DE LA SALLE UNIVERSITY - MANILA
COLLEGE OF SCIENCE
Mathematics Department

SYLLABUS

COURSE CODE: MSS505M

COURSE TITLE: Statistical Methods

CLASS DAY & TIME:

ROOM:

NAME OF FACULTY:

COURSE CREDIT: 6 units

CONTACT NO. (DEPT):   536-0270, 524-4611, loc. 420

TERM/SCHOOL YEAR:

COURSE DESCRIPTION

A course on regression, time series analysis, design of experiments, and introductory multivariate statistical methods.

PREREQUISITES: None

COURSE OBJECTIVES

- Appreciate the different statistical methods and the importance of their underlying assumptions that are needed for their appropriate use
- Demonstrate how different statistical methods are used in modeling and decision making using real life situations
- Be aware of the proper uses and applications of statistical methods in many other fields such as business, finance, economics, social science, psychology, biology, medicine, and engineering, among others
- Exhibit values like:
  - cooperation through group study;
  - honesty by claiming credit only for the work he has done;
  - patience, perseverance and diligence;
  - faith by doing what is right and giving his best in performing any assigned task;
  - self-reliance by being able to solve problems independently.

Topic/ Subtopic	Learning Strategies/ Activities	Week /Meeting
PART 1: Linear Regression Analysis	Lecture	6 Hours

Topic/ Subtopic	Learning Strategies/ Activities	Week /Meeting

1. SIMPLE LINEAR REGRESSION
    1.1 The Problem and Motivation Behind Curve Fitting
    1.2 The Least Squares Estimates
    1.3 Maximum Likelihood Estimates
 	
    1.4 Inference Regarding $\beta_0, \beta_1, \sigma^2, \beta_0 + \beta_1 x_0$
    1.5 Correlation: Inference and Relationship to Simple Linear Regression Model (SLRM)
	Class Discussion
	Problem Set
	Computer Lab Exercises	

2. MEASURES OF MODEL ADEQUACY
    2.1 Tests of Linearity
    2.2 Tests of Normality
    2.3 Tests for Homoscedasticity
    2.4 Tests for Independence
    2.5 Outliers Detection
    2.6 Transformations
		4 Hours

3. MULTIPLE LINEAR REGRESSION
    3.1 Matrix Representation and Estimation of Parameters
    3.2 Algebraic/Geometries Interpretation of Multiple Linear Regression Model (MLRM)
    3.3 Tests and Confidence Intervals Based on the T Distribution
    3.4 Full versus Reduced Model: The Partial F Test
    3.5 Extra Sum of Squares and Multicollinearity
		4 Hours

4. VARIABLE SELECTION AND MODEL BUILDING
    4.1 Criteria for Selecting Appropriate Models: MSE, C_p and Adjusted R^2
    4.2 Forward Selection, Backward Elimination and Stepwise Procedures
    4.3 Multicollinearity: the PRESS Statistic and the Hat Matrix
		3 Hours

5. ISSUES IN REGRESSION MODELLING		1 Hour

PART 2: Design of Experiments
 	
  1. Introduction
    1.1 Experimental Designs versus Survey Sampling
    1.2 Some Typical Applications of Experimental Designs
    1.3 Basic Principles
    1.4 Planning an Experiment
	Lecture
	Class Discussion
	Problem Set
	Computer Lab Exercises
		4 Hours

2. Simple Comparative Experiments
    2.1 Review of Estimation and Statistical Hypothesis Testing in Normal Populations
    2.2 Paired Comparison Designs
    2.3 Inferences About the Variances of Normal Distributions
		2 Hours

3. Experiments with a Single Factor
    3.1 Completely Randomized Designs (CRD)
    3.2 One-Way Analysis of Variance (ANOVA)
    3.3 Analysis of Fixed-Effects Models
    3.4 Diagnostic Checking and Model Adequacy Checking
    3.5 Choice of Sample Size
		2 Hours

4. Randomize Blocks, Latin Squares and Related Designs
    4.1 Randomized Complete Block Design (RCBD)
    4.2 Latin Square Design (LSD)
		2 Hours

Topic/ Subtopic	Learning Strategies/ Activities	Week /Meeting
4.3 Graeco and Hyper-Graeco Latin Square Design 4.4 Balanced Incomplete Block Designs		

5. Factorial Experiments
    5.1 Basic Definitions and Principles
    5.2 The Two-Factor Factorial Designs
    5.3 The General Factorial Design
    5.4 The 2^k Factorial Design
    5.5 Blocking and Confounding in the 2^k Factorial Design
    5.6 Two-Level Fractional Factorial Designs
		2 Hours

