Change point estimation for the telegraph process observed at discrete times

Alessandro De Gregorio, Stefano M. Iacus*

Dipartimento di Scienze Economiche, Aziendali e Statistiche
Via Conservatorio 7, 20122 Milan - Italy

February 5, 2008

Abstract

The telegraph process models a random motion with finite velocity and it is usually proposed as an alternative to diffusion models. The process describes the position of a particle moving on the real line, alternatively with constant velocity $+v$ or $-v$. The changes of direction are governed by an homogeneous Poisson process with rate $\lambda > 0$. In this paper, we consider a change point estimation problem for the rate of the underlying Poisson process by means of least squares method. The consistency and the rate of convergence for the change point estimator are obtained and its asymptotic distribution is derived. Applications to real data are also presented.

Key words: discrete observations, change point problem, volatility regime switch, telegraph process.

---
email: alessandro.degegorio@unimi.it, stefano.iacus@unimi.it

1 Introduction

The telegraph process describes a random motion with finite velocity and it is usually proposed as an alternative to classical diffusion models (see Goldstein, 1951 and Kac, 1974). The process defines the position of a particle initially located at the origin of the real line and moving alternatively with constant velocity $+v$ or $-v$. The changes of direction are governed by an homogeneous Poisson process with rate $\lambda > 0$. The telegraph process or *telegrapher's* process is defined as

$$X(t) = V(0) \int_0^t (-1)^{N(s)} \mathrm{d}s, \quad t > 0, \tag{1.1}$$

where $V(0)$ is the initial velocity taking values $\pm v$ with equal probability and independently of the Poisson process $\{N(t), t > 0\}$. Many authors analyzed probabilistic properties of the process over the years (see for example Orsingher, 1990, 1995; Foong and Kanno, 1994; Stadje and Zacks, 2004; Zacks 2004). Di Crescenzo and Pellerey (2002) proposed the geometric telegraph process as a model to describe the dynamics of the price of risky assets where $X(t)$ replaces the standard Brownian motion of the original Black-Scholes-Merton model. Conversely to the geometric Brownian motion, given that $X(t)$ is of bounded variation, so is the geometric telegraph process. This seems a realistic way to model paths of assets in the financial markets. Mazza and Rulliere (2004) linked the process (1.1) and the ruin processes in the context of risk theory. Di Masi *et al.* (1994) proposed to model the volatility of financial markets in terms of the telegraph process. Ratanov (2004, 2005) proposed to model financial markets using a telegraph process with two intensities $\lambda_\pm$ and two velocities $v_\pm$. The telegraph process has also been used in ecology to model population dynamics (see Holmes *et al.*, 1994) and the displacement of wild animals on the soil. In particular, this model is chosen because it preserves the property of animals to move at finite velocity and for a certain period along one direction (see e.g. Holmes, 1993, for an account).

For the telegraph process $\{X(t), 0 \le t \le T\}$ observed at equidistant discrete times $0 = t_0 < \ldots < t_n$, with $t_i = i\Delta_n$, $i = 0, \ldots, n$, $n\Delta_n = T$ and $\Delta_n \to 0$ as $n \to \infty$, De Gregorio and Iacus (2006) proposed pseudo-maximum likelihood and implicit moment based estimators for the rate λ of the telegraph process. Under the additional condition $n\Delta_n \to \infty$ as $n \to \infty$, Iacus and Yoshida (2007) studied the asymptotic properties of explicit moment type estimators and further propose a consistent, asymptotically gaussian and asymptotically efficient estimator based on the increments of the process.

In this paper we suppose that for a telegraph process occurs a switch of the rate from λ_1 to λ_2 at some time instant $\theta_0 \in [0, T]$ and the interest is in the estimation of the change point θ_0 and both λ_1 and λ_2.

2

The change point estimation theory has been employed widely by means of the likelihood function (see Csörgő and Horváth, 1997). Unfortunately, the likelihood function for the telegraph process is not known and the pseudo likelihood proposed in De Gregorio and Iacus (2006) is not easy to treat in this framework. We will then proceed using the alternative method based on least squares proposed in Bai (1994, 1997) and used in different contexts by many authors including Hsu (1977, 1979) for the i.i.d. case and Chen *et al.* (2005) for the mixing case. Our model is peculiar in itself for the properties of $X(t)$ described in the above, because λ is a parameter related to the mean and the variance of the Poisson process and because the mesh Δ_n plays a role in the definition of the rate of convergence of our estimators.

The paper is organized as follows. Section 2 describes the model, the observation scheme and the change point estimator. The consistency of change point estimator is discussed in Section 3 and distributional results are presented in Section 4. Finally, Section 5 contains an application of our method to real data: we will reanalyze the classical change point data sets of Dow-Jones weekly closing (see Hsu, 1977, 1979) and IBM stock prices (see Box and Jenkins, 1970).

2   Observation scheme and least squares estimator

We consider a telegraph process $\{X(t), 0 \leq t \leq T < \infty\}$ defined as in (1.1) and assume to observe its trajectory only in $n + 1$ equidistant discrete times $0 < t_1 < \ldots < t_n$, with $t_i = i\Delta_n$, $i = 1, \ldots, n$ and $n\Delta_n = T$. We assume that a rate of changes of direction shift occurs during the interval $[0, T]$ at an unknown time $\theta_0 = \tau T$, $\tau \in (0, 1)$. Therefore the changes of direction are governed by an inhomogeneous Poisson process with parameter $\lambda(t) = \lambda_1 \mathbf{1}_{\{t \leq \theta_0\}} + \lambda_2 \mathbf{1}_{\{t > \theta_0\}}$ where the positive values λ_1, λ_2 and the change point τ (or θ_0) are unknown and to be estimated given the observations $X(t_1), X(t_2), \ldots, X(t_n)$. In order to simplify the formulas we use the following notation: $X(t_i) = X(i\Delta_n) = X_i$. The asymptotic framework is the following: $\Delta_n \to 0$ and $n\Delta_n = T \to \infty$ as $n \to \infty$.

The telegraph process is not Markovian and, as mentioned in the Introduction, it is not possible to derive the explicit likelihood function of the observations X_i's, therefore we can not apply the statistical methods based on the likelihood function. To work out our estimation problem, we shall follow the approach developed in Bai (1994), which involves least squares-type estimators. The same point of view has been applied by Chen *et al.*, 2005, in a context of financial time series. For our model, the time increment Δ_n plays an active role in the study of the asymptotics of our estimators so the proofs, although in some cases along the lines of Bai (1994) require some technical, but crucial, adjustments.

