PREDICTION OF OUTSTANDING LIABILITIES
II Model variations and extensions

BY

RAGNAR NORBERG

University of Copenhagen and London School of Economics

ABSTRACT

This is a follow-up of a previous paper by the author, where claims reserving in non-life insurance is treated in the framework of a marked Poisson claims process A key result on decomposition of the process is generalized, and a number of related results are added. Their usefulness is demonstrated by examples and, in particular, the connection to the analogous discrete time model is clarified. The problem of predicting the outstanding part of reported but not settled claims is revisited and, by way of example, solved in a model where the partial payments are governed by a Dirichlet process The process of reported claims is examined, and its dual relationship to the process of occurred claims is pointed out.

KEYWORDS

Claims reserving, marked Poisson process; decomposition, thinning; amalgamation; Dirichlet process; discrete vs. continuous time

1 INTRODUCTION

A. Review of previous results

In a previous paper by the author (Norberg, 1993), henceforth referred to as (I), a continuous-time approach is taken to the problem of predicting the total liability of a non-life insurance company The model framework is a non-homogeneous marked Poisson process with position-dependent marks, the Poisson *times* representing the occurrences of claims and the corresponding *marks* representing the developments of the individual claims. Since the present paper is self-contained, we shall review the results in (I) only briefly.

The total claim amount in respect of a finite amount of risk exposure follows a compound Poisson distribution Fixing a time of consideration, the set of all claims decomposes into *s* (settled), *rns* (reported-not-settled), *inr* (incurred-not-reported), and *cni* (covered-not-incurred, corresponding to the unearned premium reserve). These four components can be viewed as arising from independent marked Poisson processes whose intensities and mark distributions have an easy interpretation. By use of this decomposition result predictors are constructed for all components of the total outstanding liability, in respect of *rns* claims, *inr* claims, and *cni* claims. In (I) also a doubly stochastic extension of the model was treated with continuous time credibility methods, but that topic shall not be pursued here.

B. Objective and plan of the present study

The present study is a follow-up of (I). In Section 2 the basic model is investigated further. Previous results are generalized and some further distribution theoretical results are added. In particular, the decomposition result is extended to quite general countable decompositions. A partial converse result on amalgamation of independent marked Poisson processes is established, generalizing a well-known result on convolution of independent compound Poisson risk processes

Section 3 provides examples of applications. In particular, the connection to the analogous discrete time model is established upon decomposing by year of occurrence and year of notification. This connection opens a way to well-reasoned parametrization of the discrete time model.

In Section 4 the problem of predicting the outstanding part of reported but not settled claims is revisited Unbiased prediction is discussed in the general set-up, and carried out in a model where the allotment of partial payments from notification until ultimate settlement is governed by a time-scaled Dirichlet process

Section 5 inquires into the process of claims reports, which is the current basis for prediction of not-reported ($nr = inr + cni$) claims. There is a duality in the situation: taking the moments of notification of claims as times and the remaining characteristics as marks, leads also to a marked Poisson process.

The style of the paper is informal in the sense that, as a rule, obvious conditions and assumptions are not spelled out. Thus, at the base of everything is some probability space $(\Omega, \mathcal{F}, P)$, which is not brought to the surface, and it is tacitly understood that sets and mappings are measurable, that expected values and other quantities displayed exist, and so on. We will dispense with every form of mathematical ceremony that would add words without adding rigour.

2 FURTHER INVESTIGATIONS IN THE BASIC MODEL

A. Notation and model assumptions

We first recapitulate and streamline a bit the notation and some basic definitions in (I) Formally, a *claim* is a pair $C = (T, Z)$, where T is the time of occurrence of the claim and Z is the so-called mark describing its development from the time of occurrence until the time of final settlement The mark is taken to be of the form $Z = (U, V, Y, \{Y'(v'); 0 \leq v' \leq V\})$ where U is the waiting time from occurrence until notification, V is the waiting time from notification until final settlement, Y is the eventual total claim amount, and $Y'(v')$ is the amount paid within v' time units after the notification, hence $Y = Y'(V)$. Henceforth we write $Y' = \{Y'(v'), 0 \leq v' \leq V\}$ in short.

We shall primarily have this situation in mind, but note that other descriptions of the claim history are possible and that the mark might be a complex entity comprising any piece of information appearing in the claim record. The space of all possible claim outcomes is $\mathcal{C} = \mathcal{T} \times \mathcal{Z}$, where $\mathcal{T} = [0, \infty)$ is the time axis (business commences at time 0) and $\mathcal{Z}$ is the set of all possible developments

The *claims process* of an insurance business is a random collection of points in the claim space, $\{(T_i, Z_i)\}_{i=1,\ldots,N}$, $N \leq \infty$, the index i indicating chronological order so that $0 < T_1 < T_2 < \ldots$ It is assumed that the times T_i are generated by an inhomogeneous Poisson process with intensity $w(t)$ at time $t > 0$ and that the marks are of the form $Z_i = Z_{T_i}$, where $\{Z_t\}_{t>0}$ is a family of random elements in $\mathcal{Z}$ that are mutually independent and independent of the Poisson process, and $Z_t \sim P_{Z|t}$. We then speak about a *marked Poisson process with intensity* $\{w(t)\}_{t>0}$ *and position-dependent marking by* $\{P_{Z|t}\}_{t>0}$, and write

$$\{(T_i, Z_i)\}_{i=1,\ldots,N} \sim Po(w(t), P_{Z|t};\ t > 0)$$

Introduce the *total risk exposure*

$$W = \int_0^\infty w(t)dt.$$

We assume throughout that $W < \infty$, having in mind the liabilities of an insurance company in respect of (the finite) business written up to the present. The case with $W = \infty$ and $\int_0^s w(t)dt < \infty$ for all finite s, is treated by just chaining together independent models for disjoint time intervals with finite exposure.

