KALMAN–BUCY FILTERING FOR LINEAR SYSTEMS DRIVEN BY THE COX PROCESS WITH SHOT NOISE INTENSITY AND ITS APPLICATION TO THE PRICING OF REINSURANCE CONTRACTS

ANGELOS DASSIOS,* *London School of Economics*

JI-WOOK JANG,** *University of New South Wales*

Abstract

In practical situations, we observe the number of claims to an insurance portfolio but not the claim intensity. It is therefore of interest to try to solve the 'filtering problem'; that is, to obtain the best estimate of the claim intensity on the basis of reported claims. In order to use the Kalman–Bucy filter, based on the Cox process incorporating a shot noise process as claim intensity, we need to approximate it by a Gaussian process. We demonstrate that, if the primary-event arrival rate of the shot noise process is reasonably large, we can then approximate the intensity, claim arrival, and aggregate loss processes by a three-dimensional Gaussian process. We establish weak-convergence results. We then use the Kalman–Bucy filter and we obtain the price of reinsurance contracts involving high-frequency events.

Keywords: Kalman–Bucy filter; Gaussian process; Cox process; shot noise process; piecewise-deterministic Markov process theory; stop-loss reinsurance contract

2000 Mathematics Subject Classification: Primary 60G35
Secondary 60F05; 60G55; 60J75; 91B30

1. Introduction

In insurance modelling, the Poisson process has been used as a claim arrival process. Extensive discussion of the Poisson process, from both applied and theoretical viewpoints, can be found in Cramér (1930), Cox and Lewis (1966), Bühlmann (1970), Çinlar (1975), Gerber (1979), and Medhi (1982). However, there is a significant volume of literature that questions the suitability of the Poisson process in insurance modelling (Seal (1983); Beard *et al.* (1984)). From a practical point of view, there is no doubt that the insurance industry needs a more suitable claim arrival process, with deterministic intensity, than the Poisson process.

As an alternative point process to generate the claim arrivals, we can employ the Cox process or a doubly stochastic Poisson process (Cox (1955); Bartlett (1963); Haight (1967); Serfozo (1972); Grandell (1976), (1991), (1997); Brémaud (1981); Consul (1989); Lando (1994)). An important book on Cox processes is that by Bening and Korolev (2002), in which various limit theorems as well as applications in both insurance and finance are discussed. The Cox process

provides us with the flexibility to allow the intensity not only to depend on time but also to be a stochastic process. In a recent paper (Dassios and Jang (2003)), the authors demonstrated how the Cox process with shot noise intensity can be used in the pricing of catastrophe reinsurance and derivatives.

As the claim intensity function within the Cox process is not observable, it can only be observed on the basis of an observed process of reported claims. Thus, we consider the filtering problem to obtain the best estimate of the claim intensity on the basis of the observed process of reported claims or observed accidents (Dassios and Jang (1998)).

We start by defining the quantities of interest; these are the doubly stochastic (with a shot noise intensity) point process of claim arrivals and the aggregate loss process. In Section 3, we prove a weak-convergence result for the three-dimensional process consisting of the intensity, claim arrival, and aggregate loss processes. In Section 4, we obtain the Kalman–Bucy filter result, which is then used, in Section 5, to price a reinsurance contract.

We employ piecewise-deterministic Markov processes, the theory of which was developed by Davis (1984), to obtain the original moments of our processes. The piecewise-deterministic Markov process theory is a powerful mathematical tool for examining nondiffusion models. For details, we refer the reader to Davis (1984); Dassios (1987); Dassios and Embrechts (1989); Jang (1998), (2004); Rolski *et al.* (1999); and Dassios and Jang (2003).

For similar results to the ones derived in this paper, we refer the reader to Gnedenko and Kolmogorov (1954); Snyder (1975); Kruglov (1976); Davis (1977); Lipster and Shiryayev (1977), (1978); Ahmed (1998); and Bening and Korolev (2002).

2. The Cox process and the shot noise process

The Cox process (or a doubly stochastic Poisson process) can be viewed as a two-step randomization procedure. A process λ_t is used to generate another process N_t by acting as its intensity. That is, N_t is a Poisson process conditional on λ_t, which itself is a stochastic process (if λ_t is deterministic then N_t is a Poisson process). Many alternative definitions of a doubly stochastic Poisson process can be given. We will offer the one adopted by Brémaud (1981).

Definition 1. Let $(\Omega, \mathcal{F}, P)$ be a probability space with information structure given by $\mathcal{F} = \{\mathfrak{I}_t, \ t \in [0, T]\}$. Let N_t be a point process adapted to $\mathcal{F}$. Let λ_t be a nonnegative process adapted to $\mathcal{F}$ such that

$$\int_0^t \lambda_s \, ds < \infty \quad \text{almost surely (no explosions).}$$

If, for all $0 \le t_1 \le t_2$ and $u \in \mathbb{R}$,

$$E(e^{iu(N_{t_2} - N_{t_1})} \mid \mathfrak{I}_{t_2}^\lambda) = \exp\left\{ (e^{iu} - 1) \int_{t_1}^{t_2} \lambda_s \, ds \right\} \tag{1}$$

then N_t is called an $\mathfrak{I}_t$-doubly stochastic Poisson process with intensity λ_t, where $\mathfrak{I}_t^\lambda$ is the σ-algebra generated by λ up to time t, i.e. $\mathfrak{I}_t^\lambda = \sigma\{\lambda_s, \ s \le t\}$.

Equation (1) gives us

$$P\{N_{t_2} - N_{t_1} = k \mid \lambda_s, \ t_1 \le s \le t_2\} = \frac{\exp(-\int_{t_1}^{t_2} \lambda_s \, ds)(\int_{t_1}^{t_2} \lambda_s \, ds)^k}{k!}. \tag{2}$$

One of the processes that can be used to measure the impact of primary events is the shot noise process (Cox and Isham (1980), (1986); Klüppelberg and Mikosch (1995)). The shot noise process is particularly useful within the claim arrival process, as it measures the frequency of primary events, their magnitude, and the time period needed to determine their effects. As time passes, the shot noise process decreases as more and more claims are settled. This decrease continues until another event occurs, which results in a positive jump in the shot noise process. Therefore, the shot noise process can be used as the parameter of a doubly stochastic Poisson process to measure the number of claims due to primary events, i.e. we will use it as a claim intensity function to generate the Cox process. We will adopt the shot noise process used by Cox and Isham (1980), given by

$$\lambda_t = \lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t - S_i)},$$

where

- λ_0 is the initial value of λ_t;

- $\{Y_i\}_{i=1,2,\ldots}$ is a sequence of independent and identically distributed random variables with distribution function $G(y)$, $y > 0$, and $\mathrm{E}(Y_i) = \mu_1$;

- $\{S_i\}_{i=1,2,\ldots}$ is the sequence representing the event times of a Poisson process M_t with constant intensity ρ;

- δ is the rate of exponential decay.

