THE ISBA BULLETIN

Vol. 15 No. 4 Dec 2008

The official bulletin of the International Society for Bayesian Analysis

A *last* MESSAGE FROM THE PRESIDENT

by Christian Robert
ISBA President
xian@ceremade.dauphine.fr

Season Greetings to all ISBA members and ISBA Bulletin readers! I hope that, despite the black clouds accumulating on various horizons, you found the time and stamina for celebration and resourcing to carry you onwards in 2009! Looking from my desk at the bright sun reflected on the snow as I am writing this final message, I still do feel confident that somehow somewhere we will find those resources within ourselves to carry on forward.

Since those are indeed my last words (as ISBA President!), I want to take the opportunity to thank all those who keep ISBA alive and thriving, at one level of involvement or another. Besides the Board members, I am very grateful to both the Treasurer Gabriel Huerta and to the Executive Secretary Robert Wolpert for effectively running ISBA. Both Robert's and Peter Green's involvements in the revamping of the IBSA website are beacons of this dedication to the society. (Accepting the heavy and unrewarding responsibility of the ISBA finances obviously sounds like the ultimate dedication!) I hope we can keep enlarging and improving the ISBA website to make it as lively and attractive as possible. In particular, I will soon write in the Bulletin about the possibilities of creating a wiki as well as syndicating Bayesian blogs. *Continue in page* 2.

A MESSAGE FROM THE EDITOR

by Raphael Gottardo
raphael.gottardo@ircm.qc.ca

Here we are, another year has passed already! I will start by wishing you all the very best for the new year. In this December issue, you will find an interesting article by Howard P. Edwards depicting the results of an interesting survey on graduate courses in Bayesian statistics. As you will see there is definite room for improvement. To complement this report, you will find the usual history, application, bibliography sections, and an interesting paragraph on "MISCONCEPTIONS ON BAYESIANISM" from our dear Christian Robert. I would also like to remind you that the ISBA bulletin exists thanks to its readers, and by that I mean YOU. In order for the bulletin to become even better, we need to make sure that 1) we have more and more readers (including more and more ISBA members) and 2) we receive more and more contributions. So as your first resolution for the new year, you can start by sharing the bulletin with whoever you think might be interested, even if they are not Bayesian! Then your second resolution will be to contribute to the Bulletin. I am sure that you have noticed that many issues of the bulletin are missing a few sections. This is simply because it is hard to find contributors and when we do find them we do not always receive the contributions on time. In short, we would love to have more contributions, so that we can be sure to always fill up all the sections of the bulletin. So feel free to contact me or any of the associate editors with your contributions. ▲

In this issue

WORDS FROM THE PRESIDENT, *Continued from page 1.* We should also think seriously of setting discussion forums linked to published papers and technical reports, starting with the obvious, namely papers published in Bayesian Analysis, on a distinct website.

I also take the availability of this tribune to congratulate Brad Carlin for doing such a great job with Bayesian Analysis. It continues to grow in visibility and in substance. The moves he made earlier this year to join forces with IMS and to get DOI's for papers published in Bayesian Analysis are important steps towards the long-term viability of the journal. An open issue is the maintenance of the BA website, which is currently operated on a single-person good-will basis, so is not viable (if working extremely well). This links with the issue of securing revenues to back up the production of BA and I wish Brad all the best in his endeavours! Looking back on 2008, I also want to repeat my earlier thanks to all the members involved in ISBA 2008 on Hamilton as this was really a terrific and unique conference that contributed so much to give the right image of our society as an active and innovative statistical community, not to mention a fun one as well. So many thanks, mates!

I also want to congratulate my friend of many years, Peter Müller, for his election to the 2010 presidency of ISBA, and I wish both Mike West and him well. I am obviously certain that, in opposition to mine!, they will have a clear and positive impact on the society and its growth in the future years. I am looking forward working with them in the coming year, especially with regards to publication and membership matters. As mentioned earlier, I think we should aim at further involving communities that are naturally using Bayesians techniques but do not consider themselves as statisticians, as for instance in Astronomy. I have no clear idea how we can do that on a large scale, but contributing by attending conferences at the interface and giving tutorials is certainly a step in this direction.

Let me conclude by repeating my best wishes for 2009 and by thanking all of you for making our Bayesian society so lively and enjoyable, and for being Bayesian such a rewarding experience! ▲

MISCONCEPTIONS ON BAYESIANISM

by Christian Robert
***ISBA* President**
xian@ceremade.dauphine.fr

It seems to me that the most common attack against Bayesianism relates to its sectarian aspects. While unjustified, this criticism is grounded and long-lasting for several reasons. The first one is the Bayesian claim to universality: no other branch of Statistics attempts to cover so generically all branches of Statistics, from estimation to testing, from design to non-parametrics, from minimax theory to graphical modelling. That Bayesian principles can integrate so smoothly all kinds of statistical optimalities may feel like propaganda to non-Bayesians, even though there are many proofs of this efficiency, from consistency to admissibility, from Dutch-book arguments to exchangeability (see de Finetti).

