q$$

com respectivos pesos,

$$w_q^+ = \left[-\frac{\partial^2}{\partial z^2}\ln[f(z)\phi(z)]|_{z=\hat{z}}\right]^{-1/2} \frac{\phi(z_q^+)}{\phi(z_q)} w_q.$$

De forma análoga à quadratura de Gauss-Hermite, a integral é aproximada por uma soma ponderada, dada por:

$$\int f(z)\phi(z)dz \approx \sum_{q=1}^{Q} w_q^+ f\left(z_q^+\right).$$

Inferência

Como os estimadores dos MLGMs são obtidos com base no método da máxima verossimilhança, toda a inferência se baseia na função de verossimilhança e nas distribuições assintóticas dos estimadores de máxima verossimilhança. Assumindo que o modelo é apropriado, os estimadores obtidos têm assintoticamente uma distribuição Normal com valor médio igual ao parâmetro a ser estimado e variância dada pelo inverso da matriz de informação de Fisher.

Assim, testes tipo Wald podem ser usados comparando as estimativas padronizadas com a distribuição $N(0,1)$. Hipóteses compostas podem ser

igualmente testadas usando uma formulação mais geral de estatística Wald que neste caso é comparada com a distribuição qui-quadrado.

Os testes de razão de verossimilhança são, em geral, os utilizados para comparar a estrutura fixa de dois modelos encaixados que tenham os mesmos efeitos aleatórios. Quando os modelos não estão encaixados usa-se o critério de informação de Akaike (AIC).

3.3) Modelo Marginal versus Modelo com Efeitos Aleatórios

Os modelos marginais e os MLGMs são ambos uma extensão dos MLGs para dados longitudinais não-normais mas possuem características diferentes, no sentido em que cada um possui diferentes especificações e pressupostos em relação à distribuição conjunta, assim como em relação à correlação ao longo das medições repetidas num mesmo indivíduo.

Quando se discutem as duas abordagens tem-se que levar em conta:

$$E[g(Y_{it}|\boldsymbol{b}_i)] \neq g[E(Y_{it}|\boldsymbol{b}_i)].$$

Considere o caso de dados binários em que o MLGM tem efeito aleatório no intercepto. De acordo com De Boeck, Naberezny e Tavares (2011), o valor esperado condicional, $E(Y_{it}|\boldsymbol{b}_i)$, em função de t é dado por:

$$E(Y_{it}|\boldsymbol{b}_i) = \frac{\exp[(\beta_0 + b_{0i}) + \beta_1 t]}{1 + \exp[(\beta_0 + b_{0i}) + \beta_1 t]}$$

enquanto que o valor esperado marginal $E(Y_{it})$ é dado por:

$$E(Y_{it}) = E[E(Y_{it}|\mathbf{b}_i)] = E\left[\frac{\exp[(\beta_0 + b_{0i}) + \beta_1 t]}{1 + \exp[(\beta_0 + b_{0i}) + \beta_1 t]}\right] \neq \frac{\exp[\beta_0 + \beta_1 t]}{1 + \exp[\beta_0 + \beta_1 t]}$$

O vetor dos parâmetros fixos nos MLGMs modelam a evolução de cada indivíduo separadamente enquanto que o vetor de parâmetros fixos no modelo marginal exprime o modo como, em média, a probabilidade de sucesso evolui na população.

Os parâmetros de regressão do modelo marginal têm uma interpretação que não depende da associação entre variáveis resposta de um mesmo indivíduo. Tem-se assim, que nos modelos marginais os parâmetros de regressão descrevem o efeito das covariáveis na variável resposta média da população. Os MLGMs têm como principal pressuposto a heterogeneidade de indivíduo para indivíduo. A função de ligação não-linear nos MLGM leva a que os parâmetros de regressão descrevem as alterações da variável resposta de cada indivíduo e a relação destas alterações com as covariáveis.

“A escolha entre o modelo marginal e o MLGM para dados longitudinais (não-normais) só pode ser feita tendo em vista o estudo e as questões a que se pretende dar resposta. Neste sentido, qualquer que seja o estudo longitudinal, diferentes questões científicas correspondem diferentes modelos” (FITZMAURICE, LAIRD e WARE, 2011).

4) ESTRUTURAS DE MATRIZES DE COVARIÂNCIAS

Existem várias estruturas para a seleção da matriz de covariância e cada uma delas tem por finalidade a acomodação com os dados em estudo (LITTELL et al, 2006). A especificação da matriz de covariância exerce grande influência na construção e na qualidade dos modelos construídos e por isso, uma atenção especial deve ser dada a esta especificação. Algumas destas estruturas são apresentadas a seguir, considerando uma situação simples em que $t = 4$, ou seja, quatro repetições no tempo:

a) Não estruturada (UN): todas as correlações são assumidas independentes e calculadas a partir do conjunto de dados,

$$\begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_{22} & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix}$$

b) Simetria composta (CS): igualdade de variâncias e covariâncias, ou seja, covariâncias constantes entre quaisquer observações de uma mesma unidade devido a erros independentes (XAVIER, 2000),

$$\begin{bmatrix} \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 \end{bmatrix}$$

c) Simetria composta heterogênea (CSH): parâmetros de variâncias diferentes para cada elemento da diagonal principal e raiz quadrada desses parâmetros nos elementos fora da diagonal principal, sendo σ_i^2 o i-ésimo parâmetro da variância e ρ o parâmetro de correlação satisfazendo $|\rho| < 1$ (XAVIER, 2000),

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

d) Autorregressiva de 1ª ordem (AR(1)): as correlações dependem da distância entre as medições (correlações adjacentes são maiores em magnitude do que as não adjacentes); elas diminuem com o aumento da distância. Adequada quando existe um tempo ou componente de ordem associados com as observações. (HOSMER, LEMESHOW e STURDIVANT, 2013). Adequada para observações igualmente espaçadas no tempo.

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

e) Autorregressiva de 1ª ordem heterogênea (ARH(1)): esta estrutura de matriz de covariância tem o mesmo padrão de correlação da estrutura AR(1), mas as variâncias podem diferir,

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

f) Toeplitz (TOEP): essa estrutura pode ser vista como uma estrutura de médias móveis com ordem igual ao tamanho da matriz (FILHO, 2002),

$$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}$$

g) Toeplitz heterogênea (TOEPH): As correlações desta estrutura estão em faixas como a estrutura TOEP, mas as variâncias podem diferir,

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_3 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{bmatrix}$$

h) Espacial (ou potência) (SP(POW)): todas as variâncias são iguais e a correlação entre as observações separadas por uma distância de d unidades é ρ^d, sendo ρ a correlação entre observações. Seu uso é aconselhado para os

casos em que as observações repetidas não são equidistantes (LITTELL et al (2006)) e seu uso só é possível se a covariável tempo for numérica (SAS INSTITUTE, 2008),

$$\sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \rho^{d_{34}} \\ \rho^{d_{41}} & \rho^{d_{42}} & \rho^{d_{43}} & 1 \end{bmatrix}$$

Somente as estruturas de matrizes de covariâncias dos itens (a), (b) e (d) estão definidas para as EEGs, enquanto que todas as estruturas de matrizes descritas estão definidas para os MLGMs.

No contexto dos MLGMs, o termo mais frequentemente utilizado para as estruturas de matrizes explicitadas acima é “matriz de covariâncias” enquanto que no contexto das EEGs o termo mais usado é “matriz de correlação de trabalho”.

Ao longo do texto iremos nos referir a estas estruturas de matrizes como sendo “estruturas de matrizes de covariâncias”, independentemente se estivermos nos referindo a elas nos MLGMs ou nas EEGs.

