Теорія Ймовір. та Матем. Статист.
Вип. 96, 2017

Theor. Probability and Math. Statist.
No. 96, 2018, Pages 45–57
https://doi.org/10.1090/tpms/1033
Article electronically published on October 5, 2018

A SEMI-MARTINGALE REPRESENTATION FOR A SEMI-MARKOV CHAIN WITH APPLICATION TO FINANCE

UDC 519.21

R. ELLIOTT, A. SWISHCHUK, AND I. Y. ZHANG

Dedicated to the 70th Anniversary of Professor Dmitrii Silvestrov

Abstract. In this paper we present the semi-martingale representation for a discrete time semi-Markov chain, and consider its application to a semi-Markov regime-switching binomial model in finance. We also introduce a semi-Markov switching Lévy process. Estimation results for Markov and semi-Markov chains are presented as well.

1. Introduction

Semi-Markov processes were first introduced by Lévy [17] in 1954. Essentially, semi-Markov processes generalize Markov jump processes by allowing holding times to be more generally distributed instead of being exponentially distributed. The theory of semi-Markov processes is based on the theory of Markov renewal processes. Smith [31] and Takacs [36] almost simultaneously developed semi-Markov processes in 1955 and 1956. The initial treatment of semi-Markov theory was given by Cinlar [7] and Pyke [24]–[27]. For an overview of hidden Markov chain processes and their financial applications see Elliott [9] and Swishchuk and Elliott [33]. For the general theories of semi-Markov processes, regime-switching, and Lévy processes, with applications in finance, see the textbooks and notes of Applebaum [2], Cohen and Elliott [6], Koroliuk and Limnios [16], Swishchuk and Wu [35], Schoutens [29], and Papapantoleon [23]. Discrete time semi-Markov random evolutions and their applications were considered in [18]. Semi-Markov processes with a discrete state space were studied in [30]. Semi-Markov switching processes in queueing systems were considered in [1].

Although the Black-Scholes formula has been quite successful in describing stock option prices, it does have well-known biases and its performance is substantially worse when pricing other derivatives. This is not surprising since the Black-Scholes model makes the strong assumption that stock returns are normally distributed with known mean and variance. The Black-Scholes formula does not depend on the mean spot return so it cannot be generalized by allowing this mean to vary. The assumption that the volatility, or instantaneous variance is constant appears wrong.

After Merton's [21] jump-diffusion models in 1976, generalized models to allow stochastic volatility were reported to be successful in explaining the prices of currency options by Melino and Turnbull [19, 20], Rumsey [28], as well as the stochastic volatility jump-diffusion models of Bates [3]. However, these papers have the disadvantage of not

2010 *Mathematics Subject Classification.* Primary 60K15, 60K10, 60G42, 60G51, 91B28.
Key words and phrases. Discrete time finite state semi-Markov chain, semi-Markov switching Lévy process, semi-martingale representation, financial derivatives, regime-switching binomial model.

having closed-form solutions and require extensive use of numerical techniques to solve two-dimensional partial differential equations.

Heston's model, developed in 1993 [14], not only allows volatility to follow a stochastic process, but the solution methods are faster than finite difference solutions to partial differential equations or integro-differential equations. This led Heston to refer to them as closed-form solutions. The famous Heston model was further developed in 1999 [15], to exploit the relationship between bond pricing models and option pricing models with stochastic volatility. A new stochastic volatility model was found with a closed-form solution for European option prices. Miltersen, Sandmann, and Sondermann [22] obtained closed-form solutions for term structure derivatives on log-normal interest rates, and for the "market model" of interest rates.

As advocated by Hamilton [13], the Markov-switching model maintains the assumption that time series data may display frequent changes in their observed behaviour and accounts for such changes through switches in states, where the data-generating process and average duration of each state are allowed to differ. Importantly, the statistical features and identification of the states are not imposed exogenously on the data, but rather are determined endogenously by an estimation procedure.

Previous empirical results have witnessed the success of the Markov-switching model in capturing observed nonlinearities. For example, Elliott and Osakwe [10] used Markov-modulated regime-switching market parameters to capture the time-inhomogeneity generated by the financial market. Goutte and Zou [12] used real foreign exchange rates data and compared the results obtained from regime-switching models with non-regime-switching models during a financial or economic crisis. Zhou and Mamon [37] also proved regime-switching models were more flexible, had better forecasting performance and provided a better fit than models without regime-switching.

There also has been considerable interest in the applications to various financial problems driven by a semi-Markov chain process. D'Amico, Janssen, and Manca [8] used a discrete time, non-homogeneous semi-Markov model for the rating evolution of the credit quality of a firm, and determined the credit default swap spread for a contract between two parties. Fodra and Pham [11] modeled microstructure noise using a semi-Markov model. Swishchuk [32] priced variance and volatility swaps for stochastic volatilities driven by semi-Markov processes.

This paper is organized as follows. Section 2 introduces discrete time finite state semi-Markov chains and presents a semi-martingale representation for the semi-Markov chains. In Section 3 we introduce semi-Markov switching Lévy processes. Section 4 deals with applications of semi-Markov chains in finance, namely, we consider the application of the semi-Markov chains to the binomial model in finance. Here, we also consider some estimation results for Markov and semi-Markov chains. The last Section 5 concludes the paper and mentions possible future work.

2. A semi-martingale representation for semi-Markov chains

In this section, we introduce discrete time finite state semi-Markov chains and give a semi-martingale representation for them.

