MASTERAL THESIS ABSTRACTS

PRINCIPAL COORDINATE SCORING PROCEDURE IN THE ESTABLISHMENT OF A PHILIPPINE MUNGBEAN *(VIGNA RADITA* (L.) WILCZEK) CORE COLLECTION

MAE P. CABABASAY*

The materials assembled and maintained in germplasm collections provide a wide genetic base for breeding programs and provide safety backups to the dangers of a highly uniform genepool. However, the growing size of the germplasm collection poses problems in its maintenance and utilization. As a solution to the growing size of the germplasm collection, a core collection, a collection of smaller size but of maximal diversity, can be established.

In this study, a procedure for the establishment of a core collection was developed. The developed procedure uses the morphoagronomic characterization of the accessions as criteria in the selection of the core entries. It is divided into three phases, namely: (1) identify the accessions with outlying values on any of the quantitative traits; (2) stratify the germplasm collection based on the outlying quantitative traits; and (3) from each stratum of the germplasm collection, select a subset of accessions that captures the qualitative traits diversity of the stratum. Identification of the outliers facilitates the nomination of the atypical accessions. Stratifying the germplasm collection into groups of related accessions and selecting a representative subset from each stratum improves the core's captured quantitative traits diversity. On the other hand, the use of qualitative variables as criteria for the selection of a stratum's representative subset facilitates the capture of the qualitative trait phenotypes.

The principal coordinate scoring (PCOS) procedure was developed to implement Phase 3 of the aforementioned procedure on core establishment. The PCOS procedure uses principal coordinate analysis to come up with the coordinates of the projection of the accessions in a reduced Euclidean space. Based on the inter-accession distance in the reduced Euclidean space, a subset of accessions that maximizes inter-accession dissimilarity is selected.

The procedure on core establishment was used to establish a core from the Asian Vegetables Research and Development Center (AVRDC) Philippine Mungbean (*Vigna radiata*) germplasm collection. Evaluation of the established core revealed that with respect to each of the quantitative variables, core entries covered the ranges defined by the minimum, 25th percentile, median, 75th percentile and maximum values of the germplasm collection. Furthermore, with respect to the qualitative trait phenotypes captured by the core collection, the principal coordinate scoring procedure was compared with random selection of accessions and was found to be better than the latter.

*Master of Science in Statistics, Institute of Statistics, UPLB (Adviser: Dr. Ann Inez N. Gironella)

EXTENSION OF THE LIMITING QUALITY INDEXED SINGLE ACCEPTING SAMPLING PLANS FOR ATTRIBUTES IN HIGH PRECISION PROCESSES

MARIA THERESA GOMEZ*

This study investigated the extension of existing single acceptance sampling plans indexed by lot tolerance percent defective (LTPD) or limiting quality (LQ) with respect to their applicability to high precision processes. LTPD/LQ indexed sampling plans were extended to cover the very low fraction defective levels of high precision processes. Plans based on the ISO 2859-2 LQ indexed plans, Dodge-Romig LTPD indexed plans and lot sensitive plans (LSP) were used as bases for the extension. Target levels for LTPD/LQ were set at defective parts per million levels ranging from 20 ppm to 5000 ppm. The performance of the extended plans was evaluated using measures relating to level of protection afforded by the plans and efficiency in terms of amount of required inspection. The extended Dodge-Romig LTPD plans showed best performance among the three plans generated. Three selected plans from Dodge-Romig LTPD sampling scheme were then subjected to simulation studies. Comparison of the theoretical and simulation data indicated that the plans behave more strictly than expected from theoretical calculations. The specific demands of high precision processes for appropriate sampling inspection plans that cover lower fraction defective levels, and at the same time satisfying the requirement for economy and efficiency of inspection were shown to be provided by this plan. These results provide a significant contribution to the manufacturing industry as it continuously strives to improve process yields and decrease fraction defective levels to respond to customer demands of better quality and improved performance.

*Master of Statistics, School of Statistics, UP Diliman (Adviser: Dr. Daniel C. Bonzo)

A COMPARISON OF SELECTED SMALL AREA ESTIMATION PROCEDURES

MA. JOSEFA ROSARIO Y. ABANILLA*

Four small area estimation procedures, namely (1) synthetic estimation; (2) weighted least squares; (3) Empirical Best Linear Unbiased Predictor (EBLUP); and (4) Empirical Bayes (EB) are compared in terms of precision using city/municipality level data relating to poverty incidence in Nueva Ecija and Pangasinan. The bootstrap standard errors of the estimates provided by the four methodologies are compared with bootstrap standard errors of the direct estimation as well. Results indicate that the EB estimator outperforms the other estimation procedure for the given data.

*Master of Science (Statistics), School of Statistics, UP Diliman (Adviser: Dr. Ana Maria L. Tabunda)

STEEL DEMAND FORECASTING: A COMPARISON OF ERROR CORRECTION, TRANSFER FUNCTION AND TIME SERIES REGRESSION MODELS

JOSEPHINE G. MARTINEZ*

Most classical time series models are based on the assumption that the series is stable over time. In the univariate case, a stochastic trend can be removed by differencing and the resulting series can be modelled using the Univariate Box-Jenkins (UBJ) techniques. However, treating nonstationary variables is not so straightforward in a multivariate context (Enders, 1995).

