HIROSHIMA MATH. J.
13 (1983), 203–214

A criterion for variable selection in multiple discriminant analysis

Yasunori FUJIKOSHI

(Received, September 17, 1982)

§1. Introduction

This paper deals with the problem of variable selection in discriminant analysis with $q+1$ populations and a multivariate linear model. The variable selection is important since there are situations where the deletion of some variables from the original variables may be preferable for the practical aim of statistical analysis. A number of stepwise procedures have been proposed for reducing the number of variables required to discriminate among the $q+1$ populations (e.g., see McCabe [7], Farmer and Freund [2]). McKay [8] has proposed a procedure for determining all subsets of variables that provide essentially as much separation among the $q+1$ populations as the original set of variables, based on a simultaneous test procedure in Gabriel's [6] sense.

In this paper we propose a criterion for determining the "best" subset of variables in the discriminant analysis whose aim is to interpret the differences among the $q+1$ populations in terms of only a few canonical discriminant variables. We obtain a creterion, based on a model fitting approach. We regard the problem of finding the "best" subset of variables as one of finding the "best" model, by introducing a family of parametric models. The parametric models are based on "no additional information hypotheses" due to Rao [10]. Our criterion is obtained by applying Akaike's information criterion (Akaike [1]) to choice of the models. The problem of finding the "best" subset of variables in a multivariate linear model is also discussed. This is a generalization of the problem of variable selection in the discriminant analysis. Asymptotic distributions of the criterion for variable selection in the multivariate linear model are obtained, resulting in generalizations of Fujikoshi [5] in the case of two-group discriminant analysis. The asymptotic distribution in the case when the original variables are ordered a priori can be reduced to a simple form.

§2. Multiple discriminant analysis

Consider $q+1$ p-variate normal population Π_α ($\alpha=1,\dots,q+1$) with means $\boldsymbol{\mu}_\alpha$ and the same covariance matrix Σ. Let $\boldsymbol{x}=(x_1,\dots,x_p)'$ be the column vector of the p variables. Assume that N_α samples from Π_α are available. We will