Numerical taxonomy

 Numerical taxonomy or taximetrics, nowadays frequently and perhaps more appropriately referred to as phenetics, refers to the application of various mathematical procedures to numerically encoded character state data for organisms under study.

Thus, it is the analysis of various types of taxonomic data by mathematical or computerized methods and numerical evaluation of the similarities or affinities between taxonomic units, which are then arranged into taxa on the basis of their affinities.

So, Numerical taxonomy is a classification system in biological systematics which deals with the grouping by numerical methods of taxonomic units based on their character states. It aims to create a taxonomy using numeric algorithms like cluster analysis rather than using subjective evaluation of their properties.

The concept was first developed by Robert R. Sokal and Peter H. A. Sneath in 1963 and later elaborated by the same authors. They divided the field into phenetics in which classifications are formed based on the patterns of overall similarities and cladistics in which classifications are based on the branching patterns of the estimated evolutionary history of the taxa.

In recent years many authors treat numerical taxonomy and phenetics as synonyms despite the distinctions made by those authors.

According to Heywood the numerical taxonomy may be defined as the numerical evaluation of the similarity between groups of organisms and the ordering of these groups into higher ranking taxa on the basis of these similarities.

 The period from 1957 to 1961 saw the development of first methods and of theory of numerical taxonomy. Plants as we all know are classified based on their characters. It was Michel Adanson, a French botanist, who for the first time put forward a plan for assigning numerical values to the similarity between organisms and proposed that equal weightage should be given to all the characters while classifying plants.

 He used as many characters as possible for the classification, and such classifications came to be known as Adansonian classifications. Numerical taxonomy was however largely developed and popularized by Sneath and Sokal.

The application of Adansonian principles and use of modern methods and electronic data processing techniques, have helped in the evolution of several new classifications of plants during the past few decades.

Principles of Numerical Taxonomy:

 Numerical taxonomy involves two aspects:

 **(a) Construction of Taxonomic Groups:**

 i. In numerical taxonomy, first, individuals are selected and their characters spotted out. There is no limitation to the number of characters to be considered. However, the larger the number of characters, better is the approach for generalization of the taxa.

ii. The resemblances among the individuals are then established on the basis of character analysis, which can often be worked out with the help of computers, the accuracy of which depends on the appropriateness in character. The best way to delimitate taxa is, to utilize maximum number of characters, with similar weightage given to all of them.

**(b) Discrimination of the Taxonomic Groups:**

 When the taxonomic groups chosen for the study show overlapping of characters, discrimination should be used to select them. Discrimination analysis can be done by various techniques, specially devised for such purposes. Numerical taxonomy is thus, based on certain principles, also called neo Adansonian principles.

**Following seven principles of numerical taxonomy have been enumerated by Sneath and Sokal:**

1. The greater the content of information in the taxa, and more the characters taken into consideration, the better a given classification system will be.
2. Every character should be given equal weightage in creating new taxa.
3. The overall similarity between any two entities is a function of the individual similarities in each of the many characters, which are considered for comparison
4. Correlation of characters differ in the groups of organisms under study. Thus distinct taxa can be recognized.
5. Phylogenetic conclusions can be drawn from the taxonomic structure of a group and from character correlations, assuming some evolutionary mechanisms and pathways.
6. The science of taxonomy is viewed and practiced as an empirical science.
7. Phenetic similarity is the base of classifications.

**Method and Procedure of Numerical Taxonomy**

* **The logical steps are:**
* (1) **Choice and Number of units to be Studied:** The basic unit of numerical taxonomy is operational taxonomic Unit (OTU). The term is given to the lowest term being studied in a particular investigation. The number of characters needed to achieve are at least 60 characters or more like 80 or 100 as desirable. In listing characters only homologous (common evolutionary origin) characters are to be compared.
* **(2) Character Selection and Taxon Matrix:** Character is a feature of an individual or taxonomic group, which can be measured, counted or assessed. In angiosperms mostly morphological, physiological or distributional ones are used as OTU, resulting in Data-matrix of attributes. (OTUs X characters or taxon).

