Closed-form and numerical computations of actuarial indicators in ruin theory and claim reserving

Alexandre Brouste, Christophe Dutang

Closed-form and numerical computations of actuarial indicators in ruin theory and claim reserving

Alexandre BROUSTE [1, 2]

Christophe DUTANG[1, 2, 3]

Abstract

Insurance reserving is a key topic for both actuaries and academics. In the present paper, we present an efficient way to compute all the key indicators in a unified approach of the ruin theory and claim reserving methods. The proposed framework allows to derive closed-form formulas for both ruin theory and claim reserves indicators. A numerical illustration of these indicators is carried out on a real dataset from a private insurer.

Keywords : *ruin theory*; *claim reserving*; *Poisson process*; *non-life insurance.*

Résumé

Le provisionnement en assurance non-vie est un sujet clé pour les actuaires et les académiques. Dans le présent article, nous présentons une méthode efficace pour calculer les indicateurs par une approche unifiée de la théorie de la ruine et du provisionnement non-vie. Le cadre proposé permet de déduire des formules fermées pour les indicateurs de provisionnement et de ruine. Une illustration de ces indicateurs est réalisée sur un jeu de données réellles.

Mots-clés : *théorie de la ruine*, *provisionnement non-vie*, *processus de Poisson*, *assurance non vie*.

1. Laboratoire Manceau de Mathématiques, Université du Maine, Le Mans, France
2. Institut du Risque de l'Assurance, Université du Maine, Le Mans, France
3. Corresponding author. Address: LMM, Université du Maine, Avenue Olivier Messiaen, F-72085 Le Mans CEDEX 9; Tel.: +33 2 43 83 31 34; Email: `christophe.dutang@univ-lemans.fr`

1 Introduction

Insurance reserving is a well-known topic for both actuaries and academics, whereas the ruin theory remains mainly the field of academics. The computation of insurance reserves being mandatory whereas ruin-related indicators are not is one of the main reasons to explain why practitioners neglect the use of ruin theory in their daily business. Nevertheless, with the upcoming risk-based regulatory requirements, the computation of solvency probabilities at different levels and different time horizons is increasingly popular in the past ten years. In the present paper, we propose a new efficient way to compute numerous key indicators in a unified approach of ruin theory and claim reserving.

Another important factor explaining the disaffection of practitioners for ruin theory when assessing reserves is the type of data to be used: the data granularity for classic reserving methods is line-of-business aggregated datasets whereas in ruin theory, individual loss level is needed, see Asmussen and Albrecher (2010) and the references therein. Reserving methods are in fact mainly for aggregated data triangles, see Wuethrich and Merz (2008) and the references therein. As pointed out by Wuethrich and Merz (2008), "most of the classical claims reserving methods do not distinguish reported claims from not-reported claims." However, there is a growing literature for micro-level or individual claim-level reserving methods.

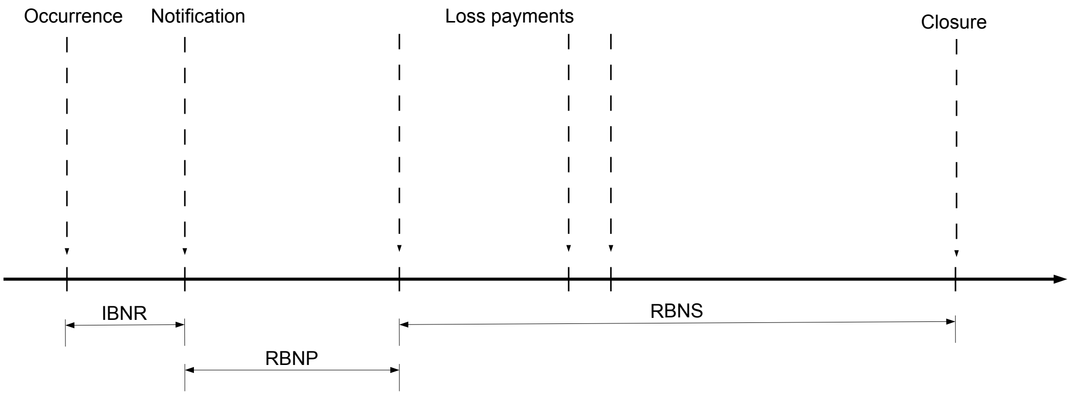

Figure 1 – Claim development process (IBNR: incurred but not reported, RBNP: reported but not paid, RBNS: reported but not settled)

Reserving in a continuous time perspective dates back to pioneer works of Karlsson (1976), Jewell (1989) and Arjas (1989). Few years after these papers, Norberg (1993) first formulates the reserving problem in a continuous time probabilistic setting by considering marked Poisson processes, see e.g. subsequent extensions Haastrup and Arjas (1996). That is, the full claim process described in Figure 1 is considered. The time between occurrence and notification corresponds to the reporting delay (IBNR in Figure 1) by the policyholder and is assumed to equal zero in this study.

The ith claim is characterized by a 4-tuple $(T_i, V_i, Y_i, Y_i'(v)_{v\in[0,V_i]})$ where T_i denotes the occurrence time, V_i the settlement time, Y_i the total claim amount and $Y_i'(.)$ the payment process. The time between notification and closure (i.e. $t \in [T_i, T_i + V_i)$) corresponds to the settlement time, which can be further subdivided into the waiting time of first payment (RBNP in Figure 1) and the payment process (RBNS in Figure 1). The claim process $(T_1, T_2, \dots)$ is governed by a non-homogeneous Poisson process $(N_t, t \geq 0)$. Other recent papers in that direction are Larsen (2007), Antonio and Plat (2014) for continuous time setting and Pigeon et al. (2013), Drieskens et al. (2012) for discrete time setting which provide estimation procedures with explanatory variables.

Currently, there exists almost only one alternative to marked Poisson processes in the actuarial literature: the Poisson shot noise processes of Klueppelberg and Mikosch (1995), further developed in Matsui and Mikosch (2010) and Matsui (2015, 2014). They consider that the ith claim is a couple $(T_i, L_i(.))$ where L_i may represent the loss process, typically independent Lévy processes.

In this paper, we follow the probabilistic framework of Norberg (1993), which is an extension of the classical Cramér-Lundberg ruin model by considering settlement times and reporting delays. The paper is structured as follows. Section 2 presents our unified-approach extended framework used in the subsequent sections. Section 3 focuses on the (un)conditional moments of the aggregate claim process and examples of settlement times. Section 4 follows with unified-approach indicators of reserving and ruin topics. Finally, Section 5 illustrates the ruin and reserving indicators on a real insurance dataset, before Section 6 concludes.

2 An extended Cramér-Lundberg framework

In this section, we consider a process closed to the marked Poisson process of Norberg (1993). Indeed, we introduce an extension of the classical Cramér-Lundberg framework (e.g. Asmussen and Albrecher (2010)) and state the model assumptions.

2.1 Notation

The surplus of an insurance company at time t is represented by the risk process $R_t = u + ct - S_t$, where S_t denotes the aggregate claim amount, u is the initial surplus, c is the premium rate. Traditionally, the aggregate claim amount S_t is the sum of claim amounts $X_1, X_2, \dots$ arrived before time t, i.e. $S_t = \sum_{i=1}^{N_t} X_i$. By considering settlement times and reporting delays, we assume that

$$S_t = \sum_{i=1}^{N_t} Z_i(t), \text{ with } Z_i(t) = \frac{X_i}{V_i}(t - T_i)\mathbb{1}_{[T_i, T_i+V_i)}(t) + X_i \mathbb{1}_{[T_i+V_i,\infty)}(t),$$

where V_i, T_i denote respectively the settlement time and the occurence of the ith claim.

In other words, $Z_i(t)$ corresponds to the claim amount paid at time t and $X_i - Z_i(t)$ is the outstanding claim amount. As the ith claim is represented by (T_i, V_i, X_i) and the implicit assumption that the payment process is an affine function of time t, we have a simplified version of Norberg's model.

Comparing risk process R_t and the no-delay-no-settlement risk process

$$\tilde{R}_t = u + ct - \sum_{i=1}^{N_t} X_i,$$

we remark that $R_t \geq \tilde{R}_t$ a.s.. Therefore, the corresponding ruin probability of the considered model is always lower than the classical setting.

For the following study, we introduce the settlement function

$$g(t, w, v) = \frac{t - w}{v}\mathbb{1}_{[w,w+v)}(t) + \mathbb{1}_{[w+v,\infty)}(t). \tag{1}$$

representing the percentage of the claim paid at time t. Thus, we have $Z_i(t) = X_i g(t, T_i, V_i)$.

2.2 Model assumptions

Keeping in mind that we want to derive explicit formulas, we make the following assumptions

A1. the claim arrival process $(N_t, t \geq 0)$ is a homogeneous Poisson process with intensity λ,

A2. the settlement times are independent and identically distributed ($(V_i)_i \overset{\text{i.i.d.}}{\sim} V$),

A3. the claim amounts are independent and identically distributed ($(X_i)_i \overset{\text{i.i.d.}}{\sim} X$ with finite variance), A4. there is independence between waiting times, settlement times and claim amounts ($T_i - T_{i-1} \perp V_i \perp X_i$).

Note that (A1) leads to exponential occurrence times $(T_i)_i \overset{\text{i.i.d.}}{\sim} \mathcal{E}(\lambda)$.

3 Main results

In this section, we present the main results of the (un-)conditional first two moments of the aggregate claim process $(S_t, t \geq 0)$. We will also focus on an efficient numerical procedure to compute these indicators in Section 5.1. Those results will then be used in the subsequent sections. In the sequel, we will need claim index sets defined as follows

$$\mathcal{C}_t^{ns} = \{i \in \{1, \dots, N_t\}, T_i \leq t < T_i + V_i\},\ \mathcal{C}_t^{s} = \{i \in \{1, \dots, N_t\}, T_i + V_i \leq t\}.$$

They represent respectively not-settled claims and settled claims. These sets are a disjoint partition of claims occurred before time t, i.e. $\mathcal{C}_t^{ns}, \mathcal{C}_t^{s} \subset \{1, \dots, N_t\}$. We introduce filtrations depending for the claim arrival process and the knowledge up to time t

$$\mathcal{F}_t^N = \sigma\left(N_s, 0 \leq s \leq t\right),\ \mathcal{F}_t^{N,V} = \sigma\left((N_s, 0 \leq s \leq t), V_1, \dots, V_{N_t}\right),$$

$$\mathcal{F}_t^{N,C} = \sigma\left((N_s, 0 \leq s \leq t), (V_i)_{i \in C_t^s}\right), \mathcal{F}_t^{N,C,X} = \sigma\left((N_s, 0 \leq s \leq t), (V_i)_{i \in C_t^s}, X_1, \dots, X_{N_t}\right).$$

In the following, we suppose that at time t the claim amount is known when reported. Only the time of settlement is assumed random.

