TRANSACTIONS OF SOCIETY OF ACTUARIES
1984 VOL. 36

MODELS FOR THE DISTRIBUTION OF AGGREGATE CLAIMS IN RISK THEORY

HARRY H. PANJER AND GORDON E. WILLMOT

ABSTRACT

This paper considers the distribution of aggregate claims of an insurer. The general form of the distribution is considered initially, and, after a few mild restrictions are imposed, it is characterized as being a compound distribution. Various models for the aggregate claims as well as the number-of-claims distribution are presented. Special emphasis is placed on characterizing the aggregate claims random variable as Compound Poisson, using the concepts of infinite divisibility and Bayesian uncertainty models. Results from stochastic process theory are applied to the modeling of the risk process faced by an insurer. The paper further provides a comprehensive review of related results that have appeared in the actuarial and statistical literature.

I. INTRODUCTION

One of the primary goals of any insurance risk endeavor is to arrive at a satisfactory model for the probability distribution of the total costs of insurance claims, usually called the distribution of aggregate claims. The collective theory of risk is based on the assumption that the counting process representing the number of claims is a Poisson process and the associated cumulative or compound process (in the terminology of Cox [8]) representing the total claim amount is thus Compound Poisson.

It has been found in many instances that the number-of-claims process is not necessarily of Poisson type, and so alternative assumptions have to be made concerning these two stochastic processes. The type of risk insured dictates the form, and it is the aim of this paper to survey and propose various models.

In many cases it can be shown that the aggregate claims process is still of Compound Poisson form (which allows for the utilization of the collective theory of risk). This property is thus desirable, and one of the primary tools used to verify this property is the theory of infinite divisibility. To this end, infinitely divisible distributions are discussed in various parts of this paper.

It is not the aim of this paper to present method of inference for these models, but rather to discuss the models themselves along with some of their basic properties.

Earlier surveys include those of Cramér [10, 11], Seal [36], Bühlmann [6], and Beekman [1]. The recent work of Gerber [18] and Bowers, Gerber, Hickman, Jones, and Nesbitt [4] should also be valuable to the reader.

The reader should note the numbering system for equations in this paper. An equation referenced by a single number, e.g. (6), refers to an equation in the current subsection; an equation referenced by two numbers, e.g. (3.6), refers to an equation in another subsection of the current section, while an equation referenced by three numbers, e.g. (III.3.6), refers to an equation in another section.

II. THE AGGREGATE CLAIMS DISTRIBUTION

II.1. *General Representation*

Initially, definitions are made of the random variables representing the number-of-claims and the aggregate claims processes.

Definition 1. The number of claims in the time interval $(0, t]$ is denoted by $N(t)$ whenever $t > 0$, and it is assumed that $N(0) = 0$.

Definition 2. The total claim amount in the interval $(0, t]$ is denoted by $X(t)$ if $t > 0$, and $X(0) = 0$.

These are the two basic processes used. In general, it is assumed that $X(t)$ has only step sample functions (i.e., the sample paths of $X(t)$ only change vertically at times of claims). Thus, as mentioned previously, $\{N(t);\ t \geqslant 0\}$ is a counting process, and $\{X(t);\ t \geqslant 0\}$ the associated cumulative process.

To derive the distribution of $X(t)$, the following definitions are needed as well.

Definition 3.

$$G_t(x) = P\{X(t) \leqslant x\}, \qquad x \geqslant 0. \tag{1}$$

Definition 4. The time of the occurrence of the ith claim is denoted by T_i for $i = 1, 2, 3, \ldots$.

Definition 5.

$$P_{n,t}(t_1, t_2, \ldots, t_n) = P\{T_1 \leqslant t_1, T_2 \leqslant t_2, \ldots, T_n \leqslant t_n, N(t) = n\}, \qquad 0 < t_1 < t_2 < \ldots < t_n < t. \tag{2}$$

Definition 6.

$$G_{n,t}(x|\{t_1, t_2, \ldots, t_n\}) = P\{X(t) \leq x | T_1 = t_1, \ldots, T_n = t_n, N(t) = n\},$$
$$0 < t_1 < t_2 < \ldots < t_n < t. \tag{3}$$

Definition 7.

$$F_{n,t}(y|x) = P\{X(t) \leq y | X(t-0) = x, N(t-0) = n, N(t) = n+1\}. \tag{4}$$

The possibility of multiple claims at any time is excluded, and hence $F_{n,t}(y|x)$ represents the conditional distribution function of a single claim at time t given that immediately before time t, there were n claims totaling x units.

THEOREM II.1.1. *The distribution function of the total claim amount distribution is given by*

$$G_t(x) = \sum_{n=0}^{\infty} \int_{0<t_1<t_2<\ldots<t_n<t} G_{n,t}(x|\{t_1, t_2, \ldots, t_n\}) dP_{n,t}(t_1, t_2, \ldots, t_n), \tag{5}$$

where

$$G_{n,t}(x|\{t_1, t_2, \ldots, t_n\}) = \int_{R^{n-1}} F_{n-1,t_n}(x|x_{n-1}) dF_{n-2,t_{n-1}}(x_{n-1}|x_{n-2}) \ldots dF_{0,t_1}(x_1|0), \tag{6}$$

and the integrals are taken to be Lebesgue-Stieltjes.

Proof. See Bühlmann ([6], p. 55).

The general representation for $G_t(x)$ is thus given by (5). In the next section some of the assumptions made are relaxed in order to arrive at the usual form of $G_t(x)$.

II.2. *The Compound Distribution*

A relatively simple form for the distribution function $G_t(x)$ may be obtained if the individual claim size distribution is independent of time. The following definition is thus made.

Definition 1. Let Y_i denote the amount of the ith claim for $i = 1, 2, 3, \ldots$.

The special form of $G_t(x)$ for time-independent claim amounts may now be derived.

THEOREM II.2.1. *If the claim size distribution is independent of time, i.e.,*

$$F_{n,t}(y|x) = F_n(y|x), \tag{1}$$

we may write

$$G_t(x) = \sum_{n=0}^{\infty} p_n(t) G(x|n), \tag{2}$$

where

$$p_n(t) = P\{N(t) = n\} \tag{3}$$

and

$$G(x|n) = P\left\{ \sum_{i=1}^{n} Y_i \leq x | N(t) = n \right\}. \tag{4}$$

Proof. Since claim amounts are independent of time, (1.6) may be written as

$$G(x|n) = \int_{R^{n-1}} F_{n-1}(x|x_{n-1}) dF_{n-2}(x_{n-1}|x_{n-2}) \ldots dF_0(x_1|0). \tag{5}$$

Taking (5) outside the integral in (1.5) and noting that

$$p_n(t) = \int_{0<t_1<t_2<\ldots<t_n<t} dP_{n,t}(t_1, t_2, \ldots, t_n) \tag{6}$$

yields (2).

The distribution (2) is often referred to as a mixed distribution. Alternatively, it is seen that

$$X(t) \equiv Y_1 + Y_2 + \ldots + Y_{N(t)}. \tag{7}$$

This is referred to by Feller [16] as a random sum. The Y_i's are not necessarily independent or identically distributed (i.i.d.). However, if they are i.i.d., then the following form $G_t(x)$ may be obtained.

THEOREM II.2.2. *If the claim sizes are identically distributed as well as independent given $N(t)$,*

$$F_n(y|x) = F(y - x), \tag{8}$$

then

$$G_t(x) = \sum_{n=0}^{\infty} p_n(t) F^{n*}(x). \tag{9}$$

Proof. We note that $G(x|n)$ in (5) becomes the n-fold convolution of F with itself, and the result follows.

The form (9) is the more useful and common representation for $G_t(x)$. Henceforth (9) will be referred to as a *compound distribution.* The variables Y_i are further assumed to be nonnegative and i.i.d.

Generating functions are now introduced in order to derive some relationships between $X(t)$, $N(t)$, and the Y_i's.

Definition 2.

$$P(s, t) = \sum_{n=0}^{\infty} p_n(t) s^n. \tag{10}$$

Definition 3.

$$\phi_t(s) = \int_0^{\infty} e^{-sx} dG_t(x). \tag{11}$$

Definition 4.

$$\psi(s) = \int_0^{\infty} e^{-sx} dF(x). \tag{12}$$

The following relationship between the above functions is satisfied.

LEMMA II.2.1.

$$\phi_t(s) = P(\psi(s), t). \tag{13}$$

Proof.

$$\begin{aligned}\phi_t(s) &= E_{N(t)}\{E[e^{-sX(t)}|N(t)]\}\\ &= E_{N(t)}\{\psi(s)^{N(t)}\} = P(\psi(s), t).\end{aligned}$$

Equation (13) also holds if $\phi_t(s)$ and $\psi(s)$ are characteristic functions, moment generating functions, or probability generating functions (if Y_i is discrete), as well. The following relationships between the moments of $X(t)$, $N(t)$, and Y_i then follow.

