THE LOGISTIC REGRESSION MODEL AND THE QUASI-COMPLETE SEPARATION CONDITION

 $\underline{E.A.Colosimo}^{\dagger 1}$, D.P.Nacle²

¹Federal University of Minas Gerais, Brazil; ²Catholic University of Minas Gerais, Brazil

[†]E-mail: *enricoc@est.ufmg.br*

Logistic regression model is the standard method of analysis for binary responses in many fields, specially in medical and biostatistical applications. Maximum likelihood is generally the estimation method of choice. However, statisticians may face real situations, named quasi-complete separation condition, where the maximum likelihood estimators do not exist. In particular, this condition occurs frequently associated to the case of rare events but it can also happens in others cases. This work presents real data sets where such a situation occurs and propose some methods to solve it. Exact logistic regression might be an interesting alternative to maximum likelihood. Another standard solution consists of adding a small constant to the data set values in order to remove it from the quasi-complete separation condition. A new approach is proposed and comparisons among these methods are presented in this work.