UNIVERSITY OF LJUBLJANA, Faculty of Social Sciences,
Institute of Social Sciences – Center of Methodology and Informatics

STATISTICAL SOCIETY OF SLOVENIA

International Conference on

METHODOLOGY AND STATISTICS

PROGRAM and ABSTRACTS

September 16 – 18, 1996
Hotel BOR and Castle HRIB
Preddvor, Slovenia

International Conference on

METHODOLOGY AND STATISTICS

ABSTRACTS

September 16 – 18, 1996

Hotel BOR and Castle HRIB

Preddvor, Slovenia

Sponsored by

Statistical Society of Slovenia

Supported by

Ministry of Science and Technology of the Republic of Slovenia

Institute of Social Sciences at Faculty of Social Sciences
University of Ljubljana

Statistical Office of the Republic of Slovenia

Organizing Committee

Anuška Ferligoj
Valentina Hlebec
Anton Kramberger
Andrej Mrvar

Scientific Program Committee

Vladimir Batagelj, Slovenia
Anuška Ferligoj (chair), Slovenia
Katarina Košmelj, Slovenia
Dagmar Krebs, Germany
Tamas Rudas, Hungary
Willem E. Saris, The Netherlands
Albert Satorra, Spain

Published by:	Center of Methodology and Informatics Institute of Social Sciences at Faculty of Social Sciences University of Ljubljana, Slovenia
Edited by:	Anuška Ferligoj and Andrej Mrvar
Printed by:	Statistical Office of Republic of Slovenia

Celebrating 10 Years of the Annual Conference on Methodology and Statistics in Slovenia

The Conference on Methodology and Statistics has been held in the middle of September first for eight years in Bled and then for the last two years in Preddvor. It is becoming a tradition. The first meeting in 1986 was organized by methodologists and statisticians at the Faculty of Social Sciences at the University of Ljubljana under the sponsorship of the Section of Methodology and Statistics at the Yugoslav Sociological Association. The Section decided to have its annual meeting in Bled. Approximately 40 participants from all parts of former Yugoslavia attended each meeting. After many discussions about the future of the Bled's meeting after 1991, Slovene methodologists and statisticians decided to continue to organize this traditional meeting under the sponsorship of the Statistical Society of Slovenia and to open it to the international scientific community. The form and the time of the previous meetings was kept. In the last five years the number of conference participants has doubled. Participants are now coming from more than ten different countries.

A three day program developed for the last five conferences consists of sections devoted to a range of topics including general methodology, research design, data collection techniques, sampling techniques, measurement, data analysis techniques, modeling, and simulation. Many of the presentations at these conferences demonstrate the usefulness of mathematical and statistical methods for solving analytical problems in empirical research. The ideas for applications stem from many different disciplines, e.g., psychology, meteorology, medicine, biology, sociology, demography, political science.

At the 1987 meeting it was decided to begin to publish a series of monographies on methodology and statistics that would include the proceedings of the conferences. The series was called 'Metodološki zvezki'. Twelve issues have now appeared in the series: four monographies and eight proceedings. The last three proceedings are published in English.

The conference organizers are grateful for the active cooperation of all participants who contributed to the success of the ten conferences. Gratefully acknowledged is the help and support all these years of the Faculty of Social Sciences, the Central Statistical Office and the Ministry of Science and Technology of the Republic of Slovenia.

Anuška Ferligoj

Program

SUNDAY, September 15, 1996

17.30 – 19.00 Registration and informal meeting in Castle Hrib

18.30 Meeting of the biometricians of Croatia and Slovenia

MONDAY, September 16, 1996

8.00 - 9.30 Registration in Castle Hrib

9.30 Opening of the conference

9.45 INVITED ADDRESS: *Chair: A. Ferligoj*

Lynne Billard, University of Georgia, USA
Inference for exponential time series

10.45 BREAK

11.15 Biostatistics 1 *Chair: L. Billard*

1. **Problems in modelling binary data: A case study**
   Mushtaq Hussain and Hermann Katz (Joanneum Research Institute for Applied Statistics, Graz, Austria)
   Herwig Friedl (Technical University, Graz, Austria)

2. **Evaluation of forest soil sampling technique**
   Polona Kalan (Slovenian Forestry Institute, Ljubljana, Slovenia)
   Katarina Košmelj (University of Ljubljana, Slovenia)

12.15 BREAK

12.30 Biostatistics 2 *Chair: H. Friedl*

1. **Information extraction from free text medical data**
   Mira Hercigonja-Szekeres (POLIMEDIKA, Zagreb, Croatia)
   Josipa Kern, Vesna Jureša, Davor Ivankovič and Silvije Vuletič (University of Zagreb, Croatia)

2. **Detection of disease clusters in time**
   Janez Stare and Andrej Blejec (University of Ljubljana, Slovenia)

13.30 LUNCH

15.00 Internet and Statistics *Chair: J. Stare*

1. **Surveying Internet users**
   Matija Remec and Vasja Vehovar (University of Ljubljana, Slovenia)

2. **Method of the WWW survey**
   Zenel Batagelj and Vasja Vehovar (University of Ljubljana, Slovenia)

3. **Three dimensional data presentations on Internet**
   Andrej Mrvar and Vladimir Batagelj (University of Ljubljana, Slovenia)

4. **The impact of Internet on teaching statistics**
   Andrej Blejec (University of Ljubljana, Slovenia)

17.00 BREAK

17.30 Applications *Chair: B. Bunting*

1. **An analysis of finantial ratios as factors of a firm's excess returns**
   Dušan Mramor and Neža Mramor Kosta (University of Ljubljana, Slovenia)

2. **How many Rs is there in 'reading'?**
   "True' or 'spurious' factors in item level factoring of a reading test
   Zdenko Lapajne (Ljubljana, Slovenia)

3. **Testing the roulette wheel**
   Mihael Perman and Jože Čibej (University of Ljubljana, Slovenia)

19.30 WELCOME RECEPTION

Offered by Tomaž Banovec, director of the Statistical Office of the Republic of Slovenia

TUESDAY, September 17, 1996

9.00 INVITED ADDRESS: *Chair: A. Ferligoj*

Edwin Diday, Universite Paris9 - Dauphine and INRIA, France
Stochastic generalization in symbolic data analysis

10.00 BREAK

10.30 Data Analysis 1 *Chair: E. Diday*

1. **Distance index - Global extreme computation**
   Matevž Bren (University of Maribor, Slovenia)
   Vladimir Batagelj (University of Ljubljana)
2. **The evaluation of the sensitivity of some clustering techniques to irrelevant variables**
   Istvan Hajnal and Geert Loosveldt (University of Leuven, Belgium)
3. **Pre-specified blockmodels: Theory and application**
   Vladimir Batagelj, Anuška Ferligoj (University of Ljubljana, Slovenia)
   Patrick Doreian (University of Pittsburgh, USA)

12.00 BREAK

12.30 Rasch Models *Chair: A. Blejec*

1. **Conditional vs. marginal likelihood estimation in Rasch models**
   Mojca Trobec (Educational Research Institut, Slovenia)
   Mihael Perman (University of Ljubljana, Slovenia)
2. **Exploring population heterogeneity when measuring attitudes and imaginations by means of mixture distribution models**
   Herbert Matschinger (University of Leipzig, Germany)

13.30 LUNCH

15.00 Statistics *Chair: M. Hussain*

1. **Missing data and probability imputations in regression models**
   Georg Heinze (University of Vienna, Austria)

2. **On some numerical characteristics of discrete probability laws and samples**
   Tibor Pogany (Maritime Faculty, Rijeka, Croatia)

16.00 BREAK

16.15 LISREL Applications *Chair: G. Coenders*

1. **Changing travel-mode choice as rational choice: Results from a longitudinal intervention study**
   Peter Schmidt (ZUMA, Mannheim, Germany)

2. **A within and between groups analysis of the effects of practice and coaching, using tests for invariance in LISREL**
   Brandan Bunting and Eugene Mooney (University of Ulster at Jordanstown, North Ireland)

17.15 BREAK

17.45 Methodology *Chair: P. Schmidt*

1. **Theoretical and empirical background for a measure assessing neighbourhood**
   Juergen H.P. Hoffmeyer-Zlotnik (ZUMA, Mannheim, Germany)

2. **Typical social characteristics of neighbourhoods**
   Dagmar Krebs (Justus-Liebig University, Giessen, Germany)
   Juergen H.P. Hoffmeyer-Zlotnik (ZUMA, Mannheim, Germany)

3. **Information tools for social research:**
   The SRM-Thesaurus and the SRM-Site as sources for information on social research methodology, statistics, and computer programs
   Maja van Logchem-Pasman and G. W. Kantebeen (Erasmus University, Rotterdam, The Netherlands)

