A PEARL OF NUMBER THEORY ON THE SHORE OF COMBINATORICS

Sukumar Das Adhikari

Rightly finding its place among the ‘pearls’ that Kinchin presents in his ‘Three pearls of Number theory’, the theorem of van der Waerden we are going to talk on, led to many interesting developments in Combinatorics and Number Theory.

We shall discuss some history, state several generalizations and mention few applications of the theorem. As we shall see, the theorem of Hales and Jewett, revealing the combinatorial nature of van der Waerden’s theorem would claim that this ‘pearl of number theory’ belonged to the ancient shore of Combinatorics. Yet, there remains very difficult unsolved problems arising out of this theme and some much deeper aspects of it belong to Number Theory.

adhikari@mri.ernet.in

EFFICIENT NON-RESPONSE-ADAPTED AND IMPUTATION-BASED REGRESSION-TYPE ESTIMATORS

M. C. Agrawal
Department of Statistics, University of Delhi, Delhi, India

with

D. C. Roy (Department of Statistics, University of Rajshahi, Bangladesh)

In the event of non-response, we have constructed, appealing to an intuitive predictive set-up under the fixed-population approach, certain imputation-based regression-type estimators of the population mean. We have established their viability and superiority vis-a-vis the existing ones proposed in the literature. We have also appraised the performance of the competing estimators under a plausible super population model.

mc_agrawal@yahoo.com

AN APPLICATION OF MEDIATION AND MODERATION IN HEALTH SCIENCES

Sangeeta Agrawal
Niedfelt Nursing Research Center, College of Nursing, University of Nebraska Medical Center, Omaha, Nebraska, U.S.A.

Hierarchical regression analyses is used to identify mediating and moderating variables in Health Sciences.
A moderating variable (as defined by Baron and Kenney, 1986) affects the direction or strength of relationship between independent and dependent variable. A mediating variable explains the relationship between Independent and Dependent variable by defusing the mechanism by which the former affects the latter. The first part of the paper explains steps involved in identifying significant mediating and moderating variables using hierarchical regression.

In the second part of the paper the mediating moderating role of self-efficacy variables will be examined in an effort to explain temporal changes in drinking behavior. 825 alcohol abusers were interviewed about their confidence, seriousness, importance and intention to quit drinking, at an assessment to assess their self-efficacy. These variables correlated significantly with drinking variables measured at 1-year pretreatment and 1-year post-treatment the Timeline follow back assessment method. A mediation - moderation model will be used to identify role of self-efficacy in explaining temporal changes in drinking behavior.

sagrawal@unmc.edu
In stratified sampling the sampler has to decide about the sample sizes from various strata before drawing a sample. In sampling literature this problem is known as the “Allocation Problem”. Equal, Proportional and Optimum allocations are well known in sampling literature. In practice any one type of allocation is selected according to the prevailing circumstances in the population and is applied to all the strata of the population. However, there are practical situations in which the prevailing circumstances markedly differ from one group of strata to other. In the present paper we propose to divide the group of strata into non-overlapping and exhaustive subgroups according to some common and useful characteristic. The use of a particular allocation is then advised in any particular subgroup depending upon the characteristic property of the subgroup. Since different allocations are to be used in different subgroups we named our proposed allocation as the “Mixed Allocation”. The variance of the estimator based on a stratified sample under mixed allocation is worked out and compared with the variances under the over all proportional and optimum allocations. The relative increase in the variance due the use of mixed allocation is studied to decide whether a mixed allocation is advisable or not in any particular situation. A numerical example is also worked out to illustrate the computational details.

ALTERNATIVE ESTIMATORS FOR MULTI-CHARACTER SURVEYS IN COMPLEX SURVEY DESIGNS

Raghunath Arnab
University of Durban-Westville, South Africa

The problems of estimating finite population mean in multi-character surveys are considered in a unified set up. It is shown that under certain superpopulation model, the sample mean fares better than any other estimator belong to the class of linear homogeneous unbiased estimator for the population mean. Alternative estimators for various sampling designs are proposed when measure of size is not well related to the study variable. Sufficient conditions for superiority of the proposed estimators over the conventional Horvitz Thompson estimator under superpopulation model have been derived. Numerical illustrations are also presented to study the performances of the proposed estimators.

APPLICATIONS OF INFORMATION MEASURES TO MEASURE THE AMBIGUITY LEVEL OF QUALITATIVE STATEMENTS

Gurdial Arora
Department of Mathematics, Xavier University of Louisiana,
New Orleans, LA 70125, U.S.A.

Many times it has been observed that there is a lot of ambiguity in the information that is being
communicated. Ambiguity in statements not only creates confusion for the reader but also it fails to serve the intended purpose of the message. Often survey developers use a pretest technique to reduce the level of ambiguity in their questionnaire. Ambiguity Index, a new measure discussed in this article, is devised to help not only survey developers but also professionals in other fields in reducing the level of ambiguity in statements. An example is discussed to illustrate the simplicity in use and understanding of this measure and the author has also provided some scenarios in which this measure can be adopted by professionals in other fields. It is also shown that the results from this new measure are quite comparable to the results obtained from different measures of information that are widely used in the literature.

gdial@xula.edu

A NOTE ON GENERATION OF PELL NUMBERS, GENERALIZED FIBONACCI NUMBERS AND TRIBONACCI NUMBERS

Gurdial Arora
Department of Mathematics, Xavier University of LA,
New Orleans, LA 70125, U.S.A.

with

Mary Moore

In this paper, we will use some unorthodox ways to generate Pell numbers, Generalized Fibonacci numbers and Tribonacci numbers. Also some new identities will be proved involving these numbers.

gdial@xula.edu

IMPROVING THE QUALITY OF A SYSTEM THROUGH ROTATABLE DESIGNS EXPERIMENTATION

Srijib Bhusan Bagchi
Burdwan University, Burdwan

with

Anis Chandra Mukhopadhyay (ISI, Kolkata, India) and Rabinra Nath Das (Burdwan University, Burdwan, India)

Statistical design of experiments has been used abundantly in reliability theory in order to improve the quality of a system. Experimentation is a powerful tool locating the optimum levels of the parameters in the production process.

In this paper we have used response surface designs (optimum rotatable designs) incorporating independent experimental errors for improving the performance of a system by increasing the reliability for a given mission time or increasing the mean life. Rotatable designs with independent and homoscedastic errors are used in estimating regression parameters by Least Squares method assuming life distribution to follow Weibull distribution and experimental errors follow independent Normal distribution.

sb_bagchi@hotmail.com
FORCE MULTIPLIER INCORPORATION IN COMBAT SIMULATION

Tapan P. Bagchi
Professor of Management, Indian Institute of Technology,
Kanpur 208016, India

with

Raktim Chakrabarty

This paper will present a new methodology for incorporating force multipliers in stochastic Lanchester equations describing two-force combat. Force multipliers, particularly those based on information support such as Battlefield Surveillance Radar (BFSR), Unmanned Arial Vehicles (UAVs) or Remotely Piloted Vehicles (RPVs), are considered major aids in warfare today in improving the effectiveness of combat without requiring additional forces. This paper will present the mathematical models and their implementation in a time-stepped simulation. Battle and attrition statistics are generated for situations with and without the presence of the force multipliers.

bagchi@iitk.ac.in

ROBUST BAYESIAN ANALYSIS OF SHIFTING NORMAL SEQUENCE: SENSITIVITY TO PRIOR

Ashok K. Bansal
University of Delhi, Delhi, India

Shift point inference problem occurs when mechanism underlying sequence of observations abruptly changes. Essentially, there are two problems associated with a changing model. First, one will want to detect a change, and second, assuming that a change has occurred, be able to estimate it as well the other parameters of the model.

Bayesian approach is often used to avoid analytic and numerical difficulties. Broemelerg and Tsurume (1987) provided solutions using normal-gamma priors for pre- and post-change mean and precision of the normal sequence. A valid criticism of the Bayesian methods is that the chosen conjugate prior may not be appropriate. It is, therefore, important to examine sensitivity of posterior inferences to potential misspecification of the prior distribution. Very little has appeared in regard to robustness and changing parametric models.

We shall discuss effects of non-normality in the post-change prior distribution of the unknown normal mean and also those of moderately non-gamma priors for changing precision of the normal sequence on posterior odds used for detection of the change also on the Bayes estimates of the change point.

dustats@del3.vsnl.net.in

ENTROPY AND INEQUALITY

M. A. K. Beg
In this paper we are dealing with the problem on income inequality and define a measure of income inequality based on a measure of information proposed by Renyi. We will also discuss the problem of reduction of this inequality through taxation.

baigmak@rediffmail.com

O. K. Belwal
Department of Statistics, H. N. B. Garhwal University, Srinagar (U.A)

The purpose of the study is to identify the best set of economic variables which are responsible for rural-urban out migration in the newly created Uttaranchal state. The study is based on population census 1991 and Statistical abstract of 13 districts of Uttaranchal.

hnbgucc_pao@sancharnet.in

Ashish Kumar Bharti
Pt. R. S. University, Raipur (C.G.)

Sambit Kumar Mishra (Samajik Vikas Sanstha, At-Gwalndih, P.O.-Godam, Distt-Raigarh)

The paper reviews the sources of market failure in financial institutions and markets and what can be done to alleviate them. It examines game-theoretic explanations for financial instability, in particular the role of asymmetric information in generating destabilising behaviour. In the area of remedies, the paper analyses the potential contribution of official safety nets and what can be done to minimize the associated moral hazard. In this context, it discusses the role of regulation and transparency with the aid of statistical techniques.

sambit_mishra_2001@yahoo.com

Sandeep Bhougal
Department of Statistics
A sequential procedure based on maximum likelihood estimator is proposed for fixed-width confidence interval estimation of mean of an inverse gaussian distribution when dispersion parameter is known. The procedure is shown to be “asymptotically efficient and consistent”. Asymptotic distribution of the stopping time and second-order approximations for average sample number are obtained for the procedure.

Rahul Gupta

PREDICTING PROPERTY AUCTION OUTCOMES USING LOGISTIC REGRESSION

Kyungjin Byun
Bond University

John S. Croucher

This paper looks at the attributes of over 140 properties that sold at auction in the Ku-Ring-Gai region of Sydney, Australia during 2001. Since this is rapidly becoming the preferred method of sale by many vendors in a number of areas of the country, a study is made to find just what factors contribute mostly to a successful outcome on the day. Logistic regression is used to determine the most significant attributes of these properties as well as the probability that a property with these qualities will actually sell. The models used can be readily applied to property auctions in any region of the world.

kbyun@staff.bond.edu.au

SAMPLE FRACTAL DIMENSION AS AN ESTIMATOR OF FRACTAL DIMENSION

Amitabha Chanda
Indian Statistical Institute, 203, B T Road, Kolkata 700 108

It is observed that in case of a monofractal object sample fractal dimension may be equal to the fractal dimension of the object. For example in the case of an irregular curve which is monofractal in nature the fractal dimension of any section of the curve is equal to the fractal dimension of the curve. But otherwise, when the curve is of multifractal nature the segments of the curve may show different fractal and multifractal dimensions. In the present paper the distribution of fractal dimensions of the sample segments is discussed.

amitabha39@yahoo.co.in

BAYESIAN RELIABILITY ESTIMATION OF SINGLE COMPONENT UNDER ACCELERATED LIFE
Several Components have a long-life under the few circumstances where these products are being used in the environmental conditions. Especially for extremely reliable Components/Items it is in general not possible to make life tests under usual stress because testing time would exceed the available time. Therefore applying the multiple stresses simultaneously on the component is more realistic for getting failure data in short time and may be considered a form of Accelerated Life Testing (ALT). ALT determines the life distribution of component on the basis of failure data and is helpful for predicting the reliability of the components/items.

This paper provides the Bayes and quasi-Bayes estimation of reliability of single component under stress-strength of Burr distribution. We consider ‘c’ and ‘k’ have independent inverted gamma prior with parameters \((a1, b1)\) and \((a2, b2)\) respectively. The Bayes and quasi-Bayes estimators of reliability are compared on the basis of relative efficiency and absolute bias of the said estimators. Numerical study has been done for the various combinations of Burr and prior parameters by generating 200 Monte-Carlo samples.

navinbhu@yahoo.co.in

RANDOMIZED PLAY THE WINNER RULE FOR ORDERED CATEGORICAL DATA – A FOLLOW-UP MODEL IN CLINICAL TRIALS

Aditya Chatterjee
Department of Statistics, The University of Burdwan, Burdwan 713104, West Bengal, India

Gopaldeb Chattopadhyay (Bureau of Applied Economics and Statistics, Government of West Bengal, India)

Comparison of a new drug with a standard or placebo drug through clinical trials is a very common problem of interest in the pharmaceutical industries. In many such trials the treatment responses are measured on an ordinal scale rather than on a continuous scale. A number of authors over the past decades have considered techniques for analyzing such type of ordered categorical data. A simple scoring system called ridits (relative to an identified distribution) that was first introduced by Bross (1958) may be used towards analyzing such data where cumulative probability scores instead of arbitrarily selected scores are considered. Brockett and Levine (1977) noticed that the ridit scores, estimated from the data, have the property that if we combine two adjacent categories and redefine the scores by the same
method, then the scores for the remaining categories remain unchanged. Ridit analysis has been successfully applied to the study of automobile accident (Bross (1960)), cancer (Wynder, Bross, Hirayama (1960)), schizophrenia (Spitzer et. al. (1965)), preference studies (Poulard et. al.(1997)).

The technique of data-dependent allocation of treatments to the patients are of paramount interest with regard to clinical trials. For example, if the subjects enter into a system sequentially, the problem of allocation of treatments among the entering subjects requires thorough scrutiny. Further, as the subjects are human beings, from ethical point of view, it is desirable to carry out a test procedure with smaller number of patients being treated by the inferior treatment in course of the decision-making. With this idea in mind Zelen (1969) introduced the concept of play-the-winner rule for dichotomous treatment responses. Later Wei and Durham (1978) and Wei (1979) modified this idea and introduced randomized-play-the-winner rule. One of the major requirements in such sequential trials is that the outcomes are known relatively quickly and the treatment responses are dichotomous. In fact the method provided by Wei (1988), holds only if the treatment responses are dichotomous and instantaneous. Following Wei, Chattopadhyay (2002) proposed a test procedure for more than two treatment response categories. However, that procedure is not suitable when the treatment response of all previously treated patients are not readily available with the clinician before treating a particular patient, i.e. when the treatment responses are not instantaneous. In practice, treatment responses are not always instantaneous and often it is required to follow up the patients after certain time period since administering the drug.

In the present paper our aim is to provide a suitable test procedure for comparing two treatments (say treatment A and treatment B) when the treatment responses are ordered categorical in nature and each patient is followed up after certain time period (say, D days) from the date of administering the treatment. On an average this rule also allows more patients to be treated by better treatment in course of decision making, thus preserving the ethical issue of clinical trial. At the same time the treatment responses are not required to be instantaneous. Various small sample and asymptotic empirical results of the test have been derived. Moreover power and ASN studies have been done by simulation to establish the claims.

ac_bustat1@yahoo.co.in

A CLASS OF SEQUENTIAL PROCEDURES WITH APPLICATIONS TO ESTIMATION AND RANKING
AND SELECTION PROBLEMS

Ajit Chaturvedi
Department of Statistics, Delhi University, Delhi

with

Rahul Gupta (Department of Statistics, University of Jammu, Jammu)

The confidence region (interval) estimation, point estimation (bounded risk, as well as, minimum risk) and ranking and selection problems are linked with suitable loss functions. Taking into consideration the relationship among 'optimal' fixed sample size solutions of various estimation and ranking and selection problems, a class of sequential procedures is developed. Under some moderate conditions on the risk
function associated with this class of sequential procedures, second-order approximations are obtained for the 'regret'. By means of examples, it is illustrated that the results obtained under this general set-up provide solutions to variety of problems.

UNBIASED ESTIMATION OF THE MSE MATRICES OF IMPROVED ESTIMATORS IN LINEAR REGRESSION

Anoop Chaturvedi
Department of Statistics, University of Allahabad, Allahabad

with

Alan T. K. Wan (Department of Management Sciences, City University of Hong Kong, Hong Kong) and
Guohua Zou (Institute of System Sciences, Chinese Academy of Sciences, China)

Stein-rule and other improved estimators have scarcely been used in empirical work. One major reason is that it is not easy to obtain precision measures for these estimators. In this paper, we derive unbiased estimators for both the mean squared error (MSE) and the scaled MSE matrices of a class of Stein-type estimators. Our derivation provides the basis for measuring the estimators' precision and constructing confidence bands. Comparisons are made between these MSE estimators and the least squares covariance estimator. For illustration, the methodology is applied to data on energy consumption.

anoopc@sancharnet.in

A NONPARAMETRIC REGRESSION SMOOTHER FOR NONNEGATIVE DATA AND ITS APPLICATION IN FINITE POPULATION SAMPLING

Yogendra P. Chaubey
Concordia University

Nonparametric regression using kernel method, local linear smoothing, splines and so on, has been used recently for estimation of mean and other parameters in finite population sampling by adapting the generalized regression estimator (see Särndal, Swensson and Wretman (1992). Model Assisted Survey Sampling, Springer-Verlag, New York.) as in Chaubey and Crisalli (2001) (Technical Report, Department of Mathematics and Statistics, Concordia University, 2001) and Breidt and Opsomer, (2000), (Annals of Statistics, 28, 1026-1053). These methods are general purpose methods and there may be scope for better methods in specific situations. In this paper we employ recently developed estimator for the regression function for non-negative data in Chaubey, Sen and Zhou (2002) in proposing non-parametric regression estimator for estimation of mean and other parameters from a finite population for non-
AN ANALYSIS OF CLUSTERED CATEGORICAL DATA - APPLICATION IN DENTAL HEALTH
Kalyan Das
Department of Statistics, Calcutta University, India
with
Asis Kr Chattopadhyay

Quite often in medical studies multiple discrete indicators are used to measure some characters that are defined only conceptually and are difficult to measure directly. Studies of this type exhibit categorical responses of dependent nature. Analysis of such categorical data appears to be extremely difficult (intractable) particularly in the presence of risk (causal) factors. In the present article, our purpose is to develop a latent mixture regression model for analyzing such multivariate categorical data. Such a mixture model accommodates correlated and over-dispersed data through the incorporation of random effects. Unfortunately, a full likelihood analysis is often hampered by the need for numerical integration. Two different procedures have been considered here. Both involve intensive computations. Numerical investigation has been carried out on the basis of a survey data covering 220 individuals from medical colleges in and around Calcutta (India). The purpose of the study is to compare tooth cleaning efficiency of brushes manufactured by different companies.

