January 20, 2010

PREDICTION IN A POISSON CLUSTER MODEL

MUNEYA MATSUI AND THOMAS MIKOSCH

Abstract. We consider a Poisson cluster model which is motivated by insurance applications. At each claim arrival time, modeled by a point of a homogeneous Poisson process, we start a cluster process which represents the number or amount of payments triggered by the arrival of a claim in a portfolio. The cluster process is a Lévy or truncated compound Poisson process. Given the observations on the process over a finite interval we consider the expected value of the number and amount of payments in a future time interval. We also give bounds for the error encountered in this prediction procedure.

1. Introduction

In this paper we consider the model

$$M(t)=\sum_{k=1}^{\infty} I_{\{T_k\le 1\}} L_k(t-T_k)\,,\quad t\ge 1\,, \tag{1.1}$$

where $0<T_1<T_2<\cdots$ are the points of a homogeneous Poisson process with intensity $\lambda>0$ and (L_k) is a sequence of iid stochastic processes independent of (T_k) and such that $L_k(t)=0$ a.s. $t\le 0$. Writing N for the counting process generated by the points T_k, $k=1,2,\ldots$, M takes on the form

$$M(t)=\sum_{k=1}^{N(1)} L_k(t-T_k)\,,\quad t\ge 1\,. \tag{1.2}$$

Processes of this type are related to Poisson shot noise which has found a multitude of applications in rather different areas. For example, in an insurance context, T_k may describe the arrival of a claim in an insurance portfolio and $(L_k(t-T_k))_{t\ge T_k}$ is the corresponding payment process from the insurer to the insured starting at time T_k. Alternatively, $(L_k(t-T_k))_{t\ge T_k}$ can be the counting process for these payments. The main focus of this paper is on applications in an insurance context.

Early on, shot noise processes have been used for modeling purposes in many fields of applied probability such as bunching in traffic (Bartlett [1]), computer failure times (Lewis [17]) and earthquake aftershocks (Vere–Jones [28]). More recently, shot noise processes have been used for modeling large computer networks such as the Internet; see for example Konstantopoulos and Lin [12], Kurtz [13] for some early work. In the context of the workload of large computer networks, shot noise processes arise as aggregated versions of the ON/OFF or infinite source Poisson models, also known as M/G/∞ model; see for example Levy and Taqqu [16], Pipiras and Taqqu [24], Mikosch et al. [22] and for further extensions Faÿ et al. [2], Mikosch and Samorodnitsky [23]. Other applications include finance (Samorodnitsky [25], Klüppelberg and Kühn [8]) and physics (Giraitis et al. [3]). Most papers mentioned above aim at an asymptotic theory for the shot noise process, deriving Gaussian or Lévy process limits, or at the asymptotics of the extremes of such processes; see also Heinrich and Schmidt [4], Hsing and Teugels [5], Klüppelberg and Mikosch [9, 10], Klüppelberg et al. [11], Lane [14, 15], McCormick [20], Stegeman [27].

1991 *Mathematics Subject Classification.* Primary 60K30, Secondary 60G25 60G55.

Key words and phrases. Poisson cluster model, prediction, conditional expectation, claims reserving, insurance, Lévy process, shot noise.

Thomas Mikosch's research is partly supported by the Danish Research Council (FNU) Grants 272-06-0442 and 09-072331. Muneya Matsui's research is supported by a JSPS Research Fellowship for Young Scientists.

The focus of this paper is not on asymptotic properties of the process M defined in (1.1) but on precise results about the prediction of the increments

$$M(t,t+s] = M(t+s) - M(t)\,, \quad t \geq 1\,, s > 0\,,$$

i.e., we will calculate $E[M(t,t+s] \mid \mathcal{F}_t]$ for some suitable σ-fields $\mathcal{F}_t$, where we do not necessarily assume that $M(t,t+s]$ has finite variance. Results of this kind are surprisingly explicit due to the Poisson structure underlying the process M.

The particular structure of the process M and the prediction problem are motivated by reserving problems in insurance. Here the points $T_k \leq 1$ describe the arrivals of claims in a portfolio in a given period, 1 year say, and $M(t)$ is the number or amount of payments for the claims arriving in $[0,1]$ and being paid off in $[0,t]$, $t \geq 1$. Correspondingly, $M(t,t+s]$ is the number or amount of payments executed in the interval $(t,t+s]$, $s > 0$.

Problems of this kind are related to *claims reserving*. Traditionally, claims reserving has dealt with a statistical model assuming conditions of the type $E[M(t+s) \mid M(t)] = f_{t,s}\, M(t)$ for suitable constants $f_{t,s}$ and then one estimates these constants based on several years of data available (so-called *chain ladder*). We refer to Chapter 11 in Mikosch [21] for an introduction to the topic. Jessen et al. [6] consider a stochastic process based model with Poisson components which allows one to derive explicit expressions for the prediction of payment numbers and amounts and the corresponding prediction errors. This model requires that time series of annual observations on payment numbers and amounts are available; the structure of this model is significantly simpler than (1.1).

In the present paper we follow a different path which was already mentioned in Section 11.3.3 of [21]. There the model (1.1) was considered for iid homogeneous Poisson processes L_k, $k = 1, 2, \ldots$, and explicit expressions for the prediction $E[M(t,t+s] \mid M(t)]$ were obtained. First, in Section 2, we show that this approach can be extended to general Lévy processes L_k. The homogeneous Poisson case is again a benchmark (Section 2.3). In this case we can also give a recursive algorithm for determining the prediction. In Section 3 we consider modifications of the model (1.2). Instead of a Lévy process L_k we allow for Lévy processes which are truncated at a random level or which take into account a delay in reporting of a claim. In Section 3 we also consider different σ-fields $\mathcal{F}_t$ and calculate the corresponding predictions $E[(M(t,t+s] \mid \mathcal{F}_t]$ and their conditional or unconditional prediction errors. Depending on $\mathcal{F}_t$ the prediction of $M(t,t+s]$ and the prediction error often get a structure which is simpler than for $E[M(t,t+s] \mid M(t)]$.

2. Prediction in a Poisson cluster process with Lévy clusters

In this section we consider the model (1.1) and we assume that an activity process L_k starts at the Poisson point $T_k \in [0,1]$. In this context, (L_k) is a sequence of iid Lévy processes with generic element L; we refer to Sato [26] for the definition and properties of Lévy processes. Then $M(t,t+s]$, $t \geq 1$, $s > 0$, can be interpreted as the measurement of activities initiated in $[0,1]$ which are still alive in $(t,t+s]$. In the insurance context mentioned in the Introduction, $M(t,t+s]$ has the interpretation as the *number* or *amount of payments* for claims that occurred in one year say. Other interpretations are possible as well. For example, $M(t,t+s]$ may describe the workload to be managed by a large computer network for sources that started an activity (such as sending packets to other sources) in the interval $[0,1]$. The Lévy process condition on L_k is assumed for convenience; in this case we will get explicit expressions for the prediction of $M(t,t+s]$ given $M(t)$, $t \geq 1$; see Section 2.2.

We start by an analysis of the moment structure of M.

2.1. The first and second moments of M.

The following result is elementary.

Lemma 2.1. *Assume the model* (1.2) *with iid Lévy processes L_k, $k = 1, 2, \ldots$, and a homogeneous Poisson process N with intensity $\lambda > 0$.*

(1) *Assume that* $\mu = E[L(1)]$ *exists and is finite. Then*

$$E[M(t)] = \lambda\,\mu\,(t-0.5)\,,\quad t\ge 1\,.$$

(2) *Assume that* $\sigma^2 = \mathrm{Var}(L(1))$ *is finite. Then, for* $1\le t_1\le t_2$,

$$\mathrm{Cov}(M(t_1),M(t_2)) = \lambda\,\sigma^2\,(t_1-0.5) + \lambda\,\mu^2\int_0^1 (t_1-s)(t_2-s)\,ds\,.$$

In particular,

$$\mathrm{Var}(M(t)) = \lambda\,\sigma^2\,(t-0.5) + \lambda\,\mu^2(t^2-t+1/3)\,,\quad t\ge 1\,.$$

We see that the process M is over-dispersed in the sense that $\lim_{t\to\infty}\mathrm{Var}(M(t))/E[M(t)] = \infty$.