6. Analysis of Covariance		2 Hours
7. Repeated Measures		2 Hours

PART 3: Time Series Analysis
 	
1. INTRODUCTION
    1.1. Definition of Terms
    1.2 Components of a Time Series
    1.3 Overview of Forecasting Methods
	Lecture
	Class Discussion
	Problem Set
	Computer Lab Exercises
		2 Hours

Topic/ Subtopic	Learning Strategies/ Activities	Week /Meeting

2. STATISTICAL FUNDAMENTALS
  2.1 Summary Statistics Used in Forecasting
  2.2 Measuring Errors
  2.3 Model – Fitting
  2.4 Review of Linear Regression
  2.5 Autocorrelation Function
  2.6 White Noise Behavior
		2 Hours

3. SIMPLE SMOOTHING METHODS
  3.1 Moving Averages
  3.2 Simple Exponential Smoothing
  3.3 Seasonal Moving Averages and Simple Exponential Smoothing
		3 Hours

4. DECOMPOSITION METHODS AND SEASONAL INDICES
  4.1 Additive and Multiplicative Seasonality
  4.2 Classical Decomposition
  4.3 Decomposition Using Regression
		3 Hours

TREND – SEASONAL SMOOTHING METHODS 5.1 Estimating Trend Using First Differences 5.2 Double Moving Average 5.3 Brown's Double Exponential Smoothing **5.4 Holt's Two – Parameter Trend Model**		3 Hours

6. UNIVARIATE ARIMA MODELING
  6.1 Autoregressive Process
  6.2 Moving Average Process
  6.3 Integrated Autoregressive Moving Average Process
  6.4 Use of ACF's and PACF's
  6.5 Parameter Estimation
  6.6 Model Checking
		3 Hours

PART 4: Introduction to Multivariate Analysis
 	
  1. PRELIMINARIES
  1.1 Some Basic Concepts of Multivariate Analysis
  1.2 Types of Multivariate Techniques
  1.3 Classification of Multivariate Techniques
  1.4 Assumption Checking
	Lecture
	Class Discussion
	Problem Set
	Computer Lab Exercises
		2 Hours

2. MULTIVARIATE ANALYSIS OF VARIANCE
 	
  2.1 Description of Multivariate Analysis of Variance
  2.2 Objectives of MANOVA
  2.3 Assumptions of ANOVA and MANOVA
  2.4 One– Way and Two– Way MANOVA
  2.5 Applications of MANOVA
  2.6 Post-hoc Analysis
		3 Hours

3. MULTIVARIATE DISCRIMINANT ANALYSIS
  3.1   Description of Discriminant Analysis
  3.2   Objectives of Discriminant Analysis
  3.3   Assumptions of Discriminant Analysis
  3.4   Linear and Quadratic Discriminant Functions
  3.5   Classification Tables
  3.6   Applications of Discriminant Analysis
		3 Hours

4. FACTOR ANALYSIS		3 Hours

Topic/ Subtopic	Learning Strategies/ Activities	Week /Meeting
4.1 Description of Factor Analysis 4.2 Objectives of Factor Analysis 4.3 Assumptions of Factor Analysis 4.4 Naming of Factors 4.5 Orthogonal and Oblique Rotations 4.6 How to Select Surrogate Variables for Subsequent Analysis 4.7 How to Use Factor Scores 4.8 Application of Factor Analysis		
5. CLUSTER ANALYSIS 5.1 Description of Cluster Analysis 5.2 Assumptions of Cluster Analysis 5.3 Similarity / Dissimilarity Measures 5.4 Types of Clustering Techniques 5.5 Applications of Cluster Analysis		3 Hours

TEACHING STRATEGIES/METHODOLOGY

1. Lecture
2. Report
3. SAS Exercises

REQUIREMENTS OF THE COURSE

1. Problem Sets   50%
2. Laboratory Exercises   50%

TEXTBOOKS

Part 1:

