

## **Course Title: Applied Statistics in Soil Science**

### **Identification**

**Code: SOL875**

**Credits: 3 (1-2)**

**Level: Master's and Doctorate**

**Professor: Ricardo Bergamo Schenato**

**Offered: Biannually**

**Course Objectives:** Deepen students' knowledge for the analysis of experimental data; enable interpretation of results; encourage discussion based on relationships between variables; enable the use of open-source software in statistical analyses.

### **Syllabus**

Significance tests; Experimental designs; Factorial experiments; Split-plot experiments; Regression; Principal Component Analysis; Factor Analysis; Canonical Correlation Analysis; Cluster Analysis; Discriminant Analysis

### **Methodology and Teaching Tools**

The course will consist of theoretical readings and discussions on analyses and problem-solving. Active participation in discussions will enhance theoretical understanding. Practical sessions will focus on problem-solving using the R statistical environment. Teaching tools include whiteboard, multimedia projector, and the R environment, which is freely available.

### **Evaluation**

Evaluation will be based on classroom participation (40% of the grade) and a final exam (60% of the grade). The final exam will be open-book and will require the use of a computer.

## **Course Outline**

### **PART 1**

#### **Unit 1 - Initial Conceptual Foundations**

- Review of basic statistics concepts
- Principles of experimentation
- Characteristics of an experiment

#### **Unit 2 - Significance Tests**

- Introduction
- Analysis of variance (ANOVA) and the F-test
- Tests for mean comparisons
- 

#### **Unit 3 - Experimental Designs**

- Completely Randomized Design (CRD)
- Randomized Complete Block Design (RCBD)

#### **Unit 4 - Factorial Experiments**

- Introduction
- Analysis and interpretation of factorial experiments

#### **Unit 5 - Split-Plot Experiments**

- Introduction
- Analysis of variance with significant interaction
- Analysis of variance with non-significant interaction

#### **Unit 6 - Group Analysis of Experiments**

- Introduction
- Analysis and interpretation procedures

#### **Unit 7 - Regression Analysis**

- Introduction
- Regression model estimation
- Parameter tests

### **PART 2**

#### **Unit 8 - Principal Component Analysis**

- Covariance matrix
- Determination of the number of components
- Variable selection
- Inferences

#### **Unit 9 - Factor Analysis**

- Factorial model
- Parameter estimation
- Factor rotation

- Inference and validation

#### **Unit 10 - Cluster Analysis**

- Measures of similarity and dissimilarity
- Hierarchical clustering
- Non-hierarchical clustering
- Determination of the number of clusters
- Multidimensional scaling

#### **Unit 11 - Canonical Correlation Analysis**

- Canonical variables and canonical correlations
- Analysis quality
- Inferences

#### **Unit 12 - Discriminant Analysis**

- Classification rules
- Identification of most important variables
- Discriminant functions

## Recommended Bibliography

- FERREIRA, D. F. **Estatística Básica**. Viçosa: UFLA, 2005, 664 p.
- BANZATTO, D. A., KRONKA, S. N. **Experimentação agrícola**. 3. Ed. Jaboticabal: FUNEP, 1995. 247 p.
- STORCK, L., GARCIA, D.C., LOPES, S. J., ESTEFANEL, V. **Experimentação vegetal**. Santa Maria: UFSM, 2000. 198 p.
- FILHO, A.C.; LÚCIO, A.C. **Análise de variância e testes complementares: teoria e aplicações**. Santa Maria: UFSM, 2008. 121 p.
- MOORE, D. S.; MCCABE, G. P. **Introdução à prática da estatística**. 3. Ed. Rio de Janeiro: LTC, 2002. 536 p.
- ZIMMERMANN, F. J. P. **Estatística Aplicada à pesquisa agrícola**. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2004. 402 p.
- OTT, R. LYMAN. **An introduction to statistical methods and data analysis**. 6th ed. Belmont, CA : Brooks/Cole Cengage Learning, 2010. 1273 p.
- MELLO, M.P.; PETERNELLI, L.A. **Conhecendo o R – uma visão mais que Estatística**. Viçosa, MG: Ed. UFV, 2013. 222p.
- HAIR, J.F.; BLACK, B.; BABIN, B.; ANDERSON, R.E.; TATHAM, R.L. **Análise Multivariada de Dados**. Porto Alegre: Bookman, 2009, 687 p.
- FERREIRA, D. F. **Estatística Multivariada**. 2. Ed. Viçosa: UFLA, 2011, 676 p.
- MINGOTI, S.A. **Análise de dados através de Métodos de Estatística Multivariada – uma abordagem aplicada**. Belo Horizonte: UFMG, 2013, 297 p.
- MANLY, B.J.F. **Métodos estatísticos Multivariados – uma introdução**. 3. Ed. Porto Alegre: Bookman, 229 p.
- WACKERNAGEL, H. **Multivariate Geostatistics**.
- MELLO, M.P.; PETERNELLI, L.A.; **Conhecendo o R – Uma visão mais que Estatística**. Viçosa: Ed. UFV, 2013, 222 p.
- R Core Team (2017)**. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL <http://www.r-project.org/>.
- HARLOW, L. L. (2014). **The Essence of Multivariate Thinking: Basic Themes and Methods** (2nd ed.). New York: Routledge.
- RENCHER, A. C. & CHRISTENSEN, W. F. (2012). *Methods of Multivariate Analysis* (3rd ed.). Hoboken, NJ: Wiley.
- TABACHNICK, B. G., & FIDELL, L. S. (2007). **Using multivariate statistics** (5th ed.). Boston, MA: Allyn & Bacon.
- SCHUMACKER, R. E. (2015). **Using R With Multivariate Statistics**. SAGE Publications.