We also define the aggregate loss process

$$C_t = \sum_{i=1}^{N_t} \aleph_i,$$

where N_t is as defined above and $\{\aleph_i\}_{i=1,2,\ldots}$ is a sequence of independent and identically distributed random variables representing the claim sizes, with distribution function $H(u)$, $u > 0$, and $m_1 = \int_0^\infty u \, dH(u)$. We assume that the Poisson process M_t and the sequences $\{Y_i\}_{i=1,2,\ldots}$ and $\{\aleph_i\}_{i=1,2,\ldots}$ are independent of each other.

As λ_t is a Markov process, the action of the generator A of the process (λ_t, t) on a function $f(\lambda, t)$ belonging to its domain is given by

$$A f(\lambda, t) = \frac{\partial f}{\partial t} - \delta \lambda \frac{\partial f}{\partial \lambda} + \rho \left\{ \int_0^\infty f(\lambda + y, t) \, dG(y) - f(\lambda, t) \right\}. \tag{3}$$

For $f(\lambda, t)$ to belong to the domain of the generator A, it is sufficient that $f(\lambda, t)$ is differentiable with respect to λ and t for all λ and t and that $|\int_0^\infty f(\lambda + y, t) \, dG(y) - f(\lambda, t)| < \infty$.

Now let us derive the mean and variance of λ_t, assuming that λ_0 is given.

Theorem 1. *The expectation of the claim intensity process λ_t, assuming that we know λ_0, is given by*

$$\mathrm{E}(\lambda_t \mid \lambda_0) = \frac{\mu_1 \rho}{\delta} + \left(\lambda_0 - \frac{\mu_1 \rho}{\delta} \right) e^{-\delta t}. \tag{4}$$

Proof. Set $f(\lambda, t) = \lambda$ in (3). Then,

$$A\lambda = -\delta\lambda + \mu_1\rho.$$

From $\mathrm{E}(\lambda_t \mid \lambda_0) - \lambda_0 = \mathrm{E}(\int_0^t \{Af(\lambda_s) \mid \lambda_0\}\, \mathrm{d}s)$,

$$\mathrm{E}(\lambda_t \mid \lambda_0) = \lambda_0 - \delta \int_0^t \mathrm{E}(\lambda_s \mid \lambda_0)\, \mathrm{d}s + \int_0^t \mu_1\rho\, \mathrm{d}s.$$

Differentiating with respect to t gives

$$\frac{\mathrm{d}\,\mathrm{E}(\lambda_t \mid \lambda_0)}{\mathrm{d}t} = -\delta\,\mathrm{E}(\lambda_t \mid \lambda_0) + \int_0^t \mu_1\rho\, \mathrm{d}s.$$

Solving this differential equation, we have

$$\mathrm{E}(\lambda_t \mid \lambda_0) = \frac{\mu_1\rho}{\delta} + \left(\lambda_0 - \frac{\mu_1\rho}{\delta}\right)\mathrm{e}^{-\delta t}.$$

Lemma 1. *The second moment of the claim intensity process λ_t is given by*

$$\mathrm{E}(\lambda_t^2 \mid \lambda_0) = \lambda_0^2 \mathrm{e}^{-2\delta t} + \frac{2\mu_1\rho}{\delta}\left(\lambda_0 - \frac{\mu_1\rho}{\delta}\right)(\mathrm{e}^{-\delta t} - \mathrm{e}^{-2\delta t}) + \left(\frac{\mu_1^2\rho^2}{\delta^2} + \frac{\mu_2\rho}{\delta}\right)(1 - \mathrm{e}^{-2\delta t}),$$

where $\mu_2 = \int_0^\infty y^2\, \mathrm{d}G(y)$.

Proof. Set $f(\lambda, t) = \lambda^2$ in (3). Then, from the proof of the previous theorem, the result follows immediately.

Corollary 1. *The variance of the claim intensity process λ_t is given by*

$$\mathrm{var}(\lambda_t \mid \lambda_0) = (1 - \mathrm{e}^{-2\delta t})\mu_2\rho/2\delta. \tag{5}$$

Proof. From $\mathrm{var}(\lambda_t \mid \lambda_0) = \mathrm{E}(\lambda_t^2 \mid \lambda_0) - \{\mathrm{E}(\lambda_t \mid \lambda_0)\}^2$, the result follows immediately.

Similarly, the asymptotic (stationary) mean and variance of λ_t can be obtained from Theorem 1 and Corollary 1.

Corollary 2. *If λ_t is stationary, that is, if λ_0 has the stationary distribution, then*

$$\mathrm{E}(\lambda_t) = \frac{\mu_1\rho}{\delta}$$

and

$$\mathrm{var}(\lambda_t) = \frac{\mu_2\rho}{2\delta}.$$

Proof. Let $t \to \infty$ in (4) and (5). Then, the results follow immediately.

From (2), we have

$$\mathrm{E}(N_t) = \mathrm{E}\left(\int_0^t \lambda_s\, \mathrm{d}s\right) = \mathrm{E}(X_t),$$

where $X_t = \int_0^t \lambda_s\, \mathrm{d}s$ (the aggregated process). Hence, assuming that λ_t is stationary, the expectation of the claim number process N_t is given by

$$\mathrm{E}(N_t) = \frac{\mu_1\rho}{\delta}t. \tag{6}$$

Similarly, assuming that we know λ_0, we can also obtain the variance of the aggregated process X_t,

$$\text{var}\left(\int_0^t \lambda_s \, ds \mid \lambda_0\right) = \text{var}(X_t \mid \lambda_0) = \left\{\frac{\mu_2}{\delta^2}t - \frac{2\mu_2}{\delta^3}(1 - e^{-\delta t}) + \frac{\mu^2}{2\delta^3}(1 - e^{-2\delta t})\right\}\rho, \quad (7)$$

and, assuming that λ_t is stationary, we have

$$\text{var}\left(\int_0^t \lambda_s \, ds\right) = \text{var}(X_t) = \left(\frac{\mu_2}{\delta^2}t - \frac{\mu_2}{\delta^3}e^{-\delta t} - \frac{\mu_2}{\delta^3}\right)\rho. \quad (8)$$

The reason for our explicit derivation of the variance of X_t will become apparent later, when we transform and approximate the Cox and shot noise processes.

