THE ISBA BULLETIN

Vol. 13 No. 1 *March 2006*

The official bulletin of the International Society for Bayesian Analysis

A MESSAGE FROM THE NEW PRESIDENT

by Alan Gelfand
ISBA President
alan@stat.duke.edu

Assuming the Presidency of a young society like ISBA brings both challenges and rewards. The rewards are obvious. We are in the midst of the most exciting time in history for Bayesian thinking. Many of us are becoming (or already are) statistical scientists, appreciating the unification of the Bayesian paradigm, enjoying the freedom that it enables with regard to exploring demanding scientific problems, exhilarated by the rich range of collaborative activities that become available, delighted by the broad acceptance within the various scientific communities, and excited about becoming broader scientists. The opportunity to play a leadership role in implementing efforts to further this evolution is special for me.

Moreover, our community itself is special. We are exceptionally talented and clever, we are passionate and enthusiastic, we are nurturing and supportive. We welcome in the most positive of ways - eagerly, nonjudgmentally, noncompetitively - asking only for a focus on advancing scientific endeavor. In particular, the various people who dedicate their time to ISBA - the members of the Executive Committee and Elected Board, the Program Council, the Publications and Journals Committees, the Constitution and Bylaws Committee, the Nominations Committee and the various Awards Committees along with the Editorial Board for Bayesian Analysis and the people who contribute to the organization of meetings, are altogether, an invaluable source of stimulation. They are providing me with ideas, directions, opportunities that I couldn't possibly muster by myself - potential for ISBA that certainly merits contemplation, airing, refining, etc. but that will surely help to move ISBA forward. ☛ Cont. in *page 2*

A MESSAGE FROM THE EDITOR

by J. Andrés Christen
jac@cimat.mx

With much pleasure I present to you this March issue of the ISBA Bulletin. I'm glad to include the message from our new president, Alan Gelfand. In the inside pages, we have a very interesting interview with Prof. Dose from the Max Planck Institute for Plasma Physics, in Germany, an article on Bayes' mathematical work, an Annotated Bibliography on some missing data problems in gene regulation and the Applications section on Bayesian Modeling on Exposure Pathways. Also we have the Student Corner section, with important information regarding postgraduate students and finally the News of the World section with upcoming events. Regarding the Student Corner, Robert Gramacy, the section's AE has just got his PhD (congratulations!) and is leaving this position. I wish to thank Robert for his enthusiastic work as AE. Any PhD student in a Bayesian area interested in taking this AE position please contact me directly. Thanks and enjoy this issue of the ISBA Bulletin!

Contents

Candidly, however, the challenges seem to exceed the rewards. These challenges are precisely what would be expected for an organization only a bit more than a decade old. How do we encourage growth? How do we want to grow? How do we create infrastructure? When and for what do we decide to pay, as opposed to relying on volunteer effort? How shall we address our financial matters? How shall we structure our dues and at what levels? What other revenue sources shall we explore? What decisions shall we make with regard to investment of our assets? How shall we devise an overall strategy for locating, organizing, planning, etc. our international meetings? Again, many of these issues will be discussed this year and will be topics for years to come. In this regard, I urge input from the membership to address these questions not just to assist me in my one-year journey but for the long-term benefit of the society. ISBA is still small enough to listen to your suggestions and to be responsive.

Finally, I feel very fortunate to be President during 2006 since it welcomes the Eighth Valencia (now Valencia/ISBA) world meeting. Not only is this a time to look forward to this wonderful gathering but also to reflect upon how far we have come since the first Valencia meeting in 1979. Though some may raise concern over the pace of our dramatic growth, we should also bask in the international embracing of what we know is so obviously the right path for Statistics. And, I do hope that many of you will join me in Benidorm for this special quadrennial week.

Alan E. Gelfand, President, ISBA, 2006. ▲

INTERVIEW WITH PROF. DR. DR. H.C. VOLKER DOSE

by Rainer Fischer
Max-Planck-Institut fuer Plasmaphysik
Rainer.Fischer@ipp.mpg.de

The German Physical Society awards the Robert-Wichard-Pohl-Prize for the year 2005 to Prof. Dr. Dr. h.c. Volker Dose, Max Planck Institute for Plasma Physics, for his outstanding interdisciplinary contributions to the physics of atomic hydrogen, to the electronic band structure of solids and to Bayesian probability theory and its application.

Volker Dose was born on 16 February 1940 in Bad Segeberg, Germany. He took his PhD at the University of Zürich in 1967 and qualified for lectureship in experimental physics two years later at the same university. In 1971 he was appointed associate professor for experimental physics at the University of Würzburg. Since 1985 he is Scientific Fellow of the Max-Planck-Institute for Plasma Physics, heading the surface science department and one of the institute's directors. 1991 he took up an additional appointment as full professor at the University of Bayreuth. In 2000 he became director of the "Centre for Interdisciplinary Plasma Science" (CIPS) together with Professor G.E. Morfill.

R.F.: What were the reasons for you as a physicist to become a Bayesian and when did this happen?

V.D.: In 1980 I met in the course of experiments in surface science the problem that data were related to the physics information through a nonlinear integral equation. The inversion problem was successfully solved in the frame work of regularization theory which bears technical similarities to the Bayesian approach. The first truly Bayesian analysis concerned the exploitation of experimental data from spin polarized inverse photoemission. The spectacular success (from a physics point of view) of this work has fostered pursuit of Bayesian data analysis.

R.F.: Which were the most influential first books or papers?

V.D.: W. von der Linden drew my attention in 1992 to the article by Steven Gull and John Skilling in IEEE proceedings 131 (1984) 647. A little later I came across the book by Myron Tribus which left me deeply impressed.

R.F.: What do you think are the recent most important developments in Bayesian probability theory?

V.D.: We learn at present how to calculate the evidence even in nasty cases of multidimensional, multimodal posteriors. In my view model selection is the most outstanding feature of Bayesian theory.

R.F.: What kind of your applications do you like most?

V.D.: My preference lies in treating real observational or experimental data. The power of

the Bayesian approach becomes paramount if these data are scarce and of poor precision.

R.F.: We all know that every person with a reasonable man's mind is a Bayesian. Why is Bayesian thinking difficult to be adapted by so many scientists?

V.D.: New paradigms in science will hardly break through by convincing established researchers but rather with the enthusiasm of the following young generation. Experience in my home institute provides a good example. More over the technical difficulties with the evaluation of the numerous sometimes really nasty integrals in the treatment of key problems in the natural sciences constitute a considerable barrier.

R.F.: A frequent statement that is ascribed to you is that reasoning should always be based on proper prior information.

V.D.: First of all recall that only proper priors can be applied in model comparison. But let me explain by example what I have in mind. The analysis of physics experiments involves frequently scaling parameters relating to the electronic equipment employed in the experiment. The scientist who is not aware of the detection limits of his electronics nor of the upper bound for an undistorted output signal is neither qualified for performing the experiment nor for the analysis of the accumulated data.

R.F.: Do you think that objective methods have never to be considered?

V.D.: I am not aware whether "objective methods" is a well defined term. Among the approaches which I have seen prior densities derived from the transformation invariance requirement appear to be the soundest.

R.F.: Germany is a country with a long scientific tradition. Nevertheless, developing and applying Bayesian ideas is still very much underrepresented. What is the reason for this and what has to be changed in Germany to catch up with the rest of the world.