3.1 Closed-form formulas on the aggregate claim distribution

We present here an efficient procedure to compute the (un-)conditional first two moments of the aggregate claim process $(S_t, t \geq 0)$.

Proposition 3.1. *The conditional expectation at time t of the aggregate claim amount knowing the information up to time $s < t$ is*

$$\mathbf{E}\left(S_t \,|\, \mathcal{F}_s^{N,C,X}\right) = \sum_{i=1, i \in \mathcal{C}_s^s}^{N_s} X_i + \sum_{i=1, i \in \mathcal{C}_s^{ns}}^{N_s} X_i G(t, T_i) + \mathbf{E}(X) e^{-\lambda(t-s)} \sum_{k=1}^{\infty} \lambda^k A_k(G)(s,t), \tag{2}$$

where $A_k(\cdot)(s,t)$ is defined as

$$A_k(G)(s,t) = \sum_{j=1}^{k} \int_s^{t_{k+1}} \dots \int_s^{t_{j+1}} G(t, t_j) \frac{(t_j - s)^{j-1}}{(j-1)!} dt_j \dots dt_k, \quad t_{k+1} = t, \quad s < t, \tag{3}$$

and G is the bivariate function defined as

$$G(t, w) = \mathbf{E}(g(t, w, V)). \tag{4}$$

Proof. Direct computation leads to

$$\begin{aligned} \mathbf{E}\left(S_t \,|\, \mathcal{F}_s^{N,C,X}\right) &= \mathbf{E}\left(\sum_{i=1}^{N_s} Z_i(t) \,|\, \mathcal{F}_s^{N,C,X}\right) + \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t) \,|\, \mathcal{F}_s^{N,C,X}\right) \\ &= \sum_{i=1}^{N_s} \mathbf{E}\left(Z_i(t) \,|\, \mathcal{F}_s^{N,C,X}\right) + \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t) \,|\, N_s\right). \end{aligned}$$

Using $\mathcal{C}_s^s \cup \mathcal{C}_s^{ns} = \{1,\dots,N_s\}$ and $Z_i(t) = X_i g(t,T_i,V_i)$, we split the first sum between settled and not-settled claims

$$\begin{aligned}
\mathbf{E}\left(Z_i(t)\,|\,\mathcal{F}_s^{N,C,X}\right) &= X_i\mathbf{E}\left(g(t,T_i,V_i)|\,\mathcal{F}_s^{N,C,X}\right)\\
&= X_i g(t,T_i,V_i)\mathbf{E}\left(\mathbb{1}_{i\in\mathcal{C}_s^s}|\,\mathcal{F}_s^{N,C,X}\right) + X_i\mathbf{E}\left(g(t,T_i,V_i)\mathbb{1}_{i\in\mathcal{C}_s^{ns}}|\,\mathcal{F}_s^{N,C,X}\right)\\
&= X_i g(t,T_i,V_i)\mathbb{1}_{i\in\mathcal{C}_s^s} + X_i\mathbb{1}_{i\in\mathcal{C}_s^{ns}}\mathbf{E}\left(g(t,T_i,V_i)|\,\mathcal{F}_s^{N,C,X}\right)\\
&= X_i\mathbb{1}_{i\in\mathcal{C}_s^s} + X_i\mathbb{1}_{i\in\mathcal{C}_s^{ns}}\mathbf{E}\left(g(t,T_i,V_i)|\,\mathcal{F}_s^{N}\right),
\end{aligned}$$

since X_i, $\mathbb{1}_{i\in\mathcal{C}_s^s}$ and $\mathbb{1}_{i\in\mathcal{C}_s^{ns}}$ are mesurable with respect to $\mathcal{F}_s^{N,C,X}$. Using $\mathbf{E}\left(g(t,T_i,V_i)|\,\mathcal{F}_s^{N}\right) = G(t,T_i)$, we obtain

$$\mathbf{E}\left(S_t\,|\,\mathcal{F}_s^{N,C,X}\right) = \sum_{i=1,i\in\mathcal{C}_s^s}^{N_s} X_i + \sum_{i=1,i\in\mathcal{C}_s^{ns}}^{N_s} X_i G(t,T_i) + \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t)\,|\,N_s\right).$$

For the second term, using $N_s+1 \le i$, $(T_i < t)$,

$$\begin{aligned}
\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t)\,|\,N_s\right) &= \mathbf{E}\left(\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t)\,|\,\mathcal{F}_t^{N,V}\right)|\,N_s\right)\\
&= \mathbf{E}\,(X)\,\mathbf{E}\left(\sum_{i=N_s+1}^{N_t}\mathbf{E}\left(g(t,T_i,V_i)\,|\,N_s,N_t\right)|\,N_s\right)\\
&= \mathbf{E}\,(X)\sum_{\ell=k}^{\infty}\sum_{i=N_s+1}^{\ell}\mathbf{E}\left(g(t,T_i,V_i)\,|\,N_s,N_t=\ell\right)P(N_t=\ell|N_s)\\
&= \mathbf{E}\,(X)\sum_{k=1}^{\infty}\sum_{i=N_s+1}^{N_s+k}\mathbf{E}\left(g(t,T_i,V_i)\,|\,N_s,N_t-N_s=k\right)P(N_t-N_s=k|N_s)\\
&= \mathbf{E}(X)\sum_{k=1}^{\infty}\frac{(\lambda(t-s))^k}{k!}e^{-\lambda(t-s)}\sum_{i=N_s+1}^{N_s+k}\mathbf{E}\left(g(t,T_i,V_i)\,|\,N_s,N_t-N_s=k\right).
\end{aligned}$$

Denoting by $\tilde{H}_k$ the distribution function of $T_{N_s+1},\dots,T_{N_s+k}$ conditionally on $N_t - N_s = k$, the inner sum gives

$$\sum_{i=N_s+1}^{N_s+k}\mathbf{E}\left(g(t,T_i,V_i)\,|\,N_s,N_t-N_s=k\right) = \int_{\mathbb{R}^k}\sum_{i=N_s+1}^{N_s+k}\mathbf{E}\left(g(t,t_i,V_i)\right)d\tilde{H}_k(t_{N_s+1},\dots,t_{N_s+k}).$$

For a Poisson process, the conditional distribution of occurrence times is perfectly known to be the order statistic of k i.i.d. uniformly distributed random variables (see e.g. Kingman (1992)), i.e. the density is $\tilde{h}_k(t_1,\dots,t_k) = \frac{k!}{(t-s)^k}\mathbb{1}_{\{s<t_{N_s+1}<\dots<t_{N_s+k}<t\}}$. Then,

$$\begin{aligned}
&\sum_{i=N_s+1}^{N_s+k}\mathbf{E}\left(g(t,T_i,V_i)\,|\,N_s,N_t-N_s=k\right)\\
&= \frac{k!}{(t-s)^k}\int_{\mathbb{R}^k}\sum_{i=N_s+1}^{N_s+k}\mathbf{E}\left(g(t,t_i,V_i)\right)\mathbb{1}_{\{s<t_{N_s+1}<\dots<t_{N_s+k}<t\}}dt_{N_s+1}\dots dt_{N_s+k}\\
&= \frac{k!}{(t-s)^k}\int_{\mathbb{R}^k}\sum_{i=1}^{k}\mathbf{E}\left(g(t,t_i,V)\right)\mathbb{1}_{\{s<t_1<\dots<t_k<t\}}dt_1\dots dt_k,
\end{aligned}$$

with $t_{k+1} = t$. Using Appendix A, the previous sum is $A_k(G)(s,t)$. Hence,

$$\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t)\,|\,N_s\right) = \mathbf{E}(X)e^{-\lambda(t-s)}\sum_{k=1}^{\infty}\lambda^k A_k(G)(s,t). \tag{5}$$

□

Proposition 3.2. *The conditional second-order moment at time t of the aggregate claim amount knowing the information up to time $s < t$ is*

$$
\begin{aligned}
\mathbf{E}\left(S_t^2 \,|\, \mathcal{F}_s^{N,C,X}\right) &= \sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_i X_j \mathbb{1}_{i,j\in\mathcal{C}_s^s} + 2\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_i X_j G(t,T_i)\mathbb{1}_{i\in\mathcal{C}_s^{ns}, j\in\mathcal{C}_s^s} \\
&\quad + \sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_i X_j G(t,T_i)G(t,T_j)\mathbb{1}_{i,j\in\mathcal{C}_s^{ns}} \\
&\quad + 2\mathbf{E}(X)e^{-\lambda(t-s)}\sum_{k=1}^{\infty}\lambda^k A_k(G)(s,t)\sum_{i=1}^{N_s}\left(X_i\mathbb{1}_{i\in\mathcal{C}_s^s} + X_i\mathbb{1}_{i\in\mathcal{C}_s^{ns}}G(t,T_i)\right) \\
&\quad + e^{-\lambda(t-s)}\sum_{k=1}^{\infty}\lambda^k\left(\mathbf{E}(X^2)A_k(G_2)(s,t) + 2\mathbf{E}(X)^2 A_k^*(\mathbf{G})(s,t)\right),
\end{aligned}
$$

where $G(t,w) = \mathbf{E}(g(t,w,V))$, $G_2(t,w) = \mathbf{E}(g(t,w,V)^2)$. Here $A_k(\cdot)$ is defined in (3) *and $A_k^*(\cdot)(s,t)$ is*

$$
A_k^*(\mathbf{G})(s,t) = \sum_{i=1}^{k-1}\int_s^{t_{k+1}}\cdots\int_s^{t_{i+1}}\frac{(t_i-s)^{i-1}}{(i-1)!}G^{i,k}(t,t_i,\ldots,t_k)dt_i\ldots dt_k \quad t_{k+1}=t, \quad s<t, \tag{6}
$$

with $\mathbf{G} = (G^{1,k},\ldots,G^{k,k})$ and $G^{i,k}(t,w_i,..,w_k) = \sum_{m=i+1}^{k} G(t,w_i)G(t,w_m)$.