In the case of small ρ, the rate of primary-event arrival, Dassios and Jang (2003) used the shot noise process as an intensity function for catastrophic events. However, if the rate of primary-event arrival is large, primary events are no longer considered to be catastrophes. Therefore, we can consider it to be an intensity function to generate the number of claims due to common events of high frequency, such as car accidents or accidents from a large collective insurance portfolio, rather than catastrophic events.

3. Convergence results

We start by introducing the following linear transformations of the processes λ_t, N_t, and C_t:

$$Z_t^{(\rho)} = \frac{\lambda_t - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} \Leftrightarrow \lambda_t = \frac{\mu_1\rho}{\delta} + Z_t^{(\rho)}\sqrt{\frac{\mu_2\rho}{2\delta}},$$

$$W_t^{(\rho)} = \frac{N_t - (\mu_1\rho/\delta)t}{\sqrt{\mu_2\rho/2\delta}} \Leftrightarrow N_t = \frac{\mu_1\rho}{\delta}t + W_t^{(\rho)}\sqrt{\frac{\mu_2\rho}{2\delta}},$$

$$U_t^{(\rho)} = \frac{C_t - m_1(\mu_1\rho/\delta)t}{\sqrt{\mu_2\rho/2\delta}} \Leftrightarrow C_t = m_1\frac{\mu_1\rho}{\delta}t + U_t^{(\rho)}\sqrt{\frac{\mu_2\rho}{2\delta}}.$$

Let us continue with a proposition by Ethier and Kurtz (1986).

Proposition 1. *For $n = 1, 2, \ldots$, let $\{\mathfrak{I}_t^n\}$ be a filtration and let M_n be an $\{\mathfrak{I}_t^n\}$-local martingale with sample paths in $D_{\mathbb{R}^d}[0, \infty)$ and $M_n(0) = 0$, where $\mathbb{R}$ is the set of real numbers and $D_{\mathbb{R}^d}[0, \infty)$ is the space of right-continuous functions from $\mathbb{R}^d$ to $\mathbb{R}$ with left limits. Let $A_n = ((A_n^{ij}))$ be symmetric $(d \times d)$-matrix-valued processes such that A_n^{ij} has sample paths in $D_{\mathbb{R}^d}[0, \infty)$ and $A_n(t) - A_n(s)$ is nonnegative definite for $0 \leq s < t$. Assume that*

$$\lim_{n \to \infty} E\left(\sup_{t \leq T}|A_n^{ij}(t) - A_n^{ij}(t-)|\right) = 0, \quad (9)$$

$$\lim_{n \to \infty} E\left(\sup_{t \leq T}|M_n(t) - M_n(t-)|^2\right) = 0, \quad (10)$$

and, for $i, j = 1, 2, \ldots, d$, that

$$M_n^i(t)M_n^j(t) - A_n^{ij}(t)$$

is an $\{\mathfrak{I}_t^n\}$-local martingale. If, for each $t \geq 0$ and $i, j = 1, 2, \ldots, d$,

$$A_n^{ij}(t) \xrightarrow{\text{P}} c_{ij}(t),$$

where '$\xrightarrow{\text{P}}$' denotes convergence in probability and $C = ((c_{ij}))$ is a continuous, symmetric, $(d \times d)$-matrix-valued function, defined on $[0, \infty)$, satisfying $C(0) = 0$ and

$$\sum_{\substack{i=1,\dots,d \\ j=1,\dots,d}} (c_{ij}(t) - c_{ij}(s))\xi_i \xi_j \geq 0, \qquad \xi \in \mathbb{R}^d,$$

then

$$M_n \Rightarrow X$$

in law, where X is a process with independent Gaussian increments such that $X_i X_j - c_{ij}$ are (local) martingales with respect to $\{\mathfrak{I}_t^n\}$.

Let us now define

$$V_t^{(\rho)} = \frac{J_t - \mu_1 \rho t}{\sqrt{\mu_2 \rho / 2\delta}}, \quad L_t^{(\rho)} = \frac{N_t - \int_0^t \lambda_s \, \mathrm{d}s}{\sqrt{\mu_2 \rho / 2\delta}} = \frac{N_t - X_t}{\sqrt{\mu_2 \rho / 2\delta}}, \quad \text{and} \quad Q_t^{(\rho)} = \frac{C_t - m_1 N_t}{\sqrt{\mu_2 \rho / 2\delta}},$$

where $J_t = \sum_{i=1}^{M_t} Y_i$.

Lemma 2. *Assuming that $\rho \to \infty$,*

$$\begin{bmatrix} V_t^{(\rho)} \\ L_t^{(\rho)} \\ Q_t^{(\rho)} \end{bmatrix} \Rightarrow \begin{bmatrix} \sqrt{2\delta} B_t^{(1)} \\ \sqrt{2\mu_1/\mu_2} B_t^{(2)} \\ \sqrt{k_2(2\mu_1/\mu_2)} B_t^{(3)} \end{bmatrix}$$

in law, where $B_t^{(1)}$, $B_t^{(2)}$, and $B_t^{(3)}$ are three independent standard Brownian motions and $k_2 = \int_0^\infty u^2 \, \mathrm{d}H(u) - (\int_0^\infty u \, \mathrm{d}H(u))^2$ (the variance of claim sizes).

Proof. The generator A of the process $V_t^{(\rho)}$ acting on a function $f(v)$ is given by

$$Af(v) = -\frac{\mu_1 \rho}{\sqrt{\mu_2 \rho / 2\delta}} \frac{\partial f}{\partial v} + \rho \left\{ \int_0^\infty f\left(v + \frac{y}{\sqrt{\mu_2 \rho / 2\delta}}\right) \mathrm{d}G(y) - f(v) \right\}. \tag{11}$$

Setting $f(v) = v^2$ in (11) gives

$$Av^2 = 2\delta.$$

The generator of the process $(X_t, N_t, C_t, \lambda_t, J_t, t)$ acting on a function $f(x, n, c, \lambda, j, t)$ is given by

$$Af(x, n, c, \lambda, j, t)$$
$$= \frac{\partial f}{\partial t} + \lambda \frac{\partial f}{\partial x} + \lambda \left\{ \int_0^\infty f(x, n+1, c+u, \lambda, j, t) \, \mathrm{d}H(u) - f(x, n, c, \lambda, j, t) \right\}$$
$$- \delta \lambda \frac{\partial f}{\partial \lambda} + \rho \left\{ \int_0^\infty f(x, n, c, \lambda+y, j+y, t) \, \mathrm{d}G(y) - f(x, n, c, \lambda, j, t) \right\}. \tag{12}$$