19.30 Meeting of the Board of Statistical Society of Slovenia and the Board of the journal *Statistika*

WEDNESDAY, September 18, 1996

9.00 INVITED ADDRESS: *Chair: A. Ferligoj*

Willem E. Saris, University of Amsterdam, The Netherlands
Research designs and models for the evaluation of survey questions on reliability and validity

10.00 BREAK

10.30 Reliability and Validity *Chair: W. Saris*

1. **On a measure of reliability of categorical variables**
   Konstantin Momirovič (Institut of criminological and sociological research, Yugoslavia)

2. **Defining and modeling additive and multiplicative method effects**
   Germa Coenders (University of Girona, Spain)
   Willem E. Saris (University of Amsterdam, The Netherlands)

11.30 BREAK

11.45 Survey Methodology *Chair: D. Krebs*

1. **On the identification of accquiescence in balanced sets of items: A structural equation approach**
   Jaak Billiet and McKee J. McClendon (University of Leuven, Belgium)

2. **Forbid/allow effect in the Slovenian case**
   Marijan Manoilov (University of Ljubljana, Slovenia)

   Confidentiality and non-response
   Vasja Vehovar (University of Ljubljana, Slovenia)
   Metka Zaletel (Statistical Office of Republic of Slovenia)

13.15 Closing

13.30 LUNCH

15.00 Bus departure to Vršič and Kobarid

THURSDAY, September 19, 1996

10.00 Departure to Piran to the Meeting of the International Research Group on Methodology and Comparative Survey Research

---

Abstracts

Pre-specified Blockmodels: Theory and Application

Vladimir Batagelj, Anuška Ferligoj, and Patrick Doreian

Social network analysis has attracted considerable interest from the social and behavioral science community in recent decades. Much of this interest can be attributed to the focus of social network analysis on *relationships* among social units, and on the patterns and implications of these relationships.

There are several topics which are treated in social network analysis. One of the main procedural goals of network analysis is to identify, in a given network, *clusters* (classes) of units that share structural characteristics defined in terms of a relation. The units within a cluster have the same or similar connection patterns to other units. They form a *clustering* which is a partition of the set of units. A clustering partitions also the relation into *blocks*. Each block consists of units belonging to two clusters and all arcs leading from the first cluster to the second cluster.

A *blockmodel* consists of structures obtained by identifying all units from the same cluster of the clustering. For an exact definition of a blockmodel we have to be precise also about which blocks produce an arc in the *reduced graph* and which do not. The reduced graph can be presented also by a matrix, called *image matrix*.

The goal of blockmodeling is to reduce a large, potentially incoherent network to a smaller comprehensible structure that can be interpreted more readily. Blockmodeling, as an empirical procedure, is based on the idea that units in a network can be grouped according to the extent to which they are equivalent, according to some *meaningful* definition of equivalence. In general, and without surprise, different definitions of equivalence lead to distinct partitions. An appropriate generalization of the equivalence idea is one where each block, of a particular partition, is free to conform to a different equivalence idea. This led Batagelj (1993) and Batagelj, Doreian and Ferligoj (1994) to the definition of several types of connection inside and between the clusters.

The problem of establishing a partition of a network in terms of a considered

equivalence is a special case of *clustering problem* that can be formulated as an optimization problem where the criterion function reflects the considered equivalence. One of the possible ways of constructing a criterion function that directly reflects the considered equivalence is to measure the fit of a clustering to an ideal one with perfect relations within each cluster and between clusters according to the considered equivalence.
Another view of blockmodeling is deductive in the sense of starting with a blockmodel that is specified in terms of substance prior to an analysis. Batagelj, Ferligoj and Doreian (1996) presented methods where a set of observed relations are fitted to a pre-specified blockmodel. In this case given a network, types of connections, and a reduced model, a solution can be determined which minimizes the criterion function. It is also possible to fit the network to a partial model and analyze the residual afterward, and to introduce different constraints on the model.
In the talk a more detailed description of the proposed approach to blockmodeling will be given and several examples presented.

Method of the WWW Survey
Zenel Batagelj and Vasja Vehovar

The age of Internet provides survey industry with a new and powerful tool for modern data collection: the WWW survey.
The major advantages of WWW surveys are the speed and extremely low costs of the data collection. No other computer assisted method for data collection (CATI, CAPI, CASI) can compare to this. For this very reasons we can expect expansion of this method. This is specifically true for the collection of the administrative data and for surveys of populations with high Internet coverage (institutions, companies, certain professions).
Another advantage of WWW surveys is its technology which enables some unique features. It is possible to precisely analyze the date and time the questionnaire was answered. It is also possible to observe the responding time and even the exact timing of each question. Thus, useful experiments are possible to decide which lay out is optimal. It is further possible to track respondent's email address and provide an easy follow up survey (this, of course, can raise serious confidentiality issues).
However, when used for sample surveys the WWW methodology suffers from

a serious drawbacks. Besides some standard, but minor, drawbacks of Computer Assisted Self-Interviewing (CASI) the non-response is the most severe disadvantage. The non-response may be so high that findings cannot be generalized not even to the population using Internet. The inference to the whole population is thus almost impossible as long as only a small segments are using Internet. On the other hand, the findings from WWW surveys can be representative for a very important and specific population: the actual users of the WWW who also react to it. In any case, tools to stimulate the respondents on WWW surveys are very important: email, letters, login messages, 'what's new' information, classic media, word of mouth.
In the paper all above mentioned aspects of the WWW surveys are discussed and illustrated in the case of national WWW survey in Slovenia. The survey was conducted in April 1996 and it included 1200 Internet users.

Inference for Exponential Time Series
Lynne Billard

Exponential time series are those for which the marginal distribution of the observations is exponentially distributed. After describing these models, we consider estimation of the parameters and the properties of these estimators. Because of underlying discontinuities, the usual methods cannot be applied. Instead we develop conditional least squares estimators. We also consider briefly prediction. We focus in particular on the exponential autoregressive model.

On the Identification of Acquiescence in Balanced Sets of Items Using a Structural Equation Approach
Jaak B. Billiet and McKee J. McClendon

Attitude surveys often use sets of items with identical response scales in order to construct scales consisting of multiple indicators of attitude constructs. These response scales are often written in a Likert format in which respondents are asked how strongly they agree or disagree with each attitude statement.
Among sociologists, working with samples from general heterogeneous popu-

lations, there is considerably evidence that such a response format can be susceptible for an agreeing-response bias called acquiescence. Several methods have been proposed for controlling for acquiescence, including forced choice items, split ballots, and balanced scales. In an important paper, Mirowsky & Ross (1991) showed how to specify structural models for the measurement and control of acquiescence in latent variable models. In this paper, a number of conditions are specified that must be met before concluding about the existence of the agreement component of the response style called 'acquiescence'. These conditions deal with hypotheses concerning the identification of a style factor within the context of structural models and with the stability over both, instruments and respondents. In the current state of our research, most of these conditions are fulfilled. The focus here is on the specification of a common style factor behind at least two independent theoretical concepts, each measured by a balanced set of item. The data come from a pilot study with interviews in a small random sample from the general population (N =183). It is argued that a model with two content factors and a style factor better represents the data than a model with two content factors. The proposed model is currently being replicated in a large scale survey (N = 2,100).

The Impact of Internet on Teaching Statistics
Blejec Andrej

Internet has great influence on many branches of everyday life. It is regarded as great opportunity for worldwide education. We will try to see its potential for statistical community, especially for teaching statistics.
Statistics is present on the Internet for a long time. Some list servers and discussion groups are very active for a long time. StatLib, the statistics library of programs, datasets and related material existed long before WWW emerged. Four years ago an electronic journal, The Journal of Statistical Education, was launched and distributed by E-mail. With the potentials offered by WWW and its graphical capabilities, Internet became attractive for wide audience.
In attempts to provide students the possibility of learning at their own pace, Internet can play an important role. There are some initiatives to prepare interactive educational material for statistics. Textbooks can be combined

with simulations, computational examples and test questions. Answers to the questions can be captured and analysed. Such materials are available worldwide, provided students have access to the Internet.

The Distance Index – Global Extreme Computation
Matevž Bren and Vladimir Batagelj

In the talks *Determining the Distance Index* I and II we had at OSDA-95 and IFCS-96 conferences we presented

- a computational approach to approximate the distance index with local optimization procedure and
- concrete results of distance index approximation for a list of 16 standard dissimilarities on binary vectors and for some usual nonlinear transformations of these dissimilarities.

Here we will present an analytic solution of this problem, the method for the distance index computation and concrete results of distance index computation for some well-known dissimilarity measures on binary vectors.

