RESEARCH ARTICLE   OPEN ACCESS

Non-life claims reserves using Dirichlet random environment

Hafedh Faires

1Department of Mathematics and Statistics Faculty of Sciences Al-Imam Muhammad Ibn Saud Islamic University P.O.Box 90950 Riyadh 11623 KSA

The purpose of this paper is to propose a stochastic extension of the Chain-Ladder model in a Dirichlet random environment to calculate the provesions for disaster payement. We study Dirichlet processes centered around the distribution of continuous-time stochastic processes such as a Brownian motion or a continuous time Markov chain. We then consider the problem of parameter estimation for a Markov-switched geometric Brownian motion (GBM) model. We assume that the prior distribution of the unobserved Markov chain driving by the drift and volatility parameters of the GBM is a Dirichlet process. We propose an estimation method based on Gibbs sampling.
Keywords: Reserving claims EDS Dirichlet prior Markov chain Gibbs sampling.

I.   Introduction

The calculation of the provisions for disaster payments is intended to allow the integral payment of the commitments to the policy-holders and the recipients of the contract. The provisions measure the commitments that the insurer still has to honor. Nevertheless, this countable concept requires a subjacent probabilistic model, since it allows one to define the ultimate claim, taking into account the disasters not yet declared but which have occurred, the disasters not sufficiently funded. Reserves are given by evaluating the provisions for each contract, IBNR (sinister not yet declared) and IBNER (sinister not sufficiently funded). Traditional methods of provisioning (by triangulation) rest on the assumption that the data are homogeneous and in sufficient quantity to ensure a certain stability and a certain credibility. The purpose of this paper is to propose a stochastic extension of the Chain-Ladder model in a Dirichlet random environment, which seems to us to be closer to reality than the existing methods of Schnieper,R [32]; Mack,T [24]; H.Liu and R.Verrall [31]; Verrall, R. J. and England, P. D.[37]. We study the estimation of claim reserves in non-life reinsurance using Drichlet processus.

In a famous paper, Ferguson [15] introduced the Dirichlet process (DP) with the parameter α, a non-negative finite measure on a measurable space V, as a random probability measure P on V whose finite-dimensional distributions are Dirichlet distributions, P being centered around the probability measure $\overline{\alpha} = \alpha/\alpha(V)$ where the positive number $\alpha(V)$ expresses a confidence degree on $\overline{\alpha}$.

There is a large literature on many interesting theoretical properties of the DP, see e.g. Ferguson [15], [16], Blackwell and MacQueen [2], Antoniak [1], Kingman [21], Sethuraman [33], Pitman and Yor [27], etc. It is also well-known that the Dirichlet distribution and the DP are extensively used in various applications, in the setting of nonparametric Bayesian statistics, see e.g. Ishwaran and Zarepour [19], Ishwaran and James [18], [20], Blei, Ng and Jordan [4], Dahl [7] etc. In all these applications, the authors deal with a probability measure $\overline{\alpha}$ defined on a finite-dimensional space and it seems that the idea of using an infinite-dimensional space, such as a functional space, first appeared in MacEachern (2000) [25] and Faires [14]. In the present paper we develop this idea in two directions. In the section below, we take $\overline{\alpha} = H$, the distribution of a continuous time Markov chain on a finite set of states. We then propose a new hierarchical model: a stochastic differential equation (SDE) in random environment with a Dirichlet prior on the path space of the chain, the states of the chain representing the environment. Our estimation method, presented in Section 6, is a generalization to continuous time of the classification method of Ishwaran - Zarepour [19] and Ishwaran - James [18] for estimating normal mixtures with a Dirichlet prior, the new problem appearing here being that we observe just one single path with dependent data. Actually, in our approach, the states of the chain are considered as classes and the classification procedure is performed along this single path. This method, which hinges on Gibbs sampling, requires some posterior distribution computations that are presented below. In Section 7, the algorithm is applied to a real dataset coming from the B.E.S.T.R.É company.

Models in which parameters move between a fixed number of regimes with switching controlled by an unobserved stochastic process, are very popular in a great variety of domains (Finance, Biology, Meteorology, Networks, etc.). This is notably due to the fact that this additional flexibility allows the model to account for random regime changes in the environment. In this paper we consider the estimation problem for a model described by a stochastic differential equation (SDE) with Markov regime-switching (MRS), i.e., with parameters controlled by a finite state continuous-time Markov chain (CTMC). Such a model was used, for example, in Deshpande and Ghosh (2008) to price options in a regime switching market. In such a setting, the parameter estimation problem poses a real challenge, mainly due to the fact that the paths of the CTMC are unobserved. A standard approach consists in using the celebrated EM algorithm (Dempster, Laird, Rubin, 1977) as proposed for example in Hamilton (1990). Elliott, Malcolm and Tsoi (2003) study this problem using a filtering approach.

In the present paper, our estimation is Bayesian, the aim being to find a pair (parameters, CTMC path) with likelihood as large as possible. approach. We refer the reader to Schnatter (2006) for a wider discussion on Markov switching models and the comparative advantages of the Bayesian approach. Standard priors are placed on the parameters space but, as the CTMC paths are unobserved, a large number of paths are drawn from a Dirichlet process placed as a prior on the path space of the CTMC. The complete model then appears as a Hierarchical Dirichlet Model (HDM), as in Ishwaran, James and Sun (2000) and Ishwaran and James (2002). The estimation procedure for the model considered in this paper requires some rather nontrivial computations of posterior distributions due to the temporal level induced by the specific SDE and the CTMC. Using the well-known stick-breaking approximation, each set of iterations selects the pair with largest likelihood and then the Dirichlet process is updated in order to look for other paths which can further improve the likelihood.

The considered SDE is that of a geometric Brownian motion, a popular model for asset prices in mathematical finance which depends on two parameters, namely the drift and volatility. It is considered in the extended MRS setting so that the CTMC transitions correspond to regime changes in the market.

The rest of the paper is structured as follows. In Section 2 we present the stock price SDE with MRS and the complete HDM. Section 3 is devoted to the posterior computations. The estimation algorithm is described in Section 4. Numerical results are

presented in Section 5 for one simulated data and on a data from the Indian market. We conclude with a summary in the last Section.

II. Bayesian Regime Switching Model

Our model is specified in a mathematical finance setting but it can be extended in a similar way to various contexts.

The following notations will be adopted:

n will denote the number of observed data and also the length of an observed path,

M will denote the number of states of the Markov chain,

The state space of the Markov chain will be denoted by $S = \{i : 1 \leq i \leq M\}$,

N will denote the number of simulated paths, Given a path of the CTMC, m will denote the number of distinct states in that path.

