XIV - Congresso Brasileiro de Automática
Natal - RN, 2 a 5 de Setembro de 2002

IDENTIFYING MULTIPLE CHANGE POINTS IN POISSON PROCESSES: A BAYESIAN APPROACH

Rosangela H. Loschi

Departamento de Estatística,
Universidade Federal de Minas Gerais,
31270-901 - Belo Horizonte - MG, Brazil
E-mail: loschi@est.ufmg.br

Frederico R. B. Cruz[a]

Department of Mechanical & Industrial Engineering,
University of Massachusetts, Amherst MA 01003, USA.
E-mail: fcruz@ecs.umass.edu

Abstract— Many subject areas, including disease and criminality mapping, medical diagnosis, industrial control, and finance share the interest in multiple change point identification problems. Algorithms based on the product partition model (PPM) are proposed to solve this important problem applied to time series of Poisson data. In order to attack the PPM, a Gibbs sampling scheme, simple and easy to implement, is derived. The algorithms are applied to the analysis of time series and the results show that the method is quite effective and makes it possible useful inferences.

Key Words— Statistics; time-series analysis; algorithms; Monte Carlo method; control oriented models.

1 Introduction

The identification of multiple change points is a problem encountered in many subject areas, including disease and criminality mapping, medical diagnosis, industrial control, and finance. Given a time series, as the one seen in Figure 1, the problem is to know whether or not change points occurred in the mean. Certainly, this is not a brand new problem and some possible tools have already been considered to tackle it, either Bayesian (Barry and Hartigan, 1993; Loschi et al., 1999; Loschi and Cruz, 2002) and non-Bayesian approaches (Hawkins, 2001; Stauffer, 2001). In particular, this paper is concerned about a Bayesian approach to the multiple change point identification problem in time series of Poisson data, more specifically, the well-known Product Partition Model (PPM).

The PPM was introduced by Hartigan (1990), as a generalization of several models (Smith, 1975; Menzefricke, 1981; Hsu, 1982). One of the advantages of using the PPM is that the number of change points in the series is a random variable and not a fix number, as considered in threshold models (Chen and Lee, 1995). Later, the PPM was considered for the identification of multiple change points in normal means (Barry and Hartigan, 1992; Barry and Hartigan, 1993; Crowley, 1997). As a direct result of previous studies (Loschi, 1998), Loschi et al. (1999) extended Barry and Hartigan's (1993) ideas, to make it possible the identification of multiple change points in the means and variances of normal data. Loschi et al. (1999) obtained the product estimates by means of a recursive algorithm by Yao (1984) and proposed a Gibbs sampling scheme to obtain the posterior distributions for the number of change points as well as for the instants when the changes have occurred.

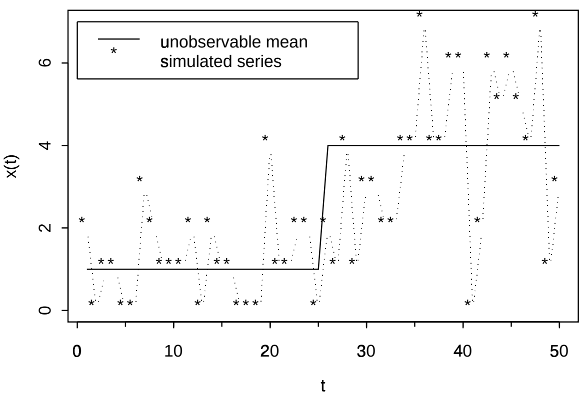

Figure 1. A Change-point Series

More recently, Loschi, Cruz, Iglesias and Arellano-Valle (2002) extended the PPM even further to include a prior specification for the probability p of having a change instead of having it fix, and also proposed a Gibbs sampling scheme to obtain the posterior relevances involved in the computation of the product estimates. As another contribution on the subject, Loschi and

[a]On sabbatical leave from the *Departamento de Estatística, Universidade Federal de Minas Gerais*, 31270-901 – Belo Horizonte – MG, Brazil. E-mail: frbc@est.ufmg.br.

Cruz (2002) asserted the adequacy of some prior distribution for the PPM applied to normal data and provided a comprehensive Monte Carlo simulation study to show the efficacy of different implementations for the PPM (Loschi, Cruz, Iglesias, Arellano-Valle and MacGregor Smith, 2002).