The second reason is that no other (major) approach to Statistics is so strongly anchored on philosophical principles. This deep connection with Philosophy sounds to me like a strong added value, in particular for analysing the nature of learning (see Savage and Dawid) and the influence of a prioris (back to Laplace and Poincaré), but the threads linking modern Statistics to Mathematics and to Informatics may make this additional link (and the argumentative discussions involved in some Bayesian papers) seem old-fashioned and un-scientific. (It is also true that the Bayesian literature abounds with philosophical arguments that are not always of the highest quality.) The essential fact that a Bayesian analysis relies on the choice of a prior distribution inevitably opens the door to the sectarian criticism, even though it is as well an (the!) inevitable part of the Bayesian principles. That two different statistical analyses of the same data could conduct to two different conclusions is seen by some as a major default in the theory, while it is and should be unavoidable. The criticism is that the use of a prior is un-scientific or un-objective (or un-falsifiable in Popperian terms) and that this choice is based on tenets only understandable to members of the sect…

A third and related reason is that Bayesians have developed along the years a real sense of community. For one thing, no other (major) branch of Statistics has members so naturally gathered under a common denomination (e.g., likelihoodists?! Basu had dubbed the Indian construct likelihood-wallah on those using likelihood, but this has obviously not stuck! Fiducians could be the closest to this fame, but I am not even sure the name exists.) There are many good things in having a feeling of community and this includes real benefits in terms of collaborative research and in keeping the unitarian perspective of the statistical approach, but one drawback of communities is that people outside the community may naturally feel dismissive, ostracised, excluded, suspicious, jealous, or, in the most extreme cases, antagonistic and belligerent, i.e. anti-Bayesians. (This is a point shared with religions and sects, most obviously, that those not "in" are automatically "out".) The fact is that this community has also developed some traditions that could be dubbed "rituals", like having meetings in sea resorts (in Spain and elsewhere), and alas rarely in cold and mountainous places (even though MCMC'ski could be the start of a new tradition!), with a strong emphasis on partying! Again, nothing wrong with adding a few extra good reasons to attending conferences, but this may not seem right to outsiders who have never attended a poster session at a Valencia meeting that starts in the hotel bar at 10am and ends up at two in the morning with people still loudly arguing around papers. Launching Bayesian Analysis was (in retrospect) a great idea, even though I remember it being fiercely debated at several Valencia meetings (and I must confess I voted against it at the time!), but it also strengthens the (wrong) impression of a closed group with its own agenda "only publishing in its own journals".

The last reason I want to point out is the fact that Bayesianism draws its name from one man, Thomas Bayes, and that, while there are good reasons for this filiation, this is also a feature shared with sects! As any other branch of Statistics, Bayesian theory has been built on the work of many and this singling out one person as the founder of the theory is unfortunate. While it seems a wee late to switch the denomination, I really think the abuse of the (maybe apocryphical) picture of the Reverend on our webpages and in our talks, and of what can be construed as a "cult of personality" (when considering that ISBA manages a fund for looking after Bayes's tomb in Bunhill cemetery) should cease to be part of our attitude. It would certain help in reducing the sectarian libels.

PS-The column Dr Fisher's casebook in the recent December issue of Significance is quite representative of these misconceptions on Bayesianism, ranking Bayesians as born-again fundamentalists...

PPS-For all those readers who want to comment on this tribune, the discussion can be conducted via the blog site http://xianblog.wordpress.com/2008/12/22/misconceptions-on-bayesianism/

ANNOTATED BIBLIOGRAPHY

CLIMATE CHANGE DETECTION AND ATTRIBUTION

Jonathan Rougier
j.c.rougier@bristol.ac.uk

Climate science is a very natural place to find Bayesian inference: there is a lot of uncertainty, not much replication, diverse sources of information, and a lot of structure and prior knowledge. But it is also very challenging: climate science is definitely 'big science', in terms of the resources required (satellites, ships, supercomputers) and of the size of the datasets acquired and generated. It is a very fertile area for statisticians interested in Computer Experiments, like myself, because it challenges the convenient view that the physical model is moderately-sized, reasonably accurate, and cheap to evaluate.

In this bibliography I have focused on one particular area in Climate Science, climate change detection and attribution, also known as 'fingerprinting'. Fingerprinting asks whether the observed patterns in climate can be attributed to particular external causes, known as 'forcings' (like changing concentrations of atmospheric CO_2) in the presence of natural climate variability. It is a Bayesian success story, with some interesting methodological questions. Several strands have emerged, so I have listed the papers by strand, rather than chronologically. I will focus on principles, rather than implementation (although much could be said about the latter). I must apologise to the authors for grossly simplifying their analyses in the interests of a compact and consistent summary.

Background and summary

F.W. Zwiers and H. von Storch, 2004. On the role of statistics in climate research, *Int. J. Climatol.*, **24**, 665–680.

This paper nicely summarises the statistical particularities of empirical climate research: "that the climate system has a large number of components and that it is impossible to conduct laboratory experiments with the earth system" (p. 674). Explains how climate datasets can be so complicated, and how that affects climate inference. We note their observation that "discussion about statistical methodology in the climate sciences is generally not very deep" (p. 675).

G.C. Hegerl and F.W. Zwiers *et al*, 2007. Understanding and attributing climate change. Ch. 9 in Solomon *et al* (eds), *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group 1 to the Fourth Assessment Report of the IPCC.

Climate Science is lucky to have regular state-of-the-art summaries as part of the Intergovernmental Panel of Climate Change (IPCC) process. This is the most up-to-date summary of fingerprinting research, both Frequentist and Bayesian. See especially sec. 9.1 for an outline, Fig. 9.2 for examples of fingerprints, and sec. 9.4.1.8 for 'Remaining uncertainties'. This is a good place to go for findings and implications.

Appendix 9.A has a **summary of fingerprinting.** Briefly, this is based around the regression

$$y = Xa + u$$

where y is the observed climate signal, the columns of X contain patterns attributable to particular external forcings (the forcings are mapped into patterns using a climate model), a is a vector of unknown amplitudes, and u is a mean-zero vector with specified covariance matrix C, representing the climate system's natural variability. *Detection* involves rejecting $H_0 : a = 0$ in favour of $H_1 : a > 0$. *Attribution* is trickier. One approach involves subsequently not rejecting $H_0 : a = 1$ in favour of $H_1 : a \neq 1$. Here it is necessary to assert that all of the external forcings are represented individually, that they combine additively, and that they are correctly scaled. The observed climate signal may be attributed to forcing i if $H_0 : a_i = 0$ is rejected in favour of $a_i > 0$.