A Within and Between Groups Analysis of the Effects of Practice and Coaching, Using Tests for Invariance in LISREL
Brendan Bunting and Eugene Mooney

Within an experimental context this study examines the effects of test familiarisation/practice and coaching. In all, seven parallel forms of a test are responded to by school children (aged 10-11 years) under two conditions. One group received three hours of coaching prior to the administration of the tests ($n = 311$). The other group only received coaching after the third test ($n = 241$). The first five tests were administered over a two week period and the last two some nine months later. The analysis was conducted using a latent variable system within the context of a step-down MANOVA model. This allows for an examination of the psychometric properties of the measures, hence, the possibility of testing assumptions which are usually left unstated in the traditional use of the MANOVA model. The effects of familiarisation/practice did not produce a significant change in the means. Coaching for a period of three hours did produce a statistically significant shift in the means, though the individuals maintained their rank order. The effects of sustained coaching over a period of nine months is shown to be substantial.

Defining and Modeling Additive and Multiplicative Method Effects
Germa Coenders and Willem E. Saris

The paper gives alternatives to Campbell and O'Connell's definitions of additive and multiplicative method effects. The alternative definitions can be formulated by means of constraints in the parameters of Marsh's Correlated Uniqueness (CU) Model and have two major advantages. First, they allow the researcher to test for additive and multiplicative method effects in a straightforward manner by simply testing the appropriate constraints. An illustration of these tests is given. Second, the alternative definitions are closely linked to the currently used Multitrait- Multimethod (MTMM) models. It can be shown that, for designs with three methods, a CU model with multiplicative constraints is equivalent to Browne's Direct Product model.

Stochastic Generalization in Symbolic Data Analysis
Edwin Diday

In standard data analysis, units are generally identified by rows of a data table where columns are attributes and each contains an unique quantitative or qualitative value. In symbolic data analysis, each cell may contain a set of values (sometime weighted), related by rules and taxonomies. These kinds of 'symbolic' data come from data bases, expert knowledge, time series and any domain where the description of a class of units is needed. For instance, a class of several different objects or the same object in an interval of time is described by such kind of unit. This unit is called 'symbolic object' if it is considered as an 'intent' (i. e. a set of properties which describe it) which 'extent' (i. e. the set of units which satisfy these properties), is calculated by a given way. A 'generalization' of a set of units is a symbolic object which extension contains this set. Several kinds of symbolic objects may be induced from data tables where each unit is described by a random variable for each attribute. In that way, standard data table my be extend to the case where each cell contains the probability distribution of each random variable. By merging the units of a class of such units, we obtain a symbolic object described by a 'stochastic generalization' of these units. First we present several mathematical results which show that 't-conorms' operators (Schweitzer and Sklar) for merging and 'capacities' (Choquet) for valuating are a natural way for describing such objects. Then, by using these tools, some basic methods of data analysis (as histograms, hierarchies, pyramids, lattices, principal component, and decision trees) may be extended in order to treat enhanced data as input and to get more explanatory results as output.

The Evaluation of the Sensitivity of some Clustering Techniques to Irrelevant Variables

Istvan Hajnal and Geert Loosveldt

In Sociology cluster analysis is often used in situations were the researcher does not have a clear theoretical view on the importance of variables in the analysis. An irrelevant variable might disguise a good clustering. Although some special purpose algorithms were proposed in the psychometric literature to deal with variable weighting and variable selection, we will focus on the classical (and widely available) algorithms. In this paper we summarize some of the results that we found in the clustering literature on the effects of irrelevant variables on the recovery in cluster analysis. Milligan (1980) found in a simulation study that all the algorithms involved were sensitive to irrelevant or noise variables. This was also found by Fowlkes (see Milligan and Cooper, 1987). In van Meter's (1984) study, on the other hand, the addition of arbitrary variables did not alter the results significantly.
We present a simulation study in which noise variables were added to a set of variables with an a priori cluster structure. The a priori cluster structure was generated with a mixture modelling framework. We included the following factors in our simulation: the number of relevant variables, the sample size in each cluster, the variance in the noise variable and the (euclidean) distance between the clusters centroids (see Hajnal and Loosveldt, 1996). These factors can affect the cluster solution. In this simulation we will evaluate the impact of these factors to some hierarchical and non hierarchical cluster al-

gorithms. We will also look whether the R-square measure (for predicting the variable from the cluster), that is typically given in k-means output, is informative for variable selection purposes (SAS Institute Inc., 1989).

Missing Data and Probability Imputation in Regression Models
Georg Heinze

This talk presents a simple way to handle missing values in categorical covariates, namely conditional probability imputation. Properties of this technique are given for various patterns of missing data in regression studies. An example shows its use in the proportional hazards model. The probability imputation technique is furthermore compared to multiple imputation and model-based approaches. It can be concluded that for certain patterns of missing data occurring typically in prognostic factor studies, the probability imputation technique has properties not inferior to more sophisticated but also more difficult to implement methods, and is outperforming standard techniques like complete case analysis or omission of covariates with missing values.

Information Extraction from Free Text Medical Data
Mira Hercigonja-Szekeres, Josipa Kern, Vesna Jureša, Davor Ivankovič, and Silvije Vuletič

Free text is the basic type of data in a GP's office. A glance at any medical record shows that it is full of words, phrases, abbreviations, very often chaotic, depending of GP's behaviour pattern: However, there is a lot of hidden but valuable information. The problem is how to find it out and how to quantify such a bulk of information and make it useful for both: patient and physician.
The aim of our study was to find a way of extracting information from free text existing in medical records which could be usable for GP's work.
The sample was consisted of 60 school children, described by data from their medical records. Some data were noticed as free text, some were coded. The free text included the data about psychomotoric development, diseases and disorders, and pre-school examination data. School children' results at the

end of the first, fifth and seventh grade of elementary school were coded as school success or failure.
Data were analysed by the software package SPAD.T for lexical analysis of textual data. Selected SPAD.T routines were applied for estimating the frequencies of lexical forms, repetition frequencies and correspondence analysis. Out of several possibilities for data analysis the correspondence analysis for tables crossing the words, from free text and results on the end of first, fifth and seventh grade of elementary school was applied.
The most remarkable feature was that the result persisting in the same success/failure group. The school failure was clearly positioned in the two-dimensional latent dimension space. The next interesting feature were the clouds of words around the success/failure category. Thus the words describing the school failure were: dyslalia, EEG, hyperactivity; strabismus, speaking disorders, pectus carinatus, myopia, Goudinough test, psychomotor impairment, otitis media, etc.
Interpretation of results was a new particularity. We were looking for proximity among individuals or groups of individuals in the space constructed of words. Sometimes, for sake of interpretation, we needed more than two dimensions and the interpretation became much more complicated. However, the reach of information, especially information hidden in textual data, could compensate all the difficulties.
As a conclusion it can be said that the correspondence analysis method adapted to text analysis can be very good tool for recognising differences in individuals and groups if such differences exist.

Theoretical and Empirical Background for a Measure Assessing Neighbourhood
Jürgen H.P. Hoffmeyer-Zlotnik

The development of a measure assessing neighbourhoods is based primarily on theory of human ecology from the Chicago school. The main insight of this theory is that social-spatial differentiation is a combination of two processes conditioning each other: first, the social differentiation of the population and second, the spatial differentiation of neighbourhoods. Therefore it should be possible to make inferences from the description of the living quarter (residential area) upon the social structure of dominant inhabitants

in the area under consideration. However, if it is possible to infer social characteristics dominant in a neighbourhood then it should also be possible, to infer characteristics of nonrespondents in a national survey.
The paper presents first, the theoretical background and operationalisation of the main variables of neighbourhood description. Second, these variables are condensed into indices describing neighbourhood and a typology of neighbourhoods is developed. Finally, the neighbourhood typologies of two different cultures - East and West Germany - are displayed.