Clearly, for $f(x, n, c, \lambda, j, t)$ to belong to the domain of the generator A, it is essential that $f(x, n, c, \lambda, j, t)$ is differentiable with respect to x, c, λ, and t for all x, n, c, λ, j, and t, and that

$$\left| \int_0^\infty f(\cdot, \cdot, \cdot, \lambda + y, \cdot, \cdot) \, \mathrm{d}G(y) - f(\cdot, \cdot, \cdot, \lambda, \cdot, \cdot) \right| < \infty,$$

$$\left| \int_0^\infty f(\cdot, \cdot, c + u, \cdot, \cdot, \cdot) \, \mathrm{d}H(u) - f(\cdot, \cdot, c, \cdot, \cdot, \cdot) \right| < \infty.$$

Setting

$$f(x, n, c, \lambda, j, t) = \left(\frac{n - x}{\sqrt{\mu_2 \rho / 2\delta}} \right)^2 \quad \text{and} \quad f(x, n, c, \lambda, j, t) = \left(\frac{c - m_1 n}{\sqrt{\mu_2 \rho / 2\delta}} \right)^2$$

in (12) gives

$$A \left(\frac{n - x}{\sqrt{\mu_2 \rho / 2\delta}} \right)^2 = \frac{2\delta}{\mu_2} \frac{\lambda}{\rho} \quad \text{and} \quad A \left(\frac{c - m_1 n}{\sqrt{\mu_2 \rho / 2\delta}} \right)^2 = k_2 \frac{2\delta}{\mu_2} \frac{\lambda}{\rho},$$

where $m_2 = \int_0^\infty u^2 \, dH(u)$ and, so, $k_2 = m_2 - m_1^2$.

Now, $f(X_t) - \int_0^t A f(X_s) \, ds$ is a martingale; therefore, Af is the solution to the 'martingale problem'. Hence, from Proposition 1,

$$\{V_t^{(\rho)}\}^2 - 2\delta t, \quad \{L_t^{(\rho)}\}^2 - \int_0^t \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds, \quad \text{and} \quad \{Q_t^{(\rho)}\}^2 - \int_0^t k_2 \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds$$

are martingales.

It is trivial to check condition (9), since

$$2\delta t, \quad \int_0^t \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds, \quad \text{and} \quad \int_0^t k_2 \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds$$

are continuous (they are proportional to either t or $\int_0^t \lambda_s \, ds$). To check condition (10), we have to look at the jumps of the processes $V_t^{(\rho)}$, $L_t^{(\rho)}$, and $Q_t^{(\rho)}$. Firstly, $L_t^{(\rho)}$ satisfies the condition rather trivially, since its jumps are the jumps of $N_t (\mu_2 \rho / 2\delta)^{-1/2}$ and they are always of size $(\mu_2 \rho / 2\delta)^{-1/2}$ and, so, converge to 0 as $\rho \to \infty$. For $V_t^{(\rho)}$ and $Q_t^{(\rho)}$, we have to check the jumps of $J_t^2 (\mu_2 \rho / 2\delta)^{-1/2}$ and $C_t^2 (\mu_2 \rho / 2\delta)^{-1/2}$. The jumps have finite expectation, since we have assumed that they have second moments. We must prove, for a sequence of nonnegative, independent, and identically distributed random variables $Z_1, Z_2, \ldots$ with finite mean, that

$$\lim_{n \to \infty} \frac{E(U_n)}{n} = 0,$$

where $U_n = \max(Z_1, Z_2, \ldots, Z_n)$. In fact, we can prove that $U_n / n \to 0$. Let F be the distribution function of Z_i. Then, for any $\varepsilon > 0$,

$$P \left(\frac{U_n}{n} > \varepsilon \right) = P(U_n > n\varepsilon) = 1 - (F(n\varepsilon))^n = 1 - (1 - \bar{F}(n\varepsilon))^n,$$

where $\bar{F}$ is the tail of the distribution of Z_i. However, since the random variables have a finite expectation, $n \bar{F}(n\varepsilon) \to 0$; hence, $(1 - \bar{F}(n\varepsilon))^n \to 1$. Now, secondly, for $C_t^2 / (\mu_2 \rho / 2\delta)$, consider the number of jumps $M^{(\rho)}$ of the Poisson process M_t in the interval $[0, T]$. From above, $C_t^2 / M^{(\rho)} \to 0$ and, since $M^{(\rho)} / \rho \to T$, condition (10) is satisfied. Finally, for $J_t^2 / (\mu_2 \rho / 2\delta)$, consider the number of jumps $N^{(\rho)}$ of the process N_t. Clearly, $J_t^2 / N^{(\rho)} \to 0$ from above and $N^{(\rho)} / \rho \to (\mu_1 / \delta) T$.

As can be seen from (7) (see also (8)), $\mathrm{var}(\int_0^t \lambda_s \, ds) = K(t)\rho$. Therefore, by Chebyshev's inequality, as $\rho \to \infty$,

$$P \left\{ \left| \int_0^t \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds - \frac{2\mu_1}{\mu_2} t \right| > \varepsilon \right\} \leq \frac{(2\delta/\mu_2)^2 \, \mathrm{var}(\int_0^t \lambda_s \, ds)}{\rho^2 \varepsilon^2} = \frac{(2\delta/\mu_2)^2 K(t)\rho}{\rho^2 \varepsilon^2} \to 0 \quad (13)$$

and

$$P\left\{\left|k_2 \int_0^t \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds - k_2 \frac{2\mu_1}{\mu_2} t\right| > \varepsilon\right\} \le \frac{k_2^2 (2\delta/\mu_2)^2 \, \text{var}(\int_0^t \lambda_s \, ds)}{\rho^2 \varepsilon^2} = \frac{k_2^2 (2\delta/\mu_2)^2 K(t)\rho}{\rho^2 \varepsilon^2} \to 0. \tag{14}$$

Therefore, from (13) and (14),

$$\int_0^t \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds \xrightarrow{\text{P}} \frac{2\mu_1}{\mu_2} t$$

and

$$\int_0^t k_2 \frac{2\delta}{\mu_2} \frac{\lambda_s}{\rho} \, ds \xrightarrow{\text{P}} k_2 \frac{2\mu_1}{\mu_2} t.$$