2.1 Hypotheses and description of the model

We suppose that the available data have a triangular form indexed by the year of accident, i, and development time, t. Given a triangle, on T years, the goal is to consider models using a minimum of parameters, in order to envisage the best possible amounts of payments of future disasters. We note the evolution of the amounts of payments of the cumulated real disasters obtained by $\{C_t^i; \quad i = 1,2,...., T; \quad t \in [1, T - i + 1]\}$.

C_t^i indicate the evolution of the cumulated real disasters indexed by the year of accident, i, and the time of development, t. We suppose that the increase in the disasters obtained $(C_t^i - C_{t-1}^i)$ is the sum of the disasters not sufficiently funded $(-D_t^i)$, and of the not declared yet claims (N_t^i). We write the following relations between C, N and D.

$$C_t^i = C_{t-1}^i - D_t^i + N_t^i, \quad 1 \leq i \leq n, \quad \forall t \in [1, +\infty[.$$

(1)

We indicate by $\mathsf{H}_s^i = \{D_t^i, N_t^i \mid i + t \leq s + 1\}$ all the variables in the triangles D and N observed until the moment s.

To simulate the future claims, it is supposed that the not sufficiently funded claims $(D_t^i)_{t \in [1, +\infty[}$, the stochastic differential equation of diffusion and the not yet incurred claims $(N_t^i)_{t \in [1, +\infty[}$ are governed by the stochastic differential equation of Black and

Scholes with jump. This assumption on the probability density function of N_t^i ensures positivity

and for D_t^i ensures the membership R, contrary to what is proposed by H.Liu and R.Verrall $HLRV$.

Conditionally with H_T^i, we simulate the distribution of N_t^i as solution of the following stochastic differential equation:

$$dN_t^i = N_{t_-}^i(\alpha^i(t,X_t,N_{t_-}^i)dt + \tau^i(t,X_t,N_{t_-}^i)dB_t + \kappa^i d\mathsf{P}_t), \qquad (2)$$

N^i_T-i+1=x,

where B_t is a standard Brownian motion on R, P_t is the Poisson process of intensity λ, X_t is a Markov process at continuous time, $\alpha^i : \mathsf{R}^+ \times \mathsf{N} \times \mathsf{R}^+ \to \mathsf{R}$, $\tau^i : \mathsf{R}^+ \times \mathsf{N} \times \mathsf{R}^+ \to \mathsf{R}$ verifying

$$\int |\alpha^i(t,X_t,N_{t_-}^i)| \, dt < \infty \quad and \quad \int |\tau^i(t,X_t,N_{t_-}^i)|^2 \, dt < \infty$$

and κ^i is a positive constant.

Conditionally with H_T^i, we suppose that the evolution of D_t^i is governed by the following diffusion:

$$dD_t^i = \rho^i(t,X_t,D_{t_-}^i,N_{t_-}^i)dt + \sigma^i(t,X_t,D_{t_-}^i,N_{t_-}^i)dB_t, \quad (3)$$

$D_{T-i+1}^i = y,$

where B_t is a standard Brownian motion on R, X_t is a Markov process at continuous time, $\mu^i : \mathsf{R}^+ \times \mathsf{N} \times \mathsf{R} \times \mathsf{R}^+ \to \mathsf{R}$ and $\sigma^i : \mathsf{R}^+ \times \mathsf{N} \times \mathsf{R} \times \mathsf{R}^+ \to \mathsf{R}$ such that $\int |\rho^i(t,X_t,D_{t_-}^i,N_{t_-}^i)| \, dt < \infty$ and $\int |\sigma^i(t,X_t,D_{t_-}^i,N_{t_-}^i)| \, dB_t < \infty$. The process (X_t) is assumed to be a continuous time Markov process taking values in the set $S = \{i : 1 \le i \le M\}$. The transition probabilities of this chain are denoted by p_{ij}, $i,j \in S$ and the transition rate matrix is $Q_0 = (q_{ij})_{i,j \in S}$ with

$$\xi_i > 0, \qquad q_{ij} = \xi_i p_{ij} \quad if \quad i \neq j, \qquad and$$

$$q_{ii} = -\sum_{j \neq i} q_{ij}, \qquad i,j \in S.$$

For any state $i = 1,2,\dots,M$, consider priors μ_i and σ_i defined as follows

$$\mu_i \,|\, \theta \overset{ind}{:} \mathsf{N}(\theta,\tau^\mu) \qquad (4)$$

$$\theta : \mathsf{N}(0,A), \qquad A > 0, \quad (5)$$

$$\sigma_i \overset{ind}{:} \Gamma(\nu_1,\nu_2) \quad (6)$$

where $\Gamma(\nu_1,\nu_2)$ denotes a Gamma distribution with shape parameter ν_1 and scale parameter ν_2.

Let us define the log-returns, $Y_t = Z_t - Z_{t-1} = \log(N_t/N_{t-1})$ for the equation (13) and $Y_t = D_t - D_{t-1}$ for the second equation (14), $t = 1,2,\dots,n$.

Given a path $X = \{X_s, 0 \le s \le n\}$. Let $T_j(t)$ be the time spent by the path X in state j in the time interval $[t-1,t]$. Define

$$\mu(t) := \sum_{j=1}^{M} \mu(j)T_j(t); \qquad \tau^2(t) := \sum_{j=1}^{M} \tau^2(j)T_j(t); \qquad (7)$$

$$\lambda(t) := \sum_{j=1}^{M} \lambda(j)T_j(t). \quad (8)$$

$$\rho(t) := \sum_{j=1}^{M} \rho(j) T_j(t); \qquad \sigma^2(t) := \sum_{j=1}^{M} \sigma^2(j) T_j(t). \qquad (9)$$

Then, conditional on the path X, the solutions of the equations (13) and (14) are respectively given by :
    • the solution of the first equation (13) is

$$-1cm Y_t = Y_0 exp\{\int_0^t (\tau(s) dB_s + \kappa d\mathsf{P}_s) + \int_0^t (\mu(s) - \frac{\tau^2(s)}{2}) ds\}$$

$$(10)$$

$$2.5cm \times \Pi_{0 < s \le t} (1 + \kappa \Delta \mathsf{P}_s) e^{-\kappa \Delta \mathsf{P}_s}$$

$$t = 1,2,\ldots,n,$$

and,

    • Y_t are i.i.d. $\mathsf{N}(\rho_t, \sigma_t)$, $t = 1,2,\ldots,n$ for the second equation (14).

2.2 Markov Regime Switching Model with Dirichlet prior

To make our model more flexible we complete it now by placing a Dirichlet prior $\mathsf{D}(\alpha\,H)$ with precision parameter $\alpha > 0$ and mean H (see Ferguson, 1973) on the path space of the CTMC (X_t). The probability measue H is the distribution of a CTMC and initially H is defined by taking as initial distribution the uniform distribution on the states, that is
$$\pi = (1/M,\ldots,1/M),$$
as transition matrix P
$$P_{ij} = 1/(M-1), i \ne j$$
and as rates
$$\lambda_i = \lambda > 0, i = 1,\ldots,M.$$
Thus the Markov chain under H will spend an exponentially distributed time with mean $1/\lambda$ in any state i and then jump to state $j \ne i$ with probability $1/(M-1)$.