THEOREM II.2.3. *The nth moment about the origin of* $X(t)$ *satisfies*

$$E\{X(t)^n\} = \sum_{j=0}^{n} \sum_{A_{n,j}} \frac{n!}{k_1!k_2! \ldots k_n!} \left\{ \prod_{i=1}^{n} \left[\frac{E(Y^i)}{i!} \right]^{k_i} \right\} E\{N(t)^{(j)}\}, \tag{14}$$

where $A_{n,j}$ is the set of nonnegative integers $(k_1, k_2, \ldots, k_n)$ such that

$$\sum_{i=1}^{n} k_i = j, \qquad \sum_{i=1}^{n} ik_i = n,$$

and Y is a single claim amount.

Proof. Kupper [23] gives the following formula for the derivative of a composite function:

$$\frac{d^n}{dx^n} f(g(x)) = \sum_{j=0}^{n} \sum_{A_{n,j}} \frac{n!}{k_1!k_2! \ldots k_n!} \left[\frac{g'(x)}{1!}\right]^{k_1} \cdots \left[\frac{g^{(n)}(x)}{n!}\right]^{k_n} \frac{d^j}{du^j} f(u)\bigg|_{u=g(x)}. \tag{15}$$

Furthermore, the following are seen to hold.

$$\frac{\partial^n}{\partial s^n}\phi_t(s)\bigg|_{s=0} = E\{X(t)^n\};$$

$$\frac{d^n}{ds^n}\psi(s)\bigg|_{s=0} = E\{Y^n\}; \tag{16}$$

$$\frac{\partial^n}{\partial s^n}P(s, t)\bigg|_{s=1} = E\{N(t)^{(n)}\}.$$

Thus, taking the nth partial derivative of (13) with respect to s and using (15) and (16), (14) follows.

In particular, for $n = 1, 2$, (14) becomes

$$E\{X(t)\} = E\{N(t)\}E\{Y\} \tag{17}$$

and

$$E\{X(t)^2\} = E\{N(t)^{(2)}\}[E\{Y\}]^2 + E\{N(t)\}E\{Y^2\}. \tag{18}$$

The following definition is now made for completeness.

Definition 5.

$$F^{0*}(x) = 1, \qquad x \geq 0. \tag{19}$$

Also, it is assumed that

$$F^{n*}(0) = 0, \qquad n > 0. \tag{20}$$

The distribution functions $G_t(x)$, $F^{n*}(x)$ are now assumed to be differentiable for $x > 0$. The following probability density functions are defined.

Definition 6.

$$g_t(x) = \frac{\partial}{\partial x} G_t(x), \qquad x > 0. \tag{21}$$

Definition 7.

$$f^{n*}(x) = \frac{d}{dx} F^{n*}(x). \tag{22}$$

The following is then an alternative representation for $G_t(x)$.

THEOREM II.2.4.

$$G_t(x) = p_0(t) + \int_0^x g_t(y)dy. \tag{23}$$

Proof. From (19),

$$G_t(x) = p_0(t) + \sum_{n=1}^{\infty} p_n(t) F^{n*}(x),$$

and so from (21)

$$g_t(x) = \sum_{n=1}^{\infty} p_n(t) f^{n*}(x). \tag{24}$$

Thus,

$$\int_0^x g_t(y)dy = \sum_{n=1}^{\infty} p_n(t) F^{n*}(x) = G_t(x) - p_0(t),$$

or

$$G_t(x) = p_0(t) + \int_0^x g_t(y)dy.$$

This result was given by Lundberg [25]. Hence $X(t)$ is a mixed discrete-continuous random variable with $P\{X(t) = 0\} = p_0(t)$.

II.3. *Some Common Examples*

In general, (2.9) and (2.23) do not allow for a simple expression for $G_t(x)$; however, in certain instances simpler forms may be obtained. Considered here are the binomial, the negative binomial (including Pascal and geometric), and the Poisson number-of-claims distributions together with exponential claim amount distributions. The incomplete gamma function is needed and is now given.

Definition 1.

$$I(k, t) = \int_0^t s^{k-1} e^{-s} ds / \Gamma(k), \qquad k>0 \tag{1}$$

is defined as the incomplete gamma function.

For the special case when k is a positive integer, this definition is equivalent to the following (verified easily by integration by parts).

Definition 1(a).

$$I(k, t) = 1 - \sum_{j=0}^{k-1} \frac{t^j e^{-t}}{j!}, \qquad k \in \{1, 2, 3, \ldots\}. \tag{2}$$

a) THE EXPONENTIAL DISTRIBUTION FOR CLAIM AMOUNTS

The exponential distribution function for single claim amounts is given by

$$F(x) = 1 - e^{-\mu x}, \qquad x \geqslant 0. \tag{3}$$

The n-fold convolution is then given by

$$F^{n*}(x) = \int_0^x \frac{\mu(\mu t)^{n-1} e^{-\mu t}}{(n-1)!} dt = I(n, \mu x). \tag{4}$$

This gamma distribution is fairly simple to work with, and explicit results for $G_t(x)$ usually assume (3) as the claim amount distribution.

b) THE POISSON DISTRIBUTION FOR CLAIM NUMBERS

The Poisson probability function for the number of claims is given by

$$p_n(t) = \frac{[\lambda(t)]^n e^{-\lambda(t)}}{n!}, \qquad n = 0, 1, 2, \ldots. \tag{5}$$

The probability-generating function is thus

$$P(s, t) = \exp\{\lambda(t)(s-1)\}. \tag{6}$$

The Laplace transform corresponding to (6) for the aggregate claims distribution is then (from (2.13))

$$\phi_t(s) = \exp\{\lambda(t)[\psi(s) - 1]\}. \tag{7}$$

The transform (7) is of vital importance in the theory of stochastic processes with stationary and independent increments as well as in the collective theory of risk. The reasons for this importance will be outlined in (II.5). For these and other reasons, (7) is often taken to be the transform for the aggregate claims distribution.

c) THE NEGATIVE BINOMIAL DISTRIBUTION FOR CLAIM NUMBERS

This number-of-claims distribution is given by

$$p_n(t)=\binom{-\beta(t)}{n}\left[\frac{1}{1+a(t)}\right]^{\beta(t)}\left[\frac{-a(t)}{1+a(t)}\right]^n, \quad n=0,\ 1,\ 2,\ \ldots, \tag{8}$$

where $\beta(t) > 0$, $0 < a(t)$. The probability-generating function is then

$$P(s,\ t)=\{1-a(t)(s-1)\}^{-\beta(t)}. \tag{9}$$

If $\beta(t)$ is an integer, (8) is sometimes called the Pascal distribution (see Johnson and Kotz [21]). Also, if $\beta(t) = 1$, (8) reduces to the geometric distribution:

$$p_n(t)=\left[\frac{1}{1+a(t)}\right]\left[\frac{a(t)}{1+a(t)}\right]^n, \quad n=0,\ 1,\ 2,\ \ldots. \tag{10}$$

The negative binomial distribution is quite commonly used, since it also arises as a result of a variety of assumptions, again to be outlined later.

d) THE BINOMIAL DISTRIBUTION FOR CLAIM NUMBERS

This distribution has probabilities given by

$$p_n(t)=\binom{N}{n}[p(t)]^n[1-p(t)]^{N-n}, \quad n=0,\ 1,\ 2,\ \ldots,\ N, \tag{11}$$

with probability-generating function

$$P(s,\ t)=\{sp(t)+1-p(t)\}^N. \tag{12}$$

This distribution is useful for several reasons, one of which is that $p_n(t) = 0$ for all $n > N$. Thus (2.9) becomes a finite sum.

With the use of the above-mentioned distributions, some expressions for $G_t(x)$ will now be derived.

Example 1: The Poisson-Exponential Model. In an identical manner to that used in Seal ([36], p. 32), the following expression for $G_t(x)$ is easily derived.

$$G_t(x)=1-e^{-\mu x}\int_0^{\lambda(t)} e^{-s}I_0[2\sqrt{(\mu sx)}]ds, \tag{13}$$

where

$$I_0(x)=\sum_{j=0}^{\infty}\frac{(x/2)^{2j}}{(j!)^2} \tag{14}$$

is a modified Bessel function.

Example 2: The Binomial-Exponential Model. Using (2) and (4),

$$\begin{aligned} G_t(x) &= p_0(t) + \sum_{k=1}^{N} p_k(t) F^{k*}(x) \\ &= p_0(t) + \sum_{k=1}^{N} p_k(t) \left\{ 1 - \sum_{j=0}^{k-1} \frac{(\mu x)^j e^{-\mu x}}{j!} \right\} \\ &= 1 - \sum_{k=1}^{N} \binom{N}{k} [p(t)]^k [1-p(t)]^{N-k} \left\{ \sum_{j=0}^{k-1} \frac{(\mu x)^j e^{-\mu x}}{j!} \right\}. \end{aligned} \tag{15}$$

A similar result to (15) is easily obtained if claims are gamma distributed with integer exponent.