The aim of this paper is to extend the application of the PPM to the identification of multiple change points in the mean θ of Poisson data, assuming a gamma prior distribution to the parameter θ and a beta prior distribution to the probability p of having a change.

The paper is organized as follows. Section 2 reviews the parametric approach to the PPM, presents inferential solutions to identify change points for Poisson random variables and details a Gibbs sampling scheme to implement the PPM. In Section 3, some computational results are presented and discussed. Section 4 closes the paper with final remarks and future topics for investigation.

2 Product Partition Model

In the parametric approach to the PPM (Barry and Hartigan, 1993), it is considered that the sequence of random variables $X_1, \ldots, X_n$, conditionally in $\theta_1, \ldots, \theta_n$, has conditional marginal densities $f_1(X_1|\theta_1), \ldots, f_n(X_n|\theta_n)$. It is assumed that given a partition $\rho = \{i_0, \ldots, i_b\}$ of the set $I \cup \{0\}$, for $I = \{1, \ldots, n\}$ and $b \in I$, such that $0 = i_0 < i_1 < \cdots < i_b = n$, one has that $\theta_i = \theta_{[i_{r-1} i_r]}$ for every $i_{r-1} < i \leq i_r$, for $r = 1, \ldots, b$, and that $\theta_{[i_0 i_1]}, \ldots, \theta_{[i_{b-1} i_b]}$ are independent, with $\theta_{[ij]}$ having (block) prior density $\pi_{[ij]}(\theta)$, where $\theta \in \Theta_{[ij]}$ and $\Theta_{[ij]}$ is the parameter space corresponding to the common parameter, say, $\theta_{[ij]} = \theta_{i+1} = \cdots = \theta_j$, which indexes the conditional density of $\mathbf{X}_{[ij]} = (X_{i+1}, \cdots, X_j)'$. Denote by $c_{[ij]}$, $i, j \in I \cup \{0\}$, $i < j$, the prior cohesion associated with the block $[ij]$ which is interpreted as the transition probability of having a change in j, given that a change takes place in i.

In this case, two observations X_i and X_j, $i \neq j$, are considered in the same block if they are identically distributed. Thus, $(X_1, \ldots, X_n, \rho)$ follows the PPM if:

i) the prior distribution of ρ is the following product distribution:

$$P(\rho = \{i_0, \ldots, i_b\}) = \frac{\prod_{j=1}^{b} c_{[i_{j-1} i_j]}}{\sum_{\mathcal{C}} \prod_{j=1}^{b} c_{[i_{j-1} i_j]}}, \quad (1)$$

in which $\mathcal{C}$ is the set of all possible partitions of the set I into b contiguous blocks with endpoints $i_1, \ldots, i_b$, satisfying the condition $0 = i_0 < i_1 < \cdots < i_b = n$, for all $b \in I$;

ii) conditionally on $\rho = \{i_0, \ldots, i_b\}$, the sequence $X_1, \ldots, X_n$ has the joint density given by:

$$f(X_1, \ldots, X_n|\rho) = \prod_{j=1}^{b} f_{[i_{j-1} i_j]}(\mathbf{X}_{[i_{j-1} i_j]}), \quad (2)$$

in which $f_{[ij]}(\mathbf{X}_{[ij]})$ is the joint density of the random vector $\mathbf{X}_{[ij]} = (X_{i+1}, \ldots, X_j)'$, given by:

$$f_{[ij]}(\mathbf{X}_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta. \quad (3)$$

Assuming the PPM, the posterior expectation (or the product estimate) of θ_k is given by:

$$E(\theta_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1} \sum_{j=k}^{n} \left(r^*_{[ij]} E(\theta_k|\mathbf{X}_{[ij]}) \right), \quad (4)$$

for $k = 1, \ldots, n$, in which the posterior relevance for the block $[ij]$ is given by:

$$r^*_{[ij]} = \frac{\lambda_{[0i]} c^*_{[ij]} \lambda_{[jn]}}{\lambda_{[0n]}}, \quad (5)$$

in which $c^*_{[ij]} = c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]})$ and $\lambda_{[ij]} = \sum \Pi_{k=1}^{b} c^*_{[i_{k-1} i_k]}$, and the summation is over all partitions of $\{i+1, \ldots, j\}$ in b blocks with endpoints $i_0, i_1, \ldots, i_b$, satisfying the condition $i = i_0 < i_1 < \cdots < i_b = j$.

