

NATIONS UNIES

COMMISSION ECONOMIQUE
POUR L'EUROPE

ОБЪЕДИНЕННЫЕ НАЦИИ

ЭКОНОМИЧЕСКАЯ КОМИССИЯ
ДЛЯ ЕВРОПЫ

UNITED NATIONS

ECONOMIC COMMISSION
FOR EUROPE

SEMINAIRE

СЕМИНАР

SEMINAR

STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN
STATISTICIANS



CES/SEM.35/S.III/2
17 May 1996

Original: ENGLISH

Seminar on Official Statistics - Past and Future
(Lisbon, Portugal, 25-27 September 1996)

SESSION 3: STATISTICIAN: A PROFESSION?

THE INTEGRATION OF THE STATISTICAL PROFESSION

Report submitted by the International Statistical Institute */

I. INTRODUCTION

1. Statisticians today are working mainly in three settings: in academia, in government, and in the private sector. For historical and other reasons these three groups, on the whole, have inadequate interaction. Growing specialization and sub-specialization also fosters the dispersal of statisticians. Therefore the intellectual and practical gains promised by more integration in the statistical profession have become all the more important.

2. The evolution of the statistical profession is a complex and fascinating subject. Fortunately in the last 10-15 years a number of valuable books and articles appeared which throw light on important aspects of this history. While in these works the attention is first of

*/ Prepared by Mr. Zoltan Kenessey, Director. The views expressed are those of the author.

GE.96-

all to the history of probabilistic thinking, owing to the interrelationships between academic and governmental statistics such studies are also helpful in understanding the evolution outside academia.

Historical Notes

3. The fountains of statistical development appear to be located in a number of areas. One was the interests of the state, which led to Achenwall's "Statistik" in Göttingen in 1749. Thinking about probability was another stream, ever since Christiaan Huygens work was first published in 1657 in Leiden, and that of Bernoulli (posthumously) in 1713 in Basel. For annuity and other insurance calculations probabilistic approaches were used by Jan de Witt (he was also a famous statesman) in 1671 in The Hague and by Johannes Hudde (he was also a mayor of Amsterdam) in 1653-4. Annuity calculations, of course, required some knowledge about the mortality of the population by age as analyzed by John Graunt in 1662 in London. The shortage of statistical information about the subject prompted Edmund Halley to obtain from Gottfried Leibnitz data from the city of Breslau for his study (1693) - (perhaps the first ever international statistical cooperation recorded in history). The names Leibnitz and Halley point to two other fountains to which statistical wisdom was very early related: mathematics and astronomy. The role of Leibnitz in mathematics and probability is too well known to need elaboration. The role of Halley (the Royal Astronomer) in a statistical endeavour is perhaps less evident for today's reader, even though the connection between astronomy and statistics continued into the 19th century, in particular in the person of Adolphe Quetelet (1796-1874). Quetelet came to statistics via astronomy (he organized and led the observatory in Belgium after his 1823 trip to the Royal Observatory in Paris). He decided to apply the probabilistic notions to the social field as well (beyond the population studies in which they were already applied).

4. However, the fountains from which various elements of statistical thinking emerged, went beyond "Statistik" (and the interests of the state), insurance, annuities, mathematics, probability, or astronomy. Important stimuli were received from philosophy as well. In the case of Quetelet, who was undoubtedly the most influential official statistician of his age (and the initiator of international statistical cooperation through the International Statistical Congresses which were held between 1853 and 1876), the influence of Victor Cousin (Paris) has been shown. [See Porter (1986), pp. 101-3]. And, last but not least, statistical efforts have been prompted by economic and social tensions. An outstanding example is the magisterial work of Sir Frederick Morton Eden, "The State of the Poor", which was published in 1796 in London, and which in 3 great volumes chronicled its subject (since the Norman Conquest) with the use of as much quantitative information as he could gather.

5. The integration of the various streams of statistical effort emerging from these many sources remained uneven. In particular, "academic statistics" and "official (government) statistics" have often remained apart. Even Adolphe Quetelet, notwithstanding his familiarity with

probability, has rarely applied such methods to social data. As Stigler (1986) formulated: "Nor was Quetelet alone in his abstention from this use of available methodology. With a few isolated and transitory exceptions, the application of probability to the measurement of uncertainty in the social sciences was unknown before the 1870's, exceedingly rare before the 1890's, and uncommon well into the twentieth century." [p. 194].