2.2. **Prediction of future increments.** In this section we derive explicit expressions for the quantities $E[M(t,t+s]\mid M(t)]$ for $t\ge 1$, $s>0$. We assume the general conditions of Lemma 2.1 and also that $\mu = E[L(1)]$ is finite.

We start with a simple observation. By the definition (1.2) of $M(t)$, $t\ge 1$, the σ-field generated by $M(t)$ is contained in the σ-field generated by $(L_i(t-T_i))$ and (T_i). Therefore, writing $L_i(t,t+s] = L_i(t+s)-L_i(t)$ for the increments of L_i, for $t\ge 1$ and $s>0$,

$$\begin{aligned}
&E[M(t,t+s]\mid M(t)]\\
&= \sum_{k=1}^{\infty} E\Big[E\big[I_{\{T_k\le 1\}}L_k(t-T_k,t+s-T_k]\mid (L_i(t-T_i)),(T_i)\big]|M(t)\Big]\\
&= \sum_{k=1}^{\infty} E\Big[I_{\{T_k\le 1\}}E\big[L_k(t-T_k,t+s-T_k]\mid L_k(t-T_k),T_k\big]\mid M(t)\Big]\\
&= \mu\,s\sum_{k=1}^{\infty} E\left[I_{\{T_k\le 1\}}\mid M(t)\right] = \mu\,s\,E\left[N(1)\mid M(t)\right]\,. \qquad (2.1)
\end{aligned}$$

In the first step we used dominated convergence and in the last one the stationary independent increments of the Lévy process L_k. Thus we are left to calculate $E[N(1)\mid M(t)]$ for $t\ge 1$. We mention at this point that, in an insurance context, the number $N(1)$ of claims arriving in the interval $[0,1]$ is in general not observable at time $t\ge 1$. It usually takes a much longer time period than just one year before the claim number for the first year $[0,1]$ is known (so-called Incurred But Not Reported or IBNR effect). Therefore the quantity $E[N(1)\mid M(t)]$, $t\ge 1$, has the intuitive meaning that one gathers information about the payments in the period $[0,t]$, represented by the quantity $M(t)$, in order to gather information about the (generally unknown) claim number $N(1)$.

It is our aim to express the prediction of $M(t,t+s]$ given $M(t)$ as explicitly as possible. Motivated by the previous calculations, we will determine

$$\widehat{M}_A(t,t+s] \;= E[M(t,t+s]\mid M(t)\in A]\,,\quad t\ge 1\,,s>0\,,$$

for any Borel set A. Later, in Section 2.3, we will specify the set A. The quantity $\widehat{M}_A(t,t+s]$ is the conditional first moment of $M(t,t+s]$. In order to get an idea of the conditional prediction error we are also interested in the conditional second moment of $M(t,t+s]$. Both moments can easily be derived from the characteristic function of $M(t,t+s]$ given $M(t)$.

Lemma 2.2. *Assume the model* (1.2) *with iid Lévy processes* L_k, $k=1,2,\dots$, *and a homogeneous Poisson process* N *with intensity* $\lambda>0$. *For any Borel set* A, *the characteristic function of* $M(t,t+s]$ *given* $\{M(t)\in A\}$ *has the form*

$$\widehat{f}_A(x) = E\left[\mathrm{e}^{\,i\,x\,M(t,t+s]}\mid M(t)\in A\right]$$

$$= \frac{E\left[(E[\mathrm{e}^{i\,x\,L(s)}])^{N(1)}\,P(L(R_{N(1)}(t))\in A\mid N(1))\right]}{P(L(R_{N(1)}(t))\in A)}\,,\quad x\in\mathbb{R}\,,\quad t\ge 1\,,s>0\,.$$

Here we assume that the denominator does not vanish and

$$R_r(t)=\sum_{i=1}^{r}(t-U_i)\,,\quad t\ge 1\,,\quad r=0,1,\dots\,, \tag{2.2}$$

for an iid $U(0,1)$ sequence (U_i) such that L, N and (U_i) are independent.

Remark 2.3. The condition $P(L(R_{N(1)}(t))\in A)>0$ is needed for a proper definition of the conditional characteristic function $\widehat{f}_A(x)$. From the proof below it follows that this condition is equivalent to $P(M(t)\in A)>0$.

Proof. The same argument leading to (2.1) yields for $x\in\mathbb{R}$, $t\ge 1$, $s>0$,

$$\begin{aligned}
&E\left[\mathrm{e}^{i\,x\,M(t,t+s]}\mid M(t)\right]\\
&= E\Big[E\big[\mathrm{e}^{i\,x\,M(t,t+s]}\mid (L_k(t-T_k))\,,\,(T_k)\big]\Big|M(t)\Big]\\
&= E\Big[E\Big[\prod_{k=1}^{\infty}\mathrm{e}^{i\,x\,I_{\{T_k\le 1\}}\,L_k(t-T_k,t+s-T_k]}\mid (L_k(t-T_k))\,,\,(T_k)\Big]\Big|M(t)\Big]\\
&= E\Big[E\Big[\prod_{k=1}^{\infty}\Big(I_{\{T_k>1\}}+I_{\{T_k\le 1\}}\,E[\mathrm{e}^{i\,x\,L(s)}]\Big)\mid (T_k)\Big]\Big|M(t)\Big]
\end{aligned}$$

$$= E\Big[\Big(E[\mathrm{e}^{i\,x\,L(s)}]\Big)^{N(1)}\Big|M(t)\Big]\,. \tag{2.3}$$

Therefore we will calculate the following quantities for any Borel set A (assuming that the denominator does not vanish)

$$\begin{aligned}
P(N(1)=r\mid M(t)\in A) &= \frac{P\Big(N(1)=r,\sum_{k=1}^{N(1)}L_k(t-T_k)\in A\Big)}{P\Big(\sum_{k=1}^{N(1)}L_k(t-T_k)\in A\Big)}\\
&= \frac{P\Big(N(1)=r,L\Big(\sum_{k=1}^{N(1)}(t-T_k)\Big)\in A\Big)}{P\Big(L\Big(\sum_{k=1}^{N(1)}(t-T_k)\Big)\in A\Big)}\,,\quad r=0,1,\dots\,.
\end{aligned}$$

In the last step we used the independent stationary increments of the iid Lévy processes L_i, $i=1,2,\dots$, conditional on (T_k). Conditioning on $N(1)$ and using the order statistics property of a homogeneous Poisson process, we obtain

$$P\Big(\sum_{k=1}^{N(1)}(t-T_k)\in\cdot\mid N(1)=r\Big)=P(R_r(t)\in\cdot),\quad r=0,1,\dots,$$

where R_r is defined in (2.2). We conclude that

$$P(N(1)=r\mid M(t)\in A)=\frac{P(N(1)=r)\,P(L(R_r(t))\in A)}{P(L(R_{N(1)}(t))\in A)}\,. \tag{2.4}$$

Now plug (2.4) in (2.3) to obtain the desired expression for $\widehat{f}_A(x)$. □

Since we know the characteristic function $\widehat{f}_A(x)$, $x \in \mathbb{R}$, of $M(t,t+s]$ given $\{M(t) \in A\}$ we can derive the moments of the prediction $\widehat{M}_A(t,t+s]$ by differentiating $\widehat{f}_A(x)$ at $x=0$ sufficiently often. The following result summarizes the analysis of the first and second conditional moments.