Problems in Modelling Binary Data: A Case Study
Mushtaq Hussain, Herwig Friedl, and Hermann Katz

Activated sludge scumming is one of the most common and widespread problem in sewage treatment plants. The activated sludge should be separated from the purified water in the secondary clarifier by sedimentation in order to have a clear secondary affluent. The sludge sometimes, however, does not settle property but floats onto the surface of the clarifier making sludge recycling troublesome or even impossible.
This paper tries, with the help of logistic regression, to model the sludge scumming. For a sewage treatment plant the activated sludge can take one of two values: $AS = 0$ if the sludge settles down, and $AS = 1$ if it flows onto the water surface. $P(AS_i = 0) = 1 - \pi_i$ reflects the probability of no sludge scumming at i-th sewage treatment plan, whereas $P(AS_i = 1) = \pi_i$ the probability of sludge scumming. Associated with each waste water treatment plant we have a number of explanatory variables $X_1, \ldots, X_k$. The principal objective of the statistical analysis is to investigate the relationship between the response probability $\pi = \pi(AS)$ and the explanatory variables $X = (X_1, \ldots, X_k)$.
To investigate the relationship between the response probability $\pi(AS)$ and the covariate vector $(X_1, \ldots, X_k)$ a formal model has been constructed which is thought capable of describing the effect on π of changes in $(X_1, \ldots, X_k)$. The dependence of π on $(X_1, \ldots, X_k)$ occurs through the linear combination.
(1) $\eta = \sum_{j=1}^{k} X_j \beta_j$ for unknown coefficients $\beta_1, \ldots, \beta_k$. Unless restrictions are imposed on β we have $-\infty < \eta < \infty$. Thus to express π as the linear combination would be inconsistent with the laws of probability. This difficulty has

been avoided using a transformation $g(\pi)$ that maps the unit interval onto the whole real line. This remedy leads to a generalised linear model in which the systematic part is

(2) $g(\pi_i) = \eta_i = \sum_{j=1}^{k} x_{ij}\beta_j;\ i = 1, \ldots, n.$ As the link function $g(\pi_i)$ the logistic function (logit link) has been used:

(3) $g(\pi) = \log \frac{\pi}{1-\pi}$. As a linear logistic model we have the form

(4) $\log \frac{\pi}{1-\pi} = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k$ for the log odds of a positive response. The parameters have been estimated using Maximum Likelihood Method.

The paper describes the problems associated with the estimating the parameters and calculating the significance of the coefficients. The model gives a plausible interpretation of sludge scumming and thus an attempt to control this phenomena.

Evaluation of Forest Soil Sampling Technique

Polona Kalan and Katarina Košmelj

In Slovenia, nine plots were designed for long term studies of oak forest site. On two of them, Cigonca and Polom, sampling of forest soil was undertaken to determine chemical properties of soil. Sampling was performed as follows: Each research plot in the size of 100 m x 100 m was divided into 25 subplots of 20 m x 20 m in size. Five subplots were selected for the collection of soil samples: four corner plots and the central one. Within each selected subplot three sampling sites were chosen, the choice being based on the professional judgment for a typical forest site. Soil samples were taken from three soil layers (0 – 5 cm, 5 – 10 cm and 10 – 20 cm). Fifteen laboratory samples and five composite laboratory samples per soil layer were prepared for chemical analyses. Soil reaction (pH), total C, N, P, K, Ca, Mg, Zn and Cd were determined twice each.

The objective of statistical analysis was to evaluate the sampling methodology used. A new sampling methodology enabling to detect the changes of mean values of chemical parameters in time is required. The criteria are as follows: the estimated average value should not differ from its true value for more than 5% for pH, for more than 10% for C, N, P, K, Ca and Mg, and not for more than 20% for Zn and Cd. These values should suffice to detect the changes of mean values in time. To reach prespecified values 3-26 samples

are required, dependent on chemical parameter under study.
Statistical analysis was divided into three parts. Firstly, the data were analysed by descriptive statistics techniques and by ANOVA of nested design (subplots, sampling sites, analytical procedure).
Secondly, two models were used to evaluate tbe sampling procedure: a simple random model and a two-stage sampling model, though the sampling methodology used does not correspond completely to any of these two. The results showed no statistically significant difference between the two models. Therefore, the sampling methodology was further examined by the simple random sampling model.
For each chemical parameter under study the variability, the relative error of average value and the number of sampling sites according to the prespecified accuracy were calculated.
We propose the following sampling methodology. At each subplot of the research plot one sampling site should be selected in its centre. Inclusion/exclusion criteria for selection of sampling sites should be taken into account. In the laboratory the required number of samples per chemical parameters must to be analysed.
High costs of the proposed methodology suggest some kind of rationalisation. The natural way is to prepare composite soil samples. A number of field samples from a particular plot are mixed and homogenisated in a composite soil sample that is analysed in laboratory.
In our research composite samples were obtained by mixing the soil from three sampling sites within a subplot. Five composite samples per research plot were analysed in the laboratory. For all chemical parameters under study, the empirical results are consistent with the expected values for standard error of the mean.
Further we developed an optimisation procedure to determine the number of subplots (n) and the number of sampling sites per subplot (m) that are required for a composite sample. The procedure accounts for the prespecified precision criteria, the cost of the fieldwork, the costs of the laboratory work and the statistical characteristics of the research plot. It minimises the global expenses per chemical parameter.
The procedure was applied to our data and compared to the methodology previously described.

Typical Social Characteristics of Neighbourhoods

Dagmar Krebs and Jürgen H.P. Hoffmeyer-Zlotnik

Based on the general hypothesis that social differentiation causes spatial differentiation, residential areas of 3.000 respondents in a national sample of the Social Science Bus have been classified with respect to type of city as well as type of neighbourhoods. Respondents from different types of residential areas have been compared with respect to their socio-economic status as well as other socio-demographic characteristics. The goal was to find out if a special type of person lives in specific residential areas or even dominates them. Cluster analysis is applied to answer this question.
The theory of social-spatial differentiation states that segregated residence in social areas is mediated by a potential system of social contacts and local social networks, having impacts on the attitudes of residents. This hypothesis is tested with a small set of attitudes. Finally, we discuss the possibility to use the description and typology of residential areas as assessment of respondent-independent characteristics to allow inference about nonrespondents in a national survey.

How Many Rs is There in 'Reading'? 'True' or 'Spurious' Factors in Item Level Factoring of a Reading Test

Zdenko Lapajne

Item level factoring. Factor analysis was originally proposed to give a parsimonious and psychologically sound description of correlations between numerous continuous test scores, assumed to be measured at interval level. When computers made calculations easier, it was too often used to factorize dichotomous variables, produced by scoring responses to individual items in psychological or educational tests, as if they were continuous. The existence of spurious 'difficulty factors' has been reported from the early forties, but numerous subject-oriented researchers without much understanding of multivariate methods usually still tend to over interpret results of item level factoring. Simulations using Monte Carlo approach showed that grouping of continuous one-factor data into two or more categories and calculating sample Pearson correlation coefficients (Φ in the case of dichotomous data) nearly always produces too many factors in exploratory factor analysis when defaults built in most popular statistical packages are used. Tetrachoric coefficients have higher standard errors and require very large random samples to be estimated. Everything is possible in studies using small ad hoc samples. Categoric data require specific methods of analysis.

Reading Comprehension. Parallel with the development in methodology there is a converging consensus among (mainly Anglo-Saxon) researchers that 'reading comprehension' (defined as the ability to construct meaning from continuous prose), is a unitary ability for students between the age of about 9 and 19 years. In the late 80's there was some research showing that 'document literacy' (ability to read complex, but not continuous verbal and symbolic messages) is a separate - though correlated - ability. In the second international IEA Reading Literacy Study (Elley, 1992; Elley, Gradišar and Lapajne, 1995; Postlethwaite and Ross, 1992) it was decided to develop three scales according to the type of text:

1. Narrative (continuous texts in which the author is telling a story),
2. Expository (continuous texts designed to describe, explain, or otherwise convey factual information or opinion to the reader), and
3. Documents (structured information displays presented in the form of charts, tables, maps, graphs, lists or set of instructions).

Data, sample; and the research problem. Reading tests for 9 and 14 year old students were carefully developed in 27 countries in the frame of Rasch model, assuming: cross-cultural comparability of reading tasks, three one dimensional abilities, equally discriminating items (in IRT meaning), and no guessing to forced-choice items. Only data of Slovenian 1991 stratified sample of 9 year old students ($N = 3,300$) will be studied in this paper to question whether three scales were really necessary to report the results. Published results caused some concern between Slovenian educators because our children scored lower on the Expository scale than on the other two scales. Reading tasks such as used in the Expository scale are considered to be essential for efficient use of textbooks. Girls scored significantly higher in most countries and the difference of 0,19 standard deviation in Slovenian sample is quite high in the international comparison.

Reliability estimates and correlations. Using all 66 forced-choice items and dichotomous scoring, Cronbach alpha reliability estimates are 0.928 for 22 item Narrative scale, 0.826 for 21 item Expository scale, and 0.794 for 23 item Documents scale. Total score of all 66 items has much higher reliability estimate of 0.928. Correlations between proposed scales are very high taking reliability in mind, and they do not support strongly the hypothesis of three separate constructs: 0.619 between Expository and Documents, 0.766 between Expository and Narrative, and 0.680 between Documents and Narrative.

Classical item analysis and construct validity. Corrected item-total correlations for three scales were compared to correlations of the item score with the other two scales. Assuming construct validity, the item should correlate much higher with its proposed scale than with the other two. The results do not support the construct validity of Expository and Narrative scale, but give some support for the Documents scale.