It is worth emphasizing that $A_k(G)(s,t)$ defined in (3) is only a particular case of the operator $A_k^*(\mathbf{G})(s,t)$ with a family of bivariate functions, namely $\mathbf{G} = \mathbf{G}_{bi} = (G_{bi}^{1,k},\ldots,G_{bi}^{k,k}) = (G(t,t_1),\ldots,G(t,t_k))$. From the first two moments, the computation of the conditional variance is immediate:

$$
\mathbf{Var}\left(S_t \,|\, \mathcal{F}_s^{N,C,X}\right) = \mathbf{E}\left((S_t)^2 \,|\, \mathcal{F}_s^{N,C,X}\right) - \mathbf{E}\left(S_t \,|\, \mathcal{F}_s^{N,C,X}\right)^2.
$$

Proof. Direct computation leads to

$$
\begin{aligned}
\mathbf{E}\left((S_t)^2 \,|\, \mathcal{F}_s^{N,C,X}\right) &= \mathbf{E}\left(\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) + \mathbf{E}\left(\sum_{i=N_s+1}^{N_t}\sum_{j=1}^{N_s} Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) \\
&\quad + \mathbf{E}\left(\sum_{i=1}^{N_s}\sum_{j=N_s+1}^{N_t} Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) + \mathbf{E}\left(\sum_{i=N_s+1}^{N_t}\sum_{j=N_s+1}^{N_t} Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) \\
&= \sum_{i=1}^{N_s}\sum_{j=1}^{N_s}\mathbf{E}\left(Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) + 2\sum_{j=1}^{N_s}\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) \\
&\quad + \mathbf{E}\left(\sum_{i=N_s+1}^{N_t}\sum_{j=N_s+1}^{N_t} Z_i(t)Z_j(t) \,|\, N_s\right).
\end{aligned}
$$

Using $\mathcal{C}_s^s \cup \mathcal{C}_s^{ns} = \{1,\ldots,N_s\}$ and $Z_i(t) = X_i\mathbb{1}_{i\in\mathcal{C}_s^s} + X_i g(t,T_i,V_i)\mathbb{1}_{i\in\mathcal{C}_s^{ns}}$, we split the first three sums. The first sum simplifies to

$$
\begin{aligned}
&\sum_{i=1}^{N_s}\sum_{j=1}^{N_s}\mathbf{E}\left(Z_i(t)Z_j(t) \,|\, \mathcal{F}_s^{N,C,X}\right) \\
=\; &\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_i X_j \mathbb{1}_{i,j\in\mathcal{C}_s^s} + 2\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_i X_j G(t,T_i)\mathbb{1}_{i\in\mathcal{C}_s^{ns}, j\in\mathcal{C}_s^s} + \sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_i X_j G(t,T_i)G(t,T_j)\mathbb{1}_{i,j\in\mathcal{C}_s^{ns}},
\end{aligned}
$$

since $\mathbf{E}\left(g(t,T_i,V_i)g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C,X}\right) = \mathbf{E}\left(g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N,C,X}\right)\mathbf{E}\left(g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C,X}\right)$ with $1 \le i,j \le N_s$.

Since for $i > N_s$, $Z_i(t) = X_i g(t, T_i, V_i)$ while for $j \leq N_s$, claims are reported (yet settled or not) $Z_j(t) = X_j \mathbb{1}_{j \in \mathcal{C}_s^s} + X_j g(t, T_j, V_j) \mathbb{1}_{j \in \mathcal{C}_s^{ns}}$, the second sum simplifies to

$$
\begin{aligned}
& \sum_{j=1}^{N_s} \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t) Z_j(t) \,|\, \mathcal{F}_s^{N,C,X} \right) \\
= & \sum_{j=1}^{N_s} \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} X_i g(t,T_i,V_i) X_j \mathbb{1}_{j\in\mathcal{C}_s^s} + X_i g(t,T_i,V_i) X_j g(t,T_j,V_j) \mathbb{1}_{j\in\mathcal{C}_s^{ns}} \,|\, \mathcal{F}_s^{N,C,X} \right) \\
= & \sum_{j=1}^{N_s} X_j \mathbb{1}_{j\in\mathcal{C}_s^s} \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} X_i g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N,C,X} \right) + \sum_{j=1}^{N_s} X_j \mathbb{1}_{j\in\mathcal{C}_s^{ns}} \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} X_i g(t,T_i,V_i) g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C,X} \right).
\end{aligned}
$$

The first term $\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} X_i g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N,C,X}\right)$ corresponds to the last term of Equation (5). For the second term, since $j \in \mathcal{C}_s^{ns}$, we obtain

$$
\begin{aligned}
& \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} X_i g(t,T_i,V_i) g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C,X} \right) \\
= & \mathbf{E}\left(\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} X_i g(t,T_i,V_i) g(t,T_j,V_j) \,|\, \mathcal{F}_t^{N,V} \right) \,|\, \mathcal{F}_s^{N,C} \right) \\
= & \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} g(t,T_i,V_i) g(t,T_j,V_j) \mathbf{E}\left(X_i \,|\, \mathcal{F}_t^{N,V} \right) \,|\, \mathcal{F}_s^{N,C} \right) \\
= & \mathbf{E}(X) \sum_{\ell=N_s+1}^{\infty} \mathbf{E}\left(\sum_{i=N_s+1}^{\ell} g(t,T_i,V_i) g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C}, N_t = \ell \right) P(N_t = \ell \,|\, N_s) \\
= & \mathbf{E}(X) \sum_{\ell=N_s+1}^{\infty} \sum_{i=N_s+1}^{\ell} \mathbf{E}\left(g(t,T_i,V_i) g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C}, N_t = \ell \right) P(N_t = \ell \,|\, N_s).
\end{aligned}
$$

However

$$
\begin{aligned}
\mathbf{E}\left(g(t,T_i,V_i) g(t,T_j,V_j) \,|\, \mathcal{F}_s^{N,C}, N_t = \ell \right) &= \mathbf{E}\left(g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N,C}, N_t = \ell \right) \mathbf{E}\left(g(t,T_j,V_j) \right) \\
&= \mathbf{E}\left(g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N}, N_t = \ell \right) \mathbf{E}\left(g(t,T_j,V) \right).
\end{aligned}
$$

This yields

$$
\begin{aligned}
= & \mathbf{E}(X) \sum_{k=1}^{\infty} \sum_{i=N_s+1}^{N_s+k} \mathbf{E}\left(g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N}, N_t - N_s = k \right) \mathbf{E}\left(g(t,T_j,V) \right) P(N_t - N_s = k \,|\, N_s) \\
= & \mathbf{E}(X) \sum_{k=1}^{\infty} \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)} \sum_{i=N_s+1}^{N_s+k} \mathbf{E}\left(g(t,T_i,V_i) \,|\, \mathcal{F}_s^{N}, N_t - N_s = k \right) \mathbf{E}\left(g(t,T_j,V) \right) \\
= & \mathbf{E}(X) G(t,T_j) \sum_{k=1}^{\infty} \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)} \sum_{i=N_s+1}^{N_s+k} \mathbf{E}\left(g(t,T_i,V_i) \,|\, N_s, N_t - N_s = k \right).
\end{aligned}
$$

The inner term has been already computed in the proof of Proposition 3.1.

$$
\sum_{i=N_s+1}^{N_s+k} \mathbf{E}\left(g(t,T_i,V_i) \,|\, N_s, N_t - N_s = k \right) = \frac{k!}{(t-s)^k} A_k(G)(s,t).
$$

Finally, the second sum is

$$
\begin{aligned}
&\sum_{j=1}^{N_s} \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t) Z_j(t) \,|\, \mathcal{F}_s^{N,C,X} \right) \\
=\; &\sum_{j=1}^{N_s} X_j \mathbb{1}_{j\in\mathcal{C}_s^s} \mathbf{E}(X) e^{-\lambda(t-s)} \sum_{k=1}^{\infty} \lambda^k A_k(G)(s,t) + \sum_{j=1}^{N_s} X_j \mathbb{1}_{j\in\mathcal{C}_s^{ns}} \mathbf{E}(X) G(t,T_j) e^{-\lambda(t-s)} \sum_{k=1}^{\infty} \lambda^k A_k(G)(s,t) \\
=\; &\mathbf{E}(X) \sum_{j=1}^{N_s} \left(X_j \mathbb{1}_{j\in\mathcal{C}_s^s} + X_j \mathbb{1}_{j\in\mathcal{C}_s^{ns}} G(t,T_j) \right) e^{-\lambda(t-s)} \sum_{k=1}^{\infty} \lambda^k A_k(G)(s,t).
\end{aligned}
$$

Using Appendices B and C, the third sum is

$$
\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} \sum_{j=N_s+1}^{N_t} Z_i(t) Z_j(t) \,|\, N_s \right) \;=\; \sum_{k=1}^{\infty} \lambda^k e^{-\lambda(t-s)} \left(\mathbf{E}(X^2) A_k(G_2)(s,t) + 2\mathbf{E}(X)^2 A_k^*(\mathbf{G})(s,t) \right),
$$

where $\mathbf{G} = (G^{1,k}, \ldots, G^{k,k})$ and $G^{i,k}(t, w_i, .., w_k) = \sum_{m=i+1}^{k} G(t, w_i, w_m)$. Here

$$
G(t, w_i, w_m) = \mathbf{E}\left(g(t, w_i, V_i) g(t, w_k, V_m) \right) = \mathbf{E}(g(t, w_i, V_i)) \mathbf{E}(g(t, w_k, V_m)) = G(t, w_i) G(t, w_m).
$$

□

From the conditional moments, the corresponding unconditional moments can be derived. From the first two moments, the computation of the unconditional variance is also immediate:

$$
\mathbf{Var}\,(S_t) = \mathbf{E}\left((S_t)^2\right) - \mathbf{E}\,(S_t)^2 .
$$

Proposition 3.3. *The expectation of the aggregate claim amount is*

$$
\mathbf{E}\,(S_t) = \mathbf{E}(X) e^{-\lambda t} \sum_{k=1}^{\infty} \lambda^k A_k(G)(0,t), \tag{7}
$$

where $A_k(G)(0,t)$ is A_k is defined in (3) *and G is the bivariate function defined as $G(t,w) = \mathbf{E}(g(t,w,V))$.*

