Región Region Région
Centro América y Caribe
Central American and Caribbean
Caraïbe et Amérique Centrale

'Biometry, the active pursuit of biological knowledge by quantitative methods."
R.A. Fisher, 1948

ABSTRACT BOOKLET

6th biennial meeting of the Central American and Caribbean Region of the International Biometric Society (IBS-RCAC)

"Why Statistics? Impact And Implications Of Statistics And Statistical Data Analysis

Mona Visitors' Lodge and Conference Centre, The University of the West Indies, Kingston, Jamaica
November 16-18, 2011

Spatial and Mixture Models and Design Considerations for Recurrent Event Processes: with Application in Health and Environmetrics

Professor Charmaine B. Dean, Dean of Science, University of Western Ontario
e-mail: sciencedean@uwo.ca

Many scientific studies involve the observation of a group of subjects where repeated measurements are taken on each subject. In fact, recently there has been substantial interest in the development of tools for the analysis of such longitudinal studies, with those in the health context being predominant. For example, studies of recurring infection or chronic disease often collect longitudinal data on the number of specific disease-related events which occur between specified follow-up periods. These studies are also important in environmetrics applications. In a forestry study of recurrent weevil infestation conducted by the Ministry of Forests in British Columbia, each tree within a plantation was examined annually for the presence or absence of infection, giving rise to longitudinal binary data. The objective of the study is to describe the movement in the status of trees between the infected and uninfected states and to characterize the variation in this transition process over space and time.

Flexible semi-parametric models for analyzing such longitudinal panel count data are presented. The emphasis here is on the need for incorporating mixture models which permit differences in the behaviour of sub-populations under investigation and on the development of methods for studying spatial variation and estimation spatial or location effects. Additionally, we consider the more general problem of contrasting efficiencies of study designs based on continuous follow-up versus follow-up at regular intervals, where counts of the number of events between follow-up times are recorded. We discuss the loss in efficiency through the analysis of such aggregate data, as well as optimal design strategies for recurrent event studies.

Several examples will illustrate the methods discussed including an epidemiological study of revascularization intervention in the province of Quebec, a forestry study of recurrent white pine weevil infestation in British Columbia, and a biological study to assess the effectiveness of a pheromone treatment in disturbing the mating habits of the cherry bark tortrix moth. The talk will also discuss the rich opportunities for research in this area.

Región Region Région
Centro América y Caribe
Central American and Caribbean
Caraïbe et Amérique Centrale

"Biometry, the active pursuit of biological
knowledge by quantitative methods."
R.A. Fisher, 1948

Mixed Model – Many Faces

Novie Younger-Coleman, Tropical Medicine Research institute, University of the West Indies, Mona, Jamaica.
e-mail: novie.younger@uwimona.edu.jm

Statistical and computer programming power has evolved alongside the increasing complexity of data and the forms in which data are collected. As life becomes more complex, data are collected to answer research questions may include selection of sampling units at many different levels.   As such, the mixed models methodology has expanded to allow for appropriate accounting for increasing numbers of levels of variability in the data. This paper introduces, particularly the end-user of such models, to the many options available to researchers who need to use mixed models. It starts with a description of two-stage modelling, one of the earliest forms of mixed models, and goes through the strengths and limitations of the different methodologies through to multi-level modelling.

Negative binomial-normal generalized linear mixed models

Raúl E. Macchiavelli*, Jairo A. Ayala Godoy, Jhonny Demey
raul.macchiavelli@upr.edu, Dept. of Crops and Agroenvironmental Sciences, University of Puerto Rico
jairoarturo.ayala@upr.edu, Department of Mathematical Sciences, University of Puerto Rico
jdemey@reacciun.ve, Instituto de Estudios Avanzados, IDEA, Venezuela

The generalized linear mixed model (GLMM) is the most widely used random effects model. It is a good alternative to traditional linear mixed models if the normal conditional distribution assumption is not satisfied. The negative binomial distribution, on the other hand, has been widely used to model counts, and it is the standard alternative for overdispersed Poisson counts. For repeated measurements and other correlated data, GLMMs using negative binomial distribution with normal random effects can be very useful to model counts, accounting for posible correlations, and for overdispersion. In this work we study some properties of this model, such as the induced marginal distribution, its moments, and the relationship between the conditional distributions defining the model and the induced marginal distribution. Many of these properties are studied using simulations, since they are analytically intractable. We apply these models to a study of seed counts collected at different locations and times in a forest in Puerto Rico between 2006 and 2008.

Importance of Limited Dependent Variable Models in Consumer and Environmental Economics

Abdullahi Abdulkadri
Department of Economics
The University of the West Indies, Mona

The talk aimed to illustrate application of statistical models of limited dependent variable in consumer and environmental economics using data from two studies. Study 1 examined how participation in MGSP (Michigan Groundwater Stewardship Programme) influences farmers' adoption of groundwater stewardship practices. Study 2 assessed public perception of genetic engineering and the impact that an individual's perception of the technology may have on their decision to purchase GM food. A probit model was used to estimate likelihood of stewardship practice adoption. A count data (Poisson) model was used to explain intensity of stewardship practice adoption. An ordered logit model was used to estimate likelihood of purchasing GM food.

