Data Analysis

6th Semester

School of Applied Mathematical and Physical Sciences

Introduction to Statistics. **Introduction to the open-source R environment and language. Descriptive Statistics**: Qualitative and categorical variables, graphs in R. **Stochastic Simulation**: Distributions in R, Goodness of fit tests, Weak law of large numbers, Central limit Theorem**. Statistical Inference**: Maximum likelihood estimation, Interval estimation, Introduction to hypothesis testing, Confidence intervals and testing statistical hypotheses, for the following cases: (i) the mean of a quantitative variable, (ii) the proportion of "success" of a binary variable, (iii) the difference of two means of a quantitative variable using matched data, (iv) the difference of two means of a quantitative variable using two independent samples and (v) the difference of two proportions of "success" of a binary variable using two independent samples. **Regression Analysis**: Simple linear models, Correlation coefficient, multiple linear models. **Analysis of Variance**: One-way and two-way analysis of variance.

Bibliography:

Agresti, A. (2002). *Categorical Data Analysis*. Wiley. New York.

Agresti, A and Franklin, C. (2007). *Statistics. The Art and Science of Learning from Data*. Prentice Hall. New Jersey.

Bain, L.J & Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*. Second Edition. Duxbury Press. Belmont, California.

Bertsimas, D. & Freund, R.M. (2000). *Data, Models and Decisions: The Fundamentals of Management Science*. South-Western College Publishing. Ohio.

Bower, A.H. & Lieberman, G.J. (1972). *Engineering Statistics*. Prentice Hall. New York.

Chatfield, C. (1995). *Problem Solving. A statistician's guide*. Chapman and Hall. London.

Cox, D.R. & Snell, E.J. (1981). *Applied Statistics. Principles and Examples*. Chapman and Hall. London.

Crawley, M.J. (2007). *The R Book*. Wiley. New York.

Dalgaard, P. (2002). *Introductory Statistics with R*. Springer-Verlag, New York.

Freedman, D., Pisani, R. and Purves R. (2007). *Statistics*. 4th Edition. W.W. Norton & Co Ltd. New York.

Gibra, I.N. (1973). *Probability and Statistical Inference for Scientists and Engineers*. Prentice Hall. Englewood Cliffs. New Jersey.

Jaynes, E.T. and Bretthorst G.L. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press. UK.

Ross, S. (2006). *A First Course in Probability*. Seventh Edition. Prentice Hall. Englewood Cliffs. New Jersey.

Snedecor, G.W & Cochran, W.G. (1989). *Statistical Methods*. Eight Edition. Iowa State University Press, Ames, Iowa.

Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley. Reading, Mass.