Para a escolha adequada da estrutura da matriz de covariância é necessário utilizar algum critério de seleção. Gbur et al. (2012) citam que o critério de informação de Akaike (AIC) pode ser utilizado para comparar as estruturas de covariâncias. Este critério é baseado no logaritmo da função de verossimilhança e dependem do número de observações e de parâmetros do modelo. Suas expressões são dadas por:

$$AIC = -2l\left(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\kappa}}\right) + 2d$$

em que l é o logaritmo da função de verossimilhança, d o número de parâmetros de efeitos fixos e aleatório estimados no modelo, $\widehat{\boldsymbol{\kappa}}$ o vetor de estimativas dos componentes de variância e n o número de observações utilizadas na estimação desses parâmetros. A estrutura de covariâncias com menor valor de AIC é considerada a mais adequada. Este critério somente é válido para os MLGMs por causa da adoção da função de verossimilhança.

Para as EEGs um critério desenvolvido por Pan (2001) e citado por Hosmer, Lemeshow e Sturdivant (2013) é definido como:

$$QIC = -2QL + 2tr(\widehat{\Omega_I}, \widehat{V_R})$$

em que QL é o logaritmo da função de quase-verossimilhança, $\widehat{V_R}$ é a estimativa robusta da matriz de covariâncias dos parâmetros e $\widehat{\Omega_I}$ é a estimativa do inverso das estimativas de covariâncias baseados nos modelos sob a suposição de independência.

Assim como acontece no AIC, menores valores de QIC indicam a “melhor” estrutura de matriz de covariância para ser utilizada na construção dos modelos.

5) RAZÃO DE CHANCES

Às vezes há o interesse em conhecer a chance de certo evento ocorrer (sucesso) de um grupo em específico, como por exemplo, desenvolver uma determinada doença em relação a outro grupo. Para esse caso estaríamos comparando dois grupos e seria feita, por exemplo, a seguinte questão: “Será que o risco (relativo) de uma determinada doença se desenvolver em pessoas expostas é maior do que para aquelas que não foram expostas?” (GORDIS, 2010). E de quanto é esse risco?

Para isso, precisa-se encontrar a chance do resultado estar presente em certo grupo de indivíduos. Allison (2001) define chance da seguinte maneira:

$$chance_j = \frac{p_{ij}}{1 - p_{ij}}$$

em que p_{ij} é a probabilidade de sucesso do indivíduo i no grupo j ($i = 1,2, \dots, n$ e $j = 0,1$).

No caso em que se tem apenas uma variável regressora, o preditor linear do modelo é dado por:

$$\hat{\eta}(x_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Então para $x_i = 1$, quando o grupo for 1 e $x_i = 0$ quando o grupo for 0, Hosmer, Lemeshow e Sturdivant (2013) e Morel e Neerchal (2012) mostram que a razão de chances (“*odds ratio*” em inglês), denotada por $\hat{O}_R$ é dada por:

$$\hat{O}_R = \frac{chance_1}{chance_0} = e^{\hat{\beta}_1}$$

Gordis (2010) diz que, como não é possível calcular o risco relativo em estudos de caso-controle, a razão de chances ($\hat{O}_R$) é uma ótima estimativa do risco relativo quando a incidência da doença na população é rara.

Sob a estimativa da razão de chance é possível calcular o intervalo de confiança (por exemplo, de 95%) para a verdadeira razão de chance. Se o intervalo não contiver o 1, isto indica que a estimativa da razão de chance é significativa ao nível de 5% (COLLETT, 2003).

Para as EEGs, a odds ratio descreve a razão de chances na população (e tem interpretação análoga ao MLG para dados binários independentes) enquanto que para os MLGMs, descreve a razão de chances para indivíduos que possuem o mesmo efeito aleatório (HOSMER, LEMESHOW e STURDIVANT, 2013; FITZMAURICE, LAIRD e WARE, 2011).

6) DESCRIÇÃO DOS DADOS

Os dados utilizados neste trabalho foram obtidos por Taffarel (2013) e publicados em sua tese de doutorado do Programa de Pós-Graduação em Medicina Veterinária da Faculdade de Medicina Veterinária e Zootecnia (FMVZ), UNESP. Os vinte e quatro equinos utilizados no experimento foram agrupados da seguinte maneira: (1) animais que somente receberam anestesia (tratamento 1), (2) animais anestesiados e que receberam analgesia prévia (tratamento 2), (3) animais anestesiados e submetidos à orquiectomia (popularmente conhecida como castração) com administração de analgésico após o processo operatório (tratamento 3) e (4) animais anestesiados e submetidos à orquiectomia com administração de analgésico antes de serem operados (tratamento 4). Os animais de (1) e (2) fizeram parte do grupo controle e os de (3) e (4) formaram o grupo tratado, configurando assim um ensaio clínico.

Os animais foram observados por cinco avaliadores em diferentes momentos: uma hora antes do procedimento cirúrgico ou anestésico (M1), quatro horas após a recuperação anestésica e antes da aplicação de analgésicos nos animais do tratamento 3 (M2), duas horas após M2 (M3) e vinte e quatro horas após a cirurgia (M4).

Recuperação anestésica é o processo pelo qual a animal reassume seus processos fisiológicos anteriores a anestesia e Taffarel (2013) padronizou o tempo da recuperação anestésica em quarenta e cinco minutos para todos os grupos do experimento.

Taffarel (2013) tinha por objetivo validar e refinar uma escala para avaliação da dor aguda pós-operatória em equinos. Para isto foram observadas várias características associadas ao processo concernente a dor tais como: posição na baia, posição do pescoço, interação com o ambiente, apetite, atividade de locomoção, resposta a palpação da área dolorosa, outros comportamentos diversos (indicativos comportamentais) e frequência cardíaca, frequência respiratória, pressão arterial sistólica e motilidade intestinal (indicativos fisiológicos).

Dentre os indicativos comportamentais está o “olhar o flanco”. Mair, Divers e Ducharme (2002) mencionam que “olhar o flanco” é um comportamento comum em equinos afetados por dores abdominais (cólicas) e pode ocorrer num processo doloroso leve até o mais intenso.

Neste trabalho o objetivo é comparar (utilizando MLGMs e EEGs) a proporção do comportamento “olhar o flanco” entre os diferentes tratamentos e momentos e para isto utilizaremos o conceito de razão de chances. Assumimos o comportamento “olhar o flanco” como sendo nossa variável resposta tomando como valores: “1” se o cavalo olha o flanco ou “0” caso contrário (distribuição Bernoulli). As variáveis explicativas consideradas foram tratamentos e momentos. Avaliadores foram considerados como blocos e a variável idade (em anos) não foi incluída no estudo pois 25% dos equinos tinham idade desconhecida.