V.D.: I can give you only my personal speculation. In Germany the true mathematics is considered to be the pure mathematics. Applied mathematics in the German sense is something indecent, let alone numerics. Maybe the general situation will profit from the up-coming discipline of computational physics, which in my view resembles in many respects the admirably successful British applied mathematics.

R.F.: Physicists are not very well educated in modern statistical methods but they have lots of very ill-conditioned problems. How can modern techniques of data analysis be made more popular in the physical community?

V.D.: The procedure is quite straightforward. Those who know the tools of Bayesian analysis have to find collaboration with physicists and solve their particular problem. There is nothing as convincing as the solution of a seemingly unsolvable problem.

R.F.: Do you have any particular suggestion to give to a physicist when he first thinks about using Bayesian methods?

V.D.: He may read a book on Bayesian methods for his personal amusement. I have found it much more efficient to choose a possibly longstanding problem from the area of ones own research and consult the Bayesian literature to the extent necessary to solve it.

R.F.: We organized the last "MaxEnt" conference in 2004 after hosting it the first time in 1998. Since many years, the ISBA community and the community of Bayesian Inference and Maximum Entropy Methods in Science and Engineering try to find basic grounds. What do you think could be appropriate measures to merge the two communities?

V.D.: A merger between the two communities would of course be desirable. Note, however, that the MaxEnt conferences deal preferentially with applied topics. If ISBA would decide to open the new electronic Bayesian journal also for this class of problems, an important first step would happen. A suitable and adequate supplementation of the editorial board would be a clear sign for such an initiative.

R.F.: You will retire end of February 2005. What are your plans for the future?

V.D.: I welcome the additional free time which will be at my disposal. Various personal interest outside science have lain dormant during the last 35 years. Scientifically I shall continue a collaboration on climate related topics. This is a realm where you deal with sparse and low precision data. Accordingly one meets aspects of Bayesian theory which never show up in the treatment of physics problems.▲

THOMAS BAYES'S MATHEMATICAL WORK

by Andrew I. Dale
dale@ukzn.ac.za

Forsaking for once discussion of the well-known *An Essay towards solving a Problem in the Doctrine of Chances*, we shall consider here some of Bayes's contributions to other fields of mathematics, attention being given to what might be considered the more important aspects of such work.

1. Berkeley and fluxions.
In 1734 Bishop George Berkeley published *The Analyst; or, a Discourse Addressed to an Infidel Mathematician*. Here he severely criticized Isaac Newton's fluxionary and differential calculi, both with respect to the methods used and the ontological status of the things considered. Bayes's criticism of *The Analyst* in the anonymous *An Introduction to the Doctrine of Fluxions* (attributed to Bayes by Augustus de Morgan) was concerned with neither the theological aspects raised by Berkeley nor the latter's comments on the fluxionary calculus and moments, but rather with the logical analysis of Newton's prime and ultimate ratios. Setting out his work in postulates, definitions, axioms and propositions, Bayes emerges as one with a logical mind keenly aware of the need for mathematical rigour. Further, one finds here more than the glimmerings of an understanding of the concept of limits (see Smith (1980)).

Later commentators are divided on whether Bayes's criticism of Berkeley and defence of Newton were successful, but I think that one may safely say that the tract was a useful comment and worthy of more attention than was paid it at the time of its appearance.

2. On a semi-convergent series.
In the short *Letter from the late Reverend Mr. Thomas Bayes* ..., the first of two posthumously published works (the other being *An Essay towards solving a Problem in the Doctrine of Chances* and the supplementary *Demonstration of the second rule in the essay* ...), Bayes considers the series expansion of $\log z!$. Giving the development up to the term in z^{-9}, with the appropriate coefficients, he notes that 'the whole series can never properly express any quantity at all; because after the 5th term the coefficients begin to increase, and they afterwards increase at a greater rate than what can be compensated by the increase of the powers of z'. The divergence of the series for $z = 1$ had earlier been noted by Euler, but Bayes seems to have been the first to notice the divergence for arbitrary z. [Mention is also made here of the divergence of the appropriate series 'for' $\log(2z-1)!!$]

3. Fluxions and differences.
In 1715 Brook Taylor gave a theorem that essentially provides an expansion of $y(x+h) - y(x)$ in terms of derivatives of y, or, as we might write it, $y{=}f(\dot{y}, \ddot{y}, \ldots)$. Many later authors repeated this result: for example, in the second (1801) edition (and perhaps also the first) of Colin MacLaurin's *A Treatise on Fluxions* we find, in Art. 751 (p. 199), the following statement:

> ... hence it appears at what rate the fluxion of y of each order contributes to produce the increment or decrement of y, since
>
> $$y - E = \dot{E} + \frac{\ddot{E}}{2} + \frac{\dddot{E}}{6} + \frac{\ddddot{E}}{24} + \&\text{c}.$$

Less common is a theorem we may give generally as $\dot{y}{=}\, f(y, y, \ldots)$, which seems to have been known by Bayes (see §4 below). This result is attributed to Lagrange by Jordan (1965), and Bellhouse suggests that it was first published by Lagrange in 1772.

4. Papers in the Stanhope collection.
Recently David Bellhouse discovered a number of items in the Stanhope of Chevening collection in the Kent County Archives in Maidstone, England, that are relevant to our topic. Here we find two preliminary papers on the series 'for' $\log z!$. One of these bears on one side the words 'Mathematical paper of M$^{\text{r}}$ Bayes's communicated Sept$^{\text{r}}$ 1$^{\text{st}}$ 1747', which certainly suggests that Bayes and Philip, Earl Stanhope were acquainted.

This collection also contains a proof by Bayes of Stirling's approximation that does not use the semi-convergent series. The proof is too complicated to mention here (see Dale (2003) for details): we note only that Bayes shows essentially (though of course not in this notation) that

$$z! = \sqrt{2\pi} z^{z+(1/2)} e^{-z}.$$

Another item is concerned with 'Mr Bayes's Demonstration of a Theorem which I [i.e. Stanhope] found lately & told him of. Sept$^{\text{r}}$ 1747'. This result runs as follows: to find $(1/x)$ when $\dot{z} = 1$ and

$$x = 1 + \frac{z}{2} + \frac{z^2}{2.3} + \frac{z^3}{2.3.4} + \frac{z^4}{2.3.4.5} + \&c.$$

Bayes deduces essentially that

$$\frac{1}{x} = \frac{z}{e^z - 1}.$$

There is also a very short Note headed 'Theorem mentioned to me at Tunbridge Wells by Mr Bayes Aug. 12. 1747', viz.

$$\dot{y} = \underset{\cdot}{y} - \frac{1}{2}\underset{\cdot\cdot}{y} + \frac{1}{3}\underset{\cdot\cdot\cdot}{y} - \frac{1}{4}\underset{::}{y} + \frac{1}{5}\underset{:\cdot}{y} - \frac{1}{6}\underset{:::}{y} + \&c$$

This is the result we have already mentioned in §3 and it is presented here without proof.

Another manuscript, in Stanhope's handwriting, is labelled 'The Reverend M^r Bayes's Paper concerning Trinomial Divisors'. This is concerned with the expansion of $x^{2n} - 2\cos(\theta)x^n + 1$, the method given here essentially being a combination of MacLaurin's geometrical argument and de Moivre's inductive generalization.

