



Statistics Seminar in the frame of ARC "Survival Analysis"

Jean-François DUPUY
IRMAR, INSA de Rennes,
Rennes, France

Logistic regression with a cure fraction

Abstract

Logistic regression is widely used in medical studies to investigate the relationship between a binary response variable Y and a set of potential predictors X . The binary response may represent, for example, the occurrence of some outcome of interest ($Y=1$ if the outcome occurred and $Y=0$ otherwise). In this talk, we consider the problem of estimating the logistic regression model with a cure fraction. A sample of observations is said to contain a cure fraction when a proportion of the study subjects (the so-called cured individuals, as opposed to the susceptibles) cannot experience the outcome of interest. One problem arising then is that it is usually unknown who are the cured and the susceptible subjects, unless the outcome of interest has been observed. In this setting, a logistic regression analysis of the relationship between X and Y among the susceptibles is no more straightforward.

We develop a maximum likelihood estimation procedure for this problem, based on the joint modeling of the binary response of interest and the cure status. We discuss and investigate the identifiability of the resulting model. The relationship between this problem and previous results obtained in the field of zero-inflated models is also discussed. Then, we consider the asymptotic properties of the proposed estimator, and we conduct a simulation study to investigate its finite-sample behavior. A real-data example illustrates the methodology.

Friday March 16, 2012 - 16:00 - at ISBA c-115