SLOPE ROTATABILITY WITH CORRELATED ERRORS
Rabindra Nath Das
Department of Statistics, Burdwan University, Burdwan, West Bengal, India

In Das (Cal. Statist. Assoc. Bull. 47, 1997, 199-214) a study of second order rotatable designs with correlated errors was initiated. Robust second order rotatable designs under autocorrelated structures was developed. In this paper, general conditions for second order slope rotatability have been derived assuming errors have a general correlated error structure. Further, these conditions have been simplified under the intra-class structure of errors and verified with uncorrelated case.

ESTIMATION OF STAGE OCCUPATION PROBABILITIES IN MULTISTAGE SYSTEMS UNDER CURRENT STATUS DATA
Somnath Datta
University of Georgia

We consider nonparametric estimation of stage occupation probabilities for a multistage system that is not necessarily Markovian. The underlying data consist of the stages occupied by the individuals in the
system each observed at a random time (current status data). We show that a kernel smoothing technique combined with weighted estimation can be used to obtain valid estimators of stage occupation probabilities.

datta@stat.uga.edu

COMPARISONS OF CLUSTERING ALGORITHMS FOR GROUPING GENES BASED OF EXPRESSION PROFILES

Susmita Datta
Georgia State University

With the advent of microarray chip technology, large data sets are now emerging containing the simultaneous expression levels of thousands of genes at various time points during a biological process. Biologists are attempting to group genes based on the temporal pattern of their expression levels. In this talk we have selected a number of clustering algorithms (of various flavors!) which can be used to group genes based on their expression profiles. We study their performance on a well known publicly available microarray dataset on sporulation of budding yeast, as well as a simulated dataset. Among other things, we formulate three reasonable validation strategies that can be used for any microarray data when temporal observations or replications are present. We evaluate each of these clustering methods with these validation measures to see, which is a good choice for a given dataset.

sdatta@mathstat.gsu.edu

DESIGNING OF SUSPENSION SYSTEM WITH SPECIAL TYPE DOUBLE SAMPLING PLAN

O. S. Deepa
Bharathiar University, Tamil Nadu, India

with

K. K. Suresh (Barathiar University, Tamil Nadu, India)

This paper is concerned with acceptance sampling systems when a small sample size are necessary or desirable. For example, when production quantities are small or when inspection is either costly or destructive. Under these conditions, a sampling plan with small sample size is not very effective, since discrimination between good and bad quality is not sufficient. Nor does the lot-by-lot inspection provide an incentive for the producer to turn out consistently good quality. Hence it is intended to adopt one plan suspension system with Special Type of Double Sampling Plan (STDS) as the reference plan.

When a single plan is used with a suspension rule, the system is called a One Plan (OP) suspension system. In OP suspension systems, a lot- by- lot sampling plan is used in the usual way to decide whether individual lots shall be accepted or rejected. The Reference Quality Level (ReQL) is the quality level which represents standards of quality, taking in to consideration production demands, specifications, costs or other significant factors. Tables are provided for easy selection of the plan. Illustrations are also provided for ready made selection of parameters for one plan suspension system.
COMPUTING THE 3-D STRUCTURE OF PROTEIN MOLECULE

Narsingh Deo
Department of Computer Science, University of Central Florida,
Orlando, Florida, U.S.A.

Determining the three-dimensional structure of a protein molecule (or any other large biomolecule) is an extremely important and computationally challenging problem. One approach involves using NMR (nuclear magnetic resonance) spectroscopy, through which a small subset of the n(n-1)/2 inter-atomic distances in an n-atom- molecule are determined. Furthermore, these distance measurements are not exact, but each lies within an upper and a lower bound. Computing the coordinates of the n points (atoms) in the 3-D Euclidean space which conforms to the NMR distance measurements is referred to as the molecular conformation problem. We use an edge-weighted, complete graph of order n, embedded in the 3-D space to model the n-atom molecule. First, the triangle inequality is applied to tighten the distance bounds. Warshall-Floyd all-pairs shortest path algorithm is used to ensure convergence and an optimal order of processing the triangles. The distance-bounds are further tightened by applying the so-called tetrangle inequality (resulting from the Cayley-Menger determinants) to all quadruples of atoms. Parallel algorithms are designed in order to keep execution times within reasonable limits. Only edge-disjoint sets of 4-node subgraphs can be processed concurrently. Finding such edge-disjoint subgraphs is equivalent to constructing a large set of 2-(n, 4, 1) packing designs. A number of other interesting problems involving design theory as well as graph theory arise in this context. After the distance bounds are tightened, a well-known result from distance-geometry is used to determine the coordinates in 3-D space (from the three largest eigenvectors of the metric matrix).


deo@cs.ucf.edu

STATISTICS: A NECESSARY EVIL

Nirmal Devi
Embry Riddle Aeronautical University, Daytona Beach, FL 32114, U.S.A.

In our time, there is no escape from figures. We simply cannot get away with the excuse that we are not interested in Statistics. The two words Statistics and Life have come to be pretty interwined. Through Statistics, modern man evaluates business activity, records social progress, elects Presidents and keeps abreast of their popularity, measures interests and aptitudes, determines which television shows will survive and which will not, decides whether to invest in stocks or not, keeps track of sports, compares the effect of new medicines over the existing ones and the list does not seem to end.

Although the quality of some statistical evidence is unimpeachable, but I am sorry to say that most of the
quoted Statistics is misleading to the common person. In this talk I will pose a diagnostic quiz. This quiz will provide a chance to see what kind of statistical eye is required to save ourselves from the evil.

devin@erau.edu

SCHOLL TRANSPORT SYSTEM IN DELHI: AN ENVIRONMENT AND ENERGY EFFICIENT APPROACH

K. K. Dewan
Department of Mathematics, Jamia Milia Islamia, New Delhi-25

with

Punita Saxena (Jamia Millia Islamia, New Delhi) and M. Mustafa (Department of Commerce and Business Studies, Jamia Millia Islamia, New Delhi)

In this paper we highlight the need for a policy approach of the Government sector to provide a uniform frame work of the facilities in the education system to every child. The study lays an emphasis on the fact that such a frame work would lead to a reduction in the vehicular density on the roads and a fall in the pollution level. It outlines a strategy to achieve an environment friendly and fuel saving transport system. The thrust of the study is to conserve energy by restructuring the education system.

fsajmi@rediff.com

DESCRIPTING OF BREASTFEEDING PRACTICES AMONG POORER SECTIONS IN CALCUTTA METROPOLITAN AREA AND ITS IMPACT ON POSTPARTUM INFECUNDITY

S. Dihidar
Indian Statistical Institute, Kolkata, India

with

K. L. Mukherjee (CMDA, Kolkata, India)

Description of breastfeeding practices among residents in slum areas of Kolkata and the influence of breast feeding on postpartum amenorrhoea is the focus of this study. Three out of four women initiate breastfeeding after one hour of birth and three out of five squeeze the first milk from the breast before breastfeeding. Though the median duration of breastfeeding is long, the duration of exclusive breastfeeding is much shorter. Breastfeeding duration the first six months has a significant negative influence on the rate of return to menses.

dshankar@isical.ac.in

CALIBRATION ESTIMATORS: A REVIEW

Madhulika Dube
Department of Statistics, M. D. University, Rohtak 124 001, India
A number of estimators, alternative to the classical and inverse calibration estimators, have been proposed in literature, each being designed to overcome certain inherent difficulties associated with the standard procedures. In this paper the objective is to review and examine various approaches to univariate statistical calibration. An attempt has also been made to study the impact of nonnormality of errors on various calibration estimators.

mdube@sancharnet.in

**SOME REGIONAL MULTI-LEVEL MODELS FOR WANTED FERTILITY IN INDIA**

Laxmi Kant Dwivedi  
International Institute for Population Sciences, Mumbai-88  
with  
F. Ram

The consistent increase in population is known problem specifically in developing countries like India. To understand the dynamics of increasing fertility that is an important component towards the population problem, understanding determinants of wanted fertility might be essential that is an important component of population changes. With this objective in mind, the present study deals with analysis of data related to factors associated with intention to have or not to have the last child/ current pregnancy to the currently married women (15-49 years). For this, the data related to most populous Indian state, Uttar Pradesh, collected under second round of the National Family Health Survey (NFHS), 1998-99, have been utilized.

The outcome, that is, dependent variable, considered under the present study is the woman’s intention to produce the last child or to carry the current pregnancy. Based on a series of exploratory analysis, to have meaningful results through sufficient frequency distribution in stratified categories, the covariates included in the data analysis are:

Individual level: Woman’s place of residence (rural/urban), religion with caste (schedule caste/schedule tribe Hindus/other hindu (non hindu), education (illiterate/literate), Husband’s education (illiterate/literate), Woman’s occupation (not working / working), Age (less than 25/greare than or equalto25), Child loss (none/one and more), Number of living sons (less than three/three and more), Discussion about family planning with partner (no/yes), Exposure to mass media (yes/no).

Primary Sampling Unit (PSU) level: Road within or outside the village/ town, Health facility within or outside the village/ town and primary school within or outside the village/ town.

In view of dichotomous dependent variable (No/Yes), the choice of traditional logistic regression analysis is obvious. Under traditional analysis we carry out individual level analysis considering disaggregation of higher level variables at individual level or aggregation of individual level variables at higher level. Under this process either we distort the important assumption of statistical independence that is essential for traditional methodology or give away the individual level variance that may be even up to 90% of the total.
optimized procedure known as Multi-level analysis has been used in the data analysis. For this, individual level variables were considered at level-I and PSU level variables at level-II. Accordingly two-level logistic regression analysis was considered in the analysis. Since this procedure considers the variables at their own level, this helps in retaining the obvious hierarchical structure of the data in the analysis. Accordingly, this procedure is expected to provide more reliable results. Further, this procedure also accounts the variability because of variables, which could not either be considered in the data analysis or could not be collected. Further, this procedure also helps in partitioning the variance of a covariate at a particular level at different levels. The residual analysis under this procedure further helps in working out the priority of intervention programme giving important clues to the public health policy planners. On account of these facts, this procedure has been used not only to analyse the data at state level but also at regional level. For this, sub-group analyses in each of the five regions of Uttar Pradesh were carried out. The importance of regional analysis is very well recognized in the existing literature that may help in planning of regional interventions that may vary from one region to another region. On account of availability of large-scale quality data available under NFHS and also availability of computational facility through specific statistical packages like MLwin, it was possible to carry out the required analysis under the present study.

The results in the present study indicate in general that the contribution of a particular covariate is varying from one region to other region. On account of observed variance of unconsidered covariates at level-II, this is also evident that there is need to consider either some more appropriate variables at level 2 or some more variables at higher level like at district level. This may be possible even through consideration of aggregation of some lower level variables at higher levels. The results obtained under the present analysis will be presented and discussed in detail.

laxmikdwivedi@rediffmail.com

CHAOTIC TIME SERIES ANALYSIS OF A SEA SURFACE TEMPERATURE TIME SERIES

Suneet Dwivedi
M. N. Saha Centre of Space Studies, Allahabad University, Allahabad

with

A. K. Mittal and A. C. Pandey

Optimized box assisted algorithm is implemented for computing fractal dimension of NINO 3 (Sea surface temperature anomaly) and IMR (Rainfall anomaly) time series data. Lyapunov exponents of these time series are also calculated.

dwivedisuneet@rediffmail.com
ESTIMATING THE DIMENSION OF TIME AVERAGED PARADIGMATIC MODEL

Suneet Dwivedi
M. N. Saha Centre of Space Studies, Allahabad University, Allahabad

with

Avinash Chandra Pandey and A. K. Mittal

Time averaged behaviour of Forced Lorenz Model is studied. The variation in the dimension of Paradigmatic Lorenz Model with forcing and with time averaging is calculated.

dwivedisuneet@rediffmail.com

PUBLICATIONS ON TUBERCULOSIS IN DEVELOPING AND DEVELOPED COUNTRIES: A QUANTITATIVE ANALYSIS

R. Elangovan
Department of Statistics, Annamalai University, India

with

M. Shanmugham

Globally, there is rapid growth of publications on Tuberculosis in the past four decades. These literatures are addressing and helping the TB health personnel to respond to the present and future challenges of TB. This study aims to determine the pattern and trend of TB publications in the last four decades by bibliometric analysis. The MEDLINE Database was used to retrieve articles on TUBERCULOSIS in the MESH fields (i.e. directly related to the articles) for the period 1966 to 2001/06 through SPIRS search Engine. A specialized software package was developed by the author using Fox Pro (ver 2.5). Various fields like Author, Country of Publications, Language, Publication Type, Year, MEdical Subject Headings (MESH) etc., were searched and stored in a text file. All the Text file contents were then merged in to a single database file, using this detailed analysis was done. Total of 72,390 articles in TUBERCULOSIS were processed. The number of TB articles available in the year 1963 was 16 only. This has abruptly raised to 1648 in the year 1965. The analysis shows that from 1966-1974 the trend has been oscillating between 2336 to 2912. In 1975, this has decreased to 1940 articles and then after it steadily decreases until the year 1994. Since then articles steadily increases till the year 2001. This has also been reflected on the publications of the developing countries. Based on the individual Journal names there were 72,390 TB articles published in 3669 Journals. These journals are from 94 countries in 38 languages. By language categorization maximum articles were published in English (35,417 articles i.e. 49%), English-French (52), English-Italian (4), English-Portuguese and English-Spanish. The remaining were in Afrikaans, Bulgarian, Czech, Chinese, Finnish, French, Germany etc. in 36 Non English Journals. Most of these are from developed countries. Considering the Journal factor, among 3,669 journals, the highest number of articles were from Probl-Tuberk (USSR,7614 articles i.e. 10.52%) followed by Kekkaku (Japan, 1914, i.e. 2.61%), both of which are Non English Journals and then by Am-Rev-Respir-Dis published in English. 72,390 articles were by published 94 countries. Of which developed countries like United States (15737), England (6372), Japan (4241), Italy (2349), Poland (2311), Germany West(1341), Germany East (1182),
had highest publications from Tuberculosis. In the developing countries highest articles was from India (1975 articles). 61.21% of the TB articles were published by Journals from developed countries. The 39.61% of the Tuberculosis articles of the developing country published in developed country journals. Strangely only 21% of the developing country articles were published in the developing country journals. The total of 27,858 (38.48%) articles are published in the same country publications Journals. The study found that publications on Tuberculosis were steady some period and oscillating in some period though there is a significantly less contributions on developing country articles even from journals published from developing country. Much work is needed to increase the publications from these countries as they bear a great burden of TB cases globally. This paper presentation will discuss in detail about developing countries and developed countries to year-to-year, language, medical subject heading, publication type, authors address, country of publications, language factor and all other factors to identify the trend in these publications.

srelangovan@rediffmail.com

THE VARIABLE SELECTION PROBLEM IN MULTIVARIATE ANALYSIS - MODEL SELECTION APPROACH

Yasunori Fujikoshi
Department of Mathematics,
Graduate School of Science,
Hiroshima University

It is seen that many problems for selection of variable in multivariate analysis can be formulated as one of model selection problems. In this formulation we need to define a variable selection model expressing that a set of variables is sufficient and the set of remainder variables is redundant. The idea of no additional information due to Rao (1948, 1973) plays an important part in the matter. Then we can use some model selection criteria such as AIC, Cp, etc. which select an appropriate model from a family of variable selection models.