In order to obtain our estimator we introduce some basic notations. Let

$$Y_i = \frac{\mathbf{1}_{\{|\eta_i| < v\Delta_n\}}}{\Delta_n}, \, i = 1, ..., n,$$

where $\eta_i = X_i - X_{i-1}$ is the increment between two consecutive observations. We indicate the mean value of Y_i with $\gamma_m = \frac{1 - e^{-\lambda_m \Delta_n}}{\Delta_n} = \lambda_m + o(1)$, $m = 1, 2$. We observe that the random variables Y_i are independent and identically distributed because depend by the increments $|\eta_i|$. Iacus and Yoshida (2007) proved that the estimators

$$\hat{\gamma}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i \qquad \text{and} \qquad \hat{\lambda}_n = -\frac{1}{\Delta_n} \log(1 - \hat{\gamma}_n \Delta_n)$$

are consistent, gaussian and asymptotically efficient estimators of γ and λ respectively. We will use these properties in the following without necessarily mentioning them.

We assume that the change occurs exactly at time $t_i = t_{k_0} = k_0 \Delta_n = \theta_0$, therefore $k_0 = [n\tau]$, $\tau \in (0, 1)$, where $[\cdot]$ represents the integer-valued function. The change point estimator is obtained as follows

$$
\begin{aligned}
\hat{k}_0 &= \arg\min_k \left(\min_{\gamma_1, \gamma_2} \left\{ \sum_{i=1}^{k} (Y_i - \gamma_1)^2 + \sum_{i=k+1}^{n} (Y_i - \gamma_2)^2 \right\} \right) \\
&= \arg\min_k \left\{ \sum_{i=1}^{k} (Y_i - \bar{Y}_k)^2 + \sum_{i=k+1}^{n} (Y_i - \bar{Y}_{k+1})^2 \right\},
\end{aligned}
\tag{2.1}
$$

where

$$\min_{\gamma_1} \sum_{i=1}^{k} (Y_i - \gamma_1)^2 = \frac{1}{k} \sum_{i=1}^{k} Y_i = \bar{Y}_k,$$

$$\min_{\gamma_2} \sum_{i=k+1}^{n} (Y_i - \gamma_2)^2 = \frac{1}{n-k} \sum_{i=k+1}^{n} Y_i = \bar{Y}_{n-k}.$$

We indicate the sum of the squares of residuals in the following manner

$$U_k^2 = \sum_{i=1}^{k} (Y_i - \bar{Y}_k)^2 + \sum_{i=k+1}^{n} (Y_i - \bar{Y}_{k+1})^2,
\tag{2.2}$$

then

$$\hat{k}_0 = \arg\min_k U_k^2,
\tag{2.3}$$

4

and $\hat{\gamma}_1 = \bar{Y}_{\hat{k}_0}$, $\hat{\gamma}_2 = \bar{Y}_{n-\hat{k}_0}$ are respectively the least squares estimators of γ_1 and γ_2. This gives the two estimators

$$\hat{\lambda}_1 = -\frac{1}{\Delta_n}\log(1 - \hat{\gamma}_1\Delta_n), \qquad \hat{\lambda}_2 = -\frac{1}{\Delta_n}\log(1 - \hat{\gamma}_2\Delta_n). \qquad (2.4)$$

By setting $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$, $S_n = \sum_{i=1}^n Y_i$, simple algebra leads to

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = U_k^2 + nV_k^2, \quad 1 \le k \le n - 1,$$

where

$$V_k \;=\; \sqrt{\frac{k(n-k)}{n^2}}(\bar{Y}_{n-k} - \bar{Y}_k) = \frac{1}{\sqrt{k(n-k)}}S_n D_k, \qquad (2.5)$$

and

$$D_k = \frac{k}{n} - \frac{S_k}{S_n}. \qquad (2.6)$$

Therefore, formula (2.3) implies that

$$\hat{k}_0 = \arg\max_k V_k^2 = \arg\max_k |D_k| = \arg\max_k \sqrt{k(n-k)}|V_k|. \qquad (2.7)$$

Our first result concerns the asymptotic distribution of $\hat{k}_0$ under the null hypothesis that $\lambda_1 = \lambda_2$. This permits us to test if a shift has taken place during the interval $[0, T]$.

Theorem 2.1. *Under H_0, i.e. $\lambda_2 = \lambda_1 = \lambda_0$, we have that for $\Delta_n \to 0$, $n\Delta_n \to \infty$ as $n \to \infty$ the following result hods*

$$\sqrt{\frac{n\Delta_n}{\lambda_0}}|D_k| \xrightarrow{d} |B^0(t)|, \qquad (2.8)$$

where $\{B^0(t), 0 \le t \le 1\}$ is a Brownian bridge.

Proof. Let $\xi_i = Y_i - \gamma_0$, $\gamma_0 = EY_i = \frac{1-e^{-\lambda_0\Delta_n}}{\Delta_n} = \lambda_0 + o(1)$. Then, $E\xi_i = 0$ and $\sigma_n^2 = Var(\xi_i) = \frac{(1-e^{-\lambda_0\Delta_n})e^{-\lambda_0\Delta_n}}{\Delta_n^2}$. We introduce the following function

$$X_n(t) = \frac{1}{\sigma_n\sqrt{n}}\mathcal{S}_{[nt]} + (nt - [nt])\frac{1}{\sigma_n\sqrt{n}}\xi_{[nt]+1}, \quad 0 < t < 1,$$

with $\mathcal{S}_n = \sum_{i=1}^n \xi_i$. We note that

$$\left| \frac{1}{\sigma_n\sqrt{n}}\sum_{i=1}^{[nt]}\xi_i - \frac{\sqrt{t}}{\sigma_n\sqrt{[nt]}}\sum_{i=1}^{[nt]}\xi_i \right| \xrightarrow{p} 0, \qquad (2.9)$$