Exploratory factor analysis. A matrix of $\frac{66*(66-1)}{2} = 2,145$ tetrachoric correlations was estimated between item scores. Student's gender was added to the variable list to get some insight in sources of gender difference at the level of items. Principal component solution (with 1 in the diagonal of communality estimates) was used, and four factors were retained for OBLIMIN rotation using Cattel's scree test and parameter $\delta = 0$. Although the first principal component with eigenvalue of 20.7 explains as much as 30.9% of total variance ($\Theta = 0.966$), the next three oblique factors can be interpreted as: (2) serial effect (probably connected with student's slowness, motivation and/or tiredness under mildly speeded conditions), (3) very slight indication for a Documents factor, and (4) gender related factor. Since text-items relations prevented items to be ordered strictly by difficulty, item's serial position within testing time is a better predictor of the second factor pattern loading ($r = -0.89$) than its difficulty (measured with p value, $r = 0.54$); while serial position and difficulty was significantly correlated ($r = -0.52$).

Conclusions. The second and the third oblique factor are strongly correlated with the first: -0.415 and 0.533 respectively. It is recommended to report only total test scores for Slovenian 9-year old students for most practical purposes. Relatively low mean score on the Expository scale in a large sample international comparison is a reliable result (for those believing that cross-cultural comparisons of reading comprehension tests are possible, useful and fair). Slight changes in the reading diet for the beginning Slovenian students cannot do much harm. Further insight in the second serial factor could be obtained only with experimental manipulation of testing time, order of items in the test, and measuring some additional variables.

Information Tools for Social Research: The SRM-Thesaurus and the SRM-Site as Sources for Information on Social Research Methodology, Statistics, and Computer Programs
Maja van Logchem-Pasman and G. W. Kantebeen

As Dutch academic Expert Centre, the SRM (Social Research Methodology) Documentation Centre in Rotterdam has focused on documentation and information concerning social research methodology, statistical analysis and computer programs for more than 25 years. During these years a bibliographic database has been built which is accessible by means of the 'SRM-Thesaurus of Social Research Methodology'.
This year a completely revised and updated version of this Thesaurus was published. An overview is given of new topics included in the various categories of the SRM-Thesaurus, following the successive stages of a research design.
Up to now the information services provided by the SRM-Documentation Centre were concentrated on the distribution of bibliographic information on CD-ROM and in print. These 'traditional' information tasks of SRM have to be extended to the field of electronic information services, in order to guarantee the function of SRM as an Expert Centre for social scientific researchers.
For individual scientists it is hardly possible to keep informed in view of the enormous supply of information on the internet. For this reason the SRM-staff decided to enlarge their task as scientific information providers with a new service: 'the SRM-Site for Research, Methodology and Statistics in the Social and Behavioural Sciences'. In short the SRM-Site, collects, selects and presents electronic information in the following domains:
1. General Resources and Directories for Social and Behavioural Sciences.
2. Researchers Collaboration: Associations, Conferences and Discussion Lists.
3. Research Programs, Current Research, and Research Institutes/Organizations.
4. Data-Archives and Data Delivery Services.
5. Teaching Research Methodology and Statistics: Courseware, Courses, and Summerschools.
6. Literature on Research Methodology and Statistics: Bibliographic Databases, Publishers, and Electronic Publishing.

7. Software for Social and Behavioural Research and Computer Programming.
Besides this ordering a link will be made to the 'SRM-Thesaurus of Social Research Methodology' by adding one or more concepts of the Thesaurus in accordance to the methodological information provided. Besides the storage and presentation of primary information, the great value of the SRM-Site is laid in enlarging the visibility and connectivity to the many existing and forthcoming WWW-home pages.
With the design and the start of the SRM-Site on the Internet, the SRM-Documentation Centre offers a shopping centre of bibliographic and electronic information for researchers, methodologists and students in the Social and Behavioural Sciences.

The Forbid/Allow Effect in the Slovenian Case
Marijan Manoilov

The forbid/allow wording effect is one of the largest ever discovered (Shuman and Presser, 1981:296) and its evidence has been present since 1940's. An early experiment by Rugg (1941) showed that small changes in wording, even in questions about the same topic, could produce dramatically different effects. He asked matched samples of respondents one of two questions: 'Do you think the United States should allow public speeches against democracy?' or 'Do you think the United States should forbid speeches against democracy?' (Sudman and others, 1996:10). When the question was one of 'allowing' public speeches, 21 percent of respondents supported free speech. But when the question was one of 'forbidding', 39 percent supported free speech.
In the 1970 Shuman and Presser replicated exactly the original forbid-allow experiment. 'Americans are much more willing to not allow speeches than they are to forbid them, although the two actions would appear to be logically equivalent in this instance. Readers who doubt the equivalence should ask themselves whether it is possible both to allow and forbid a speech at the same time, or to not allow and yet not forbid it in practice' (Schuman and Presser, 1981:276).
'Since the original study in 1940's, the effect has been successfully replicated over a wide range of issues producing differences of up to 30 percentage

points' (Hippler and Schwarz, 1986:167). Hippler and Schwarz gave also evidence, that the forbid-allow effect is stable across cultures and languages by replicating the effect in Germany.
It is not necessary that the direction of forbid/allow asymmetry is the same as reported in Rugg's experiment. There is evidence, that if applied to some issues it can be the opposite. For example Bishop and others (1988:334) reported that 'contrary to the results of all previous research on the forbid/allow effect, the respondents were significantly more likely to say that something, such as smoking in public places, should be forbidden than they were to say that it should not be allowed'.
It is also documented, that the effect is not always stable across cultures if applied to the same issues. Previous example (Bishop and others) indicates that the forbid/allow wording experiment in the United States gave opposite results than its replication in Germany. So, the forbid/allow asymmetry is stable across cultures, but there are of course some minor exceptions.
This was one of the reasons why we tested this effect in a split-ballot experiment on the Slovene population. Our main goal was to test if the forbid/allow asymmetry is present also in the Slovene language and culture. The split-ballot's sampling frame was the Telephone Directory of Slovenia. We applied a sampling method that produced two representative subsamples IPS1 ($n1 = 500$) and IPS2 ($n2 = 1000$). Then we chose telephone interviewing (CATI) as the method of data collection.
Because our split ballot experiment was a part of a research on 'Media Access' it was related to media issues like newspapers, radio and television. The 'Media Access' was a student's project of the Faculty of Social Sciences.
We tested the forbid/allow asymmetry effect for nine different question issues. The result was that the effect occurred only to four of nine question issues. Although it occurred only to some question issues, we can say that the experiment is another confirmation of the hypothesis which states that the effect is stable over time and between cultures. Our split-ballot experiment is the demonstration that the effect (for some issues) is present also in Slovenian language and culture. The effect we noticed in our 95's split-ballot had the same direction as Rugg reported in 40's.
My talk will also give a clue, that there may be a principle which discriminates between wording changes which have the forbid/allow effect and those which have not. It may be the *existence* of the issue the question is referring to. The possible hypothesis may be that the forbid/allow asymmetry is greater

the less the question issue *exists* (in sense that it is present in respondent's memory). I deduced that hypothesis because forbid/allow asymmetry effect occurred not to all question issues but only to issues that (in my opinion) have a common background. The background is that they are referring to issues which are not implemented in (Slovene) reality nor the public opinion is aware of them.
Because this classification of question issues was unexpected, we do not have enough data to prove this hypothesis, but we can deduce from some general knowledge about the Slovene media situation, that the hypothesis is plausible but further research is necessary. However a possible interpretation which covers the Schwarz's cognitive-psychological explanation and the *existence* of the question issue will be presented.

Exploring Population Heterogeneity When Measuring Attitudes and Imaginations by Means of Mixture Distribution Models
Herbert Matschinger

Instruments employed for measuring attitudes or imaginations are frequently considered to be acceptable only if they proofed to measure the construct in one and only one dimension. If the assumption of unidimensionality holds true for a particular population, the observations can be compared by a single parameter that indicates their position on the latent continuum. In most instances researchers are heavily concerned about problems of homogeneity regarding the stimuli of the item battery. In contrary we want to tackle the problem whether or not a subsample exists for which a measurement model can be employed to achieve maximum reliability and optimal unidimensionality. Mixture distribution models assume that the observed data are a mixture of data from more than one latent population (Everit & Hand, 1981; v. Davier & Rost, 1995) and different sets of model parameters are valid for these different sets of populations. The partitioning of the population is not known in advance using any manifest characteristic of the individuals. We employ a Hybrid Rasch-latent class model (Rost 1990; Rost & von Davier 1995; Yamamoto 1989) in order to separate two major classes of observations. The first class should hold for a partial credit Rasch-model (Masters 1982), the second class should hold for a partial credit latent-class model. By means of this it is possible to separate 'scalable' respondents from respondents who are homogenous with respect to a certain response paterrn. The instrument applied should exhibit a maximum reliability for the first class of respondents. The very latent-class model is demonstrated for an instrument measuring attitudes towards psychotropic drugs, which consists of 11 5-point Likert type items. These items maintain 6 desirable and 5 non-desirable effects of pychotropic drugs. The 11 items were administered to 1282 respondents in both the old and new 'Länder' of Germany in 1993.

