Binary Functional Linear Models under Choice-Based Sampling

M.S. Ahmed, M.K. Attouch, S. Dabo-Niang

PII: S2452-3062(17)30062-X
DOI: 10.1016/j.ecosta.2017.07.001
Reference: ECOSTA 69

To appear in: *Econometrics and Statistics*

Received date: 20 March 2016
Revised date: 7 July 2017
Accepted date: 7 July 2017

Please cite this article as: M.S. Ahmed, M.K. Attouch, S. Dabo-Niang, Binary Functional Linear Models under Choice-Based Sampling, *Econometrics and Statistics* (2017), doi: 10.1016/j.ecosta.2017.07.001

Binary Functional Linear Models under Choice-Based Sampling

M.S. Ahmed[a,*], M.K. Attouch[b], S. Dabo-Niang[c]

[a] *University of Lille, LEM-CNRS 9221, Villeneuve d'ascq, France.*
[b] *University of Djilalli Abbes, Algeria*
[c] *University of Lille, LEM-CNRS 9221, INRIA-MODAL, Villeneuve d'ascq, France*

Abstract

A functional binary choice model is explored in a case-control or choice-based sample design context. That is, a model is considered in which the response is binary, the explanatory variable is functional, and the sample is stratified with respect to the values of the response variable. A dimensional reduction of the space of the explanatory random function based on a Karhunen–Loève expansion is used to define a conditional maximum likelihood estimate of the model. Based on this formulation, several asymptotic properties are given. A simulation study and an application to kneading data are used to compare the proposed method with the ordinary maximum likelihood method, which ignores the nature of the sampling. The proposed model yields encouraging results. The potential of the functional choice-based sampling model for integrating special non-random features of the sample, which would have been difficult to see otherwise, is also outlined.

Keywords: Binary choice model, Functional data analysis, Choice-based sampling, Case-control.

1. Introduction

Choice models are characterized by the feature that the dependent variable is discrete instead of continuous. Examples include having a given disease or not, participation decisions, and transport choices made by individuals. In the context of choice models, the main idea in case-control or choice-based sampling design is to stratify the population with respect to the values of the categorical response. It often occurs that one or more outcomes occur infrequently in the population but are important for determining some key parameters of the model. By stratifying the population with respect to the responses, one can gather information on those infrequent outcomes at a much lower cost than would be incurred by simply increasing the size of a random sample. Equivalently, for any given sampling budget, one can increase the efficiency of predictions and parameter estimates using a suitably designed response-based sample. Such sampling designs have been independently investigated by econometricians who study choice behaviour and biostatisticians who are interested in rare diseases.

In biostatistics, case-control designs are useful for identifying the impact of several factors on the occurrence of a particular disease. The response is often binary (having the disease or not), but there may be more than two categorical responses. In a case-control study, separate samples of cases (diseased individuals) and controls (individuals without the disease) are selected, unlike in a prospective study design, in which a sample of individuals is chosen and followed through time until their responses are recorded. In the case of rare diseases, even large studies may produce only a few diseased individuals and little information about the hazard. In that case, the researcher might wish to oversample the rare disease of interest to increase the accuracy of his analysis. Therefore, compared with case-control studies, prospective studies are disadvantageous in terms of time and cost. For a general overview of medical case-control studies, see Keogh &

*Correspondence author
Email addresses: mohamed-salem.ahmed@univ-lille3.fr (M.S. Ahmed), attou_kadi@yahoo.fr (M.K. Attouch), sophie.dabo@univ-lille3.fr (S. Dabo-Niang)