Budras, Sack e Röck (2009) define flanco como sendo a porção (próxima da pelve) entre a última costela e a coxa e o site PortalEquisport como sendo “a região limitada à frente pela última costela, atrás pela anca e pela coxa, em cima pelo rim e em baixo pelo ventre”. A Figura 1 exibe a silhueta dos equinos e destaca o flanco:

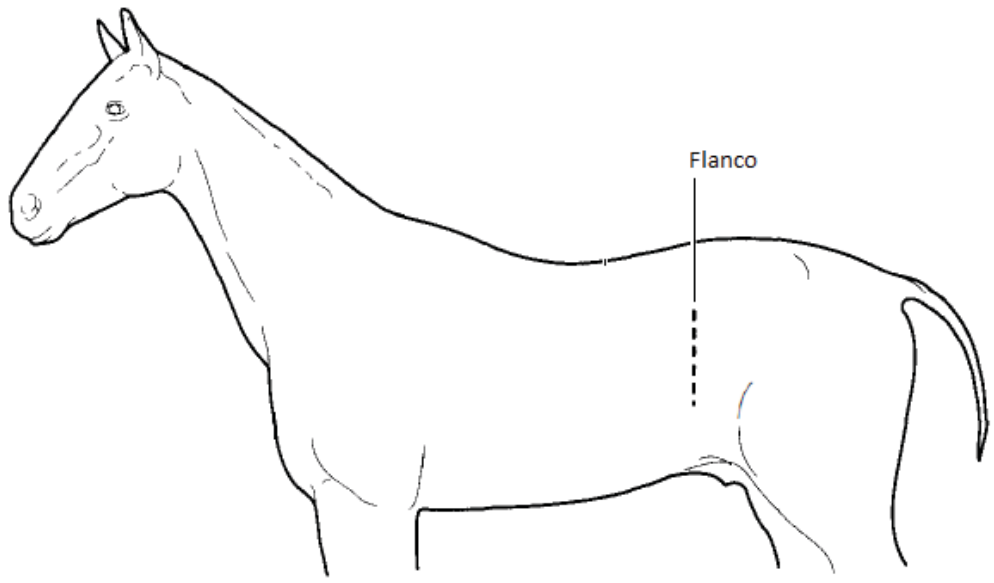

Figura 1: Silhueta dos equinos (adaptado de Mair et al (2002))

Dentre os softwares estatísticos disponíveis tanto comercialmente quanto livres, optou-se pela utilização do Statistical Analysis System (SAS) licenciado pela Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, Campus de Botucatu. Nele contém os procedimentos (PROC) utilizados neste trabalho (PROC GENMOD para as EEGs e o PROC GLIMMIX para os MLGMs) na modelagem estatística. Todos os programas estão em Anexo.

7) ANÁLISE E RESULTADOS

Para analisar os dados em estudo, foi decidido utilizar a covariável momento como categorizada (M1, M2, M3, M4) assim como foi codificada inicialmente por Taffarel (2013).

Para todos os modelos construídos a partir dos MLGMs (covariáveis como sendo de efeito fixo e o intercepto de cada indivíduo como efeito aleatório) foi utilizado o método de estimação de máxima verossimilhança. Para isso utilizou-se o método numérico da quadratura adaptativa de Gauss-Hermite (com 100 pontos de quadratura) e para ambas as metodologias (MLGMs e EEGs) foi utilizada a função de ligação $logit$ e foi construído modelos com as covariáveis tratamento, momento e a interação tratamento $\times$ momento (por se acreditar que haveria tal interação do ponto de vista biológico).

Avaliadores foram considerados como efeito de bloco nos modelos e por isso não foram analisadas as razões de chances quando estes estavam envolvidos. A variável resposta, “olhar o flanco”, assumiu os valores “1” se o animal olhou o flanco e “0” caso contrário. “Analistas de dados normalmente querem que seus modelos sejam com base na probabilidade do evento (sucesso), que muitas vezes é codificado como 1” (STOKES, DAVIS E KOCK, 2000).

Se em alguma fase da análise determinado termo não apresentasse significância estatística, este era retirado e um novo modelo era construído até que todos os efeitos apresentassem significância estatística (p-valor < 0,05).

O conjunto de dados era composto por 480 observações, sendo que 360 correspondiam ao valor “0” (fracasso) para a variável resposta, 91 ao valor “1” (sucesso) e 29 observações faltantes (“missing values” em inglês).

Os parâmetros de referência na construção dos modelos foram: avaliador 5, tratamento 3 e momento 2.

7.1) Análise Descritiva

A Tabela 1 mostra a frequência absoluta da variável resposta quando esta apresentava o valor “1” para os diferentes tratamentos analisados nos diferentes momentos observados.

Tabela 1: Frequência absoluta da variável resposta (valores “1”) por tratamento e por momento

	Momentos				
Tratamentos	M1	M2	M3	M4	Total
1	6	4	5	3	18
2	5	2	1	3	11
3	6	26	9	3	44
4	1	12	4	1	18
Total	18	44	19	10	91

A Tabela 2 mostra a frequência absoluta da variável resposta quando esta apresentava o valor “0” para os diferentes tratamentos analisados nos diferentes momentos observados (os valores entre parênteses é a frequência de valores perdidos por tratamento e por momento).

Tabela 2: Frequência absoluta da variável resposta para valores “0” e para valores perdidos (entre parênteses) por tratamento e por momento

	Momentos				
Tratamentos	M1	M2	M3	M4	Total
1	24(0)	21(5)	20(5)	27(0)	92(10)
2	25(0)	28(0)	29(0)	22(5)	104(5)
3	23(1)	04(0)	21(0)	22(5)	70(6)
4	27(2)	17(1)	26(0)	24(5)	94(8)
Total	99(3)	70(6)	96(5)	95(15)	360(29)

Nota-se pela Tabela 1 que a menor frequência (com uma observação) encontra-se nos seguintes casos: no tratamento 4 e M1, tratamento 2 e M3 e tratamento 4 e M4. A maior frequência observada está no tratamento 3 e M2 (26 observações) correspondendo a 59,1% do total (44) das realizações da variável resposta para este tratamento.

Outra informação relevante é que 44 realizações da variável resposta ocorreram no tratamento 3 correspondendo a 48,35% do total (91) das observações classificadas como sucesso (valor “1”).

A Tabela 3 exibe a proporção da variável resposta por tratamento e por momento.

Tabela 3: Proporção da variável resposta (valores “1”) por tratamento e por momento

	Momentos			
Tratamentos	M1	M2	M3	M4
1	0,20	0,13	0,16	0,10
2	0,16	0,06	0,03	0,10
3	0,20	0,86	0,30	0,10
4	0,03	0,40	0,13	0,03

Note que as duas maiores proporções ocorrem nos tratamentos 3 e 4 do momento 2 (0,86 e 0,40 respectivamente).

A Figura 2 ilustra o comportamento da proporção de sucesso da variável resposta por tratamento ao longo dos momentos complementando as informações dadas na Tabela 2:

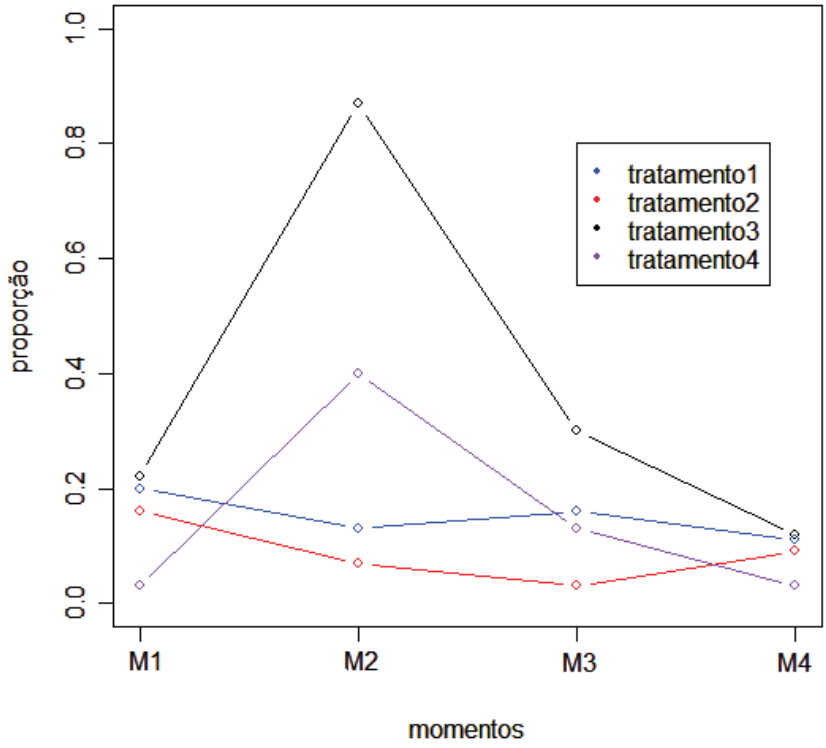

Figura 2: Gráfico da proporção da variável resposta (valores “1”) por tratamento e por momento.