In the following Po(W) denotes the Poisson distribution with parameter W, which has elementary probability function

$$po(n; W) = \frac{W^n}{n!}e^{-W}, \tag{2 1}$$

$n = 0, 1, \ldots$, and Po(W, P) denotes the corresponding compound Poisson distribution of a variate $X = \sum_{i=1}^{N} Y_i$, where $N \sim Po(W)$ and independent of the Y_i, which are independent selections from the distribution P We adopt the standard notation Pf^{-1} for the probability distribution induced by a mapping f, that is, $\mathrm{P}f^{-1}\{B\} = \mathrm{P}\{\omega; f(\omega) \in B\} = \mathrm{P}\{f^{-1}(B)\}$.

B. Alternative construction of the process

We set out by reminding of a basic result in (I).

Theorem 1 (Norberg, 1993). *The marked Poisson process $\{(T_i, Z_i)\}_{i=1,\ldots,N}$ can be constructed in two steps, first generating*

$$N \sim Po(W)$$

and, second, selecting a random sample of N pairs from the distribution P_{TZ} on C given by

$$P_{TZ}(dt, dz) = \frac{w(t)dt}{W} P_{Z|t}(dz),$$

$(t, z) \in C$, and ordering them by the chronology of the occurrences

For ease of reference we restate the proof, which just amounts to inspecting the joint probability distribution of the claims, recast as

$$P\{N = n, (T_i, Z_i) \in (dt_i, dz_i),\ i = 1, \ldots, n\} \tag{2.2}$$

$$= e^{-\int_0^\infty w(t)dt} \prod_{i=1}^{n} w(t_i)dt_i P_{Z|t_i}(dz_i) \tag{2.3}$$

$$= \frac{W^n}{n!} e^{-W} n! \prod_{i=1}^{n} P_{TZ}(dt_i, dz_i), \tag{2.4}$$

and recalling (2.1).

We easily obtain a useful generalization of a result in (I).

Corollary 1 to Theorem 1. *Let f be a real-valued function defined on C and define the random variable*

$$X_f = \sum_{i=1}^{N} f(T_i, Z_i). \tag{2.5}$$

Then $X_f \sim Po(W, P_{TZ}f^{-1})$, and the first three central moments of X_f are

$$m_{X_f}^{(k)} = \int w(t) \int f(t, z)^k P_{Z|t}(dz)dt, \quad k = 1, 2, 3. \tag{2.6}$$

Proof: The distribution result follows from the fact that the sum on the right of (2.5) does not depend on the chronological order of the claims (f is independent of i) and therefore is distributed as the sum of N replicates of $f(T, Z)$ that are mutually independent and independent of N. Then (2.6) is just a standard result about the compound Poisson law. □

The probability distribution of X_f in (2.5) may be computed by standard methods for numerical evaluation of total claims distributions.

Note the linearity property

$$X_{f_1} + \dots + X_{f_k} = X_{f_1 + \dots + f_k} \tag{2.7}$$

Corollary 2 to Theorem 1 *Let f' and f'' be real-valued functions on C and $X_{f'}$ and $X_{f''}$ the corresponding compound Poisson variates defined in accordance with (2.5). Then*

$$\operatorname{Cov}(X_{f'}, X_{f''}) = \int w(t) \int f'(t,z) f''(t,z) P_{Z|t}(dz)\,dt \tag{2.8}$$

Proof Write

$$\operatorname{Cov}(X_{f'}, X_{f''}) = \frac{1}{4}\left(\operatorname{Var}(X_{f'} + X_{f''}) - \operatorname{Var}(X_{f'} - X_{f''})\right),$$

and use the linearity property (2.7) together with (2.6) and the identity

$$\frac{1}{4}\left((f' + f'')^2 - (f' - f'')^2\right) = f'f''.$$

□

C. A general decomposition result and some complements

Let C^g, $g = 1, 2, \dots, h$ $(\leq \infty)$, be a partition of the claim space, that is, $\bigcup_{g=1}^{h} C^g = C$ and $C^{g'} \cap C^{g''} = \emptyset$ if $g' \neq g''$. Introduce

$$Z_t^g = \{z \in Z, (t,z) \in C^g\},$$

the set of developments that make a claim occurred at time t a g-claim (belonging to C^g), and

$$T^g = \left\{t \in T; P_{Z|t}\{Z_t^g\} > 0\right\},$$

the time period (or more general era) where such claims can occur. The process of g-claims is denoted $\{(T_i^g, Z_i^g)\}_{1 \leq i \leq N^g}$, $g = 1, \dots, h$, where the times T_i^g are listed in chronological order. The following result generalizes Theorem 2 in (I), which considered only finite partitions.

Theorem 2. *The component g-claims processes are independent, and*

$$\{(T_i^g, Z_i^g)\}_{i=1,\ldots,N^g} \sim Po\left(w^g(t), P^g_{Z|t}; t > 0\right),$$

with

$$w^g(t) = w(t)P_{Z|t}\{Z_t^g\}, \tag{2.9}$$

$$P^g_{Z|t}(dz) = \frac{P_{Z|t}(dz)}{P_{Z|t}\{Z_t^g\}} 1_{Z_t^g}(z) \tag{2.10}$$

Proof For finite h the proof of Theorem 2 in (I) carries over without modification. We sketch it here since it will be needed in the sequel Look back at the proof of Theorem 1. First, state the event appearing in (2.2) in terms of the component processes to rewrite the probability as

$$P\left\{\bigcap_g \{N^g = n^g, (T_i^g, Z_i^g) \in (dt_i^g, dz_i^g), i = 1, \ldots, n^g\}\right\}.$$

Second, use the fact that $\sum_g P_{Z|t}\{Z_t^g\} = 1$ for each t to rewrite (2.3) in terms of (2.9) and (2.10) as

$$\prod_{g=1}^{h}\left(e^{-\int_0^\infty w^g(t)dt}\prod_{i=1}^{n} w^g(t_i^g)dt_i^g P^g_{Z|t_i^g}(dz_i^g)\right).$$

Finally, recast each factor in this product in the same way as in (2.4) to arrive at

$$\prod_g\left(\frac{(W^g)^{n^g}}{n^g!}e^{-W^g}n^g!\prod_{i=1}^{n^g} P^g_{TZ}(dt_i^g, dz_i^g)\right), \tag{2.11}$$

where

$$W^g = \int_0^\infty w^g(t)dt, \tag{2.12}$$

with $w^g(t)$ defined by (2.9), and

$$P^g_{TZ}(dt, dz) = \frac{w^g(t)dt}{W^g} P^g_{Z|t}(dz),\ z \in Z_t^g \tag{2.13}$$

The result now follows from Theorem 1 and the product form of (2.11).