D.A. Randall and R.A. Wood *et al*, 2007. Climate models and their evaluation. Ch. 8 in Solomon *et al* (eds), *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group 1 to the Fourth Assessment Report of the IPCC.

Another great chapter from the most recent IPCC report: it helps to know a bit about climate models. It's all interesting, but see especially the Executive Summary, secs 8.1 and 8.2, and FAQ 8.1. The issue of how sub-grid-scale effects are handled is crucial, since the approximation of these effects is the main source of uncertainty when climate models are used to predict actual climate.

K. Hasselmann, 1976. Stochastic climate models part I: Theory. *Tellus*, **28**, pp 473–485.

A clear explanation of how climate models operate, and the way in which sub-grid-scale effects are represented; secs 1 and 2 especially. Prof. Hasselmann will reappear below . . .

Frequentist approaches

There are *lots* of papers using the standard Frequentist approach, based on hypothesis testing for a; I've just chosen a couple, for orientation.

P.A. Stott *et al*, 2001. Attribution of twentieth century temperature change to natural and anthropogenic causes. *Clim. Dyn.*, **17**, 1–21.

There are many choices to be made before the hypotheses are tested. This paper provides a clear statement of these choices (sec. 2, quite literally the 'small print'). Also a graphical representation of detection and attribution for two forcings (two columns in X), based on the Confidence Region (CR) of a and the Confidence Intervals (CIs) of a_1 and a_2. If the 95% CR lies in the positive quadrant, then we have detection, and if this CR contains $(1, 1)$, so much the better. If the 95% CI of a_i lies above 0, then we have an attribution to forcing i. Obviously, multicollinearity presents a challenge; cruelly, nature has arranged that the patterns due to the solar cycle and to volcanic SO_2 emissions are not orthogonal.

R.A. Levine and L.M. Berliner, 1999. Statistical principles for climate change studies. *J. Clim.*, **12**, 564–574.

This paper (which preceeded Stott *et al*, 2001) was written to help climate scientists understand the hypothesis testing framework better. It explains that inference about a can be based on the value of the GLS estimator $\hat{a}(y)$, since $\hat{a}(y) \sim \mathrm{N}(a, \{X^T C^{-1} X\}^{-1})$. Oddly, a two-tailed alternative hypothesis is proposed for detection, $H_1 : a \neq 0$. Two influential points are made: (i) it is hard to describe our confidence in $H_0 : a = 1$ when we fail to reject it; and (ii) the two-step procedure where attribution is tested after rejecting $H_0 : a = 0$ makes it hard to set the significance level of the attribution test. (In fact, there are several hypotheses tests on the way to attribution, which only strengthens this point.) The authors note that (i) can be addressed with a test for 'geoequivalence' (p. 571) for which $H_0 : a \neq 1$ and $H_1 : a = 1$. They also note that points (i) and (ii) together can be addressed by a single test of $H_0 : a = 0$ versus $H_1 : a = 1$.

Bayesian improvements

There are two strands in the Bayesian treatment of fingerprinting. One is to use Bayesian methods to make improvements to the Frequentist approach. Hence, there is a marked preference for 'more objective' priors.

S.S. Leroy, 1998. Detecting climate signals: some Bayesian aspects. *J. Clim.*, **11**, 640–651.

The coefficients a are treated as uncertain, and assigned a Gaussian prior. When combined with a Gaussian u, this gives a Gaussian posterior with mean a_m and variance A_m. A value of $(a_m)^T(A_m)^{-1}a_m \gg 1$ is indicative of detection. In the special case where the prior precision for a tends to zero, the standard test statistic for detection is recovered (two-tailed). A theoretical innovation is to consider several models of varying accuracy, leading to u being interpreted as model error plus natural variability. But in practice only internal variability is used to estimate C for each model. Different models are scored using an approximation to their marginal likelihoods (hinting at Bayes factors). It is stressed that marginal likelihoods cannot be mapped into posterior probabilities because the models do not form a partition (p. 641, expressed differently).

L.M. Berliner, R.A. Levine, and D.J. Shea, 2000. Bayesian climate change assessment, *J. Clim.*, **13**, 3805–3820.

Emphasises that confidence is not probability, and that detection is more naturally addressed by computing the *probability* that a lies outside a small region $\mathcal{D}^c$ around 0, given the data y (they write $\mathcal{D}$ for $\mathcal{D}^c$, however). Consistently with Levine and Berliner (1999), $\mathcal{D}^c$ is symmetric about zero, rather than being restricted to positive quantities. The prior for a is a mixture of two Gaussian distributions centred at 0 and 1, where the weight and the two variances quantify prior information about the causes of the observed climate signal. Attribution requires both that the probability of a lying in a region around 1 is large *and* that the posterior weight attached to the $a \approx 0$ component of the prior is small. Plug-in estimates are used for the variance matrices, and a robust approach is used to bound the posterior for a according a range of values in the prior.

T.C.K. Lee, F.W. Zwiers, G.C. Hegerl, X. Zhang, and M. Tsao, 2005. A Bayesian climate change detection and attribution assessment. *J. Clim.*, **18**, 2429–2440.