Example 3: The Pascal-Exponential Model. The Laplace transform of $G_t(x)$ is $\gamma_t(s)$, which satisfies

$$\gamma_t(s) = \phi_t(s)/s, \tag{16}$$

and that of the exponential probability density function (p.d.f.) is

$$\psi(s) = \mu/(\mu + s). \tag{17}$$

Thus, using (2.13) and (12) for the binomial-exponential model, the transform of (15) is

$$\begin{aligned} \gamma_t(s) &= \frac{1}{s} \left\{ \left(\frac{\mu}{\mu + s} \right) p(t) + [1 - p(t)] \right\}^N \\ &= \frac{1}{s} \left\{ \frac{s[1-p(t)] + \mu}{s + \mu} \right\}^N. \end{aligned} \tag{18}$$

For the Pascal-exponential model using (9) instead of (12)

$$\gamma_t(s) = \frac{1}{s} \left\{ \frac{s[1 + a(t)]^{-1} + \mu[1 + a(t)]^{-1}}{s + \mu[1 + a(t)]^{-1}} \right\}^{\beta(t)}, \tag{19}$$

which is of the same form as (18). Hence the Pascal-exponential model can be written as binomial-exponential; and, upon comparison with (15), it is seen that

$$\begin{aligned} G_t(x) = 1 - \sum_{k=1}^{\beta(t)} \binom{\beta(t)}{k} \left[\frac{a(t)}{1 + a(t)} \right]^k \left[\frac{1}{1 + a(t)} \right]^{\beta(t) - k} \\ \times \left\{ \sum_{j=0}^{k-1} \frac{1}{j!} \left[\frac{\mu x}{1 + a(t)} \right]^j e^{-\mu x/[1 + a(t)]} \right\}. \end{aligned} \tag{20}$$

The fact that $G_t(x)$ can be written as a finite sum is a somewhat surprising result and is given by Panjer and Willmot [33].

Example 4: The Geometric-Exponential Model. If $\beta(t) = 1$ in (20), then

$$G_t(x) = 1 - \left[\frac{a(t)}{1 + a(t)}\right] e^{-\mu x/[1 + a(t)]}. \tag{21}$$

When the claim frequencies are of Poisson, binomial or negative binomial form, and the claim size distribution is discrete, simple computational methods allow the aggregate claims distribution to be easily calculated. These results are given by Panjer [30, 31], Sundt and Jewell [42], Gerber [19], and Panjer and Willmot [33].

II.4. *The Compound Poisson Distribution*[1]

Any distribution with characteristic function or Laplace transform of the form (3.7) is termed a Compound Poisson distribution. It turns out that even if the number of claims is not necessarily of Poisson form, it may be the case that the distribution of aggregate claims is a Compound Poisson distribution. The reasons for this are outlined in the following subsection, which draws from the theory of infinite divisibility.

a) THE AGGREGATE CLAIMS DISTRIBUTION AS COMPOUND POISSON

Infinitely divisible distributions are now defined.

Definition 1. Let A be a (measurable) subset of R, the set of real numbers. Then a random variable (or its distribution function, or its characteristic function) is said to be infinitely divisible (i.d.) on A if its distribution is concentrated on A and if, for each positive integer n, its characteristic function $\phi(s)$ can be displayed as $\phi(s) = \{\phi_n(s)\}^n$, where $\phi_n(s)$ is the characteristic function of some distribution on A. If A is the set R, then the random variable is just said to be infinitely divisible.

It is clear that infinite divisibility on a subset A of R implies infinite divisibility. In particular, infinite divisibility on $R_+ = [0, \infty)$ implies infinite divisibility. The converse, however, is not necessarily true. Infinite divisibility, for example, does not imply infinite divisibility on $N_+ = \{0, 1, 2, \ldots\}$. (The constant 1 is infinitely divisible, but not on N_+.)

The importance of infinite divisibility in stochastic processes is due in part to its relationship with processes with stationary and independent increments in continuous time.

[1] The results of subsections II.4 and III.1 also appear in the paper "Compound Poisson Models in Actuarial Risk Theory," *Journal of Econometrics*, XXIII (1983), 63–76, and are included here for completeness.

Definition 2. A stochastic process $\{Z(t); t \geqslant 0\}$ is said to have independent increments if, for all choices of $t_0 < t_1 < \ldots < t_n$, the n random variables $Z(t_i) - Z(t_{i-1})$, $i = 1, 2, \ldots, n$ are independent.

Definition 3. A stochastic process $\{Z(t); t \geqslant 0\}$ is said to have stationary increments if the distribution of $Z(t_1 + s) - Z(t_0 + s)$ is the same as that of $Z(t_1) - Z(t_0)$ for all $t_0, t_1, s \geqslant 0$.

THEOREM II.4.1. *If the stochastic process $\{Z(t); t \geqslant 0\}$ has stationary and independent increments (s.i.i), then $Z(t)$ is infinitely divisible for any $t > 0$.*

Proof.

$$Z(t) = \sum_{i=1}^{n} \{Z(it/n) - Z((i-1)t/n)\},$$

and thus, if $\phi_t(s)$ is the characteristic function of $Z(t)$,

$$\phi_t(s) = \{\phi_{t/n}(s)\}^n, \qquad n = 1, 2, \ldots,$$

where we have assumed $Z(0) = 0$.

The next few theorems outline the close connection between infinite divisibility and the Compound Poisson law.

THEOREM II.4.2. *A Compound Poisson characteristic function is infinitely divisible.*

Proof. $\exp\{\lambda(t)[\psi(s) - 1]\} = (\exp\{\lambda(t)[\psi(s) - 1]/n\})^n$.

THEOREM II.4.3. *A characteristic function is infinitely divisible if and only if (iff) it is the pointwise limit of a sequence of Compound Poisson characteristic functions, that is, it has the form*

$$\lim_{m \to \infty} \exp\{p_m[g_m(s) - 1]\},$$

where the p_m's are positive constants and the $g_m(s)$'s are characteristic functions.

Proof. See Lukacs ([24], p. 112).

There is a very strong characterization of the distribution of aggregate claims as Compound Poisson as a result of the following theorems.

THEOREM II.4.4. *A random variable with support on $[0, \infty)$ and distribution function $F(x)$ such that $F(0) > 0$ is infinitely divisible iff it is Compound Poisson.*

Proof. See Van Harn ([48], p. 24).

Theorems II.4.1, II.4.2, and II.4.4 imply that the aggregate claims random variable $X(t)$ is Compound Poisson if the associated stochastic process has s.i.i. and $p_0(t) > 0$.

THEOREM II.4.5. *A stochastic process with step sample paths has stationary and independent increments iff it is a Compound Poisson process.*

Proof. See Delbaen and Haezendonck [12].

This certainly implies that a stochastic process with stationary and independent increments and step sample paths is Compound Poisson. Bühlmann [5] proved this result under the additional conditions that the expected number of jumps in a finite time interval is finite or that the process takes on integral values only.

THEOREM II.4.6. *A random variable is infinitely divisible on N_+ iff it is Compound Poisson with nonnegative integer terms.*

Proof. See Feller [16].

The following theorem gives another characterization of the aggregate claims distribution as being of Compound Poisson type.

THEOREM II.4.7. *The aggregate claims distribution is Compound Poisson iff the associated claim number distribution is Compound Poisson.*

Proof. See Thyrion [47].

Theorem II.4.7 states that $X(t)$ is Compound Poisson if $N(t)$ has s.i.i. Thus there is a close connection between infinite divisibility, stochastic processes with stationary and independent increments, and the Compound Poisson law. Infinitely divisible distributions are useful in another context to be discussed in the next section.

Some properties of Compound Poisson distributions will now be mentioned.

b) PROPERTIES OF THE COMPOUND POISSON DISTRIBUTION

THEOREM II.4.8. *Let W have a Compound Poisson distribution with moment-generating function*

$$\phi(t) = \exp\{\lambda[\psi(s) - 1]\}, \tag{1}$$

where $\psi(s)$ is the moment-generating function of Y. Further, let the jth cumulant of W be k_j and the jth moment about the origin of Y be p_j. Then

$$k_j = \lambda p_j, \quad j = 1, 2, 3, \ldots. \tag{2}$$

Proof. From (1), it is seen that

$$\log \phi(t) = \lambda \sum_{j=1}^{\infty} \frac{t^j}{j!} p_j = \sum_{j=1}^{\infty} \frac{t^j}{j!} k_j. \tag{3}$$

THEOREM II.4.9. *Let W_i have characteristic function*

$$\phi_i(t) = \exp\{\lambda_i[\psi_i(t) - 1]\}, \qquad i = 1, 2, \ldots, n, \tag{4}$$

and let the W_i's be independent. Then $S_n = W_1 + \ldots + W_n$ has characteristic function

$$\phi(t) = \exp\{\lambda[\psi(t) - 1]\}, \tag{5}$$

where

$$\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n \tag{6}$$

and

$$\psi(t) = \lambda^{-1}[\lambda_1\psi_1(t) + \ldots + \lambda_n\psi_n(t)]. \tag{7}$$

Proof.

$$\phi(t) = \prod_{i=1}^{n} \phi_i(t) = \exp\{\lambda[\psi(t) - 1]\}.$$

Hence the sum of a finite number of Compound Poisson random variables is itself Compound Poisson, so that, like the Poisson itself, the Compound Poisson class is closed under convolution.

The following theorems relate to random variables defined on $\{0, 1, 2, \ldots\}$.

THEOREM II.4.10. *Let W have the probability-generating function*

$$\exp\left\{\lambda\left(\sum_{i=0}^{\infty} p_i s^i - 1\right)\right\}. \tag{8}$$

Then

$$W = \sum_{i=1}^{\infty} i S_i, \tag{9}$$

where S_i is a Poisson variate with mean λp_i and S_i and S_j are independent, $i \neq j$.