Theorem 2.4. *Assume the model* (1.2) *with iid Lévy processes* L_k, $k=1,2,\dots$, *and a homogeneous Poisson process* N *with intensity* $\lambda > 0$.
(1) *Assume that* $\mu = E[L(1)]$ *exists and is finite. Then the prediction* $\widehat{M}_A(t,t+s]$ *of* $M(t,t+s]$ *given* $\{M(t) \in A\}$ *has the following form for any Borel set* A,

$$\widehat{M}_A(t,t+s] = \mu\, s \frac{E\left[N(1)\, P(L(R_{N(1)}(t)) \in A \mid N(1))\right]}{P(L(R_{N(1)}(t)) \in A)}\,, \quad t \ge 1\,, s > 0\,. \tag{2.5}$$

(2) *Assume that* $\sigma^2 = \mathrm{Var}(L(1))$ *is finite. Then the conditional variance of* $M(t,t+s]$ *given* $\{M(t) \in A\}$ *has the following form for any Borel set* A,

$$\begin{aligned} &\mathrm{Var}(M(t,t+s] \mid M(t) \in A) \\ &= \frac{\sigma^2}{\mu}\,\widehat{M}_A(t,t+s] + (\mu\, s)^2\, \frac{E\left[(N(1))^2\, P(L(R_{N(1)}(t)) \in A \mid N(1))\right]}{P(L(R_{N(1)}(t)) \in A)} - (\widehat{M}_A(t,t+s])^2\,, \\ &\quad t \ge 1\,, s > 0\,. \end{aligned} \tag{2.6}$$

Here we have assumed that the probability $P(L(R_{N(1)}(t)) \in A)$ *does not vanish.*

Remark 2.5. The variance of $M(t,t+s]$ conditional on $M(t)$ gives one a certain measure for the uncertainty of the prediction $\widehat{M}_A(t,t+s]$. In general, this conditional variance is difficult to obtain. However, if L is a homogeneous Poisson process we can derive recursive algorithms for determining this quantity; see Section 2.3. It would be desirable to get an expression for the unconditional mean square error

$$\begin{aligned} E\big[(M(t,t+s] - E[M(t,t+s] \mid M(t)])^2\big] &= E[\mathrm{Var}(M(t,t+s] \mid M(t))] \\ &= E\big[(M(t,t+s])^2\big] - E\big[(E[M(t,t+s] \mid M(t)])^2\big]\,. \end{aligned}$$

The second moment $E[(M(t,t+s])^2]$ provides an upper bound for the mean square error. It can be derived from Lemma 2.1. The quantity $E[(E[M(t,t+s] \mid M(t)])^2]$ seems not tractable.

2.3. **Poisson clusters.** In this section we assume that L is a homogeneous Poisson process with intensity $\gamma > 0$. In this case we can give explicit expressions for the prediction

$$\widehat{M}_m(t,t+s] = E[M(t,t+s] \mid M(t) = m]\,, \quad t \ge 1\,, s > 0\,, \quad m = 0,1,\dots\,,$$

and the conditional prediction error. In what follows, it will be convenient to use the Laplace-Stieltjes transform of any non-negative random variable Y:

$$\phi_Y(z) = E[\mathrm{e}^{-zY}]\,, \quad z \ge 0\,,$$

and its mth derivatives $\phi_Y^{(m)}(z) = (-1)^m\, E[Y^m\, \mathrm{e}^{-zY}]$, $m = 0,1,2,\dots$.

Theorem 2.6. *Assume that* L *is a homogeneous Poisson process with intensity* $\gamma > 0$. *Then the prediction of* $M(t,t+s]$ *given* $\{M(t) = m\}$ *has the form*

$$\widehat{M}_m(t,t+s] = \lambda\,\gamma\, s\, \frac{\phi^{(m)}_{R_{N(1)+1}(t)}(\gamma)}{\phi^{(m)}_{R_{N(1)}(t)}(\gamma)}\,, \quad t \ge 1\,, s > 0\,, \quad m = 0,1,\dots\,, \tag{2.7}$$

and the conditional mean square error is given by

$$\mathrm{Var}(M(t,t+s] \mid M(t) = m)$$

$$= (\lambda\,\gamma\,s)^2\,\frac{\phi^{(m)}_{R_{N(1)+2}}(\gamma)}{\phi^{(m)}_{R_{N(1)}}(\gamma)} + (1+\gamma\,s)\,\widehat{M}_m(t,t+s] - (\widehat{M}_m(t,t+s])^2\,,$$

$$t \ge 1\,, s > 0\,, \quad m = 0,1,\ldots\,,$$

where R_r is defined in (2.2).

Proof. According to (2.5) and (2.6), we need to evaluate $E[(N(1))^i\,P(L(R_{N(1)}(t)) = m \mid N(1))]$, $i = 0,1,2$. We have

$$\begin{aligned}
E[N(1)\,P(L(R_{N(1)}(t)) = m \mid N(1))] &= \sum_{r=1}^{\infty} r\,P(N(1)=r)\,P(L(R_r(t)) = m) \\
&= \lambda \sum_{r=0}^{\infty} \mathrm{e}^{-\lambda}\,\frac{\lambda^r}{r!}\,\frac{E[(\gamma\,R_{r+1}(t))^m\,\mathrm{e}^{-\gamma\,R_{r+1}(t)}]}{m!} \\
&= \lambda\,E\left[\frac{(\gamma\,R_{N(1)+1}(t))^m\,\mathrm{e}^{-\gamma\,R_{N(1)+1}(t)}}{m!}\right] \\
&= \frac{\lambda\,(-\gamma)^m}{m!}\,\phi^{(m)}_{R_{N(1)+1}(t)}(\gamma)\,.
\end{aligned}$$

In a similar way, one calculates

$$\begin{aligned}
P(L(R_{N(1)}(t)) = m) &= \frac{(-\gamma)^m}{m!}\,\phi^{(m)}_{R_{N(1)}(t)}(\gamma)\,, \\
E[(N(1))^2\,P(L(R_{N(1)}(t)) = m \mid N(1))] &= \frac{\lambda\,(-\gamma)^m}{m!}\,\phi^{(m)}_{R_{N(1)+1}(t)}(\gamma) + \frac{\lambda^2\,(-\gamma)^m}{m!}\,\phi^{(m)}_{R_{N(1)+2}(t)}(\gamma)\,.
\end{aligned}$$

This concludes the proof. □

In view of Theorem 2.6 it is crucial to be able to evaluate the derivatives of $\phi_{R_{N(1)}(t)}$. Although we have the representations

$$\phi_{R_{N(1)}}(\gamma) = \mathrm{e}^{-\lambda(1-\phi_{t-U}(\gamma))} \quad \text{and} \quad \phi_{t-U}(\gamma) = \gamma^{-1}\mathrm{e}^{-\gamma t}(\mathrm{e}^{\gamma}-1),$$

their derivatives are complicated and not necessarily useful. Fortunately, the following proposition yields a recursive way of determining these derivatives.

Proposition 2.7. *Let $\ell = 1,2,\ldots$ and $\phi^{(\ell)}_{t-U}(\gamma)$, $\phi^{(\ell)}_{R_{N(1)+r}(t)}(\gamma)$, $r = 0,1,2$ be the ℓth derivatives of $\phi_{t-U}, \phi_{R_{N(1)+r}(t)}$, $r = 0,1,2$. Then the following recursive relations are valid:*

$$\begin{aligned}
\phi^{(\ell)}_{t-U}(\gamma) &= (-1)^\ell \gamma^{-\ell-1}\,[\Gamma(\ell+1,\gamma\,(t-1)) - \Gamma(\ell+1,\gamma\,t)] \\
\phi^{(\ell)}_{R_{N(1)}(t)}(\gamma) &= \lambda \sum_{k=1}^{\ell} \binom{\ell-1}{k-1}\,\phi^{(k)}_{t-U}(\gamma)\,\phi^{(\ell-k)}_{R_{N(1)}(t)}(\gamma) \\
\phi^{(\ell)}_{R_{N(1)+1}(t)}(\gamma) &= \lambda \sum_{k=0}^{\ell} \binom{\ell}{k}\,\phi^{(k)}_{t-U}(\gamma)\,\phi^{(\ell-k)}_{R_{N(1)}(t)}(\gamma) \\
\phi^{(\ell)}_{R_{N(1)+2}(t)}(\gamma) &= \lambda \sum_{k=0}^{\ell} \binom{\ell}{k}\,\phi^{(k)}_{t-U}(\gamma)\,\phi^{(\ell-k)}_{R_{N(1)+1}(t)}(\gamma),
\end{aligned}$$

where $\Gamma(\alpha,x)=\int_x^\infty \mathrm{e}^{-y}\,y^{\alpha-1}\,dy$, $x>0$, *is the incomplete Gamma function.*

