Time Dependent Stick Breaking Processes

Jim Griffin

University of Kent

Joint work with Mark Steel (University of Warwick)

Outline

Introduction

We are interested in flexibly modelling distributions that change over time.

We observe data $y_1, y_2, \ldots, y_n$ at times $t_1, t_2, \ldots, t_n$.

A popular nonparametric approach assumes a time-dependent infinite mixture model

$$p(y_i|t_i) = \int \mathrm{N}(y_i|\mu, \sigma^2)\, dF_{t_i}(\mu, \sigma^2) = \sum_{j=1}^{\infty} w_j(t_i)\mathrm{N}(y_i|\mu_j, \sigma_j^2)$$

Introduction

The definition of F_{t_i} often generalizes generic constructions of standard nonparametric process such as the Dirichlet process.

For example,

- the Polya urn representation of the Dirichlet process.
- normalizing time-dependent random measures.
- the stick-breaking construction of the Dirichlet process.

Stick Breaking Processes

A random distribution F is a stick-breaking process if

$$F = \sum_{j=1}^{\infty} w_j \delta_{\theta_j}$$

where

$$w_j = V_j \prod_{k<j} (1 - V_k)$$

and

- $V_1, V_2, V_3, \ldots$ are independent and $V_j \sim \text{Be}(a_j, b_j)$.
- $\theta_1, \theta_2, \theta_3, \cdots \overset{i.i.d.}{\sim} H$.

Examples

- The Dirichlet process with concentration parameter M arises if $V_j \sim \text{Be}(1, M)$.
- The Poisson-Dirichlet (Pitman-Yor) process arises if $V_j \sim \text{Be}(1 - b, M + bj)$.

Random Walks on Discrete Distributions

Suppose that y_t is a real-value random variable then we can define a random walk in two (equivalent) ways:

$$y_{t+1} = \mathrm{N}(y_t, \sigma^2) \text{ or } y_{t+1} = y_t + \epsilon_t, \epsilon_t \sim \mathrm{N}(0, \sigma^2).$$

Wih random distribution we could write

$$F_{t+1} \sim \mathrm{DP}(M, F_t) \text{ or } F_{t+1} = V_t\, F_t + (1 - V_t)\,\epsilon_t.$$

These constructions are not equivalent since F_t is a discrete distribution.

Under the first process F_t tends to a point mass as $t \to \infty$.

Random Walks on Discrete Distributions

If ϵ_t is a distribution with a single atom at θ_t then

$$F_t = \sum_{j=-\infty}^{t} \delta_{\theta_j} V_j \prod_{k=j+1}^{t} (1 - V_k)$$

which is stick-breaking "backwards in time".

This type of process would be rather inflexible. Allowing V's to arrive according to a Poisson process and applying the stick-breaking process increases flexibility

DPAR process

$$F_t = \sum_{j=1}^{\infty} p_j(t)\delta_{\theta_j}$$

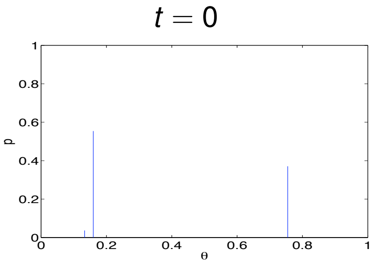
where

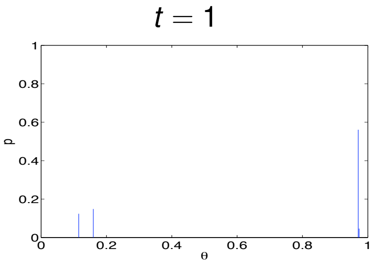
$$p_j(t) = \begin{cases} 0 & \tau_j > t \\ V_j \prod_{\{k|\tau_j<\tau_k<t\}}(1 - V_k) & \tau_j < t \end{cases}$$