A selection of a Markov chain path from the above Dirichlet prior can be generated as follows. Generate a large number, say N, of paths $X^{(i)} = \{x_s^i : 0 \le s \le n\}$, $i = 1,2,\ldots,N$, from H. The parameter λ will be chosen to be small in order to get a large variety of paths. Next generate a probability vector $(p_i, i = 1,\ldots,N)$ from a stick-breaking scheme with parameter α (see Sethuramane, 1994). Then draw a path of the Markov chain from the distribution

$$p = \sum_{i=1}^{N} p_i \delta_{X_i}. \quad (11)$$

So we have just replaced the Dirichlet process $\mathsf{D}(\alpha\,H)$ by p, this approximation is justified by Sethuramane result. Finally the prior for α is a Gamma distribution:
$$\alpha : \Gamma(\eta_1, \eta_2). \quad (12)$$
Our model is summarized by the following Dirichlet hierarchical model after choosing π, Q, N and drawing $X_1,\ldots,X_N$ from H :

$$\alpha : \Gamma(\nu_1, \nu_2)$$
$$(p_i) : SB(N, \alpha)$$
$$p = \sum_{i=1}^{N} p_i \delta_{X_i}$$
$$\theta : \mathsf{N}(0, A), \qquad A > 0,$$
$$\mu_i | \theta \overset{ind}{:} \mathsf{N}(\theta, \tau^\mu)$$

$$\sigma_i \overset{ind}{:} \Gamma(\nu_1, \nu_2)$$
$$(X_t) : p$$

$$dN_t^i \mid X = N_{t_-}^i (\alpha^i(t, X_t, N_{t_-}^i) dt + \tau^i(t, X_t, N_{t_-}^i) dB_t + \kappa^i d\mathsf{P}_t), \quad (13)$$

$D_{T-i+1}^i = x,$ and

$$dD_t^i \mid X = \rho^i(t, X_t, D_{t_-}^i, N_{t_-}^i) dt + \sigma^i(t, X_t, D_{t_-}^i, N_{t_-}^i) dB_t, \quad (14)$$

$$D_{T-i+1}^i = y.$$

III. Estimation procedures

Roughly speaking, once π, Q, N chosen, N being large, and paths $X_1, \ldots, X_N$ simulated from H, a first estimation procedure of the above model parameters given the observations is doned by using a blocked Gibbs sampling technique. This technique requires the posterior conditional distribution of each parameter given the other parameters. Drawing an initial value of α and then a value of p, a path X is drawn among $X_1, \ldots, X_N$ w.r.t. p and the parameters updated. This path and the parameters provide a likelihood. The aim the first procedure consists in selecting the pair (path, parameters) with maximal likelihood after a long run.

This first procedure is repeated with the same π, Q, N but with other simulated paths $X_1, \ldots, X_N$ in order to improve the likelihood.

Once a pair (path, parameters) with maximal likelihood is chosen, π and Q are re-estimated from this path in a standard way and the first procedure is repeated.

Let us now proceed to the detailed computations.

We denote by μ and σ, the current values of the vectors $(\mu_1, \mu_2, \ldots, \mu_n)$ and $(\sigma_1, \sigma_2, \ldots, \sigma_n)$, respectively. Let Y be the vector of observed data $(Y_1, \ldots, Y_n)$. Given the current path $X = (x_s, 0 \le s \le n)$ of the Markov chain, let $X^* = (x_1^*, \ldots, x_m^*)$ be the distinct values in X.

3.1 Modifying the observed data set

In order to obtain the conditional distribution of the parameters, we first need to extract the change in the log-returns between the jump times of the Markov chain. Let $0 = t_0 < t_1 < t_2 < \ldots < t_J$ be the times at which the path X changes state. Define the log-returns between the jump times, $W_k = \log(S_{t_k}/S_{t_{k-1}})$, $k = 1, 2, \ldots, J$. To obtain realizations of the W_k from the observed Y process, we need to simulate Gaussian random variables conditioned on their sums.

Consider any $t \in \{0, 1, \ldots, n\}$ for which the chain changes state atleast once in the time interval $[t-1, t]$. So for some p, k we have $t_{k-1} < t-1 \le t_k < \ldots < t_{k+p} \le t < t_{k+p+1}$. Let $V_t^1 = \log(S_{t_k}/S_{t-1})$ and $V_t^2 = \log(S_t/S_{t_{k+p}})$. Then,

$$Y_t = V_t^1 + \sum_{i=1}^{p} W_{k+i} + V_t^2. \quad (15)$$

Suppose that the chain X is in state j_i in the time interval $[t_{k+i-1}, t_{k+i})$, $i = 0, 1, \ldots, p+1$. Set $s_0 = t_k - (t-1)$, $s_i = t_{k+i} - t_{k+i-1}$, $i = 1, 2, \ldots, p$, and $s_{p+1} = t - t_{k+p}$. Let $m_j = \mu_{j_i} s_i$ and $v_j = \sigma_{j_i} s_i$, $i = 0, 1, \ldots, p+1$. Recall that $Y_t : \mathsf{N}(\mu(t), \sigma(t))$, where $\mu(t), \sigma(t)$ are as defined in (20). It is easy to see that the joint conditional density of $(V_t^1, W_{k+1}, \ldots, W_{k+p})$ given $Y_t = y$ will be

$$f(u_0, u_1, \ldots, u_p) = C \prod_{i=0}^{p} \exp(-\frac{1}{2} \frac{v_i + v_{p+1}}{v_i \, v_{p+1}} \qquad (16)$$

$$\left(\frac{v_{p+1} m_i + v_i (y - m_{p+1})}{v_i + v_{p+1}} \right)^2), \qquad (17)$$

where C is a constant that depends on y and the parameters. Thus, one can simulate the variables $V_t^1, W_k, W_{k+1}, \ldots, W_{k+p}$ from independent Gaussian distributions and then obtain V_t^2 using (15).

Using the above procedure, we can obtain a realization for all W_k for which $[t_{k-1}, t_k] \subseteq [t-1, t]$, for some $t \in \{0, 1, \ldots, n\}$. Now for any k for which there is a $q \geq 0$, such that $t - 1 \leq t_{k-1} < t < t+1 < \ldots < t+q \leq t_k < t+q+1$, we can obtain W_k using the relation

$$W_k = V_t^2 + \sum_{i=1}^{q} Y_{t+i} + V_{t+q+1}^1. \qquad (18)$$

Note that the W values depend on the path X and the parameter values μ, σ and hence are to be computed in each iteration of the Gibbs sampling procedure which we describe next.