2.1. Discrete time finite state semi-Markov chains. Let X_t be a discrete time, finite state semi-Markov chain defined on $(\Omega, \mathcal{F}, \mathsf{P})$. Suppose time $t \in \{0, 1, 2, \dots\}$. The state space can be taken without loss of generality to be $E = \{e_1, e_2, \dots, e_N\}$, where $e_i = (0, \dots, 1, 0, \dots, 0)' \in \mathbb{R}^N$. Suppose the jump times are $0 < \tau_1 < \tau_2 < \tau_3 < \dots$. Write $X_{\tau_n} = X_n \in E$, $\theta_{n+1} := \tau_{n+1} - \tau_n$, $\mathcal{F}_t := \sigma\{X_k : k \le t\}$. The semi-Markov property states that

$$\begin{aligned}\mathsf{P}\left(X_{n+1}=e_j,\theta_{n+1}=m \mid \mathcal{F}_{\tau_n}\right) &= \mathsf{P}(X_{n+1}=e_j,\theta_{n+1}=m \mid X_n=e_i)\\ &= \mathsf{P}(\theta_{n+1}=m \mid X_{n+1}=e_j, X_n=e_i)\,\mathsf{P}(X_{n+1}=e_j \mid X_n=e_i)\\ &= q_{ji}(m) := f_{ji}(m)P_{ji}.\end{aligned}$$

Write $\mathsf{P}(\theta_{n+1}=m \mid X_{n+1}=e_j, X_n=e_i)=f_{ji}(m)$, and $\mathsf{P}(X_{n+1}=e_j \mid X_n=e_i)=P_{ji}$. Suppose $f_{ji}(m)$ does not depend on e_j, i.e.,

$$\mathsf{P}(\theta_{n+1}=m \mid X_{n+1}=e_j, X_n=e_i)=\mathsf{P}(\theta_{n+1}=m \mid X_n=e_i)=T_i(m).$$

The process being homogeneous means these probabilities $T_i(m)$ are independent of n.

Write

$$\begin{aligned}G_i(k) &= \sum_{m=1}^{k} T_i(m)=\mathsf{P}(\theta_{n+1}\le k \mid X_n=e_i),\\ F_i(k) &= \mathsf{P}(\theta_{n+1}>k \mid X_n=e_i)=1-G_i(k),\\ F_i(k,j) &= \mathsf{P}(\theta_{n+1}>k, X_{n+1}=e_j \mid X_n=e_i)=F_i(k)P_{ji}.\end{aligned}$$

2.2. **A semi-martingale representation for a semi-Markov chain.** Consider the processes

$$\begin{aligned}p_i^n(k,j) &= \mathbb{1}_{\tau_n+k\ge\tau_{n+1}}\mathbb{1}(X_{n+1}=e_j)\mathbb{1}(X_n=e_i),\\ \tilde{p}_i^n(k,j) &= \sum_{\tau_n<\tau_n+m\le\tau_n+k\wedge\tau_{n+1}}\left(\frac{T_i(m)}{F_i(m)+T_i(m)}\right)\pi_{ji}.\end{aligned}$$

Theorem 2.1. *For $\tau_n<\tau_n+d\le\tau_n+k$,*

$$q_i^n(k,j):=p_i^n(k,j)-\tilde{p}_i^n(k,j)$$

is an $\{\mathcal{F}_k\}$ martingale.

Proof. Suppose $\tau_n<\tau_n+d\le\tau_n+k$. Then

$$\mathsf{E}[p_i^n(k,j)-\tilde{p}_i^n(d,j) \mid \mathcal{F}_d]=\mathbb{1}_{\tau_{n+1}>\tau_n+d}\left(\frac{G_i(k)-G_i(d)}{F_i(d)}\right)\pi_{ji}.$$

Also,

$$\begin{aligned}&\mathsf{E}[\tilde{p}_i^n(k,j)-\tilde{p}_i^n(d,j) \mid \mathcal{F}_d]\\ &\quad=\mathbb{1}_{\tau_{n+1}>\tau_n+d}\,\mathsf{E}\left[\sum_{\tau_n+d<\tau_n+m\le\tau_n+k\wedge\tau_{n+1}}\left(\frac{T_i(m)}{F_i(m)+T_i(m)}\right)\pi_{ji}\,\middle|\,\mathcal{F}_d\right]\\ &\quad=\mathbb{1}_{\tau_{n+1}>\tau_n+d}\,\frac{\pi_{ji}}{F_i(d)}\left[F_i(k)\sum_{\tau_n+d<\tau_n+m\le\tau_n+k}\left(\frac{T_i(m)}{F_i(m)+T_i(m)}\right)\right.\\ &\qquad\qquad\left.+\sum_{\tau_n+d<\tau_n+r\le\tau_n+k}\left(\sum_{\tau_n+d<\tau_n+m\le\tau_n+r}\frac{T_i(m)}{F_i(m)+T_i(m)}\right)T_i(r)\right].\end{aligned}$$

Now, interchanging the order in the last double sum

$$\begin{aligned}
&\sum_{\tau_n+d<\tau_n+r\le\tau_n+k}\left(\sum_{\tau_n+d<\tau_n+m\le\tau_n+r}\frac{T_i(m)}{F_i(m)+T_i(m)}\right)T_i(r)\\
&\quad=\sum_{\tau_n+d<\tau_n+m\le\tau_n+k}\left(\sum_{\tau_n+m\le\tau_n+r\le\tau_n+k}T_i(r)\right)\frac{T_i(m)}{F_i(m)+T_i(m)}\\
&\quad=\sum_{\tau_n+d<\tau_n+m\le\tau_n+k}\big(G_i(k)-G_i(m)+T_i(m)\big)\frac{T_i(m)}{F_i(m)+T_i(m)}\\
&\quad=\sum_{\tau_n+d<\tau_n+m\le\tau_n+k}\big(F_i(m)-F_i(k)+T_i(m)\big)\frac{T_i(m)}{F_i(m)+T_i(m)}\\
&\quad=\sum_{\tau_n+d<\tau_n+m\le\tau_n+k}T_i(m)-F_i(k)\sum_{\tau_n+d<\tau_n+m\le\tau_n+k}\frac{T_i(m)}{F_i(m)+T_i(m)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathsf{E}[\tilde{p}_i^n(k,j)-\tilde{p}_i^n(d,j)\mid\mathcal{F}_d]&=\nVdash_{\tau_{n+1}>\tau_n+d}\frac{\pi_{ji}}{F_i(d)}\big(G_i(k,j)-G_i(d,j)\big)\\
&=\mathsf{E}[p_i^n(k,j)-p_i^n(d,j)\mid\mathcal{F}_d].
\end{aligned}$$