Engle and Granger (1987) introduced the concepts of cointegration and error-correction mechanisms. A principal feature of cointegrated variables is that their time paths are influenced by the extent of any deviation from long-run equilibrium and if the system is to return to the long-run equilibrium, the movements of at least some of the variables must respond to the magnitude of the disequilibrium. This dynamic model is one of error-correction.

Error-correction was used to obtain forecasts of steel demand using indicators such as construction floor area and value of production index. The forecasts derived from the error-correction model were compared to those of transfer function and time series regression models using the mean prediction error (MPE) and mean absolute prediction error (MAPE).

The results of the study indicate that billet demand and construction floor area are cointegrated. The error-correction model developed shows that the behavior of billet demand and construction floor area do not drift too far apart and they move in the same direction. Deviations from one period on the long-run equilibrium can be corrected by the error-correction term in the model in the next period.

Within reasonable degree of accuracy, the error-correction model is the most efficient forecasting technique in determining future values of billet demand based on construction floor area in the long-term period. For short-term periods, transfer function and time series regression modelling have comparable fit in determining future values of billet demand.

*Master of Statistics, School of Statistics, UP Diliman (Adviser: Dr. Daniel C. Bonzo)

COMPARING THE EFFICIENCIES OF THE MULTI-LAYER PERCEPTRON, KERNEL DISCRIMINANT ANALYSIS, AND LOGISTIC REGRESSION FOR AUTOMOBILE CREDIT SCORE MODELLING

ALETA Y. LEPATAN*

Neural Networks have been extolled in the US as being more predictive than traditional statistical approaches when modelling credit scores. As a way of validating this assertion, this study sought to compare the classification and predictive efficiencies of the Multi-Layer Perceptron (MLP), the most common network used for classification and prediction problems, against the logistic regression and the kernel discriminant analysis for automobile credit score modelling.

In general, findings reveal that the rank transformation is a useful method for improving the performance of the three statistical approaches. It was responsible for increasing the discriminating power of the models under consideration when applied to a data set with no grouping patterns at all. But application of the ranks in a simulated scoring process proved to be difficult.

The specific design of the MLP used for this study, gave comparable classification results to the logistic of the kernel model. Although the MLP did not predict as well as the other two models for this particular data set, the results cannot be deemed conclusive because only a single network design was employed for the study. Time constraint and the limitation of computer resources prevented the exploration of other designs and network.

The kernel discriminant analysis, on the other hand exhibited good classification and prediction power but did not appear to lend itself very well to some of the requirements of credit scoring.

Finally, the logistic model using the rank transformation, gave the best classification and prediction performance among the models considered for this study. But the practical problems posed by the application of ranks in a simulated scoring process considerably reduced its predictive ability.

* Master of Statistics, School of Statistics, UP Diliman (Adviser: Dr. Daniel C. Bonzo)

APPLICATION OF SELECTED DISCRETE DISCRIMINANT MODELS ON THE SELF-CONCEPT OF HIGH SCHOOL SENIORS

RUEL B. GUIRINDOLA*

The performance of three discrete discriminant models namely, the Full Multinomial Model, the Second-Order Model of the Bahadur Representation and the First-Order Interaction Effect Model of the Martin-Bradley Reparametrization were used in classifying senior high school students into two (2) academic performance groups. The explanatory variables include their dichotomized scores based on the median scores on accepting attitudes, peer relations, and personal worth factors of self-concept. The models yield comparable error rates in both training and validation sets.

The models seem to allocate students displaying high positive scores (above the median) in at least two self-concept factors to high academic performance group (group with grade-point averages above the median grade) and to low academic performance group (group with grade-point averages below the median grade) otherwise. Based on the magnitude of marginal parameter estimates of the Martin-Bradley reparametrization scheme and on the way the full multinomial model and the second-order model of the Bahadur representation allocated state (1 0 0) corresponding to high in accepting attitudes factor, low in peer relations factor and low in personal worth factor, accepting attitudes seems to be the most important factor among the self-concept indicators.

*Master of Statistics, School of Statistics, UP Diliman (Adviser: Dr. Erniel B. Barrios)

SHORT RUN DEMAND MODELS FOR SELECTED PETROLEUM PRODUCTS

VIVIAN A. AMANDY*

Monthly demand for selected petroleum products was investigated using regression models. Among the petroleum products discussed were diesel fuel oil, regular and premium gasoline, aviation turbo and kerosene. Regression models for the consumption of each petroleum product were obtained using its own prices and tax, prices of related petroleum determinants and other economic variables. Autoregressive models were employed whenever appropriate.

*Master of Statistics, School of Statistics, UP Diliman (Adviser: Prof. Cristina Sotto)