A Figura 2 é o gráfico de uma série temporal discreta (os valores observados estão indicados através de pequenos círculos; o segmento de reta que os une somente tem a função de indicar o crescimento ou decrescimento da proporção de sucessos da variável resposta).

Nota-se pela Figura 2 que nos momentos M1 e M4, as proporções apresentam comportamentos semelhantes para os quatro tratamentos, enquanto que no M2 há diferenças entre as proporções (principalmente entre a proporção do tratamento 3 com as demais).

7.2) Modelagem Estatística

Nesta análise, conforme citado no início do capítulo 6, a covariável momento será considerada como sendo categorizada. A seguir, a análise será realizada inicialmente por MLGMs e, logo após, pelas EEGs.

MLGMs

Inicialmente verificou-se qual estrutura de matriz de covariância mais se adequava aos dados (para modelos com as covariáveis tratamento, momento e a interação tratamento × momento) utilizando o critério AIC. A Tabela 4 mostra os diferentes valores de AIC para as estruturas de matrizes de covariâncias descritas no capítulo 4 com exceção da estrutura SP(POW) que só é utilizada para o caso em que o momento é numérico).

Tabela 4: Valores de AIC para diferentes estruturas de matrizes de covariâncias utilizando MLGMs.

Estrutura de matriz de covariância	Valor de AIC
UN	332,11
CS	334,11
CSH	334,11
AR(1)	334,11
ARH(1)	334,11
TOEP	332,11
TOEPH	332,11

As matrizes que apresentam o menor valor de AIC (Tabela 4) foram: UN, TOEP e TOEPH. Foi escolhida a estrutura UN, pois os modelos construídos utilizando as estruturas UN, TOEP e TOEPH apresentaram os mesmos valores para todos os parâmetros ajustados (com seus respectivos erros-padrão e intervalos de confiança).

O modelo (denotado por modelo 1) construído com a estrutura de matriz de covariância UN dado pela Tabela 5, indica que as covariáveis tratamento, momento e a interação tratamento×momento foram significativas (p-valores menores que 0,05).

Tabela 5: Estatísticas de ajuste do modelo 1

Efeito	G.L* numerador	G.L* denominador	Estatística F	p-valor
Avaliador	4	411	-	-
Tratamento	3	411	3,31	0,0201
Momento	3	411	6,24	0,0004
Tratamento×Momento	9	411	3,62	0,0002

*G.L.: graus de liberdade

A Tabela 6 apresenta os valores das estimativas dos parâmetros de efeitos fixos com seus respectivos erros-padrão e p-valores do modelo 1.

Tabela 6: Estatísticas e estimativas dos parâmetros de efeitos fixos do modelo 1

Efeito	Estimativa	Erro-padrão	G.L.	Estatística t	p-valor
Intercepto	1,0473	0,8885	20	1,18	0,2523
Avaliador1	0,6381	0,6550	411	-	-
Avaliador2	2,1068	0,6170	411	-	-
Avaliador3	2,9385	0,6187	411	-	-
Avaliador4	1,7912	0,6209	411	-	-
Avaliador5	0	-	-	-	-
Tratamento1	-5,3118	1,2759	411	-4,16	<0,0001
Tratamento2	-6,3694	1,3878	411	-4,59	<0,0001
Tratamento4	-3,1675	1,1156	411	-2,84	0,0047
Tratamento3	0	-	-	-	-
Momento1	-4,4247	0,8736	411	-5,07	<0,0001
Momento3	-3,7256	0,8140	411	-4,58	<0,0001
Momento4	-4,7607	0,9491	411	-5,02	<0,0001
Momento2	0	-	-	-	-
Tratamento1×Momento1	5,0428	1,2494	411	4,04	<0,0001
Tratamento1×Momento3	4,1289	1,2191	411	3,39	0,0008
Tratamento1×Momento4	4,2024	1,3576	411	3,10	0,0021
Tratamento1×Momento2	0	-	-	-	-
Tratamento2×Momento1	5,8003	1,3542	411	4,28	<0,0001
Tratamento2×Momento3	2,8551	1,5846	411	1,80	0,0723
Tratamento2×Momento4	5,4205	1,4480	411	3,74	0,0002
Tratamento2×Momento2	0	-	-	-	-
Tratamento4×Momento1	0,6102	1,4517	411	0,42	0,6744
Tratamento4×Momento3	1,6707	1,1047	411	1,51	0,1312
Tratamento4×Momento4	1,1078	1,5213	411	0,73	0,4669
Tratamento4×Momento2	0	-	-	-	-
Tratamento3×Momento1	0	-	-	-	-
Tratamento3×Momento3	0	-	-	-	-
Tratamento3×Momento4	0	-	-	-	-
Tratamento3×Momento2	0	-	-	-	-

As estimativas dos parâmetros de efeitos aleatórios do modelo 1 estão expostas na Tabela 7.

Tabela 7: Estatísticas e estimativas dos parâmetros de efeitos aleatórios do modelo 1

Ind.	Estimativa	Erro-padrão	G.L.	Est. t	p-valor	L. I*	L. S*
1	-1,2345	1,0102	411	-1,22	0,2224	-3,2203	0,7514
2	0,7558	0,8197	411	0,92	0,3571	-0,8555	2,3670
3	-0,0051	0,8525	411	-0,01	0,9952	-1,6810	1,6708
4	-0,8558	1,0804	411	-0,79	0,4288	-2,9795	1,2680
5	2,5403	0,8292	411	3,06	0,0023	0,9102	4,1703
6	-0,5248	0,8967	411	-0,59	0,5587	-2,2875	1,2379
7	-0,7492	1,0939	411	-0,68	0,4938	-2,8995	1,4011
8	2,2983	0,8843	411	2,60	0,0097	0,5599	4,0366
9	-0,8958	1,0603	411	-0,84	0,3987	-2,9802	1,1886
10	0,5236	0,9095	411	0,58	0,5652	-1,2644	2,3115
11	-0,8958	1,0603	411	-0,84	0,3987	-2,9802	1,1886
12	0,5236	0,9095	411	0,58	0,5652	-1,2644	2,3115
13	-0,4561	0,7912	411	-0,58	0,5646	-2,0114	1,0992
14	-0,8851	0,8080	411	-1,10	0,2740	-2,4735	0,7033
15	-0,4985	0,7859	411	-0,63	0,5262	-2,0434	1,0464
16	0,5030	0,7605	411	0,66	0,5087	-0,9920	1,9980
17	-0,1418	0,7721	411	-0,18	0,8544	-1,6596	1,3760
18	1,6377	0,8324	411	1,97	0,0498	0,0015	3,2739
19	0,6310	0,8295	411	0,76	0,4473	-0,9996	2,2615
20	0,2050	0,8401	411	0,24	0,8074	-1,4465	1,8564
21	-1,5187	1,0286	411	-1,48	0,1406	-3,5406	0,5033
22	1,7097	0,8362	411	2,04	0,0415	0,0658	3,3535
23	0,2646	0,8507	411	0,31	0,7559	-1,4077	1,9369
24	-0,8510	0,9104	411	-0,93	0,3505	-2,6406	0,9387

*: L.I e L.S: Limite inferior e limite superior do intervalo de confiança (95%).