- Bapat, R. B. (2012) Linear algebra and linear models. New Delhi: Hindustan Book Agency/Springer.
- Christensen, R. (2011). Plane answers to complex questions [electronic resource]: The theory of linear models. New York, NY: Springer New York.
- Draper, N.P., and Smith, H. (1998). Applied regression analysis (3rd ed.). New York: Wiley.
- Freedman, D. (2009). Statistical models: theory and practice. Cambridge: Cambridge University Press.
- Graybill, F.A. (1976). Theory and Application of the Linear Model. Mass.: Duxbury Press.
- Kahane, L. H. (2008). Regression basics. Los Angeles: Sage Publications.
- Montgomery, D.C. and Peck, E.A. (1992). Introduction to linear regression analysis (2nd ed.). New York: Wiley.
- Neter, J., Kutner, M., Wasserman, W., and Nachtsheim, C. (1996). Applied linear regression models (3rd ed.). Chicago: Irwin.
- Neter, J., Kutner, M., Wasserman, W., and Nachtsheim, C. (1996). Applied linear statistical models (4th ed.). Chicago: Irwin.
- Searle, S. (1997). Linear Models. NY: Wiley.
- Yan, X. (2009). Linear regression analysis: theory and computing. Hackensack, NJ: World Scientific.

Part 2:

- Box, Hunter and Hunter. (1978). Statistics for Experimenters NY: Wiley.
- Cochran, W.G. and Cox, G.M. (1992). Experimental Designs (2nd ed.). New York: Wiley.
- Hair Jr., Joseph F. et. al. (2010). Multivariate Data Analysis (7th ed.). New Jersey: Prentice Hall.
- Johnson, Richard A. and Wichern, Dean W. (2007). Applied Multivariate Statistical Analysis (6th ed.). NJ: Pearson

Prentice Hall.

- Milliken, George A. and Johnson, Dallas E. (1984). Analysis of Messy Data. New York: Van Nostrand Reinhold.
- Montgomery, D.C. (2009). Design and Analysis of Experiments (7ᵗʰ ed.). New York: Wiley.
- Neter, J., Kutner, M., Wasserman, W., and Nachtsheim, C. (1996). Applied linear regression models (3ʳᵈ ed.). Chicago: Irwin.
- Thomas, R.P. (2007). Modern Experimental Design. N.J.: Wiley

Part 3:

- Abraham and Ledolter. (1993). Statistical Methods for Forecasting. J. Wiley and Sons.
- Bowerman and O'Connel. (1979). Time Series and Forecasting. PWS Pub.
- DeLurgio, S.A. (1998). Forecasting Principles and Applications. Irwin/McGraw-Hill.

  Enders, W. (2010). Applied Econometric Time Series. Hoboken, N.J.: Wiley

  Prado, R. (2010). Time Series: Modeling, Computation, and Inference. Berlin, Heidelberg: Springer Berlin Heidelberg.

  Shumway, R.H. (2011). Time Series Analysis and Its Applications with R Examples. Ny: Springer New Y

  Wei, W.W.S. (2006). Time Series Analysis : Univariate and Multivariate Methods (2ⁿᵈ ed.). Boston: Pearson/Addison-Wesley.

- Young, P.C. (2011). Recursive Estimation and Time Series Analysis- An Introduction for the Student and Practitioner. Berlin, Heidelberg: Springer Berlin Heidelberg.

Part 4:

- Delwiche, L.D, and Slaughter, S.J. (2003). The little SAS book: a primer (3ʳᵈ ed.). Cary, NC: SAS Pub.
- Everitt, B. and Hothorn, T. (2011). An introduction to applied multivariate analysis with R [electronic resource]. New York, NY; Springer New York.
- Fichet, B. (2011). Classification and multivariate analysis for complex data structures [electronic resource]. Berlin, Heidelberg: Springer Berlin Heiderberg.
- Grissom, Robert J. and Kim, John J. (2012). Effect Sizes for Research: Univariate and Multivariate Applications. New York : Routledge.
- Hair, J.F., Black, B., Babin, B., Anderson, R.E., and Tatham, R.L. (2010). Multivariate data analysis: a global perspective (7ᵗʰ ed.). Upper Saddle River, NJ: Pearson.
- Hardle, W. and Simar, L. (2012). Applied Multivariate Statistical Analysis (3ʳᵈ ed.). NY: Springer.
- Johnson, R.A., and Wichern, D.W. (2007). Applied multivariate statistical analysis (6ᵗʰ ed.). Upper Saddle River, NJ: Pearson Education International.
- Lattin, J.M., Carroll, J.D., and Green, P.E. (2003). Analyzing multivariate data. Pacific Grove, CA: Thomson Brooks/Cole.
- Marascuilo, L.A., and Levin, J.R. (1983). Multivariate statistics in the social sciences: a researcher's guide. Monterey, California: Brooks/Cole Pub. Co.
- Morrison, D.F. (1990). Multivariate statistical methods (3ʳᵈ ed.). Singapore: McGraw-Hill.
- Mukhopadhyay, P. (2009). Multivariate statistical analysis. Hackensack, NJ: World Scientific.
- Stevens, J. P. (2009). Applied multivariate statistics for the social sciences. New York: Routledge.
- Timm, N.H. (2002). Applied multivariate analysis. New York: Springer.
- Wehrens, Ron (2011). Chemometrics with R [electronic resource]: Multivariate Data Analysis in the Natural Sciences and Life Sciences. Berlin, Heidelberg: Springer Berlin Heidelberg.
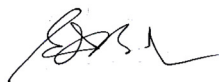
REFERENCES

