Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 52 (2008) 2623–2631
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Nonparametric estimation of conditional ROC curves: Application to discrimination tasks in computerized detection of early breast cancer

Ignacio López-de-Ullibarri[a,*], Ricardo Cao[b], Carmen Cadarso-Suárez[c], María J. Lado[d]

[a] *Department of Mathematics, Escuela Universitaria Politécnica, Universidade da Coruña, Ferrol, A Coruña, Spain*
[b] *Department of Mathematics, Facultad de Informática, Universidade da Coruña, A Coruña, Spain*
[c] *Department of Statistics and Operations Research, Facultad de Medicina, Universidade de Santiago de Compostela, Santiago de Compostela, A Coruña, Spain*
[d] *Department of Computer Science, E.S.E.I., Universidade de Vigo, Ourense, Spain*

Received 25 January 2007; received in revised form 20 July 2007; accepted 13 September 2007
Available online 22 September 2007

Abstract

A local linear method for estimating the conditional ROC curve under the presence of continuous and categorical covariates is introduced. A data driven smoothing parameter selector based on the bootstrap is proposed. The methods are illustrated with real data from a discrimination problem emerging in the context of computer-aided diagnosis. The bootstrap approach is also used to construct pointwise confidence intervals for the area under the ROC curve.

Keywords: ROC curve; Local linear estimator; Smoothing; Bootstrap; Conditional distribution; Bandwidth selection; Microcalcifications; Computer-aided diagnosis

1. Introduction

In medical and imaging research the receiver operating characteristic (ROC) curve is a well-accepted technique for assessing the accuracy of a diagnostic test that yields continuous results (Swets and Pickett, 1982; Metz, 1986; Hanley, 1989). The ROC curve is a plot of the true-positive fraction (TPF, also known as 'sensitivity') versus the false-positive fraction (FPF, also known as '1-specificity') at various threshold values for defining a positive result of a disease or condition. This allows one to choose the threshold value depending on the costs of false-positive and false-negative decisions. Frequently, the best threshold is not known, and it may vary depending on the setting in which the test is implemented. A summary measure that aggregates performance information across a range of possible thresholds is desirable. The area under the ROC curve (AUC) is the most commonly used summary index of test accuracy. AUC values close to 1 indicate a high diagnostic accuracy.

In many instances, however, the effectiveness of a diagnostic test, the ROC curve itself, and the summary index AUC may be strongly influenced by a set of continuous and/or categorical covariates. This is the case of our motivating

* Corresponding author. Tel.: +34 981 337 400; fax: +34 981 337 401.
E-mail address: ilu@udc.es (I. López-de-Ullibarri).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2007.09.013

example, which illustrates an automated system dedicated to the early detection of breast cancer in digital mammograms, and whose results may significantly vary if covariates are introduced, as it will be shown in Section 5. The main goal of this paper is to characterize the impact of the covariates on the diagnostic test accuracy by calculating covariate-specific ROC curves.

To assess possible covariate effects on the accuracy of a diagnostic test, we propose a procedure which consists of a nonparametric flexible modelling of the ROC curve through a smooth estimate of the conditional distribution of the marker for diseased and non-diseased subjects. In the absence of covariates, a variety of smoothing methods for estimating the ROC curves were proposed, including nonparametric techniques based on kernel ideas (Zou et al., 1997; Lloyd, 1998; Qiu and Le, 2001; Zhou and Harezlak, 2002; Hall and Hyndman, 2003) and local linear kernel smoothing (Fan and Gijbels, 1996; Peng and Zhou, 2004). Recently, Claeskens et al. (2003) and Hall et al. (2004) developed methods for constructing confidence intervals.

In this paper we will extend the nonparametric estimator of Peng and Zhou (2004) to a conditional setup. Advantages of this estimator include: (1) it is invariant under a monotone transformation of test results; (2) it involves only one bandwidth and can be chosen in an optimal way; and (3) this estimator can also correct the boundary problem present in the estimator of Lloyd (1998).

In this work we discuss how the Peng and Zhou (2004) methodology can be adjusted to account for covariate effects. In Section 2, we describe in detail a motivating example that deals with a computer-aided diagnosis (CAD) system dedicated to the detection of clustered microcalcifications on digital mammograms. In Section 3 we derive a nonparametric estimator of the conditional ROC curve given a vector of covariates. A method of bandwidth selection for this estimator is proposed in Section 4. In Section 5 we apply this methodology to the real data and in Section 6 we provide some concluding remarks.

2. Example: automated detection of clustered microcalcifications on digital mammograms

Breast cancer is one of the main causes of death among women, but an early detection of the abnormalities can considerably reduce the mortality rates. Several investigators have pointed out the possibility of employing CAD systems, dedicated to the detection of lesions, to help radiologists in the interpretation of mammograms (Gurcan et al., 2002; Méndez et al., 1998). Usually, a CAD system produces, as a result, suspicious areas that can be recognized as true lesions or false detections. A fundamental aspect of any computerized scheme is the reduction of the false-positive rate, employing different techniques and several statistical methods, like ROC methodology (Chan et al., 1998; Yoshida et al., 2002). However, these statistical techniques may involve several restrictions in the type of covariates that can be analyzed, or sometimes they assume normal distributions for all the variables included in the analysis, and this is not always the real situation. Due to these limitations, an alternative technique which can correctly deal with all the situations and conditions should be presented. We believe that the conditional ROC methodology can provide benefits and improve the detection results obtained by a CAD scheme. To demonstrate the usefulness of employing the conditional ROC methodology, the technique was applied to a real system, dedicated to the automated detection of clustered microcalcifications on digital mammograms, which has been previously validated elsewhere (Lado et al., 1999, 2001).

2.1. Mammogram database selection

A total of 174 mammograms containing 77 clusters of microcalcifications, proved by biopsy, were selected from the mammographic screening program that is currently underway, from 1992, in the Community of Galicia (North West Spain), among women aged 50–64 years. All the images were digitized at a resolution of 2000×2500 pixels and 12 bits precision. The mammograms were classified by an experienced radiologist into fatty (57 mammograms) and dense (117 mammograms), according to their breast tissue.

2.2. Detection of clusters of microcalcifications

The algorithm implemented to detect the clusters of microcalcifications is a five-step process that involves (Lado et al., 1999, 2001): (1) detection of the breast border, employing a tracking algorithm that computes the gradient of grey levels in the breast (Méndez et al., 1996); (2) application of one-dimensional discrete wavelet transform over