In this paper we discuss about recent developments and perspectives on the above approach for selection of variable. The selection criteria have been proposed as approximately unbiased estimators for its risk function, typically the expected log-predictive likelihood or the expected mean squared error of prediction. Our main concern is to explore modified criteria which are intended to reduce bias, focusing on a general setting between the true model and a candidate model, a distributional assumption, and a framework of asymptotic approximations.

fuji@math.sci.hiroshima-u.ac.jp
Logistic regression methods have been found useful in the estimation of ‘odds ratios’ under matched pair case-control design when the exposure variable of interest is binary or polytomous in nature. Usually, the analysis is performed using large sample approximation techniques. While carrying out the analysis with polytomous exposure variable, sometimes we encounter the situation where the numbers of discordant pairs in the resulting cells are small or the data structure sparse. In such situations, the asymptotic method of analysis is in question and hence an exact method of analysis may be more suitable. In this paper we discuss a method that uses the distribution of the conditional sufficient statistics of the logistic model parameters to perform exact inference in case of a pair-wise matched case-control data with more than two unordered exposure categories. We describe a computational method which can be used for obtaining the combinatorial coefficients involved in the distribution.

ganguly@squ.edu.om

BAYESIAN ESTIMATION OF RELIABILITY VIA LENGTH BIASED EXPONENTIAL DISTRIBUTION

Jyoti Kumar Gaur
Dept. of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur

R. S. Srivastava (DDU Gorakhpur University, Gorakhpur) and
Vijay Kumar (DDU Gorakhpur University, Gorakhpur)

In this article, an attempt has been made to obtain Bayes estimators of the reliability function of exponential lifetime distribution via length biased exponential distribution under different priors using symmetric as well as various asymmetric loss functions. The performance of the estimators has also been studied with Monte-Carlo samples.

gaur_jk@yahoo.com

STATISTICS IN MATERIAL SCIENCE AND TECHNOLOGY

Hina Amol Gokhale
Defence Metallurgical Research Laboratory, Hyderabad, India

Statistical approaches have been used for many years in various fields concerning material science and technology. The techniques applied include from simple t-test to study the homogeneity of the materials
to complex design of experiments to learn the effect of numerous factors in coating or welding experiments, and from techniques of life-time data analysis to study mechanical behaviour of materials to fractals and stochastic geometry to study microstructures of materials. And yet there are other fields still waiting to be explored statistically. The present paper provides an overview of these and other statistical applications in material science and technology, along with new possibilities for applications.

hina_gokhale@rediffmail.com

MODIFIED RATIO ESTIMATORS FOR POPULATION MEAN

K. C. Goyal
Directorate of Human Resource Management, CCS HAU, Hisar

with
Veena Manocha

In this paper, we propose modified ratio estimators of the type considered by Tin (1965) and study their properties. An attempt is made to make their efficiency comparisons with the conventional ratio estimator and Tin’s estimator.

dpme@hau.nic.in

RELATIVE EFFICIENCY OF PRODUCT ESTIMATION STRATEGIES UNDER SUPER POPULATION MODELS

Manoj Goyal
Directorate of Human Resource Management,
CCS Haryana Agricultural University,
Hisar-125004, India

with
K. C. Goyal and Veena Manocha

Singh and Singh (1997) investigated the relative efficiency of mean per unit estimator with conventional product estimator and the modified product estimators proposed by Robson (1957) and Dubey (1993) under Durbin’s super population model (1959). In this paper, it was observed that the estimators considered by them are inconsistent and that the product estimator is not efficient even than that of the mean per unit estimator whereas in sampling theory it is reported that the product estimator is always efficient than the mean per unit estimator if $-2r > \frac{C_X}{C_Y}$, the ratio of coefficient of variation in variables X and Y. It is therefore concluded that the super population model considered by them is unsuitable to study the relative efficiency of product strategies. In the present article, an alternative super population model is proposed and the relative efficiency of product estimator with mean per unit and the regression estimator have been found. The situations when the product estimator is better than the
regression estimator have been identified whereas the Srivastava (1980) found that the regression estimator can not be improved upon unto first order of approximation.

manoj@hau.nic.in

EFFICIENT FAMILIES OF CHAIN ESTIMATOR FOR RATIO AND PRODUCT OF TWO POPULATION MEANS

Lovleen Kumar Grover
G. N. D. University, Amritsar, India

with

H. S. Jhajj and M. K. Sharma

Here we proposed Efficient Families of Chain Estimator for Ratio and Product of two population Means using known information on one of the variables of the given two auxiliary variables. The bias and the mean square error of proposed families are upto first order approximation. The optimum mean square errors are also obtained. The gain in efficiency is illustrated empirically.

MATHEMATICAL MODELLING IN METALS PROCESSING

G. S. Gupta
Department of Metallurgy, Indian Institute of Science, Bangalore - 560 012

An overview of mathematical modelling in metallurgy is presented. The role of mathematical modelling in metallurgy is described by some illustrative examples such as Induction Smelting Process, Casting and carburising. Some metallurgical processes are cited where mathematical models are still in their infancy due to lack of knowledge of physical phenomena such as spray forming, liquid flow through porous media. Finally, it is stressed that mathematical models must be reachable to industries and common public in user friendly way which is illustrated citing few examples based on Graphical User Interface, statistical analysis and, analytical solutions.

govind@metalrg.iisc.ernet.in

ESTIMATORS FOR TWO PARAMETER EXPONENTIATED WEIBULL FAMILY UNDER TYPE-II CENSORING SCHEME

Pramod K. Gupta
Department of Statistics, Banaras Hindu University, Varanasi 221 005, India

with

Umesh Singh and S. K. Upadhyay (Department of Statistics, Banaras Hindu University, Varanasi 221 005, India)

Bayes and classical estimators for two parameter exponentiated Weibull distribution have been obtained in this paper when type-II censored sample is available to us. Bayes estimators have been developed under
squared error loss function as well as under LINEX loss function using non-informative type of priors for parameter. Besides the generalized maximum likelihood estimators and the usual maximum likelihood estimators have also been attempted. It has been seen that the estimators obtained are not available in nice closed forms, although they can be easily calculated for the given sample by using suitable numerical methods. The performance of the proposed estimators have been compared on the basis of their simulated risk (average loss over sample space) obtained under squared error as well as LINEX loss functions. These estimators are illustrated for a simulated data set.

pk_guptain@yahoo.com

THE COVARIANCE STRUCTURE OF THE MULTIVARIATE LIOUVILLE DISTRIBUTIONS

Rameshwar D. Gupta
Department of Computer Science & Applied Statistics,
University of New Brunswick

Donald St. P. Richards (Department of Statistics, Penn State University)

We study the covariance structure of the multivariate Liouville distributions and their power-scale transformations. For any Liouville distributions, we show that the off-diagonal entries of the corresponding correlation matrix all have the same sign; and we show that this result holds also for power-scale transformations of the Dirichlet distributions. In the case of matrix Dirichlet distributions, we derive formulas for the all moments of degree one and two, leading to the discovery of some remarkable relationships among the underlying covariances; in particular, we deduce that the resulting covariance matrices are of block-diagonal form. In the case of the matrix Liouville distributions we also drive formulas for all moments of degree one and two, thereby deducing completely their covariance structure.

gupta@unbsj.ca

CIRCUMVENTING SOCIAL DESIRABILITY RESPONSE BIAS USING RANDOMIZED RESPONSE TECHNIQUES

Sat Gupta
Department of Math/Stats & Psychology University of Southern Maine

Bill Thornton

Randomized response techniques are important tools in circumventing response bias in personal interview surveys involving sensitive questions. In this paper we discuss the use of several of these techniques in a campus survey involving students enrolled in psychology and statistics courses. The purpose of the survey was to examine a variety of social issues such as smoking, drug use, alcohol use etc. among students. The results are compared with bench-mark results based on an anonymous/confidential survey, and also
with results based on the Bogus Pipeline (BPL) technique which has been used by psychologists to elicit more accurate responses to personal questions of sensitive nature. BPL technique requires attaching electrodes to the fingers of the respondent’s non-preferred hand and wiring to an automated data acquisition computer system. Through a series of test questions the respondent is made to believe that the system can detect the truthfulness of a person’s response. Preliminary results indicate that some of the randomized response techniques are at least as effective as the BPL technique while being less intrusive.

sgupta@usm.maine.edu

ENERGY REQUIREMENT IN AGRICULTURAL SECTOR

V. K. Gupta
IASRI, Library Avenue, New Delhi - 110 012, India

with
Rajender Parsad (IASRI, New Delhi - 110 012)

The data on various aspects of energy usage in agricultural production system is being collected from the farmers of the selected villages in different agro-climatic zones under the All India Co-ordinated Research Project on Energy Requirement in Agricultural Sector. The information is collected on uses of Human Labour, Animal labour, Diesel, Electricity, Seed Rate, Farmyard Manure (FYM), Fertilizer, Chemicals, Machinery, Canal, etc. These are then converted into Mega Joule / hectare (MJ/ha) using internationally accepted conversion factors. The energy uses are also available on agricultural operations like tillage, sowing, bund making, fertilizer application, etc. Adding the energy levels from different sources generates the total energy used for crop production that forms another factor in the study. The data available on yields are converted into per hectare basis. As of now, the data is available on yield (kg/ha or MJ/ha), energy used (MJ/ha) from various sources and total energy used (MJ/ha).

The main objectives of the project are:
Ø To establish the relationship between yield and total energy; yield and other sources of energy like human labour, animal labour, diesel, electricity, FYM, fertilizers, chemical, machinery, irrigation, canal, etc.

Ø To find out the optimum values of the various energy sources for maximum productivity.

To meet these first objective, first order and second order response surfaces are fitted. A pertinent question that arises here is as to whether a single regression equation (or response surface) will adequately describe the relationship for all categories of farmers under consideration or will different regressions be required for each category of farmers? A complete description of the response (the best fit of data) would be obtained by allowing each category to have its own regression equation (or response surface). This would be inefficient, however, if the responses were similar over all categories; the researcher would be estimating more parameters than necessary. On the other hand, a single regression equation (or response surface) to represent the response for all categories will not adequately characterize any one group and could be very misleading if the relationships differed among categories.
Therefore, separate response surfaces for each category of farms was fitted and homogeneity of regression equations (or response surfaces) was also tested. Wherever, the regression equations were homogeneous, a common regression equation was fitted to the entire data set. Otherwise the analysis was carried out separately for each category. Categorization of the farmers on the basis of irrigated or rainfed, electricity use or non-use, bullock or tractor use, based on productivity levels like low (2000 Kg/ha), medium (2000 · 3250 Kg/ha), high (3250 Kg/ha), etc. or based on the ratio of total energy to yield (energy-yield ratio) like high (< 3.50), medium (3.50 · 4.00), low (4.00 · 5.00), very low 5.00 was also suggested.

To obtain the optimum energy levels for different sources like human energy, animal energy, diesel energy, electrical energy, FYM energy, fertilizer energy, machinery, irrigation, etc. to maximize the yield, a second order response surface was fitted. The co-ordinates of the stationary point were obtained by equating the first derivative of the fitted second order response surface equal to zero. Canonical analysis was performed to find the nature of the stationary point (point of maxima, minima or a saddle point). For the cases, for which the stationary point is a saddle point and lies within the input range, the response surface in the vicinity of the stationary point was explored. This exploration gives various combinations of input variables for a desired output in the vicinity of the predicted response at the stationary point. One can choose the input combination based on the practical considerations. However, using several sets of data, it has been observed that most of the time the regression coefficients are not significantly different from zero, particularly the second-degree coefficients; and/or saddle point lies outside the input range. It seems that the energy usage has not yet reached the saturation stage or plateau. In other words, the relationship of yield with energy levels of various factors appears to be linear in nature. Therefore, to obtain the levels of various inputs that maximize the yield per hectare, recourse is to be made to the use of Linear Programming (LP). In LP problem, the objective function and the constraints are very important. Therefore, one has to be cautious in defining the objective function and constraints. In the initial stages, it was thought that one should fit a multiple linear regression, and use the fitted multiple linear regression equation as an objective function and availability of the energy from different sources like human, animal, diesel, electricity, machinery, etc. as constraints. However, a close scrutiny reveals that such an objective function may be error prone like it may have large standard error of the estimated response, the regression coefficients may also have large standard errors, and moreover, many of the regression coefficients may not be significantly different from zero. Therefore, the use of such an objective function is to be avoided. The second option of the objective function is that we consider the data of energy usage and productivity of each farmer as a separate activity and define the objective function and constraints. The approach uses the maximization of yield subject to the constraints on the availability of energy from different sources like Human Labour, Animal Labour, Diesel, Electricity, Seed Rate, Farmyard Manure (FYM), Fertilizer, Chemicals, Machinery, Total Energy, etc. The procedure has also been used for minimization of total energy for obtaining a given level of yield. The concept of energy use efficiency has also been introduced. This technique is being exploited by the All India Co-ordinated Research Project on Energy Requirement in Agricultural Sector, Central Institute of Agricultural Engineering, Bhopal.

vk@iasri.delhi.nic.in
AN ASSOCIATION MODEL FOR AGRICULTURE INCOME IN EU COUNTRIES

A. Haritou
with
A. Vogiatzis

We consider and estimate the most accuracy association model for the agriculture income in EU countries from 1983-2001 instead of others econometric models.

haritou@macedonia.uom.gr

BAYESIAN ANALYSIS OF LORENZ CURVE FROM GROUPED DATA

Hikaru Hasegawa
Hokkaido University

In this article, we estimate Lorenz curve and Gini index from grouped data using the recent development of the Bayesian mixture model. Further, we present an example using real data from the National Survey of Family Income and Expenditure in Japan to illustrate our approach.

hasegawa@econ.hokudai.ac.jp

SUITABILITY OF PBIB DESIGNS FOR OBTAINING PDC PLANS

R. C. Hasija
Department of Mathematics and Statistics,
CCS Haryana Agricultural University, India
with
L. S. Kaushik

The suitability of 2 associate PBIB designs catalogued by Clatworthy (1973) and 3 associate PBIB designs has been examined for obtaining various designs on Partial Diallel Crosses (PDC). An efficiency criteria based on eigen value of C-matrix has been developed to compare the efficiencies of different designs obtainable on Partial Diallel Crosses based on PBIB designs.

rch@hau.nic.in

MULTI · CHARACTER SURVEYS USING RANDOMIZED RESPONSE SAMPLING

M. Javed
Department of Math. Stat. & Physics,
Punjab Agricultural University,
In the present investigation, alternative estimators for estimating population totals in multi-character survey sampling using randomized response technique have been suggested when certain variables have poor positive correlation and others have poor negative correlation with selection probabilities. The estimators proposed by Hansen and Hurwitz (1943), Rao (1966) and Sahoo et al. (1994) under scrambled responses are shown as special cases of the proposed estimators. The relative efficiency of the proposed estimators with respect to each other has been studied through empirical study.

sidhusukhjinder@hotmail.com

A FAMILY OF SOME CONTINUOUS DISTRIBUTIONS

J. P. Singh Joorel
Department of Statistics, University of Jammu, Jammu, J&K, 180006

with

Bilal Ahmad (Division of Agriculture Economics and Statistics, SKUAST, Jammu)

In this paper, a new class of continuous distributions is proposed with the advantage that some standard continuous distributions can be derived as its particular cases.

MINIMUM EXPECTED LOSS ESTIMATORS OF GENERALIZED BURR DISTRIBUTION

Narayan Kumar Joshi
DDU Gorakhpur University

with

R. S. Srivastava (DDU Gorakhpur University)

In this paper the minimum expected loss estimators of a shape parameter of generalised Burr distribution under squared error and asymmetric loss functions have been proposed. The asymmetric loss function used is the LINEX loss function. The approach is classical. It has been shown that the maximum likelihood estimator is also the minimum variance bound estimator. Properties of these estimators have been studied with the help of risk functions.

r_shyam_sri@yahoo.co.in

EFFICIENCY OF EXPERIMENTAL DESIGNS IN WHEAT

Kiran Kapoor
Dept. of Agricultural Economics, CCS Haryana Agricultural University, Hisar-125004

with
Efficiency of various designs were compared. It was found that RBD is more efficient in reducing error variation when plots were elongated in N-S direction over CRD. Latin Square Designs were found to be more advantageous over RBD when the columns are considered as blocks, indicating that a Latin Square layout had no advantage over an efficiently designated randomized block layout with same number of plots.

**DISCRIMINATION AND CLASSIFICATION WITH REPEATED MEASURES DATA**

Ravindra Khattree  
Oakland University  
with  
Anuradha Roy (University of Texas, San Antonio)

We study the problem of classification with multiple q-variate repeated measures on each individual. Classification rules for this situation are developed by introducing certain covariance structures. Further, the problem is studies when the population means may also have some structures. Findings are applied and illustrated on real as well as simulated data sets.

khattree@oakland.edu

**TRANSPORT IN GROUNDWATER FLOW: SIMULATIONS AND APPLICATIONS TO ENVIRONMENTAL RISK ASSESSMENT**

Franz Konecny  
Institute of Mathematics and Applied Statistics, University of Agricultural Sciences Vienna

The problem of assessing groundwater pollution has become a matter of considerable concern. For proper groundwater management, it is necessary to model the contamination mathematically in order to assess the effects of contamination and predict the transport of contaminants. Several deterministic models have been proposed and numerical procedures developed. Because of aquifer heterogenity, the spatial variation of flow properties is erratic. Therefore a stochastic model of flow regime and transport processes is more realistic.

In this talk we are concerned with modeling contaminant transport in heterogeneous porous media viewing the velocity as a random field. We apply random walk particle tracking (RWPT) for simulating transport. As a result we obtain the distribution of concentration rather than single estimates as a probabilistic description to be used in risk assessment. The risk of the contaminant transport is defined as the probability that the that the concentration exceeds a maximum acceptable upper limit. Moreover, we analyse the sensitivity of exceedance probabilities with respect to variations of certain physical
parameters of the transmissivity field, porosity and dispersivity.

konecny@mail.boku.ac.at

RANKED SET SAMPLING UNDER RANDOMIZATION FRAMEWORK

Praveen Krishna
Indian Agricultural Statistics Research Institute, Library Avenue,
New Delhi-110012

with

Anil Rai (Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi-110012)

In this article RSS procedure has been examined in randomization frame of work of survey sampling. It has been observed that this procedure falls in the category of equal probability sampling method, i.e. the probability of including every unit of the population in the sample is same. The ranked set sampling (RSS) provides efficient estimation of population mean as compared to simple random sampling (SRS). The procedure of RSS creates artificial stratification of the population; as a consequence of this it provides better representation as compared to SRS. Most of the articles available in literature are either based on infinite population theory or super population framework. In this article the estimator for estimating population mean has been proved to be unbiased and an expression of its variance has been derived in terms of variability of individual ranks. The expressions of finding inclusion probabilities of the individual sampling units and joint probabilities of including any pair of sampling units are also derived. The statistical properties of this estimator have been studied using simulation for two different cases, i.e. (i) usual case when \( N = mn^2 \) and (ii) two phase sampling when \( N \) is the size of target population, \( N > mn^2 \). Here, \( m \) denotes number of cycles and \( n \) is the sample size per cycle. Therefore, sample size of the RSS is \( mn \). For both the cases it was found that the proposed RSS estimator is always better than SRS estimator. Gain of 20 to 40 percent is achieved in RSS of equivalent sample size as compared to SRS.

pkrishna@iasri.delhi.nic.in

MAXIMUM LIKELIHOOD ESTIMATORS OF THE PARAMETER OF EXPONENTIAL DISTRIBUTION UNDER MULTIPLE TYPE-II CENSORING

Anil Kumar
Department of Statistics, B.H.U., Varanasi 221 005, India

with

Umesh Singh (Department of Statistics, B.H.U., Varanasi 221 005, India)

In this paper we address the problem of obtaining the maximum likelihood estimators under multiply type-II censoring scheme for one parameter exponential distribution. It is well known that explicit expression for the maximum likelihood estimator does not exist under this censoring scheme. However, approximate maximum likelihood estimator has been proposed by Balasubramanian and Balakrishnan (1992) using
certain approximations at the stage of solving the normal equations. We proposed to use an approximation in the likelihood function to get an approximate maximum likelihood estimator. In addition this, a numerical method is also proposed to get the maximum likelihood estimate. The estimators thus obtained are compared with the existing approximate maximum likelihood estimator on the basis of their risks and bias using Monte Carlo simulation technique.

kk_anil@rediff.com

ASSESSMENT ISSUES IN WEB-BASED AND WEB-ENHANCED COURSES
Ashok Kumar
Kentucky State University, Frankfort Kentucky 40513, U.S.A.