Proof. Let S_i be the number of jump values of size i, $i = 1, 2, 3, \ldots$.

Then if $i \neq j$,

$$P\{S_i = x,\ S_j = y\} = \sum_{n=x+y}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \binom{n}{x,\ y,\ n-x-y} p_i^x p_j^y (1 - p_i - p_j)^{n-x-y}$$

$$= \frac{e^{-\lambda}}{x!y!} (\lambda p_i)^x (\lambda p_j)^y \sum_{n=x+y}^{\infty} \frac{[\lambda(1 - p_i - p_j)]^{n-x-y}}{(n-x-y)!}$$

$$= \frac{(\lambda p_i)^x e^{-\lambda p_i}}{x!} \frac{(\lambda p_j)^y e^{-\lambda p_j}}{y!}.$$

Also,

$$P\{S_i = x\} = \sum_{y=0}^{\infty} P\{S_i = x,\ S_j = y\}$$

$$= \frac{(\lambda p_i)^x e^{-\lambda p_i}}{x!}.$$

Thus

$$P\{S_i = x,\ S_j = y\} = P\{S_i = x\} P\{S_j = y\},$$

and the result follows.

For further results on infinitely divisible random variables, see Van Harn [48], Steutel [39, 40, 41], Feller [17], or Lukacs [24].

Example 1. The Poisson distribution is Compound Poisson.

Example 2. The negative binomial distribution is Compound Poisson with logarithmic amount jump distribution. The probability-generating function may be written as

$$p(s) = \left\{ \frac{1-q}{1-sq} \right\}^{\beta}$$
$$= \exp\left\{ \left[-\beta \log(1-q) \right] \left[\frac{\log(1-sq)}{\log(1-q)} - 1 \right] \right\}. \tag{10}$$

Example 3. The Compound negative binomial distribution is Compound Poisson, as is the Compound geometric.

Example 4. The shifted logarithmic distribution is Compound Poisson (see Katti [22] for details). In fact, is a log-convex distribution and hence Compound geometric (see Van Harn [48] for details).

Example 5. The distribution defined by

$$p_n = \frac{c\rho^{n+1}}{1-\rho^{n+1}}, \qquad n = 0, 1, 2, \ldots \tag{11}$$

is Compound Poisson. See Katti [22] for details. This distribution arises in queuing theory.

Further examples of Compound Poisson distributions will be derived in later sections.

II.5. *Bayesian Uncertainty Models*

It is often of interest to model situations in which the risks are not all assumed to be identical. One method that has been used successfully is to assume that the form of the distribution is known up to the inclusion of unknown parameters, which are unique to each risk. In other words, the parameters have a distribution associated with them over the set of all risks, termed the collective. This distribution is called the risk distribution or structure function.

The actuarial interpretation of this model is that a risk is selected from the whole set of risks in accordance with the structure function, and the performance of this risk is then monitored.

The statistical interpretation is essentially Bayesian. The structure function is simply the prior distribution of the parameters.

Definition 1. θ is the random variable that characterizes the risk in the collective and has distribution function

$$U(\alpha) = P\{\theta \leq \alpha\}. \tag{1}$$

Definition 2. The following characteristics of the aggregate claims process $\{X(t);\ t \geq 0\}$ are defined conditionally:

$$G_t(x|\theta) = P\{X(t) \leq x|\theta\}; \tag{2}$$

$$\phi_t(s|\theta) = E\{e^{-sX(t)}|\theta\}; \tag{3}$$

$$\mu(\theta) = E\{X(t)|\theta\}; \tag{4}$$

$$\sigma^2(\theta) = V\{X(t)|\theta\}. \tag{5}$$

Definition 3. The following characteristics of the number-of-claims process $\{N(t;\ t \geq 0\}$ are defined conditionally:

$$p_n(t|\theta) = P\{N(t) = n|\theta\}; \tag{6}$$

$$P(s,\ t|\theta) = E\{s^{N(t)}|\theta\}. \tag{7}$$

Definition 4. For the single claim amount distribution the following definitions are made:

$$F(x|\theta) = P\{Y \leq x|\theta\}; \tag{8}$$

$$\psi(s|\theta) = E\{e^{-sY}|\theta\}. \tag{9}$$

Remark 1. Unconditionally, one obtains the above characteristics simply by weighting the conditional characteristics by the distribution function $U(\boldsymbol{\theta})$. Thus, for example,

$$G_t(x) = \int_{\boldsymbol{\theta}} G_t(x|\boldsymbol{\theta})dU(\boldsymbol{\theta}). \tag{10}$$

THEOREM II.5.1. *If* $\boldsymbol{\theta} = (\theta_1, \theta_2)$, *where* θ_1, θ_2 *are independent and*

$$p_n(t|\boldsymbol{\theta}) = p_n(t|\theta_1),$$

$$F(x|\boldsymbol{\theta}) = F(x|\theta_2),$$

then

$$G_t(x) = \sum_{n=0}^{\infty} p_n(t)F^{n*}(x), \tag{11}$$

where

$$p_n(t) = \int_{\theta_1} p_n(t|\theta_1)dU_1(\theta_1),$$

$$F(x) = \int_{\theta_2} F(x|\theta_2)dU_2(\theta_2),$$

and

$$U(\boldsymbol{\theta}) = U_1(\theta_1)U_2(\theta_2).$$

Proof.

$$\begin{aligned} G_t(x) &= \int_{\boldsymbol{\theta}} G_t(x|\boldsymbol{\theta})dU(\boldsymbol{\theta}) \\ &= \int_{\theta_2}\int_{\theta_1} \sum_{n=0}^{\infty} p_n(t|\theta_1)F^{n*}(x|\theta_2)dU_1(\theta_1)dU_2(\theta_2) \\ &= \sum_{n=0}^{\infty} \left\{ \int_{\theta_1} p_n(t|\theta_1)dU_1(\theta_1) \right\} \left\{ \int_{\theta_2} F^{n*}(x|\theta_2)dU_2(\theta_2) \right\} \\ &= \sum_{n=0}^{\infty} p_n(t)F^{n*}(x). \end{aligned}$$

COROLLARY II.5.1. *If* $\boldsymbol{\theta}$ *is a single parameter and either* $p_n(t|\boldsymbol{\theta})$ *or* $F(x|\boldsymbol{\theta})$ *is independent of* $\boldsymbol{\theta}$, *then* $G_t(x)$ *can again be written in the form (11).*

The majority of models considered in this class assume that $p_n(t|\theta)$ depends on θ and that $F(x|\theta)$ does not. This is equivalent to a mixed number-of-claims model, and further characteristics of this model will be outlined in the next chapter.

Example 1. From Example 3.4, let

$$G_t(x|\mu) = 1 - \left\{\frac{a(t)}{1+a(t)}\right\} e^{-\mu x/[1-a(t)]},$$

where

$$dU(\mu) = \frac{\alpha(\alpha\mu)^{\sigma-1}e^{-\alpha\mu}}{\Gamma(\sigma)} d\mu, \qquad \mu>0. \tag{12}$$

Then

$$G_t(x) = 1 - \left\{1 + \frac{x}{\alpha[1+a(t)]}\right\}^{-\sigma} \left\{\frac{a(t)}{1+a(t)}\right\}. \tag{13}$$

Similarly, relatively simple expressions may be obtained using the gamma density (12) for the binomial-exponential model (3.15) and hence also the Pascal-exponential model (3.20).

For a thorough discussion of these models, see Bühlmann [6].

III. THE NUMBER-OF-CLAIMS DISTRIBUTION

In this section models are considered for the number-of-claims probabilities as defined by (II.2.3). Various situations are looked at and the appropriate probabilities derived. In many of these models it is shown that $N(t)$ (and hence, by Theorem II.4.7, $X(t)$) is Compound Poisson. Perhaps the three most common models are the Poisson, negative binomial, and binomial probabilities as defined by (II.3.5), (II.3.8), and (II.3.11), respectively. In the next section Bayesian models are considered. Again the close connection between infinite divisibility and the Compound Poisson law is considered.

III.1. *Bayesian Models and Infinite Divisibility*

In the situation where risks are nonhomogeneous, it is often useful to consider number-of-claims probabilities of the form

$$p_n(t) = \int_\theta p_n(t|\theta) dU(\theta). \tag{1}$$

(See subsection II.5 for details and notation.) It is of interest to know if the unconditional probabilities (1) are from a Compound Poisson distribution. If they are, then by Theorem II.4.7, the aggregate claims distribution is also Compound Poisson. This question may sometimes be answered by the following theorem, which generalizes some results given by Bühlmann [6].

THEOREM III.1.1. *Let T and A be sets in R. For each* $\theta \in T$, *let* $\{Q_t(s)\}^\theta$ *be the characteristic function of a distribution on A. If* $U(\theta)$ *is a distribution function infinitely divisible on T, then*

$$P_t(s) = \int \{Q_t(s)\}^\theta dU(\theta) \tag{2}$$

is the characteristic function of a distribution infinitely divisible on A.