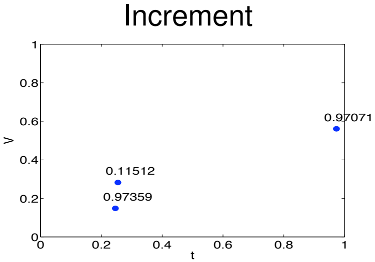
for

- $V_1, V_2, V_3, \dots \overset{i.i.d.}{\sim} \text{Be}(1, M)$.
- $\tau_1, \tau_2, \tau_3, \dots$ follow a Poisson process with intensity λ on $(-\infty, \infty)$.
- $\theta_1, \theta_2, \theta_3, \dots \overset{i.i.d.}{\sim} H$.

DPAR process

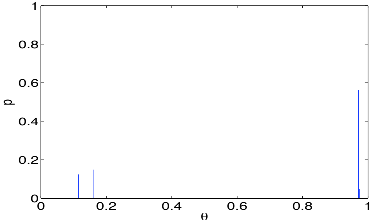




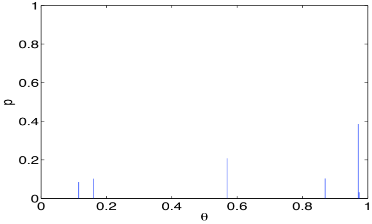


DPAR process

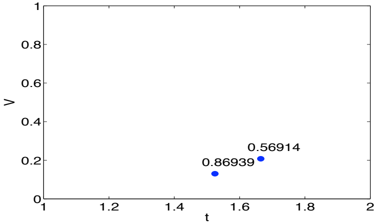
$t = 1$



$t = 2$

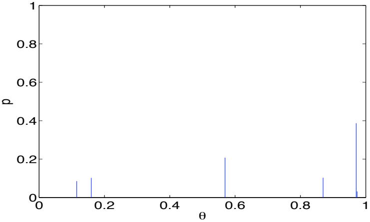


Increment

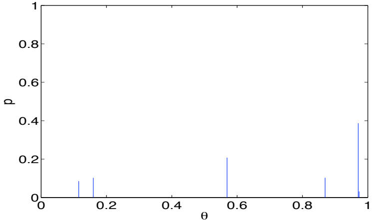


DPAR process

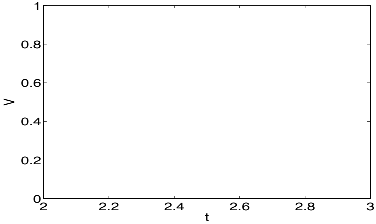
$t = 2$



$t = 3$



Increment



Properties

- F_t is DP(M, H) for all t.
- The autocorrelation is

$$\text{Corr}(F_t(B), F_{t+s}(B)) = \exp\left\{-\frac{\lambda}{M+1}s\right\}$$

Chinese restaurant representation

We integrate over the atoms which have no point associated with them to give a finite dimensional representation of the process.

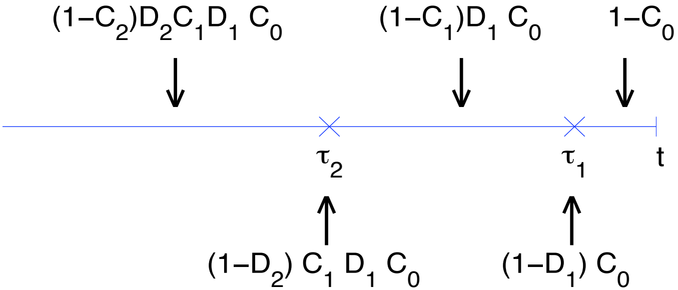
Let the active set at time r be

$$A(r) = \#\{j | \tau_{s_j} < r \leqslant t_j\}$$

$$m_j = \sum_{i=1}^{n} \mathrm{I}(s_i = j)$$

Let $\mathcal{T}_n$ be the subset of our Poisson process which have an observation allocated to them and $\mathcal{S}_n = \mathcal{T}_n \cup \{t_1, \ldots, t_n\}$.