So, $\mathsf{E}[q_i^n(k,j)\mid\mathcal{F}_d]=p_i^n(d,j)-\tilde{p}_i^n(d,j)=q_i^n(d,j)$, and q_i^n is a martingale. □

Corollary 2.1. *Write $Q(m)=\big(Q_{ji}(m),1\le i,j\le N\big)$ for the matrix with entries $Q_{ji}(m)=\frac{T_i(m)}{F_i(m)+T_i(m)}\pi_{ji}$; then for $\tau_n+k\ge1$,*

$$q^n(\tau_n+k):=\nVdash_{\tau_n+k\ge\tau_{n+1}}X_{n+1}-\sum_{\tau_n<\tau_n+m\le\tau_n+k\wedge\tau_{n+1}}Q(m)X_n\in\mathbb{R}^N$$

is an $\{\mathcal{F}_k\}$ martingale.

Proof. We note

$$p_i^n(k,j)=\nVdash_{\tau_n+k\ge\tau_{n+1}}\nVdash(X_{n+1}=e_j)\nVdash(X_n=e_i)$$

has compensator

$$\tilde{p}_i^n(k,j)=\sum_{\tau_n<\tau_n+m\le\tau_n+k\wedge\tau_{n+1}}\left(\frac{T_i(m)}{F_i(m)+T_i(m)}\right)\pi_{ji}.$$

However,

$$\nVdash_{\tau_n+k\ge\tau_{n+1}}X_{n+1}=p^n(k)=\nVdash_{\tau_n+k\ge\tau_{n+1}}\left(\sum_{i=1}^N\langle X_n,e_i\rangle\right)\left(\sum_{j=1}^N\langle X_{n+1},e_j\rangle\right)e_j.$$

So this has compensator

$$\begin{aligned}
\tilde{p}^n(k)&=\sum_{j=1}^N\sum_{i=1}^N\sum_{\tau_n<\tau_n+m\le\tau_n+k\wedge\tau_{n+1}}\frac{T_i(m)}{F_i(m)+T_i(m)}\pi_{ji}\langle X_n,e_i\rangle e_j\\
&=\sum_{\tau_n<\tau_n+m\le\tau_n+k\wedge\tau_{n+1}}Q(m)X_n.
\end{aligned}$$
□

Corollary 2.2. *Write $q(t)=\sum\nVdash(\tau_n\le\tau_n+t)q^n(t)$; then $q(t)$ is an $\{\mathcal{F}_k\}$ martingale.*

Proof. Suppose $\tau_m < \tau_n + s \leq \tau_{m+1} \leq \cdots \leq \tau_n < \tau_n + t$. Note $p^n(t)$ and $\tilde{p}^n(t)$ are only defined for $t \geq \tau_n + 1$. With $q^n(t) = p^n(t) - \tilde{p}^n(t)$, and $t \geq \tau_n + 1$,

$$\begin{aligned} \mathsf{E}[q^n(t) \mid \mathcal{F}_{\tau_n+1}] &= q^n(\tau_n+1) = p^n(\tau_n+1) - \tilde{p}^n(\tau_n+1) \\ &= \mathbb{1}_{\tau_n+1=\tau_{n+1}} X_{n+1} - Q(\tau_n+1)X_n. \end{aligned}$$

For $i \neq j$, $Q_{ji}(\tau_n+1) = \frac{T_i(\tau_n+1)}{T_i(\tau_n+1)} P_{ji}$, so $Q(\tau_n+1) = P = (P_{ji}, 1 \leq i, j \leq N)$.

Also,

$$\mathsf{E}[\mathbb{1}_{\tau_n+1} X_{n+1} \mid \mathcal{F}_{\tau_n}] = PX_n.$$

So,

$$\mathsf{E}[q^n(t) \mid \mathcal{F}_{\tau_n}] = \mathsf{E}[q^n(\tau_n+1) \mid \mathcal{F}_{\tau_n}] = 0 \in \mathbb{R}^N. \qquad \square$$

Corollary 2.3. $\sum_{n\geq 0} q^n(t) = q(t) \in \mathbb{R}^N$ *is then the martingale associated with the semi-Markov chain* $X = \{X_t, t = 0, 1, 2, \dots\}$.

We know $q(t) = \sum_{n\geq 0} p^n(t) - \sum_{n\geq 0} \tilde{p}^n(t)$. Write $Q(t) = \sum_{n\geq 0} \tilde{p}^n(t) \in \mathbb{R}^N$; then X has semi-martingale representation

$$X_t = X_0 + Q(t) + q(t) \in \mathbb{R}^N.$$

3. Semi-Markov switching Lévy processes

Here, we introduce semi-Markov switching Lévy processes which are a natural generalization of classical Lévy processes.

3.1. Lévy processes.

For completeness, we present a brief description of Lévy processes.

A Lévy process is a stochastic process representing the motion of a point whose successive random displacements are independent, and statistically identical over different time intervals of the same length. A Lévy process may thus be viewed as the continuous time analog of a random walk. The most well-known examples of Lévy processes are Brownian motion and the Poisson process. Aside from Brownian motion with drift, all other proper Lévy processes have discontinuous paths.