Other parameter considered is the number of blocks B (or the number of change points, $B - 1$) in ρ. If the PPM is assumed, the posterior distribution of B is given by:

$$P(B = b|X_1, \ldots, X_n) \propto \sum_{\mathcal{C}} \Pi_{j=1}^{b} c^*_{[i_{j-1} i_j]}. \quad (6)$$

The posterior distribution of ρ has the same form of the prior distribution given in (1), considering the posterior cohesions $c^*_{[ij]}$.

2.1 Poisson Case

For the Poisson case, it is assumed that, given $\theta_1, \ldots, \theta_n$, $X_1, \ldots, X_n$ are such that $X_k|\theta_k, \sim \mathcal{P}(\theta_k)$, for $k = 1, \ldots, n$, and that they are independent. It is also assumed that the common parameter $\theta_{[ij]}$, related to the block $[ij]$, has the conjugate gamma prior distribution denoted by:

$$\theta_{[ij]} \sim \mathcal{G}\left(\tau_{1[ij]} + 1, \tau_{0[ij]}\right),$$

where $\tau_{0[ij]} > 0$ and $\tau_{1[ij]} > -1$, and whose density function is given by:

$$f(\theta_{[ij]}|\tau_{0[ij]},\tau_{1[ij]}) = \frac{\tau_{0[ij]}^{\tau_{1[ij]}+1}}{\Gamma(\tau_{1[ij]}+1)} \times \theta_{[ij]}^{\tau_{1[ij]}} \times \exp(-\tau_{0[ij]}\theta_{[ij]}).$$

Consequently, the random vector $\mathbf{X}_{[ij]}$ follows a distribution with density function given by

$$f(\mathbf{X}_{[ij]}) = \Big(\prod_{k=i+1}^{j}\frac{1}{X_k!}\Big) \times \frac{\Gamma(\tau^*_{1[ij]})}{\Gamma(\tau_{1[ij]}+1)} \times \Big(\frac{\tau_{0[ij]}}{\tau^*_{0[ij]}}\Big)^{\tau_{1[ij]}+1} \times \Big(\frac{1}{\tau^*_{0[ij]}}\Big)^{\sum_{k=i+1}^{j}X_k}, \quad (7)$$

in which

$$\begin{cases} \tau^*_{0[ij]} = \tau_{0[ij]} + j - i, \\ \tau^*_{1[ij]} = \tau_{1[ij]} + \sum_{k=i+1}^{j} X_k + 1, \end{cases}$$

for all $i = 0,\ldots,n-1$, and $j = i+1,\ldots,n$.

Given $X_{[ij]}$, the distribution of $\theta_{[ij]}$ is the gamma distribution with parameters $\tau^*_{0[ij]}$ and $\tau^*_{1[ij]}$, that is

$$\theta_{[ij]}|X_{[ij]} \sim \mathcal{G}\Big(\tau^*_{1[ij]},\tau^*_{0[ij]}\Big).$$

Consequently, the blocks estimates are given by

$$\hat{\theta}_{[ij]} = E(\theta_{[ij]}|X_{[ij]}) = \frac{\tau^*_{1[ij]}}{\tau^*_{0[ij]}}, \quad (8)$$

and, from equations (4) and (8), it follows that the product estimates are given by

$$\hat{\theta}_k = E(\theta_k|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]}\hat{\theta}_{[ij]}, \quad (9)$$

for $k = 1,\ldots,n$. The posterior relevances $r^*_{[ij]}$ can be obtained from Eq. (5), taking into consideration the density given in Eq. (7).

2.2 *A Gibbs Sampling Scheme Applied to the PPM*

An extraordinary array of problems in Bayesian inference have been solved by Markov chain Monte Carlo (MCMC) methods since the seminal paper by Gelfand and Smith (1990) illustrated how easily a variety of intractable problems could be approximately solved. This ease of use led to an explosion of research and complex Bayesian models without analytical solution are know tractable by MCMC methods. Recent research results and overviews of the research in this area includes the papers by Besag et al. (1995), Robert (1995), and MacEachern and Peruggia (2000), to cite few. In particular, the purpose is to use Gibbs sampling (Geman and Geman, 1984) as a posterior distribution generation scheme.

