PERFORMANCE EVALUATION OF NORMAL DISTRIBUTION BASED CLASSIFICATION PROCEDURES

A.O. Adebanji[1], S. Adeyemi[2], K.S. Nokoe[3]

[1]*Department of Statistics, University of Agriculture, Abeokuta, Nigeria.*
[2]*Department of Mathematics,Obafemi Awolowo University, Ile-Ife, Nigeria*
[3]*Department of Mathematics and Computer Science, University of Science and Development, Navrongo, Ghana.*

Email: tinuadebanji@yahoo.com

Abstract: This study investigated the performance of the normal based homoscedastic Linear and heteroscedastic Quadratic classification functions (LDF) and (QDF) respectively under the non-optimal conditions of very small between group centriods separation using the Mahalanobis distance and unequal group representation (prior probabilities) in the population. A Monte Carlo study was undertaken to investigate the asymptotic behaviour of the Linear and Quadratic functions for observations from Multivariate Normal distributions. Two populations ($_i$,i=1,2) are considered with distributions $N_p(\mu_1, I)$ and $N_p(\mu_2, I)$ respectively, where μ_1=(0,0,...,0) and μ_2=(,0,...,0) are p-dimensional mean vectors, with p representing the number of variables in the simulated multivariate normal populations. The values of (square root of the Mahalanobis distance) range from 1 to 7 and p = 4, and 6. Sample size ratios (n_1:n_2) are varied as (1:1), (1:2), (1:3) and (1:4) for the $LDF_{(4)}$, $LDF_{(6)}$, $QDF_{(4)}$, $QDF_{(6)}$ models respectively. Results show that increasing sample size for the LDF models at =1 does not result in a reduction in misclassification rate of corresponding proportions. Also performance of QDF is better enhanced than LDF by increasing the number of discriminators than the LDF.