The advantage of a *probabilistic* assessment of the evidence for detection and attribution is stressed, and prior knowledge is somewhat downplayed. The non-detection region $\mathcal{D}^c$ now includes all negative values as well as small positive ones. A computer exeriment is used to select an informative prior for a (expressed as a mixture, as in Berliner *et al*, 2000), and two other priors are also tried, one more concentrated and one less so. An interesting suggestion is to report the posterior probability of detection and attribution in terms of the Bayes factor

$$B = \frac{\Pr(a \in \mathcal{D} \mid y)}{\Pr(a \notin \mathcal{D} \mid y)} \Big/ \frac{\Pr(a \in \mathcal{D})}{\Pr(a \notin \mathcal{D})},$$

using the Bayes factor scale to label the results ($B > 150$ is 'very strong' evidence). The idea is to neutralise the prior (i.e. to require stronger evidential support for strong priors), but to me this seems somewhat involved when one could simply report $\Pr(a \in \mathcal{D})$ and $\Pr(a \in \mathcal{D} \mid y)$.

Bayesian rethinking

The second strand involves a more fundamental rethink. It's characteristic of this strand that the distinction between detection and attribution is blurred, with evidence being used to arbitrate between competing theories. Also there is less emphasis on 'objectivity', and more on treating the Bayesian approach as an opportunity for experts to communicate their particular judgements.

K. Hasselmann, 1998. Conventional and Bayesian approach to climate-change detection and attribution. *Q. J. R. Meteorol. Soc.*, **124**, 2541–2565.

A paper in transition from the Frequentist to the Bayesian approach. The hypothesis H is "that there exists an anthropogenic climate-signal in the data" (p. 2555), and the evidential support for H, namely $\Pr(H \mid y)/\Pr(H)$, is expressed in terms of $\Pr(H)$ and the likelihood ratio $B = \pi(y \mid H)/\pi(y \mid \bar{H})$—Bayes factors are not mentioned explicitly. To address attribution, H is considered separately for each pattern in X, i.e. $\bar{H}$ and H are operationalised as $a_i = 0$ and $a_i = 1$ for pattern i. Problems with the tails under the conventional model (Gaussian u) require that the B is replaced by a somewhat *ad hoc* truncation. Multiple lines of evidence can be combined by taking the product of their likelihoods "since the variables are statistically independent" (p. 2559), but what this actually means is not discussed.

R. Schnur and K. Hasselmann, 2005. Optimal filtering for Bayesian detection of climate change. *Clim. Dyn.*, **24**, 45–55.

We now have two competing hypotheses, which might be that the climate signal is due to: H_0, natural forcings alone; versus H_1, natural and anthropogenic forcings. The Bayes factor $B_{10} = \Pr(y \mid H_1)/\Pr(y \mid H_0)$ is introduced explicitly. Another hypothesis H_2 is also considered, supporting the idea of attribution, in the sense that H_0 is natural forcing alone, and H_1 and H_2 are competing hypotheses about anthropogenic forcings. Bayes factors for different types of evidence (y, y', *etc*) are combined by multiplication, now with an appropriate *caveat*. Posterior probabilities for H_0, H_1, and H_2 are computed for three 'canonical' priors: 'Uninformed', 'Skeptic', and 'Advocate'. However, this would only be appropriate if the three hypotheses formed a partition, a point anticipated by Leroy (1998), so this analysis is suspect. An important innovation is the inclusion of a term for model error, which makes the model-based pattern different from the actual climate pattern; the variance matrix of this term is estimated from an ensemble of different models.

S-K Min, A. Hense, H. Paeth, and W-T Kwon, 2004. A Bayesian decision method for climate change signal analysis. *Meteorol. Z.*, **13**, 421–436.

The final step: choosing between H_0, H_1, and H_2 is treated as a decision problem. It is explained that the use of posterior probabilities to make the decision is equivalent to minimising expected loss under a 0–1 loss function. It is emphasised that choosing among the hypotheses is not the same as 'attribution', because there is no causal explanation (Judea Pearl is cited). The Bayes factors are no longer simply likelihood ratios, because more sophisticated statistical modelling means that uncertain quantities in the numerator and denominator must be integrated out. The use of H_0 (natural forcings alone) as a null hypothesis when there are multiple alternatives is discussed, particularly with reference to presenting the posterior probabilities. Plug-in estimates of the variance matrices are used, subject to an uncertain scaling, α. Bayes factors are presented (on the $\log_e$ scale), as well as posterior probabilities for different priors (again, suspect if the hypotheses do not form a partition). Figures show how the evidence favouring H_1 over H_0 has accumulated over the period 1980–2000, with a sensitivity analysis for α (which does not materially affect the results).

Reflections

The question we would like to address is whether the observed climate signal is more consistent with purely natural forcing, or with natural plus anthropogenic forcing. It is hard to see how this question can be addressed by a hypothesis test on a in a GLS regression, whether it is performed in a Frequentist or a Bayesian way. Rather, it seems natural to analyse this question using Bayes factors: where does the weight of evidence lie? Bayes factors should appeal to all camps. They are likelihood-based, and they may be compelling enough (e.g. 10 decibans or more, on I.J. Good's scale) to obviate the need for prior probabilities. As indicated above, it seems not have been appreciated that Bayes factors can be mapped into posterior probabilities only if the set of hypotheses form a partition. The construction of such a partition would certainly clarify the whole issue of detection and attribution.

One thing that I have not elaborated on is the complexity of the actual analysis: forming the y vector and the X matrix, and estimating the variance matrix C. We can see all of the choices made in these steps as auxiliary hypotheses, and they have the usual effect of blurring the conclusions we can draw about our primary hypothesis. One auxiliary hypothesis that I would like to see treated more explicitly is the link between the model evaluations and the climate system. For the regression approach and its Bayesian extension, the two seem to be treated as synonymous (subject to adjustment for location and scale). My view is that it is very difficult to interpret a result which is 'conditional upon the model being correct', when the reason that climate models are under such active development is precisely because they are not correct (or, rather, not yet correct enough). The Bayesian approach allows us to generalise, and Schnur and Hasselmann (2005) include a variance contribution for model-error. The issue of how climate models are judged to be informative about actual climate is a fascinating one, and I hope that the continuing development of Bayesian methods for detection and attribution will provide further impetus here.