In order to estimate the posterior distributions of B and ρ, and the posterior relevance of each block $[ij]$, the method proposed by Loschi, Cruz, Iglesias and Arellano-Valle (2002) is described. Let us assume the auxiliary random quantity U_i which reflects whether or not a change point occurred at the time i, that is

$$U_i = \begin{cases} 0, & \text{if } \theta_i \neq \theta_{i+1}, \\ 1, & \text{if } \theta_i = \theta_{i+1}, \end{cases}$$

for $i = 1,\ldots,n-1$.

At the kth step, the vector $\mathbf{U}^k = (U_1^k,\ldots,U_{n-1}^k)$ is generated by using the Gibbs sampling as follows. Considering a beta prior distribution for the probability p of change, $p \sim \mathcal{B}(\alpha,\beta)$, it is sufficient to consider the ratio given by the following expression, in order to generate the vectors $\mathbf{U}^k$'s:

$$R_r = \frac{f_{[xy]}(X_{[xy]})}{f_{[xr]}(X_{[xr]})f_{[ry]}(X_{[ry]})} \times \frac{\Gamma(n+\beta-b+1)\Gamma(b+\alpha-2)}{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}, \quad (10)$$

where

$$x = \begin{cases} \boxed{\begin{array}{l} \max i \\ \textbf{s.t.:}\ \ 0 < i < r, \\ \qquad U_i^k = 0, \end{array}} & \begin{array}{l}\text{if there is } U_i^k = 0, \\ \text{for some} \\ i \in \{1,\ldots,r-1\},\end{array} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$y = \begin{cases} \boxed{\begin{array}{l} \min i \\ \textbf{s.t.:}\ \ r < i < n, \\ \qquad U_i^{k-1} = 0, \end{array}} & \begin{array}{l}\text{if there is a } U_i^{k-1} = 0, \\ \text{for some} \\ i \in \{r+1,\ldots,n-1\},\end{array} \\ n, & \text{otherwise,} \end{cases}$$

since the rth element at the kth step, U_r^k, is generated from the conditional distribution

$$U_r^k|U_1^k,\ldots,U_{r-1}^k,U_{r+1}^{k-1},\ldots,U_{n-1}^{k-1};\ X_1,\ldots,X_n;p,$$

for $r = 1,\ldots,n-1$, starting from an initial vector $\mathbf{U}^0 = (U_1^0,\ldots,U_{n-1}^0)$.

Notice that, in the Poisson case, $f_{[ij]}(X_{[ij]})$ is the distribution given in Eq. (7). Consequently, the criterion for the values U_r^k becomes

$$U_r^k = \begin{cases} 1, & \text{if } R_r \geq (1-u)/u, \\ 0, & \text{otherwise,} \end{cases}$$

in which $r = 1,\ldots,n-1$, and $u \sim \mathcal{U}(0,1)$.

Notice that the posterior relevance of the block $[ij]$, for $i < j$, used in Eq. (4) to estimate θ_k can be obtained by considering the proportion of samples that presents $U_i^k = 0$, $U_{i+1}^k = \ldots = U_{j-1}^k = 1$, and $U_j^k = 0$. The random quantity ρ is perfectly identified by considering a vector of these random quantities. Consequently, one can estimate the posterior probability for each particular partition in b contiguous blocks, $\rho = \{i_0, i_1, \ldots, i_b\}$. Also notice that it is possible to use the above procedure to estimate the posterior distribution of B (or the posterior distribution of the number of change points, $B-1$) by considering that

$$B^k = 1 + \sum_{i=1}^{n-1} (1 - U_i^k).$$

Figure 2 shows the complete algorithm in pseudocode.