Chinese restaurant representation



$$D_i = \frac{1 + m_j}{1 + m_j + M + A(\tau_j)}$$

$$C_0 = \rho^{(t_{n+1} - \max\{\mathcal{S}_n\})}, \quad C_i = \rho^{\frac{M(M+1)}{(M+A(\phi_i))(1+M+A(\phi_i))}(\tau_i - \tau_{i-1})}$$

Chinese restaurant representation

If we draw a new value within (τ_{l-1}, τ_l) then the new position is $\tau^\star = \tau_l - x$ where x is exponentially distributed with parameter $\frac{\lambda M(\tau_l - \tau_{l-1})}{(M+A(\tau_l))(M+A(\tau_l)+1)}$ truncated to the region$(0, \tau_l - \tau_{l-1})$.

Π-AR processes

In the DPAR the distribution of the V's do not depend on their position in the ordering at a given time.

For more general process, such as Poisson-Dirichlet, we need the distribution of V to depend on its position in the ordering. The position is changing over time so we need V to follow a stochastic process.

Π-AR processes

$$F_t = \sum_{i=1}^{\infty} p_j(t)\delta_{\theta_j}$$

where

$$p_j(t) = \begin{cases} 0 & \tau_j > t \\ V_j(t)\prod_{\{k|\tau_j<\tau_k<t\}}(1 - V_k(t)) & \tau_j < t \end{cases}$$

where

- $V_1(t), V_2(t), V_3(t), \cdots \overset{i.i.d.}{\sim} \text{Be}(1, M)$.
- $\tau_1, \tau_2, \tau_3, \ldots$ follow a Poisson process with intensity λ on $(-\infty, \infty)$.
- $\theta_1, \theta_2, \theta_3, \cdots \overset{i.i.d.}{\sim} H$.

A stochastic process for V

If $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ are both non-decreasing sequences then the random sequence $V^{(1)}, V^{(1)}, V^{(3)}, \ldots$ defined by

$$V^{(j+1)} = w_j V^{(j)} + (1 - w_j)\epsilon_j$$

where $w_j \sim \text{Be}(a_j + b_j, a_{j+1} + b_{j+1} - a_j - b_j)$ and $\epsilon_j \sim \text{Be}(a_{j+1} - a_j, b_{j+1} - b_j)$ implies that $V^{(j)} \sim \text{Be}(a_j, b_j)$.

PDAR

We say that $\{F_t\}_{t=-\infty}^{\infty}$ follows a PDAR(M, b, λ, H) if

$$F_t = \sum_{i=1}^{\infty} p_j(t)\delta_{\theta_j}$$

where $\tau_1, \tau_2, \tau_3, \ldots$ follow a homogeneous Poisson process with intensity λ and $\theta_1, \theta_2, \theta_3, \cdots \stackrel{i.i.d.}{\sim} H$ and

$$p_j(t) = \begin{cases} 0 & \tau_j > t \\ V_j(t)\prod_{k|\tau_j<\tau_k<t}(1 - V_k(t)) & \tau_j < t \end{cases}$$

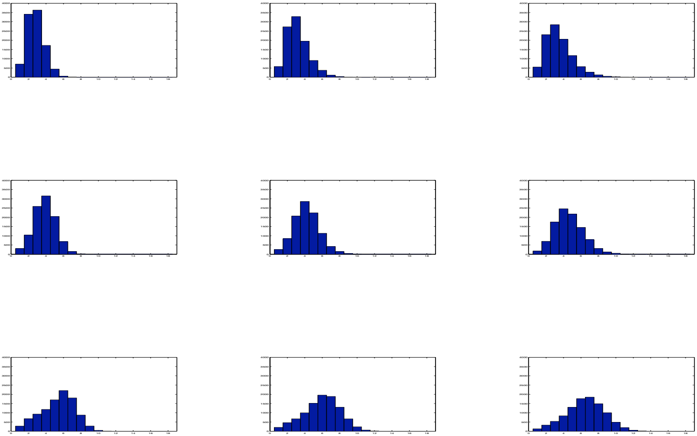
and

$$V_j(t) = \epsilon_j \prod_{m=1}^{t} w_{j,m}$$

where $\epsilon_j \sim \text{Be}(1-b, M+b)$, $w_{j,1} = 1$ and $w_{j,m} \sim \text{Be}(1 + M + b(m-1), b)$.