IV. Markov regime switching with Dirichlet prior

In this section, we take $\overline{\alpha} = H$, the distribution of a continuous time Markov chain on a finite set of states, and we propose a new hierarchical model that is specified, in non-life reinsurance. Of course, this can be similarly used in many other cases. We consider the Black-Scholes with jump SDE in random environment with a Dirichlet prior on the path space of the chain, the states of the chain representing the environment due to the disasters. We model the loss claim using a Black-Scholes with jump and standard diffusion with drift, volatility and intensity depending on the state of the disasters. The state of the claim loss is modeled as a continuous time Markov chain with a Dirichlet prior. In what follows, the notations σ^2 and τ^2 will be both used to denote the variance rather than the standard deviations.

The following notations will be adopted:
- n will denote the number of observed data and also the length of an observed path.
- M will denote the number of states of the Markov chain.
- The state space of the chain will be denoted by $S = \{i : 1 \leq i \leq M\}$.
- N will denote the number of simulated paths.
- m will denote the number of distinct states of a path.

- The IBNR claims and the IBENR claims follow respectively the following two SDE:

$$\frac{dN_t}{N_t} = \alpha(t, X_t, N_{t^-})dt + \tau(t, X_t, N_{t^-})dB_t + \kappa dP_t, \qquad t \geq 0,$$

$$dD_t = \rho(t, X_t, D_{t^-}, N_{t^-})dt + \sigma(t, X_t, D_{t^-}, N_{t^-})dB_t, \quad t \geq 0.$$

where B_t is a standard Brownian motion and P_t is a Poisson process with intensity λ. By Ito's formula, the process $Z_t = \log(N_t)$ satisfies the SDE,

$$dZ_t = \mu(X_t)dt + \tau(X_t)dB_t + \kappa dP_t, \qquad t \geq 0,$$

where $\mu(X_t) = \alpha(X_t) - \frac{1}{2}\tau^2(X_t) + \kappa.\lambda(X_t)$. The observed data are of the form $Z_0, Z_1, \ldots, Z_n$.

Nidhi verma et al. Int. Journal of Engineering Research and Applications
www.ijera.com
ISSN : 2248-9622, Vol. 4, Issue 12(Part 3), December 2014, pp.63-77

• The process (X_t) is assumed to be a continuous time Markov process taking values in the set $S = \{i : 1 \le i \le M\}$. The transition probabilities of this chain are denoted by p_{ij}, $i, j \in S$ and the transition rate matrix is $Q_0 = (q_{ij})_{i,j \in S}$ with

$$\xi_i > 0, \qquad q_{ij} = \xi_i p_{ij} \quad if \quad i \ne j, \qquad and \qquad q_{ii} = -\sum_{j \ne i} q_{ij}, \qquad i, j \in S.$$

Let us define the log-returns, $Y_t = Z_t - Z_{t-1} = \log(N_t / N_{t-1})$ for the equation (13) and $Y_t = D_t - D_{t-1}$ for the second equation (14), $t = 1, 2, \ldots, n$. Suppose we know the path $X = \{X_s, 0 \le s \le n\}$. Let $T_j(t)$ be the time spent by the path X in state j in the time interval $[t-1, t]$. Define

$$\mu(t) := \sum_{j=1}^{M} \mu(j) T_j(t); \qquad \tau^2(t) := \sum_{j=1}^{M} \tau^2(j) T_j(t); \qquad (19)$$

$$\lambda(t) := \sum_{j=1}^{M} \lambda(j) T_j(t).$$

$$\rho(t) := \sum_{j=1}^{M} \rho(j) T_j(t); \qquad \sigma^2(t) := \sum_{j=1}^{M} \sigma^2(j) T_j(t). \qquad (20)$$

Then, conditional on the path X, the solutions of the equations (13) and (14) are respectively given by :
- the solution of the first equation (13) is

$$-1cm Y_t = Y_0 exp\{\int_0^t (\tau(s) dB_s + \kappa d\mathsf{P}_s) + \int_0^t (\mu(s) - \frac{\tau^2(s)}{2}) ds\}$$

$$(21)$$

$$2.5cm \times \Pi_{0 < s \le t} (1 + \kappa \Delta \mathsf{P}_s) e^{-\kappa \Delta \mathsf{P}_s},$$

$$t = 1, 2, \ldots, n,$$

and,

- Y_t are i.i.d. $\mathsf{N}(\rho_t, \sigma_t)$, $t = 1, 2, \ldots, n$ for the second equation (14).

• For each $i = 1, 2, \ldots, M$, the priors on $\mu_i = \mu(i)$, $\rho_i = \rho(i)$, $\tau_i = \tau(i)$, $\sigma_i = \sigma(i)$ and $\lambda_i = \lambda(i)$ are specified by

$$\rho_i, \mu_i : \mathsf{N}(\theta, \tau^\mu), \quad with \quad \theta : \mathsf{N}(0, A), A > 0, \qquad (22)$$

$$\sigma_i^2, \tau_i^2 : \Gamma(\nu_1, \nu_2), \qquad (23)$$

$$\lambda_i : \Gamma(c_1, c_2). \qquad (24)$$

• The Markov chain $\{X_t, t \ge 0\}$ has prior $\mathsf{D}(\alpha H)$, where H is a probability measure on the path space of cadlag functions $D([0, \infty), S)$. The initial distribution according to H is the uniform distribution $\pi_0 = (1/M, \ldots, 1/M)$, and the transition rate matrix is Q with $p_{ij} = 1/(M-1)$ and $\lambda_i = \lambda > 0$. Thus the Markov chain under Q will spend an exponential time with mean $1/\lambda$ in any state i and then jump to state $j \ne i$ with probability $1/(M-1)$.

A realization of the Markov chain from the above prior is generated as follows: Generate a large number of paths $X_i = \{x_s^i : 0 \le s \le n\}$, $i = 1, 2, \ldots, N$, from H. Generate the vector of probabilities

$(p_i, i = 1, \ldots, N)$ from a Poisson Dirichlet distribution with the parameter $\alpha,$ using stick breaking. Then draw a realization of the Markov chain from

$$p = \sum_{i=1}^{N} p_i \delta_{X_i}, \quad (25)$$

which is a probability measure on the path space $D([0, n), S)$. The parameter λ is chosen to be small so that the variance is large and hence we obtain a large variety of paths to sample from at a later stage. The prior for α is given by,

$$\alpha : \Gamma(\eta_1, \eta_2). \quad (26)$$

V.   Estimation

Estimation is made using the simulation of a large number of paths of the Markov chain which will be selected according to a probability vector (generated by stick-breaking) and then using the blocked Gibbs sampling technique. This technique uses the posterior distribution of the various parameters.

