TWENTY-FIFTH GREGYNOG STATISTICAL CONF CE

PROGRAMME

FRIDAY

21 April 13.00 Lunch

14.15 Professor Iain Johnstone (Stanford/Bath)
Speeds of Estimation in Positron Emission Tomography and Related Inverse Problems

15.45 Tea

16.15 Dr Jim Smith (Warwick)
A Bayesian Model for Competitive Bidding for Contracts

19.00 Dinner

SATURDAY

22 April 08.30 Breakfast

09.30 Dr Janice Derr (Pennsylvania State/L.S.E.)
Statistical Challenges in Drug Development

11.00 Coffee

11.30 Dr Eric Smith (Virginia Polytechnic and State University/Kent)
Multispecies Microcosm and Mesocosm Experiments: Inference when there are more Variables than Data

13.00 Lunch

---------- AFTERNOON FREE (Walks, etc.) ----------

16.00 Tea

17.45 Dr David Spiegelhalter (MRC Biostatistics Unit, Cambridge)
Statistical Aspects of Medical Expert Systems

19.00 Dinner

SUNDAY

23 April 08.30 Breakfast

09.30 Mr Ian Hewitt (Unilever Research Laboratory, Port Sunlight)
Consumer Loyalty

11.00 Coffee

11.30 Professor Luis Pericchi (Simon Bolivar University, Venezuela/ Nottingham)
Robust Bayesian Credible Intervals and Prior Ignorance

13.00 Lunch

14.15 Professor Keith Ord (Pennsylvania State/L.S.E.)
Time Series Models for Multivariate Count Data

15.30 Tea

GREGYNOG STATISTICAL CONFERENCE 1989

Participants

Speakers

Mr Ian Hewitt	(Unilever Research Laboratory, Port Sunlight)	
Professor Iain Johnstone	(Stanford/Bath)	+1 guest To Sat teatime
Professor Keith Ord	(Pennsylvania State/LSE)	+2 guest
Professor Luis Pericchi	(Simon Bolivar University, Venezuela/Nottingham)	
Dr Eric Smith	(Virginia Polytechnic and State University/Kent)	+2 guests
Dr Jim Smith	(Warwick)	To Sat. teatime
Dr David Spiegelhalter	(MRC Biostatistics Unit, Cambridge)	+ 1 guest
Dr Janice Derr	(Pennsylvania State/LSE)	

Aberystwyth

Staff	Students	
Dr J G Basterfield	Mr I H A Al-Iathary *	Miss C Minett
Mr P H Jackson	Mr B M Assas *	Mr P Murray
Mr D A Jones	Mr Z M Chaudhry *	Miss E-J Raeburn
Dr J A Lane	Mr A M Jamil *	Mr M S Yahya *
Miss S G Lutkins	Mr R Kasap *	
Dr R J Owen	Mr R Lawrence	4 students

* non resident
o no food

Bangor

Staff	Students	
Dr J Y Kassab	Mr N Shani	2 students
Mr C J Whitaker	Mr M Al-Shamary	
Mr G C Morris		

Birmingham

Staff		Students	
Professor J B Copas	Dr P Davies	Mr J Kaluluma	
Professor A J Lawrance	~~Mr R L Holder~~ *		6 staff + student
Professor H E Daniels	Miss Karen Hurrell		* no overnight stay
Dr P V Bertrand			

Cardiff

Staff	Students	
Dr R Cheng	Miss T Davenport	
Dr F Dunstan	Mr M Makloof	4 students
Mr T Iles	Miss K Thornton	
Dr J Williams	Miss K Wheeler	

Swansea

Staff	Students	
Dr A Watkins	Mr D Bounds	1 student

ABSTRACTS

TIME SERIES MODELS FOR MULTIVARIATE COUNT DATA

Keith Ord (Pennsylvania State University, visiting London School of Economics)

(with Andrew Harvey & Christiano Fernandes (LSE))

Time series data sometimes consists of count data, where the number of events occurring in a unit time is recorded. Such data often follow a Poisson or similar distribution and Normal approximations are inappropriate when the mean is small. We develop a Kalman-type recursive scheme to describe the process over time and introduce a new form of bivariate distribution which enables us to extend the method to two or more series. The predictive distributions and the recursive likelihood function are then readily derived. An example, based on England vs. Scotland football games will be presented.

CONSUMER LOYALTY

Ian Hewitt (Unilever Research Laboratory, Port Sunlight)

Like most major consumer goods companies, Unilever spends large amounts of money in Market Research, to find out what consumers like about our products. For example, we may give a sample of housewives different prototype washing powders to use for two weeks, then get them to answer a questionnaire. From that we may discover that they like brand X best - but will they actually go out and buy it?