Proof. Let θ be a random variable with distribution function $U(\theta)$. For each positive integer n, θ can be displayed as a sum

$$\theta = \sum_{i=1}^{n} \theta_{ni} \tag{3}$$

of n i.i.d. random variables with distribution concentrated on T. Then

$$P_t(s) = E[\{Q_t(s)\}^\theta] = E[\{Q_t(s)\}^{\sum_{i=1}^n \theta_{ni}}] = (E[\{Q_t(s)\}^{\theta_{nn}}])^n. \tag{4}$$

But

$$\{P_t(s)\}^{1/n} = E[\{Q_t(s)\}^{\theta_{nn}}]$$

is a mixture of characteristic functions of distributions on A, and thus it is the characteristic function of a distribution on A.

COROLLARY III.1.1. *If the sets A and T are both chosen to be* N_+, *then* $P_t(s) = P\{Q_t(s)\}$, *where* $P\{s\}$ *is the probability-generating function of* θ.

Thus $P_t(s)$ is both a mixed and a compound distribution. (See Ord [29], p. 128.)

COROLLARY III.1.2. *If the set A is chosen to be* N_+, *then* $P_t(s)$ *is a Compound Poisson characteristic function with nonnegative terms (by Theorem II.4.6).*

Example 1: The Mixed Poisson Distribution. For the probabilities

$$p_n(t|\theta) = \frac{(\theta t)^n e^{-\theta t}}{n!} \tag{5}$$

it is seen that (2) is satisfied for $\theta > 0$, since

$$P_t(s) = \int_\theta \{e^{t(e^{is}-1)}\}^\theta dU(\theta). \tag{6}$$

Thus, from (6), the unconditional probabilities are Compound Poisson if $U(\theta)$ is infinitely divisible on a subset of R_+.

Example 2: The Mixed Negative Binomial Distribution. For the probabilities

$$p_n(t|\theta) = \binom{-\theta}{n}\left\{\frac{1}{1+a(t)}\right\}^\theta\left\{\frac{-a(t)}{1+a(t)}\right\}^n \tag{7}$$

(2) is satisfied for $\theta > 0$, since

$$P_t(s) = \int_\theta \{1 - a(t)[e^{is} - 1]\}^{-\theta} dU(\theta), \tag{8}$$

and hence the unconditional distribution is Compound Poisson if $U(\theta)$ is infinitely divisible on a subset of R_+.

Example 3: The Mixed Binomial Distribution. For the probabilities

$$p_n(t|\theta) = \binom{\theta}{n}[p(t)]^n[1 - p(t)]^{\theta-n} \tag{9}$$

(2) is satisfied if $U(\theta)$ is defined on the nonnegative integers, since

$$P_t(s) = \int_\theta \{e^{is}p(t) + 1 - p(t)\}^\theta dU(\theta), \tag{10}$$

and hence the unconditional distribution is Compound Poisson if $U(\theta)$ is a discrete Compound Poisson distribution function. This result follows easily from Corollary III.1.1 as well. The jump distribution here is also a compound distribution.

The next theorem generalizes the results of Example 1 somewhat.

THEOREM III.1.2. *If $\chi(s)$ is a characteristic function such that $-i\chi'(0) < \infty$, then*

$$\int_\lambda e^{\lambda t[\chi(s)-1]} dU(\lambda) = e^{\mu(t)[\sigma_t(s)-1]}, \tag{11}$$

where $\sigma_t(s)$ is a characteristic function, $\mu(t) \geq 0$, iff $U(\lambda)$ is infinitely divisible on a subset of R_+.

Proof. See Bühlmann and Buzzi [7].

This theorem says that a "mixed" Compound Poisson distribution (which could be the distribution of either $X(t)$ or $N(t)$) with jump distribution independent of the mixing parameter is itself Compound Poisson if and only if the mixing distribution is infinitely divisible. The conditional Poisson probabilities are assumed to be given by (5).

These two theorems suggest that if the number-of-claims probabilities (1) are mixed over an infinitely divisible mixing function, this often results in the aggregate claims distribution being of Compound Poisson form. The above theorems demonstrate the importance of infinitely divisible distributions in their own right rather than in connection with Compound Poisson distributions. Some examples of infinitely divisible distributions are given in the next subsection.

a) INFINITELY DIVISIBLE DISTRIBUTIONS

Example 4: Mixture of Exponentials.

$$F(x)=\int_0^\infty (1-e^{-\lambda x})dU(\lambda) \tag{12}$$

is an infinitely divisible distribution function for any distribution function U. See Steutel [39].

Example 5: Product of Gammas. Suppose $X(\alpha,o)$ is a random variable with density given by (II.5:12). If $X_1,\ldots,X_n$ are independent, $X_i=X(\alpha_i,o_i)$, then

$$\prod_{i=1}^{n}\{X(\alpha_i,\sigma_i)\}^{k_i} \tag{13}$$

is infinitely divisible if $|k_i| \geqslant 1$, $i=1,2,\ldots,n$. See Bondesson [2].

Example 6: Pareto Distribution.

$$F(x)=1-(1+x/\alpha)^{-\beta}, \qquad x,\alpha,\beta>0 \tag{14}$$

is infinitely divisible. In fact, $F(x)$ may be written in the form (12), where $U(\lambda)$ is a gamma distribution function. See Thorin [45].

Example 7: Log-Normal Distribution.

$$dF(x)=\exp\left\{-\frac{1}{2}\left(\frac{\log x-\mu}{\sigma}\right)^2\right\}\Big/\{\surd(2\pi)\sigma x\}dx, \qquad x>0, \tag{15}$$

is infinitely divisible. See Thorin [46].

Example 8: Weibull Distribution.

$$F(x)=1-\exp\{-(\lambda x)^p\}, \qquad x,\lambda>0,\ 0<p\leqslant 1 \tag{16}$$

is infinitely divisible. See Bondesson [3].

From Example 5, it is easily seen that the gamma and exponential distributions are infinitely divisible. Other infinitely divisible distributions include the normal, Student's *t*, *F*, logistic, Laplace, and Cauchy distributions. See Steutel [41] for a discussion of these results.

THEOREM III.1.3. *If a random variable is nondegenerate but has bounded support, then it is not infinitely divisible.*

Proof. See Lukacs ([24], pp. 258–59).

This theorem implies that the binomial, beta, and uniform distributions are not infinitely divisible.

THEOREM III.1.4. *If X and Y are independent and infinitely divisible, then X + Y is infinitely divisible.*

Proof. In an obvious notation, the characteristic functions satisfy

$$\phi_{X+Y}^{1/n}(s) = \phi_X^{1/n}(s)\phi_Y^{1/n}(s),$$

and the result follows.

b) EXAMPLES OF MIXED DISTRIBUTIONS

A few examples of mixed number-of-claims distributions are considered here. For a thorough discussion of mixing, the reader is referred to Kupper [23], and Johnson and Kotz [21].

The following example is useful in analyzing a few models.

Example 9: Binomial Mixtures. Suppose the conditional probability-generating function is

$$P(s,\ t|\theta) = \{1 + p(t)(s-1)\}^{\theta}. \tag{17}$$

If the probability-generating function of θ may be written in the form (for some function *B*)

$$P_1(s|\mu) = B[\mu(s-1)], \tag{18}$$

then

$$\begin{aligned} P(s,\ t) &= E\{P(s,\ t|\theta)\} \\ &= P_1(1 + p(t)(s-1)|\mu) \\ &= B[\mu p(t)(s-1)] \\ &= P_1(s|\mu p(t)). \end{aligned} \tag{19}$$

Three distributions that satisfy (18) are the Poisson, negative binomial, and binomial.

Example 9a: The Binomial-Poisson Model. If

$$P(\theta = n) = \lambda^n e^{-\lambda}/n!, \qquad n = 0, 1, 2, \ldots,$$

then

$$p_n(t) = \frac{\{\lambda p(t)\}^n e^{-\lambda p(t)}}{n!}, \quad n = 0, 1, 2, \ldots. \tag{20}$$

Example 9b: The Binomial–Negative Binomial Model. If

$$P(\theta = n) = \binom{-\beta}{n}\left(\frac{1}{1+a}\right)^{\beta}\left(\frac{-a}{1+a}\right)^{n}, \qquad n = 0, 1, 2, \ldots \tag{21}$$

is the probability function of θ, then from (19) it follows that

$$p_n(t) = \binom{-\beta}{n}\left(\frac{1}{1+ap(t)}\right)^{\beta}\left(\frac{-ap(t)}{1+ap(t)}\right)^{n}, \qquad n = 0, 1, 2, \ldots. \tag{22}$$

Example 9c: The Binomial-Binomial Model. If

$$P(\theta = n) = \binom{M}{n} p^n (1-p)^{M-n}, \qquad n = 0, 1, 2, \ldots, M,$$

then the unconditional distribution of $N(t)$,

$$p_n(t) = \binom{M}{n}\{p[p(t)]\}^n[1 - pp(t)]^{M-n}, \qquad n = 0, 1, 2, \ldots M, \tag{23}$$

follows easily from (19).

Kupper [23] derived (21) for the geometric case, i.e., when $\beta = 1$.