The analyses revealed that the ordered logit model did not yield a good fit as none of threshold parameters is significant and the prediction accuracy is poor. A logit model offers a better fit. Conclusion based on the results were that knowledge of groundwater issues is important in encouraging farmers to become an adopter of GSP and MGSP participation is critical in inducing a more intense adoption of groundwater stewardship practices by farmers; voluntary approaches can induce environmental stewardship in farmers; level of awareness of and perceived benefits from GE are major determinants of purchase decision; and consumers are more supportive of the use of GE on plants than livestock.

Introduction to Stata

Christine A. Walters, PhD
Office of the Dean, Faculty of medical Sciences, University of the West Indies, Mona, Jamaica

"Powerful yet easy to use, Stata lets researchers manage data their own way. Stata's built-in commands perform a broad range of basic and complex statistical analyses. Its programmable nature also permits users to tailor Stata specifically to their needs."(-StataCorp, 2011) The workshop was aimed at equipping participants to navigate the features and functions of Stata's user interface; explore the contents of a Stata dataset; produce basic descriptive univariate summaries and bivariate assessments; use simple Stata commands; transfer data between Stata and Excel files.

Abundance estimation of littorina in the Magdalenian rocky coastline (Colombia) using generalized model assisted survey estimation

Alejandra Polanco Rojas1;a, Liliana López Klein1;b, Leonardo Trujillo Oyola1;c
1Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá D.C., Colombia

Accurate and precise population abundance estimates of closed biological populations are required for ecological and environmental studies. There is a great variety of strategies in the literature considering both a sampling design and a corresponding estimator. Despite their broad application, these common strategies must follow several assumptions difficult to follow in the study of particular animal populations such as independence, low mobility and a constant population. These requirements are difficult to fulfill in ecological studies due to concomitant species, dependent capture probabilities of individuals and time seasonality of abundance, among many others. Therefore, new sampling strategies are required to estimate animal abundance when the usual assumptions do not follow and to take into account known relationships with other auxiliary variables. In this paper, we will study the abundance of some intertidal gastropods (*Littorina* sp.) in relation with some available auxiliary information at the Colombian Caribbean coastline at four localities (Cinto, Gayraca, Neguanje and Rodadero). It is know the high correlation between the abundance of Littorina with the abundance of other intertidal gastropods (such as *Nerita* sp.). Different sampling strategies were compared with a new methodology considering generalized survey assisted estimation which takes into account the available information in the aerial sampling frame. Different estimators were considered using different distributions of the errors of the generalized linear models involved due to the intrinsic nature of the data and the high encountered heterogeneity at the collection data stage.
Key words: Animal abundance, generalised linear models, *Littorina* sp., negative binomial regression, *Nerita* sp..

Simpson's Paradox: It is good for Boys, good for Girls but bad for Children

Samuel McDaniel, Department of Mathematics and Computer science, University of the West Indies, Mona, Jamaica.
Using an example based on consumption of a given food by children of different sexes, the presentation illustrates that statistically significant results in subgroup analysis may not emerge if the subgroups are combined.

Important Variables for the Periodontal Health in the Diabetics Patients: A Multivariate Approach

Sandra Vergara Cardozo1*, Mauricio Rodríguez2, Carlos Serrano3,Carlos Tadeu dos Santos Dias4
1* Statistical Departament National University of Colombia, SA - CP 111321- Bogotá- Colombia.
2,3 Oral Health Department. School of Dentistry National University of Colombia
4 USP/ESALQ;Depto de Ciências Exatas-Piracicaba, SP – Brasil.
Corresponding author *svergarac@bt.unal.edu.co

We analyzed the present of some important variables to the periodontic health in adults patients with and without diabetes. This study was developed taking account the information available in a sample of 72 patients with diabetes and 39 without diabetes. The sample was taken from the diabetic program and the odontology clinic of the National University of Colombia (2008-2009). The results are analyzed appling multivariate methods, the first exploratory analysis identify the variables which affect the way to rise the periodontic in diabetics patients and no diabetics patients, using biplot graphics methodology, as was described by Gabriel (1971). To reduce the number of the exploratory variables, and taking into account the presence of the multicollinearity. It´s use the technique of analysis of the principal components
Keywords: Multivariate analysis, biplot, diabetes, periodontal disease.