In this paper, various assessment issues involved with the fully Web-Based and Web-Enhanced courses will be explored including how to keep the integrity of the course. We will also discuss the process involved in offering a Web-Enhanced course.

akumar@gwmail.kysu.edu

PARLIAMENTARY ELECTIONS AND RULING PARTIES: EVIDENCE FROM INDIA
Jayesh Kumar
IGIDR, Mumbai, India
with
Richa Dhariwal

Parliamentary Elections and Ruling Parties: Evidence from India* In an election the probability that a ruling party wins the election and forms the government is affected by the electoral system that is, the rule for aggregating constituency wise results into a single outcome (i.e. forming the central government).

We find the probability that a particular party wins the election from a constituency given that party has won the last election from that particular constituency, under the assumption that all candidates are equally likely (i.e. random selection) given that they are from the same party. The original goal of our work is to estimate the probability of winning by ruling party (i.e. the party which has won the last election from that constituency), irrespective of whether that party forms a coalition during election, formed the last government in center or was the part of the last central government in case of coalition government.

In particular, we are not taking into account the effect of the central government’s ruling, while we are emphasizing on the effect of constituency wise ruling and the results of that particular constituency.

Finally we estimate the average probability of winning for all India’s General elections from 1971 to 1999 under given electoral system, using Binary Choice Model. JEL Classification: C25, N45.
BAYESIAN UNIT ROOT TEST FOR TIME SERIES MODELS WITH STRUCTURAL BREAKS

Jitendra Kumar
Department of Statistics, University of Allahabad, Allahabad

with

Anoop Chaturvedi

The present paper considers the autoregressive models with trend component and augmentation term in the presence of structural breaks in the deterministic trend. The issue of presence of unit root has been explored from a Bayesian perspective. The models with single known break point, single unknown break point and multiple known break points, have been considered. The posterior odds ratios for the unit root hypothesis for these models have been derived under appropriate prior assumptions.

A PROPOSED ARTIFICIAL NEURAL NETWORK CLASSIFIER TO IDENTIFY TUMOR METASTASES:
PART 1

Kuldeep Kumar
Bond University, Gold Coast, Australia

with

Sukanto Bhattacharya (School of Information Technology, Bond University, Australia)

Recursive partitioning techniques are widely used in health data. In this paper we have reviewed various recursive partitioning techniques and propose a classification scheme to isolate truly benign tumors from those that initially start off as benign but subsequently show metastases. A non-parametric artificial neural network methodology has been chosen because of the analytical difficulties associated with extraction of closed-form stochastic-likelihood parameters given the extremely complicated and possibly non-linear behavior of the state variables. This is intended as the first of a three-part research output. In this paper, we have proposed and justified the computational schema. In the second part we shall set up a working model of our schema and pilot-test it with clinical data while in the concluding part we shall give an in-depth analysis of the numerical output and model findings and compare it to existing methods of tumor growth modeling and malignancy prediction.

FORECASTING FINANCIAL DISTRESS AND CREDIT RATINGS USING STATISTICAL AND SOFT-COMPUTING METHODS

Kuldeep Kumar
Bond University, Gold Coast, Australia

In this paper we have discussed and compared various methods for forecasting financial distress and credit ratings. Some of these methods are based on statistical techniques like logistic regression and discriminant analysis. As computers have become more powerful they have permitted better testing and
forecasting the financial distress using soft computing methods like Artificial Neural Network, Chaos theory and fuzzy logic. We have taken numerous financial ratios of a combination of Australian stock exchange and used these techniques to predict financial distress and to determine which companies will be best suited for investment. Similar techniques can be used for predicting bankruptcy amongst Indian companies.

In the liberalized environment, importance of credit rating concept has increased significantly. The objective of the study is also to explore and find out the effect of the financial performance data of a firm on the credit rating of a debt issue of a firm. The study also proposes to capture the relationship, if any, between financial performance data and credit rating given by expert in an appropriate model.

Financial data relevant to a debt issue ratings is obtained from the publications of a premier credit rating agency in India. Data analysis involved building of model using conventional multiple linear dicriminant analysis and Artificial Neural Network Systems.

Findings clearly showed that financial performance data of the company before the issue has significant effect on credit rating by expert. Artificial Neural Networks (ANN) model was found superior to discriminant analysis model. ANN model can be used to increase speed and efficiency of the rating process in practical applications. In addition, if experts provide better-input data, it can be relied upon to provide automatic rating to a significant extent.

kkumar@bond.edu.au

CLUSTERING DATA WITH MEASUREMENT ERRORS

Mahesh Kumar
Massachusetts Institute of Technology

with

Nitin R. Patel (Sloan School of Management, MIT) and James B. Orlin (Sloan School of Management, MIT)

Clustering is a very well studied problem that attempts to group together similar data points from a large set of data. Most traditional clustering work assumes that the data is provided without measurement error. Real world data, however, usually does contain such errors, which in addition, can be estimated using the standard statistical methods. In the presence of such errors, popular clustering methods, like k-means and hierarchical clustering, may produce un-intuitive results. The fundamental question that this talk addresses is: “What is an appropriate clustering method in the presence of measurement errors associated with data?”

We propose using the maximum likelihood principle to obtain an objective criterion for clustering that incorporates information about the measurement errors associated with the data. The objective criterion provides a basis for several clustering algorithms that are generalizations of the popular k-means and Ward’s hierarchical clustering methods. The objective criterion has a scale-invariance property, so that the
clustering results are independent of the measurement units of the data. We also provide a heuristic solution to obtain the correct number of clusters, which in itself is a challenging problem. Finally, we show the effectiveness of our technique on simulated data, where it outperforms the k-means and hierarchical clustering methods.

maheshk@mit.edu

IMPROVING A REGRESSION TREE SOLUTION USING ERROR-BASED CLUSTERING

Mahesh Kumar
Massachusetts Institute of Technology

with

Nitin R. Patel (Sloan School of Management, MIT), Charu C. Aggarwal (IBM T. J. Watson Research Center, NY) and Philip S. Yu (IBM T. J. Watson Research Center, NY)

We consider the problem of clustering the points in a regression. The standard least squares regression, which assumes that the data comes from a single linear function, fails to work when different parts of data have different linear relationships. Regression trees have been proposed to partition the data into subsets (leaf nodes) so that the data within each subset would possess a single linear function. The regression function is then estimated using a least squares regression on each subset. The problem with this method is that, often, each subset of data has only few data points; consequently, the regression coefficient estimate for each subset will have large errors. We propose use of error-based clustering on the regression tree leaves that simultaneously determines the clusters of subsets of data and estimates the regression coefficients for each cluster. The new estimates have smaller errors, and therefore, provide a more accurate forecast on unseen data points. We justify our approach theoretically and present empirical results on both simulated and real data.

maheshk@mit.edu

SELECTION PROCEDURES FOR SCALE PARAMETERS WHEN THE POPULATIONS HAVE A COMMON KNOWN QUANTILE

Narinder Kumar
Department of Statistics, Punjab University, Chandigarh-160 014, India

with

Amar Nath Gill and G. P. Mehta

Consider $k$ independent populations $\mathcal{P}_1, ..., \mathcal{P}_k$ and let $G((x \leq \xi_i) \cap \mathcal{P}_i)$ be the cumulative distribution function (cdf) associated with population $\mathcal{P}_i$ for some unknown absolutely continuous cdf $G(\cdot)$ and unknown location (scale) parameter $\xi_i$ ($\mathcal{P}_i \neq \mathcal{P}_j$, $\mathcal{P}_i \neq \mathcal{P}_k$, $i \neq j$, $i \neq k$, $i, j, k = 1, ..., k$). We assume that these $k$ populations have the same known quantile of order $p$ ($0 < p < 1$), not necessarily equal to $\xi$, i.e., for each $i$, the known quantile $\xi^p = \xi_i + \xi G^{[1]}(p)$, where $G^{[1]}(p) \xi (\mathcal{P}_i, \xi_i, \xi)$ is
EMPIRICAL BAYES LIFE ESTIMATION FOR A MODEL ON SINGLE REPAIRABLE SYSTEM WITH AN INVERSE GAUSSIAN PRIOR

Santosh Kumar
Department of Statistics, Ewing Christian College

Empirical Bayesian analysis of the model arising from Homogenous Poisson process (a reasonable model for the repairable system) based on the life tests that are terminated after a preassigned number of failures and predetermined time is carried out under the assumption of squared error loss function and an inverse Gaussian prior density on the parameter space.

STOCHASTIC ANALYSIS OF A NON-LINEAR MODAL ON THE GROWTH OF TERRORISM

Surender Kumar
Department of Statistics & Operations Research, K.U., Kurukshetra-136119, India

with

Usha Sharma (Department of Statistics & Operations Research and K.U., Kurukshetra-136119, India)

Over the last three decades terrorism has acquired global dimension and received a great deal of international attention in the recent past. The fear expressed by Hamilton & Hamilton in there investigation that terrorists may serve to precipitate the larger disaster of internal or external war or that terrorists themselves may become capable of wrecking havoc on a war like scale, became a reality on 11th September 2001. The world from Manhattan to Manila has been painted red. It is manifesting itself in the various form such as narco-terrorism, bio-terrorism, state terrorism cross border terrorism etc. A larger number of research papers and scores of books have been authored on the topic, however these studies are confined to conceptual qualitative description of the phenomenon. The quantitative expect of the phenomenon has been hardly attempted by the researchers. A class of stochastic models for social contagion has been suggested by Hamilton & Hamilton to apply for the modeling of terrorism. Usha Sharma & Mukesh Chawla have studied a stochastic model of periodically instigated terrorism as a non-homogeneous Poisson process involving periodic variation in the occurrence rate.

A close scrutiny of the phenomenon of terrorism shows that it is quite akin to the cancer growth in a bio-
system. Both are anti-social in the sense that they destroy the very system in which they grow. A cancer is initiated by some carcinogens while the terrorism originates from sporadic violence, exploitation or false historical pride. Nature provides immune surveillance to a bio-system while civil societies build up their own security mechanism. Taking a clue from this remarkable similarity, in this piece of work, we have extended the bio-mathematical concepts used in the study of tumors and cancers, to the phenomenon of terrorism. We have proposed and examined a non-linear stochastic model for the growth of terrorism. Using linearization and random variable techniques, we have obtain generating function for the terrorism process represented by a random variable $n(t)$, denoting the number of terrorists at time $t$. We have also obtained expression for the probability of ultimate control of the terrorism and in passing have noted that even in an almost crime and exploitation free country terrorism may grow with logarithmic distribution.

thakursurender101@rediffmail.com

MINING MULTIPLE BODY MEASUREMENTS OF MALES FOR THE FEASIBILITY OF ASSIGNING A “PERSONAL CODE”

C. N. Kuruwita
Department of Statistics, University Of Colombo, Sri Lanka

with

W. N. Wickremasinghe (Department of Statistics, University of Colombo, Sri Lanka)

Due to the inherent rigid structure in the conventional size regiments in ready-made garments, a fair proportion of people find it difficult to wear them without alterations. The present study explores the possibility of assigning a “Personal Code” to an individual that could condense all his/her measurements into a single value, through an application of Statistical Data Mining. It is hoped that the new code will minimize the risk of misspecification and provide an alternative size regiment for ready-made garments.

Measurements on head circumference, collar, sleeve, waist, length of leg and foot were collected from 409 Sri Lankan males in the age group 21 to 28, and are used to identify natural constellations of individuals, with the help of statistical and artificial neural networking tools. Preliminary analysis of the data is used to show that the highest correlation exists between Collar size and Waist and that Waist size accounts for much of the total variation. Based on Principal Component analysis it is shown that first three components having meaningful descriptions of shape and structure of a person. The three components account for nearly eighty percent of the total variation. Factor analysis is also used to confirm these results. Further, it is shown that 18 different clusters can be identified as a basis for a “Personal Code”, based on the Ward’s clustering technique which is shown to outsmart other techniques used. A statistical rule that outperforms the neural networking method is devised to classify a new individual to one of those clusters with over ninety percent accuracy. Further research based on more records may be helpful in finding confirmatory results.

chinthaka77@yahoo.com

COMPARATIVE STUDY OF ROUTING ALGORITHM FOR MOBILE COMPUTING
In this paper we have tried to analyze various routing algorithms for mobile computing. The parameters used to analyze the algorithms are based on major aspects like Algorithm efficiency, Networks size, reduced routing complexity, architecture, improve delay performance, connectivity, performance and much more. Algorithms considered here are WRP (Wireless Routing Protocol), DSR (Dynamic State Routing), DDR (distributed Dynamic Routing), TORA (Temporary Ordered routing Algorithm), RDMAR (Routing Protocol for Mobile computing), GSR (Global State routing), ZRP (Zone routing protocol), Bellman Ford algorithm (DBF) and STAR (Source tree adaptive routing). Mobile Ad hoc Networks community is trying to standardize the protocols and this study will provide a great help to them.

akush20@hotmail.com

BAYESIAN ANALYSIS OF SIMPLE LINEAR ECONOMETRIC MODEL

Ram Lal
Allahabad Agricultural Institute-Deemed University, Allahabad-211 007

Here, the Bayesian analysis of a simple linear econometric model have been carried out considering suitable prior densities of the parameters involved in the model which take values in a restricted domain. Bayes estimator of the regression coefficient have been obtained under the assumptions of squared error loss function.

USAGE OF SPACINGS IN OUTLIER DETECTION

S. Lalitha
Department of Statistics, University of Allahabad, Allahabad, India

with

Anand Mishra

The minimum spacing is used as the test statistic for detection of outliers for a sample from a standard uniform distribution. The critical values for this statistic are tabulated and the power values for certain shifts are also shown. The performance of this statistic is highlighted with an example.

ADMISSIBILITY OF TEST PROCEDURE BASED ON INCOMPLETE SPECIFIED MODEL

Kusum Lata
Department of Statistics, University of Rajasthan, Jaipur, India
In this paper we have derived condition of admissibility based on three preliminary test procedures. Condition of admissibility is derived theoretically as well as we have calculated the admissible regions for different sample sizes and specifically gave admissible and inadmissible regions. Also recommendations have been made for which level of significance is our test procedure always admissible and when we get maximum space.

latak@yahoo.com

SOME OUTLIER TESTS FOR UNIFORM VARIATE

Anand Mishra
Department of Statistics, University of Allahabad, Allahabad, India

with
S. Lalitha

Two test statistics $Z_l$ for left shifted and $Z_n$ for right shifted in between observations and another two statistics $W_1$ and $W_n$ for end observations from a standard uniform distribution have been developed. A slippage alternative model has been considered. The null distribution of these statistics and their power values are calculated.

IMPROVEMENT OF SIMULATED ANNEALING ALGORITHM TO SOLVE NP COMPLETE PROBLEMS

Nutan Mishra
Department of Mathematics and Statistics, University of South Alabama, Mobile, Alabama 36688, U.S.A.

Simulated Annealing Algorithm is one of the optimization algorithms to find the near optimal solutions to NP complete problems which are combinatorial in nature. It is one of the most popular algorithms to solve the problems pertaining to design and planning. VLSI design and traveling salesman problem are couple of examples where the said algorithm has been implemented successfully. Although this is not a viable option for real time software where the response is desired in certain time limit. Simulated Annealing algorithm takes longer computer time to return the solution. Efforts have been done to improve the performance of the algorithm by reducing amount of times taken. Set of solutions are produced by the algorithms during the search of near optimal solution. This set has the markovian property among the elements which has been exploited by the Martin et. al to improve the performance. The present article reviews the different techniques to improve the algorithm in the recent years.

mishra@cis.usouthal.edu

ON RECENTLY CONDUCTED DEMOGRAPHIC SURVEYS IN INDIA
Rama S. Mishra  
Population Research Centre, Department of Statistics, Patna University, Patna

with

Arun K. Sinha

Recently, some large demographic surveys have been conducted in India. In this paper we examine critically the surveys designs and other statistical contents of these surveys.

BAYESIAN ANALYSIS OF STATE SPACE MODELS WITH PANEL DATA

Ranjita Mishra  
Department of Statistics, Banasthali Vidyapith, Rajasthan

with

Anoop Chaturvedi (Department of Statistics, University of Allahabad, Allahabad)

The panel data models are gaining popularity among applied researchers as the longitudinal nature of these data can easily handle the complexity of biological, social, economic and environmental phenomenon. This paper considers the analysis of State space models with panel data under Bayesian framework. The conditional posterior distributions of various parameters of the model have been obtained under suitable prior assumptions. Gibbs sampler scheme has been proposed to estimate the marginal posterior densities for these parameters.

ranjita_m_2000@yahoo.com

PROSPECTS & PERSPECTIVES OF MULTILATERAL TRADING SYSTEM IN INDIA WITH THE AID OF FINANCIAL STATISTICS WITH REFERENCE TO GLOBAL TRADE ANALYSIS PROJECT

Sambit Kumar Mishra  
Executive Director, Samajik Vikas Sanstha, Raigarh (C.G.)