Example 9 implies that mixing the parameter θ in the binomial over a distribution with probability-generating function of the form (18) results in the unconditional distribution being of the same form as the mixing distribution, but with different parameters. This generalizes results given by Kupper [23]. (By Corollary III.1.1, these are also compound distributions.)

Example 10: Negative Binomial Distribution. If $p_n(t|\theta)$ is given by (5) and $U(\theta)$ by II.6.12, the resulting unconditional distribution is a negative binomial one. See Bühlmann [6] for a proof of this. Since the gamma density is infinitely divisible, this provides an alternate proof of the fact that the negative binomial is a Compound Poisson distribution.

Example 11: The Generalized Waring Distribution. Seal [37] derives the generalized Waring distribution with probabilities

$$p_n(1)=\frac{\Gamma(p+k)\Gamma(p+h)\Gamma(h+n)\Gamma(k+n)}{\Gamma(p)\Gamma(n+1)\Gamma(h)\Gamma(k)\Gamma(p+h+k+n)},\qquad n=0,1,2,\ldots,\infty \tag{24}$$

by mixing a Poisson distribution twice. The probability-generating function $P(z)$ may thus be written

$$P(z)=\int_0^\infty\int_0^\infty e^{\lambda v(z-1)}\frac{h^h\lambda^{h-1}e^{-\lambda h}}{\Gamma(h)}\frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k)}h^{-k}v^{k-1}\times\left(1+\frac{v}{h}\right)^{-(p+k)}d\lambda dv$$

$$=\int_0^\infty\int_0^\infty e^{xy(z-1)}\frac{x^{h-1}e^{-x}}{\Gamma(h)}\frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k)}y^{k-1}(1+y)^{-(p+k)}dxdy.$$

Letting $s = xy$, $t = y$ gives

$$P(z)=\int_0^\infty e^{s(z-1)}\left\{\int_0^\infty\frac{(s/t)^{h-1}e^{-s/t}\Gamma(p+k)}{\Gamma(h)\ \Gamma(p)\Gamma(k)}t^{k-1}(1+t)^{-(p+k)}\frac{dt}{t}\right\}ds$$

$$=\int_0^\infty e^{s(z-1)}f(s)ds,$$

where

$$f(s)=\int_0^\infty\frac{(s/t)^{h-1}e^{-s/t}}{\Gamma(h)}\frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k)}t^{k-1}(1+t)^{-(p+k)}\frac{dt}{t}$$

can be shown to be the density of

$$X(1,\ h)X(1,\ k)/X(1,\ p)$$

in the notation of Example 5. Thus $f(s)$ is an infinitely divisible density by Example 5 and so by Example 1 with $t = 1$, the generalized Waring distribution is Compound Poisson. It can also be easily shown that this distribution can be derived by assuming that the negative binomial distribution has the transformed parameter $[1 + a(t)]^{-1}$ with a beta distribution. This result has also been discovered recently by Goovaerts and Van Wouwe [20].

Example 12: The Polya-Eggenberger Distribution. If

$$f_1(x|p)=\binom{n}{x}p^x(1-p)^{n-x},\qquad x=0,\ 1,\ 2,\ \ldots,\ n, \tag{25}$$

where

$$f_2(p)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}, \quad 0<p<1, \tag{26}$$

then

$$\begin{aligned} f(x) &= \int_0^1 f_1(x|p)f_2(p)dp \\ &= \frac{\Gamma(a+b)\Gamma(n+1)\Gamma(a+x)\Gamma(b+n-x)}{\Gamma(a)\Gamma(b)\Gamma(n-x+1)\Gamma(x+1)\Gamma(a+b+n)} \\ &= \binom{a+x-1}{x}\binom{b+n-x-1}{n-x}\Big/\binom{a+b+n-1}{n} \\ &= \binom{-a}{x}\binom{-b}{n-x}\Big/\binom{-a-b}{n}, \quad x=0, 1, 2, \ldots, n. \end{aligned} \tag{27}$$

This binomial-beta mixture is termed a Polya-Eggenberger distribution by Johnson and Kotz [21]. It is, in fact, a generalized hypergeometric distribution. By Theorem III.1.3, it is not infinitely divisible and hence not Compound Poisson.

Example 13: Neyman's Type A Distribution. If in the notation of Example 12

$$f_1(x|\lambda)=\frac{\lambda^x e^{-\lambda}}{x!}, \quad x=0, 1, 2, \ldots,$$

and

$$f_2(\lambda)=\frac{\alpha^\lambda e^{-\alpha}}{\lambda!},$$

then

$$\begin{aligned} f(x) &= \sum_{\lambda=1}^{\infty}\frac{\lambda^x e^{-\lambda}}{x!}\frac{\alpha^\lambda e^{-\alpha}}{\lambda!} \\ &= \frac{e^{-\alpha}}{x!}\sum_{\lambda=1}^{\infty}\frac{\lambda^x e^{-\lambda(1-\log\alpha)}}{\lambda!}, \quad x=1, 2, \ldots, \end{aligned} \tag{28}$$

and

$$\begin{aligned} f(0) &= e^{-\alpha}+\sum_{\lambda=1}^{\infty}e^{-\lambda}\frac{\alpha^\lambda e^{-\alpha}}{\lambda!} \\ &= e^{-\alpha}(e^{\alpha/e})=e^{-\alpha(1-e^{-1})}. \end{aligned}$$

The probability-generating function is

$$P(z) = \sum_{\lambda=0}^{\infty} e^{\lambda(z-1)} \frac{e^{-\alpha}\alpha^{\lambda}}{\lambda!} \tag{29}$$
$$= e^{\alpha(e^{z-1}-1)}.$$

This "mixed" distribution is easily seen from (29) or Corollary III.1.1 to be a Compound Poisson distribution with Poisson jump distribution. It was derived by Neyman [28] for use in accident statistics. See also Johnson and Kotz [21].

Example 14. If in the notation of Example 12

$$f_1(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

and

$$f_2(\lambda) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} e^{-\lambda a}(1-e^{-\lambda})^{b-1}, \qquad \lambda > 0, \tag{30}$$

then

$$\begin{aligned} f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{\infty} \frac{\lambda^x e^{-\lambda(a+1)}}{x!} (1-e^{-\lambda})^{b-1} d\lambda \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)x!} \int_0^{\infty} \lambda^x e^{-\lambda(a+1)} \left\{ \sum_{k=0}^{\infty} \binom{b-1}{k} (-1)^k e^{-\lambda k} \right\} d\lambda \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)x!} \left\{ \sum_{k=0}^{\infty} \binom{b-1}{k} (-1)^k \int_0^{\infty} \lambda^x e^{-\lambda(a+k+1)} d\lambda \right\} \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \binom{b-1}{k} (-1)^k (1+a+k)^{-x-1}, \qquad x = 0, 1, 2, \ldots . \end{aligned} \tag{31}$$

This is a finite sum if b is a positive integer. This distribution may be termed a Poisson-beta distribution since (30) is equivalent to $e^{-\lambda}$ having a beta distribution.

c) A LIMIT THEOREM

In general, it is difficult to choose a structure function to serve as a mixing distribution. In certain cases, information about the mixing distribution may be obtained, however. A theorem proved here provides one such case.

Definition 1.

$$\chi_t(s|\theta) = \sum_{n=0}^{\infty} e^{isn} p_n(t|\theta). \tag{32}$$

$$\chi_t(s) = \sum_{n=0}^{\infty} e^{isn} p_n(t). \tag{33}$$

The following theorem may now be proved (using the characteristic functions defined above) concerning the relationship between the distribution of the number of claims and the structure function.

THEOREM III.1.5. *If* $\chi_t(s|\theta)$ *satisfies*

$$\lim_{t\to\infty} \log \chi_t(s/t|\theta) = isk\theta, \tag{34}$$

where $k > 0$, *then* $N(t)/t$ *converges in distribution to* $k\theta$ *as* $t \longrightarrow \infty$.

Proof. The characteristic function of $N(t)/t$ is given by

$$\chi_t(s/t) = \int \chi_t(s/t|\theta) dU(\theta).$$

Thus,

$$\begin{aligned} \lim_{t\to\infty} \chi_t(s/t) &= \int \lim_{t\to\infty} \chi_t(s/t|\theta) dU(\theta) \\ &= \int e^{isk\theta} dU(\theta). \end{aligned} \tag{35}$$

The interchange of integration and limiting is justified by Lebesgue's dominated convergence theorem. The right-hand side of (35) is the characteristic function of $k\theta$. Hence, by the continuity theorem for characteristic functions, the result follows.

Remark 1. Equation (34) implies $E\{N(t)|\theta\} = k\theta t$. A sufficient condition for (34) to hold is that the first 2 moments of $N(t)|\theta$ exist and the cumulant $k_2(t|\theta)$ satisfies

$$\lim_{t\to\infty} \frac{k_2(t|\theta)}{t^2} = 0.$$

If this condition holds, then

$$\log \chi_t(s|\theta) = isk\theta t + \frac{k_2(t|\theta)}{2!}(is)^2 + O(|s|^2).$$

See Lukacs ([24], p. 26) for a derivation of this result.

This generalizes a theorem given by Bühlmann [6] and Lundberg [25], who derive a similar result for $X(t)/t$ in the mixed Poisson case (given below).

