

MI427/ME913 - Análise de dados Hierárquicos
Segundo semestre de 2020
Lista de Exercícios V

1. Resolva TODOS os exercícios deixados em sala.
2. Os dados disponíveis no site do curso sob o nome “IDH”, dizem respeito a diversas informações relativas a um tipo de pesquisa domiciliar realizado no Brasil, a chamada PNAD (Pesquisa Nacional por Amostra de Domicílios), realizados em vários anos. Nosso interesse reside no ano de 2013. Os dados compreende 5.563 municípios no Brasil (dois deles apresentam valores ausentes e exclua-os, para este exercício). A variável-resposta é um índice de pobreza definido como o percentual de pessoas com renda per capita igual ou inferior a R\$ 140,00 por mês (PMPOB). Como covariáveis considere o coeficiente de Gini (GINI) e o índice de desenvolvimento humano municipal (IDHM). No Brasil os municípios são agrupados em 27 estados. Ou seja, os estados são o nível 2, enquanto que os municípios são o nível 1. Faça análises, descritiva e inferencial (esta, sempre buscando o melhor modelo), completas e apropriadas, apresentando os comentários e conclusões, pertinentes. Veja também [aqui](#).
3. O conjunto de dados “Cefamandole”, disponível sob este nome no pacote nlme, de acordo com [Davidian e Giltinan](#) (1995, 1.1, p. 2) foi obtido durante um estudo piloto para investigar a farmacocinética da droga cefamandol. As concentrações plasmáticas do fármaco foram medidas em seis voluntários saudáveis em 14 instantes após uma dose intravenosa de 15 mg / kg, do peso corporal, de cefamandol. (Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London). Faça análises, descritiva e inferencial (esta, sempre buscando o melhor modelo), completas e apropriadas, apresentando os comentários e conclusões, pertinentes (considere todas as nuances do problema).
4. Analise, adequadamente, o conjunto de dados (disponível no R sob o nome de “Soybean”, no pacote nlme), apresentando nos [slides](#), descritiva e inferencialmente (esta, sempre buscando o melhor modelo), completa e apropriadamente, apresentando os comentários e conclusões pertinentes (considere todas as nuances do problema).
5. Para todos os modelos (M1 a M5), veja página 31 desses [slides](#), apresente todos os modelos possíveis, em relação a estrutura de regressão de segundo nível (efeitos aleatórios, sem covariáveis, como nos slides 32, 33, 34). Calcule esperanças, variâncias, covariâncias e correlações marginais. Considere sempre correlação nula entre os efeitos aleatórios.
6. Analise os conjuntos de dados vistos em sala e constantes nas listas anteriores, nesta e eventuais futuras, sob o ponto de vista bayesiano, usando o modelo que você considerar

ser o mais adequado (se a varificação da qualidade de ajuste do modelo indicar que ele não apresentou um bom ajuste, você pode considerá-lo, contanto que seja devidamente justificado).

7. Discuta/pesquise como os resultados assintóticos, dos estimadores, intervalos de confiança e estatísticas do teste, vistos nas aulas, podem ser demonstrados. Veja as referências:
 - (a) Demidenko, E. (2004). *Mixed models: theory and applications*, Hoboken, NJ: Wiley Interscience.
 - (b) Demidenko, E. (2013). *Mixed Models: Theory and Applications with R*, second edition, Hoboken, NJ: Wiley.
 - (c) Jiang, J. (1996). REML estimation: asymptotic behavior and related topics, *Ann. Statist.*, 25, 1781-1803.
 - (d) Jiang, J. (1998). Consistent estimators in generalized linear mixed models, *J. Amer. Statistic. Assoc.* 93, 720-729.
 - (e) Jian, J. (1998). Asymptotic proprieties of the empirical BLUP and BLUE in linear mixede models, *Statistics Sinica* 8, 861-885.
 - (f) Jiang, J. (2017). *Asymptotic Analysis of Mixed Effects Models: Theory, Applications, and Open Problems*, Chapman & Hall/CRC, UK.
8. Pesquisa a respeito da estimação (frequentista e beysiana) através de algortimos do tipo EM.
9. Pesquisa sobre a utilização dos pacotes DHARMA e HLMdiag (veja referências nos slides das aulas e no site do curso).