Proof. Observe that for $t\ge 1$,

$$\phi_{t-U}(\gamma)=\int_t^\infty \mathrm{e}^{-\gamma u}(\mathrm{e}^{\gamma}-1)du\,. \tag{2.8}$$

Then Leibniz's rule yields for $u>0$ and $\ell=0,1,\ldots$,

$$\begin{aligned}[\mathrm{e}^{-\gamma u}(\mathrm{e}^{\gamma}-1)]^{(\ell)} &= \sum_{k=0}^{\ell}\binom{\ell}{k}(\mathrm{e}^{-\gamma u})^{(k)}(\mathrm{e}^{\gamma}-1)^{(\ell-k)}\\ &= \mathrm{e}^{-\gamma(u-1)}(1-u)^{\ell}-(-u)^{\ell}\mathrm{e}^{-\gamma u}.\end{aligned}$$

Now, interchanging the integral and the derivative in (2.8), we have for $t\ge 1$,

$$\begin{aligned}\phi_{t-U}^{(\ell)}(\gamma) &= \int_t^\infty\Big[\mathrm{e}^{-\gamma\,(u-1)}(1-u)^{\ell}-(-u)^{\ell}\mathrm{e}^{-\gamma u}\Big]\,du\\ &= (-1)^{\ell}\int_{t-1}^{t}\mathrm{e}^{-\gamma u}\,u^{\ell}\,du\\ &= (-1)^{\ell}\,\gamma^{-\ell-1}\,[\Gamma(\ell+1,\gamma\,(t-1))-\Gamma(\ell+1,\gamma\,t)]\,.\end{aligned}$$

This is the desired formula for the derivatives of ϕ_{t-U}. By definition of $\phi_{R_{N(1)}(t)}$ observe that

$$\phi'_{R_{N(1)}(t)}=\lambda\phi'_{t-U}\,\phi_{R_{N(1)}(t)}\,.$$

Another application of Leibniz's rule yields the desired formula for the derivatives of $\phi_{R_{N(1)}(t)}$. We also observe that

$$\phi_{R_{N(1)}(t)+1}=\phi_{R_{N(1)}(t)}\phi_{t-U}\quad\text{and}\quad\phi_{R_{N(1)}(t)+2}=\phi_{R_{N(1)}(t)+1}\phi_{t-U}\,.$$

Applications of the Leibniz rule yield the desired expressions for the derivatives. □

Remark 2.8. Assume that the homogeneous Poisson process L_k has the arrivals $0<\Gamma_{k1}<\Gamma_{k2}<\cdots$, $k=1,2,\ldots$. Let $(X_{ki})_{k,i=1,2,\ldots}$ be an iid sequence of random variables such that $\nu=EX_{11}$ exists and is finite. Moreover, we assume that (T_k), (X_{ki}), $(\Gamma_{ki})_{k,i=1,2,\ldots}$ are mutually independent. Then it is possible to define the reward process

$$S(t)=\sum_{k=1}^{N(1)}\sum_{i=1}^{\infty}X_{ki}\,I_{\{T_k+\Gamma_{ki}\le t\}}\,,\quad t\ge 1\,.$$

The reward $S(t)$ can be interpreted as the total amount of payments executed at the times $T_k+\Gamma_{ki}\le t$, $t\ge 1$, $i=1,2,\ldots$, for claims arriving in an insurance portfolio at times $T_k\le 1$. A conditioning argument shows that

$$E[S(t,t+s]\mid M(t)]=\nu\,E[M(t,t+s]\mid M(t)]\,,\quad t\ge 1\,,s>0\,.$$

The iid-ness condition of the X_{ki}'s can be further relaxed. For example, if one looks at the conditional expectations $E[S(\ell,\ell+1]\mid M(\ell)]$, $\ell=1,2,\ldots$, then one may allow the distribution of the iid payments X_{ki} executed in the interval $(\ell,\ell+1]$ to depend on ℓ. The conditional variance $\mathrm{Var}(S(t,t+s]\mid M(t))$ can be calculated as well. Some of these calculations are provided in Section 11.3.3 of Mikosch [21]. We omit details.

3. Prediction with different information sets and non-Lévy clusters

3.1. **Prediction with compound Poisson clusters.** In this section we assume the model (1.2) with a sequence (L_k) of iid compound Poisson processes. We assume the representation

$$L_k(t)=\sum_{i=1}^\infty X_{ki}\,I_{\{\Gamma_{ki}\le t\}}\,,\quad t\ge 1\,,\tag{3.1}$$

where (X_{ki}) is a double array of iid random variables, independent of the double array (Γ_{ki}) and the Poisson points (T_k). The points $0<\Gamma_{k1}<\Gamma_{k2}<\cdots$ constitute the homogeneous Poisson process N_k with intensity $\gamma>0$ underlying the compound Poisson process L_k. Throughout we assume that $\mu=E[L(1)]=EX_{11}\,\gamma$ exists and is finite.

In Section 2.3 we mentioned that, in an insurance context, it is natural to assume that the number of claims $N(1)$ is unobservable at time $t\ge 1$. It is common that claims get reported long after they were incurred. In this section, we replace the condition $M(t)$ in the prediction of $M(t,t+s]$ for $t\ge 1$ and $s>0$ by the number of claims that were incurred in $[0,1]$ and reported by time $t\ge 1$. We say that the kth claim is reported if $T_k+\Gamma_{k1}\le t$, i.e., if the first payment has been executed by time t. We write for the corresponding counting process of reported claims

$$N_0(t)=\sum_{k=1}^{N(1)} I_{\{T_k+\Gamma_{k1}\le t\}}\ =\sum_{k=1}^{N(1)} I_{\{N_k(t-T_k)\ge 1\}}\,,\quad t\ge 1\,.\tag{3.2}$$

We focus on the calculation of the conditional expectation

$$\widehat{M_\ell}(t,t+s]=E[M(t,t+s]\mid N_0(t)=\ell]\,,\quad \ell=0,1,\dots\,,$$

and the corresponding prediction error.

In view of (3.2) the σ-field generated by $N_0(t)$ is contained in the σ-field $\mathcal{F}_t$ generated by $(\Gamma_{ki})_{k,i\ge 1:T_k+\Gamma_{ki}\le t}$ and (T_k). Therefore for $t\ge 1$ and $s>0$,

$$\begin{aligned}\widehat{M_\ell}(t,t+s] &= E\Big[\sum_{k=1}^\infty I_{\{T_k\le 1\}}\,E\big[L_k(t-T_k,t+s-T_k]\mid \mathcal{F}_t\big]\Big|N_0(t)=\ell\Big]\\ &= \mu\, s\, E[N(1)\mid N_0(t)=\ell]\,.\end{aligned}$$

The latter conditional expectation will be evaluated below. Since we are also interested in other conditional moments of $M(t,t+s]$ we calculate the conditional characteristic function $E[\mathrm{e}^{\,i\,xM(t,t+s]}\mid N_0(t)=\ell]$, $x\in\mathbb{R}$.