Càdlàg is a French acronym for “right-continuous with left limit”. A càdlàg, adapted, real valued stochastic process $L = (L_t)_{t\geq 0}$ with $L_0 = 0$ is a *Lévy process* if the following conditions are satisfied [1]:

(i) the process L has independent increments, i.e., $L_t - L_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s \leq t \leq T$,
(ii) the process L has stationary increments, i.e., the distribution of $L_{t+s} - L_t$ does not depend on t for any $s, t \geq 0$,
(iii) the process L is stochastically continuous, i.e., for every $t \geq 0$ and $\varepsilon > 0$:

$$\lim_{s\to t} \mathsf{P}(|L_t - L_s| > \varepsilon) = 0.$$

We now define the Lévy measure.

Let L_t be a Lévy process on $\mathbb{R}^d$. The *jump measure* μ on $\mathbb{R}^d$ defined by [21]

$$\mu(t, dz) = \sum_{s>0} \mathbb{1}_{(\Delta L_s \neq 0)} \delta_{(s,\Delta L_s)}(t, dz)$$

is called the Lévy measure of L_t; $\delta_{(s,\Delta L_s)}(t, dz)$ denotes the unit mass at $(s, \Delta L_s)$.

Any Lévy process may be decomposed into the sum of a Brownian motion, a linear drift and a pure jump process which captures all jumps of the original Lévy process. The latter can be thought of as a superposition of centred compound Poisson processes. This result is known as the Lévy–Itô decomposition. Mathematically, the Lévy–Itô decomposition for L_t is [21]

$$L_t = at + \sigma W(t) + \int_{|z|>1} \int_{]0,t]} zN(ds, dz) + \int_{|z|\le 1} \int_{]0,t]} z\widetilde{N}(ds, dz),$$

where a, σ are constants with $\sigma \ge 0$, $W(t)$ is a standard Brownian motion, N is an independent Poisson random measure $\mu(t, dz)$, and $\widetilde{N}$ is a compensated Poisson random measure, where $\widetilde{N}(dt, dz) = N(dt, dz) - m(dz)\, dt$.

The distribution of a Lévy process is characterized by its characteristic function, which is given by the Lévy–Khintchine formula: if $L = (L_t)_{t\ge 0}$ is a Lévy process, then its characteristic function $\phi_L(u)$ is given by [21]

$$\phi_t(u) = \mathsf{E}\left(e^{iuL_t}\right) = \exp\left\{ t \left(iau - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} \left(e^{iuz} - 1 - iz\mathbb{1}(|z| \le 1)\right) m(dz) \right) \right\}.$$

This is determined by the Lévy–Khintchine triplet $\left(a, \sigma^2, m(dz)\right)$.

3.2. **Semi-Markov switching Lévy process.**

For the semi-Markov process with the semi-martingale representation $X_t = X_0 + \tilde{p}(t) + q(t)$, the semi-Markov switching Lévy process Λ_t is then

$$\Lambda_t = \sum_{i=1}^{N} L_t^i \langle X_t, e_i \rangle = \langle L_t, X_t \rangle.$$

If L_t is the vector of Lévy processes $\left(L_t^1, L_t^2, \dots, L_t^N\right)$, then it gives:

$$d\Lambda_t = \langle dL_t, X_t \rangle + \langle L_t, dX_t \rangle = \sum_{i=1}^{N} dL_t^i \langle X_t, e_i \rangle + \sum_{i=1}^{N} L_t^i \langle dX_t, e_i \rangle$$

with $dX_t = d\tilde{p}(t) + dq(t)$.

4. Application of semi-Markov chains: Binomial model and estimations

In this section, we apply the semi-Markov switching binomial model in finance, and obtain the formula for the arbitrage free price of a financial asset based on our model. We also present some estimation results for Markov and semi-Markov chains.

4.1. **(B,S)-security regime-switching markets.**

Let us consider a finite state Markov chain $X = \{X_t, t = 0, 1, 2, \dots\}$. X will model the state of the economy or market. We suppose $X_t \in \{e_1, e_2, \dots, e_N\}$ (we can choose $N = 2$), $e_i = (0, \dots, 0, 1, \dots, 0)' \in \mathbb{R}^N$. Suppose there are vectors $u = (u_1, u_2, \dots, u_N)' \in \mathbb{R}^N$, $d = (d_1, d_2, \dots, d_N)' \in \mathbb{R}^N$, $e = (e_1, e_2, \dots, e_N)' \in \mathbb{R}^N$; so $u_t = \langle u, X_t \rangle$, $d_t = \langle d, X_t \rangle$, $e_t = \langle e, X_t \rangle$ (see [10]).

Suppose we have a risk-free asset, namely a bond with its value at time zero $B_0 = 1$, $B_{t+1} = (1 + e_t)B_t$. There is a second risky asset with the following price change:

$$\begin{array}{ccc} & & S_{t+1}^{+} = u_t S_t \\ & \nearrow & \\ S_t & & \\ & \searrow & \\ & & S_{t+1}^{-} = d_t S_t \end{array}$$

$\{\mathcal{F}_t\}$ is the filtration generated by price S.

Suppose V_t is a claim paid at time T. V_t is $\{\mathcal{F}_t\}$-measurable. In this binomial model, there is a perfect hedge and the arbitrage-free price of V at time t is derived by working backwards. Suppose V_{t+1} has been determined for all states of the world at time $t+1$. We wish to find V_t below:

$$\begin{tikzcd} & V_{t+1}^{+} \\ V_t \arrow[ur, dash] \arrow[dr, dash] & \\ & V_{t+1}^{-} \end{tikzcd}$$

At time t, we form a portfolio by buying α_t of S_t and β_t of B_t. The value of this portfolio is $\Pi_t = \alpha_t S_t + \beta_t B_t$. At time $t+1$ the portfolio has values:

$$\begin{aligned}
\Pi_{t+1}^{+} &= \alpha_t S_t u_t + \beta_t B_t(1+e_t) \text{ in the up state,}\\
\Pi_{t+1}^{-} &= \alpha_t S_t d_t + \beta_t B_t(1+e_t) \text{ in the down state,}
\end{aligned}$$