5

and

$$Var\left(\frac{\sqrt{t}}{\sigma_n\sqrt{[nt]}}\sum_{i=1}^{[nt]}\xi_i\right) = Var\left(\frac{\sqrt{t}}{\sigma_n\sqrt{[nt]\Delta_n}}\sum_{i=1}^{[nt]}\Delta_n\xi_i\right) \quad (2.10)$$

$$= Var\left(\frac{\sqrt{t}}{\sigma_n\sqrt{[nt]\Delta_n}}\sum_{i=1}^{[nt]}\frac{\mathbf{1}_{\{|\eta_i|<v\Delta_n\}} - (1-e^{-\lambda_0\Delta_n})}{\sqrt{\Delta_n}}\right)$$

$$= t$$

Since $\left|\frac{\mathbf{1}_{\{|\eta_i|<v\Delta_n\}} - (1-e^{\lambda_0\Delta_n})}{\sqrt{\Delta_n}}\right| < \frac{1}{\sqrt{\Delta_n}}$ the Lindeberg condition is true

$$\sum_{i=1}^{[nt]}\frac{E\left\{\mathbf{1}_{\{\sqrt{\Delta_n}|\xi_i|\geq\varepsilon\sqrt{n\Delta_n}\sigma_n\}}\xi_i^2\right\}}{\Delta_n\sigma_n^2([nt])^2} \to 0. \quad (2.11)$$

Then from (2.9), (2.10) and (2.11) we can conclude that

$$\frac{1}{\sigma_n\sqrt{n}}\mathcal{S}_{[nt]} \xrightarrow{d} N(0,t). \quad (2.12)$$

Now, by applying Donsker's theorem (invariance principle) we are able to write that

$$X_n(t) \xrightarrow{d} B(t)$$

$$\{X_n(t) - tX_n(1)\} \xrightarrow{d} B^0(t),$$

with $B(t)$ and $B^0(t)$ representing respectively a standard Brownian motion and a Brownian bridge. Let $k = [nt]$, we can write

$$X_n(t) - tX_n(1) = \frac{1}{\sigma_n\sqrt{n}}\left[\mathcal{S}_k - \frac{k}{n}\mathcal{S}_n\right] + \frac{nt-[nt]}{\sigma_n\sqrt{n}}\xi_{[nt]+1}$$

$$= \frac{1}{\sigma_n\sqrt{n}}\left[\sum_{i=1}^{k}(Y_i-\gamma_0) - \frac{k}{n}\sum_{i=1}^{n}(Y_i-\gamma_0)\right]$$

$$+ \frac{nt-[nt]}{\sigma_n\sqrt{n}}\xi_{[nt]+1}.$$

We observe that

$$\sum_{i=1}^{k}(Y_i-\gamma_0) - \frac{k}{n}\sum_{i=1}^{n}(Y_i-\gamma_0) = -D_k\sum_{i=1}^{n}Y_i,$$

and consequently

$$\sqrt{n\Delta_n}|D_k|\frac{\sum_{i=1}^{n}Y_i}{n\sqrt{\frac{(1-e^{\lambda_0\Delta_n})e^{\lambda_0\Delta_n}}{\Delta_n}}} = \left|X_n(t) - tX_n(1) - \frac{nt-[nt]}{\sigma_n\sqrt{n}}\xi_{[nt]+1}\right|. \quad (2.13)$$

It is easy to see (by Chebyshev inequality) that

$$\sup_t \left| \frac{nt - [nt]}{\sigma_n \sqrt{n}} \xi_{[nt]+1} \right| \xrightarrow{p} 0.$$

By the law of large number $\frac{1}{n} \sum_{i=1}^n Y_i \xrightarrow{p} \lambda_0$, while $\sqrt{\frac{(1-e^{\lambda_0 \Delta_n})e^{\lambda_0 \Delta_n}}{\Delta_n}} \to \sqrt{\lambda_0}$. Therefore from (2.13) follows that

$$\sqrt{\frac{n\Delta_n}{\lambda_0}} |D_k| \xrightarrow{d} |B^0(t)|.$$

$\square$

Corollary 2.1. *The same convergence result of the Theorem 2.1 follows when we consider*

$$\frac{\sqrt{n\Delta_n}}{\sqrt{\tilde{\lambda}_0}} |D_k|,$$

where $\tilde{\lambda}_0$ is any consistent estimator for λ_0.

Remark 2.1. *From Theorem 2.1 we derive immediately that for $\delta \in (0, 1/2)$*

$$\sqrt{\frac{n\Delta_n}{\lambda_0}} \sup_{\delta n \leq k \leq (1-\delta)n} |D_k| \xrightarrow{d} \sup_{\delta \leq t \leq (1-\delta)} |B^0(t)|, \qquad (2.14)$$

$$\sqrt{\frac{n\Delta_n}{\lambda_0}} \sup_{\delta n \leq k \leq (1-\delta)n} |V_k| \xrightarrow{d} \sup_{\delta \leq t \leq (1-\delta)} (t(1-t))^{-1/2} |B^0(t)|. \qquad (2.15)$$

The last asymptotic results are useful to test if doesn't exist a change point. In particular it is possible to obtain the asymptotic critical values for the distribution (2.15) by means of the same arguments used in Csörgő and Horváth (1997), pag. 25.

3   The consistency properties of the estimator

We shall study the consistency and the rate of convergence of the change point estimator (2.7). It is convenient to note that the rate of convergence is particularly important not only to describe how fast the estimator converges to the true value, but also to get the limiting distribution. The next Theorem represents our first result on the consistency.

Theorem 3.1. *The estimator $\hat{\tau} = \frac{\hat{k}_0}{n}$ satisfies*

$$|\hat{\tau} - \tau| = (n\Delta_n)^{-1/2}(\gamma_2 - \gamma_1)^{-1} O_p(\sqrt{\log n}) \qquad (3.1)$$

7

Proof. By the same arguments of Bai (1994), Section 3 and by using the formulas (10)-(14) therein, we have that

$$|\hat{\tau} - \tau| \leq C_\tau (\gamma_2 - \gamma_1)^{-1} \sup_k |V_k - EV_k|, \qquad (3.2)$$

where C_τ is a constant depending only on τ. Let $Z_i = \frac{\mathbf{1}_{\{|\eta_i| < v\Delta_n\}} - (1 - e^{\lambda\Delta_n})}{\sqrt{\Delta_n}}$, given that

$$
\begin{aligned}
V_k - EV_k &= \frac{1}{\sqrt{n\Delta_n}} \sqrt{\frac{k}{n}} \frac{1}{\sqrt{n-k}} \sum_{i=k+1}^{n} Z_i \\
&\quad + \frac{1}{\sqrt{n\Delta_n}} \sqrt{1 - \frac{k}{n}} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Z_i
\end{aligned}
$$

we obtain that

$$|V_k - EV_k| \leq \frac{1}{\sqrt{n\Delta_n}} \left\{ \bar{Z}_{n-k} + \bar{Z}_k \right\}. \qquad (3.3)$$