On a Measure of Reliability of Categorical Variables

Konstantin Momirović

Let be $E = \{e\}$ a random sample of n objects from a population P, and let be $W = \{w_j\}$ some categorical variable with finite set of m mutually exclusive categories. Let be S an indicator matrix obtained by the description of E on W, and let be J an indicator matrix obtained by the description of E on a dummy categorical variable K which randomly splits E in two equal subsamples. Let is $C = S^tSn^{-1}$ a diagonal matrix with the estimations of probabilities $p_j = p(e = w_j)$, let be $H = J^tJn^{-1}$ and let be $G = S^tJn^{-1}$ a probability matrix such that the columns of G are independent estimations of p_j based on two subsamples of E.
Then the coefficient of alienation

$$\rho = (1 - \phi^2)^{\frac{1}{2}}$$

where

$$\phi = (trace(GH^{-1}G^tC^{-1}) - 1)^{\frac{1}{2}}$$

is the single non trivial canonical correlation between W and K is the measure of similarity of estimations of probabilities p_j based on subsamples of E.
Let be $R = (r_{jk}), (j, k = 1, 2)$ the matrix with elements $r_{jj} = 1$ and $r_{12} = r_{21} = \rho$. In this experimental design the matrix $T = (t_{jk}), (j, k = 1, 2)$ with elements $t_{jk} = \rho$ is, in accordance with the classic theory of measurement, the matrix with the correlations of true scores on variable W.
Let are $\lambda = x^tRx$ and $\eta = y^tTy$ the maximal quadratic forms of R and T induced by vectors $x : x^tx = 1$ and $y : y^ty = 1$. Obviously $\lambda = 1 + \rho$ and $\eta = 2\rho$ so that from the formal definition of reliability, the coefficient

$\alpha = \frac{\eta}{\lambda} = 2\rho(1+\rho)^{-1}$ is the measure of reliability of data obtained by the description of E on W.
Since α is simple function of ϕ any test of the hypothesis $H : \alpha = \alpha^*$ can be reduced to the test of hypothesis $H : \phi = \phi^*$ with

$$\phi^* = \frac{2(1 - \alpha^{*\frac{1}{2}})}{2 - \alpha^*}$$

Because of the variable $\chi^2 = n\phi^2$ has, approximately, χ^2 distribution with $v = m - 1$ degrees of freedom, the estimation of variance of ϕ, except in trivial case $\phi = 0$, is

$$\sigma^2 = \epsilon^2 (4n\phi^2)^{-1}$$

where

$$\epsilon^2 = 4n(2\phi^2 - \phi^4)$$

is the estimation of variance of χ^2, so that the statistics

$$z = (\phi - \phi^*)\sigma^{-1}$$

is normally distributed with parameters 0 and 1.

An Analysis of Financial Ratios as Elements of a Firm's Excess Returns

Dušan Mramor and Neža Mramor Kosta

Introduction.
The purpose of this work was to construct a model which could be used for analysis, and in conditions of a stable economy also for prediction, of how far from optimal the business policy of a company is from the perspective of an equity investor.
Elements of business policy are measured by financial ratios, and an optimal business policy is reflected in an above average risk adjusted rate of return.
The underlying hypothesis is that the optimal combination of business decisions varies from industry to industry and that it is for at least a certain time period stable within an industry in a stable economy.
The model is based on cross sectional analysis and could be used in general as an alternative to evaluating firms on the basis of estimating future returns

by time series analysis. It is therefore especially useful in situations, where time series analysis is not possible. This is true for example in countries in transition like Slovenia where, due to changes in the economic and accounting systems, data between years is not comparable.

A description of the model.

The model is derived with a multiple regression of the risk adjusted rate of return on equity ($\bar{r}_e$) on a set of standard financial ratios. When companies' market rates of return are available and pricing of shares is efficient the market rate of return is already adjusted for risk and can be used directly. When market data are not available accounting return on equity (ROE) is used as a basic measure of return but it is necessary to correct this measure by a factor which includes the riskiness of the equity. We use a method for adjustment which uses accounting data so that it is applicable when market data are not available. [1]

Standard financial ratios usually used in financial analysis are functionally related in nonlinear and sometimes quite complicated ways. Therefore, it is necessary to identify a basic set of ratios which are (more or less) independent and still describe the relevant elements of business policy.

On the basis of the economic definitions as well as the values and distributions of these remaining ratios it was assumed that the dependency of $\bar{r}_e$ on an individual ratio is not always linear. Therefore, transformations of the basic variables are introduced in the model. The transformations are chosen from a set of functions which are as simple as possible and which describe the characteristic behaviors of $\bar{r}^e$ in terms of individual ratios.

Testing. The model was tested on the accounting data for Slovenian companies for the years 1992 and 1994. [2] Eleven industries were selected that had more than 40 companies with more than 10 employees, the smallest number of companies being 77, average 113 and the highest 191.

Adjusting ROE. The formula $\Delta ROE_i = -(\beta_i^e - 1)(r_m - r_f)$, where β_i^e is a relative measure for the i-th company's risk of equity, $(\beta_i^e - 1)$ is the difference between the company's and the average risk, and $(r_m - r_f)$ is the market risk premium is used to obtain the adjusted ROE_i for risk. Market and

[1] The method is developed in D. Mramor, N. Garrod: "Theoretical Relationship between Systematic Risk and Accounting Variables", 1996. Preprint.

[2] Because of turbulent changes in the economy (independence economic effects, privatization, changes in accounting system etc.) we could not use the data for 1993 and 1995.

accounting industry data for U.S. and Great Britain were used together with accounting data of each Slovenian company to obtain an estimate for β_i^e.

Defining the basic set of ratios. After several iterations (in which exact computations, covariance analysis, and theoretical considerations were used) a set of 16 ratios which are relatively independent and sufficiently informative was selected. Among these the independent variables of the model were chosen within each industry separately.

Detecting and eliminating outliers. In analysis of financial data the problem of outliers is normally present. In Slovenia this problem is much bigger due to a rapidly changing situation which forces a significant number of firms into extreme situations. We tried several methods for detecting outliers, and obtained best results on the basis of simple reasoning - any value which is beyond a logical economic interpretation is labelled as an outlier. Outliers were eliminated by truncating, since other methods seemed less appropriate in cross sectional analysis with a relatively big number of outliers in the population.

Selecting transformations. Taking each ratio as the independent variable x, linear and quadratic regression of r^e on x, and (where applicable) on $log(x)$, linear regression on $\sqrt{x}$ and on the combination $1/x$, $\log(x)/x$ was tested. The last two transformations proved to be insignificant in all cases, so one of the following transformation functions was identified as best for each ratio: $a + bx$, $a + bx + cx^2$, $a + b\log(x)$, or $a + b\log(x) + c\log^2(x)$. For each ratio one transformation function was chosen for all industries.

Selecting the relevant ratios. In the last step, multiple regression on the (appropriately standardized) ratios was used to identify the most explanatory ratios in each industry separately. The ratios were grouped according to elements of business policy which they describe (liquidity, solvency, cost efficiency, asset efficiency, etc.), and the tendency in the final choice was to include only one ratio from each group (in this way the problem of multicollinearity and nonlinear functional dependency was substantially reduced). Finally the relevant ratios were chosen for each industry separately according to the highest and most significant values of the coefficients obtained.

Results and comments. As expected, the model indicates that the ratios which contribute to excess return on equity as well as their impacts are significantly different across industries. In some cases unexpected tendencies appeared but mostly in statistically less significant ratios. These might be due on one hand to the unstable situation in Slovenia which is reflected in the data, and on the other hand on dependencies between the ratios which have not been eliminated completely. In the current stage of the testing, the model is unstable between years. Also this is probably a consequence of the unstable situation in the Slovenian economy where the testing was performed. We plan to test the stability and the predictive power on data in a stable economy.