We denote by μ, ρ, λ, σ^2 and τ^2, the current values of the vectors $(\mu_1, \mu_2, \ldots, \mu_n)$, $(\rho_1, \rho_2, \ldots, \rho_n)$, $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$ and $(\tau_1^2, \tau_2^2, \ldots, \tau_n^2)$, respectively.

Let Y be the vector of observed data $(Y_1, \ldots, Y_n)$. Let $X = (x_1, x_2, \ldots, x_n)$ be the vector of current values of the states of the Markov chain at times $t = 1, 2, \ldots, n,$ respectively. Let $X^* = (x_1^*, \ldots, x_m^*)$ be the distinct values in X.

5.1 Modifying the observed data set

In order to obtain the conditional distribution of the parameters, we first need to extract the change in the log-claims and the recourse increments between the jump times of the Markov chain. Let $0 = t_0 < t_1 < t_2 < \ldots t_J$ be the times at which the path X changes state. Define the log-claims and the recourse increments between the jump times, $W_k = \log(N_{t_k}/N_{t_{k-1}})$ and $W_k = D_{t_k} - D_{t_{k-1}}$ $k = 1, 2, \ldots, J$ respectively. To obtain realizations of the W_k from the observed Y process, we need to simulate Gaussian random variables conditioned on their sums.

Consider any $t \in \{0, 1, \ldots, n\}$ for which the chain changes state at least once in the time interval $[t-1, t]$. Let $t_{k-1} < t-1 \le t_k < \ldots < t_{k+p} < t < t_{k+p+1}$, be the jump times that lie in $[t-1, t]$, for some $p \ge 1$. Let $V_t^1 = \log(N_{t_k}/N_{t-1})$, $\hat{V}_t^1 = D_{t_k} - D_{t-1}$, $\hat{V}_t^2 = D_t - D_{t_{k+p}}$ and $V_t^2 = \log(N_t/N_{t_{k+p}})$. Then the increment of the equations (13) and (14) respectively are

$$Y_t = V_t^1 + \sum_{i=1}^{p} W_{k+i} + V_t^2 \quad (27)$$

and

$$Y_t = \hat{V}_t^1 + \sum_{i=1}^{p} W_{k+i} + \hat{V}_t^2. \quad (28)$$

Suppose for some that the chain X is in state j_i in the time interval $[t_{k+i-1}, t_{k+i})$, $i = 0, 1, \ldots, p+1$. Set $s_0 = t_k - t - 1$, $s_i = t_{k+i} - t_{k+i-1}$, $i = 1, 2, \ldots, p,$ and $s_{p+1} = t - t_{k+p}$.

Let $m_j = \mu(j_i)s_i$, $v_j = \tau^2(j_i)s_i$ and $L_j = \lambda(j_i)s_i$ $i = 0, 1, \ldots, p+1$. Remember that

$$-1.5 cm Y_t = Y_0 exp\{\int_0^t (\tau(s)dB_s + \kappa d\mathsf{P}_s) + \int_0^t (\mu(s) - \frac{\tau^2(s)}{2})ds\}$$

$$2.5 cm \times \Pi_{0 < s \le t}(1 + \kappa \Delta \mathsf{P}_s)e^{-\kappa \Delta \mathsf{P}_s}$$

$$t = 1, 2, \ldots, n,$$

where $\mu(t), \tau^2(t)$ and $\lambda(t)$ are as defined in (20).

It is easy to see that the joint conditional density of $(V_t^1, W_{k+1}, \ldots, W_{k+p})$ given $Y_t = y$ is

$$f(u_0, u_1, \ldots, u_p) = C_1 \prod_{i=0}^{p} \sum_{l=1}^{+\infty} \frac{L_i^l}{l!} \exp(-\frac{1}{2v_i}(u_i - m_i - kl)^2 - L_i), \quad (29)$$

where C_1 is a constant that depends on y and the parameters. Thus, one can simulate the variables $V_t^1, W_k, W_{k+1}, \ldots, W_{k+p}$ from independent Gaussians and then obtain V_t^2 using (27).

Let $m_j = \rho(j_i)s_i$ and $v_j = \sigma^2(j_i)s_i$, $i = 0, 1, \ldots, p+1$. Remember that $Y_t : \mathsf{N}(\rho(t), \sigma^2(t))$, where $\rho(t), \sigma^2(t)$ are as defined in (20). It is easy to see that the joint conditional density of $(\hat{V}_t^1, W_{k+1}, \ldots, W_{k+p})$ given $Y_t = y$ is

$$f(u_0, u_1, \ldots, u_p) = C_2 \prod_{i=0}^{p} \exp(-\frac{1}{2} \frac{v_i + v_{p+1}}{v_i \, v_{p+1}} \quad (30)$$

$$-\frac{v_{p+1}m_i + v_i(y - m_{p+1})}{v_i + v_{p+1}})^2),$$

where C_2 is a constant that depends on y and the parameters. Thus, one can simulate the variables $\hat{V}_t^1, W_k, W_{k+1}, \ldots, W_{k+p}$ from independent Gaussians and then obtain $\hat{V}_t^2$ using (28).

Using the above procedure, we can obtain a realization for all W_k for which $[t_{k-1}, t_k] \subseteq [t-1, t]$, for some $t \in \{0, 1, \ldots, n\}$. Now for any k for which there is a $q \geq 0$, such that $t - 1 \leq t_{k-1} < t < t+1 < \ldots < t+q \leq t_k < t+q+1$, we can obtain W_k using the relations

$$W_k = V_t^2 + \sum_{i=1}^{q} Y_{t+i} + V_{t+q+1}^1 \quad (31)$$

and

$$W_k = \hat{V}_t^2 + \sum_{i=1}^{q} Y_{t+i} + \hat{V}_{t+q+1}^1. \quad (32)$$

Note that the W values depend on the path X and need to be computed in each iteration.

5.2 The Gibbs sampling procedure

We are now ready to estimate the posterior distributions of the parameters using Gibbs sampling. Each iteration produces one realization of the parameters from their approximate posterior distribution. Each iteration consists of a large number of samples obtained recursively for each parameter conditioned on the current values of the other parameters and the data.

5.2.1 Estimation procedure of the IBNR claim parameters

- **Conditional for** μ. For each $j \in X^*$ let us draw

$$\mu_j : F_{\mu_j}$$

where the density of μ_j is calculated as follows:

$$dP_{\mu_j}(y)\{\mu_j \mid p,\mu,W\,X,\theta\} = \frac{dP_{\mu_j}\{\mu_j,p,\tau^2,W\,X\}}{\int dP_{\mu_j}\{\mu_j,p,\tau^2,W\,X\}}$$

$$= \frac{dP_{\mu_j}(y)d_W P\{W \mid H\}dP\{\theta,\tau_\mu^2,p\}}{\int dP_{\mu_j}d_W P\{W \mid H\}dP(\theta,\tau_\mu^2,p)}$$

$$= \frac{dP_{\mu_j}(y)\sum_{l=1}^{\infty}d_W P\{\eta+kN \mid H,N=l\}P(N=l)}{\int dP_{\mu_j}(y)\sum_{l=1}^{\infty}d_W P\{\eta+kN \mid H,N=l\}P(N=l)}$$

Where $H = (\tau^2,X,p,\mu_j = y\theta,\tau_\mu^2)$. For each $j \in X \setminus X^*$, independently simulate $\mu(j):\mathsf{N}(\theta,\varsigma^\mu)$.