where $\bar{Z}_k = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Z_i$ and $\bar{Z}_{n-k} = \frac{1}{\sqrt{(n-k)}} \sum_{i=k+1}^{n} Z_i$. By applying Hajék-Renyi inequality for martingales we have that

$$
\begin{aligned}
P\left\{ \max_{1 \leq k \leq n} \left| \frac{\sum_{i=1}^{k} Z_i}{c_k} \right| > \alpha \right\} &\leq \frac{1}{\alpha^2} \sum_{k=1}^{n} \frac{E(Z_k)^2}{c_k^2} \\
&= \frac{(1 - e^{-\lambda\Delta_n})e^{-\lambda\Delta_n}}{\alpha^2 \Delta_n} \sum_{k=1}^{n} \frac{1}{c_k^2} \\
&\leq \frac{\lambda\Delta_n + o(\Delta_n)}{\alpha^2 \Delta_n} \sum_{k=1}^{n} \frac{1}{c_k^2} \\
&= \frac{\lambda + o(1)}{\alpha^2} \sum_{k=1}^{n} \frac{1}{c_k^2} \qquad (3.4)
\end{aligned}
$$

Choosing $c_k = \sqrt{k}$ and observing that $\sum_{k=1}^{n} k^{-1} \leq C \log n$, for some $C > 0$ (see e.g. Bai, 1994), we have that

$$\max_{1 \leq k \leq n} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Z_i = O_p\left(\sqrt{\log n} \right). \qquad (3.5)$$

Then from the relationships (3.3) and (3.5) we obtain the result (3.1). $\qquad \square$

Remark 3.1. *By means of the law of iterated logarithm we obtain immediately the following rate of convergence which improve the previous result. We have that*

$$|\hat{\tau} - \tau| = (n\Delta_n)^{-1/2}(\gamma_2 - \gamma_1)^{-1} O_p(\sqrt{\log \log n}). \qquad (3.6)$$

8

Remark 3.2. *Theorem 3.1 implies that, under the additional hypothesis* $\Delta_n = O(n^\varepsilon)$, $\varepsilon \in (-1, 0)$ *we have also consistency, i.e.* $(n\Delta_n)^\beta(\hat\tau - \tau) \to 0$ *in probability for any* $\beta \in (0, 1/2)$.

We are able to improve the rate of convergence of $\hat\tau$.

Theorem 3.2. *We have the following result*

$$\hat\tau - \tau = O_p\left(\frac{1}{n\Delta_n(\gamma_2 - \gamma_1)^2}\right). \tag{3.7}$$

Proof. We use the same framework of the proof of the Proposition 3 in Bai (1994), Section 4, therefore we omit the details.

We choose a $\delta > 0$ such that $\tau \in (\delta, 1 - \delta)$. Since $\hat k/n$ is consistent for τ, for every $\varepsilon > 0$, $Pr\{\hat k/n \notin (\delta, 1 - \delta)\} < \varepsilon$ when n is large. In order to prove (3.7) it is sufficient to show that $Pr\{|\hat\tau - \tau| > M(n\Delta_n\gamma_n^2)^{-1}\}$ is small when n and M are large, where $\gamma_n = \gamma_2 - \gamma_1$. We are interested to study the behavior of V_k for $n\delta \le k \le n(1 - \delta)$, $0 < \delta < 1$. We define for any $M > 0$ the set $D_{n,M} = \{k : n\delta \le k \le n(1 - \delta), |k - k_0| > M\Delta_n^{-1}\gamma_n^{-2}\}$. Then we have that

$$Pr\{|\hat\tau - \tau| > M(n\Delta_n\gamma_n^2)^{-1}\} \le \varepsilon + Pr\{\sup_{k \in D_{n,M}} |V_k| \ge |V_{k_0}|\},$$

for every $\varepsilon > 0$. Thus we study the behavior of $Pr\{\sup_{k \in D_{n,M}} |V_k| \ge |V_{k_0}|\}$. It is possible to prove that

$$
\begin{aligned}
Pr\left\{\sup_{k \in D_{n,M}} |V_k| \ge |V_{k_0}|\right\} \le{}& Pr\left\{\sup_{k \in D_{n,M}} V_k - V_{k_0} \ge 0\right\} \\
&+ Pr\left\{\sup_{k \in D_{n,M}} V_k + V_{k_0} \le 0\right\} \\
={}& P + Q
\end{aligned}
\tag{3.8}
$$

Furthermore

$$
\begin{aligned}
Q \le{}& 2Pr\left\{\sup_{k \le n(1-\delta)} \frac{1}{n-k}\left|\sum_{i=k+1}^{n}(Y_i - \gamma_2)\right| \ge \frac{1}{4}EV_{k_0}\right\} \\
&+ 2Pr\left\{\sup_{k \ge n\delta} \frac{1}{k}\left|\sum_{i=1}^{k}(Y_i - \gamma_1)\right| \ge \frac{1}{4}EV_{k_0}\right\}.
\end{aligned}
\tag{3.9}
$$

By observing that $\sum_{i=m}^\infty i^{-2} = O(m^{-1})$, the Hajék-Renyi inequality yields

$$P\left\{\max_{k \ge m} \left|\frac{1}{k}\sum_{i=1}^{k} Z_i\right| > \alpha\right\} \le \frac{1}{\alpha^2 m}\frac{(1 - e^{\lambda\Delta_n})e^{-\lambda\Delta_n}}{\sqrt{\Delta_n}}, \tag{3.10}$$

9

where r.v.'s Z_i are defined in the proof of Theorem 3.1. The inequality (3.10) implies that (3.9) tends to zero as n tends to infinity. Let $b(k) = \sqrt{((k/n)(1 - k/n))}, k = 1, 2, ..., n,$ for the first term in the right-hand of (3.8) we have that

$$
\begin{aligned}
P \leq \quad & Pr \left\{ \sup_{k \in D_{n,M}} \frac{n}{|k_0 - k|} |G(k)| > \frac{\gamma_n C_\tau}{2} \right\} \\
& + Pr \left\{ \sup_{k \in D_{n,M}} \frac{n}{|k_0 - k|} |H(k)| > \frac{\gamma_n C_\tau}{2} \right\} \\
= \quad & P_1 + P_2,
\end{aligned}
\tag{3.11}
$$

where

$$
G(k) = b(k_0) \frac{1}{k_0} \sum_{i=1}^{k_0} (Y_i - \gamma_1) - b(k) \frac{1}{k} \sum_{i=1}^{k} (Y_i - \gamma_1)
\tag{3.12}
$$

$$
H(k) = b(k) \frac{1}{n-k} \sum_{i=k+1}^{n} (Y_i - \gamma_2) - b(k_0) \frac{1}{n-k_0} \sum_{i=k_0+1}^{n} (Y_i - \gamma_2)
\tag{3.13}
$$