Example 15. For the mixed Poisson distribution (5),

$$\chi_t(s|\theta) = \exp\{\theta t(e^{is} - 1)\}.$$

Thus

$$\log \chi_t(s/t|\theta) = \theta t \sum_{n=1}^{\infty} (is/t)^n/n!$$

and

$$\lim_{t\to\infty} \chi_t(s/t|\theta) = is\theta.$$

In this case $k = 1$ and $N(t)/t$ converges to θ in distribution.

Example 16. For the negative binomial characteristic function given by

$$\chi_t(s|\theta) = \{1 - a(e^{is} - 1)\}^{-\theta t} \tag{36}$$

it is easily seen that

$$\lim_{t\to\infty} \chi_t(s/t|\theta) = \lim_{t\to\infty} \left\{1 - \frac{isa}{t} - a\sum_{j=2}^{\infty}\frac{(is/t)^j}{j!}\right\}^{-\theta t}$$
$$= \exp\{isa\theta\}.$$

Here $k = a$ and $N(t)/t$ converges to $a\theta$ in distribution.

Many variations on these results are possible. Lundberg [25] discusses many properties of the mixed Poisson distribution with generating function given by (6), whereas Kupper [23] derives many distributions by mixing.

III.2. *Point Processes*

In the previous subsection models were considered in which the set at risk was suspected to be nonhomogeneous. In this subsection and the remaining subsections of this section, the situation where the occurrence of a claim influences the probability distribution of the occurrence of other claims is considered. This dependency of sorts is known as "contagion." For a discussion of this phenomenon see Feller [15].

Methods used to handle this situation include urn schemes, in which different colored balls are placed in an urn. Balls are drawn from the urn one at a time, and, after each draw, a set of balls is added to the urn. This set is dependent upon the color of the ball drawn. For a thorough discussion of these models, see Kupper [23]. It should be noted that in many instances the asymptotic distribution of the number of balls drawn of a certain color can be shown to be of negative binomial or, more generally, of Compound Poisson form.

Other models will now be considered to handle contagion, beginning with the number of claims as a general point process. Included in this general class of models are renewal processes, alternating renewal processes, and semi-Markov or Markov renewal processes, among others.

Definition 1. Let $W_n = T_n - T_n - T_{n-1}$ be the nth claim interoccurrence time, $n = 1, 2, 3, \ldots$. Let the distribution and density functions of T_n be $S_n(t)$ and $s_n(t)$, respectively.

The distribution of the number of claims is given by the following relationship.

THEOREM III.2.1.

$$p_n(t) = S_n(t) - S_{n+1}(t). \tag{1}$$

Proof.

$$\begin{aligned} p_n(t) &= P\{N(t) < n+1\} - P\{N(t) < n\} \\ &= \{1 - S_{n+1}(t)\} - \{1 - S_n(t)\} \\ &= S_n(t) - S_{n+1}(t). \end{aligned}$$

Hence the probability of a claim occurring at any time may be dependent upon the number of claims that have occurred and the time since the last claim.

The mean and variance of $N(t)$ are given in the following corollary.

COROLLARY III.2.1.

$$E\{N(t)\} = \sum_{n=1}^{\infty} S_n(t). \tag{2}$$

$$V\{N(t)\} = 2\sum_{n=1}^{\infty} nS_n(t) - \sum_{n=1}^{\infty} S_n(t) - \left\{\sum_{n=1}^{\infty} S_n(t)\right\}^2. \tag{3}$$

Proof.

$$E\{N(t)\} = \sum_{n=0}^{\infty} n\{S_n(t) - S_{n+1}(t)\} = \sum_{n=1}^{\infty} S_n(t).$$

$$\begin{aligned} V\{N(t)\} &= E\{N(t)[N(t)+1]\} - E\{N(t)\}\{E(N(t))\}^2 \\ &= 2\sum_{n=1}^{\infty} nS_n(t) - \sum_{n=1}^{\infty} S_n(t) - \left\{\sum_{n=1}^{\infty} S_n(t)\right\}^2. \end{aligned}$$

If $E\{N(t)\}$ is proportional to t, then the process is called a stationary point process. See McFadden [27] for details.

a) THE RENEWAL PROCESS

When the claim interoccurrence times are independent and identically distributed, the point process becomes a renewal process. This model was first considered by Sparre-Andersen [38]. The probability of a claim occurring at any time in this case only depends upon the elapsed time since the last claim. since the last claim.

Definition 2. Let the common distribution function and density function of the W_i's be denoted by $S(x)$ and $s(x)$, respectively.

Then (1) easily becomes

$$p_n(t) = S^{n*}(t) - S^{(n+1)*}(t) \tag{4}$$

in terms of convolutions. For a detailed discussion of renewal processes, see, for example, Cox [8].

Example 1: Gamma Interoccurrence Times. If

$$s(x) = \frac{\lambda(\lambda x)^{r-1}e^{-\lambda x}}{(r-1)!}, \qquad x>0, \tag{5}$$

then

$$\begin{aligned} p_n(t) &= \int_0^t \frac{\lambda(\lambda x)^{nr-1}e^{-\lambda x}dx}{(nr-1)!} - \int_0^t \frac{\lambda(\lambda x)^{(n+1)r-1}e^{-\lambda x}dx}{[(n+1)r-1]!} \\ &= \sum_{j=nr}^{(n+1)r-1} \frac{(\lambda t)^j e^{-\lambda t}}{j!}, \end{aligned} \tag{6}$$

using (II.3.2). Tellenbach [44] considered this model. When $r = 1$, (5) becomes the exponential density and (6) becomes the simple Poisson. By the memoryless property of the exponential, it is easy to see that in this special instance the occurrence of a claim is independent of the time since the last claim, so that the Poisson process does not have contagious properties.

b) THE MARKOV RENEWAL PROCESS

An alternative model is to consider the number-of-claims process to be a Markov renewal or a semi-Markov process. For simplicity, it is assumed that there are two claim types and two corresponding claim interoccurrence densities. The following definitions are now made.

Definition 3. For $i, j = 1, 2$, let $f_i(t)$ be the density of the time until the next claim given that it will be of type i; p_i be the probability that a claim of type i will occur next, given that a claim of type i just occurred; $k_{ij}(t)$ be the density of the time until the first type j claim given that a claim of type i just occurred.

Definition 4. Let p be the probability that the first claim interoccurrence density is $f_1(t)$.

THEOREM III.2.2. *In an obvious notation, the Laplace transform of the density of the time until the occurrence of the* $(n+1)$*st type 1 claim is given by*

$$k_1^*(s)\{k_{11}^*(s)\}^n, \tag{7}$$

where

$$k_1^*(s)=f_1^*(s)\left\{p+(1-p)\frac{(1-p_2)f_2^*(s)}{1-p_2f_2^*(s)}\right\} \tag{8}$$

and

$$k_{11}^*(s)=f_1^*(s)\left\{p_1+(1-p_1)\frac{(1-p_2)f_2^*(s)}{1-p_2f_2^*(s)}\right\}. \tag{9}$$

Proof. It is easily seen from probabilistic arguments that

$$k_{11}(t)=p_1f_1(t)+(1-p_1)\int_0^t f_2(\tau)k_{21}(t-\tau)d\tau.$$

Thus,

$$k_{11}^*(s)=p_1f_1^*(s)+(1-p_1)f_2^*(s)k_{21}^*(s).$$

Similarly,

$$k_{21}(t)=(1-p_2)f_1(t)+p_2\int_0^t f_2(\tau)k_{21}(t-\tau)d\tau,$$

or

$$k_{21}^*(s)=(1-p_2)f_1^*(s)+p_2f_2^*(s)k_{21}^*(s),$$

which implies

$$k_{21}^*(s)=\frac{(1-p_2)f_1^*(s)}{1-p_2f_2^*(s)}, \tag{10}$$

and (9) follows easily. By a similar argument (8) is derived. Since $k_1(t)$ is the time until the first type 1 claim and $k_{11}(t)$ for each successive type 1 claim, expression (7) follows.

Remark 1. The probability distribution for the number of type 1 claims can be obtained using Theorem III.2.2 and Theorem III.2.1. To obtain the distribution of type 2 claims, simply replace p by $(1 - p)$ and interchange the subscripts 1 and 2 in Theorem III.2.2.