Proposition 3.4. *The second-order moment of the aggregate claim amount is*

$$
\mathbf{E}(S_t^2) = e^{-\lambda t} \sum_{k=1}^{\infty} \lambda^k \left(\mathbf{E}(X^2) A_k(G_2)(0,t) + 2\mathbf{E}(X)^2 A_k^*\left(\mathbf{G}\right)(0,t) \right),
$$

where A_k, A_k^* are defined in (3) *and* (6) *respectively and G_2 is previously defined.*

Proof. The proof of Propositions 3.3 and 3.4 are obtained by setting $s = 0$ (i.e. $\mathcal{C}_s^s = \mathcal{C}_s^{ns} = \emptyset$) in Propositions 3.1 and 3.2 respectively. □

3.2 Relevant examples of settlement-linked functions

In this subsection, we present two examples of settlement functions g. Direct computation of (4) leads to

$$
G(t,w) = \mathbf{E}(g(t,w,V)) = (t-w) \int_{t-w}^{\infty} \frac{dF_V(x)}{x} + F_V(t-w),
$$

for $t \geq w \geq 0$. In order to compute the second order moment, similar computations lead to

$$
G_2(t,w) = \mathbf{E}(g(t,w,V)^2) = (t-w)^2 \int_{t-w}^{\infty} \frac{dF_V(x)}{x^2} + F_V(t-w).
$$

The ith component of $\mathbf{G} = (G^{1,k}, \dots, G^{k,k})$ consists in summing G functions, namely

$$G^{i,k}(t, w_i, .., w_k) = \sum_{m=i+1}^{k} G(t, w_i)G(t, w_m).$$

Let us start with the usual case of immediate settlement. If $V = 0$ a.s., then $G(t, w) = \mathbb{1}_{\{w \le t\}}$, leading to $A_k(G)(t) = t^k/(k-1)!$. Therefore, Proposition 3.3 gets back to a well known result $\mathbf{E}(S_t) = \lambda t \mathbf{E}(X)$. Consequently, we also have $G_2(t, w) = \mathbb{1}_{\{w \le t\}}$. Then, $A_k(G_2)(t) = t^k/(k-1)!$. Furthermore,

$$G^{i,k,\Sigma}(t, w_i, .., w_k) = \mathbb{1}_{\{w_i \le t\}} \sum_{m=i+1}^{k} \mathbb{1}_{\{w_m \le t\}} \Rightarrow A_k^*(\mathbf{G})(0, t) = \sum_{i=1}^{k-1} (k-i)\frac{t^k}{k!} = \frac{t^k}{2(k-2)!}.$$

Thus Proposition 3.4 gives another well-known result of a compound Poisson process

$$\begin{aligned}
\mathbf{E}(S_t^2) &= e^{-\lambda t}\sum_{k=1}^{\infty}\lambda^k\left(\mathbf{E}(X^2)\frac{t^k}{(k-1)!} + \mathbf{E}(X)^2\frac{t^k}{(k-2)!}\right) \\
&= \mathbf{E}(X^2)\lambda t e^{-\lambda t}\sum_{k=1}^{\infty}\frac{(\lambda t)^{k-1}}{(k-1)!} + e^{-\lambda t}\mathbf{E}(X)^2(\lambda t)^2\sum_{k=1}^{\infty}\lambda^{k-2}\frac{t^{k-2}}{(k-2)!} \\
&= \mathbf{E}(X^2)\lambda t + \mathbf{E}(X)^2(\lambda t)^2.
\end{aligned}$$

Short-tailed business (such as material damages for motor and household insurance with settlement generally within four or five years) corresponds to line of business where the settlement time is either quick or immediate. As the opposite, long-tailed business such as third-party liability (especially medical malpractice or liability for lawyers) experiences very long development of claims (generally more than to 20 years, see e.g. Partrat et al. (2008)). Hence, modeling the settlement process depends heavily on the studied guarantee. This paper first attempt to model such a process using the zero-inflated exponential distribution. In the numerical section, we will split the dataset between long and short tailed guarantees leading to distinct values of parameters of the two situations. We postpone the use of more complex distributions (such as Gamma or Weibull and their zero-inflated version) to future research.

Considering a zero-inflated exponential distribution for V (i.e. a mixture of a geometric distribution and a Dirac distribution at 0) yields to

$$F_{ZIE}(x) = (p + (1-p)(1 - e^{-\mu x}))\mathbb{1}_{[0,+\infty)}(x).$$

In other words with probability p, the claim is settled immediately, otherwise (with probability $1-p$) the settlement time is strictly positive. Hence, for $t > w$,

$$\begin{aligned}
G_{ZIE}(t, w) &= p + (1-p)(t-w)\int_{t-w}^{\infty}\frac{\mu e^{-\mu x}}{x}dx + (1-p)(1 - \exp(-\mu(t-w))) \\
&= 1 - (1-p)\exp(-\mu(t-w)) + (1-p)\mu(t-w)\mathrm{E}_1(\mu(t-w)),
\end{aligned}$$

where E_1 denotes the exponential integral, see e.g. (Olver et al., 2010, Chap. 6).

$$\begin{aligned}
G_{ZIE,2}(t, w) &= p + (1-p)(t-w)^2\int_{t-w}^{\infty}\frac{\mu e^{-\mu x}dx}{x^2} + (1-p)(1 - \exp(-\mu(t-w))) \\
&= 1 - (1-p)\exp(-\mu(t-w)) + (1-p)\mu(t-w)E_2(\mu(t-w)),
\end{aligned}$$

where E_2 denotes the generalized exponential integral, see e.g. (Olver et al., 2010, Chap. 8). Finally,

$$G_{ZIE}^{i,k,\Sigma}(t, w_i, .., w_k) = \sum_{m=i+1}^{k} G_{ZIE}(t, w_i)G_{ZIE}(t, w_m).$$

Of course, the case of the exponential distribution is obtained by setting $p = 0$ in the previous expressions of G_{ZIE} and $G_{ZIE,2}$.

4 Computing classic actuarial indicators

In this section, we present different indicators starting with insurance reserving and then ruin theory.

4.1 Reserving topics

From a reserving perspective, we now ignore the initial capital u and the premium rate c and focus on the aggregate claim amount S_t at time t. Classical methods for claim reserving are designed for aggregated data for which claim amounts are aggregated per accident year and per development year, see e.g. Wuethrich and Merz (2008). Therefore, claims are sorted per accident year and cumulated per development year to get a so-called claims development triangle.

At individual claim level, the accident year k of a claim occurred at time T is the year of occurrence, i.e. $k = \lfloor T \rfloor$ (where $\lfloor . \rfloor$ denotes the integer part). The jth development year of a claim occurred at time T corresponds to payments done in interval $(\lfloor T \rfloor + j - 1, \lfloor T \rfloor + j)$. Let $k = 0, \dots, K$ be an accident year and $j = 0, \dots, J$ a development year. As before, we want to deal with reserving topics, and we introduce the claim set of accident year k reported at time t

$$\mathcal{C}_{t,k} = \{i \in \{1, \dots, N_t\}, k = \lfloor T_i \rfloor\}.$$

Note that the current time is $k + j + 1$ since both k and j starts from 0. Let us define the aggregate (paid) claim amount for accident year k and development year j

$$S_{k,j} = \sum_{i \in \mathcal{C}_{k+j+1,k}} Z_i(j + k + 1) = \sum_{i=1}^{N_{j+k+1}} Z_i(j + k + 1) \mathbb{1}_{i \in \mathcal{C}_{k+j+1,k}}.$$

The sum can be expressed as in the previous subsection using $Z_i(t) = X_i g(t, T_i, V_i)$

$$S_{k,j} = \sum_{i=1}^{N_{j+k+1}} X_i g(j + k + 1, T_i, V_i) \mathbb{1}_{\{k \le T_i < k+1\}}.$$

Denoting $\tilde{g}_k(y, t, v) = g(y, t, v) \mathbb{1}_{\{k \le t < k+1\}}$, we get back to a sum similar the aggregate claim S_t at time $t = j + k + 1$ with a new settlement function $\tilde{g}_k(y, t, v)$. This leads to the following property.

In order to deal with conditional expectation, we split the claim set into two subsets

$$\mathcal{C}^{ns}_{t,k} = \{i \in \{1, \dots, N_{j+k+1}\}, k = \lfloor T_i \rfloor, T_i \le t < T_i + V_i\},$$

$$\mathcal{C}^{s}_{t,k} = \{i \in \{1, \dots, N_{j+k+1}\}, k = \lfloor T_i \rfloor, T_i + V_i \le t\}.$$

They represent claims of accident year k not-settled and settled at time t.

Let us define the aggregate (paid) claim amount for accident year k and development year $j+m$ given that the current time is $k + j + 1$

$$S_{k,j+m} = \sum_{i=1}^{N_{j+k+1}} X_i g(j + m + k + 1, T_i, V_i) \mathbb{1}_{\{k \le T_i < k+1\}}.$$

Corollary 4.1. *The conditional expectation of the aggregate claim after $j+m$ development years amount knowing the information up to time $s = k + j + 1$ is*

$$\mathbf{E}\left(S_{k,j+m} \,|\, \mathcal{F}^{N,C,X}_s\right) = \sum_{i=1, i \in \mathcal{C}^s_{s,k}}^{N_{k+j+1}} X_i + \sum_{i=1, i \in \mathcal{C}^{ns}_{s,k}}^{N_{k+j+1}} X_i \tilde{G}_k(s + m, T_i) \quad (8)$$

where $\tilde{G}_k(t, u) = \mathbf{E}\left(g(y, t, V) \mathbb{1}_{\{k \le t < k+1\}}\right)$.