BAYESIAN HISTORY

Brazilian Bayesians

by Hedibert Freitas Lopes
The University of Chicago Booth School of Business
hlopes@ChicagoBooth.edu

The December 2000 issue of the ISBA Bulletin (volume 7, number 4) included an article entitled *Bayesians in Brazil* by Sergio Wechsler and Basílio de Bragança Pereira. The article ends with the following paragraph.

> *Bayesian research activities are now very active in Brazilian universities with strong interaction with universities abroad. Last year the Brazilian Bayesians decided to create a Brazilian Chapter of ISBA. The new Chapter made the decision to host the First Latin American Bayesian Meeting (I COBAL) to be held very probably on January 2002 in Brazil.*

This paragraph succinctly and precisely described what would be the backbone of the Bayesian community in Brazil during the first decade of the XXI century. Indeed, the First COBAL, along with the Sixth Brazilian Meeting for Bayesian Statistics (EBEB VI, in Portuguese), was held in Ubatuba, in São Paulo State in February 2002. The Brazilian Chapter of ISBA, named *ISBRA* and chaired by Márcia Branco sponsored three other EBEB took place during those eight years: EBEB VII in São Carlos (SP, 2004), EBEB VIII in Búzios (RJ, 2006) and EBEB IX in Maresias (SP, 2008) respectively. The last two editions honored two of the most influential Bayesians in Brazil, Hélio Migon and Carlos Pereira, respectively, during their sixtieth birthdays.

The Bayesian paradigm has been gaining an increasing number of supporters worldwide, especially amongst practitioners both within and outside academia, for a number of reasons. The trend in Brazil follows suit and more researchers are adhering to (several levels of) Bayesianism. An encouraging example is the recently founded PhD Program in Statistics in the Federal University of Rio de Janeiro (UFRJ), led primarily by Hélio Migon and Dani Gamerman, where 12 PhD theses and 30 MSc theses were granted between 2004 and 2008 and 2001 and 2008, respectively. Additional well known and successful research groups are, amongst a few others, Carlos Pereira's Bayesian Biostatistics group at the University of São Paulo (USP), Renato Assunção and Rosangela Loschi at the Federal University of Minas Gerais (UFMG) and Josemar Rodriguez and Francisco Louzada-Neto at the Federal University of São Carlos (UFSCAR). Between 2001 and 2008, Carlos graduated 20 students (8 PhD and 12 MSc), Renato and Rosangela graduated 25 MSc students and Josemar and Francisco graduated 37 students (6 PhD and 31 MSc). These more established Graduate Programs and Research Groups are paving the way for new ones including, amongst a few others, Marinho Filho from USP's Institute of Mathematical and Computational Sciences in São Carlos (ICMC-USP) with 16 students (3 PhD and 16 MSc), Roseli Leandro from USP's Luiz de Queiroz School of Agriculture in Piracicaba (ESALQ-USP) with 12 students (4 PhD and 8 MSc) and the Department of Electrical Engineering, Pontifical Catholic University in Rio de Janeiro (DEE-PUC-RIO). When compared to the US, for instance, 30 PhD degrees and 100 MSc degrees over a decade might look minuscule. For example, the Department of Statistical Sciences, former Institute of Statistical and Decision Sciences, at Duke University, one of the world's most important Bayesian think-tanks, granted 40 PhD degrees from 2001 to 2008. Nonetheless, these numbers point towards an extraordinary and encouraging improvement when compared to the number of Brazilian graduates over the last quarter of the twentieth century.

At least five important points highlight the recent positive receptiveness and respect of the international Bayesian community regarding a more mature and organized Bayesian community in Brazil. Firstly, more students go to the United States and Europe (predominantly England) to pursue their PhD degrees and are fully financed by universities abroad. For several generations Brazilians studying abroad were primarily financed by the Brazilian government through its foundations used to foment scientific and technological development, such as CAPES and CNPq. Secondly, researchers who obtained their PhD degrees in Brazil are able to find postdoctoral positions abroad both in academia and in the industry. Thirdly, Brazilians have also served as Officers, Board Members, Committee Members and Bulletin Editor of ISBA. For instance, Dani Gamerman, Lurdes Inoue and Alexandra Schmidt were, or still are, Nominating, Savage and Lindley Award and Board Committee members. Alexandra Schmidt is also the 2010 Program Chair. Forthly, more Brazilians researchers are members of editorial committees of first tier scientific journals. Finally, many Brazilian Bayesians were welcomed into US and European institutions, mainly US academia. These include Carlos Carvalho (The University of Chicago), Marco Ferreira (University of Missouri at Columbia), Lurdes Inoue (University of Washington), Telba Irony (Food and Drug Administration, USA) and Guilherme Rosa (University of Wisconsin at Madison), to name but a few.

We focused mainly on the growth of Bayesianity in Brazil over the last 8 years. We believe that, when coupled with Wechsler and Pereira's aforementioned article, this short note will help convince and alert the unaware reader, particularly the new generations of Bayesians in Brazil, that high level research is being done by Brazilian Bayesians as we speak and that such achievement represents the combined and continuous efforts of several generations of researchers over the last 40 years. I would like to thank my mentors and friends Helio Migon and Dani Gamerman, also two of the most influential Brazilian Bayesians, whose invaluable remarks have considerably improved this note.

