

## A distance-based, misclassification rate adjusted classifier for multiclass, high-dimensional data

Makoto Aoshima · Kazuyoshi Yata

Received: 5 November 2012 / Revised: 20 August 2013 / Published online: 29 October 2013  
© The Institute of Statistical Mathematics, Tokyo 2013

**Abstract** In this paper, we consider a scale adjusted-type distance-based classifier for high-dimensional data. We first give such a classifier that can ensure high accuracy in misclassification rates for two-class classification. We show that the classifier is not only consistent but also asymptotically normal for high-dimensional data. We provide sample size determination so that misclassification rates are no more than a prespecified value. We propose a classification procedure called the *misclassification rate adjusted classifier*. We further develop the classifier to multiclass classification. We show that the classifier can still enjoy asymptotic properties and ensure high accuracy in misclassification rates for multiclass classification. Finally, we demonstrate the proposed classifier in actual data analyses by using a microarray data set.

**Keywords** Asymptotic normality · Distance-based classifier · HDLSS · Sample size determination · Two-stage procedure