algorithm
  read *all prior specifications*
  read $X_1, \ldots, X_n$
  for $k = 1$ **to** SAMPLES **do**
    generate $\mathbf{U}^k$
  end for
  for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
    $r^*_{[ij]} \leftarrow$ proportion of samples such that $U_i^k = 0, U_{i+1}^k = \cdots = U_{j-1}^k = 1, U_j^k = 0$
  end for
  for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
    $\tau^*_{0[ij]} \leftarrow \tau_{0[ij]} + j - i$
    $\tau^*_{1[ij]} \leftarrow \tau_{1[ij]} + \sum_{k=i+1}^{j} X_k + 1$
  end for
  for $k = 1$ **to** n **do**
    $E(\theta_k | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \hat{\theta}_{[ij]}$
  end for
  write $E(\theta_1), \ldots, E(\theta_n)$
  write B^k
end algorithm

Figure 2. PPM Gibbs Sampling Algorithm

3 Computational Experiments

Because of its computationally intensive nature, the algorithm presented in Figure 2 was coded in C++. All tests were performed on a PC, Pentium processor 400 MHz, 256 MB RAM, taking less than one minute of CPU time. In order to estimate the posterior relevances $r^*_{[ij]}$ and the posterior distribution of B (or the number of change points, $B-1$), 4,600 samples of 0–1 values were generated with the dimension of the time series, starting from a sequence of zeros. The initial 100 iterations were discharged for *burn-in* and a lag of one was selected. Discussion about the number of iterations to be discharged, as well as the lag to be taken, can be easily found in the literature (see, for instance, Gamerman, 1997).

3.1 Prior Specifications

In order to verify the accuracy of the approach, the computational experiments were conducted considering the artificial (simulated) time series shown in Figure 1. The time-series observations are assumed to be conditionally independent and distributed according to the Poisson distribution $\mathcal{P}(\theta_{[ij]})$. We considered the natural conjugate prior distribution for the parameters $\theta_{[ij]}$, which is in this case a gamma distribution. These assumptions are not too restrictive, since the Poisson distribution is appropriated for many practical applications. Additionally, the gamma distribution is rich enough to describe the uncertainty about the parameters under many practical circumstances.

For the present case, it seems reasonable to consider that

$$\begin{cases} \tau_{0[ij]} = 1.0, \\ \tau_{1[ij]} = 0.0. \end{cases} \tag{11}$$

About this subjective choice, the gamma distribution considered, $\mathcal{G}(1, 1)$, plotted in Figure 3, concentrates its mass in a low value and it is also as flat as our uncertainty about this parameter. Other similar settings for $\tau_{0[ij]}$ and $\tau_{1[ij]}$ where considered but the results (not shown) do not differ quite much.

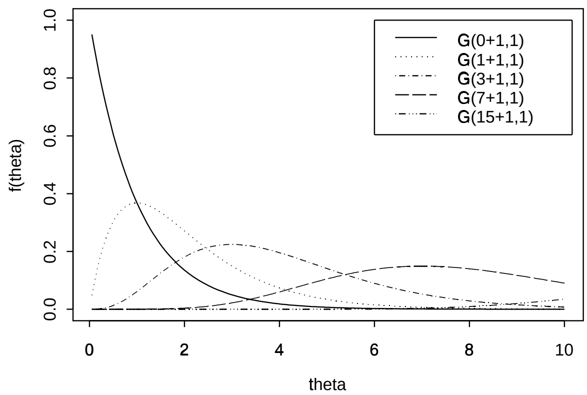

Figure 3. Probability Density $\mathcal{G}(\tau_1, \tau_0)$

The truncated geometric distribution with parameter p, $p \in (0, 1)$, is considered as prior cohesions since it is assumed that past change points are non-informative about future change points. Thus, one last decision that has to be made concerns the probability p of having a changing point. It is assumed that $p \sim \mathcal{B}(2, 8)$, plotted in Figure 4, since a small number of changes is expected in the simulated series.

