NS.SCI.37

A STATISTICAL TOOL TO CHARACTERIZE SUB-POPULATION DATA OF BIOLOGICAL SPECIES

I. G. C. G. Idamawatta[1], R. D. Yapa[1], S. D. S. S. Sooriyapathirana[2]

[1]Department of Statistics and Computer Science,
[2]Department of Molecular Biology and Biotechnology,
Faculty of Science, University of Peradeniya

Patterns of morphological variations of biological species are different for their sub populations. The morphological characteristics of a particular species may change due to environmental and genetic effects. There are several statistical techniques to summarize, analyze and visualize both numerical and categorical data. In statistics, there are several multivariate methods to characterize these sub-population data, such as Cluster Analysis, Principal Component Analysis, Factor Analysis, model ANOVA (Analysis of Variance) *etc*. By these statistical analyses, similar groups, correlated characteristics, difference between sub populations *etc* can be identified. Further by using pie charts, bar charts, box plots, scatter plots, *etc*. data can be visualized and made out to several patterns and distributions of each variable. In many biological studies data have to be analyzed statistically. Though several statistical software are available to analyze the data, most of them are not open source and are complex for biologists. Further, statistical analysis may be difficult for most of the biological researchers and they may not be familiar with handling statistical functioning and computations. Therefore, availability of software or tool that is convenient to understand and handle statistical techniques is a necessity. Further, accessibility of several help files and some assistances would be more convenient to the users. The main goal of this research was to introduce such a statistical tool that could be used to characterize geographically or spatially separated sub-population data of biological species.

In this study, MATLAB was used to implement the standalone statistical tool called Morphological Analysis, which is user friendly and has several user guidance and help files to understand the way of handling statistical methods. The graphical summaries can be viewed by two or three dimensional ways. The menu items and step by step guidance procedures help to obtain the desire output without typing any functions or programmes. Thus the user can perform statistical analyses without having much knowledge about statistics. Therefore, this is useful in data analyzing purposes for biology researchers.