Para verificar se o modelo 1 é adequado aos dados, faz-se necessário verificar a suposição de normalidade dos efeitos aleatórios. A Tabela 8 exibe algumas estatísticas descritivas para a distribuição das estimativas dos efeitos aleatórios do modelo 1.

Tabela 8: Estatísticas descritivas das estimativas dos efeitos aleatórios do modelo 1.

Estatística descritiva	Valor
Mínimo	-1,5187
1º quartil	-0,8534
Mediana	-0,0734
Média	0,0866
3º quartil	0,5773
Máximo	2,5400
Variância	1,2187

Observando a Tabela 8, nota-se que a média (0,0866) está próxima de zero e a variância (1,2187) próxima de 1. Além disso, a variância indica que há uma pequena variabilidade das estimativas dos efeitos aleatórios do modelo 1 em torno da média. Estas variam de -1,5187 (mínimo) até 2,5400 (máximo) com amplitude total de 2,7374.

A Figura 3 mostra o histograma e o box-plot da distribuição das estimativas dos parâmetros de efeitos aleatórios do modelo 1.

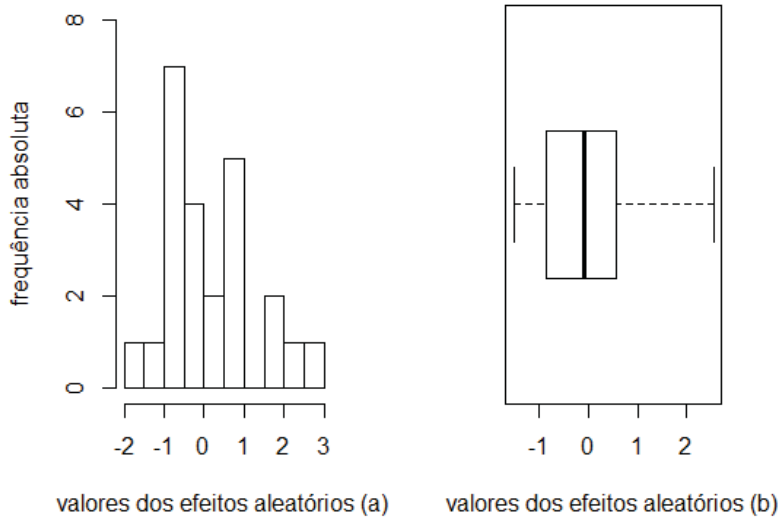

Figura 3: Histograma (a) e box-plot (b) das estimativas dos parâmetros de efeitos aleatórios do modelo 1.

Nota-se pela Figura 3 (a) e (b) que a distribuição das estimativas dos efeitos aleatórios do modelo 1 apresenta um ligeira assimetria à direita e (b) mostra que 50% delas estão em torno do valor zero.

A Tabela 9 exibe as estatísticas e p-valores de testes de normalidade para a distribuição das estimativas dos parâmetros de efeitos aleatórios do modelo 1. Nota-se que para os quatro testes realizados, seus respectivos p-valores são maiores que 0,05, indicando assim que podemos assumir distribuição Normal como estrutura probabilística à distribuição das estimativas dos parâmetros de efeitos aleatórios do modelo 1.

Tabela 9: Testes de normalidade para a distribuição das estimativas dos parâmetros de efeitos aleatórios do modelo 1

Teste de normalidade	Estatística do teste	p-valor
Shapiro-Wilk	0,9246	0,0741
Kolmogorov-Smirnov	0,1468	0,1500
Cramer-von Mises	0,0976	0,1165
Anderson Darling	0,6545	0,0807

A Figura 4 exibe os gráficos usados na análise de resíduos do modelo 1.

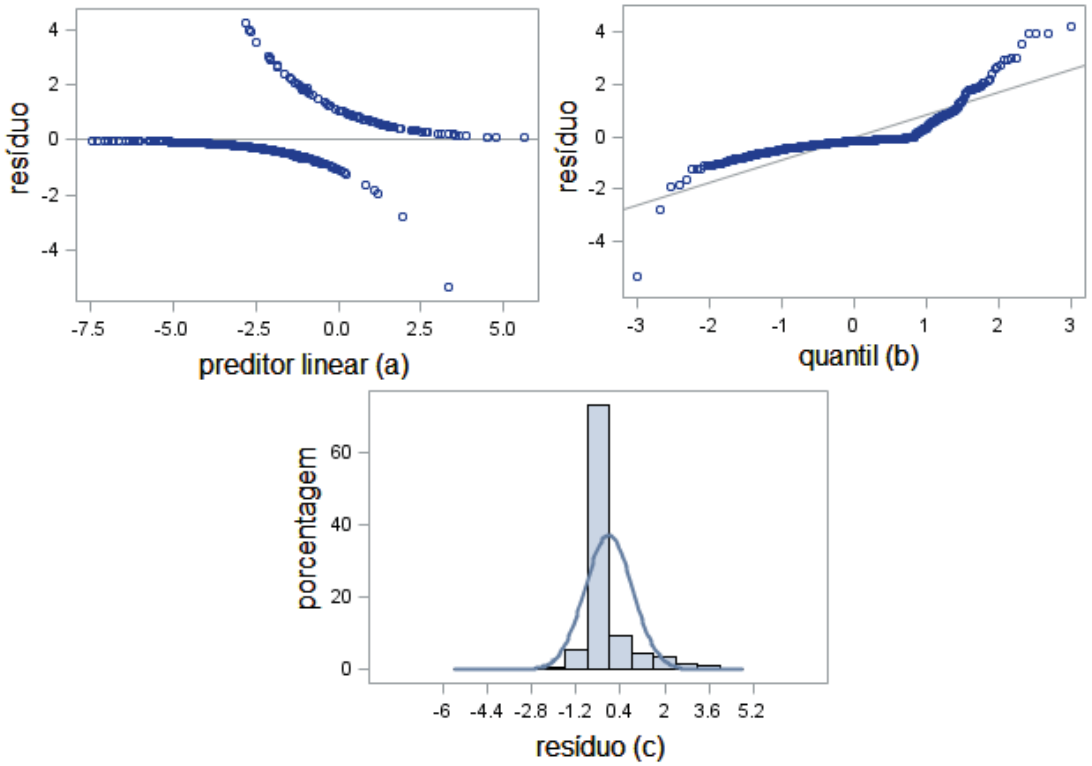

Figura 4: Gráficos dos resíduos condicionais estudentizados, (a) em função dos valores preditos, (b) quantil – quantil (c) histograma.

Assim o modelo 1 mostrou-se adequado para explicar a variável resposta “olhar o flanco”.

A Tabela 10 exibe as estimativas das razões de chances e seus intervalos de confiança (95%) para o modelo 1. Como vários termos da interação tratamento×momento não foram significativos, a interpretação das estimativas das razões de chances foi feita somente para os efeitos principais.