This paper uses information on bound & applied tariffs from the agricultural market access database, & the global trade analysis project (GTAP) computable general equilibrium model to assess the effects of tariff reforms under alternative assumptions. Effective tariffs as the minimum of bound rates after negotiated tariff cuts & applied rates are calculated & integrated into the GTAP database. Based on this information, several trade policy reform scenarios evaluated, comparing the outcome of model runs in which tariff reductions are based on effective rates with ones in which the reform are represented through cuts in either applied tariffs or bound tariffs. The results as reference to get highlights the commodities & countries for which a detailed representation of tariff schedules in economic models seems particularly warranted.

sambit_kumar_mishra@rediffmail.com

RECRUITMENT PROGRAMMES WITH THE AID OF OPTIMAL TIME INTERVAL FOR RESTRUCTURING RECRUITMENT CHALLENGES OF INDIAN INDUSTRY
In any organization recruitment of persons for training before placements is a common practice. Training of personnel involves some cost. The exit or waste of trained personnel also involves some cost. Placement of trained personnel is in phased manner depending upon the vacancies that arise. In this paper the optimal time period between successive recruitment programmes is obtained under specific assumptions of cost. Cost of shortage due to non availability of trained personnel is also discussed. Numerical illustrations are also provided with graphical illustrations.

ON ESTIMATING THE SCALE PARAMETER OF THE SELECTED GAMMA POPULATION UNDER THE SCALE INVARIANT SQUARED ERROR LOSS FUNCTION

Neeraj Misra
Department of Mathematics, Indian Institute of Technology,
Kanpur 208 016, India

with

E. C. Van der Meulen (Department of Mathematics, Katholieke University, Leuven, B-3001 Heverlee, Belgium) and Karlien V. Branden (Department of Mathematics, Katholieke University, Leuven, B-3001 Heverlee, Belgium)

Let $X_1$ and $X_2$ be two independent random variables representing the populations $P_i_1$ and $P_i_2$, respectively, and suppose that the random variable $X_i$ has a gamma distribution with shape parameter $p$, same for both the populations, and unknown scale parameter $\theta_i$, $i = 1, 2$. Define, $M = 1$, if $X_1 > X_2$, $M = 2$, if $X_2 > X_1$ and $J = 3 - M$. We consider the component wise estimation of random parameters $\theta_M$ and $\theta_J$, under the scale invariant squared error loss functions $L_1(\theta_M) = (\theta_M/\theta_M)^2$ and $L_2(\theta_J) = (\theta_J/\theta_J)^2$, respectively. Sufficient conditions for the inadmissibility of equivariant estimators of $\theta_M$ and $\theta_J$ are derived. As a consequence, various natural estimators are shown to be inadmissible and better estimators are obtained.

A STUDY OF SANITARY PRACTICES AND ITS ASSOCIATION WITH DIARRHOEA IN RURAL CHILDREN / ROLE OF EDUCATION

R. P. Misra
Department of Community Medicine, KG Medical College, Lucknow and Institute of Applied Statistics and Development Studies, Lucknow

with
The present study was carried out on 159 children whose parents were educated about the importance of sanitary practices and its role in reducing childhood diarrhea through a campaign approach along with messages related to good hygiene. Another 165 children were taken as control where no such education was imparted. The education mainly focused on the use of sanitary latrines and washing of hand before meals, after defecations and after cleaning infant’s faces. The impact of the health education was assessed at the end of one year. A significant reduction in diarrhea was found at the end of one year (from 14.2 to 7.4%) in those families where the education was imparted as compared to those where no such education was given. The logistic regression was applied to study the relative importance of various factors in reducing the diarrhea, the details of which will be discussed.

dr_aknigam@yahoo.com

**BIOINFORMATICS-APPLICATION IN BASIC AND APPLIED BIOLOGY OF FISH**

Bimal P. Mohanty  
Central Inland Fisheries Research Institute- Riverine Division, 24 Panna Lal Road, Allahabad, India  

with  
S. Mohanty (School of Life Sciences, Jawaharlal Nehru University, New Delhi, India) and D. Karunakaran (Central Inland Fisheries Research Institute- Riverine Division, 24 Panna Lal Road, Allahabad - 211 002, India)

Bioinformatics is a comparatively younger discipline that bridges the life sciences and computer sciences. The explosive growth of biological sequence information has made it imperative to integrate these two disciplines. Organization and analysis of biological data are the main activities of bioinformatics. Algorithms to create, maintain and access the sequence databases are among the most important contributions that bioinformatics has made for the life sciences. In the flow of genetic information from sequence to function, the stored information is translated twice: first from DNA to mRNA in the process of transcription, then from mRNA to protein in the process of translation. DNA and protein sequence comparisons have become routine steps in biochemical characterization, from newly cloned proteins to entire genomes. Genomics attempts to make a complete inventory of genes and nucleic acid sequences. In contrast to genomics approach, proteomics attempts to study the expressed proteins. Protein manifest physiological as well as pathophysiological processes in a cell or an organism and proteomics describe the complete inventory of proteins in dependence on in vivo parameters. Proteomics is complementing genomics as a tool to study life sciences. The two key technologies in experimental proteomics are: 1) 2-D PAGE with image analysis and 2) biological mass spectrometry (MS) with database searching. 2-D PAGE technique is finding application in fisheries for identification of serum/plasma proteins that might be involved in the constitutive resistance to infections, muscle protein characterization, and biochemical analysis of cross-reactive antigens, understanding the molecular pathogenesis and genetics of disease resistance. We are developing 2D-reference maps of commercially important fish and shellfish and plan to
construct an index of the piscine proteins, by the construction of 2D-database that may be useful in identification of quantitative trait loci (QTL).

bimalmohanty12@rediffmail.com

POWER FUNCTION AND ADMISSIBILITY OF THE TEST PROCEDURE OF AVERAGE LIFE OF EXPONENTIAL DISTRIBUTION IN TYPE-II CENSORED DATA

S. Mudgal
University of Rajasthan, Jaipur

with

S. K. Saxena and S. K. Gupta

This paper presents power and admissibility of test procedure of average life of exponential distribution in type II censored data. A preliminary test is used to decide whether to use a one or two parameter exponential distribution. The power of the test procedure for model three are calculated and discussed. The power of model one and two and admissibility conditions for all three models are obtained.

sharad_mudgal@indiatimes.com

DISTRIBUTION OF URBAN POPULATION IN INDIA

S. Muthurathnam
Department of Biostatistics, Christian Medical College, Vellore- 632 002

The study of distribution of urban population is useful for planning development programs. As revealed by the recent census of India, 285 million people live in urban areas. Nearly 108 million of them reside in the first four urban agglomerations. 68p.c of this population is concentrated in the central (municipal corporation) areas. The extent of such concentration in the largest urban agglomerations is highlighted. These statistics are compared with the pattern of distribution of urban population in 1991.

smuthusaras@rediffmail.co.in

ECHELONS AND PHASES AS CANDIDATES FOR RANKING IN SPATIAL RANKED SET SAMPLING

Wayne L. Myers
The Pennsylvania State University

Ranked set sampling improves efficiency of estimation by ensuring better representation across the distributional range than conventional occurrence-based randomization. Geospatial investigations offer opportunities for using ranked set sampling in a double sampling sense if synoptic information is available for purposes of ranking. Databases for geographic information systems (GIS) typically constitute the more readily available sources of spatial variables by which to accomplish ranking. Once approach considered here facilitates cluster-based combination of several spatial variables for joint use in the ranking operation. Multi-band image data and other pseudo-band information can thereby be mobilized for the
purpose of geospatial ranked set sampling. A second analytical approach called 'echelons' is applicable for surface or pseudo-surface variables. Echelons serve to give rankings relative to regional context and spatial complexity. By using echelons, one can segregate rankings for spatial elements that are part of upper-level formations from those that have similar values but occur in the context of lower-level formations.

INFRASTRUCTURE DEVELOPMENT INDEX: AN ANALYSIS FOR 17 MAJOR INDIAN STATES
(1990-91 TO 1996-97)

A. L. Nagar
National Institute of Public Finance and Policy, New Delhi, India
with
Sudip Ranjan Basu (National Institute of Public Finance and Policy, New Delhi, India)

In this paper we propose a method to compute a composite measure of infrastructure development by combining the available services of physical infrastructure. We estimate the value of infrastructure development index (IDI) for 17 major Indian states for the period 1990-91 to 1996-97. We also obtain the weights to be attached to different physical infrastructural services. The telecommunication services turn out to be the most dominant among the chosen infrastructure services, followed by the transportation facility and availability of energy/power services. A positive relationship has been established between the IDI and per capita net state domestic product (PCNSDP).

MODELING THE MALARIA ACTION IN THE BODY

Bello Hassan Onimisi
Department of Pharmaceutical Chemistry, University of Lagos,
Idi-Araba, Lagos, Nigeria

Malaria has been a major diseases in Africa for decades and up till now researchers have work and are still working to stem the menace of this deadly diseases down and if possibly come u with a vaccine but all of these effort seems to no avail.

In this paper I have propose to bring up models that would probably provide a lead in the search for such vaccine.

ONIBELL’S ALTERNATE INFERENTIAL STATISTICS THEORY

Bello Hassan Onimisi
Department of Pharmaceutical Chemistry, University of Lagos,
Idi-Araba, Lagos, Nigeria
The paper dwell much on the interference of probability of event with its normal distribution when applied to combination.

I have observed some quite some interesting result when the combination of an event or object I traced in line of its probability under its distribution.

Further works would reveal how this theory can be use to fashion out models that would help policy makers when selection and probability effect collide in policy making.

hassanbello2001@yahoo.com

RELIABILITY ESTIMATION: FROM CLASSICAL TO BAYESIAN APPROACH

M. Pandey
Department of Zoology, Banaras Hindu University, India

Statistical reliability is comparatively a recent discipline in the area of statistical theory. Reliability theory concerns in particular in the evaluation of the performance of a system or a subsystem or a component. The initial literature mostly describes classical estimation procedures (cf. Mann Schafer and Singpurwalla (1974), Barlow and Proschan (1975) and Lawless (1982)) and mentions in brief Bayesian techniques also. Martz and Waller (1982) gave the first detailed Bayesian approach to reliability estimation. This paper tries to give the evolution of the reliability estimation methodology starting from Pandey (1982) to Pandey and Chandra (2002) and Pandey and Shomvanshi (2002) from classical to Bayesian through preliminary test estimators, adaptive estimators, Bayes estimators and concatenated adaptive Bayesian estimators. The reliability problems considered are with well known failure models, some new failure models, stress-strength models, a form of accelerated life testing problems using multiple stresses and in software reliability growth models.

manju@bhu.ac.in

WATER QUALITY INDEX OF RIVER GANGA AND MULTIPLE REGRESSION AND DIRECTIONAL MEASURES OF ITS EFFECT ON INCIDENCE OF WATER BORNE ENTERIC DISEASES IN VARANASI CITY

M. Pandey
Department of Computer Science and Department of Zoology, BHU, Varanasi, India

From the secondary data on 5 water quality parameters a Water Quality Index (WQI) has been obtained for river Ganga at Varanasi. Based on rating curves, sub indices are obtained for each parameter from January 1995 to June 1999 and are aggregated using the harmonic mean square formula. Tend of water quality of river Ganga is obtained using regression plots and Seasonal Kendall test. Despite more than 15 years of implementation of Ganga Action Plan has elapsed, the quality is still deteriorated. The multiple linear regression analysis shows a high multiple correlation implying that incidence of water borne enteric
diseases in Varanasi is significantly affected by water quality of river Ganga at Varanasi.

The directional data analysis indicates a variation in incidence of said diseases in various areas of Varanasi City.

WAGE AND EMPLOYMENT CONDITIONS OF RURAL LABOURERS IN ALLAHABAD DISTRICTS

Rakesh Kumar Pandey
Lecture Govt. P.G. College Khatima, U. S. Nagar, Uttaranchal

The specific objectives of the present study are to investigate the problems of under employment and low wages for rural labors in Allahabad District, the wage discrimination by sex wise and age wise among rural labors in Allahabad, the future of employment opportunities for rural laborers in Allahabad and to propose a strategy for economic and social betterment of rural laborers, which are based on the hypothesis that maximum of the agricultural laborers is working on farms due to their backward social and economic conditions, a very important impact of worst socio-economic conditions of rural laborers on their work and wages, the enough discrimination in wages of rural laborers and the social welfare schemes for rural laborers run by government are very effective.

The present study is based on multi-stage purposive sample survey method. By tabulating the data on the basis of income, assets, sex wise/age wise wage, sex wise/age wise employment days, consumption expenditure etc. of rural laborers, we analyse above mentioned specific objectives.

AN ALTERNATIVE SAMPLING SCHEME BASED ON CONTROLLED SELECTION OF SAMPLES
USING LINEAR PROGRAMMING APPROACH

Archana Sinha
Department of Statistics, Lucknow University

with

Sheela Mishra and S. K. Pandey

An alternative sampling scheme named A-S (Arithmetic-Standard Deviation) Scheme based on the technique of controlled selection given by Rao and Nigam (1990) has been proposed and is applied on 19 natural populations taken from Cochrane, Parimal Mukhopadhyay and Sukhatme and Sukhatme. In each of these 19 natural populations, the proposed A-S scheme gives more efficient estimator for the estimation of population mean as compared to that given by Midzuno-Sen Horvitz Thompson estimator.

LOCALLY MOST POWERFUL AND OTHER RANK TESTS FOR INDEPENDENCE-WITH A CONTAMINATED WEIGHTED ALTERNATIVE

Parameshwar V. Pandit
The quantification of the concept of stochastic dependence for bivariate distribution has been attempted by several authors in the literature. Two weighted rank tests for testing independence against a weighted contamination alternative is proposed and their distributional properties are studied. We also derived a locally most powerful rank test for the alternative setting. The rank tests proposed are shown to be asymptotic locally most powerful for specific distributions.

DESIGNS FOR CROP SEQUENCE EXPERIMENTS: SOME CASE STUDIES

Rajender Parsad
IASRI, Library Avenue, New Delhi · 110 012, India

with

V. K. Gupta (IASRI, New Delhi · 110 012) and R. Srivastava (IASRI, New Delhi · 110 012)

Crop sequence experiments are conducted with two sets of treatments. One set of treatments is applied to Kharif crop and the other set of treatments is applied to Rabi crop. It is desired to compare the direct effects of treatments applied in Kharif and Rabi, residual effects of Kharif treatments and the interaction between the residual effects of Kharif treatments and direct effects of Rabi treatments. These experiments are run in a randomized complete block (RCB) design for Kharif treatments and for application of Rabi treatments the plots of Kharif experiment are subdivided into as many sub plots as the number of treatments in the Rabi crop. In this layout the main plots represent the residual effect of Kharif treatments and sub plots represent the direct effect of Rabi treatments. It is assumed that the residual effects persist for only one season. The treatments applied to Kharif crop have a residual effect on Rabi crop and the treatments applied to Rabi crop have a residual effect on Kharif crop. Usually such experiments are continued for two or more years. In the second year the Kharif treatments are again applied on main plots and observations are recorded sub plot wise. The sub plot differences now give the residual effects of Rabi treatments and main plots give the direct effects of Kharif treatments and the procedure is continued. This procedure, however, has the following drawbacks: (i) The effects in sub-plots and interactions are estimated with good precision but the precision for main plot treatments is generally poor, (ii) A split plot design is nothing but a RCB design with convenience, and in these experiments due to large number of subplots the replication becomes large and it is not possible to maintain homogeneity within replications. This results into large intra block variances and hence poor precision on treatment comparisons, (iii) The combined analysis of experiments conducted over years is a problem particularly when the error variances are heterogeneous. Therefore, balanced confounded factorial experiments or the extended group divisible designs may be an alternative for such experimental situations. These designs provide flexibility in choosing a design for a given experimental situation depending upon the precision required for each of the factorial effects. Considerable reduction in block size is possible through these designs. In some situations it is possible to get a design with smaller number of replications. It is also possible to carry out the combined analysis of the data generated through these experiments. These
alternatives are valid even when the treatments in both or any crop have the factorial structure.

Some such designs have been recommended in the real life situations. One such real life situation is described in the sequel.

Experimental Situation: Consider an experimental situation related to crop sequences where 5 herbicidal treatments were to be applied to the Kharif crop and 4 herbicidal treatments to the Rabi crop conceived by one Ph.D. student of Division of Agronomy, IARI, New Delhi. The interest is to compare the direct effects of Kharif and Rabi treatments, residual effects of Kharif treatments and interaction between the residual effects of Kharif treatments and direct effects of Rabi treatments. If we denote the Kharif treatments as 1, 2, 3, 4, 5 and Rabi Treatments as a, b, c, d, then the layout of the design suitable for the above experimental situation is given as:

Block 1 1a 2a 3a 4a 5a 1b 2b 3b 4b 5b Block 2 1a 2a 3a 4a 5a 1c 2c 3c 4c 5c Block 3 1a 2a 3a 4a 5a 1d 2d 3d 4d 5d Block 4 1b 2b 3b 4b 5b 1c 2c 3c 4c 5c Block 5 1b 2b 3b 4b 5b 1d 2d 3d 4d 5d Block 6 1c 2c 3c 4c 5c 1d 2d 3d 4d 5d This is an extended group divisible design with 20 (5 x 4) treatments and is based on the association scheme 1a 1b 1c 1d 2a 2b 2c 2d 3a 3b 3c 3d 4a 4b 4c 4d 5a 5b 5c 5d The two treatments are (01)th associates if first factor is at same levels i.e. in the same row, (10)th associates if the second factor is at same levels i.e. in the same column and rest are (11)th associates. As a result the parameters of the design are \( v = 5 \times 4 \times 20 \), \( b = 6 \), \( r = 3 \), \( k = 10 \), \( \lambda(01) = 1 \), \( \lambda(10) = 3 \), \( \lambda(11) = 1 \), and the efficiencies of different factorial effects as compared to the randomized complete block design are \( \bar{\lambda}(10) = 1 \), \( \bar{\lambda}(01) = 2/3 \), \( \bar{\lambda}(11) = 1 \), respectively. Here \( \bar{\lambda}(10) \), \( \bar{\lambda}(01) \), \( \bar{\lambda}(11) \) denote respectively the efficiencies of main effect of first factor, main effect of second factor and interaction of first and second factors respectively. The analysis of the data can be done using the contrast analysis.