Set

$$f(x, n, c, \lambda, j, t) = \left(\frac{n - x}{\sqrt{\mu_2 \rho/2\delta}}\right)\left(\frac{j - \mu_1 \rho t}{\sqrt{\mu_2 \rho/2\delta}}\right),$$

$$f(x, n, c, \lambda, j, t) = \left(\frac{c - m_1 n}{\sqrt{\mu_2 \rho/2\delta}}\right)\left(\frac{j - \mu_1 \rho t}{\sqrt{\mu_2 \rho/2\delta}}\right),$$

and

$$f(x, n, c, \lambda, j, t) = \left(\frac{c - m_1 n}{\sqrt{\mu_2 \rho/2\delta}}\right)\left(\frac{n - x}{\sqrt{\mu_2 \rho/2\delta}}\right)$$

in turn in (12). Then,

$$A\left(\frac{n - x}{\sqrt{\mu_2 \rho/2\delta}}\right)\left(\frac{j - \mu_1 \rho t}{\sqrt{\mu_2 \rho/2\delta}}\right) = 0,$$

$$A\left(\frac{c - m_1 n}{\sqrt{\mu_2 \rho/2\delta}}\right)\left(\frac{j - \mu_1 \rho t}{\sqrt{\mu_2 \rho/2\delta}}\right) = 0,$$

$$A\left(\frac{c - m_1 n}{\sqrt{\mu_2 \rho/2\delta}}\right)\left(\frac{n - x}{\sqrt{\mu_2 \rho/2\delta}}\right) = 0.$$

Therefore, from Proposition 1,

$$V_t^{(\rho)} = \frac{J_t - \mu_1 \rho t}{\sqrt{\mu_2 \rho/2\delta}} \Rightarrow \sqrt{2\delta} B_t^{(1)},$$

$$L_t^{(\rho)} = \frac{N_t - \int_0^t \lambda_s \, ds}{\sqrt{\mu_2 \rho/2\delta}} \Rightarrow \sqrt{\frac{2\mu_1}{\mu_2}} B_t^{(2)},$$

and

$$Q_t^{(\rho)} = \frac{C_t - m_1 N_t}{\sqrt{\mu_2 \rho/2\delta}} \Rightarrow \sqrt{k_2 \frac{2\mu_1}{\mu_2}} B_t^{(3)}$$

in law, where $B_t^{(1)}$, $B_t^{(2)}$, and $B_t^{(3)}$ are three independent standard Brownian motions, as required.

Let us now state and prove the main result of this section.

Theorem 2. *Assume that $\rho \to \infty$ and that λ_0 is a random variable, independent of everything else, such that $(\lambda_0 - (\mu_1\rho/\delta))(\mu_2\rho/2\delta)^{-1}$ converges in distribution to Z_0. Then, $Z_t^{(\rho)}$, $W_t^{(\rho)}$, and $U_t^{(\rho)}$ converge in law to Z_t, W_t, and U_t, where*

$$\mathrm{d}Z_t = -\delta Z_t \,\mathrm{d}t + \sqrt{2\delta}\,\mathrm{d}B_t^{(1)}, \qquad \mathrm{d}W_t = Z_t \,\mathrm{d}t + \sqrt{\frac{2\mu_1}{\mu_2}}\,\mathrm{d}B_t^{(2)},$$

$$\mathrm{d}U_t = m_1 \,\mathrm{d}W_t + \sqrt{k_2 \frac{2\mu_1}{\mu_2}}\,\mathrm{d}B_t^{(3)} = m_1 Z_t \,\mathrm{d}t + \sqrt{m_2 \frac{2\mu_1}{\mu_2}}\,\mathrm{d}B_t^{(4)}. \qquad (15)$$

Here, $B_t^{(1)}$, $B_t^{(2)}$, and $B_t^{(3)}$ are three independent standard Brownian motions, and

$$B_t^{(4)} = \frac{m_1\sqrt{2\mu_1/\mu_2}\,B_t^{(2)} + \sqrt{k_2 2\mu_1/\mu_2}\,B_t^{(3)}}{\sqrt{(m_1^2 + k_2)2\mu_1/\mu_2}}$$

(also a standard Brownian motion).

Proof. $Z_t^{(\rho)}$, $W_t^{(\rho)}$, and $U_t^{(\rho)}$ can be written as follows:

$$\begin{aligned}
Z_t^{(\rho)} &= \frac{\lambda_t - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} \\
&= \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}\mathrm{e}^{-\delta t} - \frac{\mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}(1 - \mathrm{e}^{-\delta t}) + \frac{J_t}{\sqrt{\mu_2\rho/2\delta}} - \delta\int_0^t \mathrm{e}^{-\delta(t-u)}\frac{J_u}{\sqrt{\mu_2\rho/2\delta}}\,\mathrm{d}u \\
&= \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}\mathrm{e}^{-\delta t} + \frac{J_t - \mu_1\rho t}{\sqrt{\mu_2\rho/2\delta}} - \delta\int_0^t \mathrm{e}^{-\delta(t-u)}\frac{J_u - \mu_1\rho u}{\sqrt{\mu_2\rho/2\delta}}\,\mathrm{d}u, \qquad (16)
\end{aligned}$$

since $\delta\int_0^t u\mathrm{e}^{-\delta(t-u)}\,\mathrm{d}u = t - (1 - \mathrm{e}^{-\delta t})/\delta$;

$$W_t^{(\rho)} = \frac{N_t - (\mu_1\rho/\delta)t}{\sqrt{\mu_2\rho/2\delta}} = \frac{N_t - \int_0^t \lambda_s \,\mathrm{d}s}{\sqrt{\mu_2\rho/2\delta}} + \int_0^t \frac{\lambda_s - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}\,\mathrm{d}s; \qquad (17)$$

and

$$U_t^{(\rho)} = \frac{C_t - m_1(\mu_1\rho/\delta)t}{\sqrt{\mu_2\rho/2\delta}} = \frac{C_t - m_1 N_t}{\sqrt{\mu_2\rho/2\delta}} + m_1\left(\frac{N_t - (\mu_1\rho/\delta)t}{\sqrt{\mu_2\rho/2\delta}}\right). \qquad (18)$$

Therefore, by the continuous mapping theorem (Billingsley (1968)) and Lemma 2, (16), (17), and (18) converge to

$$\begin{aligned}
Z_t &= Z_0\mathrm{e}^{-\delta t} + \sqrt{2\delta}\left(B_t^{(1)} - \int_0^t \mathrm{e}^{-\delta(t-s)}B_s^{(1)}\,\mathrm{d}s\right) \\
&= Z_0\mathrm{e}^{-\delta t} + \sqrt{2\delta}\int_0^t \mathrm{e}^{-\delta(t-s)}\,\mathrm{d}B_s^{(1)}, \qquad (19)
\end{aligned}$$

$$W_t = \int_0^t Z_s \,\mathrm{d}s + \sqrt{\frac{2\mu_1}{\mu_2}}B_t^{(2)}, \qquad (20)$$

and

$$U_t = m_1 W_t + \sqrt{k_2 \frac{2\mu_1}{\mu_2}} B_t^{(3)}, \tag{21}$$

respectively.