For $h = \infty$, consider any finite set of categories $g_1, \ldots, g_q$, and lump all the remaining categories into one category g_0, say. The result for the finite case applies to these $q + 1$ categories, and it follows that the q component processes are independent marked Poisson processes as specified in (2.9) – (2.10). Since the probability measure is determined by the finite-dimensional distributions, the result follows. □

The result says that g-claims occur with an intensity which is the claim intensity times the probability that the claim belongs to the category g, and that the development of the claim is governed by the conditional distribution of the mark, given that it is a g-claim. Accordingly, the quantity W^g in (2.12) may be termed the total exposure in respect of claims of category g or just the *g-exposure*. Observe that P^g_{TZ} in (2.13) is the conditional distribution of (T, Z), given that it is a g-claim

$$P^g_{TZ}(dt, dz) = \frac{P_{TZ}(dt, dz)}{W^g/W} 1_{C^g}(t, z).$$

Theorem 2 may be seen as a general result on so-called thinning of Poisson processes, which in its simplest form amounts to throwing out a certain proportion of the occurrences by some coin-tossing mechanism independent of the process itself. See e.g. Karr (1991).

The following result is a direct consequence of Theorem 2 (and previous results):

Corollary 1 to Theorem 2. *Let C^g, $g = 1, 2, \ldots$, be a partition of C and, for each $g = 1, 2, \ldots$ let f^g be a real-valued function on C^g. Then the corresponding compound Poisson variates*

$$X^g = \sum_{i=1}^{N^g} f^g(T^g_i, Z^g_i),$$

$g = 1, 2, \ldots$, *are mutually independent.*

The following reformulation presents an interest of its own.

Corollary 2 to Theorem 2. *Let f^g, $g = 1, 2, \ldots$ be a sequence of real-valued functions on C satisfying $f^{g'}(t, z) f^{g''}(t, z) = 0$ for $g' \neq g''$. Then the corresponding compound Poisson variates X_{f^g}, $g = 1, 2, \ldots$, are mutually independent.*

Remark. By Corollary 2 to Theorem 1, we knew beforehand that the X_{f^g} are uncorrelated.

Proof. Define $C^g = \{(t, z); f^g(t, z) \neq 0\}$, $g = 1, 2, \ldots$, and note that these sets together with $C^0 = \{(t, z); f^g(t, z) = 0, \; g = 1, 2, \ldots\}$ form a partition of C. The result follows upon rewriting each X_{f^g} as

$$X_{f^g} = \sum_{i=1}^{N^g} f^g(T^g_i, Z^g_i)$$

and invoking Theorem 2 and Corollary 1 to Theorem 1. □

Before proceeding, we present a small auxiliary lemma whose proof is obvious.

Lemma. *Suppose* $\{(T_i, Z_i^*)\}_{i=1,\dots,N} \sim Po(w(t), P^*_{Z^*|t}; t > 0)$, *a marked Poisson process on* $\mathcal{T} \times \mathcal{Z}^*$. *Let* ζ *be a function defined on* $\mathcal{Z}^*$ *and with values in* $\mathcal{Z}$, *and denote the transformed marks by* $Z_i = \zeta(Z_i^*)$. *Then*

$$\{(T_i, Z_i)\}_{i=1,\dots,N} \sim Po(w(t), P_{Z|t}; t > 0),$$

a marked Poisson process on $\mathcal{T} \times \mathcal{Z}$, *with*

$$P_{Z|t} = P^*_{Z^*|t}\zeta^{-1} \tag{2.14}$$

A standard result, known as the amalgamation theorem for compound Poisson claims processes, goes as follows: Let X^g, $g = 1, \dots, h\ (< \infty)$, be independent compound Poisson processes, that is, each X^g is of the form $X^g(t) = \sum_{j=1}^{N^g(t)} Y_j^g$, where N^g is a homogeneous Poisson process with claim intensity w^g, and the individual claims amounts Y_j^g are independent selections from a claim size distribution P^g and, moreover, independent of N^g. Then the process $X = \sum_{g=1}^h X^g$ is a compound Poisson process with claim intensity $w = \sum_{g=1}^h w^g$ and claim size distribution $P = w^{-1} \sum_{g=1}^h w^g P^g$. This generalizes to the following, which appears as a partial converse of the decomposition Theorem 2, but really is implied by it:

Theorem 3. *Suppose* $\{(T_i^g, Z_i^g)\}_{i=1,\dots,N^g} \sim Po(w^g(t), P^g_{Z|t}; t > 0)$, $g = 1, \dots, h$, *are a finite number of mutually independent marked Poisson processes on* $\mathcal{T} \times \mathcal{Z}$. *Then the amalgamated process* $\{(T_i, Z_i)\}_{i=1,\dots,N}$, *obtained by assembling the claims of the individual processes, is also a marked Poisson process on* $\mathcal{T} \times \mathcal{Z}$, *and* $\{(T_i, Z_i)\}_{i=1,\dots,N} \sim Po(w(t), P_{Z|t}, t > 0)$, *with*

$$w(t) = \sum_{g=1}^{h} w^g(t), \tag{2.15}$$

$$P_{Z|t}(dz) = \frac{1}{w(t)} \sum_{g=1}^{h} w^g(t) P^g_{Z|t}(dz). \tag{2.16}$$

Remark: The claimed result is precisely what one would expect. The property of "independent partitions" carries over from the individual processes to the amalgamated one and suggests the Poisson property of the occurrences of the latter. Furthermore, (2.15) says that the total probability of a claim occurrence in a small time interval is the sum of the corresponding probabilities for the individual processes, and (2.16) states that a claim occurred at time t is from the g-th individual process with probability $w^g(t)/w(t)$, in which case the mark is generated by the mark distribution of that process. □