In this presentation, several such real life situations will be discussed. Several existing association schemes will be shown as particular cases of Extended Group Divisible Association Scheme. Some methods of construction of extended group divisible designs shall also be presented.

Some crop sequence experiments are also conducted to develop suitable integrated nutrient supply system of a crop sequence. In these experiments, the treatments do not comprise of a complete factorial structure. The experimenter is interested in estimating the residual and direct effect of the treatments along with their cumulative effects. An application of structurally incomplete row-column designs for obtaining designs for these experimental situations shall also be discussed.

rajender@iasri.delhi.nic.in

ROBUST BAYESIAN RELIABILITY ANALYSIS OF RAYLEIGH MODEL

Manaswini Pati
Department of Statistics, University of Allahabad, Allahabad, India

with

Anoop Chaturvedi
The robust Bayesian analysis based on type-II censored data from a Rayleigh distribution has been carried out assuming a contamination class of prior distributions for the parameters. The robust Bayes estimators for the mean life, reliability function and failure rate have been obtained under the squared loss function (SELF) and the LINEX loss function.

**CURRENT DISPERSION PATTERN IN AN EPHEMERAL RIVER BEND, INDIA**

Barendra Purkait  
Geological Survey of India Petrology Division, 9th Floor, Marine Wing Block-DK-6, Sector-II, Salt Lake  
Kolkata-700091, India

Detailed plane table mapping on 1cm=5m scale of a point bar immediately after the monsoon during the winter months (October-March) was carried out in a bend of a low sinuous, ephemeral river Pathri, a tributary of the Usri river of the state of Jharkhand, Eastern India. Nature of crest lines of the surface ripples were plotted in details as per their physical occurrences. Sediment transport directions perpendicular to the crest of the ripples were measured. Water flow directions adjacent to the bar were also measured directly perpendicular to the ripple crest lines developed in the channel bed.

As the bar is crescentic in shape, sediment transport directions as well as the water flow directions of the adjacent channel swing gradually with the bend. Hence the entire bend has been divided into six sectors based on the curvature. In each sector, the vector means of the sediment transport directions in the bar and the water flow directions of the adjacent channel water were compared. Finally, the vector means of the sediment transport directions were also compared with the average direction of the channel central line in each sector.

It is observed that the vector means of the sediment transport directions in the bar lie within 4.5010 to 28.8780 of the vector means of the water flow directions of the adjacent channel. The sediment transport direction in the bar has a diverse pattern. Moreover, when the vector means of the sediment transport direction in the bar are compared with the average direction of the channel central line, the difference lies within 2.1160 to 9.6910.

Watson’s U2-statistic is employed for a meaningful interpretation of the difference. The results of the test statistic indicate that the sediment transport directions in the bar do not always differ significantly from the water flow direction of the adjacent channel but in each sector the distribution pattern of the sediment transport direction in the bar is same as the azimuthal distribution pattern of the channel central line.

The findings might be helpful in reconstructing the ancient river course, particularly to delineate the channel central line for a low sinuous river with reasonable accuracy.

marinew@vsnl.com

**PERSISTENCE AND GLOBAL STABILITY IN THREE POPULATION ECOLOGICAL COMMUNITIES**
We shall consider Dynamical Systems Models represented by a system of Autonomous Ordinary Differential Equations. An important feature of the ecological communities is that of persistence, i.e., the survival of all populations as opposed to the extinction of some of them. We shall consider models modelling Predation, Competition and Symbiotic interactions and give criteria for persistence and global stability.

**BAYESIAN ESTIMATION OF THE SCALE PARAMETER OF GAMMA DISTRIBUTION UNDER LINEX LOSS FUNCTION WITH CENSORING**

**Arun Kumar Rao**  
Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur

**R. S. Srivastava** (Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur)

In this article, we have obtained the Baye estimators of the scale parameter of gamma distribution. The loss function used is LINEX. A number of prior distributions have been considered.

vkgkp@rediffmail.com

**BAYESIAN ESTIMATION OF THE SHAPE PARAMETER AND RELIABILITY FUNCTION OF GENERALIZED PARETO DISTRIBUTION USING LINEX LOSS FUNCTION**

**Arun Kumar Rao**  
Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur

**Vijay Kumar and R. S. Srivastava** (DDU Gorkhpur University)

In this paper the Bayes estimator of the shape parameter and reliability function of the generalized Pareto distribution(GPD) have been obtained under LINEX loss function using quasi, natural conjugate and uniform prior distributions for shape parameter and non-informative and beta priors for estimation of reliability function. These estimators are compared with corresponding Bayes estimators under squared error loss function.

vkgkp@rediffmail.com

**CUMULATIVE SUM CONTROL CHART FOR LOG-LOGISTIC DISTRIBUTION**

**G. Srinivasa Rao**  
Department of Statistics, JKC Collage, Guntur-522006, India

**R. R. L. Kantam** (Department of statistics, Nagarjuna University)

The sequential probability ratio procedures of Statistical Inference are made use of in construction of a Cumulative Sum Control Chart for a variable process characteristic. The distribution of process variate is
log-logistic distribution. The construction of V mask and values of average run length are also presented. Such a chart is useful in early detection of shifts in the process average.

gaddesrao@yahoo.com

OPTIMUM AGE AND REPAIR REPLACEMENT POLICY

S. R. Rattihalli
The New College, Kolhapur

In this paper we consider a replacement policy which depends on both age and repair times of a single unit system. The failure of the system can be classified as either a Type I failure or a Type II failure. A Type I failure causes only a minor damage and can be set right by a minor repair. After a minor repair the unit is as good as that is at the time of failure and minor repair time is negligible. A Type II failure requires a major repair. If the repair time exceeds a fixed time, say $T_r$ then failed unit is replaced by a new unit. Otherwise the repaired unit is put into operation. If a working unit attains age $T_a$ without Type II failure the unit is exchanged by a new unit. Further it is assumed that the system needs maintenance while working. We derive an expression for expected long run cost per unit time based on a cycle associated with a regenerative process. The optimum values of $T_r$ and $T_a$ are obtained. These results generalise some existing results.

FAMILIES OF DISTRIBUTION FUNCTIONS

R. N. Rattihalli
Department of Statistics, Shivaji University, Kolhapur 416004, India

Class properties of some families of parametric distributions on $[0, 1]$ are defined and these families are used to enlarge a given family of distributions. The problems of testing and estimation have been discussed for a particular family of distributions on $[0, 1]$.

rnr5@rediffmail.com

HIND SIGHT ANALYSIS OF SOME TIME SERIES DATA

Atul N. Roy
Montgomery College and University of Maryland System

This presentation will show illustration and comparison of some different methods and computing packages for forecasting suitable for typical users.

aroy@mc.cc.md.us

BIVARIATE EXPONENTIAL DISTRIBUTUON: A REGRESSION APPROACH

Dilip Roy
Department of Business Administration,
Bivariate modeling of exponential distribution has remained a problem of interest for model builders for long, as there has been no convergence of thoughts. The basic problem arises out of the fact that majority of attempts were centred around celebrated lack of memory property. In the bivariate case, lack of memory property in the strongest sense gives rise to unique determination of independent exponential marginals. Since it does not admit a structure of dependence various alternative versions of bivariate lack of memory property have come up with some merits and some demerits. Sometimes even the marginal exponentiality was ignored for the sake of satisfying a lack of memory property.

The present work revisits this modeling problem from a regression approach. A bivariate exponential distribution has been proposed with marginal exponentiality and linear regression. The natures of dependence, conditional distribution and bivariate characteristic function have been studied. The linear regression established herein can be of use for prediction of one variable from the other.

We also propose a general approach for bivariate extension with linear regression under some mild restrictions.

diliproy@vsnl.com

GENERATING LIFE TABLES FOR THE DISTRICTS OF WEST BENGAL: A MODEL APPROACH

Anuj Kumar Saha
Indian Statistical Institute

with

Anuj Kumar Saha (Indian Statistical Institute, Kolkata, India) and Debasish Roy (University of Auckland, New Zealand)

The present study aims at focussing on the age pattern of mortality by sex for the districts of West Bengal, a constituent state of India. The districts do not have any idea about the age pattern of mortality since partition of India in 1947. Registrar General (R.G.), India, did not provide life tables for the districts even to date which might be due to unavailability of death statistics by age distribution of the population. We have only two different inputs namely, children overborne and surviving by sex classified by age group of mother, and widowhood status of the male/female spouse by age group respectively. At the first instance, the study attempts to estimate infant and child mortality, and adult mortality and, finally it has tried to focus on the age pattern of mortality by blending these two sets of survivorship probabilities namely, (1) child survivorship probabilities and (2) adult survivorship probabilities with a view to deriving life tables by sex generated by Brass Two Parameter Logit Life Tables Systems.

sahaanuj@hotmail.com

A TRIPLE $g$-CLASS OF ESTIMATORS OF STRUCTURAL COEFFICIENTS

Rajiv Saksena
A triple g-class of Stein-type estimators of coefficients in a structural equation of a simultaneous equations model is proposed and its properties are analyzed when disturbances are small. The range of dominance conditions of the triple g-class of estimators over the two stage least squares estimator is wider than that of the dominance conditions of double g-class estimators over the two stage least squares estimator; hence, in the extended range of the dominance conditions of the triple g-class of estimators over two stage least squares estimator, the triple g-class of estimators dominates both the double g-class and the two stage least squares estimators.

rajvaksena@indiatimes.com

A NEW APPROACH TO PPS SAMPLING SCHEME-II

Alpna Saxena
Department of Mathematics and Statistics, J.N.V. University, Jodhpur, India

with

Manju Saxena

In the present paper, we propose another Sampling Scheme II using auxiliary variable for cluster sampling. The efficiency of a proposed sampling scheme is compared with those of usual cluster sampling scheme, sampling scheme I where auxiliary variable is used in selecting the elements in the clusters and SRSWOR. It has been observed that among all the schemes considered for comparison, the performance of sampling scheme proposed earlier, i.e., Sampling Scheme-I is the best. But keeping in view the factors such as simplicity of selection, operational convenience and other factors, the proposed sampling scheme II can be used with advantage in practice.

ON SELECTION OF SAMPLES WITH INCLUSION PROBABILITIES PROPORTIONAL TO SIZE

Manju Saxena
Department of Mathematics and Statistics, Jai Narain Vyas University, Jodhpur, Rajasthan, India

In this paper a sampling scheme with inclusion probabilities proportional to size for sample size greater than two is suggested. The sampling scheme, besides being simple has the advantage of providing a non-negative and stable variance estimator. The performance of the proposed sampling scheme relative to alternative schemes such as samford’s II (pie) PS systematic sampling scheme and PPS with replacement sampling scheme, is very satisfactory. The sampling scheme is given for even and odd sample sizes both.

manjusairam@rediffmail.com

MARKOV PROCESSES, TIME-SPACE HARMONIC FUNCTIONS AND POLYNOMIALS

Arindam Sengupta
For a stochastic process $M = \{M_t\}$, a time-space harmonic function is a function $f$ of two variables such that $\{ft, M_t\}$ is a martingale with respect to the natural filtration of $M$. We investigate some criteria of richness of the class of time-space harmonic functions that yield Markov property for $M$. A consequence for the Markov property for a polynomially harmonisable process when the functions can be chosen to be polynomials is obtained. A few methods for the construction of time-space harmonic polynomials for certain Markov processes are described and illustrated with examples.

arindam@iitg.ernet.in

**PREDICTION IN RESTRICTED REGRESSION MODELS**

Shalabh
Punjab University

This paper has considered the problem of predicting both the actual and average values of study variable in a linear regression model subject to a set of exact linear restrictions on regression coefficients. Three types of predictions arising from restricted regression and Stein-rule methods are presented for the values of study variable within the sample and outside the sample, and their performance properties are analyzed.

shalabh1@yahoo.com

**LIFE TEST ACCEPTANCE SAMPLING PLAN WITH SPECIFIED WARRANTEE PERIOD**

Gauri Shankar
Pt. Ravishankar Shukla University, Raipur (C.G.), India

with

D. K. Gangeshwar

The electronics industries have heavily prioritized enhancing the quality and life time of electronic goods or components. We come across many items where there is a requirement of a minimum warranted life time, say $t (t > 0)$ at (from) which the unit is regarded as satisfactory. The units which fail during warrantee period or in burn-in period are discarded and replaced by a new one. This paper presents an acceptance sampling plan to ensure warranted life time of units in the accepted lots. For simplicity, the life time distribution is taken to be exponential. The OC function of the sampling plan have been derived and illustrated numerically.

gaurishankar@rediffmail.com

**PUBLICATIONS ON TUBERCULOSIS IN DEVELOPING AND DEVELOPED COUNTRIES: A QUANTITATIVE ANALYSIS**
SCRA 2002-FIM IX

M. Shanmugham
The Tamil Nadu Dr. M.G.R. Medical University, Chennai, Tamil Nadu, India

with

S. R. Elangovan (Annamalai University, Tamil Nadu)

Globally, there is rapid growth of publications on Tuberculosis in the past four decades. These literatures are addressing and helping the TB health personnel to respond to the present and future challenges of TB. This study aims to determine the pattern and trend of TB publications in the last four decades by bibliometric analysis. The MEDLINE Database was used to retrieve articles on TUBERCULOSIS in the MESH fields (i.e. directly related to the articles) for the period 1966 to 2001/06 through SPIRS search Engine. A specialized software package was developed by the author using Fox Pro (ver. 2.5). Various fields like Author, Country of Publications, Language, Publication Type, Year, Medical Subject Headings (MESH) etc., were searched and stored in a text file. All the Text file contents were then merged in to a single database file, using this detailed analysis was done. Total of 72, 390 articles in TUBERCULOSIS were processed. The number of TB articles available in the year 1963 was 16 only. This has abruptly raised to 1648 in the year 1965. The analysis shows that from 1966-1974 the trend has been oscillating between 2336 to 2912. In 1975, this has decreased to 1940 articles and then after it steadily decreases until the year 1994. Since then articles steadily increases till the year 2001. This has also been reflected on the publications of the developing countries. Based on the individual Journal names there were 72, 390 TB articles published in 3669 Journals. These journals are from 94 countries in 38 languages. By language categorization maximum articles were published in English (35, 417 articles i.e. 49English-Italian (4), English-Portuguese and English-Spanish. The remaining were in Afrikaans, Bulgarian, Czech, Chinese, Finnish, French, Germany etc in 36 Non English Journals. Most of these are from developed countries. Considering the Journal factor, among 3,669 journals, the highest number of articles were from Probl-Tuberk (USSR, 7614 articles i.e. 10.52 Keckakku (Japan, 1914, i.e.2.61 Journals and then by Am-Rev-Respir-Dis published in English. 72, 390 articles were by published 94 countries. Of which developed countries like United States (15737), England (6372), Japan (4241), Italy (2349), Poland (2311), Germany West (1341), Germany East (1182), had highest publications from Tuberculosis. In the developing countries highest articles was from India (1975 articles). 61.21 were published by Journals from developed countries. The 39.61 the Tuberculosis articles of the developing country published in developed country journals. Strangely only 21 country articles were published in the developing country journals. The total of 27,858 (38.48 country) publications Journals. The study found that publications on Tuberculosis were steady some period and oscillating in some period though there is a significantly less contributions on developing country articles even from journals published from developing country. Much work is needed to increase the publications from these countries as they bear a great burden of TB cases globally. The presentation will discuss in detail about developing countries and developed countries to year-to-year, language, medical subject heading, publication type, authors address, countries of publications, language factor and all other factors to identify the trend in these publications.

murushanmugham@yahoo.com
WHY ESTIMATION OF SOFTWARE RELIABILITY IS SO DIFFICULT?