The final result we shall mention here is Bayes's expansion of $x = (\arcsin z)^n$. This is given, for unit radius, as a series in $z^k, k \in \{n-1, n+1, n+3, \ldots\}$, which is found by expanding $x^{n-1}/\sqrt{1-z^2}$.

There is one other result that perhaps bears some mention. This occurs in a letter, with seal, addressed to Stanhope where it is shown that

$$\dot{y}/\dot{x} = (\underset{\cdot}{y} - \frac{1}{2}\underset{\cdot\cdot}{y} + \frac{1}{3}\underset{\cdot\cdot\cdot}{y} - \frac{1}{4}\underset{::}{y})/\underset{\cdot}{x} + \&c.$$

The proof (written by Stanhope?) bears at the end, and written with a different pen, the words 'This is a Theorem shewn to me by the late M^r Bayes.'

5. Bayes & Simpson.
Among the Bayes papers in the Library of the Royal Society is a letter from Bayes to John Canton. The letter starts 'You may rem[ember] a few days ago we were speaking of Mr Simpson attempt to show y^e great advantage of taking y^e mean between several Astron. observations rather than trusting to a single observation carefully made in order to diminish y^e error arising fro[m] y^e imp[er]fection of instrum^{ts} and y^e organs of sense.' Bayes astutely observed that Simpson's suggestion was a useful one only when the distribution of errors was essentially symmetric and, moreover, that it was contingent on the perfection of the measuring instrument.

Thomas Simpson's first publication on this matter was in 1756 in a letter to George, Earl of Macclesfield, published in volume 49 of the *Philosophical Transactions*, the material being repeated, with minor modifications, in his *Miscellaneous Tracts* of 1757. Bayes's letter is undated: if it was written after the first of these publications, then either the matter was not brought to Simpson's attention or he chose to pay little, if any, attention to it.

It is perhaps interesting to note, as having some slight relevance to this letter, two more recent observations: the first, by Jeffreys, 'a single observation can tell us nothing about its own precision' (1933, p. 532) (cf. Bayes's 'rather than trusting to a single observation carefully made') and the second by Savage, who wrote '... it seems important to mention that, in principle, a single measurement with an instrument of known high accuracy nearly induces the same normal posterior for everyone' (1962, p. 101).

6. On the Beta and Normal Distributions.
The contributions of Bayes and Richard Price to our knowledge of the beta probability integral are discussed by Dutka (1981) and Hald (1990, 1998). Briefly put, in his *Essay* Bayes obtained an approximation to the two-sided beta probability integral (thus presenting the first use of the incomplete beta function in a probabilistic setting). Price effected improvements, the result being considerably better than the Normal approximation. 'The Bayes-Price results are obtained by approximating the skew beta density by a symmetric beta density times a factor tending to unity for $n \longrightarrow \infty$, the two functions having the same maximum and the same points of inflection' (Hald, 1990, p. 139). Hald also notes (loc. cit.) that from this approximation all Laplace's results based on the Normal distribution follow easily.

In the *Demonstration of the Second Rule in the Essay*... (also by Bayes, but with considerable amplification by Price), published in volume 54 of the *Philosophical Transactions*, Price in fact did something more: he derived a result that is equivalent to the approximation of the posterior distribution by the Normal, essentially expressing the approximation in terms of the power series

$$\frac{2}{\sqrt{\pi}} \sum_{i=0}^{\infty} \frac{(-1)^i u^{2i+1}}{i!\,(2i+1)}.$$

7. Conclusion.
The above items essentially cover all of Bayes's mathematical work. Some further jottings are to be found in a Notebook attributed to him and at present in the Equitable Life Assurance Society in London: see Chapter 11 of Dale (2003) for details.▲

REFERENCES.

[1] Bellhouse, D.R. (2002). "On some recently discovered manuscripts of Thomas Bayes." *Historia Mathematica*, **29**, 383–394.

[2] Dale, A.I. (2003). *Most Honourable Remembrance: The Life and Work of Thomas Bayes*. Springer-Verlag, New York.

[3] Dutka, J. (1981). "The incomplete beta function — a historical profile." *Archive for History of Exact Sciences*, **24**, 11-29.

[4] Hald, A. (1990) "Evaluations of the beta probability integral by Bayes and Price." *Archive for History of Exact Sciences*, **41**, 139–156.

[5] ——(1998). *A History of Mathematical Statistics from 1750 to 1930*. Wiley, New York.

[6] Jeffreys, H. (1933). "Probability, statistics, and the theory of errors." *Proceedings of the Royal Society of London, A*, **140**, 523–535.

[7] Jordan, C. (1965). *Calculus of Finite Differences*. 3rd ed. Chelsea, New York.

[8] Savage, L.J. et al. (1962). *The Foundations of Statistical Inference*. Methuen & Co. Ltd, London.

[9] Smith, G.C. (1980). "Thomas Bayes and fluxions." *Historia Mathematica*, **7**, 379–388.

BAYESIAN METHODS FOR SOME MISSING DATA PROBLEMS IN GENE REGULATION

by Mayetri Gupta and Joseph G. Ibrahim
gupta@bios.unc.edu and
ibrahim@bios.unc.edu

Latent variable problems in genomics, often combined with high dimensionality and massive data sets, present an attractive scenario for Bayesian modeling combined with powerful Monte Carlo computational tools. Some typical "missingness" issues include sequence alignment problems such as global alignments for determining evolutionary relationships between proteins and short repetitive patterns within long sequences, arising in regulatory motif discovery. Other examples include discovering groupings of genes that belong to the same functional pathway (gene "clusters" being unobservable); determining which genes (or gene groups) are differentially expressed in a microarray experiment; estimating the phylogenetic tree that reflects the evolutionary relationship between species. Due to space limitations, we will focus on missing data problems arising in functional genomics with an emphasis on gene regulation, and some of the proposed Bayesian solutions.

Multiple sequence alignment

Some of the initial Bayesian approaches to sequence alignment and motif discovery problems are found in the following.

1. Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F., and Wootton, J. C. (1993). Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science*, **262**(5131), 208–14. This is one of the first approaches to the motif discovery problem based on a Bayesian product multinomial model, with Dirichlet priors, and using a Gibbs sampling approach to impute the sites corresponding to a single motif.

2. Liu, J. S., Neuwald, A. F., and Lawrence, C. E. (1995). Bayesian models for multiple local sequence alignment and Gibbs sampling strategies. *J. American Statistical Association*, **90**, 1156–1170. The single-motif model is extended to allow multiple motifs and a more efficient Gibbs sampling algorithm, collapsed Gibbs, is proposed.

3. Zhu, J., Liu, J. S. and Lawrence, C. E. (1998) Bayesian adaptive sequence alignment algorithms. *Bioinformatics*, **14**, 25-39. Relaxes the traditional requirement of a fixed set of scoring matrices and gap penalty parameters in pairwise alignment. Proposes a Bayesian solution providing the posterior distribution of all alignments considering a range of gaps and scoring matrices.