Lemma 3.1. *The conditional characteristic function of $M(t,t+s]$ given $\{N_0(t)=\ell\}$ is given by*

$$E[\mathrm{e}^{\,i\,xM(t,t+s]}\mid N_0(t)=\ell]=(E[\mathrm{e}^{\,i\,x\,L(s)}])^\ell\,\mathrm{e}^{\,-\lambda_0(t)\,(1-E[\mathrm{e}^{\,i\,x\,L(s)}])}\,,\quad x\in\mathbb{R}\,,\tag{3.3}$$

where

$$\lambda_0(t)=\lambda\,\frac{\mathrm{e}^{\,-\gamma\,t}}{\gamma}\,(\mathrm{e}^{\,\gamma}-1)\,.\tag{3.4}$$

Proof. We proceed similarly to the proof of Lemma 2.2:

$$\begin{aligned}E[\mathrm{e}^{\,i\,xM(t,t+s]}\mid N_0(t)] &= E\Big[E[\mathrm{e}^{\,i\,xM(t,t+s]}\mid\mathcal{F}_t]\Big|N_0(t)\Big]\\ &= E\Big[\Big(E\Big[\mathrm{e}^{\,i\,x\,L(s)}\Big]\Big)^{N(1)}\Big|N_0(t)\Big]\\ &= \Big(E\Big[\mathrm{e}^{\,i\,x\,L(s)}\Big]\Big)^{N_0(t)}\,E\Big[\Big(E\Big[\mathrm{e}^{\,i\,x\,L(s)}\Big]\Big)^{N(1)-N_0(t)}\Big|N_0(t)\Big]\,.\end{aligned}\tag{3.5}$$

Let Q be a Poisson random measure on the state space $E = [0,1] \times [0,\infty)$ with mean measure $\nu = \lambda \,\mathrm{Leb} \times F$, where F denotes the distribution of Γ_{k1}. Then $N(1)$ and $N_0(t)$ have the Poisson integral representations

$$\begin{aligned} N(1) &= \int_E Q(ds,dy) \\ &= \int_{s=0}^1 \int_{y=0}^{t-s} Q(ds,dy) + \int_{s=0}^1 \int_{y=t-s}^{\infty} Q(ds,dy) \\ &= N_0(t) + [N(1) - N_0(t)]\,. \end{aligned}$$

Due to the splitting property of the Poisson process and since the integrals above are defined on disjoint subsets of the state space, the random variables $N_0(t)$ and $N(1) - N_0(t)$ are independent and Poisson distributed with parameters

$$E[N(1) - N_0(t)] = \lambda \int_0^1 \int_{t-s}^{\infty} F(dy)\,ds = \lambda \,\frac{\mathrm{e}^{-\gamma t}}{\gamma}\,(\mathrm{e}^{\gamma} - 1) = \lambda_0(t) \quad \text{and} \quad E[N_0(t)] = \lambda - \lambda_0(t)\,.$$

Therefore we conclude from (3.5) that

$$E[\mathrm{e}^{\,i\,xM(t,t+s]} \mid N_0(t)] \;=\; \Big(E\left[\mathrm{e}^{\,i\,x\,L(s)}\right]\Big)^{N_0(t)} \; E\Big[\Big(E\Big[\mathrm{e}^{\,i\,x\,L(s)}\Big]\Big)^{N(1)-N_0(t)}\Big]\,.$$

The latter relation yields the desired conditional characteristic function. □

The following result can now be obtained by differentiating the characteristic function (3.3) sufficiently often and then considering the derivatives at zero.

Theorem 3.2. *Assume that L is a compound Poisson process with underlying Poisson intensity γ.*
(1) *If $\mu = EX_{11}\,\gamma$ exists and is finite then the prediction of $M(t,t+s]$ given $\{N_0(t) = \ell\}$ has the form*

$$\widehat{M}_\ell(t,t+s] \;=\; \mu\, s\,[\lambda_0(t) + \ell]\,, \quad t \ge 1\,, s > 0\,, \quad \ell = 0,1,\ldots\,, \tag{3.6}$$

where $\lambda_0(t)$ is given in (3.4).
(2) *If in addition $\sigma^2 = \mathrm{Var}(L(1)) = E[X_{11}^2]\,\gamma < \infty$ then*

$$\begin{aligned} &\mathrm{Var}(M(t,t+s] \mid N_0(t) = \ell) \\ &= \sigma^2\, s\,(\ell + \lambda_0(t)) + (\mu\, s)^2\, \lambda_0(t)\,, \quad t \ge 1\,, s > 0\,, \quad \ell = 0,1,\ldots. \end{aligned} \tag{3.7}$$

Remark 3.3. Notice that the unconditional prediction error of $M(t,t+s]$ given $N_0(t)$ is given by

$$\begin{aligned} E[\mathrm{Var}(M(t,t+s] \mid N_0(t))] &= (\sigma^2\, s + (\mu\, s)^2)\,\lambda_0(t) + \sigma^2\, s\, E[N_0(t)] \\ &= \sigma^2\, s\, \lambda + (\mu\, s)^2\, \lambda_0(t)\,. \end{aligned}$$

Remark 3.4. In a next step we include more information in the condition of the expected value of $M(t,t+s]$, i.e., we focus on $E[M(t,t+s] \mid M(t)\,, N_0(t)]$. Both processes $M(t)$ and $N_0(t)$ are assumed to be observable at time t. Then the σ-field $\mathcal{G}_t$ generated by (T_k), $((\Gamma_{ki}, X_{ki}))_{k,i\ge 1\,,T_k+\Gamma_{ki}\le t}$ is larger than the σ-field generated by $M(t)$ and $N_0(t)$. Therefore

$$\begin{aligned} E[M(t,t+s] \mid M(t)\,, N_0(t)] &= \mu\, s\, E\Big[E[N(1) \mid \mathcal{G}_t]\Big| M(t)\,, N_0(t)\Big] \\ &= \mu\, s\, E\Big[N(1) \mid M(t)\,, N_0(t)\Big]\,. \end{aligned} \tag{3.8}$$

Consider the Poisson random measure $\widetilde{Q}$ with points $(T_k, (\Gamma_{ki})_{i=1,2,\ldots}, (X_{ki})_{i=1,2,\ldots})$ on the state space $[0,\infty)\times[0,\infty)^\infty\times\mathbb{R}^\infty$. Then $N(1)$, $N_0(t)$ and $M(t)$ have the representation as Poisson integrals with respect to $\widetilde{Q}$:

$$\begin{aligned}
N(1) &= \int_{[0,1]\times[0,\infty)^\infty\times\mathbb{R}^\infty} \widetilde{Q}(ds, d(\gamma_i), d(x_i))\,,\\
N_0(t) &= \int_{[0,1]\times[0,\infty)^\infty\times\mathbb{R}^\infty} I_{\{s+\gamma_1\le t\}}((s,(\gamma_i),(x_i)))\,\widetilde{Q}(ds, d(\gamma_i), d(x_i))\,,\\
M(t) &= \int_{[0,1]\times[0,\infty)^\infty\times\mathbb{R}^\infty} \sum_{i=1}^\infty x_i\, I_{\{s+(\gamma_1+\cdots+\gamma_i)\le t\}}((s,(\gamma_i),(x_i)))\,\widetilde{Q}(ds, d(\gamma_i), d(x_i))\,.
\end{aligned}$$

Notice that

$$N(1)-N_0(t) = \int_{[0,1]\times[0,\infty)^\infty\times\mathbb{R}^\infty} I_{\{s+\gamma_1> t\}}((s,(\gamma_i),(x_i)))\,\widetilde{Q}(ds, d(\gamma_i), d(x_i))\,.$$

The support of the integrand in $N(1)-N_0(t)$ is disjoint from the supports of the integrands in $N_0(t)$ and $M(t)$ and therefore $N(1)-N_0(t)$ is independent of $N_0(t)$ and $M(t)$. We conclude from (3.8) that

$$\begin{aligned}
E[M(t,t+s]\mid M(t)\,,N_0(t)] &= \mu\, s\big[E[N(1)-N_0(t)\mid M(t),N_0(t)] + E[N_0(t)\mid M(t)\,,N_0(t)]\big]\\
&= \mu\, s\big[E[N(1)-N_0(t)] + N_0(t)\big] = \mu\, s\big[\lambda_0(t)+N_0(t)\big]\,. \qquad (3.9)
\end{aligned}$$

Surprisingly, this is the same formula (3.6) as for $E[M(t,t+s]\mid N_0(t)]$. Hence taking into account information additional to $N_0(t)$ does not change the prediction of $M(t,t+s]$. A similar calculation shows that the prediction error remains the same. Also notice that we may conclude from (3.9) that

$$E[M(t,t+s]\mid M(t)] \;=\; \mu\, s\big[\lambda_0(t)+E[N_0(t)\mid M(t)]\big]\,.$$

The latter relation sheds some light on the prediction formula (2.5).