We prove that P_1 tends to zero when n and M are large. Thus we consider only $k \leq k_0$ or more precisely those values of k such that $n\delta \leq k \leq n\tau - M\Delta_n^{-1}\gamma_n^{-2}$. For $k \geq n\delta$, we have

$$
|G(k)| \leq \frac{k_0 - k}{n\delta k_0} \left| \sum_{i=1}^{k_0} (Y_i - \gamma_1) \right| + B\frac{k_0 - k}{n} \frac{1}{n\delta} \left| \sum_{i=1}^{k} (Y_i - \gamma_1) \right| + \frac{1}{n\delta} \left| \sum_{i=k+1}^{k_0} (Y_i - \gamma_1) \right|,
\tag{3.14}
$$

where $B \geq 0$ satisfies $|b(k_0) - b(k)| \leq B|k_0 - k|/n$. By means of (3.4), (3.10) and (3.14), we obtain

$$
\begin{aligned}
P_1 \leq \quad & Pr \left\{ \frac{1}{n\tau} \left| \sum_{i=1}^{[n\tau]} (Y_i - \gamma_1) \right| > \frac{\delta\gamma_n C_\tau}{6} \right\} \\
& + Pr \left\{ \sup_{1 \leq k \leq n} \frac{1}{n} \left| \sum_{i=1}^{k} (Y_i - \gamma_1) \right| > \frac{\delta\gamma_n C_\tau}{6B} \right\} \\
& + Pr \left\{ \sup_{k \leq n\tau - M\Delta_n^{-1}\gamma_n^{-2}} \frac{1}{n\tau - k} \left| \sum_{i=k+1}^{[n\tau]} (Y_i - \gamma_1) \right| > \frac{\delta\gamma_n C_\tau}{6} \right\} \\
\leq \quad & \frac{36D}{(\delta C_\tau)^2 \tau n \Delta_n \gamma_n^2} + \frac{36B^2 D}{(\delta C_\tau)^2 n \Delta_n \gamma_n^2} + \frac{36D}{\delta C_\tau^2 M}
\end{aligned}
$$

where $D = \frac{(1 - e^{-\lambda\Delta_n})e^{-\lambda\Delta_n}}{\Delta_n} \leq \lambda + o(1)$. When n and M are large the last three terms are negligible. Analogously we derive the proof of P_2.

$\square$

4 Asymptotic distributions

We want to study in this Section the asymptotic distribution of $\hat{\tau}$ under our limiting framework for small variations of the rate of change of the direction. The case $\lambda_n = \lambda_2 - \lambda_1$ equal to a constant is less interesting because when λ_n is large the estimate of k_0 is quite precise.

We note that $\lambda_n = \lambda_2 - \lambda_1 \to 0$ implies $\gamma_n = \gamma_2 - \gamma_1 \to 0$. By adding the condition

$$\lambda_n \to 0, \quad \frac{\sqrt{n\Delta_n}\gamma_n}{\sqrt{\log n}} \to \infty, \tag{4.1}$$

the consistency of $\hat{\tau}$ follows immediately either from Theorem 3.1 or Theorem 3.2. In order to obtain the main result of this Section, it is useful to observe that

$$\hat{k}_0 = \arg\max_k V_k^2 = \arg\max_k n\Delta_n(V_k^2 - V_{k_0}^2) \tag{4.2}$$

and to define a two-sided Brownian motion $W(v)$ in the following manner

$$W(u) = \begin{cases} W_1(-u), & u < 0 \\ W_2(u), & u \geq 0 \end{cases} \tag{4.3}$$

where W_1, W_2 are two independent Brownian motions. Now we present the following convergence in distribution result.

Theorem 4.1. *Under assumption* (4.1)*, for $n\Delta_n \to \infty, \Delta_n \to 0$ as $n \to \infty$, we have that*

$$\frac{n\Delta_n\gamma_n^2(\hat{\tau} - \tau)}{\tilde{\lambda}} \xrightarrow{d} \arg\max_v \left\{ W(v) - \frac{|v|}{2} \right\}, \tag{4.4}$$

where $W(v)$ is a two-sided Brownian motion and $\tilde{\lambda}$ is any consistent estimator for λ_1 or λ_2.

Proof. The proof follows the same steps in Bai (1994), Theorem 1, hence we only sketch the parts of the proof that differ. We consider only $v \leq 0$ because of symmetry. Let $K_n(v) = \{k : k = [k_0 + v\Delta_n^{-1}\gamma_n^{-2}], -M \leq v \leq 0, M > 0\}$ and

$$\Lambda_n(v) = n\Delta_n(V_k^2 - V_{k_0}^2) \tag{4.5}$$

with $k \in K_n(v)$. We note that

$$\begin{aligned} n\Delta_n(V_k^2 - V_{k_0}^2) &= 2n\Delta_n EV_{k_0}(V_k - V_{k_0}) \\ &\quad + 2n\Delta_n(V_{k_0} - EV_{k_0})(V_k - V_{k_0}) \\ &\quad + n\Delta_n(V_k - V_{k_0})^2 \end{aligned} \tag{4.6}$$

11

The last two terms in (4.6) are negligible on $K_n(v)$. Since $\sqrt{n\Delta_n}(V_{k_0} - EV_{k_0})$ is bounded by (3.2), we have to show that $\sqrt{n\Delta_n}|V_k - V_{k_0}|$ is bounded. In particular, we can write

$$\sqrt{n\Delta_n}|V_k - V_{k_0}| \leq \sqrt{n\Delta_n}|G(k) + H(k)| + \sqrt{n\Delta_n}|EV_k - EV_{k_0}|,$$

where $G(k)$ and $H(k)$ are defined respectively in (3.12) and (3.13). The upper bound (3.14) is $o_p(1)$, because the first term is such that

$$\begin{aligned}
\sqrt{n\Delta_n}\frac{k_0 - k}{n\delta k_0}\left|\sum_{i=1}^{k_0}(Y_i - \gamma_1)\right| &\leq \frac{M}{\delta\tau n\Delta_n\gamma_n^2}\left|\sum_{i=1}^{k_0}\sqrt{\Delta_n}(Y_i - \gamma_1)\right| \\
&= \frac{O_p(1)}{n\Delta_n\gamma_n^2} = o_p(1),
\end{aligned} \tag{4.7}$$

similarly for the second term and for the third term we apply the invariance principle (2.12). Now we explicit the limiting distribution for

$$2n\Delta_n EV_{k_0}(V_k - V_{k_0}) = 2\sqrt{\tau(1-\tau)}n\Delta_n\gamma_n(V_{[k_0+v\Delta_n^{-1}\lambda_n^{-2}]} - V_{k_0}). \tag{4.8}$$