4. Liu, J., Neuwald, A. F. and Lawrence, C. E. (1999) Markovian structures in biological sequence alignments. *J. American Statistical Association*, **94**, 1-15. Provides a novel multiple sequence alignment methodology in a Bayesian framework. Insertions are characterized by a propagation model that combines the sensitivity of the block-based motif model with the flexibility of the hidden Markov model framework.

Generalizing motif models

For detecting motifs in the large sequence search space of complex genomes, more sensitive motif models are necessary. Work has been done in generalizing the product multinomial motif model by incorporating

insertions and deletions; modeling dependencies between motif columns; modeling dependence between motif occurrences in sequences to find spatial clusters of motif sites (regulatory modules); and relaxing between-sequence independence assumptions to explicitly model evolutionary relationships.

5. Gupta, M. and Liu, J. S. (2003). Discovery of conserved sequence patterns using a stochastic dictionary model. *J. American Statistical Association*, **98**(461), 55-66. Introduces a dictionary-based model with motifs as stochastically varying words and a recursive data augmentation scheme for sampling motif sites. Motifs with insertions and deletions, and of varying lengths are allowed. A Maximal a Posteriori model selection criterion is used to determine the total number of motifs.

6. Kechris, K. J., van Zwet, E., Bickel, P. J., and Eisen, M. B. (2004). Detecting DNA regulatory motifs by incorporating positional trends in information content. *Genome Biology*, **5**(7), R50. To model the structure of motifs that have a unimodal profile, the authors penalize deviations from a conserved profile using a normal or double exponential prior, instead of the usual Dirichlet distribution. An EM algorithm with a modified M-step is used to estimate the parameters.

7. Zhao, X., Huang, H., and Speed, T. P. (2004). Finding short DNA motifs using permuted Markov models. *RECOMB proceedings*, pages 68–75. A variable-length permuted Markov model is proposed to model motif sites. It is assumed that an unobserved permutation has acted on the positions of all the motif sites, the original ordered positions being described by a Markov chain. The Markov chain in different contexts at different positions of a motif are allowed to have variable lengths of memory, controlling the size of the parameter space.

8. Zhou, Q. and Liu, J. S. (2004). Modeling within-motif dependence for transcription factor binding site predictions, *Bioinformatics*, **20**(6), 909–916. The product multinomial model is allowed one or more correlated column pairs, under the restriction that no two pairs of correlated columns can share a common column. A Metropolis-Hastings step is added to the Gibbs sampler to decide whether to add or delete a pair of correlated columns at each iteration. The posterior distribution is collapsed over the motif probability matrix during the Metropolis-Hastings step to avoid a parameter space of varying dimensions for different numbers of correlated columns.

Long-range dependence

9. Thompson, W., Palumbo, M. J., Wasserman, W. W., Liu, J. S., and Lawrence, C. E. (2004). Decoding human regulatory circuits. *Genome Research*, **10**, 1967–74. A hidden Markov model (HMM) is used to represent a regulatory module, the hidden states representing a fixed total number of motif types, and the ordering of motif sites within a module depending on the transition probabilities of the Markov chain. Gibbs sampling is used to sample locations of motif sites and find posterior parameter estimates.

10. Xing, E. P., Wu, W., Jordan, M. I. and Karp, R. M. (2004). LOGOS: A modular Bayesian model for de novo motif detection. *J. Bioinformatics and Computational Biology*, **2**(1), 127-154. The model consists of a combination of two sub-models- a local model that uses a hierarchical Dirichlet mixture to reflect subjective prior knowledge and positional dependence within the motif structure, and a global sequence model that models frequencies and dependencies of motif occurrences. Model parameters are fit in an empirical Bayes framework using a variational EM algorithm.

11. Gupta, M. and Liu, J. S. (2005) De novo cis-regulatory module elicitation for eukaryotic genomes. *Proc. National Academy of Sciences USA*, **102**, 7079-84. The hidden Markov model approach for modules is extended in two ways:(i) a length distribution is imposed on the distances between site occurrences in the cluster, and (ii) the assumption that the total number of motif types must be known in advance is relaxed. For the resulting "semi-HMM" with an unknown number of states, the model selection problem is addressed by means of combining evolutionary Monte Carlo techniques with data augmentation.

12. Li, X. and Wong, W.H. (2005). Sampling motifs on phylogenetic trees. *Proc. National Academy of Sciences USA*, **102**, 9481-9486. Provides a motif discovery algorithm for sequences of multiple species related through an (unobserved) evolutionary tree. Motifs

and background (non-motif) sequence are incorporated in a Bayesian framework characterized by two different continuous-time Markov chain evolutionary models. Motif instances are sampled, one species at a time, using a Gibbs sampling algorithm from a likelihood that incorporates the evolutionary relationship.

Motif scoring and clustering

13. Liu, X., Brutlag, D. L., and Liu, J. S. (2001). Bioprospector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pacific Symposium on Biocomputing*, p. 127–138. Noting that the Gibbs motif sampler updating step can be reformulated as an approximate "score" function contrasting the motif and background, the authors use a Metropolis-like approach to scan the sequences and determine a site to be a motif occurrence if its score exceeds a "high" threshold, while sites within a high and low threshold are given a chance to be sampled into the final set. It also allows more flexible motif models such as two-block motifs.

14. Jensen, S. T., Liu, X. S., Zhou, Q. and Liu, J. S. (2004). Computational discovery of gene regulatory binding motifs: a Bayesian perspective. *Statistical Science*, **19**, 188–204. Presents a more accurate approximation to the Bioprospector scoring function and a simulated annealing procedure to optimize this function.

15. Qin, Z. S., McCue, L. A., Thompson, W., Mayerhofer, L., Lawrence, C. E., and Liu, J. S. (2003). Identification of co-regulated genes through Bayesian clustering of predicted regulatory binding sites. *Nature Biotechnology*, **21**(4), 435-39. Addresses the problem of clustering similar motif sequences to determine possible co-regulated genes. A hierarchical Bayesian clustering scheme, using Gibbs sampling is developed, using a generalization of a Dirichlet process prior model.

Differential gene expression

16. Ibrahim, J. G., Chen, M. H. and Gray, R. (2002). Bayesian models for gene expression with DNA microarray data. *J. American Statistical Association*, **97**, 88-99. A log-normal model with a truncation threshold is developed for differential gene expression. Prior elicitation strategies using empirical Bayes methods and gene selection algorithms are proposed.

17. Tadesse, M. G., Ibrahim, J. G., and Mutter, G. (2003). Identification of differentially expressed genes in high-density oligoneucleotide arrays accounting for the quantification limits of the technology. *Biometrics*, **59**, 542–554. A two-way hierarchical ANOVA gene expression model is proposed treating low expression levels as left censored data.

18. Tadesse, M. G., Ibrahim, J. G., Gentleman, R., Chiaretti, S., Ritz, J., and Foa, R. (2005). A Bayesian error-in-variable survival model for the analysis of Genechip arrays. *Biometrics*, **61**, 488–497. A measurement error Li-Wong model is proposed for the expression data which is then linked to a piecewise exponential survival model, leading to a joint model for gene expression and survival data.

19. Newton, M. A., Kendziorski, C. M., Richmond, C. S., Blattner, F. R., and Tsui, K. W. (2001). On differential variability of expression ratios: improving statistical inference about gene expression changes from microarray data. *J. Computational Biology*, **8**, 37-52. Estimates of gene expression changes are derived with a hierarchical gamma model for cDNA microaaray data and significant gene expression changes are identified by deriving posterior odds.