α_t and β_t are to be chosen so that $\Pi_{t+1}^{+} = V_{t+1}^{+}$ and $\Pi_{t+1}^{-} = V_{t+1}^{-}$.
Π_t must then be the price of V_t at time t. In fact, this gives

$$V_t = \frac{1}{(1+e_t)}\left[\frac{((1+e_t)-d_t)}{u_t-d_t}V_{t+1}^{+} + \frac{(u_t-(1+e_t))}{u_t-d_t}V_{t+1}^{-}\right].$$

If we introduce the "risk-neutral" probability $\pi_t = \frac{(1+e_t)-d_t}{u_t-d_t}$, $1-\pi_t = \frac{u_t-(1+e_t)}{u_t-d_t}$, then

$$V_t = \frac{1}{(1+e_t)}\big[\pi_t V_{t+1}^{+} + (1-\pi_t)V_{t+1}^{-}\big].$$

Write $\Pi_i = \frac{(1+e_i)-d_i}{u_i-d_i}$, and $\Pi = (\pi_1, \pi_2, \dots, \pi_N)' \in \mathbb{R}^N$.
Then $\pi_t := \frac{(1+e_t)-d_t}{u_t-d_t} = \langle \pi, X_t\rangle$.

4.2. **Regime-switching binomial model.** Now consider a Bernoulli random variable $\mathcal{X}_t = 0$ or 1 with probability $p_t = \mathsf{P}(X_t = 1 \mid \mathcal{F}_{t-1})$. P_t is $\mathcal{F}_t$-measurable where $\mathcal{F}_t = \sigma\{\mathcal{X}_s, 0 \le s \le t\}$.

We are considering a binomial model but with risk-neutral probabilities π_t, $t = 0, 1, \dots$, so the pricing still works with

$$V_t = \frac{1}{1+e_t}\big(\pi_t V_{t+1}^{+} + (1-\pi_t)V_{t+1}^{-}\big).$$

In the binomial model, the hedging and unique price depends on there being two states of the world at the next time and two assets to hedge. This hedge works if the u_t, d_t, and e_t are adapted to $\{\mathcal{F}_t\}$. That is, at time t, u_t, d_t, and e_t are $\mathcal{F}_t$-measurable.

Suppose $S_{t+1} = u_t S_t$ if $\mathcal{X}_{t+1} = 1$, $S_{t+1} = d_t S_t$ if $\mathcal{X}_{t+1} = 0$; the state of the economy is modelled by a finite state Markov chain $X = \{X_t : t = 0, 1, \dots, T\}$. Consider vectors

$$u = (u_1, \dots, u_N)' \in \mathbb{R}^N, \qquad d = (d_1, \dots, d_N)' \in \mathbb{R}^N, \qquad r = (r_1, \dots, r_N)' \in \mathbb{R}^N;$$

so $u_t = \langle u, X_t\rangle$, $d_t = \langle d, X_t\rangle$, $r_t = \langle r, X_t\rangle$.

Suppose initially we know the trajectory of the chain X_t, for $t = 0, 1, \dots, T$, i.e., suppose we know $X_0 = e_{i_0}$, $X_1 = e_{i_1}$, …, $X_T = e_{i_T}$. Then, $u_t = u_{i_t}$, $d_t = d_{i_t}$, $r_t = r_{i_t}$.

Then, there is a risk-neutral measure Q given by

$$\begin{aligned}\left.\frac{dQ}{d\mathsf{P}}\right|_{\mathcal{F}_T \vee \mathcal{F}_T^X} &= \sum_{n=0}^{T-1} \left(\frac{q_n}{p_n}\right)^{\mathcal{X}_{n+1}} \left(\frac{1-q_n}{1-p_n}\right)^{1-\mathcal{X}_{n+1}} \\ &= \sum_{n=0}^{T-1} \left(\frac{\langle q, X_n\rangle}{p_n}\right)^{\mathcal{X}_{n+1}} \left(\frac{1-\langle q, X_n\rangle}{1-p_n}\right)^{1-\mathcal{X}_{n+1}} \\ &= \sum_{n=0}^{T-1} \left(\frac{q_{i_n}}{p_n}\right)^{\mathcal{X}_{n+1}} \left(\frac{1-q_{i_n}}{1-p_n}\right)^{1-\mathcal{X}_{n+1}}.\end{aligned}$$

If $X_0 = e_{i_0}$, $X_1 = e_{i_1}$, $\dots$, $X_T = e_{i_T}$, knowing this trajectory of X to time T, if G is an $\mathcal{F}_t$-measurable random variable, the arbitrage free price of G is

$$V_t = \mathsf{E}^Q\left[G\left(\sum_{n=t}^{T-1}\frac{1}{r_{i_n}}\right)\,\middle|\, \mathcal{F}_t \vee \mathcal{F}_T^X\right].$$

Putting in $\frac{dQ}{d\mathsf{P}}$, knowing $X_0, X_1, \dots, X_T$, the arbitrage free price of G is

$$\begin{aligned}V_t &= \mathsf{E}\left[G\left(\sum_{n=t}^{T-1}\left(\frac{\langle q, X_n\rangle}{p_n}\right)^{\mathcal{X}_{n+1}} \left(\frac{1-\langle q, X_n\rangle}{1-p_n}\right)^{1-\mathcal{X}_{n+1}} \frac{1}{\langle r, X_n\rangle}\right)\,\middle|\, \mathcal{F}_t \vee \mathcal{F}_T^X\right] \\ &= \mathsf{E}\left[G\left(\sum_{n=t}^{T-1}\left(\frac{q_{i_n}}{p_n}\right)^{\mathcal{X}_{n+1}} \left(\frac{1-q_{i_n}}{1-p_n}\right)^{1-\mathcal{X}_{n+1}} \frac{1}{r_{i_n}}\right)\,\middle|\, \mathcal{F}_t \vee \mathcal{F}_T^X\right]\hat{\pi}(t:T),\end{aligned}$$

where $\hat{\pi}(t:T) = \prod_{n=t}^{T-1}\langle e_{i_n}, X_n\rangle$.