Proof: Anticipating the result, start from a marked Poisson process

$$\{(T_i, Z_i^*)\}_{i=1,\ldots,N} \sim Po(w(t), P^*_{Z^*|t}; t > 0) \tag{2.17}$$

on $C^* = \mathcal{T} \times \mathcal{Z}^*$, where $\mathcal{Z}^* = \{1, \ldots, h\} \times \mathcal{Z}$ and

$$P^*_{Z^*|t}(g, dz) = \frac{w^g(t)}{w(t)} P^g_{Z|t}(dz). \tag{2.18}$$

The generic mark of this process is $Z^* = (G, Z)$, the original mark augmented with an index for "type of claim". It is seen from (2.18) that a claim occurred at time t is of type $G = g$ with probability $w^g(t)/w(t)$ and, given this, the Z-part of the mark is generated from $P^g_{Z|t}$.

Now, on the one hand, applying Theorem 2 to the decomposition of C^* by claim type, $C^{*g} = \{(t', g', z'), \; g' = g\}$, $g = 1, \ldots, h$, we readily find that the component processes have the distribution properties of the individual processes as specified in the assumptions of the present theorem, and so we can as well let the latter be constructed as the component processes in the present model (2.17) – (2.18).

On the other hand, it is realized that in the present model the amalgamated process is obtained from $\{(T_i, Z_i^*)\}_{i=1,\ldots,N}$ upon leaving the type G unobserved or, in the terms of the lemma above, considering the process with marks transformed by $\zeta(g, z) = z$. Under this simple mapping the probability distribution in (2.14) is just the marginal distribution of Z in the distribution of (G, Z) given by (2.18), which is precisely the one defined in (2.16). Thus, the lemma completes the proof. □

We round off this paragraph with an alternative proof of Corollary 2 to Theorem 1. It makes use of the decomposition theorem and, moreover, serves to demonstrate a useful technique:

Second Proof of Corollary 2 to Theorem 1. Suppose the results holds for indicator functions f' and f''. Then, by the bilinearity of the covariance operator, it also holds for linear combinations of indicator functions. Since every (measurable) non-negative function is the monotone limit of linear combinations of simple functions, the result extends to non-negative functions f' and f'' by the monotone convergence theorem. Finally, since any function is the difference of its positive and negative parts, the result extends to general functions.

Thus, it suffices to prove the result for $f' = 1_{C'}$ and $f'' = 1_{C''}$, where C' and C'' are subsets of C. Since f' and f'' are binary, the functions $f'f''$, $f'(1 - f'')$, and $f''(1 - f')$ satisfy the "orthogonality" condition in Corollary 2 to Theorem 2, hence the corresponding compound Poisson variates are independent. By the linearity property (2.7) $X_{f'} = X_{f'f''} + X_{f'(1-f'')}$ and $X_{f''} = X_{f'f''} + X_{f''(1-f')}$. These things together imply that

$$\mathrm{Cov}(X_{f'}, X_{f''}) = \mathrm{Var}(X_{f'f''}) = \int w(t) \int (f'(t,z) f''(t,z))^2 P_{Z|t}(dz)dt,$$

where we have made use of (2.6). Now, since $f'f''$ is binary, it equals its square, and we arrive at (2.8). □

Results like Corollary 2 to Theorem 1 are valid for more general marked pointed processes, see e g Karr (1991) The present proofs are worth reporting since they are simple thanks to the fact that "everything is independent of everything else" in the Poisson scenario.

3 APPLICATIONS

A. Notation

In the following we will frequently use notation pertaining to the situation where (U, V, Y) belonging to a claim occurred in time t has a joint density $p_{UVY|t}(u, v, y)$ with respect to Lebesgue measure In a self-explaining way we denote marginal densities by e.g. $p_{Y|t}(y)$ and conditional densities by e.g. $p_{U|t,y}(u)$ If $P_{Z|t}$ is independent of t, we speak about *time-independent marks* and drop t from the subscript

B. Decomposition by claim amount; franchise and reinsurance

Fix $0 = m_0 < m_1 < ...$ and, for each $g = 1, 2, ...,$ let

$$C^g = \{(t, z); m_{g-1} < y \leq m_g\}$$

be the set of claims with amount in the interval $(m_{g-1}, m_g]$. The corresponding component processes are independent, with intensities and mark distributions given by

$$w^g(t) = w(t) \int_{m_{g-1}}^{m_g} p_{Y|t}(y)dy,$$

$$P^g_{Z|t}(dz) = \frac{P_{Z|t}(dz)}{\int_{m_{g-1}}^{m_g} p_{Y|t}(y)dy} 1_{(m_{g-1}, m_g]}(y).$$

In particular, for a fixed m, let the two sets $C^s = \{(t, z), y \leq m\}$ and $C^\ell = \{(t, z), y > m\}$ decompose the business into "small" and "large" claims We may interpret m as the deductible part by minimum franchise or first risk in the context of direct insurance or as the retention level in the context of excess of loss reinsurance.

Pursuing the latter interpretation, consider a reinsurance treaty under which the cedent and the reinsurer cover $f'(t, z)$ and $f''(t, z)$, respectively, of a claim occurred at time t and with mark z. The covariance between their total losses is given by (2.8), and their means and variances are given in Corollary 1 to Theorem 1

For instance, for quota share reinsurance we have $f'(t,z) = ky$ and $f''(t,z) = (1-k)y$ so that, with time-independent marks, the covariance is simply

$$Wk(1-k)\mathrm{E}[Y^2].$$

For excess of loss reinsurance we have $f'(t,z) = \min(y,m)$ and $f''(t,z) = \max(y-m,0)$ and, since the product of these functions is $1_{(m,\infty)}m(y-m)$, the covariance is

$$Wm\mathrm{E}[1_{(m,\infty)}(Y-m)] = Wm\int_m^\infty (1-P_Y(y))dy.$$

The results carry over to business in respect of limited periods of exposure by just letting the integral with respect to t range over a suitable period of time. This aspect comes up next.