3.2 Numerical Results

For the sake of conciseness, only results for the artificial (simulated) time series presented in Figure 1 are shown. Another simulations were carried

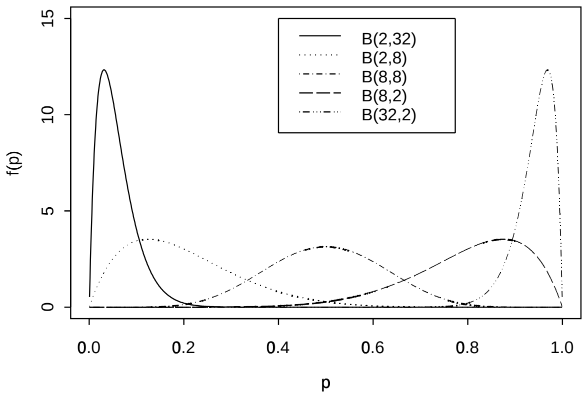

Figure 4. Probability Density $\mathcal{B}(\alpha,\beta)$

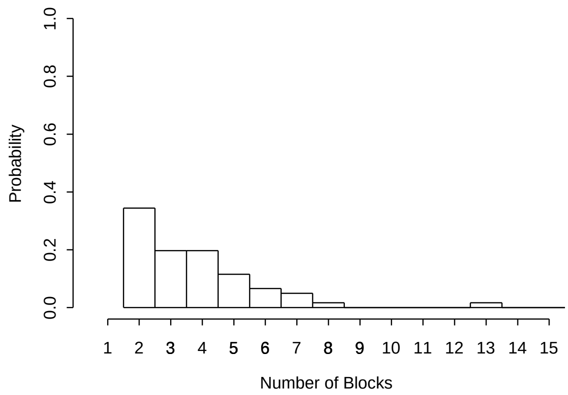

Figure 6. Posterior Distribution of B

with similar simulated series and the results (not shown) do not differ significantly. The main advantage of such an analysis is that one can control for errors in the method since the actual (unobservable) mean is known. In Figure 5, the posterior estimates for the expected θ are presented. Note the close agreement among unobservable mean and posterior estimates. The method certainly is not able to immediately fell the border between observations 25th and 26th but after only 10 observations the estimates picks it up.

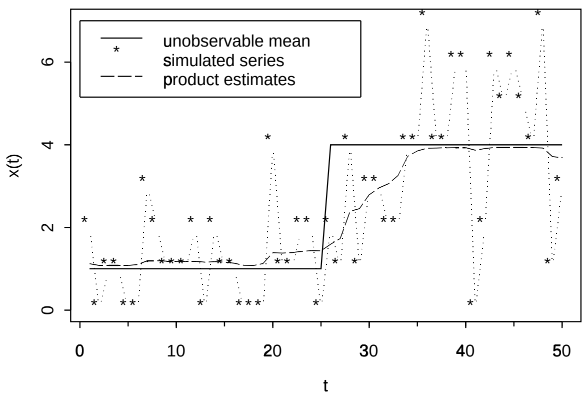

Figure 5. Posterior Estimates

However, much more information are available through the method and also possible it is to study the posterior distribution of the number of blocks. The results are shown in Figure 6. Here, we can see that certainly the most probable number of blocks is just 2, with the probability of 32%. As expected, the posterior distribution of the number of blocks concentrates its mass on small values.

Another important information that we can easily drawn is the most probable partition which is shown in Figure 7. The method identify the 27th observation as a change point, just one observation away from the 'real' point, the 26th.

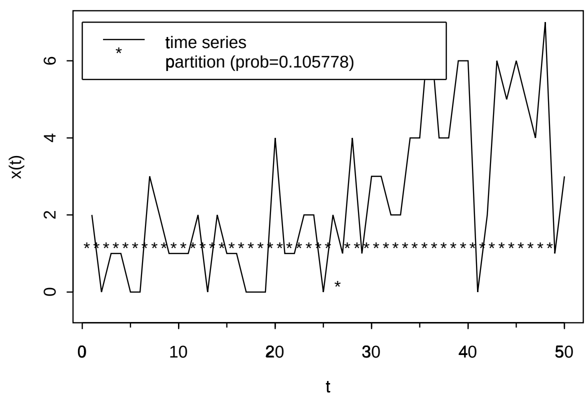

Figure 7. Most Probable Partition

4 Conclusions and Final Remarks

The problem of identifying multiple change point in Poisson data was considered under the product partition model (PPM), a Bayesian framework. The PPM was described and its importance to change point problems was stressed, particularly to analyze time series. A Gibbs sampling scheme was proposed to implement the PPM and to avoid its computational difficulties. Coded and tested, the algorithms proposed proved to be an efficient and useful tool for time series analysis of Poisson data. In the simulated time series in which it was applied, the method worked quite satisfactory.