For simplicity we shall assume that $k_0 + v\Delta_n^{-1}\gamma_n^{-2}$ and $v\Delta_n^{-1}\gamma_n^{-2}$ are integers. We observe that

$$n\Delta_n\gamma_n(V_k - V_{k_0}) = n\Delta_n\gamma_n(G(k) + H(k)) - n\Delta_n\gamma_n(EV_{k_0} - EV_k), \tag{4.9}$$

where $G(k), H(k)$ are defined in the expressions (3.12), (3.13). We can rewrite $G(k)$ as follws

$$\begin{aligned}
G(k) = \quad &b(k_0)\frac{k - k_0}{kk_0}\sum_{i=1}^{k_0}(Y_i - \gamma_1) + \frac{b(k_0) - b(k)}{k}\sum_{i=1}^{k}(Y_i - \gamma_1) \\
&+ b(k_0)\frac{1}{k}\sum_{i=k+1}^{k_0}(Y_i - \gamma_1).
\end{aligned} \tag{4.10}$$

By the same arguments used to prove (4.7) we can show that the first two terms in (4.10) multiplied by $n\Delta_n\gamma_n$ are negligible on $K_n(M)$. Furthermore $b(k_0) =$

$\sqrt{\tau(1-\tau)}$ and $n/k \to 1/\tau$ for $k \in K_n(M)$, then we get that

$$
\begin{aligned}
n\Delta_n\gamma_n G(k_0 \ + \ v\Delta_n^{-1}\gamma_n^{-2}) &= n\Delta_n\gamma_n b(k_0)\frac{1}{k}\sum_{i=k+1}^{k_0}(Y_i - \gamma_1) + o_p(1) \\
&= b(k_0)\frac{n}{k}\left\{\gamma_n\sqrt{\Delta_n}\sum_{i=k+1}^{k_0}\sqrt{\Delta_n}(Y_i - \gamma_1)\right\} + o_p(1) \\
&= b(k_0)\frac{n}{k}\left\{\gamma_n\sqrt{\Delta_n}\sum_{i=1}^{|v|\Delta_n^{-1}\gamma_n^{-2}}\sqrt{\Delta_n}(Y_{i+k} - \gamma_1)\right\} + o_p(1) \\
&\xrightarrow{d} \frac{\sqrt{(1-\tau)\tau}}{\tau}\sqrt{\lambda_1}W_1(-v) \quad\quad\quad\quad\quad\quad\quad (4.11)
\end{aligned}
$$

where in the last step we have used the invariance principle (2.12). Analogously we show that

$$
n\Delta_n\gamma_n H(k_0 + v\Delta_n^{-1}\gamma_n^{-2}) \xrightarrow{d} \frac{\sqrt{(1-\tau)\tau}}{1-\tau}\sqrt{\lambda_1}W_1(-v). \quad\quad (4.12)
$$

Since

$$
n\Delta_n\gamma_n(EV_{k_0} - EV_k) \to \frac{|v|}{1\sqrt{\tau(1-\tau)}} \quad\quad\quad\quad\quad (4.13)
$$

we obtain that

$$
\Lambda_n(v) \xrightarrow{d} 2\left\{\sqrt{\lambda_1}W_1(-v) - \frac{|v|}{2}\right\}. \quad\quad\quad\quad\quad (4.14)
$$

In the same way, for $v > 0$, we can prove that

$$
\Lambda_n(v) \xrightarrow{d} 2\left\{\sqrt{\lambda_1}W_2(v) - \frac{|v|}{2}\right\}. \quad\quad\quad\quad\quad (4.15)
$$

By applying the continuous mapping theorem and Theorem 3.2.

$$
\frac{n\Delta_n\gamma_n^2(\hat{\tau} - \tau)}{\hat{\lambda}} \xrightarrow{d} \frac{1}{\lambda_1}\arg\max_v \Lambda_n(v). \quad\quad\quad\quad (4.16)
$$

Since $aW(v) \overset{d}{=} W(a^2 v), a \in \mathbb{R}$, a change in variable transforms $\arg\max_v \Lambda_n(v)$ into $\lambda_1 \arg\max_v\left\{W(v) - \frac{|v|}{2}\right\}$, which concludes the proof. $\qquad\square$

Using the consistency result, we are able to obtain the asymptotic distributions for the estimators $\hat{\lambda}_1, \hat{\lambda}_2$, defined in (2.4).

Theorem 4.2. *Under the assumption* (4.1) *we have that*

$$\sqrt{n\Delta_n} \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix} \xrightarrow{d} N(0, \Sigma), \tag{4.17}$$

where

$$\Sigma = \begin{pmatrix} \tau^{-1}\lambda_1 & 0 \\ 0 & (1-\tau)^{-1}\lambda_2 \end{pmatrix}. \tag{4.18}$$

Proof. We start noticing that

$$\sqrt{n\Delta_n}(\hat{\gamma}_1(\hat{k}) - \hat{\gamma}_1(k_0)) \tag{4.19}$$

$$= \sqrt{n\Delta_n} \left(\frac{1}{\hat{k}} \sum_{i=1}^{\hat{k}} Y_i - \frac{1}{k_0} \sum_{i=1}^{k_0} Y_i \right)$$

$$= \mathbf{1}_{\{\hat{k} \le k_0\}} \left(\sqrt{n\Delta_n} \frac{k_0 - \hat{k}}{k_0 \hat{k}} \sum_{i=1}^{k_0} (Y_i - \gamma_1) - \sqrt{n\Delta_n} \frac{1}{\hat{k}} \sum_{i=\hat{k}}^{k_0} (Y_i - \gamma_1) \right)$$

$$+ \mathbf{1}_{\{\hat{k} > k_0\}} \left(\sqrt{n\Delta_n} \frac{k_0 - \hat{k}}{k_0 \hat{k}} \sum_{i=1}^{k_0} (Y_i - \gamma_1) + \sqrt{n\Delta_n} \frac{1}{\hat{k}} \sum_{i=k_0}^{\hat{k}} (Y_i - \gamma_2) \right.$$