From (20) and (21), we have

$$dU_t = m_1 \, dW_t + \sqrt{k_2 \frac{2\mu_1}{\mu_2}} B_t^{(3)} = m_1 Z_t \, dt + m_1 \sqrt{\frac{2\mu_1}{\mu_2}} \, dB_t^{(2)} + \sqrt{k_2 \frac{2\mu_1}{\mu_2}} \, dB_t^{(3)}.$$

Since the sum of two independent standard Brownian motions is also a standard Brownian motion, this completes the proof of the theorem.

Theorem 2 implies that Z_t, W_t, and U_t are normally distributed. Therefore, we can define $\tilde{\lambda}_t$, $\tilde{N}_t$, and $\tilde{C}_t$ as Gaussian approximations of λ_t, N_t, and C_t, as follows:

$$\tilde{\lambda}_t = \frac{\mu_1 \rho}{\delta} + Z_t \sqrt{\frac{\mu_2 \rho}{2\delta}} \Leftrightarrow Z_t = \frac{\tilde{\lambda}_t - \mu_1 \rho/\delta}{\sqrt{\mu_2 \rho/2\delta}},$$

$$\tilde{N}_t = \frac{\mu_1 \rho}{\delta} + W_t \sqrt{\frac{\mu_2 \rho}{2\delta}} \Leftrightarrow W_t = \frac{\tilde{N}_1 - (\mu_1 \rho/\delta)t}{\sqrt{\mu_2 \rho/2\delta}}, \tag{22}$$

$$\tilde{C}_t = m_1 \frac{\mu_1 \rho}{\delta} + U_t \sqrt{\frac{\mu_2 \rho}{2\delta}} \Leftrightarrow U_t = \frac{\tilde{C}_t - m_1(\mu_1 \rho/\delta)t}{\sqrt{\mu_2 \rho/2\delta}}. \tag{23}$$

4. The Kalman–Bucy filter and the distribution of Z_t

Let us derive the conditional distribution of Z_t, given $\{W_s, \ 0 \le s \le t\}$, using the Kalman–Bucy filter, where

$$dZ_t = -\delta Z_t \, dt + \sqrt{2\delta} \, dB_t^{(1)} \tag{24}$$

and

$$dW_t = Z_t \, dt + \sqrt{\frac{2\mu_1}{\mu_2}} \, dB_t^{(2)}. \tag{25}$$

To do so, we begin with a proposition from Theorem 6.10, Chapter IV, of Øksendal (1992).

Proposition 2. *The solution* $\hat{Z}_t = \mathrm{E}(Z_t \mid W_s, \ 0 \le s \le t)$ *of the one-dimensional linear filtering problem*

$$dZ_t = F(t) Z_t \, dt + C(t) \, dB_t^{(1)}, \qquad F(t), C(t) \in \mathbb{R},$$

$$dW_t = G(t) Z_t \, dt + D(t) \, dB_t^{(2)}, \qquad G(t), D(t) \in \mathbb{R},$$

satisfies the stochastic differential equation

$$d\hat{Z}_t = \left\{ F(t) - \frac{G^2(t) S(t)}{D^2(t)} \right\} \hat{Z}_t \, dt + \frac{G(t) S(t)}{D^2(t)} \, dW_t, \qquad \hat{Z}_0 = \mathrm{E}(Z_0), \tag{26}$$

where $S(t) = \mathrm{E}((Z_t - \hat{Z}_t)^2)$ *satisfies the Riccati equation:*

$$\frac{dS}{dt} = 2F(t) S(t) - \frac{G^2(t)}{D^2(t)} S^2(t) + C^2(t), \qquad S(0) = \mathrm{E}(\{Z_0 - \mathrm{E}(Z_0)\}^2) = \mathrm{var}(Z_0). \tag{27}$$

Theorem 3. *Let (Z_t, W_t) be a two-dimensional normal process satisfying the system of equations given by (24) and (25). Then, the estimate of Z_t based on the observations $\{W_s, \ 0 \le s \le t\}$ is*

$$\hat{Z}_t = \mathrm{E}(Z_t \mid W_s, \ 0 \le s \le t) = \exp\left\{\int_0^t \Psi(s)\,\mathrm{d}s\right\}\hat{Z}_0 + \frac{\mu_2}{2\mu_1}\int_0^t \exp\left\{\int_s^t \Psi(u)\,\mathrm{d}u\right\}S(s)\,\mathrm{d}W_s,$$

(28)

where

$$S(s) = \frac{\xi(1 + \eta(s))}{\eta(s) - 1} - 2\delta\frac{\mu_1}{\mu_2}$$

and

$$\Psi(s) = -\frac{\xi(1 + \eta(s))}{(2\mu_1/\mu_2)(\eta(s) - 1)},$$

with

$$\xi = \sqrt{\frac{2\mu_1}{\mu_2}}\sqrt{\delta\left(\frac{2\delta\mu_1}{\mu_2} + 2\right)},$$

$$\eta(s) = \frac{a^2 + 2\delta\mu_1/\mu_2 + \sqrt{2\mu_1/\mu_2}\sqrt{\delta(2\delta\mu_1/\mu_2 + 2)}}{a^2 + 2\delta\mu_1/\mu_2 - \sqrt{2\mu_1/\mu_2}\sqrt{\delta(2\delta\mu_1/\mu_2 + 2)}}$$
$$\times \exp\left\{\frac{\sqrt{2\mu_1/\mu_2}\sqrt{\delta(2\delta\mu_1/\mu_2 + 2)}}{\mu_1/\mu_2}s\right\},$$

and $S(0) = a^2$.

Proof. Let $S(0) = a^2$. Then, the Riccati equation (27) has the solution

$$S(t) = \frac{\xi(1 + \varphi)}{\varphi - 1} - 2\delta\frac{\mu_1}{\mu_2},$$

(29)

where ξ is as given above and $\varphi = \eta(t)$. Therefore, from (26), (29) offers a solution for $\hat{Z}_t$ of the form

$$\hat{Z}_t = \mathrm{E}(Z_t \mid W_s, \ 0 \le s \le t) = \exp\left\{\int_0^t \Psi(s)\,\mathrm{d}s\right\}\hat{Z}_0 + \frac{\mu_2}{2\mu_1}\int_0^t \exp\left\{\int_s^t \Psi(u)\,\mathrm{d}u\right\}S(s)\,\mathrm{d}W_s,$$

where

$$\Psi(s) = -\frac{\xi(1 + \eta(s))}{(2\mu_1/\mu_2)(\eta(s) - 1)},$$

as required.