20. Newton, M. A., Noueiry, A., Sarkar, D., and Ahlquist, P (2004). Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics*, **5**, 155-176. A hierarchical mixture model is developed that is sensitive in detecting differential expression and flexible to account for the complex variability of normalized microarray data.

21. Baldi, P. and Long, A. D. (2001). A Bayesian framework for the analysis of microarray expression data: regularized *t*-test and statistical inferences of gene changes. *Bioinformatics*, **17**, 509-519. A gene-specific *t*-test is developed within a hierarchical model where the variance of each gene is calculated by a weighted average of the empirical variance and a local background variance associated with the neighboring genes.

22. Do, K.-A., Mueller, P., and Tang, F. (2005). A Bayesian mixture model for differential gene expression. *Applied Statistics*, **54**, 611-626. A nonparametric Bayesian model is proposed for the distribution of gene intensities under various conditions. The approach is similar to empirical Bayes, but the avoidance of plug-in estimates facilitates the evaluation of posterior expected false discovery rates.

23. Hein, A-M., Richardson, S., Cuaston, H. C., Graeme, A. K., and Green, P. J. (2005). BGX: A fully Bayesian integrated approach to the analysis of Affymetrix genechip data. *Biostatistics*, **6**, 349-373. Hierarchical models for Affymetrix GeneChip data are proposed in which the processing steps of the raw data are modeled and integrated into a common statistical framework. The full posterior distribution of the gene expression indices are derived. Multiplicative and additive error models are considered.

Classification by variable selection

24. Ishwaran, H. and Rao, J. S. (2003). Detecting differentially expressed genes in microarrays using model selection, *J. American Statistical Association*, **98**, 438-455. Differentially expression genes are detected using high-dimensional model selection, termed Bayesian ANOVA for Microarrays (BAM). The approach involves a weighted average of generalized ridge regression estimates, providing the benefits of using shrinkage estimation combined with model averaging.

25. Parmigiani, G., Garrett, E. S., Anbazhagan, R., and Gabrielson, E. (2002). A statistical framework for expression-based molecular classification in cancer. *J. Royal Statistical Society, Ser. B*, **64**, 717-736. A modeling framework is proposed to inform and organize the development of exploratory tools for classification. Uses latent categories to provide both a statistical definition for differential expression and a precise experiment-dependent definition of a molecular profile.

26. Ishwaran, H. and Rao, S. (2005). Spike and slab gene selection for multigroup microarray data. *J. American Statistical Association*, **100**, 764-780. A spike and slab hierarchical model is proposed for the multigroup gene selection problem. The variable selection procedure extends the "spike and slab" idea of a two-point mixture distribution with a uniform flat distribution and a degenerate distribution at zero to a more general multivariate normal scale mixture distribution specified through the prior for the hypervariance.

27. Tadesse, M. G., Sha, N., and Vannucci, M. (2005). Bayesian variable selection in clustering high dimensional data, *J. American Statistical Association*, **100**, 602-617. A method for simultaneous clustering and variable selection is proposed. The clustering problem is formulated in terms of a multivariate normal mixture model with an unknown number of components and the reversible jump algorithm is used to move between different dimensional spaces. The predictors are selected by introducing a binary latent vector which gets updated via stochastic search techniques.

28. West, M. (2003). Bayesian factor analysis regression for models in the Large p, Small n paradigm. In Bayesian statistics **7**, J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West (Eds)., Oxford: Oxford University Press, pp. 733-742. Factor analysis models are developed for the large p small n paradigm with applications to gene expression data. Latent factor models for high dimensional variables are proposed along with regression approaches to in which low dimensional latent factors are the predictor variables.

Multiple testing

29. Efron, B. Tibshirani, R. Storey, J., and Tusher, V. G. (2001). Empirical Bayes analysis of a microarray experiment. *J. American Statistical Association*, **96**, 1151-1160. A nonparametric empirical Bayes method is introduced to guide the efficient reduction of the data to a single summary statistic per gene, and also make simultaneous inferences concerning which genes are affected by a treatment. The empirical Bayes inferences are closely related to the frequentist false discovery rate (FDR) criterion.

30. Mueller, P., Parmagiani, G., Robert, C, Rousseau, J. (2004). Optimal sample size for multiple testing: the case of gene expression microarrays. *J. American Statistical Association*, **99**, 990-1001. A decision theoretic approach is developed for optimal sample size

in multiple-comparison problems, with applications to the choice of the number of microarray experiments to be carried out when learning about differential gene expression. The decision rule that emerges takes the same form of the rules proposed for controlling the posterior expected false discovery rate.

Using gene expression in motif discovery

There have been a few recent approaches in trying to combine information from sequence and gene expression to get a more complete picture of gene regulation.

31. Conlon, E.M., Liu, X.S., Lieb, J.D., Liu, J.S. (2003). Integrating regulatory motif discovery and genome-wide expression analysis. *Proc. National Academy of Sciences USA*, **100**, 3339-3344. A multiple linear regression is used to model the relationship between the logarithm of the differential expression values and a Bayesian sequence motif score. A stepwise regression approach is used to determine significantly related motifs.

32. Tadesse, M.G., Vannucci, M. and Lio, P. (2004) Identification of regulatory motifs using Bayesian variable selection. *Bioinformatics*, **20**, 2553-2561. The Conlon et al. procedure is extended by selecting significant regulatory motifs using Bayesian variable selection techniques.

33. Ji, H. and Wong, W. H. (2005), TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics*, **21**, 3629-3636. A *t*-like statistic with an empirical Bayes variance estimate is used to identify peaks in chromatin immunoprecipitation (ChIP-chip) tiling array data, and the false discovery rate (FDR) for each peak is calculated using a nonparametric mixture formulation.

Books and review papers

34. Liu, J. S. and Logvinenko, T. (2003). Bayesian Methods in Biological Sequence Analysis, in *Handbook of Statistical Genetics*, 2nd Ed., D.J. Balding, M. Bishop and C. Cannings (eds), J. Wiley & Sons.

35. Gupta, M. and Liu, J. S. (2006). Bayesian modeling and inference for motif discovery, in *Bayesian inference for gene expression and proteomics*, Kim-Anh Do, Peter Mueller and Marina Vannucci (Eds). Cambridge University Press. Exposure Pathways

36. Sebastiani, P., Gussoni, E., Kohane, I., and Ramoni, M. F. (2003). Statistical challenges in functional genomics, Statistical Science (with discussion), **18**, 33-70.

Biology references

37. Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., Darnell, J. E. (1999). *Molecular Cell Biology*, 4th ed., New York: W. H. Freeman & Co. Ch7,10. `http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowTOC&rid=mcb.TOC&depth=10`

▲

BAYESIAN MODELING OF EXPOSURE PATHWAYS

Catherine Calder
`calder@stat.ohio-state.edu`

A team of researchers at The Ohio State University and Battelle, a research and development enterprise headquartered in Columbus, OH, are developing a hierarchical Bayesian modeling framework for analyzing pathways of exposure to toxic substances. Involved in the "Sources to Biomarkers" (STB) study are four members of the faculty of the Department of Statistics at Ohio State (Catherine Calder, Peter Craigmile, Noel Cressie, and Thomas Santner), three senior researchers from Battelle (Bruce Buxton, Nancy McMillan, and Michele Morara), and several current and former Ohio State statistics Ph.D. students (Crystal Dong, Hongfei Li, Ke Wang, and Jian Zhang) and Battelle researchers (Vincent Agboto, Jessica Sanford, Greg Young). Funding for the project resulted from a submission to the EPA's FY2003 STAR Grant program, which was jointly funded by the EPA's National Center for Environmental Research (NCER) and the American Chemistry Council (ACC).