6. It is almost as if probability and mathematical statistics, on the one hand, and official statistics, on the other have historically evolved separately. Admittedly, there have been national differences even within the development of official statistics in various countries; in certain countries descriptive statistics, in others materials more amenable for "political arithmetic" (or mathematical treatment in general) may have been developed earlier. Notwithstanding these historical differences (and their decreasing relevance by our times), interactions between mathematical statisticians and official statisticians, even today, are rather uneven. It may be argued, that modern developments in mathematical statistics and probability have not penetrated official statistical activities sufficiently. Conversely, it may be argued that probabilists and mathematical statisticians have paid inadequate attention to the problems encountered by official statisticians. One may also suppose that there is some truth in both assumptions.

7. There is, apparently, a long history of this type of dichotomy in various countries as well as internationally. Indeed, even the historiography of statistics has reflected this dichotomy for a long time. Philip Kreager's penetrating review study (to which we cannot do justice at this juncture), put this matter as follows: "...scant treatment of quantification [lack of attention to data gathering aspects] reflects a division of labour that has long troubled the history of mathematical ideas. On the one hand, a common procedure is to write histories of statistical subjects in which no mathematical formulae or quantitative illustrations are used (e.g. Westergaard, Lorimer, and Greenwood)...On the other hand, the history of statistics until recently was dominated by another approach, largely the province of mathematical specialist." [Kreager, 1993] Kreager reviews, in particular, three new historical studies issued between 1975 and 1991 [Daston (1988), Hald (1990), and Hacking (1975)]. Kreager suggests that Hald among these 'qualifies as a major new event' in the direction of integrating the history of the relevant intellectual movements over the recent centuries [Kreager (1993), p. 537].

8. Kreager cites complex reasons why the history of statistical activities and statistical methods evolved on different lines, which occasionally intersected, but typically remained parallel. He suggests, that the new type of historical studies by Hacking, Daston and Hald "makes an impressive case for the fundamental role of early modern population

inquiry in the history of probability, and thereby in the rise of social mathematics." [Kraeger (1993), p. 539]. Whether his suggestion will be as fruitful as he hopes, remains to be seen. Yet the prospect of joining the intellectual evolution of statistical activities, probability, population studies, 'political arithmetic', and national statistical activities may indeed open up the possibility for a more integrated historical evaluation of all the somewhat separate developments. And a more integrated historical perspective may be very illuminating for understanding the special reasons for the 'fragmentation' of statistical fields that still persists, and provide some guidance as to the integration needs and possibilities of the statistical profession as we contemplate it today in all its diversity.

9. In short, the history of statistical thought and activity points to early interactions and historical influences coming from many important streams originating in a number of areas. Yet the history of statistics also reveals complex and early reasons for a certain lack of 'professional integration' in our discipline. New historical research explains some of the reasons for this; other causes still appear to be only partly understood. Given the intellectual and historical complexity of this matter, further research results will be needed (and encouraged) for a more complete understanding of the problem. Indeed, it would be in the best interest of statisticians active in all main areas if they asked the historians of statistical thought and of statistical activity to further examine the issue of integration (or the lack of it) in our profession.

II. OFFICIAL STATISTICS WITHIN STATISTICAL SCIENCE

10. As suggested earlier, it may be argued, that modern developments in mathematical statistics and probability have not penetrated official statistical activities sufficiently and that probabilists and mathematical statisticians have paid inadequate attention to the problems encountered by official statisticians.

11. Thirty years ago I had the opportunity to publish a paper in the Review of the International Statistical Institute with the title 'Some Questions of the Interpretation of Statistics as a Science with Special Regard to Official Statistics'. At the time I suggested "...that official statistical development programmes to a greater than necessary extent are based on intuition and traditional lines. There seem to emerge new possibilities to apply a growing measure of scientific rigour for the direction, organization and development of modern statistical services...". Notwithstanding the progress made since the 1960's, there is still some truth to that statement. At the same time, one can also hear the concerns that academic statistics has moved too far towards creating its own 'ivory tower', which may be not only a loss to the users

of new statistical approaches, the applied statisticians. After all, during the history of statistics, 'real life' problems have stimulated many important advances in statistical theory; the connections between 'theoretical' and 'applied' statistics certainly do not move along a one-way street alone. Indeed, a hierarchial perception of the different fields of knowledge is counter-productive on both intellectual and practical grounds.