Now we can easily obtain the conditional distribution of Z_t, given $\{W_s, \ 0 \le s \le t\}$, as we have obtained $\mathrm{E}(Z_t \mid W_s, \ 0 \le s \le t) = \hat{Z}_t$.

Corollary 3. *Let Z_t, W_t, $\hat{Z}_t$, and $S(t)$ be as defined above. Then, the conditional distribution of Z_t, given $\{W_s, \ 0 \le s \le t\}$, is given by*

$$\mathrm{E}(\mathrm{e}^{-\gamma Z_t} \mid W_s, \ 0 \le s \le t) = \exp\{-\gamma\hat{Z}_t + \tfrac{1}{2}\gamma^2 S(t)\}.$$

Proof. From Theorem 3 and the facts that $\mathrm{var}(Z_t \mid W_s, \ 0 \le s \le t) = S(t)$ and Z_t is normally distributed, the result follows immediately.

It will be interesting to examine the filtering problem for the Cox process with shot noise intensity when the rate of primary-event arrival, ρ, is small. It should be necessary for us to remove the linearity and derive a nonlinear filter to obtain premiums for primary events of low frequency, using numerical techniques.

5. Pricing of a reinsurance contract using the Kalman–Bucy filter

We have transformed and approximated λ_t and N_t as normal variables Z_t and W_t, from which we have obtained the conditional distribution of Z_t, given $\{W_s, \ 0 \le s \le t\}$. Now, let us derive the pricing model for a stop-loss reinsurance contract using the normal variables Z_t and W_t. As mentioned earlier, as we have assumed that $\rho \to \infty$, this approach can be used for the pricing of common events of high frequency, such as car accidents or accidents from a large collective insurance portfolio.

Let $\aleph_i, \ i \ = \ 1, 2, \ldots$, be the claim amounts, which are assumed to be independent and identically distributed with distribution function $H(u)$. The actuarial stop-loss reinsurance premium at time t is

$$\mathrm{E}\left(\left(\sum_{i=1}^{N_T - N_t} \aleph_i - b\right)^+ \ \bigg| \ N_s, \ 0 \le s \le t\right), \tag{30}$$

where b is a suitably large retention limit. In particular, we define β such that

$$b = \sqrt{\frac{\mu_2 \rho}{2\delta}}\beta + m_1 \frac{\mu_1 \rho}{\delta}(T - t).$$

Let $C_T - C_t$ be the total number of claims between times T and t. Then, from (30), the stop-loss reinsurance premium at time t becomes

$$\mathrm{E}(\{(C_T - C_t) - b\}^+ \mid N_s, \ 0 \le s \le t). \tag{31}$$

Since we have obtained $\tilde{C}_t$ and $\tilde{N}_t$ (see (22) and (23)), which are the Gaussian approximations of C_t and N_t, we will employ them here. Substituting (23) into (31) gives

$$\mathrm{E}(\{(\tilde{C}_T - \tilde{C}_t) - b\}^+ \mid \tilde{N}_s, 0 \le s \le t) = \sqrt{\frac{\mu_2 \rho}{2\delta}}\,\mathrm{E}(\{U_T - U_t - \beta\}^+ \mid W_s, 0 \le s \le t). \tag{32}$$

Let us derive the expectation and variance of $U_T - U_t$, as they need to be determined in order to obtain the stop-loss reinsurance premium based on (32).

Lemma 3. *The expectation of $U_T - U_t$ is given by*

$$\Upsilon = \mathrm{E}(U_T - U_t \mid W_s, \ 0 \le s \le t) = m_1 \frac{1 - \mathrm{e}^{-\delta(T-t)}}{\delta}\hat{Z}_t \tag{33}$$

and the variance of $U_T - U_t$ is given by

$$\begin{aligned}
\Sigma &= \mathrm{var}(U_T - U_t \mid W_s, \ 0 \le s \le t) \\
&= 2\left(\frac{m_1^2}{\delta} + \frac{m_2 \mu_1}{\mu_2}\right)(T - t) \\
&\quad + \left(\frac{m_1}{\delta}\right)^2 [\{1 - \mathrm{e}^{-\delta(T-t)}\}^2 S(t) - \mathrm{e}^{-2\delta(T-t)} + 4\mathrm{e}^{-\delta(T-t)} - 3].
\end{aligned} \tag{34}$$

Proof. From (15),

$$U_T - U_t = m_1 \int_t^T Z_s \, \mathrm{d}s + \sqrt{m_2 \frac{2\mu_1}{\mu_2}} \int_t^T \mathrm{d}B_s^{(4)}. \tag{35}$$

Substituting (19) into (35) then gives

$$U_T - U_t = m_1 \frac{1 - \mathrm{e}^{-\delta(T-t)}}{\delta} Z_t + m_1 \sqrt{2\delta} \int_t^T \frac{1 - \mathrm{e}^{-\delta(T-u)}}{\delta} \, \mathrm{d}B_u^{(1)} + \sqrt{m_2 \frac{2\mu_1}{\mu_2}} \int_t^T \mathrm{d}B_s^{(4)}. \tag{36}$$

Taking the expectation of (36) results immediately in (33). Also, from

$$\mathrm{var}(Z_t \mid W_s, \ 0 \le s \le t) = S(t),$$

we recover (34).

We can now easily find the stop-loss reinsurance premium at time t based on the observations $\{W_s, \ 0 \le s \le t\}$.

Theorem 4. *The stop-loss reinsurance premium at time t based on the observations $\{W_s, 0 \le s \le t\}$ is given by*

$$\mathrm{E}(\{(\tilde{C}_T - \tilde{C}_t) - b\}^+ \mid W_s, \ 0 \le s \le t) = \sqrt{\frac{\mu_2 \rho \Sigma}{4\delta\pi}} \mathrm{e}^{-L^2/2} + \sqrt{\frac{\mu_2 \rho}{2\delta}}(\Upsilon - \beta)\Phi(-L), \tag{37}$$

where $L = (\beta - \Upsilon)\Sigma^{-1/2}$ and $\Phi(\cdot)$ is the cumulative normal distribution function.

Proof. From (32), we have

$$\mathrm{E}((U_T - U_t - \beta)^+ \mid W_s, \ 0 \le s \le t) = \int_\beta^\infty (\upsilon - \beta) \frac{1}{\sqrt{2\pi\Sigma}} \mathrm{e}^{-(\upsilon - \Upsilon)^2/2\Sigma} \, \mathrm{d}\upsilon. \tag{38}$$

Setting $y = (\upsilon - \Upsilon)\Sigma^{-1/2}$ in (38) and multiplying both sides by $(\mu_2\rho/2\delta)^{1/2}$ gives (37).