Proof. Immediate by taking $s = j + k + 1$, $t = j + k + 1 + m = s + m$ and $\tilde{g}$ in Proposition 3.1 in which the term $A_n(\tilde{G}_k)(s, s + m)$ cancels. □

Within this notation, a reserving triangle looks like (for $s = 3$)

AY k \ DY j	0	1	2
0	$S_{0,0}$	$S_{0,1}$	$S_{0,2}$
1	$S_{1,0}$	$S_{1,1}$	$\mathbf{E}\left(S_{1,2} \mid \mathcal{F}_3^{N,C,X}\right)$
2	$S_{2,0}$	$\mathbf{E}\left(S_{2,1} \mid \mathcal{F}_3^{N,C,X}\right)$	$\mathbf{E}\left(S_{2,2} \mid \mathcal{F}_3^{N,C,X}\right)$

Corollary 4.2. *The conditional second-order moment of the aggregate claim after $j + m$ development years amount knowing the information up to time $s = k + j + 1$ is*

$$\mathbf{E}\left(S_{k,j+m}^2 \mid \mathcal{F}_s^{N,C,X}\right) = \sum_{i=1}^{N_s}\sum_{j=1}^{N_s}(X_iX_j \mathbb{1}_{i,j\in\mathcal{C}_{s,k}^s} + 2X_iX_j\tilde{G}_k(s+m, T_i)\mathbb{1}_{i\in\mathcal{C}_{s,k}^{ns}, j\in\mathcal{C}_{s,k}^s}) + \sum_{i=1}^{N_s}\sum_{j=1}^{N_s} X_iX_j\tilde{G}_k(s+m, T_i)\tilde{G}_k(s+m, T_j)\mathbb{1}_{i,j\in\mathcal{C}_{s,k}^{ns}}, \quad (9)$$

where $\tilde{G}_k(t, u) = \mathbf{E}\left(g(y, t, V)\mathbb{1}_{\{k\le t<k+1\}}\right)$.

Proof. Immediate by taking $s = j + k + 1$, $t = j + k + 1 + m = s + m$ and $\tilde{g}$ in Proposition 3.2 in which the terms $A_n(\tilde{G}_k)(s, s + m)$, $A_n(\tilde{G}_{2,k})(s, s + m)$, $A_m^*(\mathbf{G}_k)(s, s + m)$ cancel. □

4.2 Ruin topics

Within regulatory frameworks, the computation of solvency probabilities for different line of business and different time horizons are increasingly studied. Generally, the solvency probability at time t is defined as

$$\phi_t(u) = P(R_t > 0 \mid R_0 = u).$$

This quantity depends on the initial capital u, the premium rate c and the probability law of R_t. This indicator allows to calibrate the premium rate c for a risk management perspective.

The computation of the solvency or equivalently of the ruin probability has been studied for decades in the literature, see Asmussen and Albrecher (2010). In the case of independence between claim arrivals and claim amounts, there are mainly two cases to distinguish depending on the tail heaviness of the claim distribution. In the light-tailed case, exact formulas are available when the claim distribution is an exponential distribution or a phase-type distribution (which includes the Erlang distribution and mixtures of Erlang distributions), see e.g. Asmussen and Rolski (1991). In the heavy-tailed case, exact formulas are rare and we must rely on integrated tail approximation or bounds of the ruin probability. Furthermore in the light-tailed case, the solvency probability can be approximated by a normal distribution (Asmussen and Albrecher, 2010, Chap. 16) or a translated gamma distribution ((Dickson, 2005, Chap. 4)) based on two or more moments of the claim distribution.

In this paper, we simply use the two-moment normal approximation of the distribution of the risk process R_t at time t. That is, the solvency probability is approximating only through the mean $\mathbf{E}(R_t \mid R_0 = u)$ and the variance $\mathbf{Var}(R_t \mid R_0 = u)$. These previous characteristics can be computed in our model via Propositions 3.3 and 3.4. Indeed,

$$\mathbf{E}(R_t \mid R_0 = u) = u + ct - \mathbf{E}(S_t),\ \mathbf{Var}(R_t \mid R_0 = u) = \mathbf{Var}(S_t).$$

In this context, the solvency probability can be expressed as

$$\phi_t(u) \approx 1 - \Phi\left(\frac{0 - (u + ct - \mathbf{E}(S_t))}{\sqrt{\mathbf{Var}(S_t)}}\right), \quad (10)$$

where Φ denotes the cumulative distribution function of the standard Gaussian distribution.

5 Numerical illustrations

In this section, we will present the numerical illustrations of the previous results and we compare them with the classical ruin theory model and chain-ladder method for claim reserving. On a real dataset, we show the numerical computation of the ruin probability and claim reserving. In this example, we choose to model settlement times with zero-inflated exponential distribution of parameter p and μ.

We consider an actuarial dataset from an unknown private insurer on a portfolio of general third-party liability policies for private individuals. 332,892 claims were reported between January 1990 and December 1999, which are all closed on December 31 2008. We distinguish two types of claims: material damage (material) and bodily injuries (injury). We randomly select 6000 claims in the previous dataset. Then, we prepare the training set with claims reported before December 1997.

Parameters of the model are calibrated with classical statistical methods. We estimate the parameters λ, $\nu = \mathbf{E}(X)$, p and μ with the estimators $\widehat{\lambda} = \frac{N_s}{s}$, $\widehat{\nu} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\widehat{p} = \frac{1}{n}\sum_{i=1}^{N} \mathbb{1}_{\{V_i=0\}}$ and $\widehat{\mu} = \left(\frac{1}{n}\sum_{i=1}^{n} V_i\right)^{-1}(1-\widehat{p})$ respectively. $\widehat{p}$, $\widehat{\mu}$ are maximum likelihood estimators, whereas $\widehat{\lambda}$ and $\widehat{\nu}$ are moment-based estimators.

	ν	λ	p	μ	$(1-p)/\mu$
material	1178	498.6	0.032	0.269	3.606
body	21520	105.5	0.004	0.233	4.272

Table 1 – Fitted parameters of the training set

In Table 1, we can observe the key characteristics of each type of claims. The mean value of the claim amount is much bigger for bodily injuries ($\nu_{injury} > \nu_{material}$). On this dataset, material damages are more frequent than bodily injuries ($\lambda_{material} > \lambda_{injury}$). Furthermore with our model, we can see that bodily injuries are rarely paid immediately compared to material damages, since p_{injury} is much smaller than $p_{material}$. By direct computation (using $(1-\widehat{p})/\widehat{\mu}$), the expected settlement time of bodily injuries is bigger than for material damages (see corresponding column in Table 1).

In Figure 2, we plot the distribution function (both the empirical and the fitted functions respectively in solid and dashed lines) for the two claim types (left for material and right for injury). We also observe that the zero-inflated exponential distribution, considered in our model (see section 3.2), fits reasonably well the settlement times, yet distributions with more parameters will better fit the empirical distribution.

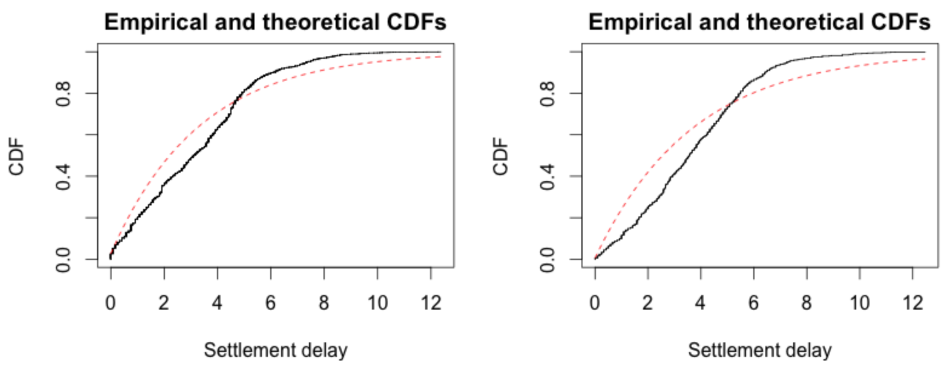

Figure 2 – Settlement delays $(V_i)_i$ (left for material and right for injury)

5.1 Computational aspects

In this numerical section, we use the statistical software R (R Core Team (2016)), a personal library of dedicated functions hosted on a git repository at https://portail.math.cnrs.fr/, the ChainLadder and the fitdistrplus packages (see Gesmann et al. (2015) and Delignette-Muller and Dutang (2015)). We use numerical and/or Monte-Carlo approaches in order to compute the key indicators previously presented in Sections 3 and 4.

Firstly, for the numerical approach we fully used the efficient formulation of indicators. In fact, those formulations of the indicators rely on the inversion of sums and multiple integrals which simplifies the computation (see Appendices A, B and C). These multiple integrals can be approximated via the rectangle rule, see Appendix D for details. Benchmarks of the computation of theses indicators have been carried out in the case of immediate settlement (that is $G(t, w) = \mathbb{1}_{\{w \leq t\}}$).

Secondly, the Monte-Carlo approach consists in simulating both the claim occurence times (T_i) and the settlement time (V_i), then mean and variance are replaced by their empirical versions. There is no issue to simulate T_i as they are i.i.d. uniformly distributed between $[s, t]$ knowing N_t and to simulate V_i since they are zero-inflated exponentially distributed. In the following subsections, we choose a number of simulations equals to 1000.

For large values of λt, we prefer the Monte-Carlo approach because the numerical approach needs a sharp discretization grid in time combined with a dedicated library. Indeed, computing large binomial sums needs the use high precision floating-point arithmetic libraries such as the GMP library of Grandlund Torbjoern & the GMP Devel. Team (2015), or the MPFR library of Fousse et al. (2007). These libraries are available thanks to the R package Rmpfr of Maechler (2016). This makes the numerical approach slow even with parallel computations.

5.2 Reserving

For illustration purposes, the reserving triangles (with cumulative incurred amounts) are displayed in Tables 5 and 6 in Appendix E. The predicted claim charge has been computed by the numerical approach using Equation (8) in Table 2 for damage cover and Table 3 for bodily injury cover. Each table contains the latest known value in the first column, the predicted claim charge by the numerical approach and the Chain Ladder method (see e.g. Wuethrich and Merz (2008)) in the second and third columns, the claim reserves denoted by IBNR corresponding to the numerical approach and the Chain Ladder method in the fourth and fifth columns.