Aditya K. S. Shomvansi
Department of Statistics, Banaras Hindu University Varanasi – 221005, India

This work is concerned to the difficulties that arise in the estimation of the software reliability. An attempt has been made to highlight the nature of software failure data generating mechanism, which preventing use of the exponential model, which plays a central role in the reliability analysis as the normal distribution in classical statistics.

akss9@yahoo.co.in

A MULTIVARIATE CLASS OF CHAIN ESTIMATORS FOR ESTIMATING POPULATION PARAMETERS

Diwakar Shukla
Department of Mathematics and Statistics, H. S. Gour University,
Sagar, M.P., India

Suppose a variable is under study and several other variables are also available, each correlated with the study variable and mutually too, then an intelligent use of auxiliary information may improve upon the efficiency of estimators. Many estimators could be defined for making a use of such resource by changing variety of sampling schemes. A group of estimators, possessing similar properties generates a class. This paper presents a class of estimators for estimating ratio and product of two population means, using information on multiple supplementary variables through a double sampling scheme. The asymptotic expressions for bias and mean squared error of the proposed class are derived in this text. Many properties are discussed along with the case of optimum mean squared error. The class is exhaustive enough to incorporate several existing and possible non-existing estimators as particular cases.

diwakar@epatra.com

HOW MANY MEASUREMENTS FOR EXPOSURE ASSESSMENT? BALANCING COST AND PRECISION FOR AN OPTIMUM ENVIRONMENTAL SAMPLING OF AMBIENT AEROSOL

Rakesh Shukla
Division of Biostatistics and Epidemiology, Department of Environmental Health, University of Cincinnati

Studies requiring ambient exposure assessments must ask the following question: How often should measurements be taken? Answers to such questions are dictated by budgetary constraints as well as spatial-temporal considerations. For example, do we obtain samples during all seasons, all months within seasons, weeks within months and days within weeks and so on? On the one hand, one can obtain a snapshot sample and regard it as an estimate of the “true” exposure mean. On the other hand, one can obtain a large number of samples and then average those to represent the “true” exposure mean. The former estimate is the least expensive but also the most imprecise. The latter, may be very precise but prohibitively costly. We propose a solution. First conduct a pilot study with a feasible and promising
sampling plan and apply the statistical methodology of Variance Component Analysis (VCA) to the exposure data. We demonstrate that an optimum sampling design will maximize the precision of our exposure estimate for a pre-specified total relative cost (relative to the pilot sampling design) of sampling. Alternatively, we can minimize the sampling costs for a certain pre-specified relative precision of the estimate (relative to the precision in the pilot design). Our approach is illustrated with an on-going study of assessing exposure to diesel particulates in a birth cohort. We show that a pilot study followed by the VCA analysis, will lead to considerable savings and can also provide precise estimates for the subsequent full-fledged study based on the optimum design arrived at by the VCA approach.

rakesh.shukla@uc.edu

MULTI-CHARACTER SURVEYS USING RANDOMIZED RESPONSE SAMPLING

S. S. Sidhu
Department of Mathematics, Statistics & Physics, Punjab Agricultural University, Punjab Agricultural University
with
M. Javed and I. S. Grewal

In the present investigation, alternative estimators for estimating population totals in multi-character survey sampling using randomized response technique have been suggested when certain variables have poor positive correlation and others have poor negative correlation with selection probabilities. The estimators proposed by Hansen and Hurwitz (1943), Rao (1966) and Sahoo et al. (1994) under scrambled responses are shown as special cases of the proposed estimators. The relative efficiency of the proposed estimators with respect to each other has been studied through empirical study.

husukhjinder@hotmail.com

FORECASTING MORTALITY IN INDIA

Abhishek Singh
International Institute for Population Sciences, Mumbai
with
F. Ram (International Institute for Population Sciences)

In the present paper time series methods has been used to make long-term forecasts of age specific death rates in India for the period 2000-2015. Lee Carter model is used to forecast the level of mortality in future. The logs of age specific death rates are modeled as a linear function of an unobserved time varying index k, with parameters depending on age. The model is fit to the matrix of age specific death rates of India for the period 1970-1998. Whereas life expectancy at birth has increased at a decreasing rate over the years, the decline in k may be taken to be roughly linear. k is next modeled as a time series. The forecasting of k has been done by using Autoregressive Integrated Moving Average Model. Forecasts of
age specific death rates are obtained from the forecasts of $k$ and the other life table functions are derived and presented in the paper.

By the year 2015, the life expectancy at birth is forecasted to increase by 5 years to 65.78 years for males and by 8 years to 69.92 years for females. For the period 1993-97, 45 % of male births and 52 % of the female births survive up to age 70. This percentage is expected to increase to 51 % and 61 % respectively by the year 2015. By the year 2015, 92 % of the births irrespective of the sex are expected to survive up to age 5. Comparison with the Registrar general Estimates reveals that the forecasts given by LC model for females is more or less the same as that of RGI. But for males, the forecasts given by the LC model is found to be lower than that of RGI.

a_singh18@rediffmail.com

MACHINE REPAIR MODEL WITH SPARES AND TWO MODE FAILURE

Charan Jeet Singh
Department of Mathematics, G.N.D. University, Amritsar, India

with

Madhu Jain (Department of Mathematics, I.B.S. Agra, India)

This investigation deals with M/M/R machine repair problem with warm standby spares. The operating/spare machines may fail in two modes. The failure times of the machine and repair times of the repair men are assumed to follow a negative exponential distribution. There is a provision of additional repairmen in case of heavy traffic in order to reduce the backlog. We obtain queue size distribution at equilibrium in explicit form. Several performance indices are facilitated in explicit form by using queue size distribution. A cost model is developed in order to determine the optimal number of repairmen and spares simultaneously satisfying a specified level of machine availability. Numerical experiment is performed to illustrate the optimization problem posed. Sensitivity analysis is also done to analyse the effect of various system parameters.

USE OF LINEAR TRANSFORMATION FOR IMPROVING THE PRODUCT METHOD OF ESTIMATION IN SAMPLE SURVEYS

G. N. Singh
Department of Applied Mathematics, Indian School of Mines, Dhanbad-826 004, India

For estimating a finite population mean under product method of estimation, simple linear transformations using the known coefficient of skewness, coefficient of kurtosis and standard deviation of the auxiliary variable have been considered. The suggested transformations provide efficient product estimators with
less absolute bias than conventional product and other product type estimators.

gnsingh_ism@yahoo.com

BAYESIAN INFERENCE OF DISTRIBUTIONS WITH PERIODIC FAILURE RATES

Randhir Singh
Ewing Christian College, Allahabad, India

This paper deals with the Bayesian estimation of the unknown parameter 'alpha' and the reliability of the distributions with periodic failure rates. The estimation has been done by taking the Gamma prior for 'alpha' and various types of loss functions.

randhir@mri.ernet.in

STOCHASTIC MODELLING AND ANALYSIS OF A TWIN HEARTH FURNACE SYSTEM OF STEEL MELTING SHOP AREA OF AN INTEGRATED STEEL PLANT

S. K. Singh
S. O. S. in Statistics, Pt. Ravishankar Shukla University

with

S. N. Singh (Department of Mathematics, University of Transkei, Utamata, Eastern Cape, Private Bag XI, South Africa, e-mail: singh@getafix.utr.ac.za)

This paper deals with the stochastic modeling and analysis of a Twin Hearth Furnace system of steel melting shop area of Bhilai Steel Plant. System consists of a twin hearth furnace with one charging machine in standby configuration. The furnace has two hearths connected to each other with a passage to enable the waste gas flow from one hearth to the other. Three oxygen lancers are provided on the roof of each hearth for intense blowing. Besides regular repair, schedule maintenance is also performed. Failure time distributions of all units are taken to be negative exponential whereas repair and maintenance time distributions are taken to be arbitrary. Using regenerative point technique, several system characteristics such as mean time to system failure (MTSF), availability, busy period analysis of the repairman which are useful to the system managers and engineers, are evaluated. At last some graphs are plotted in order to highlight the important results.

drksingh@raipurmail.com

MODELING RANDOM FLUCTUATIONS IN ROTATIONAL ANGLES OF MOLECULES

Harshinder Singh
West Virginia University, U.S.A.

Probabilistic modeling of rotational (torsional) angles of a molecule is important in molecular sciences for evaluation of conformational entropy: a measure of stability of the conformation. Many pivotal
conformational changes such as protein folding and ligand binding are caused by reduction in the conformational entropy and thus the evaluation of entropy is important in understanding the factors that cause change of one conformation to another. We discuss circular probability modeling approach to modeling of torsional angles in molecules. Application of von Mises distribution is illustrated by modeling the torsional angle of methanol molecule. We discuss a probabilistic model for two dependent circular variables which is a natural torus analogue of the bivariate normal distribution. Its usefulness is illustrated by modeling of two torsional angles of a pentapeptide. We also present a nonparametric approach to estimation of entropy which is based on kth nearest neighbor distances between the n sample points, where k (< n) is a fixed positive integer.

hsingh@stat.wvu.edu

A SELECTION PROCEDURE FOR SELECTING GOOD LOGISTIC POPULATIONS

Parminder Singh
Guru Nanak Dev University, Amritsar, India

with

Amar Nath Gill (Punjab University, Chandigarh, India)

In the present paper a selection procedure based on sample median is proposed to select a subset of k logistic populations which include all the good populations with probability at least P* (a pre assigned value). Simultaneous confidence intervals for the difference of location parameters, which can be derived with the help of proposed procedures, have been discussed. A selection procedure is proposed so that the probability of either selection a bad population or omitting a good population is at most 1–P*.

pjasmeet@go.com

APPROXIMATE MMSE ESTIMATION FOR EXPONENTIAL DISTRIBUTION FROM ACCELERATED TEST DATA UNDER ARREHENIUS REACTION RATE MODEL

Pranveer Pratap Singh
DDU Gorakhpur University, Gorakhpur

with

R. S. Srivastava (Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur)

In this article, we have considered the problem of minimum mean squared error (MMSE) estimator for exponential distribution from accelerated life test data obeying Arrehenius Reaction Rate Model.

A GENERALIZED MIXED REGRESSION ESTIMATOR WHEN DISTURBANCES ARE NON-NORMAL

R. Karan Singh
Department of Statistics, Lucknow University, Lucknow-226007, India

with

N. Rastogi
A generalized mixed regression estimator for the estimation of the regression coefficients in a linear regression model with incomplete prior information is proposed and its properties are studied when the disturbances are small and non-normal. Risk under the general quadratic loss function is considered for the comparison of estimators. Some estimators of the generalized class of mixed regression estimators are shown having their range of dominance conditions over the mixed regression estimator wider than that of the dominance conditions of available existing estimators over the mixed regression estimator.

rkaransingh@yahoo.com

STATISTICAL MODELS IN MARKETING RESEARCH - A REVIEW

Sandeep Singh
School of Management Sciences, Varanasi, India

with

Pradeep Singh (Southeast Missouri State University, USA) and
A. R. Mishra (School of Management Sciences, Varanasi)

Past few years have witnessed unparalleled developments in research methods. The technology of marketing research also has reached a stage of unusual complexity and diversification. This has lead to the development of advanced techniques. This paper will review various important and practically relevant statistical models, which can be used in marketing research.

prad_singh@yahoo.com

A TWO-STAGE PROCEDURE FOR POINT ESTIMATION IN A RANDOM ONE WAY MODEL

Sanjeeb Singh
Department of Statistics, University of Jammu,
Jammu, J&K, 180006, India

with

Rahul Gupta

A two stage procedure is proposed for the minimum risk point estimation of the mean of a random one-way model after proving the failure of the fixed sample size procedure to handle it. Second-order asymptotics for the first two moments of the stopping rule and the ‘regret’ are also obtained for the proposed ‘Two-stage’ procedure.

sanjeebsingh@rediffmail.com

ON BAYESIAN ESTIMATION OF MAXWELL DISTRIBUTION UNDER PRECAUTIONARY LOSS FUNCTION

Satya Prakash Singh
Department of Mathematics & Statistics,
In this paper, we have considered the problem of Bayesian estimation of scale parameter of Maxwell distribution using different priors with type II censoring. An attempt has been made to study the performance of the estimators.

**STOCHASTIC MODELING AND ANALYSIS OF SLAB CASTER SYSTEM OF CONTINUOUS CASTING SHOP AREA OF AN INTEGRATED STEEL PLANT**

Satyendra Prasad Singh  
S. O. S. in Statistics,  
Pt. Ravishankar Shukla University,  
Raipur, Chhattisgarh - 492010, India

with  
S. K. Singh (S. O. S. in Statistics, Pt. Ravishankar Shukla University, Raipur, Chhattisgarh - 492010, India)

This paper deals with the stochastic modeling and analysis of slab caster system of continuous casting shop area of Bhilai steel plant. System consists of two tundishes, one mould with rolling belt and gas cutter each. Initially one tundish is in preparation while other is kept in the cold standby. Once the functioning of all the required units gets started it is continued till the completion of the cycle, after this whole system is sent to scheduled maintenance. Slabs obtained are cut into the required sizes with the help of gas cutters and are sent to slab storage yard. Failure time distributions of all units are taken to be negative exponential whereas repair and maintenance time distributions are taken to be arbitrary. Using regenerative point technique, several system characteristics such as mean time to system failure (MTSF), availability, busy period analysis of the repairman which are useful to the system managers and engineers, are evaluated. At last some graphs are plotted in order to highlight the important results.

spsingh1232@rediffmail.com

**MAXIMUM LIKELIHOOD ESTIMATORS OF THE PARAMETER OF EXPONENTIAL DISTRIBUTION UNDER MULTIPLY TYPE-II CENSORING**

Umesh Singh  
Department of Statistics,  
Banaras Hindu University,  
Varanasi 221 005, India

with  
Anil Kumar and S. K Upadhyay (Department of Statistics, Banaras Hindu University, Varanasi 221 005, India)
In this paper we address the problem of obtaining the maximum likelihood estimators under multiply type-II censoring scheme for one parameter exponential distribution. It is well known that explicit expression for the maximum likelihood estimator does not exist under this censoring scheme. However, approximate maximum likelihood estimator has been proposed by Balasubramanian and Balakrishnan (1992) using certain approximations at the stage of solving the normal equations. We proposed to use an approximation in the likelihood function to get an approximate maximum likelihood estimator. In addition this, a numerical method is also proposed to get the maximum likelihood estimate. The estimators thus obtained are compared with the existing approximate maximum likelihood estimator on the basis of Monte Carlo Simulation technique.

AN ALTERNATIVE SAMPLING SCHEME BASED ON CONTROLLED SELECTION OF SAMPLES USING LINEAR PROGRAMMING APPROACH

Archana Sinha
Department of Statistics, Lucknow University
with
Sheela Mishra and S. K. Pandey

An alternative sampling scheme named A-S (Arithmetic-Standard Deviation) Scheme based on the technique of controlled selection given by Rao and Nigam (1990) has been proposed and is applied on 19 natural populations taken from Cochrane, Parimal Mukhopadhyay and Sukhatme and Sukhatme. In each of these 19 natural populations, the proposed A-S scheme gives more efficient estimator for the estimation of population mean as compared to that given by Midzuno-Sen Horvitz Thompson estimator.

sunil_abha@rediffmail.com

THE COST-EFFICIENCY OF RANKED SET SAMPLING

Arun K. Sinha
Department of Statistics, Patna University, Patna 800 005, Bihar

Ranked set sampling (RSS) is a cost-effective sampling method that is used to improve upon standard estimators of various population parameters of interest. In view of this fact this paper investigates the cost considerations involved with this scheme and compares the cost-efficiency of some RSS estimators with other corresponding standard estimators. This work may be of some particular interest to those who look for a cost-effective survey sampling technique.

arunkrsinha@yahoo.com

ROBUST SEQUENTIAL DESIGNS FOR NONLINEAR REGRESSION

Sanjoy K. Sinha
University of Winnipeg, Canada
The authors introduce the formal notion of an approximately specified nonlinear regression model, and investigate sequential design methodologies when the fitted model is possibly of an incorrect parametric form. They present small-sample simulation studies which indicate that their new designs can be very successful, relative to some common competitors, in reducing mean squared error due to model misspecification and to heteroscedastic variation. Their simulations also suggest that standard normal-theory inference procedures remain approximately valid under the sequential sampling schemes. The methods are illustrated both by simulation and in an example using data from an experiment described in the chemical engineering literature.

s.sinha@uwinnipeg.ca

AN APPLICATION OF A BAYESIAN HIERARCHICAL MODEL FOR ITEM FAMILY CALIBRATION

Sandip Sinharay
Educational Testing Service, Princeton, New Jersey, U.S.A.

with
Matthew Johnson (Educational Testing Service, Princeton, New Jersey, U.S.A.)

For complex educational assessments, there is an increasing use of ‘item families’, which are groups of items related to each other in some way. For example, in automatic item generation (AIG) systems, a test may consist of multiple items generated from a number of item models. However, item calibration or scoring for such an assessment requires fitting models that can take into account the dependence structure inherent among the items that belong to the same item family. Glas and van der Linden (2001) suggest a Bayesian hierarchical model to analyze data involving item families. We fit that hierarchical model using the Markov chain Monte Carlo (MCMC) algorithm. We find that formulating the MCMC algorithm in a situation like this provides little additional difficulty (compared to fitting a simple item response theory model) even with the additional complexity in the form of hierarchy in the model. We show that the model can indeed take into account the dependence structure inherent among the items and hence is an improvement over the models currently used in similar situations. We introduce the notion of the family expected response function (FERF) as a way to summarize the probability of a correct response to an item randomly generated from an item family. We also suggest a way to estimate the FERF’s. Our work is a step towards creating a tool with which one can save a significant amount of resources for tests involving item families because it may now be enough to calibrate only the item families rather than calibrating each item in the families separately.

ssinharay@ets.org

ON SELECTING THE BEST COMPONENT IN A MULTIVARIATE NORMAL POPULATION

Tumulesh Solanky
For a multivariate normal population we consider the problem of selecting the best component. Various multistage procedures are considered and their asymptotic properties are derived. Simulation studies are carried out to study the small and moderate sample performances.

tsolanky@math.uno.edu

METHODOLOGY FOR MEASURING ENVIRONMENTAL EFFICIENCY

Alpana Srivastava
Division of Agricultural Economics, IARI, New Delhi-12, India

R. S. Shrivastava (CIFRI, 24, Panna Lal Road, Allahabad, U.P. India)

This paper attempts to provide a scientifically sound methodology to measure environmental efficiency with special reference to Indian Agriculture. The adoption of new technologies, though on one hand, resulted into enhanced productivity, but on the other gave rise to problems of environmental degradation because of intensive use of modern inputs. Further, the shift to new cropping strategy (rice-wheat cropping pattern) gave rise to depletion of natural farm resources VIZ; excessive use of irrigation-decline in water table, adverse impact on soil health (soil salinity & alkalinity), increased diversification of diseases and insect pests, intensive and imbalance use of chemicals (fertilizers and pesticides)-depletion of major and micro-nutrients(declining soil fertility and nutrient mining), absorption of nutrients (disproportionate C:N ratio), pestcides residue, effect on biological microorganism (N & P fixing bacteria), water logging and so on. Of course, recent eco-friendly technological innovations like integrated pest management, biological pest control, bio-fertilizer, organic manuring and cultivation of fertility enhancing crops can be an integral part of farming strategy for correcting the impact of environmental degradation required for sustained productivity. The first types of environmental parameters are named as environmentally detrimental inputs and second type are named as environmentally recuperative inputs. Conventionally, the inputs like land, labor and capital essentially required for any production process are named as basic inputs. The phenomena of shifting to new cropping strategy affected the efficiency of the farms as a unit and also the inefficiency of the above three inputs.