$$\left. + \sqrt{n\Delta_n} \gamma_n \frac{\hat{k} - k_0}{\hat{k}} \right).$$

Since $k_0 = [\tau n]$, $\hat{k} = k_0 + O_p(\Delta_n^{-1}\gamma_n^{-2})$, and $n\Delta_n\gamma_n^2 \to \infty$, we have that (4.19) is $(\sqrt{n}\gamma_n)^{-1}O_p(1)$, which converges to zero in probability. Then $\hat{\lambda}_1(\hat{k}) = -\frac{1}{\Delta_n}\log(1 - \hat{\gamma}_1(\hat{k})\Delta_n)$ and $\hat{\lambda}_1(k_0) = -\frac{1}{\Delta_n}\log(1 - \hat{\gamma}_1(k_0)\Delta_n)$ have the same limiting distribution. Obviously the same result holds for $\hat{\lambda}_2$. By Theorem 4.1 in Iacus and Yoshida (2007), the convergence result (4.17) follows. $\qquad\square$

5 Application to real data

In this section we consider an application of our model to two well known real data sets. The first data set is about the Dow-Jones industrial average and the second one is the IBM stock prices. In both cases, the data mesh Δ_n is not close to zero, hence the asymptotics of our set up does not hold. Nevertheless, our findings seems to confirm the results of previous analyses.
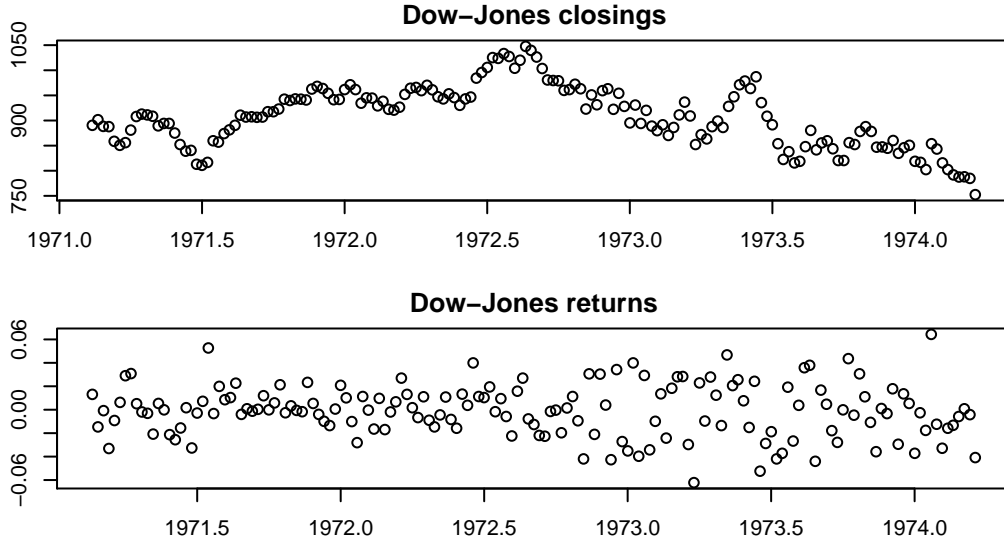
Figure 1: Weekly closings of the Dow-Jones industrial average July 1971 - Aug 1974 (up) and correspoind returns (down).

5.1 Dow-Jones data

This data set contains the weekly closings of the Dow-Jones industrial average in the period July 1971 - Aug 1974. These data have been proposed by Hsu (1977, 1979) and used by many other authors to test change point estimators. There are 162 data and the main evidence found by several authors is that a change in the variance occurred at point 89th which corresponds to the third week of March 1973. Instead of working on the values we transform the data into returns as usual $X(t_i) = (W(t_i) - W(t_{i-1}))/W(t_{i-1})$, $i = 1, \ldots, n$ with W the series of Dow-Jones closings and X the returns. We assume that X follows a telegraph process. In this application, the data are not sampled at high frequency, i.e. Δ_n is not close to zero, hence we test our estimator of the change point even if the asymptotics is not realized. Further, and for the same reason, we cannot assume as known the velocity v of the process hence we first estimate v by the average of the rescaled increments. i.e.

$$\hat{v}_n = \frac{1}{n} \sum_{i=1}^{n} \frac{\eta_i}{\Delta_n}$$

This is a consistent estimator of v, hence we construct the estimator of γ as follows

$$\hat{\gamma} = \frac{1}{n} \sum_{i=1}^{n} Y_i \qquad \text{where} \qquad Y_i = \frac{\mathbf{1}_{\{|\eta_i| < \hat{v}_n \Delta_n\}}}{\Delta_n}$$
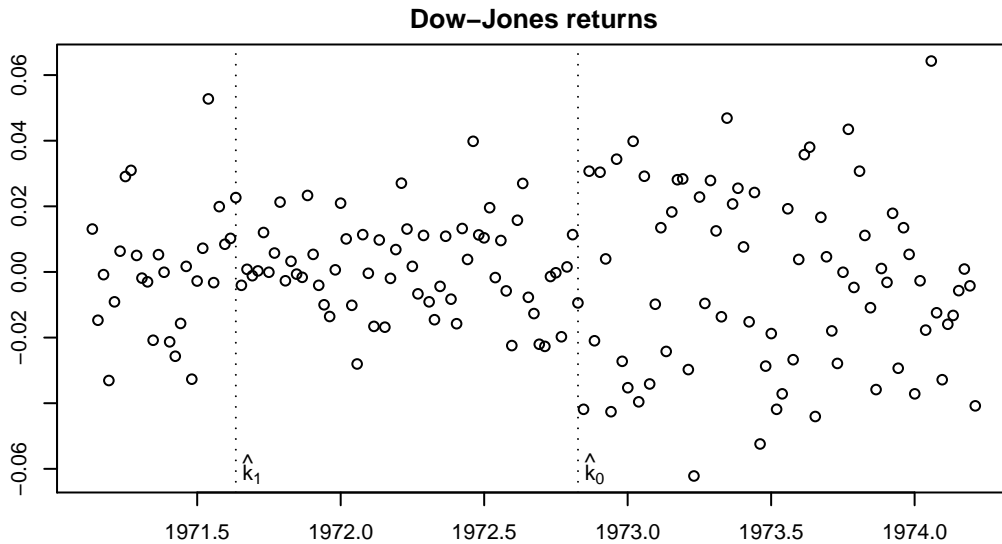
15

Dow–Jones returns

Figure 2: Change point estimates on the returns of the weekly closings of the Dow-Jones industrial average July 1971 - Aug 1974. Major change point estimate $\hat{k}_0 = 89$ which corresponds to the 3rd week of March 1973; second change point estimate $\hat{k}_1 = 27$, August 1971.