Some open research questions remain. How long would the treatable series be? How well does the methodology fit for other subject areas? These and other similar questions are interesting and relevant topics for future research in this area.

Acknowledgments

Rosangela H. Loschi acknowledges the CAPES, PRPq-UFMG (Fundo-2002) and Fundación Andes (Chile), for a partial allowance to her research. Frederico R. B. Cruz has been partially funded by the CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) of the Ministry for Science and Technology of Brazil (grants 301809/96-8 and 201046/94-6), FAPEMIG (grants CEX-289/98 and CEX-855/98), and PRPq-UFMG (grant 4081-UFMG/RTR/FUNDO/PRPq/99).

References

Barry, D. and Hartigan, J. A. (1992). Product partition models for change point problems, *The Annals of Statistics* **20**(1): 260–279.

Barry, D. and Hartigan, J. A. (1993). A Bayesian analysis for change point problem, *Journal of the American Statistical Association* **88**(421): 309–319.

Besag, J., P. Green, D. H. and Mengersen, K. (1995). Bayesian computation and stochastic systems, *Statistical Science* **10**: 3–66.

Chen, C. W. S. and Lee, J. C. (1995). Bayesian inference of threshold autoregressive models, *Journal of Time Series Analysis* **16**(5): 483–492.

Crowley, E. M. (1997). Product partition models for normal means, *Journal of the American Statistical Association* **92**(437): 192–198.

Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman & Hall, London, UK.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**: 398–409.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution and Bayesian restoration of images, *IEEE Transactions on Pattern Analysis & Machine Intelligence* **6**: 721–741.

Hartigan, J. A. (1990). Partition models, *Communication in Statistics - Theory & Method* **19**(8): 2745–2756.

Hawkins, D. M. (2001). Fitting multiple change-point models to data, *Computational Statistics & Data Analysis* **1**: 323–341.

Hsu, D. A. (1982). A Bayesian robust detection of shift in the risk structure of stock market returns, *Journal of the American Statistical Association* **77**(2): 29–39.

Loschi, R. H. (1998). *Imprevistos e suas Conseqüências (Unpredictabilities and Consequences)*, PhD thesis, Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil (in Portuguese).

Loschi, R. H. and Cruz, F. R. B. (2002). An analysis of the influence of some prior specifications in the identification of change points via product partition model, *Computational Statistics & Data Analysis* **39**(4): 477–501.
*ftp://est.ufmg.br/pub/fcruz/pub/LoCr02.pdf

Loschi, R. H., Cruz, F. R. B., Iglesias, P. L. and Arellano-Valle, R. B. (2002). A Gibbs sampling scheme to product partition model: An application to change point problems, *Computers & Operations Research* (in press).
*ftp://est.ufmg.br/pub/fcruz/pub/CAOR_886.pdf

Loschi, R. H., Cruz, F. R. B., Iglesias, P. L., Arellano-Valle, R. B. and MacGregor Smith, J. (2002). Monte Carlo simulation for product partition model: Atypical data and sample size cases. Submitted for publication.
*ftp://est.ufmg.br/pub/fcruz/pub/mc.pdf

Loschi, R. H., Iglesias, P. L. and Arellano-Valle, R. B. (1999). Bayesian detection of change points in the Chilean stock market, *Proceedings of the Section on Bayesian Statistical Science*, Annual Meeting of American Statistical Association, Baltimore, MD, USA, pp. 160–165.

MacEachern, S. N. and Peruggia, M. (2000). Subsampling the Gibbs sampler: Variance reduction, *Statistics & Probability Letters* **47**: 91–98.

Menzefricke, U. (1981). A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point, *Applied Statistics* **30**(2): 141–146.

Robert, C. (1995). Convergence control methods for Markov chain Monte Carlo algorithms, *Statistical Science* **10**: 231–253.

Smith, A. F. M. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables, *Biometrika* **62**(2): 407–416.

Stauffer, D. (2001). Percolation models of financial market dynamics, *Advances in Complex Systems* **4**(1): 19–27.

Yao, Y. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches, *The Annal of Statistics* **12**(4): 1434–1447.