3.2. **Prediction with delay in reporting.** In this section we assume that the cluster process starting at T_k has the form $L_k(t-T_k)=R_k(t-T_k-D_k)$, where R_k is a Lévy process on $[0,\infty)$ with the convention that $R_k(s)=0$ a.s. for $s\le 0$ and D_k is a positive random variable with distribution F_D. We write $R=R_0$ for a generic element of the sequence (R_k). We also assume that the iid sequence (D_k), the sequence of the claim arrivals (T_k) and the iid sequence $(R_k)_{k=0,1,\ldots}$ of Lévy processes are independent. We interpret D_k as the time that elapses between the arrival time T_k of the kth claim and its reporting time T_k+D_k. Thus the inhomogeneous Poisson process

$$\widetilde{N}(t) = \sharp\{k\ge 1: T_k+D_k\le t\,, T_k\in[0,1]\}\,,\quad t\ge 1\,,$$

is observable at time $t\ge 1$, whereas the claim number $N(1)$ is not necessarily observable. In what follows, we will give expressions for the prediction of $M(t,t+s]$ given $\widetilde{N}(t)$:

$$\widetilde{M_\ell}(t,t+s] = E[M(t,t+s]\mid \widetilde{N}(t)=\ell]\,,\quad \ell=0,1,\ldots,\quad t\ge 1\,,s>0\,.$$

The key to the derivation of the prediction and the prediction error is again an expression for the characteristic function of $M(t,t+s]$ given $\widetilde{N}(t)$.

Lemma 3.5. *The conditional characteristic function of $M(t,t+s]$ given $\{\widetilde{N}(t)=\ell\}$ is given by*

$$\begin{aligned}
&E\big[\mathrm{e}^{\,i\,x\,M(t,t+s]}\mid \widetilde{N}(t)=\ell\big]\\
&= \Big(E[\mathrm{e}^{\,i\,x\,R(s)}]\Big)^\ell \exp\Big\{-\lambda\int_{v=0}^1\int_{r=t-v}^{t+s-v}\Big(1-E\big[\mathrm{e}^{\,i\,x\,R(t+s-v-r)}\big]\Big)\,dv\,F_D(dr)\Big\}\,, \qquad (3.10)
\end{aligned}$$

$$\ell = 0,1,\dots\,, \quad t \ge 1\,, s > 0\,.$$

Proof. We start by calculating the characteristic function of $M(t,t+s]$ conditional on (T_j) and (D_j):

$$\begin{aligned}
E[\mathrm{e}^{\,i\,x\,M(t,t+s]} \mid (T_j),(D_j)] &= E\Big[\prod_{k=1}^{\infty} \mathrm{e}^{\,i\,x\,R_k(t-T_k-D_k,t+s-T_k-D_k]I_{\{T_k\le 1\}}} \mid (T_j),(D_j)\Big] \\
&= \prod_{k=1}^{N(1)} E\Big[\mathrm{e}^{\,i\,x\,R(t-T_k-D_k,t+s-T_k-D_k]} \mid T_k, D_k\Big] \\
&= \Big(E[\mathrm{e}^{\,i\,x\,R(s)}]\Big)^{\widetilde N(t)} \prod_{k:T_k\le 1\,,T_k+D_k>t} E\Big[\mathrm{e}^{\,i\,x\,R(t+s-T_k-D_k)} \mid T_k, D_k\Big]\,. \qquad (3.11)
\end{aligned}$$

In the last step we used that $R(t-T_k-D_k)=0$ a.s. for k such that $T_k+D_k\ge t$ a.s. Write $\widetilde Q$ for the Poisson process of the points (T_k,D_k) on the state space $[0,1]\times(0,\infty)$ with mean measure $\lambda\mathrm{Leb}\times F_D$. Then the Poisson process $\widetilde N$ has representation

$$\widetilde N(t) = \int_{v=0}^{1}\int_{r=0}^{t-v} \widetilde Q(dv,dr)\,, \quad t\ge 1\,. \qquad (3.12)$$

Since the σ-field generated by $\widetilde N(t)$ is contained in the σ-field generated by (T_k) and (D_k) we have in view of (3.11),

$$\begin{aligned}
&E[\mathrm{e}^{\,i\,x\,M(t,t+s]} \mid \widetilde N(t)] \\
&= \Big(E[\mathrm{e}^{\,i\,x\,R(s)}]\Big)^{\widetilde N(t)} E\Big[\prod_{k:T_k\le 1,T_k+D_k>t} E\Big[\mathrm{e}^{\,i\,x\,R(t+s-T_k-D_k)} \mid T_k,D_k\Big] \mid \widetilde N(t)\Big] \\
&= \Big(E[\mathrm{e}^{\,i\,x\,R(s)}]\Big)^{\widetilde N(t)} E\Big[\exp\Big\{\sum_{k:t<T_k+D_k\le t+s} \log\, E\Big[\mathrm{e}^{\,i\,x\,R(t+s-T_k-D_k)} \mid T_k,D_k\Big]\Big\} \mid \widetilde N(t)\Big] \\
&= \Big(E[\mathrm{e}^{\,i\,x\,R(s)}]\Big)^{\widetilde N(t)} E\Big[\exp\Big\{\int_{v=0}^{1}\int_{r=t-v}^{t+s-v} \log\, E\Big[\mathrm{e}^{\,i\,x\,R(t+s-v-r)}\Big]\widetilde Q(dv,dr)\Big\} \mid \widetilde N(t)\Big]\,.
\end{aligned}$$

We observe that the Poisson integrals in the last expression and in (3.12) have disjoint supports, hence they are independent and

$$\begin{aligned}
&E[\mathrm{e}^{\,i\,x\,M(t,t+s]} \mid \widetilde N(t)] \\
&= \Big(E[\mathrm{e}^{\,i\,x\,R(s)}]\Big)^{\widetilde N(t)} E\Big[\exp\Big\{\int_{v=0}^{1}\int_{r=t-v}^{t+s-v} \log\, E\Big[\mathrm{e}^{\,i\,x\,R(t+s-v-r)}\Big]\widetilde Q(dv,dr)\Big\}\Big]\,.
\end{aligned}$$

Direct calculation for $t\ge 1$, $s>0$ yields (3.10). □

Remark 3.6. Notice that for $t\ge 1$, $s>0$,

$$\int_{v=0}^{1}\int_{r=t-v}^{t+s-v} \Big(1-E\Big[\mathrm{e}^{\,i\,x\,R(t+s-v-r)}\Big]\Big)\,dv\,F_D(dr) = \int_t^{t+s}(1-E[\mathrm{e}^{\,i\,x\,R(t+s-z)}])\,G(dz)$$

where $G=F_U*F_D$. Therefore

$$\exp\Big\{-\lambda\int_{v=0}^{1}\int_{r=t-v}^{t+s-v}\Big(1-E\Big[\mathrm{e}^{\,i\,x\,R(t+s-v-r)}\Big]\Big)\,dv\,F_D(dr)\Big\}$$

$$= \exp\Big\{-E[\widetilde N(t,t+s]]\int_t^{t+s}(1-E[\mathrm{e}^{\,i\,x\,R(t+s-z)}])\,(G(dz)/E[\widetilde N(t,t+s]])\Big\}$$

$$= E\Big[\Big(E\Big[\mathrm{e}^{\,i\,x\,R(t+s-Z)}\Big]\Big)^{\widetilde N(t,t+s]}\Big]\,,$$

where Z is independent of R and has distribution $G(dz)/E[\widetilde N(t,t+s]]$ on $(t,t+s]$. This expression is an alternative formula for the second term in (3.10).

Differentiation of the conditional characteristic function at $x=0$ sufficiently often yields the following result.