Tabela 10: Estimativas das razões de chances do modelo 1

Efeito	Efeito	Estimativa	G.L.	L.I*	L.S*
Tratamento1	Tratamento2	2,416	411	0,332	17,595
Tratamento1	Tratamento4	1,422	411	0,202	10,022
Tratamento3	Tratamento1	7,142	411	1,126	45,454#
Tratamento4	Tratamento2	1,697	411	0,218	13,157
Tratamento3	Tratamento2	17,24	411	2,403	125#
Tratamento3	Tratamento4	10,204	411	1,515	66,666#
Momento1	Momento3	1,001	411	0,361	2,773
Momento1	Momento4	1,676	411	0,571	4,925
Momento2	Momento1	4,761	411	1,785	12,65#
Momento3	Momento4	1,675	411	0,566	4,960
Momento2	Momento3	4,76	411	1,782	12,820#
Momento2	Momento4	8	411	2,777	22,727#

*: L.I e L.S: Limite inferior e limite superior do intervalo de confiança (95%).
#: Limite superior de intervalos de confiança (95%) que não contém o valor 1.

Conforme a Tabela 10, verifica-se que para os intervalos de confiança com o símbolo (#) na coluna do limite superior, estes não contém o valor 1 e por isso pode-se dizer que as estimativas das razões de chances são significativas ao nível de 5%.

Portanto pode-se afirmar que, segundo os MLGMs, os equinos do tratamento 3 têm em torno de sete vezes mais chance de “olhar o flanco” do que os do tratamento 1, 17 vezes mais chance do que os do tratamento 2 e 10 vezes mais chance do que os equinos do tratamento 4.

Quanto aos momentos, a chance de “olhar o flanco” ocorre em torno de quatro vezes mais no M2 quando comparado com os M1 e M3 e oito vezes mais do que no M4.

Estas interpretações das razões de chance nos MLGMs se dá somente para indivíduos que possuem o mesmo efeito aleatório (HOSMER, LEMESHOW e STURDIVANT, 2013; FITZMAURICE, LAIRD e WARE, 2011).

EEGs

Ao analisar os dados segundo as EEGs, primeiramente foi testado somente duas estruturas de matrizes de covariâncias: a CS e a AR(1)

(considerando modelos com as covariáveis tratamento, momento e a interação tratamento × momento). Para sua seleção, utilizou-se o critério QIC. A Tabela 11 exibe os valores de QIC para as estruturas CS e AR(1):

Tabela 11: Valores de QIC para as estruturas de matrizes de covariâncias CS e AR(1).

Estrutura de matriz de covariância	Valor de QIC
CS	388,2144
AR(1)	386,6207

Para os modelos construídos utilizando as matrizes CS e AR(1), as covariáveis tratamento, momento e a interação tratamento × momento não foram significativas ao nível de 0,05. Logo, foi retirado a interação tratamento × momento e calculado novamente os valores de QIC conforme Tabela 12.

Tabela 12: Valores de QIC para as estruturas de matrizes de covariâncias CS e AR(1) para o modelo sem a interação tratamento × momento.

Estrutura de matriz de covariância	Valor de QIC
CS	407,8876
AR(1)	403,7100

Decidiu-se utilizar na construção dos modelos a estrutura AR(1) pois apresentou o menor valor de QIC (conforme Tabela 12). Logo, o modelo (denotado por modelo 2) construído a partir da estrutura de covariância AR(1) é dado pela Tabela 13:

Tabela 13: Estatísticas de ajuste do modelo 2

Efeito	G.L	Estatística qui-quadrado	p-valor
Avaliador	4	-	-
Tratamento	3	7,46	0,0586
Momento	3	9,28	0,0258

A covariável tratamento não apresentou significância estatística ao nível de 0,05. Mas calculado o valor de QIC para o modelo com estrutura AR(1) sem a covariável tratamento (somente com a covariáveis momento e avaliador como efeito de bloco), este foi de 420,902 (maior do que QIC para a estrutura AR(1) da Tabela 12 que foi de 403,71).

Assim adotou-se como modelo final o exposto na Tabela 13 (modelo 2) visto que a covariável tratamento é de grande importância no experimento e, além disso, seu p-valor está muito próximo de 0,05.

As estimativas dos parâmetros do modelo 2 estão exibidas na Tabela 14 e a estimativa das razões de chances pela Tabela 15. A Figura 5 mostra gráficos de diagnósticos referentes ao modelo 2.

Tabela 14: Estatísticas e estimativas dos parâmetros do modelo 2

Efeito	Estimativa	Erro-padrão	L.I.	L.S.	Estatística Z	p-valor
Intercepto	-0,2870	0,5864	-1,4363	0,8623	-0,49	0,6245
Avaliador1	0,3740	0,3229	-0,2590	1,0069	1,16	-
Avaliador2	1,3286	0,3421	0,6581	1,9990	3,88	-
Avaliador3	1,8882	0,4088	1,0869	2,6895	4,62	-
Avaliador4	1,1483	0,3053	0,5499	1,7467	3,76	-
Avaliador5	0	0	0	0	-	-
Tratamento1	-1,4200	0,7262	-2,8432	0,0033	-1,96	0,0505
Tratamento2	-2,2656	0,7631	-3,7613	-0,7699	-2,97	0,0030
Tratamento4	-1,4083	0,4712	-2,3319	-0,4847	-2,99	0,0028
Tratamento3	0	0	0	0	-	-
Momento1	-1,4978	0,6652	-2,8015	-0,1940	-2,25	0,0243
Momento3	-1,4012	0,5296	-2,4391	-0,3633	-2,65	0,0081
Momento4	-2,0282	0,5946	-3,1936	-0,8627	-3,41	0,0006
Momento2	0	0	0	0	-	-

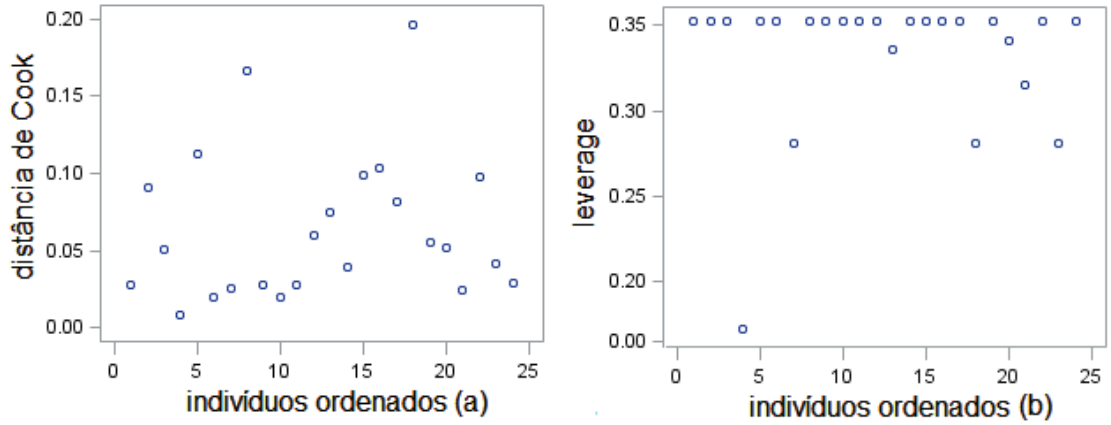

Figura 5: Gráficos de diagnósticos, (a) distância de Cook, (b) leverage.

Tabela 15: Estimativas das razões de chances do modelo 2

Efeito	Efeito	Estimativa	L.I*	L.S*
Tratamento2	Tratamento1	0,4238	0,0656	2,7385
Tratamento3	Tratamento1	4,0189	0,9689	16,6692
Tratamento4	Tratamento1	1,0043	0,2156	4,6772
Tratamento3	Tratamento2	9,4820	2,1316	42,1797#
Tratamento4	Tratamento2	2,3698	0,5034	11,1560
Tratamento3	Tratamento4	4,0012	1,5876	10,0838#
Momento2	Momento1	4,5349	1,2269	16,7620#
Momento3	Momento1	1,1364	0,3916	3,2981
Momento4	Momento1	0,6015	0,3217	1,1243
Momento2	Momento3	3,9904	1,4113	11,2829#
Momento2	Momento4	7,5398	2,3536	24,1537#
Momento4	Momento3	0,5292	0,1895	1,4777

*: L.I e L.S: Limite inferior e limite superior do intervalo de confiança (95%).
#: Limite superior de intervalos de confiança (95%) que não contém o valor 1.