Characterizing routes of exposure to toxic substances is a difficult task. While levels of toxic substances in environmental media and in human biomarkers can be measured, large amounts of individual-level data are not readily available

due to the burdensome nature of the data collection process. As a result, there is a need to supplement individual-level exposure data with additional sources of information such as spatially referenced measurements of the levels of toxics in environmental media. Bayesian hierarchical modeling allows these diverse data sources to be synthesized in a scientifically interpretable manner that accounts for all sources of uncertainty coherently.

The goal of the STB project is to develop a multiscale (areal and individual) statistical model that describes how multiple media (*e.g.*, air, soil, dust, food, water) contribute to direct routes of exposure (inhalation, ingestion, dermal). The hierarchical modeling framework includes five specific levels starting from sources of pollutants to the manifestation of the pollutants in human bodies, as measured by biomarkers:

$$\begin{array}{c} \textit{Biomarkers} \\ \Uparrow \\ \textit{Personal Exposure} \\ \Uparrow \\ \textit{Local Environment} \\ \Uparrow \\ \textit{Global Environment} \\ \Uparrow \\ \textit{Sources} \end{array}$$

While the structure of the STB modeling framework can be readily adapted to study exposure to a variety of pollutants, the current project focuses on four toxic metals (arsenic, cadmium, chromium, and lead) in EPA Region 5 (Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin) and Arizona. These two regions coincide with the population surveyed by the National Human Exposure Assessment Survey (NHEXAS), a residential-based exposure survey conducted from 1995-1998 (see Lebowitz *et al.*, 1995, Pellizzari et al., 1995, for details and Clayton *et al*, 2002, for a structural-equation analysis of exposure pathways using NHEXAS data). NHEXAS provides various biomarker and environmental media measurements for a stratified random sample of individuals who were monitored for a period of seven days. Additionally, demographic, occupational, and activity information, as well as food diaries for the participants are provided. These individual-specific data drive the *Local Environment* to *Personal Exposure* to *Biomarkers* stages of the STB model.

Despite the wealth of information on individual-specific exposure routes provided by NHEXAS, a major difficulty in making population-level inferences is that the NHEXAS sampling design provides limited geographic coverage. NHEXAS data alone cannot be used to make inferences on exposure routes for the general population since there is substantial spatial variation in the both the naturally occurring and anthropogenic levels of metals in the *Global Environment*. In order to generalize the information provided by the NHEXAS data, a variety of different types of additional data (*e.g.*, emission inventories, dietary patterns, and levels of metals in ambient air, source and treated water, soil, stream sediment) are incorporated into the model for individual exposure pathways. These data are used in the *Sources* and *Global Environment* stages of the hierarchy. Extending the pathways framework from individual-specific *Local Environment* back to the *Global Environment* requires modeling data collected at misaligned spatial scales and according to different protocols. As an example of one of the *Global Environment* components of the STB model, stream sediment data associated with watersheds are used to characterize the levels of metals in soil across counties. Figure 1 shows the location of measurements of arsenic in stream sediment across Region 5, along with the posterior mean of the *log* arsenic concentration in topsoil at the watershed level. NHEXAS data alone are not able to capture the spatial variation in the arsenic concentration of soil across Region 5. Consequently, it is essential that *Global Environment* information be used in order to characterize variation in individual-specific *Personal Exposure* across the region.

In addition to providing a better understanding of pathways of exposure to toxic metals, a long term objective of the STB project is to explore the link between exposure to toxics and human health outcomes. Since the biological effects of toxic exposure are not well understood, the aim is to characterize the variation in individuals' exposures within a population, which in turn can be related to health patterns in the population. Finally, the pathways modeling approach is amenable to assessing the potential impacts of environmental policy. Given different emission reduction or environmental remediation scenarios, predictions of the reduction of individual-specific exposure can be obtained, along with corresponding uncertainty statements. Quantifying the impact of different scenarios on the outcome of interest (*i.e.*, personal exposure) is of great value when performing cost-benefit analyses of changes in environmental policy.

For further information on the STB modeling framework, see Cressie *et al.* (2005). ▲

References

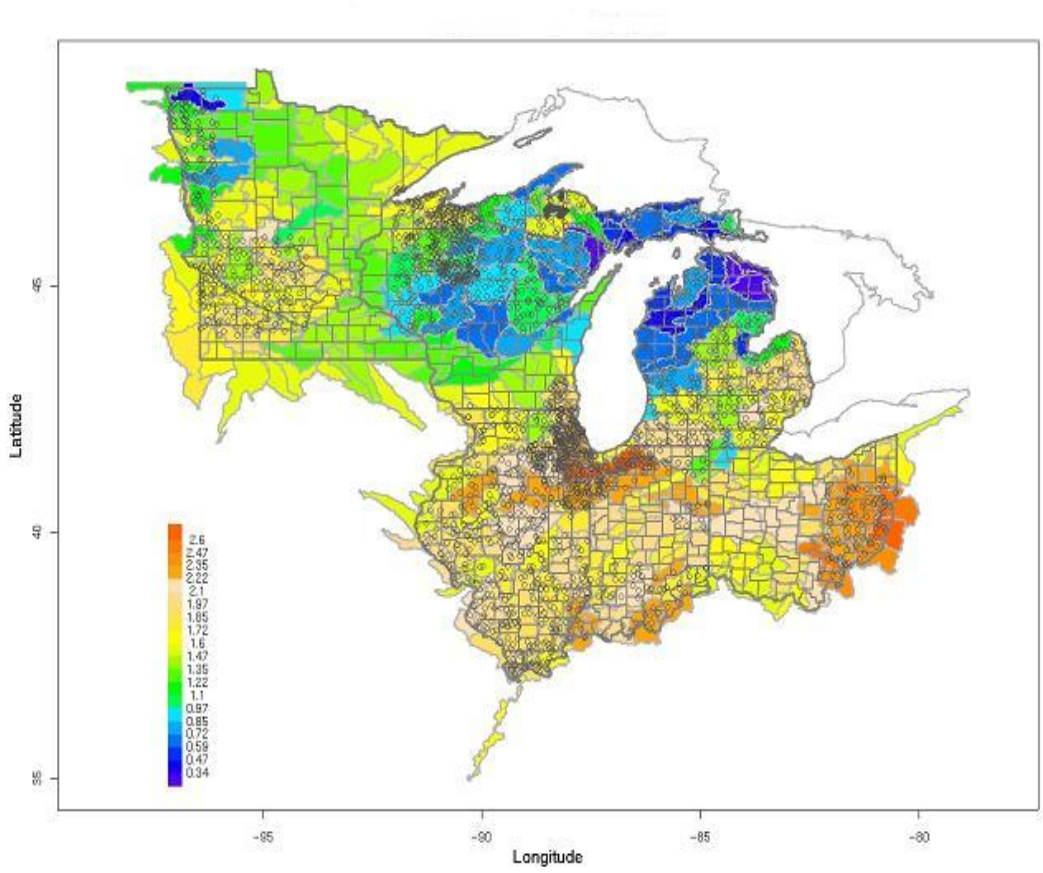
- Clayton, C.A., Pellizzari, E.D., and Quakenboss, J.J. (2002), "National Human Exposure Assessment Survey: Analysis of Exposure Pathways and Routes for Arsenic and Lead in EPA Region 5", *Journal of Exposure Analysis and Environmental Epidemiology*, **12**, 29-43.