- **Conditional for** τ. For each $j \in X^*$ let us draw

$$\tau(j):F_{\tau_j}$$

where the density of τ_j is calculated as follows:

$$dP_{\tau(j)} \quad (y) \quad \{\tau^2(j) \mid p,\mu,W\,X,\theta\} = \frac{dP_{\tau_j}\{\tau_j^2,p,\mu,X \mid \tau_j^2 = y\}}{\int dP_{\tau_j^2}\{\tau_j^2,p,\mu,X \mid \tau_j^2 = y\}}$$

$$= \frac{dP_{\tau_j}(y)d_W P\{W \mid \tau^2,X,p,\tau_j^2 = y\theta,\tau_\mu^2\}dP\{\theta,\tau_\mu^2,p\}}{\int dP_{\tau_j}d_W P\{W \mid \tau^2,X,p,\tau_j^2 = y\theta,\tau_{\tau^2}^2\}dP(\theta,\tau_\mu^2,p)}$$

$$= \frac{dP_{\tau_j}(y)\sum_{l=1}^{\infty}d_W P\{\eta+kN \mid F,N=l\}P(N=l)}{\int dP_{\tau_j}(y)\sum_{l=1}^{\infty}d_W P\{\eta+kN \mid F,N=l\}P(N=l)}$$

$$= \frac{\dfrac{y^{v_1}exp(\dfrac{-y}{v_2})}{\Gamma(v_1)v_2^{v_1}}\sum_{l=1}^{\infty}\Pi_{\{k:X_{t_k}=j\}}N_{k,l}}{\int \dfrac{y^{v_1}exp(\dfrac{-y}{v_2})}{\Gamma(v_1)v_2^{v_1}}\sum_{l=1}^{\infty}\Pi_{\{k:X_{t_k}=j\}}N_{k,l}}$$

Where $F = (X,p,\tau_j^2 = y\theta,\tau_\mu^2)$. and $N_{k,l} = \dfrac{1}{\sqrt{2\pi\tau^2(t_k-t_{k-1})}}exp(w_k-(y+kl))^2\dfrac{(\Delta\lambda)^l exp(-\Delta\lambda)}{l!}$

Also for each $j \in X \setminus X^*$, independently simulate $\tau^2(j):\Gamma(v_1,v_2)$.

- **Conditional for** X.

$$(X \mid p): \sum_{i=1}^{N} p_i^* \delta_{X_i}, \qquad (33)$$

where

$$P\{X = X_i \mid p,\mu,\tau,W\} = P\{W \mid p,\tau,X = X_i,\mu\} \times P\{X = X_i \mid \tau,\mu,p\}P\{\mu,\sigma\} \propto \Sigma_{l=1}^{\infty} \Pi_{\{k:X_{t_k} = j\}} \frac{1}{\sqrt{2\pi\tau^2(t_k - t_{k-1})}} \times exp(w_k - (y + kl))^2 \frac{(\Delta\lambda)^l \times \exp(-\Delta\lambda)}{l!}) \, p_i$$

where $\{x_1^{i,*},\ldots,x_m^{i,*}\}$ denote the current $m = m(i)$ unique values of the states and t_k^i, W_k^i are as defined in subsection 5.1 for the path X_i, $i = 1,\ldots,N$.

- **Conditional for** p.

$$p_1 = V_1^*, \; and \; p_k = (1 - V_1^*)\cdots(1 - V_{k-1}^*)V_k^*, \; k = 2,3,\ldots,N-1, \text{(34)}$$

where

$$V_k^*: \beta(1 + r_k, \alpha),$$

r_k equal 1 if $i = k$ and 0 else.

- **Conditional for** α.

$$(\alpha \mid p): \Gamma\left(N + \eta_1 - 1, \eta_2 - \sum_{i=1}^{N-1} \log(1 - V_i^*) \right),$$

where the V^* values are those obtained in the simulation of p in the above step.

- **Conditional for** θ.

$$(\theta \mid \mu): \mathsf{N}(\theta^*, \varsigma^*), \qquad (35)$$

where

$$\theta^* = \frac{\tau^*}{\varsigma^\mu} \sum_{j=1}^{M} \mu_j,$$

and

$$\varsigma^* = \left(\frac{M}{\tau^\mu} + \frac{1}{A} \right)^{-1}.$$

5.2.2 Estimation procedure of the IBNER claim parameters

We are now ready to estimate the posterior distributions of the parameters using Gibbs sampling. Each iteration produces one realization of the parameters from their approximate posterior distribution. Each iteration consists of a large number of samples obtained recursively for each parameter conditioned on the current values of the other parameters and the data.

- **Conditional for** ρ. For each $j \in X^*$ let us draw

$$(\rho_j \mid \theta, \tau^\rho, \sigma, X, W): \mathsf{N}(\rho_j^*, \sigma_j^*), \qquad (36)$$

where

$$\rho_j^* = \sigma_j^* \left(\sum_{k:X_{t_{k-1}} = j} \frac{W_k}{\sigma_j(t_k - t_{k-1})} + \frac{\theta}{\tau^\rho} \right),$$

$$\sigma_j^* = \left(\frac{n_j}{\sigma_j} + \frac{1}{\tau^\rho} \right)^{-1},$$

and n_j being the number of times j occurs in X. For each $j \in X \setminus X^*$, independently simulate $\rho_j : \mathsf{N}(\theta, \tau^\rho)$.

- **Conditional for σ.** For each $j \in X^*$ let us draw

$$\left(\sigma_j \mid \rho, \nu, X, W \right) : \Gamma(\nu_1 + \frac{n_j}{2}, \nu_{2,j}^*), \qquad (37)$$

where

$$\nu_{2,j}^* = \nu_{2,j} + \sum_{k: X_{t_{k-1}} = j} \frac{(W_k - \rho_j(t_k - t_{k-1}))^2}{2(t_k - t_{k-1})}.$$

Also for each $j \in X \setminus X^*$, independently simulate $\sigma_j : \Gamma(\nu_1, \nu_2)$.