Pela Tabela 15 pode-se dizer que equinos do tratamento 3 teve em torno de 9 vezes mais chance de “olhar o flanco” quando comparados com equinos do tratamento 2 e 4 vezes mais chance do que os equinos do tratamento 4.

Em relação aos momentos, a chance dos equinos “olhar o flanco” ocorreu cerca de 4,5 vezes mais no M2 do que no M1, 4 vezes mais do que no M3 e 7,5 vezes mais do que no M4.

A interpretação das razões de chances acima se dá comparando as proporções de sucesso da variável reposta na população nos diferentes níveis da covariável de interesse enquanto todas as outras covariáveis são consideradas fixas (HOSMER, LEMESHOW e STURDIVANT, 2013; FITZMAURICE, LAIRD e WARE, 2011).

Conforme Taffarel (2013) os animais dos grupos (1) e (2) fizeram parte do controle e os de (3) e (4) do tratado. Assim foi construído contrastes com o objetivo de verificar se houve diferença (quanto a “olhar o flanco”) entre controle e tratado, diferença entre grupos que fizeram parte do controle e diferença entre grupos que fizeram parte do tratado. A Tabela 16 exibe as estatísticas e p-valores dos contrastes mencionados acima.

Tabela 16: Estatísticas e p-valores dos contrastes

Contraste	G.L.	Est. qui-quadrado	p-valor
Diferença entre controle e tratado	1	4,20	0,0405
Diferença entre grupos que fizeram parte do controle	1	0,63	0,4290
Diferença entre grupos que fizeram parte do tratado	1	5,89	0,0152

Pela Tabela 16, nota-se que houve diferença estatisticamente significativa entre controle e tratado (p-valor < 0,05) e entre grupos que fizeram parte do tratado (p-valor < 0,05), mas não entre grupos pertencentes ao controle (p-valor > 0,05).

8) CONCLUSÕES

A Tabela 1, Tabela 2 e a Figura 2 deixaram evidente que as maiores ocorrências da variável “olhar o flanco” ocorreram justamente nos tratamentos 3 e 4, aos quais os equinos foram submetidos ao processo cirúrgico (grupo tratado).

Os modelos 1 e 2 (construídos com base nos MLGMs e EEGs respectivamente) resultaram em razões de chances que mostraram que a ocorrência da variável “olhar o flanco” era maior no tratamento 3 do que nos demais. Isto pode ser explicado pelo fato de que o tratamento 3 foi considerado o mais invasivo visto que os animais envolvidos neste tratamento foram anestesiados, operados e no pós-operatório receberam medicação (analgésicos) quatro horas após a recuperação anestésica.

Sobre os momentos, M2 foi o que apresentou as maiores chances de ocorrência da variável “olhar o flanco” nos modelos 1 e 2. Uma possível explicação para o ocorrido é que o M2 é o momento mais próximo (quatro horas após a recuperação anestésica) depois do processo cirúrgico e, consequentemente, o mais doloroso.

O contraste (Tabela 16) entre animais do controle e tratado mostrou que há diferença estatisticamente significativa (p-valor = 0,0405) entre eles quanto a “olhar o flanco”.

Assim, a reação do cavalo de “olhar o flanco” pode ser usada pelos médicos veterinários como indicativo de dor também para operações de orquiectomia eletiva (castração).

Ambos os modelos estatísticos ajustados aos dados (modelos 1 e 2) obtidos através das EEGs e dos MLGMs mostraram-se ser uma poderosa ferramenta para explicar a variável resposta “olhar o flanco” principalmente em relação a seleção das covariáveis estatisticamente significativas.

Nas EEGs, como não se exige a especificação da distribuição dos vetores-resposta, esta tornou-se uma metodologia estatística bastante

vantajosa, principalmente para dados binários como os analisados neste trabalho.

Os MLGMs, por terem como uma de suas suposição a heterogeneidade de indivíduo para indivíduo conforme a Tabela 7 indicou, estes se mostraram adequados quando o objetivo de inferência é sobre um indivíduo específico.

Um esforço considerado, neste trabalho, foi destinado à escolha da estrutura da matriz de covariância que se acomodasse bem a natureza dos dados. Apesar de as EEGs terem a propriedade de produzirem estimadores consistentes mesmo quando a estrutura da matriz de covariância for mal especificada, uma especificação muito grave dessa matriz pode afetar a eficiência dos estimadores.

Uma limitação do estudo conforme relata Taffarel (2013) foi “a inclusão de machos castrados e fêmeas nos grupos sem dor e o uso de machos inteiros nos grupos com dor. [...] poderia se esperar [...] diferenças comportamentais características de equinos castrados e fêmeas quando comparados a animais não castrados, comumente mais agitados.” Desta forma não houve uma homogeneização dos grupos o que pode ter afetado as estimativas dos parâmetros dos modelos, consequentemente afetando também as estimativas das razões de chances e suas interpretações.

Sobre os modelos estatísticos utilizados no conjunto de dados, houve uma superioridade das EEGs (em relação aos resultados obtidos) pois nos MLGMs as razões de chances só fazem sentido para animais que possuem o mesmo efeito aleatório o que ocorreu somente entre os animais 9 e 11 e entre os animais 10 e 12 (conforme Tabela 7) e estes fazem parte do grupo que receberam o tratamento 2.

Outro fato sobre o conjunto de dados é que este apresentou 75% das observações como “fracasso”, ou seja, excessivas observações com valor “0”. Esta característica afeta as estimativas dos erros-padrão dos parâmetros de regressão (consequentemente os intervalos de confiança das razões de chances conforme Tabela 10 e Tabela 15 pois apresentaram intervalos de confiança muito longos prejudicando assim a precisão do estudo) e estudos

futuros sobre modelos inflacionados de zeros com resposta binária correlacionada aplicados ao conjunto de dados utilizado nesta dissertação são necessários.

REFERÊNCIAS BIBLIOGRÁFICAS

ALLISON, P.D. **Logistic regression using the SAS System: Theory and application.** Cary, NC: SAS Institute, 1999.

BRESLOW, N.; CLAYTON, D. **Approximate inference in generalized linear mixed models.** Journal of the American Statistical Association, v. 88, 1993.

BUDRAS, K.D.; SACK, W.O.; RÖCK, S. **Anatomy of the horse.** 5 ed. Hannover: Schlütersche, 2009.

COLLETT, D.; **Modelling binary data.** 2 ed. Florida: Chapman e Hall, 2003.

DAVIS, C. S, **Statistical methods for the analysis of repeated measurements.** New York: Springer, 2002.

DE BOECK, P.; AZEVEDO, C.L.N.; TAVARES, H.R. **Linear and nonlinear generalized mixed models: inference and applications**. Fortaleza: XII Escola de Modelos de Regressão, 2010.

DIGGLE, P.J.; HEAGERTY, P.; LIANG, K.Y.; ZEGER, S.L. **Analysis of longitudinal data**. 2 ed. Oxford: Oxford Science, 2002.

DIGGLE, P. J.; LIANG, K. Y.; ZEGER, S. L. **Analysis of longitudinal data**. Oxford: Clarendon Press, 1994.

