

Universidade Federal do Paraná - Departamento de Estatística

CE075 - Análise de Dados Longitudinais

Prof. José Luiz Padilha da Silva

Lista 3

Obs: não precisa ser entregue.

Questão 1. Considere um estudo aleatorizado com o objetivo de comparar dois tratamentos orais (A e B) para micose de unha. Os pacientes foram aleatorizados com relação ao grau de onicólise (desprendimento da unha) no baseline (semana 0), e nas semanas 4, 8, 12, 24, 36 e 48. A variável resposta (grau de onicólise) é binária (nenhuma ou amena versus moderada ou severa) e foi avaliada em 294 pacientes compreendendo um total de 1908 medidas. O principal objetivo da análise é comparar os efeitos dos tratamentos A e B nas mudanças das probabilidades da resposta ao longo da duração do estudo. Os dados estão no arquivo `toenail.dta`.

- a) Considere um modelo marginal para as *log odds* de onicólise moderada ou severa. Usando o GEE, ajuste um modelo que assume tendências lineares para as *log odds* ao longo do tempo, com intercepto comum para os dois tratamentos mas com inclinações distintas:

$$\text{logit}\{E(Y_{ij})\} = \beta_1 + \beta_2 \text{Month}_{ij} + \beta_3 \text{Treatment}_i \times \text{Month}_{ij}.$$

- b) Qual é a interpretação de β_2 neste modelo?
- c) Qual é a interpretação de β_3 neste modelo?
- d) Quais conclusões você tira da análise? Forneça resultados que embasem as suas conclusões.
- e) Considere agora um modelo linear generalizado misto com interceptos aleatórios. Usando máxima verossimilhança, ajuste um modelo com tendências lineares para as *log odds* no tempo e com inclinações que dependem do grupo tratamento:

$$\text{logit}\{E(Y_{ij}|b_i)\} = (\beta_1 + b_i) + \beta_2 \text{Month}_{ij} + \beta_3 \text{Treatment}_i \times \text{Month}_{ij}.$$

em que, dado b_i , Y_{ij} segue uma distribuição de Bernoulli. Assuma que $b_i \sim N(0, \sigma_b^2)$.

- f) Qual é a estimativa de σ_b^2 ? Dê uma interpretação da magnitude da variância estimada.
- g) Qual é a interpretação da estimativa de β_2 ?
- h) Qual é a interpretação da estimativa de β_3 ?
- i) Compare e contraste as estimativas de β_3 dos modelos marginal e misto. Por que elas diferem?
- j) Repita a análise sequencialmente aumentando o número de pontos de quadratura usado no ajuste. Compare as estimativas e erros padrões dos parâmetros do modelo quando o número de pontos de quadratura é 2, 5, 10, 20, 30, e 50. Os resultados dependem do número de pontos de quadratura?

Questão 2. Retorne aos dados do arquivo `rats.xlsx` analisados com o estimador GLS na Lista 2. Os dados são referentes a observações de 50 ratos em terapia hormonal. Os ratos foram divididos em três grupos (controle, dose baixa e dose alta) e foram avaliados em 7 ocasiões (50, 60, 70, 80, 90, 100 e 110 dias). A resposta medida foi a distância crânio-facial no raio-x (em pixels) e os objetivos são: i) comparar os grupos, e ii) avaliar as mudanças temporais.

- a) Ajuste modelos marginais usando a função `geeglm()` usando as estruturas de correlação independente, simetria composta, AR(1) e não estruturada.
- b) Quais são as suposições distribucionais feitos no ajuste GEE?

- c) Discuta, para cada estrutura de correlação, as diferenças encontradas nas inferências em relação aos ajustes equivalentes obtidos pela função `gls()`.
- d) Interprete os resultados para o ajuste GEE que você considera mais adequado. Existem diferenças entre os grupos? A evolução temporal depende do grupo?

Questão 3. Utilize os mesmos dados da questão anterior.

- a) Ajuste modelos mistos por meio da função `lme()` considerando i) efeitos aleatórios na intercepto apenas, e ii) efeitos aleatórios no intercepto e tempo. Quais suposições distribucionais são feitas?
- b) Discuta como as correlações são determinadas pelos efeitos aleatórios. Quais estruturas de correlação são decorrentes de cada um dos dois modelos?
- c) Qual é o número total de parâmetros estimados nos modelos ajustados no item a)? Compare com o total de parâmetros estimados em cada estrutura de correlação no GEE. O modelo misto é mais parcimonioso? O que acontece com o número de parâmetros se dobrarmos a quantidade de medidas repetidas?
- d) Qual dos modelos mistos é preferível? Discuta por meio de testes e medidas como AIC/BIC.
- e) Apresente as previsões para os efeitos aleatórios para este modelo e discuta a magnitude da variabilidade estimada. Qual é a correlação marginal da resposta estimada pelo modelo?
- f) Interprete os resultados dos ajustes para os efeitos fixos. Existem diferenças entre os grupos? A evolução temporal depende do grupo?
- g) Realize o diagnóstico do modelo escolhido por meio de gráficos de resíduos.
- h) Por fim, você considera o modelo misto mais adequado que o modelo marginal? Quando você escolheria o GEE (e não o modelo misto) para análise?