We would like to finish with a paragraph that could be used a few years down the road by younger Brazilian Bayesians to confirm or trash our predictions. We forecast more structured interactions and collaborations amongst the above research groups at various educational and research levels, as well as their active roles in the worldwide promotion of the Bayesian paradigm. Also, *No one can stop us* (quoting Peter Müller), for instance, from speculating that, maybe after a few more EBEB's, an ISBA world meeting will be held in the (still) "Marvelous City" of Rio de Janeiro. We humbly ask for the forgiveness of our friends from São Paulo, Minas Gerais and other parts of Brazil for being biased, but again we can not help ourselves from being Bayesians.

Useful websites:
ISBRA: http://www.ime.usp.br/~isbra

EBEB IX:
http://www.ime.usp.br/~isbra/ebeb/9ebeb/en
EBEB VIII: http://www.im.ufrj.br/ebeb8/Eng
EBEB VII:
http://www.ufscar.br/~des/ISBRA/7ebeb.htm

UFRJ: dme.ufrj.br/index_EN.html
USP: www.ime.usp.br/
UFMG: est.ufmg.br/portal
UFSCAR: www.ufscar.br/~des/
ICMC-USP: www.icmc.usp.br/
ESALQ-USP:
www.lce.esalq.usp.br/indexingles.html
DEE-PUC-RIO: www.ele.puc-rio.br/

STUDENTS' CORNER

Call for Dissertation Abstracts

Luke Bornn
l.bornn@stat.ubc.ca

Recent Ph.D graduates, having your dissertation abstract published is as simple as emailing it to the email address above. Publishing your abstract will not only provide exposure for your research, but it may potentially lead to collaborations with future colleagues. In addition, you are providing an important service to the Bayesian community by giving established researchers a taste of the interests of young researchers. Faculty, please encourage your students' participation.

NEWS FROM THE WORLD

Events

2009 Bayesian Biostatistics, Houston, Texas, 26-28th Jan. 2009.

Current and prospective users of Bayesian biostatistics are invited to join experts in the field for a three-day conference sponsored by the Department of Biostatistics at The University of Texas M. D. Anderson Cancer Center in Houston, Texas, USA. Attendees will have the opportunity to attend two courses on the first day of the conference (Monday). The Use of Bayesian Statistics in Clinical Trials, and Applications of Bayesian Methods to Drug and Medical Device Development. On Tuesday and Wednesday, invited presentations will cover a variety of topics, possibly including comprehensive decision modeling; prior distributions in clinical studies and drug development; what Bayesian methods can provide that traditional methods cannot provide; Bayesian methods in medical journals; Bayesian methods in epidemiology; Bayesian methods and medical ethics; how to build a cadre of Bayesian experts; why Bayesian methods are not more widely used; how to assure good quality in Bayesian methods; and guidelines for publishing Bayesian analyses.

$500 travel grants will be awarded to qualifying pre-doctoral students (post-docs are not eligible for this grant). For more information visit the website, http://www.mdanderson.org/departments/biostats/, or contact Lydia Davis *lbdavis@mdanderson.org*.

BISP6 - Bayesian Inference in Stochastic Processes, Bressanone/Brixen, Italy, 18-20th Jun. 2009.

In this workshop, we will bring together experts in the field to review, discuss and explore directions of development of Bayesian Inference in Stochastic Processes and in the use of Stochastic Processes for Bayesian Inference. There will be sessions on Markov processes, state-space models, spatial, empirical, birth-death and branching processes. Theoretical and applied contributions (for example queueing, population modelling, signal processing) are both welcome. The workshop will thus be of interest to workers in both Bayesian Inference and Stochastic Processes. For more information visit the website, http://www.mi.imati.cnr.it/conferences/bisp6.html.

6th Workshop on Bayesian Nonparametrics, Moncalieri (Turin), Italy, 21-25th Jun. 2009.

The aim of the 7th Workshop on Bayesian Nonparametrics is to highlight the latest developments in Bayesian Nonparametrics covering a wide variety of both theoretical and applied topics. The meeting will be held at the Collegio Carlo Alberto, a Research Institution housed in an historical building located in Moncalieri on the outskirts of Turin, Italy. For more information visit the website, http://bnpworkshop.carloalberto.org, or contact Pierpaolo De Blasi *bnp@carloalberto.org*.

O-Bayes09, University of Pennsylvania, Philadelphia, 6-9th Jun. 2009.

O-Bayes09, the 2009 International Workshop on Objective Bayes Methodology will take place at the Wharton School of the University of Pennsylvania, Philadelphia, PA, USA. The principal objectives of O-Bayes09 are to facilitate the exchange of recent research developments in objective Bayes methodology, to provide opportunities for new researchers to shine, and to establish new collaborations and partnerships that will channel efforts into pending problems and open new directions for further study. O-Bayes09 will also serve to further crystallize objective Bayes methodology as an established area for statistical research.

The workshop will consist of a series of invited talks followed by discussion and one or more sessions dedicated to contributed posters.For more information visit the website, http://stat.wharton.upenn.edu/statweb/Conference/OBayes09/OBayes.html, or contact Linda Zhao *lzhao@wharton.upenn.edu*.