The modelling of consumer purchase behaviour has a fairly long history. The most notable work has been that of Ehrenberg and his co-workers, culminating in their 1984 paper to the RSS on the Dirichlet model. This describes the act of buying and the choice of brands in a simple model which fits the data fairly well.

However, one of the deficiencies of the model is that it predicts that successive purchases will be independent, which is quite clearly not the case. A modification of the model to incorporate some dependence leads to a natural measure of consumer loyalty, which is essentially a measure of the tendency for the consumer to carry on buying the same product.

We can estimate consumers' loyalty from the vast quantities of data at our disposal, and study some of its properties. Most important of these is that loyalty appears to be a property of the consumer, rather than a property of the brand. Consumers who always purchase the same washing powder are likely to purchase the same washing-up liquid at the same time. They are also likely to shop at the same store (store loyalty being an analagous concept measured in a similar way, and of interest to many Geographers).

ROBUST BAYESIAN CREDIBLE INTERVALS AND PRIOR IGNORANCE

Luis Pericchi (Simon Bolivar University, Venezuela, visiting Nottingham) (with Peter Walley)

In this paper we survey and compare various classes of probability densities that may be used to model weak prior information. We focus on the simple problem of constructing credible intervals for an unknown Normal mean. We distinguish two main types of models: *neighbourhood* models, which can be used to "robustify" a strict Bayesian analysis in problems where there is substantial prior information about location, and *translation-invariant* models, which are appropriate in problems where there is little prior information. The neighbourhood models examined here include examples of ε-contamination neighbourhoods (previously studied by Huber and Berger) and intervals of measures (DeRobertis and Hartigan). Of the translation-invariant models, we examine classes of conjugate priors, classes of double exponential densities, and constant odds-ratio neighbourhoods of these. We argue that a reasonable class of prior densities is not the same as a class of reasonable prior densities and we discuss various desiderata for a "reasonable" class, including coherence, near-ignorance, translation-invariance, and the dependence of inferences on sample size. The "noninformative" uniform prior is rejected because it does not satisfy the desiderata of near-ignorance and coherence. Of the seven models studied in detail, a type of interval of measures seems most suitable for robust Bayesian analysis, and a translation-invariant class of double exponential priors (or a neighbourhood of this class) is favoured for modelling little prior information.

STATISTICAL ASPECTS OF MEDICAL EXPERT SYSTEMS

David Spiegelhalter (MRC Biostatistics Unit, Cambridge)

Statistical models to be used in decision-support have been carefully developed and validated, but have only exploited clinical knowledge to a limited extent. Artificial intelligence emphasises the use of 'deep models' that contain an understanding of the disease process, and causal networks are frequently recommended as an appropriate knowledge-representation. However, handling the inevitable uncertainty on the network presents a problem and, in particular, the issue of how to combine expert opinion with information from databases has not been seriously faced.

We describe a methodology intended to handle both the propagation of evidence through the network as data on a new case is obtained, and to allow monitoring, criticism and adaptation of both quantitative and qualitative aspects of the model as the database accumulates. We take advantage of a number of alternative representations of a joint probability distribution related to graphical structure, and in particular show how certain topological changes to the assumed causal network correspond to an embedding of the initial probability assessments into a decomposable log-linear model. Straightforward algorithms then exist for absorption and propagation of evidence, suitable for programming in an object-oriented environment.

Initial imprecise probabilities in the network can be updated in the light of data using standard Bayesian reasoning, thus providing a smooth transition between expert opinion and 'hard' data. Furthermore, since a statistical system can always make probabilistic predictions about the information it will receive, predictive ability may be monitored and areas of poor performance detected. Substantial changes to either quantitative or qualitative aspects may then be indicated, although it is probably neither feasible nor desirable that this process should be completely automatic.

Clinical applications will be described.

STATISTICAL CHALLENGES IN DRUG DEVELOPMENT

Janice Derr (Pennsylvania State University, visiting London School of Economics)

In 1988/89 I spent a portion of my sabbatical activities as a consultant to the Biological Sciences unit of a major pharmaceutical company. My role was to evaluate and improve protocols for data analysis used in the early (pre-clinical) stages of drug development. I found this experience to be very stimulating. Statistical issues included robustness in linear and nonlinear models, the analysis of ordinal responses, and quality control in bioassay. My particular research focus concerned the specification of standard errors of calibration (or inverse regression) estimates for linear and nonlinear models. I discovered that this issue was important in a wide range of circumstances. A special challenge for all statistical issues was to incorporate concepts of quality control, robustness, and correctly specified standard errors into on-line data analysis systems which could be readily used by the non-statistician in the routine screening of potential drugs.