Theorem 3.7. *Assume the model (1.2) with the delayed Lévy processes $L_k(t)=R_k(t-D_k)$, $k=1,2,\ldots$, as cluster processes, where by convention $R_k(s)=0$ a.s. for $s\le 0$, (D_k) constitutes an iid sequence of positive random variables with distribution F_D and the sequences (T_k), (D_k) and (R_k) are independent.*
(1) *Assume that $\mu=E[R(1)]$ exists and is finite. Then the prediction $\widetilde{M_\ell}(t,t+s]$ of $M(t,t+s]$ given $\{\widetilde N(t)=\ell\}$ has the following form for $\ell=0,1,\ldots$*

$$\widetilde{M_\ell}(t,t+s] \;=\; \mu\,s\,\ell+\mu\,J_1\,,\quad t\ge 1\,,s>0\,, \tag{3.13}$$

where

$$J_i=J_i(t,s)=\lambda\int_{v=0}^1\int_{r=t-v}^{t+s-v}(t+s-r-v)^i\,F_D(dr)\,dv\,,\quad i=1,2\,.$$

(2) *If in addition $\sigma^2=\mathrm{Var}(R(1))<\infty$ then*

$$\mathrm{Var}(M(t,t+s]\mid \widetilde N(t)=\ell) \;=\; \ell\,s\,\sigma^2+\sigma^2\,J_1+\mu^2\,J_2\,. \tag{3.14}$$

Remark 3.8. The prediction formulas (3.13) and (3.6) are rather similar. Both are linear functions of $\widetilde N(t)=\ell$ or $\widehat N(t)=\ell$, respectively. This is agreement with the assumptions of the chain ladder which is a standard technique for claims reserving; see Mack [18, 19]. A particularly interesting case occurs when the delay in reporting variable D is $U(0,a)$ distributed for some $a>0$. Then

$$\widetilde{M_\ell}(t,t+s] \;=\; \mu\,s\,\ell+\mu\,\lambda\,a^{-1}\int_0^1\int_{r=(t-v)\wedge a}^{(t+s-v)\wedge a}(t+s-r-v)\,dr\,dv\,.$$

A comparison of the conditional prediction errors (3.7) and (3.14) shows that both are linear function of ℓ as well. Since $\widetilde N(t)$ is Poisson distributed with parameter

$$E[\widetilde N(t)] \;=\; \lambda\int_{v=0}^1\int_{r=0}^{t-v}dv\,F_D(dr)=\lambda\int_{t-1}^t F_D(v)\,dv\,,\quad t\ge 1\,,$$

straightforward calculation yields the unconditional prediction error

$$E[\mathrm{Var}(M(t,t+s]\mid \widetilde N(t))]=\lambda\,\sigma^2\,s\int_{t-1}^t F_D(v)\,dv+\sigma^2\,J_1+\mu^2\,J_2\,.$$

3.3. **Truncated compound Poisson clusters.** In this section we consider another modification of the Poisson cluster process. We again consider the model (1.2) but the cluster processes L_k are not Lévy processes. We assume that the iid processes L_i, $i=1,2,\ldots$, have the following structure

$$L_i(t)=K_i\wedge R_i(t)\,,\quad t\ge 1\,, \tag{3.15}$$

where K_i is a non-negative random variable and R_i is a non-negative compound Poisson process with representation given on the right-hand side of (3.1). Moreover, we assume that $(K_i)_{i=0,1,\ldots}$ is an iid sequence independent of the iid sequence $(R_i)_{i=0,1,\ldots}$. We also write N_i for the homogeneous Poisson process with intensity $\gamma>0$ which underlies the compound Poisson process R_i. If $N_i=R_i$

and K_i is a non-negative integer-valued random variable, K_i has the interpretation as the total number of payments for the ith claim. In particular, if $K_i = k_0$ is a constant integer there are exactly k_0 payments for each claim.

As explained before, in practice one is often not informed at time $t = 1$ whether a claim has happened, i.e., the arrival times T_k are often unknown until some future instant of time. Therefore we assumed in Section 3.1 that we take into account only those claims for which $T_k + \Gamma_{k1} \le t$, and in Section 3.2 we take into account only those claims for which $T_k + D_k \le t$. Here we will consider prediction of $M(t, t+s]$ given $N_0(t)$ defined in (3.2) and $(R_k(t - T_k))_{k:T_k \le 1, T_k+\Gamma_{k1} \le t}$. We write $\mathcal{H}_t$ for the σ-field generated by these quantities which are observable at time t. Here and in what follows we use the notation of Section 3.1.

Theorem 3.9. *Assume that the iid sequence (L_i) has the structure described in* (3.15)*.*

(1) *If $E[K_1] < \infty$ then the prediction of $M(t, t+s]$ given $\mathcal{H}_t$, $t \ge 1$, $s > 0$, has the form*

$$E(M(t,t+s] \mid \mathcal{H}_t) = \lambda_0(t)\, E[K_0 \wedge R_0(s)] + \sum_{k:T_k\le 1\,, T_k+\Gamma_{k1}\le t} E[K_0 \wedge (R_0(s) + R_k(t-T_k)) - K_0 \wedge R_k(t-T_k) \mid R_k(t-T_k)]\,,$$

where $\lambda_0(t)$ is defined in (3.4)*. Here $(R_k)_{k=0,1,\dots}$, (T_k) and $(K_k)_{k=0,1,\dots}$ are independent.*

(2) *If in addition* $\mathrm{Var}(K_1) < \infty$ *then for $t \ge 1$, $s > 0$,*

$$\mathrm{Var}(M(t,t+s] \mid \mathcal{H}_t) = \lambda_0(t)\, E[(K_0 \wedge R_0(s))^2] + \sum_{k:T_k\le 1\,, T_k+\Gamma_{k1}\le t} \mathrm{Var}\big(K_0 \wedge (R_0(s) + R_k(t-T_k)) - K_0 \wedge R_k(t-T_k) \mid R_k(t-T_k)\big)\,.$$

This result again follows by differentiation at zero of the characteristic function of $M(t, t+s]$ given $\mathcal{H}_t$.

Lemma 3.10. *The conditional characteristic function of $M(t, t+s]$ given $\mathcal{H}_t$ has the form*

$$E[\mathrm{e}^{\,i\,x\,M(t,t+s]} \mid \mathcal{H}_t] \;=\; \prod_{k:T_k\le 1, T_k+\Gamma_{k1}\le t} E\Big[\mathrm{e}^{\,i\,x\,\big[K_0\wedge(R_0(s)+R_k(t-T_k))-K_0\wedge R_k(t-T_k)\big]} \mid R_k(t-T_k)\Big]$$

$$\mathrm{e}^{\,-\lambda_0(t)\,(1-E\mathrm{e}^{\,i\,x\,K_0\wedge R_0(s)})}\,. \tag{3.16}$$

Proof. We start by observing that

$$\begin{aligned} M(t,t+s] &= \sum_{k:T_k\le 1\,, T_k+\Gamma_{k1}\le t} \big[K_k \wedge R_k(t+s-T_k) - K_k \wedge R_k(t-T_k)\big] \\ &\quad + \sum_{k:T_k\le 1\,, T_k+\Gamma_{k1}> t} K_k \wedge R_k(t+s-T_k) \\ &= \sum_{k:T_k\le 1\,, T_k+\Gamma_{k1}\le t} \big[K_k \wedge R_k(t+s-T_k) - K_k \wedge R_k(t-T_k)\big] \\ &\quad + \sum_{k:T_k\le 1\,, T_k+\Gamma_{k1}> t} K_k \wedge R_k(t-T_k, t+s-T_k]\,. \end{aligned}$$