- Cressie, N., Buxton, B.E., Calder, C.A., Craigmile, P.F., Dong, C., McMillan, N.J., Morara, M., Santner, T.J., Wang, K., Young, G., and Zhang, J. (2005), "From Sources to Biomarkers: A Bayesian Approach to Human Exposure Modeling", Department of Statistics Preprint No. 765, The Ohio State University. (Under review).

- Lebowitz, M.D., O'Rourke, M.K., Gordon, S., Moschandreas, D.J., Buckley, T., and Nishioka, M. (1995), "Population-Based Exposure Measurements in Arizona: A Phase I Field Study in Support of the National Human Exposure Assessment Survey", *Journal of Exposure Analysis and Environmental Epidemiology*, **5**, 297-325.

- Pellizzari, E., Lioy, P., Quankenboss, J., Whitmore, R., Clayton, A., Freeman, N., Waldman, J., Thomas, K., Rodes, C., and Wilcosky, T. (1995), "Population-Based Exposure Measurements in EPA Region 5: A Phase I Field Study in Support of the National Human Exposure Assessment Survey", *Journal of Exposure Analysis and Environmental Epidemiology*, **5**, 327-358.

Figure 1: The map coloring corresponds to the posterior mean of the *Global Environment* level of arsenic in soil across EPA Region 5 watersheds. The circles represent the locations of the 5221 stream sediment samples from the USGS's National Geochemical Survey (NGS) which are jointly modeled with the 249 NHEXAS participants' soil samples. This figure illustrates the potential for understanding the mechanisms driving the spatial variation in arsenic exposure across a large geographic area by supplementing individual exposure data with information on the background environmental levels of pollutants.



GET YOUR ISBA/VALENCIA TRIP FOR FREE!

by Robert B. Gramacy
rbgramacy@statslab.cam.ac.uk

First and foremost, I'd like to call attention to funding available to students and young researchers for ISBA/Valencia 2006. There are three different applications, depending on whether residence is claimed in North America, Europe, or a developing country. The application deadlines are 15, 31, and 31 March, respectively. Residents of North America and Europe must be presenting a contributed paper or poster in order to be considered. Abstracts for the for poster contributions are due by 31 March. Please see the conference web page for more details:

http://www.uv.es/valenciameeting

From my own personal experience, the biggest obstacle to getting a grant to go to an ISBA meeting, or any meeting for that matter, is in filling out the application. And even that is pretty easy—you just have to do it! Regardless of the stage in your degree, or progress on your dissertation, I strongly urge you to apply for funding.

Dissertation Abstracts:

A primary feature of the Student Corner section is the publication of dissertation abstracts. If you have recently defended your Ph.D. thesis, please email the abstract to rbgramacy@statslab.cam.ac.uk. This month we have one abstract.

CONTRIBUTIONS TO BAYESIAN STATISTICAL ANALYSIS: MODEL SPECIFICATION AND NONPARAMETRIC INFERENCE

by Milovan Krnjajic
milovan@ams.ucsc.edu
Department of Applied Math & Statistics
University of California, Santa Cruz
Advisor: David Draper

This dissertation concerns two topics in Bayesian statistical analysis: model specification and nonparametric inference. We follow de Finetti in our understanding that the (most) general way to define the Bayesian model may be by regarding it as a joint predictive distribution for data that have not yet been observed. In order to identify preferred models we (1) argue for comparing their performance in the posterior predictive space, (2) explore versions of log-scoring criterion for that purpose, and (3) show how to calibrate the log-score scale. We further undertake a simulation study to explore the ability of Bayesian parametric and nonparametric models, based on Dirichlet process (DP) mixtures, to provide an adequate fit to count data and find that the nonparametric models are able to flexibly adapt to the data, enable rich posterior inference, and provide, in a variety of settings, more accurate predictive inference than parametric models. We also present novel Bayesian nonparametric methodology for quantile regression developing DP mixture models for the error distribution in an additive regression formulation. The models allow the shape of the error density to adapt to the data and thus provide more reliable predictive inference than models based on parametric error distributions. We also consider model extensions for data sets with censored observations. Moreover, we employ dependent Dirichlet processes to develop quantile regression models that allow the error distribution to change nonparametrically with the covariates.

Milovan is now a post-doctoral researcher at Lawrence Livermore National Laboratory, EETD, Systems and Decision Sciences Section. He and David Draper will be speaking on related topics at ISBA/Valencia 2006, covering points (1–3) above in particular.

On a related note, I too have finished my Ph.D. (Yay!) However, rather than paste my abstract here, let me refer you to my applications article in December 2004 edition of the ISBA bulletin. In February I started a job as post-doctoral researcher in the Statistical Laboratory at the University of Cambridge, under Steve Brooks. My post is funded by the newly formed National Centre for Statistical Ecology—a collaboration between the universities at Cambridge, Kent and St. Andrews.

Sadly, this means that this is probably my last stint in the role of *Student Corner* associate editor of the ISBA Bulletin. It has been fun! I hope to see you all at ISBA/Valencia. Get your funding applications in!▲

NEWS FROM THE WORLD

by Alexandra M. Schmidt
alex@im.ufrj.br

I would like to encourage those who are organizing any event around the World, to get in touch with me to announce it here.

Events

Seminar on Bayesian Inference in Econometrics and Statistics, University of Iowa, EUA, April, 28th-29th 2006.

The Seminar on Bayesian Inference in Econometrics and Statistics, founded by Arnold Zellner, has reconvened with Siddhartha Chib as Director. All interested ISBA members are invited to participate. The call for papers and information about registration and financial support can be found at http://www.biz.uiowa.edu/econ/sbies/.

Spatial Epidemiology Conference, Imperial College, London, UK, May, 23rd-25th, 2006.

This conference will bring together international expertise with a particular interest in geographical variations in environmental health. Participants from all parts of the Public Health community (researchers, public health specialists, policymakers etc) are welcomed. A short course on Bayesian methods in spatial epidemiology using GeoBugs will be held on the 22nd and 23rd of May. Then a post conference GIS workshop will happen on the the 25th and 26th of May. See http://www.spatepiconf.org/ for more details.

Bayesian Inference in Complex Stochastic Systems, University of Warwick, Coventry, UK, May, 28th-30th, 2006.

The workshop will be organised into five keynote presentations and fifteen related presentations. There will be three related presentations for each keynote speaker and a poster session. The timing of the meeting was deliberately chosen to allow participants a convenient transition to the Valencia meeting.

