

Modelação da insuficiência cardíaca congestiva por modelos lineares generalizados com variável resposta binária

Ricardo Miguel Vieira de São João

Departamento de Informática e Métodos Quantitativos, Escola Superior de Gestão e Tecnologia de Santarém - ricardo.sjoao@esg.ipsantarem.pt

Resumo: Este trabalho trata a modelação da Insuficiência Cardíaca Congestiva (ICC) por uma classe restrita de Modelos Lineares Generalizados (MLGs): MLGs com variável resposta binária. Os MLGs foram aplicados a um conjunto real de dados com a finalidade de modelar a ICC. Na modelação da ICC foram usadas diversas funções de ligação, dentre as quais a Aranda-Ordaz. Dentre os modelos obtidos, o modelo logístico revelou-se o mais adequado. Para aferir da qualidade do ajustamento foram empregues o teste de Hosmer e Lemeshow, entre outras técnicas de diagnóstico. O poder discriminante do modelo foi avaliado com recurso à curva ROC (Receiver Operating Characteristic curve). Foram alvo de estudo discrepâncias nas probabilidades ajustadas pelo modelo. Foi avaliado o impacto do diagnóstico médico no modelo obtido.

Palavras-chave: Curva ROC, insuficiência cardíaca congestiva, modelos lineares generalizados, regressão logística, teste de Hosmer e Lemeshow.

Abstract: This work is concerned with modeling Congestive Heart Failure (CHF) using a restrict class of Generalized Linear Models (GLM's): GLM's with binary response variable. The GLM's were applied on a real data set with the main goal of modelling CHF. Different link functions were used, namely Aranda-Ordaz. The logistic model reveals better adjustment than other models. To check the goodness-of-fit, various diagnostic techniques including the Hosmer and Lemeshow test were used. The discriminant power of the model was evaluated using ROC (Receiver Operating Characteristic curve) curve analysis. Discrepancies in fitted probabilities by the model were analysed. The impact of medical diagnosis was assessed.

Keywords: ROC curve, congestive heart failure, generalized linear models, logistic regression, Hosmer and Lemeshow test.

1 Introdução

O coração é o motor que dá movimento ao sangue. Um pouco maior que o nosso punho fechado, bombeia em média seis mil litros de sangue por dia a um ritmo médio de setenta e oito batimentos por minuto. O seu mau funcionamento implica uma circulação deficitária que origina danos e distúrbios nas funções dos tecidos e órgãos do corpo humano. A Insuficiência Cardíaca Congestiva (ICC)