Graduate Courses in Bayesian Statistics: Results of an Online Survey

by Howard P. Edwards
Institute of Information and Mathematical Sciences, Massey University, Auckland, New Zealand
http://iims.massey.ac.nz/people/h.edwards/

Keywords: Bayesian statistics, statistics education

Introduction

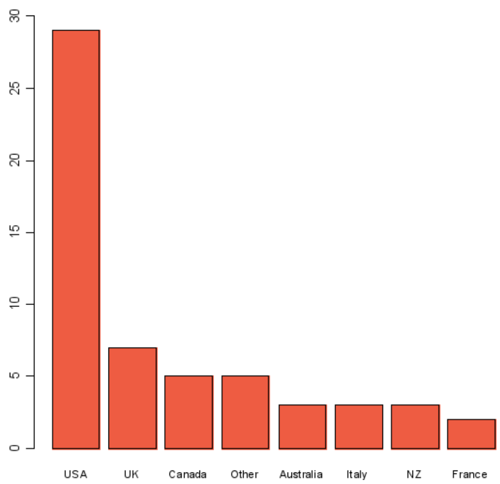

Figure 1: Number of Courses by Country

University level courses in Bayesian statistics or including Bayesian statistics as part of a larger topic (e.g. statistical inference) have been offered for many years from most universities with a Statistics department or group of equivalent size (eg a Mathematics department). For example, the author has taught a graduate level paper in Bayesian statistics at Massey University in New Zealand since 1979. However it is only in the last ten or so years that Bayesian methodology has become widespread and applicable to an increasing number of disciplines. Much of this is due to computational methods such as Markov chain Monte Carlo and the availability of software such as WinBUGS. Therefore it is timely to consider the teaching of Bayesian statistics and how it has been affected by these and other changes. From anecdotal evidence it appears that most teaching of Bayesian statistics takes place at the graduate level, either at Masters or PhD level, but to the best of my knowledge no attempt has ever been made to assess the scope and level of Bayesian teaching taking place. The purpose of this study was to try and take an international "snapshot" of graduate level courses in Bayesian statistics.

Methodology

In order to gain rapid responses across an international population in a fairly short time period, an online survey was used. The survey was opened on 13 March 2008 with notices posted to the following electronic mailing lists: bayes-news, allstat and anzstat. A total of 41 responses were received from this initial notice. The same notice was posted to the edstat mailing list on 23 April 2008 and a reminder notice was posted to the first three lists on 28 April 2008. This produced a further 16 responses yielding a total of 57 responses received by 19 May 2008 when this analysis was conducted. As this was a self-selecting survey with obvious sampling biases, the results are presented as summaries together with comments.

Result

Country

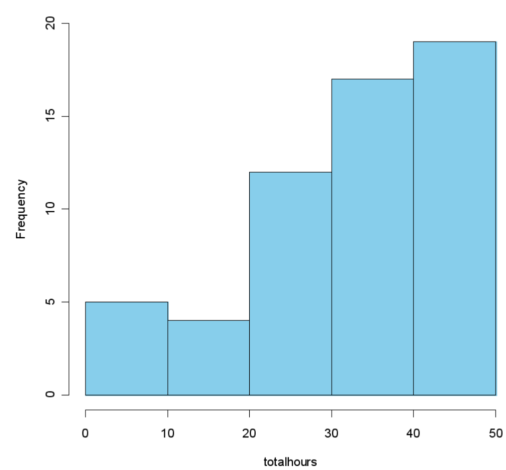

Figure 2: Total Number of Hours

Just over half of respondents (51%) were from the USA. This was followed by the UK and Canada with 12% and 9% respectively, with smaller percentages from Australia, Italy, New Zealand and France. Countries with a single respondent (merged as "Other" in Figure 1) were Aruba, Austria, Brazil, Mexico and Nigeria (although the individual responses from the Aruba submission suggested that this was a coding error and probably from Australia).

Entry Requirements and Student Background

Most courses (71%) required graduate status as an entry requirement. About 20% required little or no formal statistics background, about 40% required an intermediate level statistics background and the remaining 40% required either a Statistics major background or graduate level Statistics status.

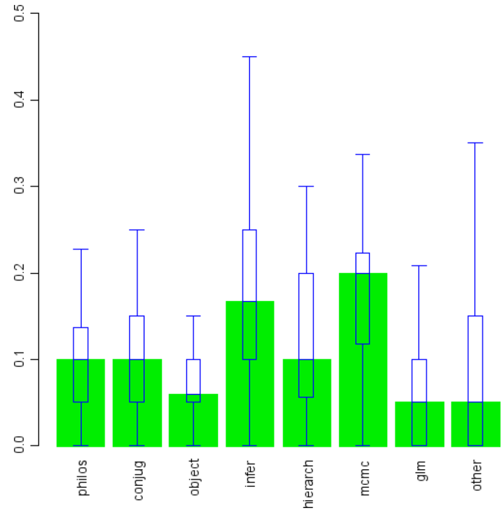

Figure 3: Distribution of Topics

Student Characteristics

Table 1: "Other" topics covered

Topic	No. of times topic mentioned
Modelling	9
Applications	6
Inference	5
Asymptotics/large sample theory	3
Bayesian networks	2
Missing data	2

About three quarters (77%) of courses catered for a specific group of students. Specifically, 42% catered for Statistics students and 21% catered for Biostatistics and Epidemiology students, with smaller numbers catering for Economics and Business students (7%) and Computer Science students (4%). It should be noted that some courses catered for more than one group, for example Statistics and Biostatistics students (5%).

Just over half (51%) of the courses consisted of Masters students or equivalent, with 39% consisting of PhD students. Courses with an "other" response consisted of both groups of students, which meant that 60% of courses included Masters students and 41% included PhD students. Only one course was aimed at working professionals which suggests that university departments are not delivering this material through short courses or other professional development programmes but rather by formal enrolment in a graduate level university programme.