- **Conditional for X.**

$$\left(X \mid p \right) : \sum_{i=1}^{N} p_i^* \delta_{X_i}, \qquad (38)$$

where

$$p_i^* \propto \prod_{j=1}^{m} \left(\prod_{\{k : x_i^{i,*} = j\} \atop t_{k-1}^i} \frac{1}{(2\pi\sigma_j(t_k - t_{k-1}))^{1/2}} \right.$$
$$\qquad (39)$$

$$\times e^{-\frac{1}{2\sigma_j}(W_k^i - \rho_j(t_k - t_{k-1}))^2} \right) p_i, \qquad (40)$$

where $\{x_1^{i,*}, \ldots, x_m^{i,*}\}$ denote the current $m = m(i)$ unique values of the states and t_k^i, W_k^i are as defined in subsection 5.1 for the path X_i, $i = 1, \ldots, N$.

- **Conditional for p.**

$$p_1 = V_1^*, \text{ and } p_k = (1 - V_1^*) \cdots (1 - V_{k-1}^*) V_k^*, \; k = 2, 3, \ldots, N-1, \; (41)$$

where

$$V_k^* : \beta(1 + r_k, \alpha),$$

$r_k = 1$ if $i = k$ and 0 otherwise.

- **Conditional for α.**

$$\left(\alpha \mid p \right) : \Gamma \left(N + \eta_1 - 1, \eta_2 - \sum_{i=1}^{N-1} \log(1 - V_i^*) \right),$$

where the V^* values are those obtained in the simulation of p in the above step.

- **Conditional for θ.**

$$\left(\theta \mid \rho \right) : \mathsf{N}(\theta^*, \tau^*), \qquad (42)$$

where

$$\theta^* = \frac{\tau^*}{\tau^\rho} \sum_{j=1}^{M} \rho_j,$$

and

$$\tau^* = \left(\frac{M}{\tau^\rho} + \frac{1}{A} \right)^{-1}.$$

VI.   Implementation

The algorithm presented in the previous section was implemented in C language. The implementation includes

- functions that simulate standard probability distributions Uniform, Normal, Gamma, Beta, Exponential.

- a function that returns an index $\in \{1,\dots,n\}$ according to a vector of probability $p_1,\dots,p_n$.

- a function that simulates a probability vector according to stick-breaking scheme.

- a function that simulates n paths of a Markov chain.

- a function that records the number of times a state appears in a path.

- a function that chooses one of the paths according to a vector of probability.

- a function that modifies the parameters of prior distributions according to the formulas of the posteriori distributions.

After having simulated a number of paths, we perform the iterations. At each iteration a path is randomly selected and the parameters are updated according to posteriori formulas. At the end of each iteration of the Gibbs sampling, we obtain a path X of the Markov chain. From this, the parameters π and Q_0 can be re-estimated. From Q_0 the parameters λ_i and p_{ij} can be derived.

6.1 Reinsurance data of the B.E.S.T.R.É company

We have also applied our algorithm to the reinsurance data of the B.E.S.T.R.É company from $21/12/2006$ to $15/11/2007$. For this dataset of Incurred But Not Enough Reported claims reserve D_{ij}. We have, $n = 250$, $\Delta t = 1$, and we deal with $N = 100$ paths while Gamma$(2,4)$ is the prior for α.

With the above choice, we obtain four regimes for which the estimates for the mean, variance and stationary probabilities are as follows:

	R 1	R 2	R 3	R 4
μ	-740.4	1110.6	10612.4	-1727.6
σ	5.9	7.2166	12.3023	4.3800
π	43 %	13 %	10 %	34 %

The parameters λ_is and the matrix of transition probability $(p_{i,j})_{i,j=1,\dots,64, i \neq j}$ of the most frequent Markov chain path are respectively equal to:

λ_1	λ_2	λ_3	λ_4
.5	1.2	1	1.12

$$\begin{pmatrix} & 0.5 & 0.03 & 0.47 \\ 0.04 & & 0.61 & 0.35 \\ 0.26 & 0.2 & & 0.54 \\ 0.375 & 0.325 & 0.3 & \end{pmatrix}$$

On the other hand, for the dataset of Incurred But Not Reported claims reserve N_{ij}. We have, $n = 250$, $\Delta t = 1$, and we deal with $N = 100$ paths while Gamma$(2,4)$ is the prior for α.

With the above choice, we obtain six regimes for which the estimates for the mean, variance and stationary probabilities are as follows:

	R 1	R 2	R 3	R 4	R 5	R 6
μ	9225	35254	863.2	285421	13254.2	22581
σ	9.132	7.2166	2.3023	7.3800	1.186	3.3372 e-05
π	20 %	3 %	29 %	5 %	10 %	33 %

The most frequent Markov chain path, its parameters λ_i s and the matrix of the transition probability $(p_{i,j})_{i,j=1,\dots,6,i\neq j}$ are respectively equal to:

3 5 3 6 3 6 3 6 1 6 5 1 3 6 3 5 3 3 6 6 5 6 3 6 1 1 4 1 6 1 3 3 6 6 6 3 1 3 3 3 6 3 3 3 4 5 6 6 6 6 4 6 1 1 1 6 6 6 6 6 1 3 3 3 1 6 1 3 3 5 6 3 3 1 6 5 4 1 3 6 4 6 3 3 5 6 3 6 2 3 6 1 3 3 6 1 6 6 5 5 1 1 5 3 5 3 3 6 1 6 5 6 1 6 6 3 1 6 3 1 1 6 2 3 6 6 6 3 3 2 6 6 6 1 3 3 6 6 3 1 3 6 6 1 6 6 1 1 6 1 5 3 5 1 3 5 3 4 1 3 3 5 3 1 3 6 6 6 1 3 5 6 5 3 3 6 3 6 1 3 5 6 6 6 5 1 6 3 3 1 1 6 6 6 3 6 1 3 6 3 6 6 6 6 3 6 3 6 6 4 6 3 6 1 1 6 4 6 1 3 4 3 6.

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
.25	1	1.42	1	1.05	1.73

$$\begin{pmatrix} & 0 & 0.48 & 0.03 & 0.06 & 0.42 \\ 0 & & 0.66 & 0 & 0 & 0.33 \\ 0.16 & 0.02 & & 0.062 & 0.2 & 0.54 \\ 0.375 & 0 & 0 & & 0.125 & 0.5 \\ 0.157 & 0 & 0.42 & 0.052 & & 0.36 \\ 0.36 & 0.038 & 0.384 & 0.077 & 0.134 & \end{pmatrix}$$

VII.   Conclusion

A Bayesian approach to estimation for a regime switching geometric Brownian motion is proposed. The algorithm while being computationally intensive is able to segregate the different regimes based on the drift and volatility, thus giving useful insights into the behavior of the sinister. It has been observed empiricaly that sinisters fluctuate between periods of high, moderate and low volatilities. The above estimation procedure provides a clear quantitative picture of the number of regimes and an estimate of the drifts and volatilities in these regimes. Estimation of current sinister state is also easier using the algorithm proposed compared to models using continuous stochastic volatility models. Given an estimate of the regime, the algorithm also gives an idea of likely duration for which the regime is likely to persist and the distribution of the regimes that may follow.