The σ-field $\mathcal{H}_t$ is contained in the σ-field $\mathcal{G}_t$ generated by $(T_k), ((\Gamma_{ki}, X_{ki}))_{k,i\ge 1, T_k+\Gamma_{ki}\le t}$. Therefore

$$E[\mathrm{e}^{\,i\,x\,M(t,t+s]} \mid \mathcal{G}_t]$$

$$= \prod_{k:T_k\le 1,T_k+\Gamma_{k1}\le t} E\Big[\mathrm{e}^{\,i\,x\,\big[K_0\wedge(R_0(s)+R_k(t-T_k))-K_0\wedge R_k(t-T_k)\big]}\mid R_k(t-T_k)\Big]$$

$$\prod_{k:T_k\le 1,T_k+\Gamma_{k1}> t} E\Big[\mathrm{e}^{\,i\,x\,K_0\wedge R_k(t-T_k,t+s-T_k]}\mid T_k\,,R_k(t-T_k)\Big]$$

$$= \prod_{k:T_k\le 1,T_k+\Gamma_{k1}\le t} E\Big[\mathrm{e}^{\,i\,x\,\big[K_0\wedge(R_0(s)+R_k(t-T_k))-K_0\wedge R_k(t-T_k)\big]}\mid R_k(t-T_k)\Big]$$

$$\Big(E\Big[\mathrm{e}^{\,i\,x\,K_0\wedge R_0(s)}\Big]\Big)^{N(1)-N_0(t)}\,.$$

The first factor is measurable with respect to $\mathcal{H}_t$, and $\mathcal{H}_t$ is independent of $N(1)-N_0(t)$. Therefore we have

$$E[\mathrm{e}^{\,i\,x\,M(t,t+s]}\mid \mathcal{H}_t] = \prod_{k:T_k\le 1,T_k+\Gamma_{k1}\le t} E\Big[\mathrm{e}^{\,i\,x\,\big[K_0\wedge(R_0(s)+R_k(t-T_k))-K_0\wedge R_k(t-T_k)\big]}\mid R_k(t-T_k)\Big]$$

$$E\Big[\Big(E\Big[\mathrm{e}^{\,i\,x\,K_0\wedge R_0(s)}\Big]\Big)^{N(1)-N_0(t)}\Big]\,.$$

The latter relation implies the desired result (3.16). □

Remark 3.11. The unconditional prediction error is given by

$$E[\mathrm{Var}(M(t,t+s]\mid\mathcal{H}_t)] = \lambda_0(t)\,E[(K_0\wedge R_0(s))^2]$$

$$+E\Bigg[\sum_{k=1}^{N(1)} I_{\{N_k(t-T_k)\ge 1\}}\mathrm{Var}\big(K_0\wedge(R_0(s)+R_k(t-T_k))-K_0\wedge R_k(t-T_k)\mid R_k(t-T_k)\big)\Bigg]\,.$$

Using the order statistics property of the homogeneous Poisson process N, we obtain

$$E[\mathrm{Var}(M(t,t+s]\mid\mathcal{H}_t)]-\lambda_0(t)\,E[(K_0\wedge R_0(s))^2]$$

$$=\lambda\int_0^1 E\big[I_{\{N_1(t-y)\ge 1\}}\mathrm{Var}\big(K_0\wedge(R_0(s)+R_1(t-y))-K_0\wedge R_1(t-y)\mid R_1(t-y)\big)\big]\,dy\,.$$

Acknowledgment. A major part of this paper was written when Muneya Matsui visited the Department of Mathematics of the University of Copenhagen. He would like to thank for hospitality and for financial support from the FNU Grant 272-06-0442.

References

[1] Bartlett, M.S. (1963) The spectral analysis of point processes. *J. Royal Statist. Soc. Ser. B* **25**, 264–296.

[2] Faÿ, G., González-Arévalo, B., Mikosch, T. and Samorodnitsky, G. (2006) Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Systems Theory Appl.* **54**, 121–140.

[3] Giraitis, L., Molchanov, S.A. and Surgailis, D. (1993) Long memory shot noises and limit theorems with application to Burgers's equation. In: *New Directions in Time Series Analysis,* Part II, pp. 153–176. *IMA Vol. Math. Appl.* **46**, Springer, New York.

[4] Heinrich, L. and Schmidt, V. (1985) Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. Appl. Probab.* **17**, 709–730.

[5] Hsing, T. and Teugels, J.L. (1989) Extremal properties of shot noise processes. *Adv. Appl. Probab.* **21**, 513–525.

[6] Jessen, A.H., Mikosch, T. and Samorodnitsky, G. (2009) Prediction of outstanding payments in a Poisson cluster model. Available under
www.math.ku.dk/∼mikosch/Preprint/Prediction/prediction260309.pdf

[7] Kallenberg, O. (1983) *Random Measures.* Third edition. Akademie-Verlag, Berlin.

[8] Klüppelberg, C. and Kühn, C. (2004) Fractional Brownian motion as a weak limit of Poisson shot noise processes—with applications to finance. *Stoch. Process. Appl.* **113**, 333–351.

[9] Klüppelberg, C. and Mikosch, T. (1995) Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli* **1**, 125–147.

[10] Klüppelberg, C. and Mikosch, T. (1995) Delay in claim settlement and ruin probability approximations. *Scand. Actuar. J.* 154–168.

[11] Klüppelberg, C., Mikosch, T. and Schärf, A. (2003) Regular variation in the mean and stable limits for Poisson shot noise. *Bernoulli* **9**, 467–496.

[12] Konstantopoulos, T. and Lin, Si-Jian (1998) Macroscopic models for long-range dependent network traffic. *Queueing Systems Theory Appl.* **28**, 215–243.

[13] Kurtz, T.G. (1997) Limit theorems for workload input models. In: Kelly, F.P., Zachary, S. and Ziedins, I. (Eds.) *Stochastic Networks. Theory and Applications,* pp. 119–139. Oxford University Press, Oxford, UK.

[14] Lane, J.A. (1984) The central limit theorem for the Poisson shot-noise process. *J. Appl. Probab.* **21**, 287–301.

[15] Lane, J.A. The Berry–Esseen bound for the Poisson shot-noise. *Adv. Appl. Probab.* **19**, 512–514.

[16] Levy, J.B. and Taqqu, M.S. (2000) Renewal reward processes with heavy-tailed interrenewal times and heavy-tailed rewards. *Bernoulli* **6**, 23–44.

[17] Lewis, P.A. (1964) A branching Poisson process model for the analysis of computer failure patterns. *J. Royal Statist. Soc. Ser. B* **26**, 398–456.

[18] Mack, T. (1993) Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin* **23**, 213–225.

[19] Mack, T. (1994) Which stochastic model is underlying the chain ladder method? *Insurance: Mathematics & Economics* **15**, 133–138.

[20] McCormick, W.P. (1997) Extremes for shot noise processes with heavy tailed amplitudes. *J. Appl. Probab.* **34**, 643–656.

[21] Mikosch, T. (2009) *Non-Life Insurance Mathematics. An Introduction with the Poisson Process.* 2nd Edition. Springer, Heidelberg.

[22] Mikosch, T., Resnick, S.I., Rootzén, H. and Stegeman, A. (2002) Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.* **12**, 23–68.

[23] Mikosch, T. and Samorodnitsky, G. (2007) Scaling limits for cumulative input processes. *Math. Oper. Res.* **32**, 890–919.

[24] Pipiras, V. and Taqqu, M.S. (2000) The limit of a renewal-reward process with heavy-tailed rewards is not a linear fractional stable motion. *Bernoulli* **6**, 607–614.

[25] Samorodnitsky, G. (1996) A class of shot noise models for financial applications. In: *Athens Conference on Applied Probability and Time Series Analysis,* Vol. I, pp. 332–353. *Lecture Notes in Statistics* **114**, Springer, New York.

[26] Sato, K.-i. (1999) *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge, UK.

[27] Stegeman, A. (2002) Extremal behavior of heavy-tailed on-periods in a superposition of ON/OFF processes. *Adv. Appl. Probab.* **34**, 179–204.

[28] Vere–Jones, D. (1970) Stochastic models for earthquake occurrences. *J. Royal Statist. Soc. Ser. B*, **32**, 1–62.

Department of Mathematics, Keio University, 3-14-1 Hiyoshi Kohoku-ku, Yokohama 223-8522, Japan
E-mail address: mmuneya@gmail.com

Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark
E-mail address: mikosch@math.ku.dk