C. Decomposition by year of occurrence

As accounts are typically kept on an annual basis, we shall now decompose by year, and take calendar year j to mean the time interval $(j-1,j]$. The "cohort" of claims occurred in year j is

$$C^j = \{(t,z); j-1 < t \leq j\}.$$

The total claim amount in respect of such claims is a compound Poisson variate with frequency parameter

$$W^j = \int_{j-1}^{j} w(t)dt$$

and claim size density

$$P^j_Y(y) = \frac{1}{W^j}\int_{j-1}^{j} w(t)p_{Y|t}(y)dt.$$

In particular, in the *homogeneous case* with time-independent marks and constant Poisson intensity, $w(t) = w$, we have

$$W^j = w, \; p^j_Y = p_Y. \tag{3.1}$$

Decomposition by cohort pertains to reinsurance on the basis of underwriting year Under a contract specifying that the reinsurer covers $f(y)$ of any claim of size y occurring in year j, the reinsured part of the total claim amount is distributed in accordance with $(W^j, P^j_Y f^{-1})$

D. Decomposition by year of notification

Next, consider claims reported in year j,

$$C^j = \{(t,z),\ 0 < t \leq j,\ j-1-t < u \leq j-t\}$$

If claims are settled immediately upon notification, this decomposition pertains to reinsurance on the basis of accounting year.

The total claim amount in respect of these claims is a compound Poisson variate with frequency parameter

$$W^j = \int_0^j w(t) \int_{j-1-t}^{j-t} p_{U|t}(u)du\, dt$$

and claim size density

$$p_Y^j(y) = \frac{1}{W^j} \int_0^j w(t) \int_{j-1-t}^{j-t} p_{UY|t}(u,y)du\, dt$$

Interchanging the order of the integrations in the expression for $W^{u,j}$ above, we find

$$W^j = \int_0^j p_{U|t}(u) \int_{j-1-u}^{j-u} w(t)dt\, du.$$

Similarly we recast p_Y^j as

$$p_Y^j(y) = \frac{1}{W^j} \int_0^j p_{UY|t}(u,y) \int_{j-1-u}^{j-u} w(t)dt\, du.$$

Consider again the homogeneous case with time-independent marks and constant Poisson intensity, w. Letting j increase, the expressions above tend to

$$W^j = w,\ p_Y^j(y) = p_Y(y).$$

Comparing with (3.1) we conclude, loosely speaking, that for a stationary insurance business the liability in respect of occurrence year is the same as the liability in respect of accounting year. This conclusion carries over to the reinsurance businesses that motivated the two types of decomposition.

E. Decomposition by year of occurrence and year of notification

The set of claims occurred in year j and reported in year $j + d$ is

$$C^{jd} = \{(t,z);\ j-1 < t \leq j,\ j+d-1-t < u \leq j+d-t\}$$

The total claim amount in respect of such claims is a compound Poisson variate with frequency parameter

$$W^{jd} = \int_{j-1}^{j} w(t) \int_{j+d-1-t}^{j+d-t} p_{U|t}(u)du\, dt$$

and claim size density

$$p_Y^{jd}(y) = \frac{1}{W^{jd}} \int_{j-1}^{j} w(t) \int_{j+d-1-t}^{j+d-t} p_{UY|t}(u,y)du\, dt.$$

Note that, even if $P_{Z|t}$ should be independent of t, p_Y^{jd} may vary with j for fixed d due to possible variations in the shape of the intensity $w(t)$ from one year to another. This effect has been studied by Hesselager (1995).

F. Connection to the discrete time model

In Norberg (1986) the author launched a model which is a discrete time rudiment of the present one. It was assumed that claims are settled immediately upon notification (or rather in the same year). An issue in that set-up was how to specify the distribution of the size of a claim that occurs in year j and is reported d years later Leaving the possible dependence on j aside, we need to specify claim size distributions P_Y^d for delay times $d = 1, 2,$ It appears that the discrete time set-up allows for no other approach than just specifying these distribution directly, possibly starting from some standard parametric claim size distribution and letting the parameters be some parametric functions of d

The continuous time model creates another and, from an aesthetic viewpoint, more pleasing possibility. A parametric specification of the continuous time model, which may be supported by physical reasoning, will automatically induce a parametrization of the discrete time model We shall illustrate this by a simple example, assuming now that $w(t)$ is constant.

Let $Ga(\alpha, \beta)$ denote the gamma distribution with shape parameter α and inverse scale parameter β, both positive, which has density

$$ga(y;\ \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} 1_{(0,\infty)}(y). \tag{3 2}$$

Assume that the joint distribution of $(U,\ Y)$ is such that

$$p_Y(y) = ga(y;\ \alpha, \beta)$$

and

$$p_{U|y}(u) = ga(u,\ 1, \mu y) = \mu y e^{-\mu y u} 1_{(0,\infty)}(u)$$

(an exponential distribution), implying that large claims tend to be reported more promptly than small claims We easily find that U has a scaled and shifted Pareto distribution with density

$$p_U(u) = \frac{\mu\alpha\beta^\alpha}{(\mu u + \beta)^{\alpha+1}} 1_{(0,\infty)}(u).$$

(It is seen that $\mathrm{E}[U^k] < \infty$ for $-1 < k < \alpha$.)

Some easy calculations lead to the following expression for the distribution of Y for a jd-claim as defined in the previous paragraph (by assumption it does not depend on j):

$$p_Y^d(y) = \frac{2\phi_d\gamma_d(y) - \phi_{d+1}\gamma_{d+1}(y) - \phi_{d-1}\gamma_{d-1}(y)}{2\phi_d - \phi_{d+1} - \phi_{d-1}},$$

where

$$\phi_d = (\mu d + \beta)^{1-\alpha},$$
$$\gamma_d(y) = ga(y, \ \alpha - 1, \ \mu d + \beta).$$

Thus, we end up with a mixture of gamma distributions, which is mathematically tractable.