References
[1]   Antoniak, C.E.. Mixtures of Dirichlet processes. *Ann. Statist.* **2**, 6, 1152-1174.(1974)
[2]   Blackwell, D., and MacQueen, J. B. . Ferguson distributions via Polya urn schemes. *Ann. Statist.* **2**,1, 353-355.(1973)
[3]   Brunner, L.J. and Lo, A. Y.. Bayesian classification To appear.(2002)
[4]   Blei, D., Ng, A., and Jordan, M.. Latent Dirichlet allocation. *Journal of Machine Learning Research,* 3, 993-1022.(2003)
[5]   Cifarelli, D. M. and Melilli, E.. Some new results for Dirichlet priors, *Journal of the American Statistical Association, Ann. Statist.* 28 1390-1413. (2000).
[6]   David Blei and Michael I. Jordan. Variational inference for Dirichlet process mixtures. *Bayesian Analysis, 1(1)* 121-144. (2005).
[7]   Dahl, D. B. Modeling differential gene expression using a Dirichlet Process mixture model, *in Proceedings of the American Statistical Association, Bayesian Statistical Sciences Section, American Statistical Association, Alexandria, VA.*(2003a).

[8]   Damien, P. , Wakefield,J. C. and S. G.Walker. . S. G.Walker. Gibbs sampling for Bayesian nonconjugate and hierarchical models using auxiliary variables.*Journal of the Royal Statistical Society Series B,* **61:** 331-344. (1999).

[9]   Deshpande, A. and Ghosh, M. K.. Risk Minimizing Option Pricing in a Regime Switching Market. *Stochastic Analysis and Applications,* Vol. 28, 2008. To appear. (2007).

[10]   Diaconis, P. and Kemperman, J. Some new tools for Dirichlet priors, *In Bayesian Statistics 5 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.)* (1996)

[11]   Duan, J. A. Guindani, M. and Gelfand, A.E. Generalized spatial Dirichlet process models. *Biometrika* **94**, 4, 809-825. (2007).

[12]   Dunson, D. B. and Park, J.-H.. Kernel stick-breaking processes. Tech. rep., *Institute of Statistics and Decision Sciences* Duke University. (2007).

[13]   Emilion, R. . Process of random distributions, *Afrika Stat*, vol 1, 1, pp. 27-46. http://www.ufrsat.org/jas (contenus). (2005).

[14]   Faires,H (2012). SDEs in Dirichlet Random Environment *IJSS* **1**, 55–66.

[15]   Ferguson, T.S. A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209-230. (1973)

[16]   Ferguson, T.S.. Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615-629. (1974).

[17]   Ferguson, T. S.. Bayesian density estimation by mixtures of normal distributions. *Recent Advances in Statistics (H. Rizvi and J. Rustagi, eds.), New York: Academic Press.* (1983).

[18]   Ishwaran H., James, L.F and Sun, J. Bayesian Model Selection in Finite Mixtures by Marginal Density Decompositions,*conditionally accepted by Journal of the American Statistical Association.*(2000).

[19]   Ishwaran, H. and Zarepour, M. Markov Chain Monte Carlo in Approximate Dirichlet and Beta Two-Parameter Process Hierarchical Models, *EJP* **1**, 1-28. (2000).

[20]   Ishwaran, H. and James, L.F.. Approximate Dirichlet processes computing in finite normal mixtures: smoothing and prior information. *J. Comp. Graph. Stat.* **11** 209-230. (2002).

[21]   Kingman, J. F. C. Poisson Processes. Clarendon, Oxford University Press. (1993).

[22]   Kingman, J. F. C. Random discrete distributions. *J. Roy. Statist. Soc. B*, **37,** 1-22. (1975).

[23]   Kottas, A, - Duan, J. A., and Gelfand, A.E. Modeling disease incidence data with spatial and spatio-temporal Dirichlet mixtures. *Biometrical J.* **49**, 5, 1-14. (2007).

[24]   Mack, T. Distribution-free Calculation of the Standard Error of Chain ladder Reserve Estimates. *ASTIN Bulletin 23 (1)* 214-225.(2000)

[25]   MacEachern, S.N. Dependent Dirichlet processes. *Technical Report Ohio State University.* Dependent Dirichlet processes. Technical Report. Ohio State University. (2000)

[26]   Muller, P., Quintana, F., and Rosner, G.. Hierarchical meta-analysis over related non-parametric Bayesian models. *Journal of Royal Statistical Society.* Series B , **66** , 735-749. (2004).

[27]   Pitman, J. and Yor, M.. Random discrete distributions derived from self-similar random sets. *EJP* **1**, 1-28. (1996).

[28]   Pitman, J. and Yor, M.. The two-parameter Poisson-Dirichlet distribution derived from stable subordiantors. *Ann. Proba* **25**, 2, 855-900.(1996)

[29]   Pitman, J. Poisson-Kingman partitions. In: Science and Statistics: A Festschrift for Terry Speed. D. R. Goldstein editor. Lecture Notes - Monograph Series 30 1-34 *Institute of Mathematical Statistics Hayward, California*. **2**Lecture Notes - Monograph Series 30 1-3. (2003).

[30]   Rodriguez, A. B. DUNSON, D. and GELFAND, A. E.. The Nested Dirichlet process. Preprint JASA. (2002)

[31]   Liu, H. and and Verrall, R.J.. Predictive Distributions for Reserves which Separate true *IBNR*and *IBNER*Claims.http://www.actuaries.org/ASTIN/Colloquia/Orlando/Papers/Liu.pdf. (2002).

[32]   Schnieper, R. (1991).Separating True *IBNR* and *IBNER* Claims. *ASTIN Bulletin 21* **1** 111-127.

[33]   Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica* **4** 639-650.

[34]   Sethuraman, J. and Tiwari, R. C. (1982).Convergence of Dirichlet Measures and the Interpretation of Their Parameters, *Statistical Decision Theory and Related Topics* III, 2, 305-315.

[35]   Sethuraman, J. and West, M.(1994).Bayesian Density Estimation and Inference Using Mixtures, *Journal of the American Statistical Association, 90, pp.* , 577-588.

[36]   Teh, Y. W., Beal, M. J., and and Blei, D. M.(2006). Sharing clusters among related groups: Hierarchical Dirichlet processes. *Journal of the American Statistical Association,* **101**, 1566-1581.

[37]   Verrall, R.J. and England, P.D.(2006). Comments on: "A comparison of stochastic models that reproduce chain ladder reserve estimates", by Mack and Venter. *Insurance: Mathematics and Economics,* **101**, 26, 1, 109-111. (2000). http://ideas.repec.org/a/eee/insuma/v26y2000i1p109-111.html.