

MI 634 - Análise de dados longitudinais
Segundo semestre de 2018
Lista de Exercícios IV

1. Resolva TODOS os exercícios deixados em sala.
2. Considere o seguinte MLGM: $Y_{ij}|b_j \stackrel{ind.}{\sim} \text{gama}(\mu_{ij}, \phi)$ (veja a forma da densidade no livro Paula (2013)), em que $\mathcal{E}(Y_{ij}|b_j) = \mu_{ij}$, $\mathcal{V}(Y_{ij}|b_j) = \mu_{ij}^2/\phi$, $\ln \mu_{ij} = \beta_0 + \beta_1(x_{ij} - \bar{x}_i) + b_j$, $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$ e $b_j \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$. Encontre as expressões analíticas para $\mathcal{E}(Y_{ij})$, $\mathcal{V}(Y_{ij})$, $Cov(Y_{ij}, Y_{i'j})$ e $Corr(Y_{ij}, Y_{i'j})$.
3. (Paula (2013) No arquivo “ratosgee.dat” estão os dados de um experimento em que 30 ratos tiveram uma condição de leucemia induzida. Três drogas quimio-terápicas foram utilizadas no tratamento dos animais. Foram coletadas de cada animal a quantidade de células brancas (WBC), a quantidade de células vermelhas (RBV) e o número de colônias de células cancerosas (RESP) em quatro períodos diferentes. Faça uma análise descritiva apropriada dos dados. Assuma distribuição de Poisson para RESP em cada período e verifique através de um modelo log-linear misto se existe diferenças significativas entre os três tratamentos considerando WBC, RBC e o período como variáveis explicativas. Reduza, se possível, o modelo inicial. Faça uma análise de resíduos (do modelo inicial e do modelo final). A análise não pode exceder 15 páginas.
4. (O arquivo de dados está disponível no R no pacote “prLogistic” sob o nome “Toenail”). Os dados são de um estudo comparando dois tratamentos orais para a infecção do dedo do pé-unha, incluindo informações para 294 pacientes medidas em 7 semanas, compreendendo um total de 1908 medições. A variável binária resultado “onycholysis” indica o grau de separação da lâmina ungueal da unha-cama (nenhuma ou leve contra moderada ou grave). Os pacientes foram avaliados no início do estudo (semana 0) e nas semanas 4, 8, 12, 24, 36, 48 e daí em diante. As variáveis no arquivo de dados são: ID - identificação do paciente, Response - Grau de onycholysis (0 se nenhum ou leve, 1 se moderada ou grave), Treatment - Indicador oral tratamento (1 = Tratamento A, 0 = Tratamento B), Month - Tempo exato das medições, em meses, Visit - (números da visita, 1-7, correspondendo as visitas agendadas a 0, 4, 8, 12, 24, 36 e 48 semanas). Faça uma análise descritiva apropriada dos dados. Assuma distribuição de Bernoulli para a variável resposta em cada período e verifique através de um modelo logístico misto se existe diferenças significativas entre os dois tratamentos considerando o período (número da visita) e o tempo como variáveis explicativas. Reduza, se possível, o modelo inicial. Faça uma análise de resíduos (do modelo inicial e do modelo final). A análise não pode exceder 15 páginas.
5. (Conjunto de dados disponível no pacote R “bild” sob o nome de locust): O conjunto de dados foi apresentado e analisado por MacDonald e Raubenheimer (1995), e posterior-

mente examinado por outros autores, incluindo Gonçalves e Azzalini (2008). Os dados foram coletados para estudar o efeito da fome sobre o comportamento de locomoção de gafanhotos (*Locusta migratoria*). Especificamente, 24 gafanhotos foram observados em 161 instantes de tempo, em intervalos de trinta segundos; as unidades experimentais foram divididas em dois grupos de tratamento (feed e unfeed), e dentro de cada um dos dois grupos, os sujeitos eram alternativamente masculinos e femininos. Para o propósito desta análise, as categorias da variável de resposta estavam em movimento e não em movimento (veja também a descrição disponível no R). Analise o conjunto de dados, de forma apropriada, considerando todas as variáveis, tendo a variável “move”, como resposta.

6. Este conjunto de dados faz parte do conjunto de dados longitudinais coletados pelo Hospital Geral da cidade de St. Johns, no Canadá. Consideramos aqui os dados de contagem longitudinal que contêm os registros completos de 144 indivíduos por quatro anos ($n = 4$) de 1985 a 1988. O número de visitas a um médico por cada indivíduo durante um determinado ano foi registrado (resposta), durante esse período. Além disso, as informações sobre quatro covariáveis, a saber, gênero, número de condições crônicas, nível de escolaridade e idade, foram registrados para cada indivíduo. Análise de forma apropriada o conjunto de dados (veja também https://www.springer.com/cda/content/document/cda_downloaddocument/9781441983411-c1.pdf?SGWID=0-0-45-1107741-p174083597, dados: planilha health care utilization data.ods).
7. Considere pelo menos um conjunto de dados longitudinais, entre os vistos e não vistos em aula e proponha um modelo mais apropriado possível, analisando tal conjunto de forma apropriada.