G. Inflation and discounting

As a final example of the applicability of the general theory, suppose the insurer currently invests (or borrows) at a fixed rate of interest δ. Then, taking our stand at a given time τ, it may be relevant to consider the value of the claims payments in $[0, \tau]$ accumulated with compound interest,

$$X^a = \sum_i f(T_i, Z_i),$$

where

$$f(T, Z) = 1_{[0,\tau]}(T + U) \int_0^{\tau - T - U} \exp\{(\tau - T - U - v')\delta\} dY'(v').$$

Again we can conclude that X^a is a compound Poisson variate, which in principle is simple The claim size distribution may in this case be a bit complicated, though, but it could be simulated in any case.

Inflation at rate δ can be accommodated in the model e.g. by letting $P_{Y'|tuvy}$ be the distribution of $Y'(v') = \int_{[0,v']} \exp\{(t+u+v'')\delta\} dY^\circ(v''), 0 \le v' \le v$, where Y° is a process with some distribution $P_{Y^\circ|uv}$, independent of t.

4 PREDICTING THE OUTSTANDING PART OF REPORTED CLAIMS

A. Modelling the claim developments; general considerations

We now turn to the issue of modelling the mark distribution $P_{Z|t}$. To avoid blurring the picture, let us assume independence of t and denote by P_Z the distribution of the generic mark $Z = (Y, V, Y, Y')$. (Various forms of time dependence due to trends in risk conditions and inflation can be obtained by trivial reparametrization and scaling.)

Presumably, it will be felt that (U, V, Y) are the primary characteristics of the claim (they tell us "what kind of claim it is") and that the partial payments Y' are secondary, more or less explained by (U, V, Y) Then it is natural to construct P_Z in two steps, specifying first the marginal distribution of (U, V, Y) and, second, the conditional distribution of the process Y', given (U, V, Y).

One convenient choice of P_{UVY} is the trivariate lognormal distribution. It has 9 parameters (3 means and 6 variances or covariances) and may be viewed as a fit model based on moments up to second order. If experience and physical reasoning would dictate a more sophisticated model, one would typically regard Y as the basic entity and specify first the marginal distributions P_Y and, second, the conditional distribution $P_{UV|Y}$. The candidate models are countless and, having no particular application in mind, it does not make any sense to list some dozens of them here.

We shall focus on modelling the conditional distribution of Y', given (U, V, Y). One possible way of building this model is to put

$$Y'(v') = Q(v'/V)\,Y, \tag{4.1}$$

where $\{Q(s),\ 0 \leq s \leq 1\}$ is some stochastic distribution function on $[0, 1]$, stochastically independent of (U, V, Y). This kind of model is suitable if the shape of the partial payments process is independent of other claim characteristics, roughly speaking. Again there are many candidates; any stochastic process X that is non-decreasing, right-continuous, and such that $0 = X(-\infty) < X(\infty) < \infty$, produces a stochastic distribution function Q on the real line $\mathcal{R}$ defined by

$$Q(s) = X(s)/X(\infty) \tag{4.2}$$

B. The Dirichlet process

A convenient choice of X in (4 2) is the gamma process defined as follows. Let α be a scaled distribution function on $\mathcal{R}$ (i.e. $\alpha(s)/\alpha(\infty)$ is a distribution function), and let X have independent increments such that

$$X(s) - X(r) \sim \mathrm{Ga}(\alpha(s) - \alpha(r), \beta)$$

for $r \leq s$ (confer (3.2)). That X is well-defined this way follows from Kolmogorov's consistency condition and the convolution property of the gamma distribution (to be described below). The inverse scale parameter β is immaterial in the construction of Q by (4 2), of course, and could be set to 1.

Now, let $-\infty = s_0 < s_1 < \ldots < s_k = \infty$ be a finite partition of $\mathcal{R}$, and abbreviate $\alpha_i = \alpha(s_i) - \alpha(s_{i-1})$, $i = 1, \ldots, k$. Starting from the independent gamma variates $X_i = X(s_i) - X(s_{i-1}) \sim \mathrm{Ga}(\alpha_i, \beta)$, $i = 1, \ldots, k$, one easily finds that the fractions

$$Q_i = Q(s_i) - Q(s_{i-1}) = \frac{X(s_i) - X(s_{i-1})}{X(\infty)},$$

$i = 1, \ldots, k$, are independent of $X(\infty)$, that $X(\infty) \sim \mathrm{Ga}(\alpha(\infty), \beta)$ (of course), and that $(Q_1, \ldots, Q_k) \sim \mathrm{Dir}(\alpha_1, \ldots, \alpha_k)$, the Dirichlet distribution with density

$$\mathrm{dir}(q_1, \ldots, q_k;\ \alpha_1, \ldots, \alpha_k) = \frac{\Gamma(\sum_{j=1}^k \alpha_j)}{\prod_{j=1}^k \Gamma(\alpha_j)} \prod_{j=1}^k q_j^{\alpha_j - 1},$$

$q_j > 0$, $j = 1, \ldots, k$, $q_1 + \ldots + q_k = 1$. In particular (taking $k = 2$), $Q(s) \sim \mathrm{Be}(\alpha(s), \alpha(\infty) - \alpha(s))$, where $\mathrm{Be}(\alpha, \beta)$ is the beta distribution with density

$$\mathrm{be}(q;\ \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} q^{\alpha - 1}(1 - q)^{\beta - 1},$$

$0 < q < 1$. The stochastic process Q thus defined is called the Dirichlet process with parameter $\alpha = \{\alpha(s),\ s \in \mathcal{R}\}$, and we write $Q \sim \mathrm{Dir}(\alpha)$. The Dirichlet process plays an important role in nonparametric Bayesian analysis, see Ferguson (1972).