 If one is classifying 30 OTUs and using 100 characters, the data matrix will consist of 3000 attributes. A computer is essential to calculate the data because of number and variety of operational units. The use of a computer necessitates the codification of the attributes in some simple form which can be feed into the computer, and this presents great problem. The simplest codification is a Binary or two state system, e.g., ± and – or 0 and 1 where each character exists in only two states

If one is classifying 30 OTUs and using 100 characters, the data matrix will consist of 3000 attributes. A computer is essential to calculate the data because of number and variety of operational units. The use of a computer necessitates the

codification of the attributes in some simple form which can be feed into the computer, and this presents great problem. The simplest codification is a Binary or two state system, e.g., ± and – or 0 and 1 where each character exists in only two states.



 **3. Merits of Numerical Taxonomy**

 According to Sokal and Sneath, numerical taxonomy has the following advantages over conventional taxonomy:

 a. The data of conventional taxonomy is improved by numerical taxonomy as it utilizes better and more number of described characters. The data are collected from a variety of sources, such as morphology, chemistry, physiology, etc.

b. As numerical methods are more sensitive in delimiting taxa, the data obtained can be efficiently used in the construction of better keys and classification systems, creation of maps, descriptions, catalogues, etc. with the help of electronic data processing systems. Numerical taxonomy has in fact suggested several fundamental changes in the conventional classification systems.

c. The number of existing biological concepts have been reinterpreted in the light of numerical taxonomy.

d. Numerical taxonomy allows more taxonomic work to be done by less highly skilled workers.

 **4. Demerits of Numerical Taxonomy**

 Numerical taxonomy can however prove to be disadvantageous from the following points of view:

 a. The numerical methods are useful in phenetic classifications and not phylogenetic classifications.

b. The proponents of “biological” species concept, may not accept the specific limits bound by these methods.

c. Character selection is the greatest disadvantage in this approach. If characters chosen for comparison are inadequate, the statistical methods may give less satisfactory solution.

 d. According to Steam, different taxonometric procedures may yield different results. A major difficulty is to choose a procedure for the purpose and the number of characters needed in order to obtain satisfactory results by these mechanical aids. It is necessary to ascertain whether a large number of characters would really give satisfactory results than those using a smaller number.

**5. Applications of Numerical Taxonomy:**

 Numerical taxonomy has been successfully applied in the following studies:

a. Study of similarities and differences in bacteria, other micro-organisms and several animal groups.

b. Delimitation of several angiospermic genera like Oryza, Sarcostemma Solarium, and other groups including Farinosae of Engler and a few others.

 c. In the study of several other angiospermic genera including Apocynum, Chenopodium, Crotalaria, Cucurbita, Oenothera, Salix, Zinnia, wheat cultivars, Maize cultivars, etc.

 d. Phytochemical data from seed protein and mitochondrial DNA RELP studies has been numerically analyzed by Mondal et al. to study the interspecitic variations among eight species of *Cassia* L. Based on the results of electrophoretic patterns, the degree of pairing affinity (PA) or similarity index was calculated by the following formula, according to the method of Sokal & sneth and Romero Lopes et al.:

 

Separate dendograms expressing the average linkage were computed using the cluster method UPGMA, which showed that the eight species could be placed into two categories or clusters (Fig. 9.6) with C. alata, C siamea, C. fistula and C. reginera, all being trees or large shrubs and characterized by the absence of foliar glands on petiole or rachis and presence of dense axillary terminal racemes greater than 30 cm long, being clustered into one group, whereas the other four species, i.e., C. occidentalis, C. sophera, C. mimosoides and C. tor a, forming the other cluster, all being herbs or undershrub’s and characterized by the presence of short corymbose racemes less than 10 cm long and with foliar glands, either on petiole or rachis.

*Dr Habibur Rahman, Associate Professor, J. N. College, Boko*