PDAR

	$b = 0$	$b = 0.1$	$b = 0.2$
$C = 2$			
$C = 4$			
$C = 8$			



Computation

Markov chain Monte Carlo are fairly straightforward by extending the Retrospective Sampling methods of Papaspiliopoulos and Roberts (2008).

The breaks of the PDAR process have a product form and the posterior can be sampled using an extension of methods for Matrix Stick Breaking Processes.

Stochastic volatility

Let $p_1, p_2, \ldots, p_T$ be the daily values of a stock index

$$y_t = \log p_{t+1} - \log p_t = \sqrt{h_t}\epsilon_t$$

where ϵ_t are independently drawn from some returns distribution and the conditional variance h_t is modelled

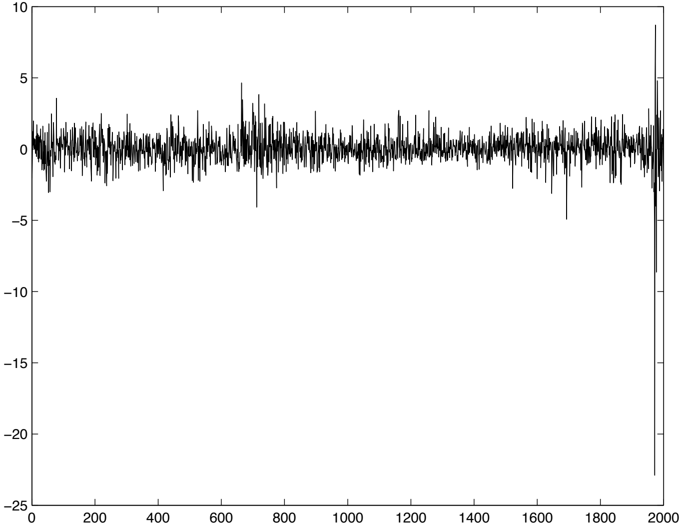
$$\log h_t \sim \mathrm{N}(\delta \log h_{t-1}, \sigma_v^2).$$