- Chen, X., Ender, P., Mitchell, M. and Wells, C. (2003). Regression with SAS: http://www.ats.ucla.edu/stat/sas/webbooks/reg/default.html
- Lock, R. WWW Resources in teaching Statistics: http://it.stlawu.edu/~rlock/tise98/onepage.html
- StatSoft, Inc. Electronic Statistics Textbook. Tulsa, OK: StatSoft. Web: http://www.statsoft.com/textbook/
- West, R. Regression Applet: http://www.stat.sc.edu/~west/javahtml/Regression.html
- Concepts of Experimental Design: Design Institute for Six Sigma: http://support.sas.com/resources/papers/sixsigma1.pdf

- Basic Experimental Design:
  http://liutaiomottola.com/myth/expdesig.html
- DoE & Analysis of Experimental Data (using R):
  http://cran.r-project.org/web/views/ExperimentalDesign.html
- What is Experimental Design?:
  http://www.itl.nist.gov/div898/handbook/pri/section1/pri11.htm
- A Field Guide to Experimental Designs:
  http://www.tfrec.wsu.edu/anova/index.html

FACULTY OUTPUT

- Arcilla, R., Co, F. and Ocampo, S. (2011). *"Correlates of Poverty: Evidence from the Community-Based Monitoring System (CBMS) Data"*. DLSU Business and Economics Review, Vol. 20, No. 2, January 2011, pp. 33-43 (ISSN 0116-7111, http://www.philjol.info/philjol/index.php/BER/article/view/1912).
- Beltrano, Elline Jade, Leong, Robert Neil F., and Co, Frumencio F. (2013). Regression Analyses of the Philippine Birth Weight Distribution. *The Philippine Statistician, 62*(2), 31-52.
- Carandang, J. and Co, F. (2012). *"Some factors affecting the student evaluation ratings of Biology faculty at DLSU"*. Proceedings of the 3rd International DLSU Education Congress, DLSU College of Education, Manila, September 2012.
- Co, F., Arcilla, R., and Ocampo, S. (2012). *"Correlates of Hunger: Evidence from the CBMS Data of Pasay City"*. Proceedings of the 2012 Philippine Statistical Association Annual Conference, Quezon City, August 2012.
- Janairo, J.I.B., Janairo, G.C., Yu, D.E.C. and F. Co. (2010). *"Regression Analysis on the Chemical Descriptors of a Selected Class of DPP4 Inhibitors"*. Studies in Mathematical Sciences, Vol. 1, No. 1, 2010, pp. 01-06 (ISSN 1923-8444-Print; ISSN 1923-8452 – Online, www.cscanada.net).
- Janairo, J.I.B., Janairo, G.C., Yu, D.E.C. and F. Co. (2011). **"***Assessing the Binding Affinity of a Selected Class of DPP4 Inhibitors using Chemical Descriptor-Based Multiple Linear Regression"*. Orbital (The Electronic Journal of Chemistry), Vol. 3, No. 1, January – March 2011, pp. 01-06 (ISSN 1984-6428 , http://www.or-bital.ufms.br/inpress/inpress.htm).
- Ocampo, S., Arcilla, R., Co, F., Jumangit, R. and F. J. Diokno. (2011). *"Exploring Latent Factors Using Non-Bayesian and Bayesian Factor Analyses"*. Proceedings of the DLSU Science and Technology Congress, DLSU, Manila, February 2011.
- Ocampo, S., Arcilla, R., Co, F., Jumangit, R. and F.J. Diokno. (2013). *"Enthusing students towards statistical literacy using transformative learning paradigm: Implementation and Appraisal"* . Proceedings of the 2013 IASE/IAOS Conference, IASE/IAOS, Hong Kong/Macau, China, August 2013.

Noted by:

Dr. Isagani B. Jos
Chair, Department of Mathematics

Dr. Jose Santos R. Carandang VI
Dean, College of Science