The moments of $Q(s)$ are easily calculated. In particular, $\mathrm{E}[Q(s)] = \alpha(s)/\alpha(\infty)$, showing that the expected value of Q is just α normed to a probability distribution, and $\mathrm{Var}[Q(s)] = \mathrm{E}[Q(s)](1 - \mathrm{E}[Q(s)])/(\alpha(\infty) + 1)$, showing that the total mass of α is a measure of the precision of the process Q; a large value of $\alpha(\infty)$ means little randomness in Q

The conditional Q-distribution on an interval $(a, b]$ is

$$Q(s|(a, b]) = \frac{Q(s) - Q(a)}{Q(b) - Q(a)} = \frac{X(s) - X(a)}{X(b) - X(a)},\ a < s \leq b.$$

Putting $X(s) - X(a)$ and $X(b) - X(a)$ in the roles of $X(s)$ and $X(\infty)$, respectively, in the construction above, the whole story repeats itself We find that $Q(s|(a, b]) \sim \mathrm{Dir}(\alpha_{(a,b]})$, where $\alpha_{(a,b]}$ is the restriction of α to $(a, b]$, and that it is independent of $X(b) - X(a)$ and $X(r)$ for $r \notin (a, b]$ Thus, conditional Q-distributions on disjoint intervals are independent Dirichlet processes.

C. Predicting the outstanding part of Dirichlet type payments

We adopt the general model in Paragraph A above with partial payments Y' of the form (4.1), where $Q \sim \mathrm{Dir}(\alpha)$. Of course, in this context α is concentrated on the unit interval [0, 1], i.e. $0 = \alpha(0-) < \alpha(1) = \alpha(\infty)$.

Let τ denote the present time and consider a reported but not settled claim occurred at time $t < \tau$, notified with a delay $U = u < \tau - t$, hence $V > v' = \tau - t - u$, and for which we have observed the partial cumulative payments $Y'(v_j)$ at development times $0 \leq v_1 < .. < v_k = v'$. Denote all this information by $\mathcal{F}'$. The natural predictor of the outstanding payments on the claim is

$$\mathrm{E}[Y|\mathcal{F}'] - Y'(v'). \tag{4.3}$$

It is unbiased per definition. More generally, by the law of iterated expectation, any predictor of the form $E[Y|\mathcal{F}''] - Y'(v')$, with $\mathcal{F}'' \subset \mathcal{F}'$, is unbiased.

To obtain an expression for (4.3), let us derive the joint distribution of the random variables involved. The quantities

$$U, V, Y, \{Y'(v_j);\ j = 1, \quad , k\}$$

correspond one-to-one with

$$U, V, Y, \{Q_j;\ j = 1, ..., k\}, Y'(v'), \tag{4 4}$$

where

$$Q_j = \frac{Y'(v_j) - Y'(v_{j-1})}{Y'(v')} = \frac{Q(v_j/V) - Q(v_{j-1}/V)}{Q(v'/V)}$$

(recall (4 1)), with the interpretation $v_0 = -\infty$. By use of the results in the previous paragraph, we find that the joint density of the variates in (4 4) is

$$\begin{aligned} p_{v_1, \quad , v_k}(u, v, y, q_1, ..., q_k, y') &= p(u, v, y) \times \\ &\mathrm{dir}\left(q_1, ..., q_k;\ \alpha\left(\frac{v_1}{v}\right) - \alpha\left(\frac{v_0}{v}\right), .., \alpha\left(\frac{v_k}{v}\right) - \alpha\left(\frac{v_{k-1}}{v}\right)\right) \times \\ &\mathrm{be}\left(\frac{y'}{y},\ \alpha\left(\frac{v_k}{v}\right), \alpha(1) - \alpha\left(\frac{v_k}{v}\right)\right)\frac{1}{y}, \end{aligned} \tag{4 5}$$

$0 < u$, $0 < v$, $0 < y' \leq y$, $q_j > 0$, $j = 1, ..., k$ and $q_1 + ... + q_k = 1$, for $0 \leq v_1 < \quad < v_k = v'$ Thus, we obtain the following expression for the first term in (4 3)·

$$\mathrm{E}[Y|\mathcal{F}'] = \frac{\int_{y>y'} \int_{v>v'} y p_{v_1, \quad , v_k}(u, v, y, q_1, .., q_k, y') dv\, dy}{\int_{y>y'} \int_{v>v'} p_{v_1, \quad , v_k}(u, v, y, q_1, ..., q_k, y') dv\, dy}.$$

Numerical techniques are required to compute this fairly complex expression. It is the double integrals that represent the hard part of the problem, and so it does not bring any great computational relief to skip the information contained in the fractions Q_j, $j = 1, \ldots, k$. In fact, since the Q_j are expected to reproduce their conditional means $(\alpha(v_j/V) - \alpha(v_{j-1}/V))/\alpha(v'/V)$, they may provide valuable information on V and, thereby, also on Y if the shape of α differs significantly from the uniform distribution

Pursuing these considerations, we note that also the remaining time until settlement may be predicted on the basis of $\mathcal{F}'$ The predictive distribution of V has density

$$\frac{\int_{y>y'} p_{v_1,\ldots,v_k}(u, v, y, q_1, \ldots, q_k, y')dy}{\int_{y>y'} \int_{v''>v'} p_{v_1,\ldots,v_k}(u, v'', y, q_1, \ldots, q_k, y')dv''dy}$$

We round off this paragraph with a few words about the aptness of the Dirichlet process as a description of the partial payment process. The Dirichlet process is purely discrete and has infinitely many jumps in every interval where the continuous part of α has strictly positive mass, see Ferguson (1972) Admittedly, such path properties do not comply with the behaviour of real life payment streams, which certainly also are purely discrete, but have isolated jumps. However, such myopic considerations may be subordinate to the important fact that the Dirichlet process is able to depict virtually any conceivable pattern of payments by suitable choice of α.