Three Dimensional Data Presentations on Internet
Andrej Mrvar and Vladimir Batagelj

VRML (Virtual Reality Modeling Language) and freely available browsers for it enable three dimensional presentations to become very popular also on personal computers.
We can use VRML to describe 3D objects, for example multivariate data in data analysis, 3D graphs in mathamatical graphics and chemistry...
In a presentation of multivariate data we can use several VRML elements: position (x, y, z), color, size, shape (sphere, cube, cone, cylinder, plane, ...), pattern (texture), direction, text. Beside these we can include some other effects, for example lights (light sources, shadowing, transparency, reflections,...).
One of the most important features of VRML is the possibility of moving in the obtained space (called scene). This possibility is something special if we compare it to the classical graphical presentations.
In the paper we present some examples of 3D presentations of multivariate data.
The draft of the paper in HTML/VRML can be found on:
`http://vlado.mat.uni-lj.si/vrml/dsi.96/defengl.htm`

Testing the Roulette Wheel
Mihael Perman and Jože Čibej

With more and more statistical research relying on simulation random number generation is a sprawling industry. This talk, however, will not talk about new methods of generating random numbers with a given distribution but focus on a rather oldfashioned random number generator: the roulette wheel.

In principle, the roulette wheel should produce random numbers that are uniformly distributed on the set $\{0, 1, 2, ..., 36\}$ but small mechanical flaws or wilful manipulation can cause the roulette wheel to deviate from the uniform distribution.

Gambling houses (in this case) usually have equipment that automatically records long series of outcomes on their roulette wheels and it is their desire to have statistical methods to monitor the outcomes and test for possible deviations from uniformity. The reasons are twofold: (1) gambling houses use a human weakness for potentially hefty profits. As the addiction is bad enough (or even disastrous for individuals like for example my great–great–grand father) one should make sure that the odds offered really are what they are advertised to be. (2) according to gambling house operators there are highly sophisticated gangs of professional gamblers who are able to detect small deviations from the advertised odds and use betting coalitions to take advantage of them. Which of the two reasons may be the true motivation for the casinos is left to the reader to decide.

In mathematical notation the problem is the following: one observes i.i.d. random variables $X_1, X_2, \ldots$ taking values in the set $\{0, 1, 2, \ldots, 36\}$ and one wishes to test whether the observations come from a uniform distribution on the above set. One idea that immediately comes to mind is to use the χ^2–test with 36 degrees of freedom. This test is in fact based on the observation that the vector $\sqrt{n}(\hat{F}_0^n - F_0, \ldots, \hat{F}_{36}^n - F_{36})$ where the F_i are the expected relative frequencies (in our case $1/37$) and the $\hat{F}_i^n$ are the observed relative frequencies in n spins of the roulette wheel converges weakly to a multivariate normal distribution. One does, however, need to ask a few practical questions:

(1) How fast is the convergence to the multivariate normal distribution? Is a series of a few hundred observations enough to apply normal approximation?

(2) What is the power of the χ^2–test? What function of the alternative hypotesis is the power? How many observations would we need?

(3) What other tests could one use?

In the talk I will attempt to answer the above questions using a mixture of theoretical results and simulation. In particular, I will discuss the quality of multivariate normal approximation, the power of the χ^2–test, and compare it to the power of possible alternatives. In the end I will present a collection of test statistics that has actually been chosen to monitor the wheels in real life casinos.

On Some Numerical Characteristics of Discrete Probability Laws and Samples
Tibor Pogany

Let us consider a finite real sequence $a = (a_i)_{i=1}^{i=n}$. Then we have an inequality by Mitrinović (compare Problem *E 2032, Amer. Math. Monthly 76 (1996)*, pp. 691-692), reads as follows

$$\delta^2(a) := \min_{1\leq j<k\leq n} (a_k - a_j)^2 \leq \frac{12}{n(n^2-1)} \sum_{j=1}^{j=n} a_j^2$$

This inequality, combined with the Gatti-Birnbaum inequality (see e.g. in M. de Novellis, Some applications and developments of the Gatti-Birnbaum inequality, *Metron 19 (1958)* pp. 245-247) will be the main tool in deriving certain estimates involving the variance and the mean of a discrete, finite probability law. Moreover an upper bound is given for the sample variance observing just the minimal distance between two consecutive sample values. Denote during the paper $\overline{X}_n$, $\overline{S}_n^2$ the sample mean, the sample variance and $\mathbf{E}X$, $\mathbf{D}X$ the mean value and the variance of a random variable X respectively.
Now, we list the main results of the paper.

Theorem 1 *Consider the discrete random variable* $X \sim \mathbf{P}\{X = x_n\} = p_k, k = 1, \ldots, n$. *Then it is*

$$\min_{1\leq j<k\leq n} (x_k p_k - x_j p_j + (p_j - p_k)\mathbf{E}X)^2 \leq \frac{12\mathbf{D}X}{n(n^2-1)}.$$

Theorem 2 *Consider a real sample* $x = (x_1, \ldots, x_n)$. *Then we have the following upper bound estimate of the sample variance*

$$\delta^2(x) \leq \frac{12}{n^2-1}\overline{S}_n^2.$$

When x is nonnegative then it holds

$$\sum_{k=1}^{k=n} |x_k - \overline{X}_n| > \delta(x)\sqrt{\frac{n+1}{6n}}.$$

Finally certain examples are discussed with the special attention to the binomially distributed $\mathcal{B}(n;p)$ $p \in (0,1)$ random variable.

Surveying Internet Users

Matija Remec and Vasja Vehovar

The use of the Internet is expanding very rapidly. The segment of population using Internet is very specific and it is becoming larger and larger. However, when measuring this segment we face some serious difficulties. When surveying Internet users all the standard methodological problems re-appear, however, they take a very specific form. The paper concentrates on the following issues:

- The problem of the operationalization of the key terms arises from the technology which is relatively sophisticated. The respondents may understand it in a very specific way. The final wording of the questions is thus extremely sensible to any changing of the meanings.

- The issue of non-response is specifically dramatic because Internet users are only a small and specific segment of the population. There exist a strong evidence that non-response behavior is correlated to the substantive questions.

- The international comparisons suffer from the specific national settings of the Internet providers. The comparisons also suffer from the rapid growth. For example, the two digit monthly growth of the number of Inernet users can interfere with a mail survey employing full TDM method.

The paper discusses all above issues in the case of RIS (Research on Internet in Slovenia) project. The project screened 7000 households and surveyed 500 Internet users (telephone survey). Additionally to this some 5000 companies and 800 institutions were surveyed by mail. The Internet users were also invited to participated on WWW survey. The key questions on Internet use were incorporated also in all major social and marketing surveys. Surprisingly, the telephone surveys, probability sampling surveys and quota surveys gave similar estimates of the proportion of the population using Internet.

Research Designs and Models for the Evaluation of Survey Questions on Reliability and Validity
Willem E. Saris

In designing questionnaires many decisions have to be made in order to formulate a specific question: it has to be decided if one uses an long or short introduction, a statement or a question, a don't know filter or not etc. Many studies have shown that such choices are not arbitrary but effect the results which are obtained with these questions. Therefore it is important to know the effects of the different options and of the combinations of the different options. In the literature different procedures can be found to evaluate the consequences of the different choices. For example Statistical offices often use re-interviewing designs to evaluate the quality of their questions. Schuman and Presser and Billiet et al have published many splitt ballot experiments showing the effects of different component, Alwin and Crosnick have published a study using panel studies and a quasi simplex model to evaluate the reliability of survey questions. Molenaar has studied the differences in response distributions due to similar choices. Finally Andrwews,Andrews and Rodgers, Koltringer, Scherpenzeel have published studies using the MTMM research design and model to evaluate the validity and reliability of questions.Each of these approaches has its own assumptions which are not immediately obvious but play a role in the results which are obtained.
In this presentation I want to show what the assumptions are and how severe they are. I will also present results of an experiment which was designed to compare the different methods.

Changing Travel-Mode Choice as Rational Choice: Results from a Longitudinal Intervention Study
Peter Schmidt

The explanation of ecological-related behavior by different versions of the rational choice approach becomes increasingly popular (see for example Diekmann/Franzen 1994). The use of this approach implies that activities like energy consumption, recycling of old glass or the use of public transportation is determined by preferences and restrictions. This presupposes that ecological related behavior is both goal oriented and consistent.