12. According to the Kendall & Buckland Dictionary of Statistical Terms the definition of statistics is as follows: "Numerical data relating to an aggregate of individuals; the science of collecting, analyzing and interpreting such data". [Kendall and Buckland (1957), p. 279]. This definition as corresponding to practical experience and as a theoretically correct starting point yields help to the systematization of statistical science. I believe it is appropriate to recognize - or develop - as a part of statistical science all those fields of knowledge which from the point of view of the above cited definition of statistics can be taken into consideration. Statistics seems to be a group of sciences. It embraces statistical theory as well as the measurement, structural investigation, the study of relations of the moments characterizing the factors and the operations of stochastic processes, cybernetical systems, etc. which exist in the frame of real (concrete) populations.

13. According to my view in the thinking about official statistics it is not fully recognized that certain fields of knowledge, related to the experiences of official statistical activity should be also considered and developed as proper parts of the broader group of statistical sciences. This is corresponding not only to the cited Kendall-Buckland definition on statistics (...the science of collecting, analyzing and interpreting...numerical data relating to an aggregate of individuals...) and to the modern approach on science. It is also calling for growing scientific rigour in the development of the fields of knowledge, regarding applied statistical activities everywhere, both in government and in the private sector.

14. The general point I am making is that an inclusive interpretation of the scope of the statistical science is preferable to an exclusive approach. [Kenessey (1966) and (1995)]. All fields of application, including official and private sector statistical activities, should be acknowledged - along the many statistical applications in all other scientific disciplines - as actual or potential branches of statistical science.

15. Dispersion and cohesion in the statistical profession have important and practical ramifications and therefore deserve attention and serious study by statisticians. I suggest that dispersion is fostered and cohesion is hindered if an exclusive interpretation of statistics as a science reigns. In order to keep the potential intellectual and practical

benefits of increased interaction and strengthened cohesion in the profession, the adherence to an inclusive interpretation of statistics as a science is preferred. The research results in all relevant fields should be also acknowledged as actual or potential areas of the inclusive interpretation of statistical knowledge.

IV. SPECIALIZATION WITHIN STATISTICS

16. The growing specialization (and sub-specialization) within statistics is a well documented phenomenon. The number of publications, working groups, task forces, etc. dedicated in official statistics to the many different fields are too well known to be documented here. For the present audience, however, it may be of interest to cite the phenomenal growth of academic journals dealing with mathematical statistics, stochastics and probability. An indication of specialization may be gleaned from the appendix to this paper which shows the classification scheme used by the ISI journal "Statistical Theory and Method Abstracts".

17. Naturally, specialization, and the evolution of sub-specialties, is about as characteristic of the statistical profession as of medicine, law, economics or the physical sciences. As knowledge expands in a given area, it has become difficult or impossible for professionals to cover it entirely. Thus, restricting efforts to certain fields within an area becomes advantageous and even unavoidable. This has to do, one can say, with the fruitfulness of certain principles of economics (such as the division of labor) - however it is not the purpose of this piece to elaborate on this matter.

18. The community of statisticians, within countries and internationally, became very much aware of the needs to foster **specialization** in statistics. The International Statistical Institute (ISI), for example, established such special sections as the Bernoulli Society for Mathematical Statistics and Probability, the International Association for Official Statistics, the International Association for Statistical Computing, the International Association for Statistical Education, and the International Association of Survey Statisticians. Each of these specialized areas developed important research programs, exchanges of ideas, and publications dedicated to their problems. It seems to be true that many of the advances in modern statistics were achieved in the framework of increased and ever intensified specialization in the statistical profession.

19. Yet, it is generally recognized that the relentless advancement of specialization also produces undesirable side-effects. This has been understood in many fields of knowledge for some years. Most of the important problems are complex in nature and can be seldom solved by relying on knowledge from only one specialty or sub-specialty. Moreover, researchers have found that very fruitful areas of investigations can

emerge on the boundaries of two (or more) fields of knowledge. There are so many examples for this process, that it is superfluous to list the numerous successes such as seen by the evolution of biochemistry, astrophysics and many other fields.