We are then left to condition out the product

$$\prod_{n=t}^{T-1}\langle e_{i_n}, X_n\rangle = \langle e_{i_t}, X_t\rangle\langle e_{i_{t+1}}, X_{t+1}\rangle \dots \langle e_{i_{T-1}}, X_{T-1}\rangle = \hat{\pi}(t:T)$$

given X_t. This will be

$$\begin{aligned}\mathsf{E}[\pi(t:T) \mid X_t = e_{i_t}] &= \langle e_{i_t}, X_t\rangle\langle e_{i_{t+1}}, AX_t\rangle \dots \left\langle e_{i_{T-1}}, A^{T-1-t}X_t\right\rangle \\ &= a_{i_{t+1}i_t} a_{i_{t+2}i_{t+1}} a_{i_{t+3}i_{t+2}} \dots a_{i_{T-t-1}i_{T-t-2}} \\ &= \hat{\pi}(t:T),\end{aligned}$$

where $A = (a_{ji}, 1 \le i, j \le N)$ and $a_{ji} = \mathsf{P}(X_{t+1} = e_j \mid X_t = e_i)$. The final result will be the sum over all possible paths $e_{i_t}, e_{i_t+1}, \dots, e_{i_{T-1}}$ for $X_t, \dots, X_{T-1}$,

$$\widehat{V}_t = \sum \mathsf{E}\left[G\left(\sum_{n=t}^{T-1}\left(\frac{q_{i_n}}{p_n}\right)^{\mathcal{X}_{n+1}} \left(\frac{1-q_{i_n}}{1-p_n}\right)^{1-\mathcal{X}_{n+1}} \frac{1}{r_{i_n}}\right)\right]\hat{\pi}(t:T),$$

of which each of the paths has a probability $\hat{\pi}(t:T)$, which is the product of the transition probabilities of the steps in the path

$$\hat{\pi}(t:T) = a_{i_{t+1}i_t} a_{i_{t+2}i_{t+1}} \dots a_{i_{T-t-1}i_{T-t-2}}.$$

4.3. **Estimates for Markov and semi-Markov chains.**

4.3.1. *Estimates for a Markov chain.* We suppose the state X_t is known or observed at each time t.

Suppose the chain X has transition probabilities

$$p_{ji} = \mathsf{P}(X_{t+1} = e_j \mid X_t = e_i) = \mathsf{P}(X_1 = e_j \mid X_0 = e_i).$$

If the chain is homogeneous, write $A = (p_{ji}, 1 \leq i, j \leq N)$; then

$$X_{t+1} = AX_t + M_{t+1},$$

where $\mathsf{E}[M_{t+1} \mid \mathcal{F}_t] = 0 = (0, 0, \dots, 0)' \in \mathbb{R}^N$.

The likelihood ratio is

$$\Lambda_t = \prod_{l=0}^{t} \lambda_l,$$

where $\lambda_0 = \langle l_0, X_0 \rangle$; and for $l \geq 1$, $\lambda_l = \langle X_l, AX_{l-1} \rangle$. We also have $p_{ji} \geq 0$ and $\sum_{j=1}^N p_{ji} = 1$.

Now

$$\log \Lambda_t = \sum_{l=0}^{t} \lambda_l = \sum_{l=1}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} \log p_{ji} \langle X_l, e_j \rangle \langle X_{l-1}, e_i \rangle.$$

We wish to maximize this subject to $\sum_{j+1}^N p_{ji} = 1$.

Write λ for the Lagrange multiplier and consider

$$L_t := \sum_{j=1}^{N} \sum_{i=1}^{N} \log p_{ji} \nu^{ij}(t) + \lambda \left(\sum_{j=1}^{N} p_{ji} - 1 \right),$$

where $\nu^{ji}(t) = \sum_{l=1}^t \langle X_l, e_j \rangle \langle X_{l-1}, e_i \rangle$ is the number of jumps from e_i to e_j up to time t.

First order conditions given for a_{ji} is

$$\frac{1}{p_{ji}} \nu^{ij}(t) + \lambda = 0$$

and

$$\sum_{j=1}^{N} p_{ji} = 1.$$

Then

$$\lambda p_{ji} = -\nu^{ji}(t)$$

and summing over j,

$$\lambda = -\sum_{j=1, i=1}^{N} \nu^{ij}(t) = \nu^i(t),$$

where $\nu^i(t) = \sum_{l=1}^t \langle X_{l-1}, e_i \rangle$ is the amount of time the chain X has spent in state e_i up to time t.

Therefore, observing X_t for $0, 1, 2, \dots, t$, the estimate of p_{ji} is $\frac{\nu^{ji}(t)}{\nu^i(t)}$.

4.3.2. *Estimates for a semi-Markov chain.* Now suppose $X = \{X_t, t = 0, 1, \dots\}$ is a finite state semi-Markov chain with jumps $0 < \tau_1 < \tau_2 < \dots$. We shall write $X_{\tau_n} = X_n$, $\theta_{n+1} := \tau_{n+1} - \tau_n$. Recall that the semi-Markov property states

$$\begin{aligned} \mathsf{P}(X_{n+1} = e_j, \theta_{n+1} = m \mid \mathcal{F}_{\tau_n}) &= \mathsf{P}(X_{n+1} = e_j, \theta_{n+1} = m \mid X_n = e_i) \\ &= q_{ji}(m) := P_{ji} f_{ji}(m). \end{aligned}$$

Write $h_l(X_l) = k$. When $X_{l-k} \neq e_i$ but $X_{l-k+1} = e_i$, $X_{l-k+2} = e_i$, $\dots$, $X_l = e_i$; then $h_l(X_l) = 1 + \langle X_l, X_{l-1} \rangle h_{l-1}(X_{l-1})$.