The following example illustrates the calculation of premiums for a stop-loss reinsurance contract for high-frequency events, using the pricing model derived above.

Example. The numerical values used to simulate the claim arrival process are $\delta = 0.5$, $\lambda_0 = 200$. We will assume that $\rho = 100$, i.e. that the interarrival time between jumps is exponential with mean 0.01 and that the jump size is exponentially distributed with mean 1, i.e. $Y_i \sim \mathrm{Exp}(1)$. S-PLUS® was used to generate random values and to simulate the claim arrival process. The numerical values used to calculate (28) and (37) are

$$\hat{Z}_0 = 0, \qquad\qquad\qquad S(0) = 0,$$
$$\mu_1 = 1, \qquad\qquad\qquad \mu_2 = 2,$$
$$m_1 = 1, \qquad\qquad\qquad m_2 = 3,$$
$$t = 1, \qquad\qquad\qquad T = 2,$$
$$b = 0, 180, 190, 200, 210, 220, \qquad \theta = 0.1,$$

TABLE 1: Stop-loss reinsurance premiums.

Retention level b	Net reinsurance premium ($\theta = 0$)	Risk reinsurance premium ($\theta = 0.1$)
0	206.21	226.83
180	26.58	29.24
190	18.06	19.87
200	11.00	12.10
210	5.77	6.35
220	2.41	2.66

where $(1+\theta)$ is a security loading factor by which risk premium can be calculated by multiplying it by net premium (37), and where

$$\mathrm{E}(C_T - C_t) = \mathrm{E}(N_T - N_t)\,\mathrm{E}(\aleph_i) = \frac{\mu_1 \rho}{\delta} m_1 = 200.$$

The stop-loss reinsurance premiums for high-frequency events at each retention level b – with and without a relative security loading factor θ – were calculated by computing (28) and (37) using MAPLE® and S-PLUS (with $\hat{Z}_1 = 0.557\,9152$) and are shown in Table 1.

References

AHMED, N. U. (1998). *Linear and Nonlinear Filtering for Scientists and Engineers*. World Scientific, Singapore.

BARTLETT, M. S. (1963). The spectral analysis of point processes. *J. R. Statist. Soc. B* **25,** 264–296.

BEARD, R. E., PENTIKAINEN, T. AND PESONEN, E. (1984). *Risk Theory,* 3rd edn. Chapman and Hall, London.

BENING, E. AND KOROLEV, V. Y. (2002). *Generalised Poisson Models and Their Applications in Insurance and Finance*. VSP, Utrecht.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.

BRÉMAUD, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York.

BÜHLMANN, H. (1970). *Mathematical Methods in Risk Theory*. Springer, Berlin.

ÇINLAR, E. (1975). *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.

CONSUL, P. C. (1989). *Generalized Poisson Distributions*. Marcel Dekker, New York.

COX, D. R. (1955). Some statistical methods connected with series of events. *J. R. Statist. Soc. B* **17,** 129–164.

COX, D. R. AND ISHAM, V. (1980). *Point Processes*. Chapman and Hall, London.

COX, D. R. AND ISHAM, V. (1986). The virtual waiting time and related processes. *Adv. Appl. Prob.* **18,** 558–573.

COX, D. R. AND LEWIS, P. A. W. (1966). *The Statistical Analysis of Series of Events*. Methuen, London.

CRAMÉR, H. (1930). *On the Mathematical Theory of Risk*. Skand. Jubilee Volume, Stockholm.

DASSIOS, A. (1987). Insurance, storage and point process: an approach via piecewise deterministic Markov processes. Doctoral Thesis, Imperial College, London.

DASSIOS, A. AND EMBRECHTS, P. (1989). Martingales and insurance risk. *Commun. Statist. Stoch. Models* **5,** 181–217.

DASSIOS, A. AND JANG, J. (1998). Linear filtering in reinsurance. Working Paper LSERR 41, Department of Statistics, The London School of Economics and Political Science.

DASSIOS, A. AND JANG, J. (2003). Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. *Finance Stoch.* **7,** 73–95.

DAVIS, M. H. A. (1977). *Linear Estimation and Stochastic Control*. Chapman and Hall, London.

DAVIS, M. H. A. (1984). Piecewise deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Statist. Soc. B* **46,** 353–388.

ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.

GERBER, H. U. (1979). *An Introduction to Mathematical Risk Theory*. S. S. Huebner Foundation for Insurance Education, Philadelphia, PA.

GNEDENKO, B. V. AND KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, MA.

GRANDELL, J. (1976). *Doubly Stochastic Poisson Processes*. Springer, Berlin.

GRANDELL, J. (1991). *Aspects of Risk Theory*. Springer, New York.

GRANDELL, J. (1997). *Mixed Poisson Processes*. Chapman and Hall, London.

HAIGHT, F. A. (1967). *Handbook of the Poisson Distribution*. John Wiley, New York.

JANG, J. (1998). Doubly stochastic point processes in reinsurance and the pricing of catastrophe insurance derivatives. Doctoral Thesis, The London School of Economics and Political Science.

JANG, J. (2004). Martingale approach for moments of discounted aggregate claims. *J. Risk Insurance* **71,** 201–211.

KLÜPPELBERG, C. AND MIKOSCH, T. (1995). Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli* **1,** 125–147.

KRUGLOV, V. M. (1976). Method of accompanying infinitely divisible distributions. In *Proc. Third Japan–USSR Symp. Prob. Theory* (Tashkent, 1975; Lecture Notes Math. **550**), Springer, Berlin, pp. 316–323.

LANDO, D. (1994). On Cox processes and credit risky bonds. Preprint, Department of Theoretical Statistics, University of Copenhagen.

LIPSTER, R. S. AND SHIRYAYEV, A. N. (1977). *Statistics of Random Processes*, Vol. 1, *General Theory*. Springer, New York.

LIPSTER, R. S. AND SHIRYAYEV, A. N. (1978). *Statistics of Random Processes*, Vol. 2, *Applications*. Springer, New York.

MEDHI, J. (1982). *Stochastic Processes*. John Wiley, New Delhi.

ØKSENDAL, B. (1992). *Stochastic Differential Equations*. Springer, Berlin.

ROLSKI, T., SCHMIDLI, H., SCHMIDT, V. AND TEUGELS, J. (1999). *Stochastic Processes for Insurance and Finance*. John Wiley, Chichester.

SEAL, H. L. (1983). The Poisson process: its failure in risk theory. *Insurance Math. Econom.* **2,** 287–288.

SERFOZO, R. F. (1972). Conditional Poisson processes. *J. Appl. Prob.* **9,** 288–302.

SNYDER, D. L. (1975). *Random Point Processes*. John Wiley, New York.