Available online at www.sciencedirect.com
ScienceDirect
European Journal of Operational Research 180 (2007) 282–291
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/ejor

Stochastics and Statistics

Full predictivistic modeling of stock market data: Application to change point problems

R.H. Loschi [a], P.L. Iglesias [b], R.B. Arellano-Valle [b], F.R.B. Cruz [a,*]

[a] *Department of Statistics, Federal University of Minas Gerais, 31270-901 – Belo Horizonte – MG, Brazil*
[b] *Departamento de Estadística, Facultad de Matematicas, Pontificia Universidad Católica de Chile, Santiago, Chile*

Received 28 June 2005; accepted 5 April 2006
Available online 5 June 2006

Abstract

In change point problems in general we should answer three questions: how many changes are there? Where are they? And, what is the distribution of the data within the blocks? In this paper, we develop a new full predictivistic approach for modeling observations within the same block of observation and consider the product partition model (PPM) for treating the change point problem. The PPM brings more flexibility into the change point problem because it considers the number of changes and the instants when the changes occurred as random variables. A full predictivistic characterization of the model can provide a more tractable way to elicit the prior distribution of the parameters of interest, once prior opinions will be required only about observable quantities. We also present an application to the problem of identifying multiple change points in the mean and variance of a stock market return time series.

Keywords: Uncertainty modeling; Normal-inverse-gamma distribution; Product partition model; Student-*t* distribution

1. Introduction

The student-*t* distribution is a class of model that can be obtained as a location and scale mixture of the normal distribution in which the mixing measure is the normal-inverse-gamma distribution. Consequently, the student-*t* distribution can be constructed in two stages. Firstly, given the location and scale parameters, a conditional normal distribution is considered. Secondly, a prior distribution for these location and scale parameters is specified.

In a predictivistic characterization of the process, the first stage is replaced by an invariance or sufficiency assumption over an infinite sequence of potentially observable random quantities. For the student-*t* model above mentioned the first stage is replaced by the orthogonal invariance, which preserves the vectors of ones (see, for example, the works of Smith (1981) and Diaconis et al. (1992), among others). However, this invariance condition is not enough to identify the mixing measure. Thus some extra conditions on observable random variables are necessary to obtain the mixing measure (see the papers by Diaconis and Ylvisaker (1979) and Arellano-Valle et al. (1994), for instance). In a full predictivistic approach, Loschi et al. (2003b)

* Corresponding author. Tel.: +55 31 3499 5929; fax: +55 31 3499 5924.
E-mail address: fcruz@est.ufmg.br (F.R.B. Cruz).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2006.04.016

stated the conditions to characterize such a student-t distribution.

This paper basically revisits previous contributions (Loschi et al., 2003a; Loschi et al., 2003b; Loschi et al., in press) used as a basis for an original predictivist justification to the successful methods developed in the past for multiple change point identification, but without any formal justification such as the one present here. Additionally, the paper shows how this apparently hard way of modeling could be applied to model the behavior of stock market returns and other similar data sequences. The multiple change point problem is treated here in the light of the well-known product partition model (PPM), proposed by Barry and Hartigan (1993). However, it is worthwhile noting that some other models also deal with the change point problem and, in some aspects, do it in a much more natural way, the so-called Dynamic Linear Model, which can be seen in the book by West and Harrison (1999), along with other classes of dynamic models in Bayesian time series analysis and forecasting.

Some of the advantages of the PPM are that it allows the identification of multiple change points in the parameters, as well as in the functional form of the distribution function itself. Besides, the PPM brings flexibility into the analysis because the number of change points is a random variable, unlike threshold models (Chen and Lee, 1995) and others (Hawkins, 2001; Zmeškal, 2005) that consider the number of change points fixed. Firstly, Barry and Hartigan (1993) applied the PPM to the identification of multiple change points in the means of normal random variables with common variances. Afterwards, Crowley (1997) provided a new implementation of a Gibbs sampling scheme for the PPM in order to estimate the normal means, which was extended by Loschi et al. (2003a) to estimate both the means and the variances. Extensions of the PPM to a more general context can also be found in the papers by Quintana and Iglesias (2003) and Loschi et al. (in press).

In order to illustrate the methodology, we apply the predictivistic model developed to identify multiple change points both in the means μ and variances σ^2 of normal data sequentially observed. A conjugate prior distribution is considered for the parameters, μ and σ^2, which is justified within a full predictivistic setting. In fact, a more tractable way to elicit the prior distribution of μ and σ^2 is proposed, once opinions are required only about observable quantities. Yao's algorithm Yao, 1984 is used to compute the posterior estimates and a Gibbs sampling scheme is applied to estimate the posterior distributions of the number of change points and the instants when the changes occurred. The algorithms were implemented in C++ and the code is available from the authors upon request. The method was applied to identify multiple change points in the means and variances of a series of returns of the Chilean stock market. Different prior distributions were considered for the probability that a change occurs in any instant of the time and a sensitivity analysis was provided. As a result, it was seen that the returns in the Chilean stock market are characterized by changes in the expected returns (means) and volatilities (measured here as variances).

The paper is organized as follows. In Section 2, the PPM is briefly reviewed and a predictivistic characterization of the student-t PPM is provided, which explains in an alternative way the choices adopted for the prior distributions. In Section 3, the methodology is illustrated to the identification of change points in the mean return and volatility of ENDESA returns (the Chilean National Electricity Company). The return behavior within each block is modeled taking into account the predictivistic approach. A sensitivity analysis to the PPM is also provided. Finally, Section 4 closes the paper with final concluding remarks.

2. The student-t PPM

In this section we apply the PPM to identify multiple change points in the mean and variance of normal data observed sequentially through time. We consider a conjugate analysis and present a new full predictivistic characterization to the complete model (the likelihood function and prior distribution).

2.1. The product partition model (PPM)

Let $X_1, \ldots, X_n$ be a data sequence. Consider a random partition ρ of the set $I = \{1, \ldots, n\}$ and a random variable B that represents the number of blocks in ρ. Consider that each partition $\rho = \{i_0, i_1, \ldots, i_b\}$, $0 = i_0 < i_1 < \cdots < i_b = n$, divides the sequence $X_1, \ldots, X_n$ into $B = b$, $b \in I$, contiguous subsequences, which will be denoted by $\mathbf{X}_{[i_{(r-1)} i_r]} = (X_{i_{(r-1)}+1}, \ldots, X_{i_r})'$, $r = 1, \ldots, b$. Let $c_{[ij]}$ be the prior cohesion associated to the block