$h_l(X_l)$ counts the number of consecutive states the process has been in state X_l.

Write

$$p_i(k) := \mathsf{P}(\theta_{n+1} = k \mid Z_n = e_i),$$
$$F_i(k) := \mathsf{P}(\theta_{n+1} \geq k \mid Z_n = e_i) = \sum_{l=k}^{\infty} p_i(l).$$

Lemma 4.1. *Suppose* $X_t = e_i$ *and* $\mathsf{P}(h_t(X_t) = k \mid X_t = e_i) > 0$; *then*

$$\mathsf{P}(X_{t+1} \neq e_i \mid X_t = e_i, h_t(X_t) = k) = \frac{p_i(k)}{F_i(k)}.$$

Corollary 4.1. *Suppose* $X_t = e_i$ *and* $\mathsf{P}(h_t(X_t) = k \mid X_t = e_i) > 0$; *then*

$$\mathsf{P}(X_{t+1} = e_i \mid X_t = e_i, h_t(X_t) = k) = \frac{F_i(k+1)}{F_i(k)}.$$

Write

$$\frac{p(k)}{F(k)} := \left(\frac{p_1(k)}{F_1(k)}, \frac{p_2(k)}{F_2(k)}, \dots, \frac{p_N(k)}{F_N(k)}\right),$$
$$\frac{F(k+1)}{F(k)} := \left(\frac{F_1(k+1)}{F_1(k)}, \frac{F_2(k+1)}{F_2(k)}, \dots, \frac{F_N(k+1)}{F_N(k)}\right).$$

For $j \neq i$,

$$A_{ji} = \mathsf{P}(X_{n+1} = e_j \mid X_n = e_i),$$
$$A = (A_{ji}), \qquad 1 \leq i, j \leq N.$$

We have defined

$$\begin{aligned} q_{ji}(m) &= \mathsf{P}(X_{n+1} = e_j, \theta_{n+1} = m \mid X_n = e_i) \\ &= \mathsf{P}(\theta_{n+1} = m \mid X_{n+1} = e_j, X_n = e_i)\,\mathsf{P}(X_{n+1} = e_j, X_n = e_i) \\ &= T_{ji}(m) A_{ji}, \end{aligned}$$

where $T_{ji}(m) = \mathsf{P}(\theta_{n+1} = m \mid X_{n+1} = e_j, X_n = e_i)$.

We shall suppose $T_{ji}(m)$ is independent of e_j, so

$$\mathsf{P}(\theta_{n+1} = m \mid X_{n+1} = e_j, X_n = e_i) = \mathsf{P}(\theta_{n+1} = m \mid X_n = e_i) = T_i(m).$$

Then $q_{ji}(m) = T_i(m) A_{ji}$.

Consider the matrix

$$B_t(X_t) := \left\langle X_t, \frac{p(h_t(X_t))}{F(h_t(X_t))} \right\rangle A + \left\langle X_t, \frac{F(h_t(X_t)+1)}{F(h_t(X_t))} \right\rangle I.$$

Lemma 4.2. *For all* $t \geq 0$,

$$\mathsf{E}[X_{t+1} \mid \mathcal{F}_t] = B_t(X_t) X_t \in \mathbb{R}^N.$$

Write $\lambda_0 = \langle X_0, p_0 \rangle$, where p_0 is the distribution of X_0.

For $l \geq 1$, $\lambda_{l+1} = \langle B_l(X_l) X_l, X_{l+1} \rangle$. The likelihood ratio is then

$$\Lambda_k := \prod_{l+1}^{k} \lambda_l.$$

Again we have $\sum_{j=1}^{N} p_{ji} = 1$, and by construction we have $p_{ii} = 0$. Then

$$\log \Lambda_k = \sum_{t=0}^{k} \sum_{j=1}^{N} \sum_{i=1}^{N} \log \left[\frac{p_i(h_t(X_t))}{F_i(h_t(X_t))} a_{ji} + \frac{F_i(h_t(X_t)+1)}{F_i(h_t(X_t))} \right] \langle X_{t+1}, e_j \rangle \langle X_t, e_i \rangle.$$

As $p_{ii} = 0$, this is equal to

$$\sum_{t=0}^{k}\sum_{j=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\log\left[\frac{p_i(h_t(X_t))}{F_i(h_t(X_t))}p_{ji}\right]\langle X_{t+1},e_j\rangle\langle X_t,e_i\rangle$$
$$+\sum_{t=0}^{k}\sum_{i=1}^{N}\log\left[\frac{F_i(h_i(X_t)+1)}{F_i(h_i(X_t))}\right]\langle X_{t+1},e_j\rangle\langle X_t,e_i\rangle.$$

Now

$$\log\left[\frac{p_i(h_t(X_t))}{F_i(h_t(X_t))}p_{ji}\right] = \log(p_i(h_t(X_t))) - \log F_i(h_t(X_t)) + \log p_{ji}.$$

So the first order conditions again give the estimates

$$p_{ji} = \frac{\nu^{ij}(t)}{\nu^i(t)},$$

where

$$\nu^{ij}(t) = \sum_{\tau_{n+1}\leq t}\langle X_{n+1},e_j\rangle\langle X_n,e_i\rangle$$

and

$$\nu^i(t) = \sum_{\substack{j=1\\ j\neq i}}^{N}\nu^{ji}(t).$$

5. Conclusion

In this paper, we introduced a new model of discrete time switching semi-Markov chains, derived its semi-martingale representation, and applied a semi-Markov switching binomial model in finance. Semi-Markov switching models will be extended to the continuous case and other derivative pricing in finance in our future papers. We shall also apply semi-Markov switching Lévy processes to derivative pricing and other problems in finance in our future works. Semi-Markov processes are also widely applied in other models, including limit order books in finance [34], computer science, sociology, biology, and medicine [35].