20. Recognizing the additional value of integration for the statistical profession, which has so much benefitted from specialization, ISI also has been fostering increased interaction among the branches of statistics. The biennial meetings of the ISI, for example, provide opportunities for specialists to meet statisticians outside their own areas, and try to promote the interaction amongst the various groups. Also, each issue of the ISI Newsletter - which is received by all members of the ISI and the Sections - covers the activities of all ISI Sections, and provide an opportunity for specialists in one field to get news about activities not only in their own areas but in the others as well. The new editorial policy of the International Statistical Review, which is the flagship journal of the ISI, also aims at complementing the effects of ever narrower specialization by providing reviews of statistical developments of a wider interest to readers.

21. It should be noted that over the years the international statistical community recognized problems regarding both specialization and interaction in statistics. In respect of specialization the main problem appears to be inadequate international attention to certain specialized statistical activities. For example the currently operating associations (five sections) of the ISI do not necessarily provide a forum to certain groups of statisticians. This may be related to the way the specialized groups evolved within ISI: more or less by 'function' and less by 'field' of application. (It is interesting to compare, in this context, the sections of the ISI with the sections of the Royal Statistical Society, which are the following: Research; Business and Industry; Social Statistics; Official Statistics; Statistical Computing; Medical Statistics; and General Applications). Also, the numerous sections of the American Statistical Association provide a more thorough 'field' coverage than the ISI does. Within the ISI, the problems of 'coverage' have been addressed, in part, by special committees for certain fields of application, such as industry, environment, agriculture, geographic information systems. In addition, explorations have been made about the possibility of creating a 'general' section (somewhat similar to the RSS 'general applications' section) even though there have been some doubts about this endeavour.

22. Another lack in attention to specialized interests has been the earlier neglect of the needs of statisticians working in the private sector. Private sector statistical activities are historically speaking relatively new pursuits. But today it is recognized the need to provide more 'space' within ISI for statisticians active in the private sector.

In general, ISI traditionally focused more on academic and official statisticians than on those employed in the private sector. The numbers of the latter group have been growing; according to some estimates about 30% of ASA members work in the private sector. Notwithstanding the inadequacies of coverage of certain specialized areas within the ISI, it seems that the problems recognized regarding the relatively weak **interaction** among the specialized areas of statistics are equally, and perhaps more, pressing than the issues cited regarding inadequate attention to **specialization** within statistics. The difficulties of international statistics in this regard seem to be similar to the problems experienced among specialized fields of knowledge in general. The Director General of UNESCO, Federico Mayor in his article "The Role of the Social Sciences in a Changing Europe" pointed out the following in this regard: "Firstly, the interdisciplinary as well as comparative approaches must be promoted decisively and extensively. Conversely, the disciplinary, paradigmatic, epistemological and ontological fragmentation and debates, important as they are, must be down-played... The aim is to render the social sciences more operational, more relevant in terms of data and evidence and as a knowledge base for policy-making... Secondly, interdisciplinary and comparative research must be based on good quality statistical data. This is absolutely crucial for the further advancement of the social sciences." [Mayor (1992), p. 279]

23. Good as it is to read Mayor's suggestions about the use of statistics in the social sciences, one has to think that the problems he sees in the social sciences exist in statistics as well: that story is also about us ('de te fabula narratur...'). In order to be able to help all the other fields effectively, the inter-disciplinary, and indeed the intra-statistical affairs within our own increasingly sub-specialized field also require thoughtful attention.

24. Our discussion, of course, assumes that increased interaction among specialized fields of knowledge in general, and among the various areas of statistics in particular, can have beneficial results which can usefully augment the knowledge acquired in each specialty. It also assumes, that the cost of such interaction - the efforts needed for increased communication, etc. - are smaller than the benefits which may be derived. It is not easy to prove that these assumptions are valid. However, the success of many emerging fields of knowledge on the boundaries of different disciplines makes such assumptions plausible for statistics as well. Yet, if considerable net benefits are expected, what has been preventing the existing specialties in statistics from reaping the rewards of more interaction? It seems that certain technical factors, which are experienced in many other fields of knowledge are operative in the statistical fields as well such as the circumstance that the chances of professional recognition are much better within the specialties.