Accident year	Latest	Ultimate CL	Ultimate New	IBNR CL	IBNR New
1990	602 261	602 261	602 261	0	0
1991	774 350	800 779	861 315	26 429	86 965
1992	580 910	648 991	593 044	68 081	12 134
1993	419 019	519 118	484 517	100 099	65 498
1994	568 651	840 366	642 925	271 715	74 274
1995	285 542	529 908	348 506	244 366	62 964
1996	331 037	866 481	644 304	535 444	313 267
1997	93 018	499 976	427 523	406 958	334 507
Total	3 654 788	5 307 880	4 604 394	1 653 091	949 609

Table 2 – Results for damage cover

In Table 2, we observe that the numerical approach proposed in this paper mostly underestimate the ultimate predicted claim charge for both guarantees for all accident year (except for 1991). This leads to an underestimation of the total claim charge (5 307 880 vs. 4 604 394). Obviously, when subtracting the latest known claim value, the claim reserve (IBNR columns) are also lower for the numerical approach than for the Chain Ladder method.

Accident year	Latest	Ultimate CL	Ultimate New	IBNR CL	IBNR New
1990	4 365 142	4 365 142	4 365 142	0	0
1991	3 466 759	3 639 763	3 518 086	173 004	51 327
1992	2 474 422	2 779 462	2 530 368	305 040	55 946
1993	797 925	1 007 187	906 636	209 262	108 711
1994	1 047 918	1 608 889	1 246 146	560 972	198 228
1995	626 053	1 269 117	995 574	643 064	369 521
1996	819 551	2 542 603	1 764 573	1 723 052	945 021
1997	404 867	3 045 051	1 552 357	2 640 184	1 147 490
Total	14 002 636	20 257 213	16 878 881	6 254 577	2 876 246

Table 3 – Results for body cover

In Table 3, we observe that this effect is even more pronounced. For the bodily injury cover, the ultimate predicted claim is lower with the numerical approach than with the Chain Ladder method for every accident year without exception. Chain Ladder reserve's estimate are generally two times or three times bigger than the numerical approach.

These results must be taken with care (especially for the bodily injury cover) because the presented estimation do not take into account development after the seventh year. For long-tail business such as bodily injury cover, this is not recommended to do so. Furthermore, the claim triangles (see Appendix E) present some accident year effect probably due to the portfolio size under exposure: for the damage cover, accident year 1994's claim charges are particularly large, whereas for the bodily injury cover, accident years 1990-91's claim charge are heavy.

However, these two effects probably do not explain all differences. The remaining differences may be explained by two reasons: the proposed model has less parameters than the Chain Ladder method (2 vs. 7); the numerical approach does not take into account claims that are incurred but not year reported (IBNYR).

These differences reduce when looking to a smaller time horizon. In fact using the 1998's claim information (for validation), we compare the claim charges observed for accident years 1991-1997 with the prediction by the numerical approach and the Chain Ladder method. For the damage cover, the numerical approach and the Chain Ladder method produce similar estimates: the numerical approach is even better. For the bodily injury cover, the predictions for the numerical approach are dubious especially for the most recent accident years where the uncertainty implied by IBNYR claims is prominent. This is logical because there are almost no IBNYR claims for the damage guarantee but many late IBNYR for the bodily injury guarantee.

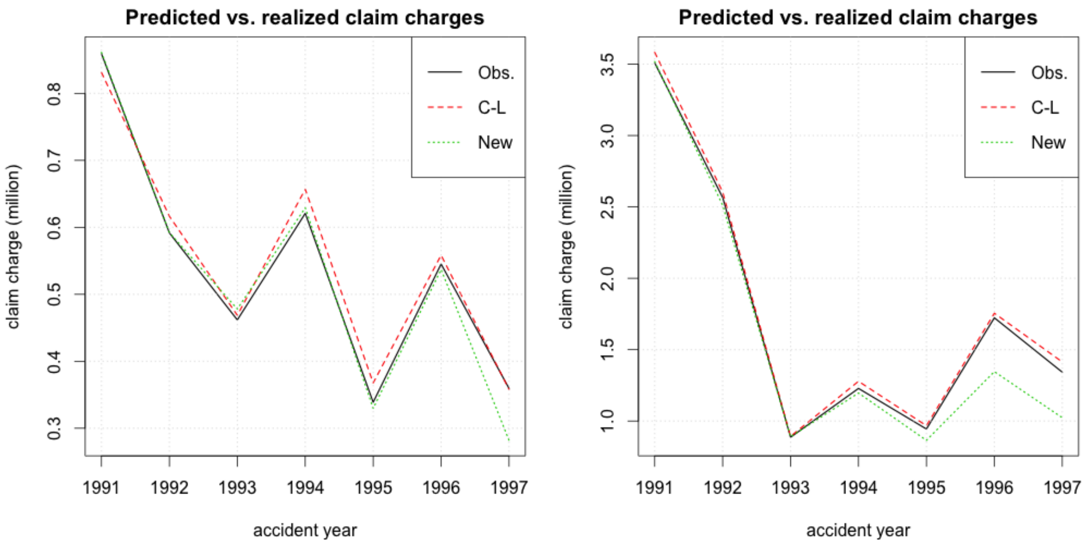

Figure 3 – Prediction of claim charge over one-year time horizon

5.3 Ruin

In this section, we consider the solvency probability approximation using Equation (10). As generally in numerical illustrations of ruin-related quantities, $\phi_t(u)$ is plotted as a function of the initial capital u for different time values t. On Figure 4, we plot the solvency probability in 8 situations by considering (i) two time horizon $t = 1$ and $t = 2$, (ii) two variance values for the claim distribution $Var(X) = 1/2$ and $Var(X) = 2$ (whereas the mean is $\mathbf{E}(X) = 1$) and (iii) two type of settlement Dirac and exponential. We do not plot the zero-inflated case since it is an intermediate situation between these two situations. Finally, the loading factor is assumed to be 1%, i.e. $c = \mathbf{E}(X)(1 + 1\%)$.

For each situation, the solvency probability has an exponential convergence towards 1 since Equation (10) use the distribution function of the standard normal distribution. As expected, increasing the time horizon given a value of initial capital leads to a (sharp) decrease of the solvency probability. Also as one would expect, increasing the variance of the claim amounts decrease the solvency probabilities: the slopes are flatter in the right-hand graph.

Furthermore, the two types of settlements either immediate with a Dirac distribution or gradually with an exponential distribution (of mean 1/2) have a large impact of the solvency probability. As the ruin process in these two situations are stochastically ordered, having a random settlement time leads to higher solvency probabilities.

In Table 4, we compute the solvency capital at 99.5% level, that is the difference between the quantile at 99.5% level and the mean of the aggregate claim. Increasing the time horizon leads to a larger increase of required capital when the settlement time are random: for the exponential distribution from 2.38 to 6.53 compared to the Dirac distribution from 3.14 to 7. However, the effect of random settlement times is reduced as the time horizon increases.

6 Conclusion

We propose in this paper an efficient way to compute insurance indicators in a unified framework of ruin theory and claim reserving. In an extended Cramér-Lundberg framework, we derive efficiently computable closed-form formulas for the key indicators. We illustrate these methods on a real insurance

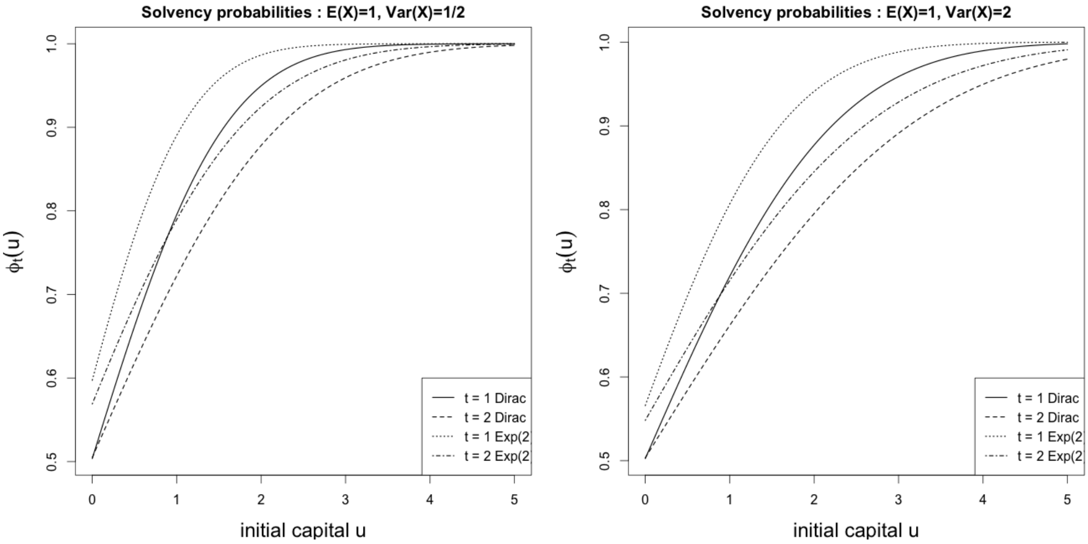

Figure 4 – Solvency probability

	$Var(X) = 1/2$		$Var(X) = 2$	
Time horizon	Dirac	Exp	Dirac	Exp
$t = 1$	3.14	2.38	4.45	3.46
$t = 2$	4.44	3.84	6.29	5.51
$t = 3$	5.43	4.88	7.70	7.00
$t = 4$	6.27	5.73	8.88	8.25
$t = 5$	7.00	6.53	9.93	9.35

Table 4 – Solvency capital at 99.5% level

dataset. This numerical application reveals that the proposed framework underestimates the ultimate claim charges (assuming the Chain-Ladder method is the most appropriate method). On a one-year time horizon, the backtesting procedure shows that the new method to estimate claim charges performs reasonably well. Regarding ruin-theory topics, we retrieve that taking into account settlement times naturally increases solvency probabilities, yet this effect diminishes for longer time horizons.

This extended model is relatively simple and merits further research. Multiple directions can be considered. The next step should attempt to take into account non-null reporting delays as well as random claim charge in order to better assess the reserving risk. By considering random reporting delays, the observed claim process is not a Poisson process, yet the unobserved claim settlement process and reporting process are Poisson processes. This latent model could be better tackled in a general renewal process for the claim process, see (Asmussen and Albrecher, 2010, Chap. 6) or with inhomogeneous Poisson process, see (Wuethrich and Merz, 2008, Chap. 10). Other directions for future research may include the study of more general settlement functions, uncertain claim charges, the dependence between claim sizes and claim inter-arrival times, claim sizes and settlement times, and the asymptotic behavior of the proposed estimators.