Acknowledgments

The first and second authors wish to thank NSERC for continuing support. The third author wishes to thank Dr. R. Elliott and Dr. A. Swishchuk for continuing support.

Bibliography

[1] V. Anisimov, *Switching Processes in Queueing Models*, Wiley, 2013. MR2437051

[2] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, 2004. MR2072890

[3] D. S. Bates, *Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options*, Rev. Financ. Stud. **9** (1996), 69–107.

[4] P. Bremaud, *Point Processes and Queues: Martingale Dynamics*, Springer, 1981. MR636252

[5] J. Buffington and R. J. Elliott, *American options with regime switching*, Int. J. Theor. Appl. Finance **5** (2002), 497–514. MR1916958

[6] S. Cohen and R. Elliott, *Stochastic Calculus and Applications*, 2nd edition, Birkhauser, 2015. MR3443368

[7] E. Cinlar, *Time dependence of queues with semi-Markov services*, J. Appl. Probab. **4** (1967), 356–364. MR0214157

[8] G. D'Amico, J. Janssen, and R. Manca, *Valuing credit default swap in a non-homogeneous semi-Markovian rating based model*, Comput. Econ. **29** (2007), no. 2, 119–138.

[9] R. J. Elliott, *Stochastic Calculus and Applications*, Springer, Berlin–Heidelberg–New York, 1982. MR678919

[10] R. J. Elliott and C. J. Osakwe, *Option pricing for pure jump processes with Markov switching compensators*, Finance Stoch. **10** (2006), 250–275. MR2223098

[11] P. Fodra and H. Pham, *Semi-Markov model for market microstructure*, Appl. Math. Finance **22** (2013), no. 3, 261–295. MR3367230

[12] S. Goutte and B. Zou, *Foreign exchange rates under Markov regime switching model*, CREA Discussion paper series, University of Luxemburg, 2011.

[13] J. D. Hamilton, *A new approach to the economic analysis of nonstationary time series and the business cycle*, Econometrica **57** (1989), no. 2, 357–384. MR996941

[14] S. L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Financ. Stud. **6** (1993), 327–343.

[15] S. L. Heston, *A simple new formula for options with stochastic volatility*, Washington University of St. Louis Working Paper, 1999.

[16] V. Koroliuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Scientific, 2005. MR2205562

[17] P. Levy, *Processus Semi-Markoviens*, Proceedings of International Congress of Mathematics **3** (1954), 416–426. MR0088105

[18] N. Liminios and A. Swishchuk, *Discrete-time semi-Markov random evolutions and their applications*, Adv. Appl. Probab. **45** (2013), no. 1, 214–240. MR3077547

[19] A. Melino and S. M. Turnbull, *Pricing foreign currency options with stochastic volatility*, J. Econometrics **45** (1990), 239–265.

[20] A. Melino and S. M. Turnbull, *The pricing of foreign-currency options*, Canad. J. Econ. **24** (1991), 251–181.

[21] R. C. Merton, *Option pricing when underlying stock returns are discontinuous*, J. Finance Econ. **3** (1976), 125–144.

[22] K. Miltersen, K. Sandmann, and D. Sondermann, *Closed form solutions for term structure derivatives with log-normal interest rates*, J. Finance **52** (1997), 409–430.

[23] A. Papapantoleon, *An Introduction to Lévy Processes with Applications to Mathematical Finance. Lecture notes*, arXiv:0804.0482, 2000.

[24] R. Pyke, *Markov renewal processes: definitions and preliminary properties*, Ann. Math. Statistics **32** (1961), 1231–1242. MR0133888

[25] R. Pyke, *Markov renewal processes with finitely many states*, Ann. Math. Statistics **32** (1961), 1243–1259. MR0154324

[26] R. Pyke, and R. Schaufele, *Limit theorems for Markov renewal processes*, Ann. Math. Statistics **35** (1964), 1746–1764. MR0168026

[27] R. Pyke and R. Schaufele, *The existence and uniqueness of stationary measures for Markov renewal processes*, Ann. Math. Statistics **37** (1966), 1439–1462. MR0203811

[28] J. Rumsey, *Pricing cross-currency options*, J. Futures Markets **11** (1991), 89–93.

[29] W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives*, Wiley, 2003.

[30] D. S. Silvestrov, *Semi-Markov Processes with a Discrete State Space*, Library for an Engineer in Reliability, Sovetskoe Radio, Moscow, 1980. MR612026

[31] W. L. Smith, *Regenerative stochastic processes*, Proceedings of the Royal Society of London Archives **232** (1955), 6–31. MR0073877

[32] A. Swishchuk, *Pricing of variance and volatility swaps with semi-Markov volatilities*, Can. Appl. Math. Q. **18** (2010). MR2858147

[33] A. Swishchuk and R. Elliott, *Pricing Options. Hidden Markov Models in Finance*, Springer, 2007.

[34] A. Swishchuk and N. Vadori, *A semi-Markovian modelling of limit order markets*, SIAM J. Finan. Math. (2017). MR3625801

[35] A. Swishchuk and J. Wu, *Evolution of Biological Systems in Random Media: Limit Theorems and Stability*, Springer, 2003. MR2031169

[36] L. Takacs, *On secondary processes generated by recurrent processes*, Archiv der Mathematik **7** (1956), 17–29. MR0078584

[37] N. Zhou and R. S. Mamon, *An accessible implementation of interest rate models with regime switching*, Expert Syst. Appl. **39** (2012), no. 5, 4679–4689.

University of Calgary, Calgary, Canada and University of South Australia
Email address: relliott@ucalgary.ca

University of Calgary, Calgary, Canada
Email address: aswish@ucalgary.ca

University of Calgary, Calgary, Canada
Email address: yi.zhang@ucalgary.ca

Received 27/FEB/2017
Originally published in English