With these quantities, we construct the statistics D_k in (2.6) and maximize it. The maximum is reached at $\hat{k}_0 = 89$ which confirms the evidence in Hsu (1974, 1979). Once we obtained the estimation of the change point, we re-estimate the velocity in both part of the series (before and after point 89th) and the two lambda's. We obtained respectively $v_1 = 0.61$, $\lambda_1 = 48.53$ and $v_2 = 1.24$ and $\lambda_2 = 34.61$ which confirms the intuition from the graphical inspection of the returns (i.e. in the first period there is a high number of switches but with low velocity which correspond to low variance of the returns; conversely for the second period). Looking better at the first part of the series, we observe that variance is not stable, so we re-run the procedure and obtained a new change point $\hat{k}_1 = 27$ around august 1971. Figure 2 contains the two change point estimates plotted against the Dow-Jones returns.

5.2 IBM stock prices

This data set contains 369 closing stock prices of the IBM as,set. They have been analyzed in Box and Jenkins (1970) and further by Wichern *et al.* (1976) in order to discover change points. Box and Jenkins (1970) fitted an ARIMA(0,1,1) on the first order difference and discover heteroschedasticity; Wichern *et al.* (1976) fitted an AR(1) model on the first differences of the logarithms. We consider instead the

Figure 3: Return of the IBM stock closings (see e.g. Box and Jenkins, 1970). The major change point occurs at index $\hat{k}_0 = 235$, the other two at $\hat{k}_1 = 18$ and $\hat{k}_2 = 309$.

returns as in previous example and apply the same sequential procedure. Data are reported in Figure 3 along with a couple of change points discovered by our estimates. The first change point was found at point $\hat{k}_0 = 235$ which confirms the findings of Wichern *et al.* (1976). We further discovered another change point at time index $\hat{k}_1 = 18$ on the time series on the left to $\hat{k}_0$ and a second change point on the right-hand series at time $\hat{k}_2 = 309$.

References

[1] Bai, J. (1994) Least squares estimation of a shift in linear processes, *Journal of Times Series Analysis*, **15**, 453-472.

[2] Bai, J. (1997) Estimation of a change point in multiple regression models, *The Review of Economics and Statistics*, **79**, 551-563.

[3] Box, G.E.P., Jenkins, G.M. (1970) *Time series analysis: forecasting and control*, San Francisco, Holden-Day.

[4] Chen, G., Choi, Y.K., Zhou, Y. (2005) Nonparametric estimation of structural change points in volatility models for time series, *Journal of Econometrics*, **126**, 79-144.

[5] Csörgő, M., Horváth, L. (1997) *Limit Theorems in Change-point Analysis*, Wiley and Sons, New York.

[6] De Gregorio, A., Iacus, S.M. (2006) Parametric estimation for the telegraph process observed at discrete times. `http://services.bepress.com/unimi/statistics/art14`

[7] Di Crescenzo A, Pellerey F. (2002) On prices' evolutions based on geometric telegrapher's process, *Applied Stochastic Models in Bussiness and Industry*, **18**, 171-184.

[8] Di Masi, G.B, Kabanov, Y.M., Runggaldier, W.J. (1994) Mean-variance hedging of options on stocks with Markov volatilities, *Theory of Probability and its Applications*, **39**, 172-182.

[9] Fong S.K., Kanno, S. (1994) Properties of the telegrapher's random process with or without a trap, *Stochastic Processes and their Applications*, **53**, 147-173.

[10] Goldstein S. (1951) On diffusion by discontinuous movements and the telegraph equation, *The Quarterly Journal of Mechanics and Applied Mathematics*, **4**, 129-156.

[11] Holmes, E. E. (1993) Is diffusion too simple? Comparisons with a telegraph model of dispersal, *American Naturalist*, **142**, 779-796.

[12] Holmes, E. E., Lewis, M.A., Banks, J.E., Veit, R.R. (1994) Partial differential equations in ecology: spatial interactions and population dynamics, *Ecology*, **75**(1), 17-29.

[13] Hsu, D.A. (1977) Tests for variance shift at an unknown time point, *Appl. Statist.*, **26**(3), 279-284.

[14] Hsu, D.A. (1979) Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis, *Journal American Stat. Ass.*, **74**(365), 31-40.

[15] Iacus, S.M., Yoshida, N. (2007) Estimation for the telegraph process observed to discrete times, to appear in *Theory of Probability and Mathematical Statistics*.

[16] Inclan, C., Tiao, G.C. (1994) Use of cumulative sums of squares for retrospective detection of change of variance, *Journal of the American Statistical Association*, **89**, 913-923.

[17] Kac M. (1974) A stochastic model related to the telegrapher's equation, *Rocky Mountain Journal of Mathematics*, **4**, 497-509.

[18] Mazza C., Rulliére D. (2004) A link between wave governed random motions and ruin processes, *Insurance: Mathematics and Economics*, **35**, 205-222.

[19] Orsingher E. (1990) Probability law, flow function, maximun distribution of wave-governed random motions and their connections with Kirchoff's laws, *Stochastic Processes and their Applications*, **34**, 49-66.

[20] Orsingher E. (1995) Motions with reflecting and absorbing barriers driven by the telegraph equation, *Random Operators and Stochastic Equations*, **3**, 9-21.

[21] Ratanov, N. (2004) A Jump Telegraph Model for Option Pricing, to appear in *Quantitative Finance*.

[22] Ratanov, N. (2005) Quantile Hedging for Telegraph Markets and Its Applications To a Pricing of Equity-Linked Life Insurance Contracts. `http://www.urosario.edu.co/FASE1/economia/documentos/pdf/bi62.pdf`

[23] Stadje W., Zacks S. (2004) Telegraph processes with random velocities, *Journal of Applied Probability*, **41**, 665-678.

[24] Wichern, D.W., Miller, R.B., Hsu, D.A. (1976) Changes of Variance in first-order autoregressive time series models with an application, *Applied Stat.*, **25**(3), 248-256.

[25] Zacks (2004) Generalized integrated telegraph processes and the distribution of related stopping times, *Journal of Applied Probability*, **41**, 497-507.