where

$$p(\epsilon_t) = \int \mathrm{N}(\mu, \sigma^2)\, dF_t(\mu, \sigma^2)$$

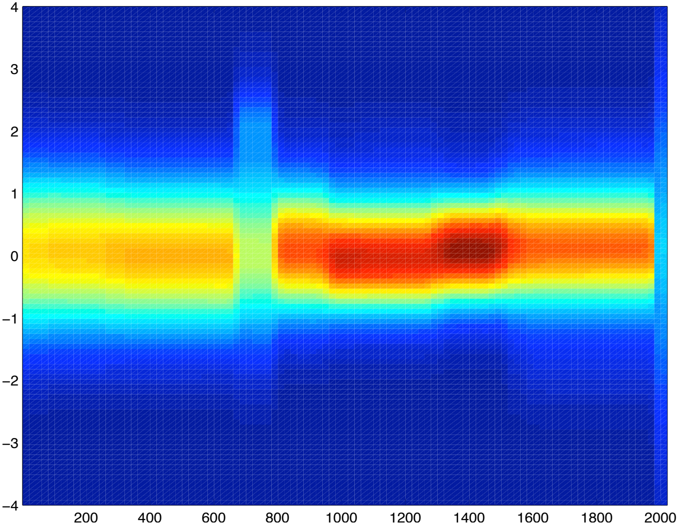
Stochastic volatility

Standard and Poors index from 1/1/1980 until 31/12/1988



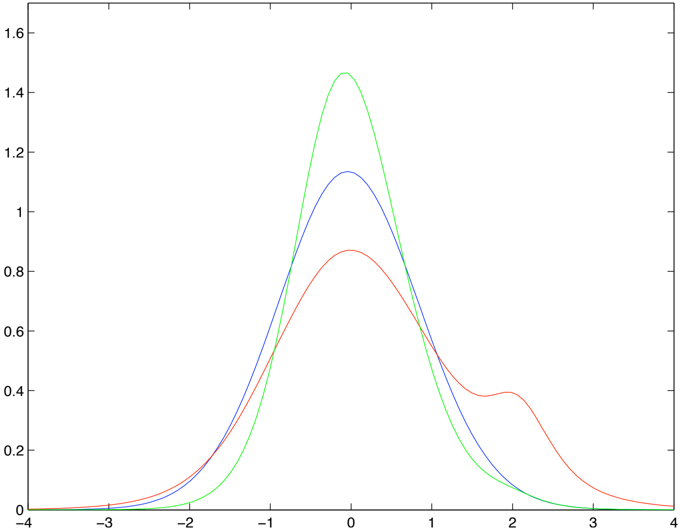
Stochastic volatility

Fitted return distributions over time



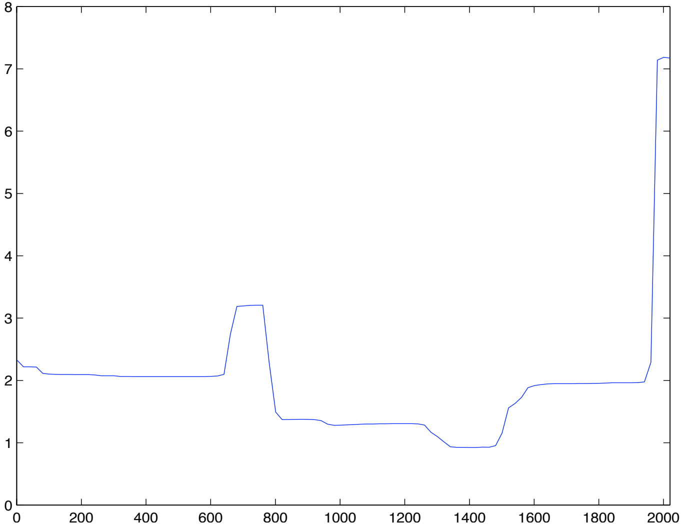
Stochastic volatility

Fitterd return distributions at various times



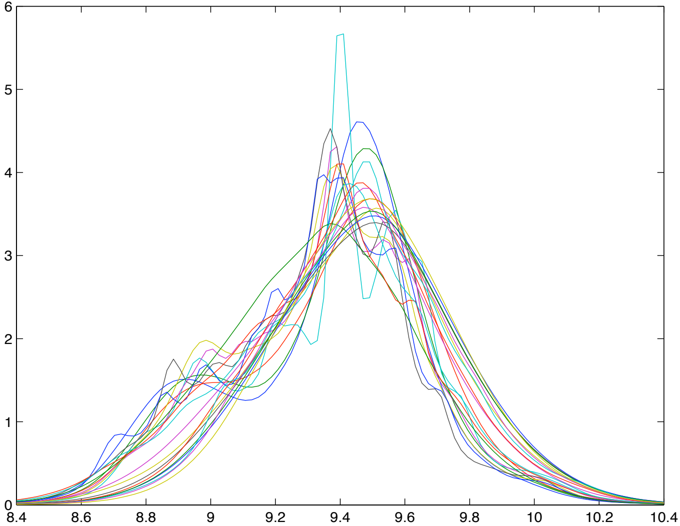
Stochastic volatility

Estimated variance over time



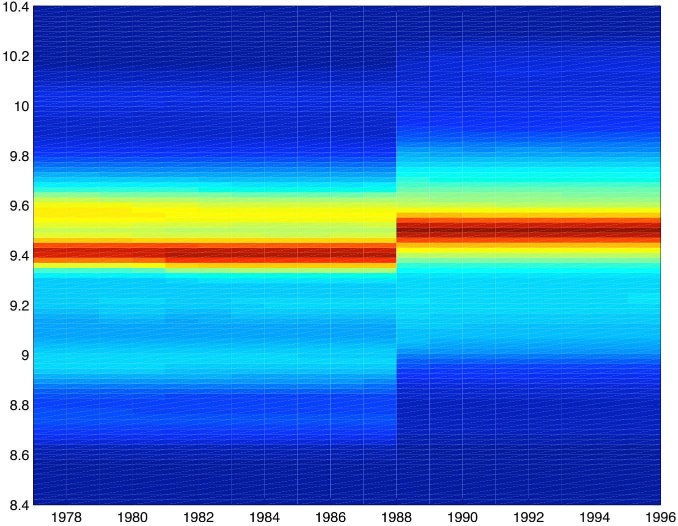
Time-dependent density estimation

We model real (log) per capita GDP over 110 EU regions from 1977 to 1996.



