



UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
DEPARTAMENTO DE TECNOLOGIA DE ALIMENTOS

Campus Universitário – Viçosa, MG – 36570-000 – Telefone (31)3899-2226 – fax: (31) 3899-2208 - E-mail: dta@ufv.br

TAL 797 – Seminário

Data: 28/06/2017

Aluno: Victor de Souza Tavares

Orientador: Ronaldo Perez (Departamento de Tecnologia de Alimentos)

Análise estatística de agrupamento (*cluster*): Aspectos teóricos e práticos
Statistical cluster analysis: Theoretical and practical aspects

Multivariate Statistical Analysis consists of several statistical methods used to analyze and interpret large data sets, which are usually collected and measured through using many variables. One of these methodologies is the *cluster analysis*, whose objective is mainly to group objects based on their characteristics. In *cluster analysis*, the homogeneity within the groups must be maximized, as well as the heterogeneity between different groups. Basically, the steps of the methodology consisting in choosing a criterion of similarity, to form groups, defining the number of groups (it can be a at first or at end of the process), validating the groups obtained by appropriated statistical tests and finally interpreting the results obtained. This methodology can be applied using appropriate software such as IBM® SPSS® Statistics or STATA® Statistical Data Analysis for example. *Cluster analysis* is applied, for example, for the market analysis of food products in order to identify groups of consumers with similar needs, attitudes and perceptions such as for organics products; another possibility is to group food based on their physicochemical characteristics. It is noted that this is a methodology with applications in different areas and that is possible to obtain different solutions for the same problem, so it is concluded that the choice of the best solution depends on the research problem, on the proposed objectives and mainly on the knowledge of the studied subject.

Bibliographic references:

GAYIN, J., CHANDI, G. K., MANFUL, J., SEETHARAMAN, K. Classification of rice based on statistical analysis of pasting properties and apparent amylose content: The case of *Oryza glaberrima* accessions from Africa. **Cereal Chemistry**, v. 94, n. 2, p. 22-28, 2015.

HAIR JR., J. F. et al. **Análise Multivariada de Dados**. 5ª edição. Porto Alegre: Bookman, 2005. 593p.

JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 6rd ed. Englewood Cliffs, New Jersey: Prentice Hall, 1992. 642p

MINGOTI, S. A. **Análise de Dados Através de Métodos de Estatística Multivariada – Uma Abordagem Aplicada**. Belo Horizonte: Editora UFMG, 2005. 295p.

ZAKOWSKA-BIEMANS, S. Polish consumer food choices and beliefs about organic food. **British Food Journal**, v. 113, n. 1, p. 122-137, 2011.