The aim of this new methodology is to provide estimates of farm efficiencies due to adverse impact of environmentally detrimental inputs, positive (eco-friendly) impact of environmentally recuperative inputs and impact of basic inputs. The methodology will also provide a unique and best combination of all three types of inputs yielding enhanced productivity and sustainability of modern agriculture in India. This will ultimately provide a sustainable farming strategy to farmers and policy makers. The methodology may also proved to be a turning point to theorists for future research in environmental modeling.
In conventional specification of production function models single output is function of only variable inputs (basic inputs viz; land, labor and capital) and the response of inputs are measured by single input matrix specifying linear or non linear single equation model. The functions were usually estimated by OLS technique. Contrary to above, the model in new methodology provides specification of vector of output as a function of input matrix. The input matrix is partitioned into three components reflecting the response of basic inputs, environmentally detrimental inputs and environmentally recuperative inputs, respectively. The function thus specified is proposed to be modeled considering all the three inputs together using single equation model or simultaneous equation model. The model will be estimated by using stochastic frontier production technique. Single equation stochastic frontier production function model may be written as \( Y_n = f(X_n, Z_n, F_n \exp(V \cdot U_i)); \) where for all farms are indexed with a subscript \( i \) and all years are indexed with a subscript \( t \); \( Y_n \) is the vector output; \( X_n \) is the vector of basic inputs; \( Z_n \) is a vector of environmentally detrimental inputs and \( F_n \) is vector of environmentally recuperative inputs. Environmental efficiency for detrimental input is defined as the ratio of minimum feasible use to observed use of environmentally detrimental inputs, conditional on observed levels of the desirable output and the conventional inputs. So defined, environmental efficiency is a non-radial input-oriented measure of technical efficiency that allows for a radial reduction of the environmentally detrimental inputs. Similarly, the environmental efficiency for recuperative inputs is defined as the ratio of maximum feasible use to observed use of environmentally recuperative inputs, conditional on observed levels of the desirable output and the conventional inputs. So defined, environmental efficiency for recuperative inputs is a non-radial input-oriented measure of technical efficiency that allows for a radial increase of the environmentally recuperative inputs. Isoquant function in \( E_n \) space for the three inputs can be derived from estimated production function. The environmental efficiency for each types of inputs along with unique combination of inputs in multi-input partitioned matrix can be derived representing an eco-friendly optimal cropping strategy.

rss10@hclinfinet.com

**MULTIVARIATE ANALYSIS WITH FEWER OBSERVATIONS THAN THE DIMENSION**

**M. S. Srivastava**

University of Toronto

In this paper we develop multivariate theory for analyzing multivariate datasets with fewer observations than the dimension. Such data arise, for example, in DNA microarray where there are observations on thousands of genes but only on few patients. Methods of drawing inference such as testing of hypotheses and confidence intervals are presented. Methods for detecting outliers and imputing missing observations are also given.

A sample measure of distance between two populations is defined. The sample squared distance is used in classifying an observation into several groups by the method of minimum distance.

srivast@utstat.toronto.edu
BAYESIAN ESTIMATION OF THE SCALE OF WEIBULL DISTRIBUTION USING PRECAUTIONARY LOSS FUNCTION

R. S. Srivastava
Dept. of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur

In this paper Bayes estimators of the scale parameter of Weibull distribution have been obtained by taking quasi, inverted gamma and uniform prior distributions using the precautionary loss function. These are compared with corresponding Bayes estimators under squared error loss function.

PROCESS CONTROL IN THE PRESENCE OF SYSTEMATIC ASSIGNABLE CAUSE

J. Subramani
Department of Mathematics, Bannariamman Institute of Technology, Sathiyamangalam - 638401, Erode Dt., Tamil Nadu, India

The present paper deals with the use of systematic sampling and diagonal systematic sampling in industrial applications for on line process control. The relative performance of systematic and diagonal systematic sample means and the simple random sample mean is assessed for a natural population obtained from an automobile ancillary industry, in which there exists a linear trend among the observed values due to the tool wear. The efficiency of the different sampling schemes are measured for estimating the mean of quality characteristics; for controlling and monitoring the manufacturing process and also for the effective computation of process capability indices. Consequently it has been recommended to use the diagonal systematic sampling for the selection of random samples to assess the quality characteristics of the components produced in a short production run of engineering industries in the presence of systematic assignable cause.

drjsubramani@sify.com

ON THE CLASSES OF THREE STAGE ESTIMATION PROCEDURES WITH APPLICATION TO RELIABILITY MODELS

Neeraj Tiwari
Department of Statistics, Kumaon University Campus, Almora-263601, India

with

Ajit Chaturvedi (Department of Statistics, Delhi University, Delhi, India)

Taking into account the common distributional properties of the estimator(s) of the parameter(s) to be estimated under different continuous probability models and those of nuisance parameter(s) involved therein, classes of three-stage estimation procedures are developed. The problems of constructing fixed-size confidence regions, as well as, point estimation are considered. Asymptotic properties of the proposed estimators are presented. By means of examples, we illustrate that these classes cover many estimation problems including those related to various reliability models.
LARGE SAMPLE ASYMMPTOTIC PROPERTIES OF FAMILIES OF k-CLASS ESTIMATORS IN SIMULTANEOUS EQUATION MODEL

R. Tiwari
Department of Statistics, University of Jammu, Jammu, J&K, 180006

This article discusses the large sample asymptotic properties of k-class estimators, viz., OLS, 2SLS and LIIML for the coefficients of single structural equation. For the estimation of the coefficient in a single structural equation of a complete simultaneous equations model, Nagar (1959) has considered k-class estimators and has derived large sample asymptotic approximations for their bias vector and mean squared error matrix. He has considered those k-class estimators which are consistent according to large sample asymptotic theory. Such estimators include the case of 2SLS estimator but fail to consider two popular estimators such as OLS and LIIML estimators. The present paper considers the large sample asymptotic approximations of the bias vector and mean squared error matrix for these two estimators and to make a comparative study of the performance properties of these estimators (OLS, 2SLS and LIIML) using the large sample asymptotic theory.

rjtiwari@yahoo.com

PREDICTING U.S. CANCER MORTALITY USING STATE SPACE MODELS

Ram C. Tiwari
National Cancer Institute

Accurate prediction of cancer mortality figures for the current and upcoming year are extremely essential for public health planning and evaluation. Due to delay in reporting of cause-specific mortality for the US, there is a 3-year lag between the latest year for which such figures are available and the current year. The American Cancer Society (ACS) currently predicts cancer mortality counts by first fitting a time series model with quadratic trend and autoregressive error to the past data and then projecting this model into the future. In this paper, we present an alternative to this procedure. The proposed method uses a local quadratic trend with random time-varying coefficients to model the mortality rates (as opposed to counts) and is called a State Space Model (SSM). The SSM reduces to a model similar to the current one when the error variances are assumed to be zero. The SSM generally provides better predictions, and is able to quickly adjust to sudden changes in the observed trend. The application of the ACS method and the proposed method to a number of cancer sites will be demonstrated.

tiwarir@mail.nih.gov

BIOLOGICAL OSCILLATORS IN UNISON

Sunil Tiwari
Sonoma State University
A mathematical model for synchronous firing of biological oscillators has been derived and studied. When a given oscillator fires, it pulls the others up by a fixed amount or brings them to the firing threshold, whichever is less. The main result is that for almost all initial conditions, the population evolves to a state in which all the oscillators are firing synchronously. Examples include synchronously flashing fireflies, group of women whose menstrual cycles become mutually synchronized, and crickets that chirp in unison.

Sanjeev K. Tomar
Department of Microbiology, Ch. Charan Singh University, Meerut, India

with

Ajit Chaturvedi (Department of Statistics, Delhi University, Delhi, India)

For the probabilistic model considered by Chaturvedi and Tomar (2002), the problems of estimating the powers of the parameters, reliability function and \( P\{X > Y\} \) are considered. Bayes estimators of these parametric functions are obtained under squared error loss function and general entropy loss function. A study of the performance of estimators is done.

Ved P. Tripathi
Department of Statistics, Delhi University, Delhi, India

with

Diwakar Khare

Poverty is the state or condition under which a unit (may be an individual, household, group of individuals etc.) may suffer from deficiency of certain minimum required essential provisions of life as per the norms set in a particular society. The package of this ‘Bare minimum essential provisions of life’ has however been a point of great controversy and heated debate among various social scientist particularly economists, sociologists and statisticians. This controversy has not been centered only on minimum level of requirement but there has been a controversy also on the unit, that should be subjected to the measurement of minimum requirements, say whether poverty should be measured for an individual or household. In Indian conditions the most popularly accepted unit is household but even in the case when one accepts household as an assessment unit, the question remains to be answered is whether there can be the same cut off point of minimum requirement for all odd house holds which obviously differ vigorously in their size, structure and composition? The present paper deals with this question or issue at length. The present paper reveals that the poverty cutoff point is \( > \) bound to be different for different households owing to the difference in their size, composition and nature of work or activity undertaken by the members of a particular house hold. In essence no accurate or functional poverty measurement is possible if one ignores the structure and volume/nature of activity undertaken by households. Here household structure or composition takes some vital points into consideration such as age, and sex profile.
of members of households, volume, nature or type of occupational activity etc. which are largely
responsible to determine their minimum requirements.

In this paper a sample of 180 households has been taken to reveal how minimum requirement level (or to
say cut off point) may significantly vary from one household to another on account of difference in
structure/composition and activity patterns of the household. Along with a few vital aspects of poverty
measurement in relation to household structure and activity, the present also depicts a close relationship
between the size of the household and level of poverty; an aspect bearing great significance in relation to
family planning and family welfare policies and programme of the Government of India.

BAYESIAN ESTIMATION IN STEP ACCELERATED AND GROUPED ACCELERATED LIFE TEST
MODELS

R. K. Tyagi
CIFRI, 24 Panna Lal Road, Allahabad

with

Anoop Chaturvedi and Nand Kishore Singh

The present paper considers the Bayesian analysis of Step Accelerated Life Test (SALT) and grouped
Accelerated Life Test (GALT) with life-time distribution of units following an exponential distribution. The
Bayes estimators of the parameters of the model, mean life and reliability are obtained under squared
error loss function.

rktyagi@hclinfinet.com

STATISTICAL MODELLING IN FISH POPULATION DYNAMICS

R. K. Tyagi
CIFRI, 24 Panna Lal Road, Allahabad

with

K. N. Kurup

Fishery resources are natural renewable resources, having their own dynamics. In recent years fisheries
resources are threatened with deteriorating stock position owing to indiscriminate exploitation and over
capitalization. Hence there is an urgent need to employ stock management measures to conserve the
resources and to provide equitable returns to its stakeholders. Management of fisheries needs extensive
data on biology of fishes, exploitation pattern and assessment of their impact on the resources. Presently
mathematical models are employed to describe growth, mortality, recruitment patterns, and stock
evaluation. The estimates of parameters of these models are obtained by empirical methods often not
supported by distribution theory. Thus analytical tools used in fisheries call for sound statistical
procedures and study, and have wide scope of application in studies of fish population dynamics. Though
the subject is vast, the present paper is an attempt to provide a bird’s eye view on the dimensions of the
problem, which needs earnest attention of mathematicians and statisticians.
MEASURING CHRONIC AND TRANSIENT COMPONENTS OF POVERTY: A BAYESIAN APPROACH

Kazuhiro Ueda
Nihon Fukushi University

with

Hikaru Hasegawa (Hokkaido University)

After the publication of Ravallion’s seminal work on chronic and transient poverty in 1988, wide attention is given on them. In this article, we propose a Bayesian model with unknown number of normal mixtures to measure them. Further, we present an example using the Panel Study of Income Dynamics (PSID) data to illustrate our approach.

UEDA@N-FUKUSHI.AC.JP

BAYESIAN ANALYSIS FOR OUTLIERS IN SURVEY SAMPLING

N. K. Unnikrishnan
Assistant General Manager, Dept. of External Investments and Operations, Central Office Building, Reserve Bank of India

Outliers in survey sampling have some specific features. A few points which separate outliers in survey sampling distinct from that of parametric inference are (i) robustification against outliers may not be always meaningful; (ii) outliers may correspond to ‘hot spots’; (iii) the role of prior information. Besides, there can be observed as well as unobserved outliers; the paper assumes that there are no unobserved outliers. We use a generalized linear model (GLM) with higher variances for the outlying units. Discrete observation models such as Poisson are treated through overdispersed GLM of Gelfand and Dalal (1990). The link function is assumed to have scale-mixtures of normal distributions as their priors. The framework covers both standard survey sampling and small area estimation problems. The numbers, as well as the set of outliers, are assumed to be unknown. Posterior joint distribution is found using there versatile jump Markov chain and Metropolis - Hastings algorithm. The basic framework can be extended to various models appropriate in survey sampling context such as double sampling and conditional autoregressive models. The method is illustrated using leukaemia patients data of Cox and Snell (1981), Scottish lip cancer data and Baltimore census data.

NKU@INDIA.COM

BAYESIAN MODEL COMPARISON/VALIDATION WHEN THE PRIORS ARE NON-INFORMATIVE

S. K. Upadhyay
Department of Statistics, Banaras Hindu University, Varanasi, India

Non-informative priors are often considered when the experimenters have either very little or no a priori information about the concerned model parameters. In such situations the usual Bayesian tool-kits for
model comparison or validation have to be modified in order to provide a valid conclusion. The present talk focuses on few important versions of Bayesian p-values and Bayes factors suitable for the situations when the priors are non-informative. Suitable illustrations are provided for few important models used in reliability context.

skupadhaya@satyam.net.in

A CLUSTERING APPROACH FOR CENTRALIZED ROUTING: ALGORITHM AND VERIFICATION
Shuchita Upadhyaya
Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, India

with
K. K. Aggarwal (Guru Gobind Singh Indraprastha University, Delhi, India) and Ashok Kumar (Kurukshetra University, Kurukshetra, India)

Topological design of a network is an important aspect related with the efficiency of routing. Complexity of centralized routing control strategies may be reduced by partitioning a network into smaller sub networks or clusters such that the routing functionality leads to efficient network performance. Thus, for centralized routing, a network is to be partitioned into clusters with each cluster having one Network routing control center (NRC). Existing hierarchical and partitional clustering are based upon creating groups consisting of elements that are close to each other with respect to the distance or parameter considered in the application. These algorithms are not appropriate and adequate as such to carry out the required partitioning of a network for a centralized hierarchical routing strategy.

This paper presents an algorithm for clustering a large network by first choosing appropriate nodes that can act as NRC’s. Subsequently, the clusters are created by associating the rest of the nodes to their nearest NRC’s. The algorithm is a modification to the NEWCLUST algorithm proposed by Dysart and Georganas for the design of multidrop teleprocessing networks with concentrators. Verification for the correctness of the algorithm is then carried out.

shuchita_bhasin@yahoo.com

L^P INEQUALITIES FOR THE POLAR DERIVATIVE OF A POLYNOMIAL
C. M. Upadhye
Department of Mathematics, Jamia Millia Islamia, New Delhi

with
K. K. Dewan (Department of Mathematics, Jamia Millia Islamia, New Delhi)

Let \( p(z) \) be a polynomial of degree and \( D_\square \{ p(z) \} \) its polar derivative govil, Hyuydinkony and Tameru [J. Math. Anal. Appl. 254, 618-626 (2001)] proved that, if \( p(z) \) has no zeros in \( |z| < 1 \), then for \( \square \geq 1 \)
and \(|a| = 1\),

\[
\left| D_n \left( e^{(d)} \right) \right|^2 \Delta \left| d \right|^{1/2} n \left( \left| d \right| \right) + 1 \right) F_1 \frac{1}{2} \frac{d}{d} \left| e^{(d)} \right|^2 \Delta \left| d \right|^{1/2} ,
\]

where \(F_1 = \frac{2}{2} \left| 1 + e^{(d)} \Delta \left| d \right|^{1/2} \right| .

In this paper we will obtain generalization of this result for the class of polynomials having no zeros in the disk \(|z| < k, k \geq 1\). We also obtain analogous inequalities for the polynomials having all their zeros in \(|z| > k \geq 1\).

MULTIPLE REGRESSION ANALYSIS FOR YIELD IMPROVEMENT IN WHEAT

Gayatri Vishwakarma

with

H. L. Sharma and Pramod Kumar (SRF in IARI)

The present study is an attempt to fit the models for biological yield and its qualitative characters in wheat. An experiment was conducted [Kumar, P. (2000)] at the area under Wheat Improvement Project - (WIP), Department of Plant Breeding and Genetics, J.N.K.V.V., Jabalpur during Rabi season, 1998-99. Multiple regression analysis serves an effective solution for improvement of wheat yield through some adequate models. Four adequate models were selected and most important character affecting the biological yield per plant was spike weight per plant. It has explained the variation itself towards wheat yield (more than 60%) followed by 1000-kernel weight (18%) and spikelets per spike (17%), while kernels per spike describe the variation less than 5%. Overall, these four factors represent the wheat yield to a satisfactory level.

WAVELET WITH A SIMPLE THRESHOLD FUNCTION AND ITS APPLICATION TO EEG DATA

Takashi Yanagawa

Graduate School of Mathematics, Kyushu University, Japan

with

Toru Fujii

Considering \(X_t = f(X(t\|)), \ldots, X(t\| d)) + e_t\), a Wavelet estimator with simple thresholding function is developed for \(f\) without assuming stationarity. Characteristic of the estimator is studied,
including the convergence rate. The estimator is applied to EEG data from an epilepsy patient. It is shown that the proposed method works well; a simple underlying dynamic is discovered in the data.

yanagawa@math.kyushu-u.ac.jp