Parametric and Non-Parametric Bootstrapping for Estimation of Resource Selection Probability Functions

Sandra Vergara Cardozo1* , Bryan F.J. Manly2, Carlos Tadeu dos Santos Dias3
1*Departament of Statistical University National of Colombia, SA - CP 111321- Bogotá- Colombia.
2Western EcoSystems Technology Inc., Cheyenne, Wyoming, USA.
3 USP/ESALQ;Depto de Ciências Exatas-Piracicaba, SP – Brasil.
Corresponding author < svergarac@bt.unal.edu.co >

Resource selection functions (RSFs) are used for quantify how animals are selective in the use of the habitat period or food. A Resource Selection Probability Function (RSPF) can be estimated if N, the total number of units in the population, and n1 the total number of used units in the study period are both known and small. An approximation of the RSPF can then be estimated using any standard program for logistic regression but the variances of the estimates of the parameters are too small. Three methods of bootstrap sampling, parametric, non-parametric and a modified parametric method are proposed for the estimation of variances, with a discussion about the limitations of logistic regression for estimating RSPF. The method for estimating the RSPF described here has potential applications in medicine, ecology and other areas.
Keywords: Resource selection functions (RSFs); Resource selection probability function (RSPF); Bootstrap; Logistic regression

F TESTS in STRIP-SPLIT-PLOT DESIGN

Días P Daniel[1], López P Luis A[2] y Zimermmann Francisco[3]
1 Biostatiscts Program University of Florida-USA, 2 Departamento de Estadística-Universidad Nacional Bogotá lalopezp@unal.edu.co y Universidad Católica de Goias-Brasil

There are experimental situation where a researcher needs to know the behaviour of a factor in relation to one and/or two additional factors. When this happens, it is usually considered a split plot design adding subplots into the former plots. This is done partly ,   to the great development reached by this type of experiments .An alternative at this design is the Strip-split-plot -design (SSPD) ; that is an extension of strip block designs such that each plot on the intersection is subdivided into subplots in order to insert a third factor. This new factor will be more accurate on its measurement due to the high number of its observations and interactions; which is the more important feature of the design.In this work, we presents the structure of the F tests, the variance

components and the approximate degrees of freedom for each of the eight possible mixed models of the strip-split plot design.

Key words: Experimental design, F tests, Strip-split-Plots.

The Performance of the Jamaica Stock Exchange index (JSE) and its Impact on Economic Growth

Luwayne Thomas
Department of Economics, University of the West Indies, Mona Campus

This study examines the statistical relationship between the performance of the JSE index and growth of the Jamaican economy (GDP). There have been several studies done on the relationship between stock market development and economic growth of countries; this phenomenon is yet to be investigated for Jamaica however. **Methodology**: Co-integration and Error Correction Models (ECM) complimented with Granger causality tests were utilized to unearth statistical relationship between the independent and dependent variables. The Engle and Granger co-integrating procedure was used to determine the existence of a long run relationship between the variables. While the Granger causality test was employed to scrutinize the directional dynamics of the relationship between GDP growth rates and the performance of the JSE index. **Results**: The findings showed a long run relationship between the variables. There was also a statistically significant causality, running from the JSE index performance to GDP growth. **Conclusion**: The performance of the JSE index positively affects economic growth over the long run.
Keywords: Stock Exchange Performance, Economic Growth, Granger causality, model
JEL classification: E44, G11, G18

Poster Presentations
Discriminating community assembly processes through functional diversity indices

Laura Pla[1], Fernando Casanoves[2], María del Carmen Romero[3] and Julio Di Rienzo[4]
[1] *laura@reacciun.ve, Universidad Nacional Experimental Francisco de Miranda, Venezuela;*
[2] *casanoves@catie.ac.cr, Centro Agronómico Tropical de Investigación y Enseñanza, Costa Rica;*
[3] *mariadelc.romero@gmail.com Universidad Nacional del Centro de la Provincia de Buenos Aires;*
[4] *dirienzo@agro.unc.edu.ar Universidad Nacional de Córdoba, Argentina*

Three main ecological assembly processes are responsible for the pattern of species composition in a community. This study quantifies the capability of functional diversity indices to detect the effect of these processes using trait´s information. We use three assembly processes to generate patterns of coexisting species from 250 random species pools assuming that traits are phylogenetically conserved: limiting similarity to produce even dispersion, habitat filtering to produce clustered dispersion and neutral assembly to produce random dispersion. We used lognormal distribution to model abundances. Using 11 richness (from 20 to 200) we estimate eight functional diversity indices and compare confidence intervals (CI) for the three community assembly processes.
The ability of indices that do not use abundance to detect assembly is better when include the 100% of species. For example, FD with an error of 53.7% improves its ability and has 24.2% classification error rate if considering information on trait values for all the species. With 90% of species when there are no abundances, but using FD, MFAD and FRic together has an acceptable 4.4% classification error rate. For the combination of all indices that include abundance, the classification error rates were 2.1% with all the species and 0.8% with those that covered the 90%. We may conclude that if the aim is to detect possible differences in assembly processes, indices based on trait and abundance values perform better than those that do not include estimation of species contribution to the community.
Key Word: environmental filtering, neutral assembly, limiting similarity, discriminant analysis, confidence intervals