Acknowledgements

The authors thank the editor and a referee for their helpful comments which led to significant improvements of this article. The remaining errors, of course, should be attributed to the authors alone.

References

Antonio, K. and Plat, R. (2014), 'Micro-level stochastic loss reserving for general insurance', *Scandinavian Actuarial Journal* **2014**(7), 649–669.

Arjas, E. (1989), 'The claims reserving problem in non-life insurance: some structural ideas', *ASTIN Bulletin* **19**(2), 139–152.

Asmussen, S. and Albrecher, H. (2010), *Ruin Probabilities*, 2nd edn, World Scientific New Jersey.

Asmussen, S. and Rolski, T. (1991), 'Computational method in risk theory: a matrix-algorithmic approach', *Insurance: Mathematics and Economics* **10**, 259–274.

Delignette-Muller, M.L. and Dutang, C. (2015), 'fitdistrplus: An R Package for Fitting Distributions', *Journal of Statistical Software* **64**(4), 1–34.

Dickson, D.C.M. (2005), *Insurance Risk and Ruin*, Cambridge University Press.

Drieskens, D., Henry, M., Walhin, J. and Wielandts, J. (2012), 'Stochastic projection for large individual losses', *Scandinavian Actuarial Journal* **2012**(1), 1–39.

Fousse, L., Hanrot, G., Lefèvre, V., Pélissier, P. and Zimmermann, P. (2011), *MPFR: A multiple-precision binary floating-point library with correct rounding.*
URL: *http://mpfr.org/*

Gesmann, M., Murphy, D., Zhang, W., Carrato, A., Crupi, G., Wuthrich, M. and Concina, F. (2015), ChainLadder: Statistical Methods and Models for Claims Reserving in General Insurance.', *R package version 0.2.1.*

Graham, R., Knuth, D. and Patashnik, O. (1994), *Concrete Mathematics: A Foundation for Computer Science*, Addison-Wesley.

Grandlund Torbjoern & the GMP Devel. Team (2015), *GNU MP - The GNU Multiple Precision Arithmetic Library.*
URL: *http://gmplib.org/*

Haastrup, S. and Arjas, E. (1996), 'Claims reserving in continuous time: a nonparametric bayesian approach', *ASTIN Bulletin* **26**(2), 139–164.

Jewell, W. (1989), 'Predicting IBNYR events and delays. I. continuous time', *ASTIN Bulletin* **19**(1), 25–55.

Karlsson, J. (1976), 'The expected value of IBNR-claims', *Scandinavian Actuarial Journal* **1976**(2), 108–110.

Kingman, J. (1992), *Poisson Processes*, Oxford University Press.

Klueppelberg, C. and Mikosch, T. (1995), 'Explosive Poisson shot noise processess with applications to risk reserves', *Bernoulli* **1**(1/2), 125–147.

Larsen, C. (2007), 'An individual claims reserving model', *ASTIN Bulletin* **37**(1), 113–132.

Maechler, M. (2016), *Rmpfr: R MPFR - Multiple Precision Floating-Point Reliable*, ETH Zurich.

Matsui, M. (2015), 'Prediction in a Poisson cluster model with multiple cluster processses', *Scandinavian Actuarial Journal* **2015**(1), 1–31.

Matsui, M. (2014), 'Prediction in a non-homogeneous Poisson cluster model', *Insurance: Mathematics and Economics* **55**, 10–17.

Matsui, M. and Mikosch, T. (2010), 'Prediction in a Poisson cluster model', *Journal of Applied Probability* **47**(2), 350–366.

Norberg, R. (1993), 'Prediction of outstanding liabilities in non-life insurance', *ASTIN Bulletin* **23**(1), 95–115.

Norberg, R. (1999), 'Prediction of outstanding liabilities ii. model variations and extensions', *ASTIN Bulletin* **29**(1), 5–25.

Olver, F. W. J., Lozier, D. W., Boisvert, R. F. and Clark, C. W., eds (2010), *NIST Handbook of Mathematical Functions*, Cambridge University Press.
URL: *http://dlmf.nist.gov/*

Partrat, C., Lecoeur, E., Nessi, J.-M., Nisipasu, E. and Reiz, O. (2008), *Provisionnement technique en assurance non-vie : Perspectives actuarielles modernes*, Economica.

Pigeon, M., Antonio, K. and Denuit, M. (2013), 'Individual loss reserving with the multivariate skew normal framework', *ASTIN Bulletin* **43**(3), 399–428.

R Core Team (2016), 'R: A language and environment for statistical computing', *R Foundation for Statistical Computing, Vienna, Austria.*

Wuethrich, M. and Merz, M. (2008), *Stochastic Claims Reserving Methods in Insurance*, Wiley Finance.

A Simplification of operator $A_k(G)(s,t)$

Let $t_{k+1} = t$ and $0 \le s < t$. Iterating the following splitting procedure

$$\int_0^{t_2} \sum_{i=1}^k G(t,t_i)\mathbb{1}_{]s,t[}(t_1,\dots,t_k)dt_1 = \int_s^{t_2} G(t,t_1)dt_1 + \sum_{i=2}^k G(t,t_i)\mathbb{1}_{]s,t[}(t_2,\dots,t_k)(t_2-s)$$

$$\int_0^{t_3} \sum_{i=2}^k G(t,t_i)\mathbb{1}_{]s,t[}(t_2,\dots,t_k)(t_2-s)dt_2 = \int_s^{t_3} G(t,t_2)(t_2-s)dt_2 + \sum_{i=3}^k G(t,t_i)\mathbb{1}_{]s,t[}(t_3,\dots,t_k)\frac{(t_3-s)^2}{2!},$$

leads to

$$\begin{aligned}
A_k(G)(s,t) &= \int_0^t \int_0^{t_k} \dots \int_0^{t_2} \mathbb{1}_{\{s<t_1<\dots<t_k<t\}} \sum_{i=1}^k G(t,t_i)dt_1\dots dt_k \\
&= \int_s^{t_{k+1}} \dots \int_s^{t_2} G(t,t_1)dt_1\dots dt_k + \int_0^{t_{k+1}} \int_0^{t_k} \dots \int_0^{t_3} G(t,t_2)\mathbb{1}_{]s,t[}(t_2,\dots,t_k)(t_2-s)dt_2\dots dt_k \\
&\quad + \int_0^{t_{k+1}} \int_0^{t_k} \dots \int_0^{t_4} \sum_{i=3}^k G(t,t_i)\mathbb{1}_{]s,t[}(t_3,\dots,t_k) \int_s^{t_3} (t_2-s)dt_2\dots dt_k \\
&= \int_s^{t_{k+1}} \dots \int_s^{t_2} G(t,t_1)dt_1\dots dt_k + \int_s^{t_{k+1}} \dots \int_s^{t_3} G(t,t_2)(t_2-s)dt_2\dots dt_k \\
&\quad + \int_0^{t_{k+1}} \int_0^{t_k} \dots \int_0^{t_4} \sum_{i=3}^k G(t,t_i)\mathbb{1}_{]s,t[}(t_3,\dots,t_k)\frac{(t_3-s)^2}{2!}dt_2\dots dt_k \\
&= \sum_{j=1}^k \int_s^{t_{k+1}} \dots \int_s^{t_{j+1}} G(t,t_j)\frac{(t_j-s)^{j-1}}{(j-1)!}dt_j\dots dt_k.
\end{aligned}$$

In other words, the operator A_k is

$$A_k(G)(s,t) = \sum_{j=1}^k B_{j,k}(G)(s,t) \text{ where } B_{j,k}(G)(s,t) = \int_s^{t_{k+1}} \dots \int_s^{t_{j+1}} G(t,t_j)\frac{(t_j-s)^{j-1}}{(j-1)!}dt_j\dots dt_k.$$

B Computation of the second order moment

Let us compute the following expectation which simplifies to $\mathbf{E}(S_t^2)$ when $s = 0$. Let $G_i(t) = g(t, T_i, V_i)$. Splitting the inner sum yields to

$$\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i(t) \sum_{j=N_s+1}^{N_t} Z_j(t) \mid N_s\right) = \mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i^2(t) \mid N_s\right) + 2\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} \sum_{j=j+1}^{N_t} Z_i(t)Z_j(t) \mid N_s\right).$$

The first sum can be computed as in Proposition 3.1. Consequently,

$$\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} Z_i^2(t) \mid N_s\right) = \mathbf{E}(X^2)e^{-\lambda(t-s)} \sum_{k=1}^{\infty} \lambda^k A_k(G_2)(s,t)$$

with the function $G_2 : (t, w) \mapsto \mathbf{E}(g(t, w, V)^2)$. With similar conditioning, the second term is

$$\begin{aligned}
&2\mathbf{E}\left(\sum_{i=N_s+1}^{N_t} \sum_{j=i+1}^{N_t} Z_i(t)Z_j(t) \mid N_s\right) \\
&= 2\mathbf{E}\left(X^2\right) \sum_{\ell=N_s+2}^{\infty} \sum_{i=N_s+1}^{\ell-1} \sum_{j=N_s+1}^{\ell} \mathbf{E}\left(g(t,T_i,V_i)g(t,T_j,V_j) \,|\, N_s, N_t = \ell\right) P\left(N_t = \ell \mid N_s\right) \\
&= 2\mathbf{E}\left(X^2\right) \sum_{k=2}^{\infty} \sum_{i=N_s+1}^{N_s+k-1} \sum_{j=N_s+1}^{N_s+k} \mathbf{E}\left(g(t,T_i,V_i)g(t,T_j,V_j) \,|\, N_s, N_t - N_s = k\right) P\left(N_t - N_s = k \mid N_s\right).
\end{aligned}$$

With $G(t, t_i, t_j) = \mathbf{E}\left(g(t, t_i, V_i)g(t, t_j, V_j)\right) = G(t, t_i)G(t, t_j)$ (by assumption A3), the double sum is

$$\begin{aligned}
&\sum_{i=N_s+1}^{N_s+k-1} \sum_{j=i+1}^{N_s+k} \mathbf{E}\left(g(t,t_i,V_i)g(t,t_j,V_j) \,|\, N_s, N_t - N_s = k\right) \\
= \;& \int_{\mathbb{R}^k} \sum_{i=N_s+1}^{N_s+k-1} \sum_{j=i+1}^{N_s+k} G(t,t_i,t_j) dH_k^s(t_{N_s+1}, \ldots, t_{N_s+k}) = \frac{k!}{(t-s)^k} \int_s^{t_{k+1}} \cdots \int_s^{t_2} \sum_{n=1}^{k-1} \sum_{m=n+1}^{k} G(t,t_n,t_m) dt_1 \ldots dt_k,
\end{aligned}$$

with $t_{k+1} = t$. Using Appendix C, we have

$$\int_s^{t_{k+1}} \cdots \int_s^{t_2} \sum_{n=1}^{k-1} \sum_{m=n+1}^{k} G(t,t_n,t_m) dt_1 \ldots dt_k, = A_k^*(\mathbf{G})(s,t).$$

Combining with $P(N_t = k|N_s) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}$ leads to the desired result.