D. Mixed business; a brief sketch

In some practical applications of the theory the data analysis suggested that the claim size distribution P_Y be bimodal and in some cases even multimodal. A closer examination of claim records uncovered that claims were of different types, e g in accident insurance they could be permanent injury (disablement), medical bills, or tooth damage. It also turned out in some cases that the claim type would be established at some time between notification and final settlement. Such situations can be dealt with by augmenting the mark with an index G for type as in the proof of Theorem 3 and, possibly, also a waiting time V^* from notification until the type comes to the case-handler's knowledge ($V^* \leq V$)

Prediction of *nr*-and *cn*-claims goes as before. Prediction of outstanding payments on *rns*-claims goes basically along the same lines as in the previous paragraph. After the type is known one uses the predictor above, only with type $G = g$ and $V^* = v^*$ included in the conditioning. Until the type is known one has to integrate out these variates, summing over all g and integrating over $v^* > v'$ along with integrating over $v > v'$.

Finally, we note that observable covariates can be included in the analysis by a suitable extension of the mark, confer Section 7 in (1).

5 THE DUAL ROLES OF OCCURRENCES AND REPORTS

A. The process of reported claims

Just to keep notation simple, let us assume for the time being that the waiting time distribution $P_{U|t}$ is absolutely continuous with density $p_{U|t}$

We now change the point of view and order the claims by time of notification. Thus, for the generic claim (T, U, V, Y, Y') we take the time of report $\tilde{T} = T + U$ as *the time* and, accordingly, let the remainder of the claim characteristics, $\tilde{Z} = (T, V, Y, Y')$, constitute the mark. This way we get a process

$$\{(\tilde{T}_i, \tilde{Z}_i)\}_{i=1,\dots,N} \tag{5.1}$$

on $\mathcal{T} \times \tilde{\mathcal{Z}}$, defined in an obvious manner. Very conveniently, we have the following result, which is easy to interpret:

Theorem 4. *The process in (5.1) is of marked Poisson type,*

$$\{(\tilde{T}_i, \tilde{Z}_i)\}_{i=1,\dots,N} \sim Po\Big(\tilde{w}(\tilde{t}), \tilde{P}_{\tilde{Z}|\tilde{t}}, \tilde{t} > 0\Big),$$

with intensity

$$\tilde{w}(\tilde{t}) = \int_0^{\tilde{t}} w(s) p_{U|s}(\tilde{t} - s)\, ds, \tag{5.2}$$

and time-dependent mark distribution given by

$$\tilde{P}_{\tilde{Z}|\tilde{t}}(d\tilde{z}) = \frac{\tilde{w}(\tilde{t})}{W} \tilde{p}_{T|\tilde{t}}(t)\, dt\ P_{VYY'|t,\tilde{t}-t}(dv, dy, dy'), \tag{5.3}$$

where

$$\tilde{p}_{T|\tilde{t}}(t) = \frac{w(t) p_{U|t}(\tilde{t} - t)}{\int_0^{\tilde{t}} w(s) p_{U|s}(\tilde{t} - s)\, ds}. \tag{5.4}$$

Proof The key to the proof is the representation Theorem 1. In the setting of that theorem, the "tilde" process is obtained by just representing the generic claim $C = (T, U, V, Y, Y')$ equivalently as $\tilde{C} = (\tilde{T}, T, V, Y, Y')$, that is, transforming (T, U) to $(\tilde{T}, T) = (T + U, T)$. Under this transform the density of (T, U), which is

$$\frac{w(t)}{W} p_{U|t}(u),\ t > 0,\ u > 0,$$

gives rise to the following density of $(\tilde{T}, T)$ at $(\tilde{t}, t)$, $0 < t < \tilde{t}$:

$$\frac{w(t)}{W} p_{U|t}(\tilde{t} - t) = \frac{\tilde{w}(\tilde{t})}{W} \tilde{p}_{T|\tilde{t}}(t).$$

This establishes (5.3), since the conditional distribution of the rest of the mark is unaffected by the transformation – just insert $(t, u) = (t, \tilde{t} - t)$ in $P_{VYY'|tu}$. Inspecting (5.3), noting that $\int_0^\infty \tilde{w}(\tilde{t}) d\tilde{t} = W$, and comparing with Theorem 1, we arrive at the conclusion. □

Another route to (5.2) goes as follows. Starting from the original process, fixing $\tilde{t}$, and applying the decomposition theorem to the claims reported within time $\tilde{t}$, we know that they form a marked Poisson process with total exposure

$$W^r(\tilde{t}) = \int_0^{\tilde{t}} w(t) P_{U|t}(\tilde{t} - t) dt$$

Differentiating with respect to $\tilde{t}$ and using the fact that $P_{U|t}(0) = 0$, we find again that the intensity of reports at any time $\tilde{t}$ is given by (5.2)

B. The chicken and the egg

People with a statistical background might be inclined to take the flow of observable events as basic and, accordingly, claim that one should start from the tilde process and let the $\tilde{T}_i$, not the T_i, take the role of the *times* Paragraph A above tells us that, from a mathematical point of view, either way is fine, and that it is not important to discuss which came first, occurrence or notification; we remain in the marked Poisson scenario anyway The author's opinion is that, at the stage of specifying the distributions, it is easier to let imagination start from the occurrences and build the model from there.

References

FERGUSON, T (1972) A Bayesian analysis of some nonparametric problems *Ann Statist* 1 209-230

HESSELAGER, O (1995) Modelling of discretized claim numbers in loss reserving *ASTIN Bull* 25 119-135

KARR, A F (1991) *Point Processes and Their Statistical Inference* 2nd ed Marcel Decker

NORBERG, R (1986) A contribution to modelling of IBNR claims *Scand Actuarial J* 1986 155-203

NORBERG, R (1993) Prediction of outstanding liabilities in non-life insurance *ASTIN Bull* 23 95-115

Ragnar Norberg
Laboratory of Actuarial Mathematics,
Universitetsparken 5,
DK-2100 Copenhagen Ø,
Denmark
Statistics Department,
London School of Economics and Political Science,
Houghton Street,
London WC2A 2AE,
United Kingdom