V. CONCLUSIONS

25. It is believed that studying the evolution of statistics provides insights into the diversity and the communality of those working in the many fields of statistics. Indeed further historical research into the subject should be encouraged.

26. The knowledge accumulated regarding official statistics should be recognized as part and parcel of statistical science. Those parts of this knowledge which even by the end of the 20th century are transmitted only orally to the new generations should become the subject of books and articles and should find their way into the literature.

27. In addition to recognizing the value of growing specialization and sub-specialization, the integration of the statistical profession should be given emphasis as well. There are numerous intellectual and practical gains promised by such an approach.

28. Last but not least the inclusive interpretation of statistics as a discipline requires not only more interaction between academic and official statisticians, but it also mandates organized interactions with the ever growing number of statisticians active in the private sector. Also, there are many "statisticians" (some of them do not think themselves being in our group of professionals) who toil in fields other than academic, government or private sector "statistics". They may be active in astronomy, psychometrics, literary studies, law, geography, physics, etc. For the first time in 1997 the Istanbul Session of the ISI will provide an organized opportunity for such persons to present their research results to the wider statistical community. We hope that both persons from the "regular" statistical circles and from these related areas will benefit from the interaction.

29. Statistics is an inherently interdisciplinary activity. We need to recognize this in the planning of our activities and in our attempts to foster the integration of statistics as a profession.

REFERENCES

Daston, L. (1988), **Classical Probability in the Enlightenment**, Princeton: Princeton University Press.

Hacking, I. (1975), **The Emergence of Probability**. Cambridge: Cambridge University Press.

Hald, A. (1990), **The History of Probability and Statistics and their Applications before 1750**. New York: John Wiley and Sons.

Kendall, G.M. & Buckland, W.R. (1957). **A Dictionary of Statistical Terms**. Published for the ISI by Oliver and Boyd, London.

Kenessey, Z. (1966) "Some Questions of the Interpretation of Statistics as a Science with Special Regard to Official Statistics". **Review of the International Statistical Institute**, Vol. 34:2, Pp.156-164.

Kenessey, Z. ed. (1995). **The Future of Statistics**. An International Perspective. Editions Voorburg.

Kreager, P. (1993), "Histories of Demography: A Review Article", **Population Studies**, 47, pp. 536-537.

Mayor, F. (1992), "The Role of the Social Sciences in a Changing Europe". **International Social Science Journal**, May Issue, p. 279.

Porter, Theodore M. (1986), **The Rise of Statistical Thinking, 1820-1900**, Princeton, Princeton University Press.

Stigler, Stephen M. (1986), **The History of Statistics**. The Measurement of Uncertainty before 1900. Cambridge, Harvard University Press.

ANNEX

CLASSIFICATION SCHEME
of Statistical Theory and Method Abstracts

00 MATHEMATICAL METHODS
00:000 General
00:010 Theory of sets, fuzzy sets
00:020 Algebraic structures
00:030 Combinatorial analysis; graph theory
00:040 Number theory
00:050 Linear algebra; matrices and vectors
00:055 Theory of functions, real analysis
00:060 Functional analysis, topological spaces
00:070 Measure theory
00:080 Potential theory
00:090 Fourier and harmonic analysis
00:100 Special functions and transforms
00:110 Functions of matrix arguments
00:120 Zonal polynomials
00:130 Functional equations
00:140 Calculus of finite differences
00:150 Optimisation
00:160 Numerical analysis
00:170 (Differential) geometry
00:900 Algorithms, computer programs and
computations
00:990 Other

01 PROBABILITY THEORY
01:000 General
01:010 Classical probability theory
01:020 Subjective probability
01:025 Bayesian approach
01:030 Geometric probability
01:040 Combinatorial problems
01:050 Measure theoretic approach
01:060 Probability on algebraic and topological
structures
01:070 Random variables, expectations and
parameters
01:071 Random vectors and matrices
01:075 Sufficiency, completeness, ancillarity
01:080 Independence
01:082 Exchangeability
01:090 Generating and characteristic functions
01:100 Transformations
01:110 Inequalities
01:120 Convergence in probability
01:130 Almost sure convergence