Total Hours

Respondents were asked how many hours of formal course contact time were required as a total (rather than as e.g. hours /week). This total did not include optional activities or work done in the students' own time. Responses were categorised as 0 - 10 hours etc in ten hour classes with a final "40 hours or more" class. Figure 2 shows a histogram of the responses using class midpoints (5,15,25,35,45). The majority of offerings (63%) required 30 or more hours contact time which is consistent with the findings in 3.3 above, in that many graduate level university courses are taught using 3-4 hours of lectures per week over a 12 week semester. It is conceivable that some of the respondents reporting fewer than 10 hours total may have misread this item and reported a weekly value instead.

Topics Covered (and Not Covered)

Respondents were given a list of topics and asked to estimate the proportion of their course that was devoted to each topic. These topics were: Bayesian philosophy, conjugate families, objective Bayes, statistical inference, hierarchical models, MCMC models, generalised linear models, together with an "other" category - respondents selecting "other" were asked to list the relevant topic(s). Figure 3 shows the distribution of each proportion (as boxplots) together with the distribution of median proportions. "Other" covered a range of topics from measure theory and decision theory through to applications such as health economics, educational psychology models and critiquing of published journal articles using Bayesian methodology. These are summarised in Table 1. When asked "What topic(s) would you most like to include but currently don't have the time to cover?", the most common responses related to computer programming or use of computer software. Generalised linear models, model comparison and Markov Chain Monte Carlo were the next most commonly occurring responses. These are summarised in Table 2.

Table 2: Topics that respondents would like to cover but don't have time

Topic	No. of times topic mentioned
(Further) computational methods	10
Generalised linear models	6
MCMC	5
Model comparison and model fit	5
Spatial applications	4
Hierarchical models	3
Missing data	2

Textbooks

Respondents were asked to list up to three textbooks used in their course. 9 respondents either did not answer this question or else indicated that a textbook was not used either by stating this explicitly or by stating that their own notes were used. Many of the respondents listed more than one text, and as these were sometimes listed in alphabetical order it could not be assumed that the ordering related to usage or preference. The responses are summarised in two ways: (a) total number of "mentions" of each textbook and (b) a total score based on the scoring rule "mentioned first" = 3, "mentioned second" = 2 and "mentioned third" = 1. These summaries are given in Table 3. Both summaries indicated that Gelman et. al. (2004) was the most popular text, with both number of mentions and total score more than twice that of the second most popular text (Robert(2007)). No attempt was made to distinguish between editions of a textbook, and in the references only the most recent edition of each text is given.

Table 3: Prescribed textbooks

Text	No. of times text prescribed	Total score using "321" scoring
Gelman Carlin Stern & Rubin	19	55
Robert	7	17
Berger	5	13
Bernardo & Smith	4	11
Carlin & Louis	4	10
Congdon (either)	4	9
Lee	4	8
Spiegelhalter	4	4
Abrams & Myles		
Albert	2	4
Gamerman & Lopes	2	3
nonBayesian	2	
Other	12	
None	14	

Software Usage

Of the 57 responses received, only 7 did not indicate any software usage. Not surprisingly perhaps the most popular packages used were R (37 responses) and WinBUGS (28 responses). OpenBUGS, R2WinBUGS and BRugs were all mentioned, suggesting that the "R and some sort of BUGS" combination is quite popular. See Table 4 for details.

Table 4: Statistical software used

Package	No. of times package mentioned
R	37
WinBUGS	28
Matlab/Mathemetica	8
OpenBUGS	6
FirstBayes	5
R2WinBUGs	5
BRugs	3
JAGS	2
Other	4
No response	7

Electronic and Online Resources

Just over half (55%) of the courses provided electronic resources online and about three-quarters (77%) had a web page link that provided at least some description of the course.

Student Feedback

Respondents were asked to summarise any student feedback relating to the course. As might be expected, responses were many and varied. In general, students seemed to appreciate Bayesian methodology and enjoyed their course of study. Several respondents noted:

1. A desire for more aspects of computation such as use of WinBUGS to be included;
2. Students with applied backgrounds sometimes found the theoretical aspects of a course difficult;
3. Classes with a mixture of student backgrounds were difficult to cater for.

All of these comments are familiar to instructors of Statistics courses in general, not just Bayesian ones. Finally, it should be noted that the information gathered here is: a) highly subjective and may reflect on the respondent's status as an instructor; b) subject to recall and summary; and c) relates to a series of different courses, not just one.

Acknowledgements

I would like to thank all the respondents who took the time and trouble to fill out the survey form. Many replied with a large amount of information and the passion they had for their teaching and subject matter was clearly evident. On the basis of this rather unscientific sample I am reasonably confident that the future of Bayesian teaching at the graduate level is in safe hands!

References

[1] Albert, J. (2007), *Bayesian Computation with R*, New York: Springer.

[2] Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, New York: Springer.

[3] Bernardo, J. M. and Smith, A. F. M. (2000), Bayesian Theory, New York: Wiley.

[4] Carlin, B. P. and Louis, T. A. (2000), Bayes and Empirical Bayes Methods for Data Analysis (2nd ed), Boca Raton: Chapman and Hall.

[5] Congdon, P. (2003), *Applied Bayesian Modelling*, New York: Wiley.

[6] Congdon, P. (2001), *Bayesian Statistical Modelling*, New York: Wiley.

[7] Gamerman, D. and Lopes, H.F. (2006), *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Boca Raton: Chapman and Hall.

[8] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2004), *Bayesian Data Analysis (2nd ed)*, Boca Raton: Chapman and Hall.

[9] Lee, P. M. (2004), *Bayesian Statistics: An Introduction (3rd ed)*, London: Hodder

[10] Robert, C. P. (2007), *The Bayesian Choice (2nd ed)*, New York: Springer.

[11] Spiegelhalter, D.J., Abrams, K.R. and Myles, J.P. (2004), *Bayesian Approaches to Clinical Trials and Health-care Evaluation*, New York; Wiley.