The abstract core of the basic model has not included a proposal for a measurement model e.g. operationalization of these concepts, but has left this to the individual applications and researchers. This may be one of the reasons that lead to the following statement of Green and Shapiro (1994) in their monograph on Pathologies of Rational Choice Theory – a Critique of Applications in Political Sciences: We focused our attention, therefore, on the empirical rational choice literatures and gradually came to the view that exceedingly little has been learned. Part of the difficulty stems from the sheer paucity of empirical applications; proponents of rational choice seem to be most interested in theory elaboration, leaving for later, or others the messy business of empirical testing.
In the first part we deal with the concept of theory driven evaluation, its relation to Intervention and the concept of bridge assumptions between conceptual theory and action theory.This is followed by a concise description of the theory of planned behavior as a rational actor theory. The concepts of this theory are then confronted with the concepts of the rational choice approach. In the third subsection we show how the intervention which we study that is the introduction of the semester-ticket in a region in Germany can be used to formulate a set of bridge assumptions between conceptual theory and action theory.
Section 3 contains information about our design, the sample including panel attrition and a short description of the measurement instruments. This is followed by the most interesting descriptive results referring to the change in the modal-split of traffic mode over the two time points before and after intervention and t-tests for changes in the means over time in the measured variables.
Section 5 deals with the specification, test and evaluation of a structural equation model for the two-wave panel study including latent means. We demonstrate in detail how the conceptual theory, the bridge assumptions and the measurement theory are transformed in the respective statistical model and which a- priori restrictions follow from that for the model to be tested. Then we discuss the global fit of the model, detailed fit measures and the interpretation of parameters. Finally we summarize our main findings and draw the the most relevant conclusions.

Detection of Disease Clusters in Time

Janez Stare and Andrej Blejec

A so called 'scan statistic' has been widely used in epidemiology for detection of clusters of disease in time. The statistic evaluates whether an apparent cluster could be due to chance. Its use requires that a window of length l is moved along the time axis and the probability of observing a cluster of a certain size is calculated. This statistic can therefore detect epidemics of certain size, but can not detect repetitions of smaller epidemics.

Being faced with the problem of deciding whether cases of newborns with harelips occur in small clusters we have decided to use a different (Monte Carlo) approach to the problem, which we here briefly describe:

1. define some measure of closeness (distance for example)
2. calculate this measure on our data
3. generate samples of the same size, distributed according to the null (uniform in our case)
4. calculate the measure on this samples, estimate the empirical distribution and see where the value of the measure calculated in 2. lies in this distribution

The most important part seems to be 1., choosing a suitable distance. For example, one can calculate the distance of every point to its closest point, or to its k closest points.

We will present results of experiments with different measures and comparison of our results to those obtained by the scan statistic.

Conditional Vs. Marginal Likelihood Estimation in Rasch Models

Mojca Trobec and Mihael Perman

Rasch models have become the standard statistical tool to model binary data that typically arise as responses to test questions (usually referred to as items) by students (usually referred to as persons). Typically not all persons would respond to all questions so the data can be thought of as an incomplete matrix of 0's and 1's. Rasch models have many properties that have contributed to their widespread use in testing situations and they have proved to be the right measurement vehicle for the estimation and comparison of performance of different groups on the same test.

One example where the evaluation of data is based on Rasch models and its generalisations is the Third International Mathematics and Science Study (TIMSS) which is an assessment of two curriculum areas in more than 45 countries or educational systems (including Slovenia where the testing is adminstered by the Educational Research Institute). The study is conducted under the auspices of the International Association for the Evaluation of Educational Achievement and it encompasses three distinct populations of students. Although the definitions of populations are age-based, in most countries the population 1 are 3&4-grade students, population 2 are 7&8-grade students, and population 3 are students enrolled in the final year of secondary schooling.

TIMSS seeks to measure the student achievement in the light of their attitudes, home background, the nature of the mathematics and science curricula, instructional practices of their teachers and other classroom activities. Student achievement in mathematics and science was measured through written tests requiring to select appropriate responses or solve problems and answer questions in a free-response format.

To return to Rasch models, a short description of what is being assumed in the simplest case is that the probability of a person responding to a given item correctly depends only on that person's ability and the difficulty of the item. Furthermore, we assume that given the person's ability the responses to different items are independent (the assumption is usually referred to as local independence), and that the total scores of persons are sufficient statistics for the ability parameter. These assumptions lead to a family of equivalent models of which the one now commonly referred to as the Rasch model is defined as follows: let the response of person i to item j be denoted by X_{ij}, the ability parameter of person i by β_i and the difficulty of item j by δ_j. In the simple case when all persons respond to all questions the Rasch model states that

$$P[\,X_{ij} = x_{ij}, 1 \leq i \leq n, 1 \leq j \leq m\,] = \prod_{i=1}^{n} \prod_{j=1}^{m} \frac{\exp\{(\beta_i - \delta_j)x_{ij}\}}{1 + \exp\{\beta_i - \delta_j\}}\,.$$

Here n is the number of persons being tested and m is the number of items presented to the n persons. For details see for example G. Fischer and I. Molenaar, Rasch models: Foundations, Recent Developments and Applications, Ch. 2.

There are several approaches to parameter estimation in Rasch models, particularly when the person parameters need to be estimated. One difficulty is that the above model is an example of what is called a Scott–Neyman problem in that the number of parameters increases as we sample more and more persons. For this reason the usual maximum likelihood approach fails. To circumvent this difficulty most estimation procedures first focus on estimation of item parameters.

Essentially two methods of estimation have emerged to solve the problem of the increasing number of parameters in the model:

1. The conditional likelihood is based on the built in property of Rasch models that the total scores are sufficient statistics for person parameters. Taking the conditional likelihood given the scores $X_{i+} = \sum_{j=1}^{m} X_{ij}$ of all persons eliminates the person parameters and the conditional loglikelihood is of the form

$$\log P[X_{ij} = x_{ij},\ 1 \leq i \leq n, 1 \leq j \leq m \mid X_{i+} = x_{i+}, 1 \leq i \leq n] = \\ = \sum_{j=1}^{m}(-\delta_j x_{+j}) - \sum_{t=0}^{m} n_t \ \log \gamma_t(\delta_1, \ldots, \delta_m)$$

where n_t represents the number of persons with score t, and

$$\gamma_t(\delta_1, \ldots, \delta_m) = \sum_{u_{i1}+\ldots+u_{im}=t} \exp\{-\sum_{j=1}^{m} \delta_j u_{ij}\}\ .$$

It turns out that the conditional maximum likelihood estimates (CML estimates) are consistent and asymptotically normal under mild conditions which can be expected to hold in most applications.

2. The second method to eliminate the large number of person parameters is to assume the form of the sampling distribution of the abilities and integrate out the nuisance parameters thus basing the estimation on the marginal distributions of the response vectors. The assumed form of the sampling distribution is usually a parametric family $\{f(\beta|\theta)\}$, and the estimation is carried out by maximising the marginal joint likelihood based on marginal likelihoods of the form

$$P[\,X_{ij} = x_{ij}, 1 \leq j \leq m\,] = \int \prod_{j=1}^{m} \frac{\exp\{(\beta - \delta_j)x_{ij}\}}{1 + \exp\{\beta - \delta_j\}} f(d\beta|\theta)$$

which does not involve the person parameters to obtain the marginal maximum likelihood estimates (MML).
In the talk we will review the relative merits of the two methods and present what is known about their asymptotic behaviour. While it is known that the CML estimates will virtually always be consistent, that cannot be said of the MML estimates if the "true" sampling distribution of the abilities that enter the model differs from the assumed form of this distribution. This leads naturally to the question about how much bias can be introduced if the model specification is false. As the method used in TIMSS essentially assumes a normal sampling distribution for the ability parameters we will try to answer the question about the magnitude of bias when the actual distribution of abilities in the sample differs from the normal. The approach will be based partially on asymptotic formulae and partially on simulation results.
The presentation will draw on papers by Fischer (1981), Lindsay, Clogg and Grego (1991), de Leeuw and Verhelst (1986), Follman (1988) and Lord (1983).

Confidentiality and Non-response
Vasja Vehovar and Metka Zaletel

The pattern of growing non-response can be observed in many European countries. In Slovenia the noticeable increase occurs in late 80's. It is continuing in 90's when a number of surveys expanded dramatically. Specifically sharp was the increase in the refusal rates.
The paper asks the question whether the confidentiality issues are important factors in the process of deteriorating response rates.
For this purpose the data from different surveys were analyzed.

- General attitude toward confidentiality issues and the respondent's future cooperation in the surveys were asked in the General Social Survey and Labour Force Survey.

- Specific confidentiality questions were asked the item (income) non-respondents in the Consumer Attitude Survey.

- Confidentiality issue was also raised in the Internet Survey.

The empirical analysis showed that the confidentiality concern exist. The survey topic and the agency conducting the survey were found highly correlated with the existing confidentiality concern. However, there is no clear evidence that the concern for confidentiality would decrease the co-operation in a survey. This can be partially explained with the fact that the confidentiality issues were not discussed extensively in the public yet. The awareness of the confidentiality is not particularly correlated with cooperation in the survey since no serious abuses of the survey data are known to the public. Respondents thus don't correlate confidentiality with the cooperation in the survey. Their concern for confidentiality awakens only when there would be a strong and direct request for revealing a transparently treating issue.