01:140 Convergence in distribution, limit
theorems, invariance principles
01:150 Asymptotic expansions, rate of
convergence
01:160 Sums of independent random vectors
01:170 Large deviations and extreme values
01:180 Decision theory
01:190 Information theory
01:900 Algorithms, computer programs and
computations
01:990 Other

02 PROBABILITY DISTRIBUTIONS
02:000 General
02:010 Univariate discrete distributions
02:020 Univariate continuous distributions
02:050 Multivariate discrete distributions
02:070 Multivariate continuous distributions
02:080 Extreme value distributions
02:090 Exponential families
02:100 Infinitely divisible and stable
distributions
02:110 Limit distributions
02:120 Relations between distributions,
deconvolution
02:130 Truncated and mixed distributions
02:140 Approximations and asymptotic results
02:150 Lifetime distributions
02:160 Systems of distributions (Pearson,
Johnson, etc.)
02:900 Algorithms, computer programs and
computations
02:990 Other

03 SAMPLING DISTRIBUTIONS
03:000 General
03:010 Empirical distributions and descriptive
measures
03:020 Correlation and regression coefficients,
residuals
03:030 Student's t ; F , X^2 distributions
03:040 Order statistics
03:050 Empirical moments and quantiles
03:060 Quadratic forms

- 03:070 Null distributions of test statistics
- 03:080 Non-null distributions of test statistics
- 03:085 Distribution of estimators
- 03:090 Tables of percentage points
- 03:100 Non-central distributions
- 03:110 Non-normal populations
- 03:120 Multivariate distributions
- 03:130 Limit distributions
- 03:140 Approximations and asymptotic results
- 03:150 Truncated distributions
- 03:900 Algorithms, computer programs and computations
- 03:990 Other

- 04 ESTIMATION
- 04:000 General
- 04:010 Method of maximum likelihood
- 04:015 Other parametric methods
- 04:020 Exact properties of estimators
- 04:025 Combining estimators
- 04:030 Asymptotic properties of estimators
- 04:035 Comparison of estimators
- 04:040 Simultaneous estimation
- 04:045 Estimation of distribution functions
- 04:050 Linear estimation
- 04:060 Confidence regions
- 04:062 Fiducial intervals
- 04:064 Statistical tolerance intervals, prediction intervals
- 04:070 Bayesian estimation
- 04:075 Semi-parametric estimation
- 04:080 Non-parametric estimation
- 04:081 U -statistics and R -estimators
- 04:085 Estimating equations
- 04:090 Sequential and multistage estimation
- 04:095 Change-point model and signal detection
- 04:100 Stochastic approximation
- 04:110 Robust and adaptive estimation
- 04:111 L - and M -estimators
- 04:120 Sample reuse techniques
- 04:130 Censoring
- 04:140 Outliers
- 04:150 Finite population parameter estimation
- 04:160 Estimation in multivariate problems
- 04:165 Shrinkage estimation
- 04:170 Estimation under constraints, isotonic regression
- 04:180 Density estimation and related topics
- 04:190 Statistical information and inference

- 04:900 Algorithms, computer programs and computations
- 04:990 Other

- 05 HYPOTHESIS TESTING
- 05:000 General
- 05:010 Test construction: likelihood-based approach
- 05:020 Test construction: other approaches
- 05:030 Exact properties of tests
- 05:040 Asymptotic properties of tests
- 05:045 Preliminary testing, testimators
- 05:050 Linear hypothesis and contrasts
- 05:060 Bayesian tests
- 05:070 Non-parametric tests, rank tests
- 05:075 Semi-parametric tests
- 05:080 Sequential and multistage tests
- 05:090 Robust and adaptive procedures
- 05:100 Outliers and slippage tests
- 05:105 Censoring
- 05:110 Goodness-of-fit tests
- 05:120 Multiparameter hypothesis testing
- 05:130 Ranking and selection, paired comparisons
- 05:140 Combining tests
- 05:150 Comparing tests
- 05:160 Sample reuse techniques
- 05:170 Testing under constraints, isotonic regression
- 05:900 Algorithms, computer programs and computations
- 05:990 Other