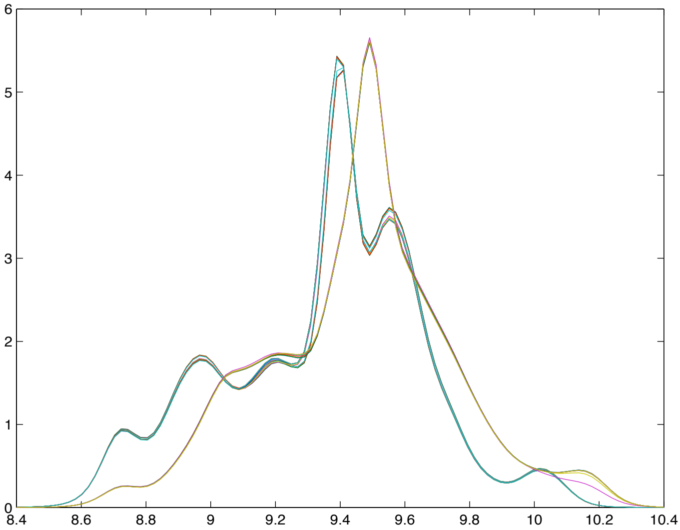
Time-dependent density estimation

Fitted distributions over time

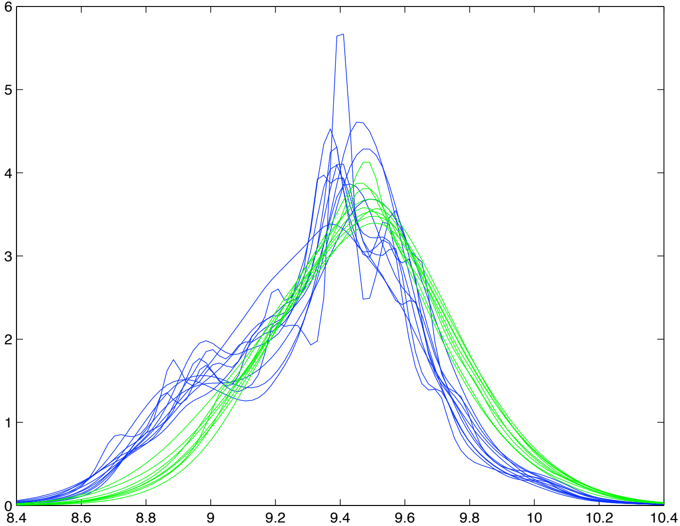


Time-dependent density estimation

Fitted distributions for each year



Time-dependent density estimation



Discussion

- Π-AR process allow us to construct stochastic process on probability measures with a wide-class of stick-breaking processes as marginals.
- DPAR has a “Chinese restaurant”-type representation which may be useful for non-MCMC estimation methods.
- The processes define jump processes on probability measures and are useful for finding change-points.