Keynote speakers are Jim Berger (ISDS, Duke University), Alan Gelfand (ISDS, Duke University), Ed George (The Wharton School, University of Pennsylvania), Peter Muller (MD Anderson Cancer Centre, University of Texas) and Jon Wakefield (Department of Biostatistics, University of Washington). Registration is now open. Full details (including programme and confirmed speakers) are available at www.warwick.ac.uk/go/bicss .

International workshop on bayesian inference and maximum entropy methods in science and engineering, MAXENT 2006, CNRS, Paris, France, July, 8th-13th, 2006.

The Tweny sixth International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering will be held in France under the auspices of Centre national de la recherche scientifique (CNRS), Universit de Paris-sud, Orsay and cole suprieure d'lectricit (Suplec). MaxEnt 2006 strives to present Bayesian inference and Maximum Entropy methods in data analysis, information processing and inverse problems from a broad range of diverse disciplines: Astronomy and Astrophysics, Geophysics, Medical imaging, Non Destructive Evaluation, Particle Physics, Physical and Chemical measurement techniques, Economics and Econometrics. Special interest will be given to Bayesian inference applications in Inverse problems, Time series and image analysis, Source Separation and Data and information fusion with application areas such as X-ray, Diffractive, Diffusive Imaging and Quantum Tomographic. The workshop includes a one day tutorial session, state of the art lectures, invited papers, contributed papers, and poster presentations. The official languages will be French and English. Selected papers by the program committee will be edited and published in a book. All the papers will be in English. Contributed papers relating the above topics are being solicited. Especially encouraged are papers whose content is novel, either as to approach or area of application. Abstracts (one page of about 400 words) of the proposed papers should be received by the conference secretariat on Mars 01, 2006. See http://www.maxent2006.org for more details.

2nd SIPTA SCHOOL ON IMPRECISE PROBABILITIES, Madrid, July 24-28, 2006.

The Second SIPTA Summer School on Imprecise Probabilities will take place in the Headquarters of the Rey Juan Carlos University Foundation, in Madrid (Spain), on July 24-28, 2006. The school is intended as a wide and deep introduction to imprecise probability topics, both theoretical and applied. The topics covered will be:

- The Imprecise Dirichlet Model (Jean-Marc Bernard, Université Paris V).
- Predictive inference with imprecise probabilities (Gert de Cooman, Ghent University).

- Non-additive measures and applications on decision theory (Jean-Yves Jaffray, Université Paris VI).
- Coherent lower previsions and their behavioural interpretation (Enrique Miranda, Rey Juan Carlos University).
- Knowledge discovery from data sets under weak assumptions: the case of prior ignorance and incomplete data (Marco Zaffalon, IDSIA).

The deadline for pre-registration is March 31, 2006. Notification of acceptance will be made shortly after. People interested in participating should also submit a short CV (no longer than 2 pages). You can find all the relevant information on http://bayes.escet.urjc.es/ emiranda/sipta . If you have any questions or remarks, please contact Enrique Miranda, at `enrique.miranda@urjc.es` .

New issue of the eletronic journal BAYESIAN ANALYSIS:

A new issue of the new electronic journal Bayesian Analysis has been published at `http://ba.stat.cmu.edu`. The issue includes the following articles:

- Deconvolution in High-Energy Astrophysics: Science, Instrumentation, and Methods, by David A. van Dyk, Alanna Connors, David N. Esch, Peter Freeman, Hosung Kang, Margarita Karovska, Vinay Kashyap, Aneta Siemiginowska, Andreas Zezas.
  - Comment on article by van Dyk et al., by Ji Meng Loh and Andrew Gelman.
  - Rejoinder, by David A. van Dyk and Hosung Kang.
- Inferring Particle Distribution in a Proton Accelerator Experiment, by Herbert K. H. Lee, Bruno Sansó, Weining Zhou, David M. Higdon.
- Bayesian nonparametric estimation of the radiocarbon calibration curve, by Caitlin E. Buck, Delil Gomez Portugal Aguilar, Cliff D. Litton and Anthony O'Hagan.
- Who Wrote Ronald Reagan's Radio Addresses?, by Edoardo M. Airoldi, Annelise G. Anderson, Stephen E. Fienberg, and Kiron K. Skinner.
- Model-based subspace clustering, by Peter D. Hoff.
- A One-Pass Sequential Monte Carlo Method for Bayesian Analysis of Massive Datasets, by Suhrid Balakrishnan and David Madigan.
- Conjugate Analysis of the Conway-Maxwell-Poisson Distribution, by Joseph B. Kadane, Galit Shmueli, Thomas P. Minka, Sharad Borle, and Peter Boatwright.
- Misinformation in the conjugate prior for the linear model with implications for free-knot spline modelling, by Christopher J. Paciorek.

The journal is sponsored by the International Society for Bayesian Analysis (ISBA). Its founding editors are Alicia Carriquiry, Phil Dawid, David Heckerman, Xiao-Li Meng, Christian Robert, Fabrizio Ruggeri, and Dalene Stangl. Rob Kass is serving as Editor-in-Chief, Herbie Lee is Managing Editor, Marina Vanucci is Deputy Editor, and Pantelis Vlachos is System Managing Editor.

Bayesian Analysis seeks to publish a wide range of articles that demonstrate or discuss Bayesian methods in some theoretical or applied context. The journal welcomes submissions involving presentation of new computational and statistical methods; reviews, criticism, and discussion of existing approaches; historical perspectives; description of important scientific or policy application areas; case studies; and methods for experimental design, data collection, data sharing, or data mining. Evaluation of submissions is based on importance of content and effectiveness of communication.

The aim is to provide reports to authors within 10 weeks of submission on at least 80% of articles submitted.▲

INTERNATIONAL SOCIETY FOR BAYESIAN ANALYSIS

Executive Committee

President: Alan Gelfand
Past President: Sylvia Richardson
President Elect: Peter Green
Treasurer: Bruno Sansó
Executive Secretary: Deborah Ashby

Program Council

Chair: Kerrie Mengersen
Vice Chair: Peter Müller
Past Chair: José Miguel Bernardo

Web page:
http://www.bayesian.org

Board Members

Carmen Fernandez, Valen Johnson, Peter Müller, Fernando Quintana, Brad Carlin, Merlise Clyde, David Higdon, David Madigan, Michael Goldstein, Jun Liu, Christian Robert, Marina Vannucci.

EDITORIAL BOARD

Editor

J. Andrés Christen <jac@cimat.mx>

Associate Editors

Annotated Bibliography
Marina Vannucci <mvannucci@stat.tamu.edu>

Applications
Catherine Calder <calder@stat.ohio-state.edu>

Interviews
Brunero Liseo<brunero.liseo@uniroma1.it>

News from the World
Alexandra M. Schmidt <alex@im.ufrj.br>

Software Review
Ramses Mena <ramses@sigma.iimas.unam.mx>

Student's Corner
Robert Gramacy <rbgramacy@soe.ucsc.edu>

Bayesian History
Antonio Pievatolo <marco@mi.imati.cnr.it>

ISBA/SBSS Archive for Abstracts

All authors of statistics papers and speakers giving conference presentations with substantial Bayesian content should consider submitting an abstract of the paper or talk to the ISBA/SBSS Bayesian Abstract Archive. Links to e-prints are encouraged. To submit an abstract, or to search existing abstracts by author, title, or keywords, follow the instructions at the abstract's web site,

http://www.isds.duke.edu/isba-sbss/