FILHO, J. A. C. **Modelos lineares mistos: estruturas de matrizes de variâncias e covariâncias e seleção de modelos**. 2002. Tese de Doutorado em Agronomia: Área de Concentração: Estatística e Experimentação Agronômica, Esalq, USP, Piracicaba, 2002.

FITZMAURICE, G.M.; LAIRD, N.M.; WARE, J.H. **Applied longitudinal analysis**. 2 ed. Massachusetts: Wiley, 2011.

GBUR, E. E.; STROUP, W.W.; McCARTER, K.S.; DURHAM, S.; YOUNG. L.J.; CHRISTMAN. M.; WEST. M.; KRAMER, M. **Analysis of generalized linear mixed models in the agricultural and natural resources sciences.** Madison: American Society of Agronomy, 2012.

GORDIS, L. **Epidemiologia**. 4 ed. Rio de Janeiro: Revinter, 2010.

HOSMER, D.W.; LEMESHOW, S.; STURDIVANT. R.X. **Applied logistic regression**. 3 ed. Hoboken: Wiley, 2013.

LIANG, K. Y.; ZEGER, S. **Longitudinal data analysis using generalized linear models**. Biometrika. V. 73, 1986.

LITTELL, R.; MILLIKEN, G.; STROUP, W.; WOLFINGER, R.; SCHABENBERGER. O. **SAS for mixed models.** New York, 6 ed. Cary, N.C.: SAS Institute, 2006.

MAIR, T.; DIVERS, T.; DUCHARME, N. **Manual of equine gastroenterology.** London: W B Saunders, 2002.

McCULLAGH, C.E.; SEARLE, S. R.; NEUHAUS, J.M. **Generalized, linear, and mixed models**. 2 ed. Hoboken: Wiley, 2008.

McCULLAGH, P.; NELDER. J.A. **Generalized linear models**. 2 ed. Chicago: Chapman e Hall, 1989.

MOREL, J. G.; NEERCHAL, N. K. **Overdispersion Models in SAS.** New York, 1 ed. Cary, N.C.: SAS Institute, 2012.

MYERS, R. H.; MONTGOMERY. D.C.; VINING. G.G.; ROBINSON. T.J. **Generalized linear models with applications in engineering and the sciences**. 2 ed. Hoboken: Wiley, 2010.

NELDER, J. A.; WEDDERBURN, R. W. M. **Generalized linear models**. Journal of the Royal Statistical Society, Series A, London, 1972.

PORTAL EQUISPORT. 2014. **Apresenta informações sobre algumas estruturas anatômicas dos equinos**. Disponível em: <http://www.equisport.pt/pt/artigos/a-volta-do-cavalo-lusitano-parte-iii/a-volta-do-cavalo-lusitano-iii-contd/a-volta-do-cavalo-lusitano-part-iii-contd>. Acesso em: 09 set. 2014.

PRENTICE, R.; ZHAO, L. **Estimating equations for parameters in means and covariance of multivariate discrete and continuous response**. Biometrics 47: 825-839. 1991.

SAS INSTITUTE. SAS/STAT 9.2 **User's guide. The GLIMMIX procedure (Book excerpt).** Cary, 2008. Disponível em: <http://www.support.sas.com/documentation> Acesso em: 19 jun. 2013.

STOKES, M. E.; DAVIS, C. S.; KOCH, G. G. **Categorical data analysis using the SAS system.** 6 ed. Cary, N.C.: SAS Institute, 2000.

TAFFAREL, M. O. **Proposição de escala clínica para a avaliação da dor em equinos.** Tese (doutorado), Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Medicina Veterinária e Zootecnia de Botucatu. 106 f. 2013.

XAVIER, L. H. **Modelos univariados e multivariados para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação**. Dissertação (mestrado), Universidade de São Paulo, Escola superior de agricultura "Luiz de Queiroz". 91 f. 2000.

ZEGER, S. L.; KARIM, M. R. **Generalized linear model with random effects: gibbs sampling approach.** Journal of the American Statistical Association – Theory and Methods, v. 86, 1991.

ZUUR, A.F.; IENO, E.N.; WALKER, N.J.; SAVELIEV, A.A.; SMITH, G.M. **Mixed effects models and extensions in ecology with R**. New York: Springer, 2009.

ANEXOS

```
*-------------------------------------Figura 2 ------------------------------------------------*

momento <- c(0,6,8,25)
trat1 <- c(0.20,0.16,0.20,0.10)
trat2 <- c(0.17,0.07,0.03,0.12)
trat3 <- c(0.21,0.87,0.30,0.13)
trat4 <- c(0.04,0.41,0.13,0.04)
plot(momento,trat1,type="b",main=" ",xlab="momento
(horas)",ylab="proporção",col="blue",ylim=c(0,1),lty=1)
lines(momento, trat2,col="red", type="b",lty=1)
lines(momento, trat3,col="black", type="b")
lines(momento, trat4,col="purple", type="b")
legend(17,0.8,c("tratamento1","tratamento2","tratamento3","tratamento4"),col
=c("blue","red","black","purple"), pch=rep(20,2))


/*------------------------ Modelagem estatística ---------------------------------------*/


proc import out=work.a /* momento categorizado*/
          datafile="C:\Users\maicon.galdino\Desktop\anestesia1.xlsx"
          dbms=excel replace;
     sheet="Plan4";
     getnames=yes;
     mixed=yes;
     usedate=yes;
     scantime=yes;
run;


proc glimmix data=a method=quad(qpoints=100) plots=(residualpanel
pearsonpanel studentpanel); /* modelo 1 escolhido */
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= UN cl;
output out=new pred(ilink)= predi stderr(ilink)= sepredi pred=pred stderr=sepred
resid=resid student=student;
run;


proc glimmix data=a method=quad(qpoints=100);
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= CS cl;
run;
```

```
proc glimmix data=a method=quad(qpoints=100);
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= CSH cl;
run;

proc glimmix data=a method=quad(qpoints=100);
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= AR(1)cl;
run;

proc glimmix data=a method=quad(qpoints=100);
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= ARH(1) cl;
run;

proc glimmix data=a method=quad(qpoints=100);
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= TOEP cl;
run;

proc glimmix data=a method=quad(qpoints=100);
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / solution
dist=bin link=logit s oddsratio(diff=all);
random intercept /  subject=animal type= TOEPH cl;
run;

proc genmod data=a descending;
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / dist=bin
link=logit type3;
repeated subject=animal/ type=CS;
run;

proc genmod data=a descending;
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento tratamento*momento / dist=bin
link=logit type3;
repeated subject=animal/ type=AR(1);
run;
```

```
proc genmod data=a descending;
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento / dist=bin link=logit type3;
repeated subject=animal/ type=CS;
run;

ods graphics on;
proc genmod data=a descending plot=all; /* modelo 2 escolhido */
class tratamento (REF="3") animal avaliador (REF="5") momento(REF="M2");
model flanco = avaliador tratamento momento / dist=bin link=logit type3;
repeated subject=animal/ type=AR(1);

contrast 'diferenca de grupo'
                tratamento 1 1 -1 -1;
estimate 'diferenca de grupo'
                tratamento 1 1 -1 -1;

contrast 'diferenca de controle'
                tratamento 1 -1 0 0;
estimate 'diferenca de controle'
                tratamento 1 -1 0 0;

contrast 'diferenca de tratado'
                tratamento 0 0 1 -1;
estimate 'diferenca de tratado'
                tratamento 0 0 1 -1;
run;
ods graphics off;
```