C Simplification of operator $A_k^*(\mathbf{G})(s,t)$

Still with $t_{k+1} = t$, we use similar reasoning as Appendix A

$$\int_s^{t_2} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} G(t,t_i,t_j) dt_1 = \int_s^{t_2} \sum_{j=2}^{k} G(t,t_1,t_j) dt_1 + \sum_{i=2}^{k-1} \sum_{j=i+1}^{k} G(t,t_i,t_j)(t_2 - s),$$

$$\int_s^{t_3} \sum_{i=2}^{k-1} \sum_{j=i+1}^{k} G(t,t_i,t_j)(t_2 - s) dt_2 = \int_s^{t_3} \sum_{j=3}^{k} G(t,t_2,t_j)(t_2 - s) dt_2 + \sum_{i=3}^{k-1} \sum_{j=i+1}^{k} G(t,t_i,t_j) \frac{(t_3 - s)^2}{2!},$$

leads to

$$\int_s^{t_{k+1}} \dots \int_s^{t_2} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} G(t,t_i,t_j) dt_1 \dots dt_k$$

$$= \int_s^{t_{k+1}} \dots \int_s^{t_2} \sum_{j=2}^{k} G(t,t_1,t_j) dt_1 \dots dt_k + \int_s^{t_{k+1}} \dots \int_s^{t_3} (t_2 - s) \sum_{j=3}^{k} G(t,t_2,t_j) dt_2 \dots dt_k$$

$$+ \int_s^{t_{k+1}} \dots \int_s^{t_4} \frac{(t_3-s)^2}{2!} \sum_{i=3}^{k-1} \sum_{j=i+1}^{k} G(t,t_i,t_j) dt_3 \dots dt_k$$

$$= \sum_{i=1}^{k-1} \int_s^{t_{k+1}} \dots \int_s^{t_{i+1}} \frac{(t_i-s)^{i-1}}{(i-1)!} \sum_{j=i+1}^{k} G(t,t_i,t_j) dt_i \dots dt_k = A_k^*(\mathbf{G})(s,t)$$

The operator A_k is defined as

$$\begin{array}{llcl} A_k^*: & C^1_{[s,t]^{k+1} \mapsto \mathbb{R}^k} & \longmapsto & A_k^*(\mathbf{F}) \in C^0_{\mathbb{R}_+^2} \\ & (t,t_1,..,t_k) \mapsto \begin{pmatrix} F_1(t,t_1,..,t_k) \\ \vdots \\ F_k(t,t_k) \end{pmatrix} & \longmapsto & (s,t) \mapsto \mathbb{1}_{0<s<t} \sum_{i=1}^{k} \int_s^{t_{k+1}} \dots \int_s^{t_{i+1}} \frac{(t_i-s)^{i-1}}{(i-1)!} F_i(t,t_i,..,t_k) dt_i..dt_k \end{array}$$

Two functions are considered. For the conditional expectation

$$\mathbf{G}_{bi}(t,t_1,..,t_k) = \begin{pmatrix} G(t,t_1) \\ \vdots \\ G(t,t_k) \end{pmatrix} = \begin{pmatrix} \mathbf{E}(g(t,t_1,V)) \\ \vdots \\ \mathbf{E}(g(t,t_k,V)) \end{pmatrix}.$$

For the conditional second-order moment

$$\mathbf{G}(t,t_1,..,t_k) = \begin{pmatrix} \sum_{j=2}^{k} G(t,t_1,t_j) \\ \vdots \\ G(t,t_{k-1},t_k) \\ 0 \end{pmatrix} = \begin{pmatrix} \sum_{j=2}^{k} \mathbf{E}(g(t,t_1,V))\mathbf{E}(g(t,t_j,V)) \\ \vdots \\ \mathbf{E}(g(t,t_{k-1},V))\mathbf{E}(g(t,t_k,V)) \\ 0 \end{pmatrix}.$$

D Heuristic computation for $B_{jk} = \int_s^{t_{k+1}} \dots \int_s^{t_{j+1}} f(t_j) dt_j..dt_k$

Let $\delta(s,t) = \frac{t-s}{n}, t_l(s,t) = s + (l-1)\delta, l = 1,\dots,n$. We consider an approximation of the successive integral based on the rectangle rule.

— $j = k$: a sum approximation based on the rectangle rule

$$B_{kk}(f)(s,t) \approx \delta \sum_{i=1}^{n} f(t_i) \times 1,$$

— $j = k-1$: a cumsum approximation

$$B_{k-1,k}(f)(s,t) \approx \delta^2 \sum_{l=1}^{n} \sum_{i=1}^{l} f(t_i) = \delta^2 \sum_{i=1}^{n} f(t_i) \sum_{l=i}^{n} 1 = \delta^2 \sum_{i=1}^{n} f(t_i) c^1_{i,n},$$

— $j = k-2$: a double cumsum approximation

$$B_{k-2,k}(f)(s,t) \approx \delta^3 \sum_{l=1}^{n} \sum_{i=1}^{l} f(t_i) c^1_{i,l} = \delta^3 \sum_{i=1}^{n} f(t_i) \sum_{l=i}^{n} c^1_{i,l} = \delta^3 \sum_{i=1}^{n} f(t_i) c^2_{i,n},$$

— general j: multiple cumulative sum approximation

$$B_{jk}(f)(s,t) \approx \delta^{k-i+1} \sum_{i=1}^{n} f(t_i) c_{i,n}^{k-j} \text{ where } c_{i,n}^{j} = \sum_{l=i}^{n} c_{i,l}^{j-1}, c_{i,n}^{0} = 1.$$

Computing the first terms, we notice that

$$\begin{aligned} c_{i,n}^0 &= 1 = \binom{n-l}{0}, \quad c_{i,n}^1 = \sum_{l=i}^n 1 = n-i+1 = \binom{n-l+1}{1}, \\ c_{i,n}^2 &= \sum_{l=i}^n (n-l+1) = \sum_{l=1}^{n-l+1} l = \frac{(n-i+1)(n-i+2)}{2} = \binom{n-i+2}{2}, \\ c_{i,n}^3 &= \sum_{l=i}^n \frac{(n-l+1)(n-l+2)}{2} = \sum_{l=1}^{n-l+1} l(l+1)/2 = \frac{(n-i+1)(n-i+2)(n-i+3)}{6} = \binom{n-i+3}{3}. \end{aligned}$$

That is $c_{i,n}^1$ is the sum of 1, $c_{i,n}^2$ is the sum of integers, $c_{i,n}^3$ is the sum of square integers. We now use the well-known parallel summation identity $\sum_{l=0}^n \binom{r+l}{l} = \binom{r+n+1}{n}$, e.g. from Graham et al. (1994). We have $c_{i,n}^0 = \binom{n-l+0}{0}$ Assume that $c_{i,n}^{k-1} = \binom{n-i+k-1}{k-1}$. Summing over l gives

$$c_{i,n}^k = \sum_{l=i}^n \binom{n-l+k-1}{k-1} = \sum_{l=i}^n \binom{n-l+k-1}{n-l} = \sum_{j=0}^{n-i} \binom{k-1}{j} = \binom{n-i+k-1+1}{n-l} = \binom{n-i+k}{k}.$$

This ends the recurrence. So $B_{jk}(f)(s,t) \approx \delta^{k-j+1} \sum_{i=1}^n f(t_i) \binom{n-i+k-j}{k-j}$. Therefore, an heuristic computation for $A_k = \sum_{j=1}^k \frac{1}{(j-1)!} \int_s^{t_{k+1}} \cdots \int_s^{t_{j+1}} f(t_j) dt_j..dt_k$ can be obtained. Using previous approximation, we get

$$A_k(f)(s,t) \approx \sum_{j=1}^k \frac{1}{(j-1)!} \delta^{k-j+1} \sum_{i=1}^n f(t_i) \binom{n-i+k-j}{k-j} = \sum_{i=1}^n f(t_i) \sum_{j=1}^k \frac{\delta^{k-j+1}}{(j-1)!} \binom{n-i+k-j}{k-j}$$

yielding to

$$e^{-\lambda(t-s)} \sum_{k\geq 1} \lambda^k A_k(f)(s,t) \approx \sum_{k=1}^{k_{max}} e^{-\lambda(t-s)} \lambda^k \sum_{i=1}^n f(t_i) \sum_{j=1}^k \frac{\delta^{k-j+1}}{(j-1)!} \binom{n-i+k-j}{k-j}.$$

The Trapezoidal rule is obtained by replacing $f(t_i)$ by $\frac{f(t_i)+f(t_{i+1})}{2}$.

E Claim triangles

	Development year							
Accident year	0	1	2	3	4	5	6	7
1990	37482	139760	242037	344315	446593	534979	582384	602261
1991	67954	215479	363005	510531	643364	720880	774350	
1992	114975	262831	410686	511030	566393	580910		
1993	90355	202967	302796	373944	419019			
1994	216343	442578	519775	568651				
1995	178740	242198	285542					
1996	188638	331037						
1997	93015							

Table 5 – Triangle for damage cover

	Development year							
Accident year	0	1	2	3	4	5	6	7
1990	1141816	1836488	2531161	3225834	3796269	4157660	4365142	4480327
1991	1210060	1904389	2598718	3090223	3330113	3466759	3490643	
1992	874032	1368811	1841729	2227655	2474422	2527571		
1993	379682	577927	708582	797925	844412			
1994	713520	922794	1047918	1143308				
1995	409297	626053	805173					
1996	819551	1332359						
1997	915450							

Table 6 – Triangle for bodily injury cover