- 06 ASSOCIATION AND DEPENDENCE
- 06:000 General
- 06:005 Exploratory data analysis, diagnostics
- 06:010 Association and dependence of quantitative data
- 06:015 Rank correlation
- 06:020 Association and dependence of qualitative data
- 06:030 Categorical data analysis, contingency tables
- 06:040 Canonical analysis, canonical correlation
- 06:050 Discriminant analysis
- 06:070 Principal component analysis
- 06:080 Factor analysis
- 06:085 Correspondence analysis
- 06:090 Scaling methods

06:100 Pattern recognition and scene analysis
06:105 Image processing and remote sensing
06:120 Cluster analysis and classification
06:900 Algorithms, computer programs and computations
06:990 Other

07 REGRESSION ANALYSIS

07:000 General
07:010 Model building and selection
07:012 Analysis of residuals
07:020 Least squares methods
07:021 Maximum likelihood methods
07:030 Least absolute value methods
07:040 Orthogonal regression
07:050 Minimax regression
07:060 Ridge regression, regression over constraints
07:069 Other methods
07:070 Selection of predictors (order and number)
07:071 Prediction in regression
07:075 Influential observations, breakpoint analysis
07:080 Calibration, optimal coding
07:085 Poisson regression, overdispersion
07:090 Logistic and loglinear model, quantal response
07:095 Generalised linear models
07:100 Errors in model input variables
07:110 Missing observations
07:111 Linear models with general error distribution
07:120 Nonparametric methods, cross-validation
07:130 Bayesian and related methods
07:140 Nonlinear regression
07:142 Transformed models
07:143 Proportional hazards models
07:146 Numerical methods, splines
07:148 Logit, probit and tobit analysis
07:150 Robustness, sensitivity
07:160 Smoothing procedures
07:165 Response curve estimation
07:180 Simultaneous equations
07:190 Seemingly unrelated regressions
07:900 Algorithms, computer programs and computations
07:990 Other

08 ANALYSIS OF VARIANCE

08:000 General
08:010 Fixed effects models
08:020 Variance components models
08:030 Mixed and other models
08:040 Analysis of covariance
08:050 Non-parametric methods
08:060 Sequential and multistage methods
08:070 Robust and adaptive methods
08:080 Non-orthogonal data and missing values
08:090 Multiple comparisons; multiple decision procedures
08:100 Non-linear models
08:110 Multivariate analysis of variance
08:112 Growth curve models
08:120 Asymptotic theory
08:130 Non-normality of errors or statistical dependence
08:900 Algorithms, computer programs and computations
08:990 Other

09 SAMPLING

09:000 General
09:010 Classical methodology
09:020 Use of auxiliary information in estimation
09:025 Sampling on successive occasions
09:030 Unequal probability sampling strategies
09:040 Optimal sampling strategies
09:050 Super-population approach; robustness studies
09:060 Bayesian inference
09:070 Natural (human, animal, environmental, ecological and biological) populations
09:072 Capture-recapture methods
09:074 Areal sampling, line transect sampling
09:080 Asymptotic theory
09:090 Randomised response techniques
09:100 Non-sampling problems, non-response
09:110 Survey sampling
09:120 Market and opinion research
09:130 Quality and process control
09:140 Sampling of continuous populations; bulk sampling
09:900 Algorithms, computer programs and computations
09:990 Other

10 DESIGN OF EXPERIMENTS

10:000 General
10:005 Optimal designs
10:010 Block designs, group-divisible designs

10:015 Latin and Graeco-Latin squares, Youden designs
 10:020 Factorial designs, orthogonal arrays
 10:022 Cross-over designs
 10:024 Split-plot, nested and weighing designs
 10:030 Other designs
 10:040 Response surface designs, rotatable designs
 10:050 Repeated measurements and sequential designs
 10:060 Clinical trials and bio-assays
 10:070 Sensitivity analysis
 10:080 Mixture designs
 10:090 Screening tests
 10:900 Algorithms, computer programs and computations
 10:990 Other

