

COMPARAÇÃO DE MÉTODOS DE ESTIMAÇÃO EM MODELOS LOGÍSTICOS MULTINÍVEIS PARA DADOS LONGITUDINAIS: UM ESTUDO DE SIMULAÇÃO

Renata de Miranda Esquivel

Mestranda do Programa de Modelagem Computacional, SENAI CIMATEC,
Av. Orlando Gomes, 1845- Piatã – Salvador- BA
esquivel.renata@hotmail.com

Leila Denise Alves Ferreira Amorim

Departamento de Estatística da UFBA
Av. Adhemar de Barros, S/N- Ondina – Salvador- BA
leiladen@ufba.br

Rosemeire Leovigildo Fiaccone

Departamento de Estatística da UFBA
Av. Adhemar de Barros, S/N- Ondina – Salvador- BA
fiaccone@ufba.br

RESUMO

O desenvolvimento de métodos estatísticos para análise de dados obtidos em situações em que as observações são dependentes tem apresentado crescimento relevante nas últimas décadas e, em especial, na análise de dados provenientes de estudos longitudinais. No contexto dos modelos multiníveis lineares generalizados (GLMM), a estimação pode ser realizada sob dois enfoques: métodos de aproximação da verossimilhança ou por integração numérica. Esse trabalho objetiva comparar a performance dos métodos de estimação disponíveis no *software* R para o ajuste de modelos logísticos multiníveis na análise de dados longitudinais. Para isto foram conduzidos estudos de simulação, considerando-se diferentes graus de dependência, número de observações repetidas e número de indivíduos. Resultados dos estudos de simulação apontam para diferenças mais acentuadas de performance dos métodos (em termos de vício, variabilidade dos estimadores e probabilidade de cobertura) em situações com reduzido número de mensurações ao longo do tempo e alto coeficiente de correlação intraclasse. Neste caso, alguns métodos produziram estimativas menos viesadas, enquanto outros foram mais acurados. No entanto, o aumento conjunto do número de indivíduos sob estudo e do número de medidas repetidas apontam que a performance destes métodos de estimação tendem a ser muito similares. Desta forma, a escolha do método a ser utilizado no ajuste do modelo logístico multinível dependerá da disponibilidade em software estatístico.

PALAVRAS CHAVE. Modelos Logísticos Multiníveis, Dados longitudinais, Simulação.

Área principal (Estatística).

ABSTRACT

The development of statistical methods for analyzing data in situations where the observations are dependent has experienced significant growth in recent decades and, in particular, the analysis of data from longitudinal studies. In the context of generalized linear multilevel models (GLMM), the estimation can be performed from two perspectives: likelihood approximation methods or numerical integration. This paper aims to compare the performance of estimation methods available in software R for multilevel logistic models in the analysis of longitudinal data. To that end, we conducted simulation studies, considering different degrees of dependence, number of repeated observations and number of individuals. Results of simulation studies indicate differences more pronounced performance of the methods (in terms of bias, variability of the estimators and likelihood of coverage) in situations with a reduced number of measurements over time and high intraclass correlation coefficient. In this case, some methods produce less biased estimates, while others were more accurate. However, the combined increase in the number of individuals under study and the number of repeated measurements show that the performance of these estimation methods tends to be very similar. Thus, the choice of method to be used in adjusting the multilevel logistic model depends on the availability of statistical software.

KEYWORDS. Multilevel Logistic Models. Longitudinal Data. Simulation.

Main area (Statistics).