11 THEORY OF STOCHASTIC PROCESSES

11:000 General
 11:010 Stationary processes
 11:012 Ergodic theory
 11:015 Auto-regressive and moving average processes
 11:020 Gaussian and related processes
 11:030 Generalised stochastic processes, e.g. white noise, Hilbert-space valued
 11:035 Stochastic geometry, stereology
 11:040 Martingales, semi- and sub-martingales, amarts
 11:050 Stochastic analysis: differential equations, integration
 11:060 Stochastic control
 11:070 Information theory: entropy, information measures, capacity of channels
 11:080 Jump processes
 11:085 Poisson processes
 11:090 Markov chains
 11:100 Markov processes in continuous time
 11:110 Semi-Markov processes
 11:115 Regenerative processes
 11:120 Renewal theory and processes
 11:130 Theory of queues
 11:140 Random walks and applications
 11:150 Brownian motion; Wiener processes
 11:155 Processes related to Brownian motion
 11:160 Diffusion and related processes
 11:170 Birth-death processes
 11:180 Branching and related processes

11:185 Local times
 11:190 Stopping times, optimal stopping problems; gambling theory
 11:200 Point processes
 11:210 Random measures
 11:220 Random fields and multiparameter processes
 11:230 Statistical mechanics type models, percolation theory, Gibbs states
 11:240 Infinitely divisible, stable, semi-stable, Lévy and self-similar processes
 11:245 Extremal processes
 11:250 Convergence of processes
 11:255 Empirical and quantile processes
 11:260 Sample path properties
 11:900 Algorithms, computer programs and computations
 11:990 Other

12 INFERENCE FOR STOCHASTIC PROCESSES

12:000 General
 12:010 Inference for Markov processes
 12:020 Inference for diffusion processes
 12:025 Inference for point processes
 12:030 Inference for other univariate processes
 12:040 Inference for other multivariate processes
 12:050 Inference for multiparameter processes
 12:060 Univariate time series: classification and identification
 12:065 Univariate time series: hypothesis testing
 12:070 Univariate time series: estimation
 12:080 Inference for multivariate time series
 12:090 Spectral analysis
 12:100 Filtering
 12:105 Seasonal adjustment
 12:110 Smoothing
 12:120 Prediction, forecasting
 12:130 Sampling
 12:140 Aliasing
 12:150 Sensibility
 12:900 Algorithms, computer programs and computations
 12:990 Other

13 OPERATIONS RESEARCH

13:000 General
 13:010 Graph theory and combinatorics
 13:020 Mathematical programming

- 13:030 Dynamic programming
 13:040 Control theory
 13:050 Transportation problems
 13:060 Scheduling
 13:070 Theory of replacement; reliability; life testing
 13:080 Queues
 13:090 Inventory, risk analysis, dams
 13:100 Game theory
 13:110 Utility theory
 13:120 Learning theory
 13:125 Searching methods, optimal search
 13:130 Simulation; Monte Carlo methods
 13:140 Mathematical models
 13:150 E-M algorithms, nearest neighbour methods
 13:160 Expert systems
 13:900 Algorithms, computer programs and computations
 13:990 Other
- 14 SPECIAL TOPICS: APPLICATIONS
- 14:000 General
 14:010 Geostatistics, physics and chemistry
 14:015 Astronomy and astrophysics
 14:020 Meteorology
 14:030 Biostatistics and biomathematics
 14:032 Genetics, eugenics
 14:035 Statistics in agriculture
 14:040 Medical and epidemiological statistics, case control studies
 14:045 Sports
 14:050 Demography, population models
 14:060 Ecology and environmental statistics
 14:070 Economics
 14:071 Index numbers
 14:075 Finance and auditing
 14:080 Actuarial mathematics
 14:090 Psychometrics
 14:095 Neural networks; artificial intelligence
 14:100 Social sciences
 14:110 Linguistics
 14:120 Engineering
 14:125 Marketing
 14:130 Survival analysis
 14:140 Data banks
 14:145 Data collection
 14:150 Computer science and automata
 14:160 Politics and voting systems
 14:170 Traffic
 14:180 Education
- 14:190 Legal & Forensic Statistics
 14:200 Management Science
 14:900 Algorithms, computer programs and computations
 14:990 Other
- 15 GENERAL FEATURES
- 15:010 General methodology and philosophy of statistics
 15:020 Teaching and training methods
 15:030 History and biographies
 15:040 Bibliographies
 15:050 Statistical tables
 15:060 Graphical methods
 15:070 Computer software
 15:080 Construction of algorithms
 15:090 Computers and operating systems
 15:100 Random number generators
 15:110 Congress proceedings and other collective works
 15:900 Algorithms, computer programs and computations
 15:990 Other