Example 2: Exponential Interoccurrence Times. Let

$$f_i(t) = \lambda_i e^{-\lambda_i t}, \qquad t>0. \tag{11}$$

Then (9) becomes

$$k_{11}^*(s) = \frac{\lambda_1}{\lambda_1 + s}\left\{p_1 + (1-p_1)\frac{\lambda_2(1-p_2)}{\lambda_2(1-p_2)+s}\right\}, \tag{12}$$

from which it follows that (7) is

$$\left\{\frac{\lambda_1}{\lambda_1+s}\right\}^{n+1}\left\{p+(1-p)\frac{\lambda_2(1-p_2)}{s+\lambda_2(1-p_2)}\right\}\left\{p_1+(1-p_1)\frac{\lambda_2(1-p_2)}{s+\lambda_2(1-p_2)}\right\}^n. \tag{13}$$

This last expression is recognized as the convolution of a gamma density, a compound Bernoulli-exponential, and a compound binomial-exponential (see Example II.3.2). Alternatively, note that it may be rewritten as a linear combination of convolutions of gamma densities. Formally (writing $\lambda_2^* = \lambda_2(1-p_2)$), (13) becomes

$$\sum_{k=0}^{n}\binom{n}{k}(1-p_1)^k p_1^{n-k}\left\{p\left(\frac{\lambda_1}{\lambda_1+s}\right)^{n+1}\left(\frac{\lambda_2^*}{\lambda_2^*+s}\right)^k + (1-p)\left(\frac{\lambda_1}{\lambda_1+s}\right)^{n+1}\left(\frac{\lambda_2^*}{\lambda_2^*+s}\right)^{k+1}\right\}, \tag{14}$$

which upon inversion yields the following density of the time until the $(n + 1)$st claim:

$$\sum_{k=0}^{n}\binom{n}{k}(1-p_1)^k p_1^{n-k}\left\{\int_0^t \frac{\lambda_1[\lambda_1(t-x)]^n e^{-\lambda_1(t-x)}}{n!} \times\left[p\frac{\lambda_2^*(\lambda_2^* x)^{k-1}e^{-\lambda_2^* x}}{(k-1)!} + (1-p)\frac{\lambda_2^*(\lambda_2^* x)^k e^{-\lambda_2^* x}}{k!}\right]dx\right\}.$$

From this may be obtained the corresponding distribution function and hence $p_n(t)$ by Theorem III.2.1.

Example 3: The Renewal Process. Putting p, $p_1 = 1$ in Theorem III.2.2 yields $k_{11}^*(s) = k_1^*(s) = f_1^*(s)$, and the ordinary renewal process follows.

Example 4: The Alternating Renewal Process. If $p_1 = p_2 = 0$ in Theorem III.2.2, then claim types alternate. This model has a direct interpretation in terms of disability insurance. The type 1 density could be viewed as the time until a claim occurs and the type 2 as the time until recovery. Hence the number of claims is given by the number of type 1 claims.

From Theorem III.2.2, the transform of the density of the time until occurrence of the $(n + 1)$st type 1 claim is seen to be

$$\{f_1^*(s)f_2^*(s)\}^n\{pf_1^*(s)+(1-p)f_1^*(s)f_2^*(s)\}. \tag{15}$$

Seal [36] assumes the healthy period to be of exponential form. If it is assumed that

$$f_1(t)=\lambda e^{-\lambda t}$$

and

$$f_2(t)=\frac{\lambda(\lambda t)^{r-1}e^{-\lambda t}}{(r-1)!},$$

then (15) becomes

$$p\left(\frac{\lambda}{\lambda+s}\right)^{n(r+1)+1}+(1-p)\left(\frac{\lambda}{\lambda+s}\right)^{(n+1)(r+1)},$$

which yields upon inversion

$$p\frac{\lambda(\lambda t)^{n(r+1)}e^{-\lambda t}}{[n(r+1)]!}+(1-p)\frac{\lambda(\lambda t)^{nr+n+r}e^{-\lambda t}}{(nr+n+r)!}, \tag{16}$$

i.e., a weighted average of 2 gamma densities. From Theorem III.2.1, it is easily verified that

$$\begin{aligned} p_n(t)&=p\sum_{j=nr+n-r}^{n(r+1)}\frac{(\lambda t)^je^{-\lambda t}}{j!}+(1-p)\sum_{j=n(r+1)}^{nr+n+r}\frac{(\lambda t)^je^{-\lambda t}}{j!}, \qquad n>0;\\ p_0(t)&=pe^{-\lambda t}+(1-p)\sum_{j=0}^{r}\frac{(\lambda t)^je^{-\lambda t}}{j!}. \end{aligned} \tag{17}$$

The generalization to more than 2 types of claim densities is obvious, but the algebra quickly becomes cumbersome. For a further discussion of semi-Markov processes, see, for example, Cox and Miller [9] or Ross [35] for details.

III.3. *Nonhomogeneous Birth Processes*

An alternative model is to consider the number-of-claims process to be a nonhomogeneous birth process. The probability of a claim occurring at any time is dependent on that time and the total number of claims up to that time. The following definition is due to Parzen [34].

Definition 1. An integer-valued stochastic process $\{N(t);\, t \geqslant 0\}$ is said to be a nonhomogeneous pure birth process with transition probabilities

$$p_{n,m}(t,\, s) = P\{N(s) = m | N(t) = n\}, \tag{1}$$

which satisfy

$$\left.\begin{aligned} \lim_{h \to 0} \frac{p_{n,n+1}(t,\, t+h)}{h} &= c_n(t) \\ \lim_{h \to 0} \frac{1 - p_{n,n}(t,\, t+h)}{h} &= c_n(t) \end{aligned}\right\}, \quad n = 0,\, 1,\, 2,\, \ldots, \tag{2}$$

where $c_n(t) \geqslant 0$, $n = 0, 1, 2, \ldots$.

The functions $c_n(t)$ are referred to as intensity functions. The following lemma is instrumental in the solution of the transition probabilities (1).

LEMMA III.3.1. *The transition probabilities satisfy*

$$\begin{aligned} \frac{\partial}{\partial s} p_{n,n}(t,\, s) &= -c_n(s) p_{n,n}(t,\, s), \\ \frac{\partial}{\partial s} p_{n,m}(t,\, s) &= -c_m(s) p_{n,m}(t,\, s) + c_{m-1}(s) p_{n,m-1}(t,\, s), \qquad m > n, \end{aligned} \tag{3}$$

for all $n = 0, 1, 2, \ldots$ *and all* $s \geqslant t \geqslant 0$.

Proof. See Parzen [34].

Explicit solutions are obtainable from (3) which are of a recursive nature.

THEOREM III.3.1. *The solution to the system (3) is given by*

$$\begin{aligned} p_{n,n}(t,\, s) &= \exp\left\{ -\int_t^s c_n(\tau)\, d\tau \right\}, \\ p_{n,m}(t,\, s) &= \left[\int_t^s c_{m-1}(\tau) p_{n,m-1}(t,\, \tau) \exp\left\{ \int_t^\tau c_m(\mu)\, d\mu \right\} d\tau \right] \\ &\quad \times \exp\left\{ -\int_t^s c_m(\tau)\, d\tau \right\}, \qquad m > n. \end{aligned} \tag{4}$$

Proof. Substitution of (4) in system (3) yields the desired result.

It appears that no simpler solution than (4) is obtainable unless further assumptions about the intensities of frequency $c_n(t)$ are made.

Assumption 1. Let the claim intensities of frequencies admit the factorization

$$c_n(t) = \lambda_n \beta(t), \tag{5}$$

where

$$\beta(t) > 0, \qquad t > 0.$$

For a discussion of this assumption, see Bühlmann [6].
Definition 2.

$$p^*_{n,m}(\tau_1, \tau_2) = p_{n,m}(t, s), \tag{6}$$

where

$$\tau_1 = \int_0^t \beta(\mu)d\mu, \qquad \tau_2 = \int_0^s \beta(\mu)d\mu. \tag{7}$$

This change to "operational time" allows for a simplification of the system of equations (3), given in the following lemma.

LEMMA III.3.2. *The functions* $p^*_{n,m}(\tau_1, \tau_2)$ *satisfy*

$$\begin{aligned} &\frac{\partial}{\partial \tau_2} p^*_{n,n}(\tau_1, \tau_2) = -\lambda_n p^*_{n,n}(\tau_1, \tau_2), \\ &\frac{\partial}{\partial \tau_2} p^*_{n,m}(\tau_1, \tau_2) = -\lambda_m p^*_{n,m}(\tau_1, \tau_2) + \lambda_{m-1} p^*_{n,m-1}(\tau_1, \tau_2), \qquad m > n. \end{aligned} \tag{8}$$

Proof. It follows from Definition 2 that

$$\frac{\partial}{\partial s} p_{n,m}(t, s) = \frac{\partial}{\partial \tau_2} p^*_{n,m}(\tau_1, \tau_2)\beta(s), \tag{9}$$

and the result follows upon substitution of (5) and (6) in system (3).

The introduction of operational time allows for the elimination of the function $\beta(t)$ from the system (3). Clearly, $\lambda_n > 0$ for all n, and the following theorem gives some idea of the magnitude of λ_n.

THEOREM III.3.2. *The process* $\{N(t);\ t \geq 0\}$ *with transition probabilities* $p^*_{n,m}(\tau_1, \tau_2) = p_{n,m}(t, s)$ *is dishonest iff*

$$\sum_{n=1}^{\infty} \lambda_n^{-1} < \infty,$$

that is,

$$\sum_{m=n}^{\infty} p^*_{n,m}(\tau_1, \tau_2) < 1$$

if and only if

$$\sum_{n=1}^{\infty} \lambda_n^{-1} < \infty.$$

Proof. See Feller [16], pp. 452–53).

In light of Theorem III.3.2, the following assumption is made.

Assumption 2.

$$c_n(t) = (a + bn)\beta(t), \qquad n = 0, 1, 2, \ldots, N, \tag{10}$$

where N could be infinite.

Explicit solutions for the transition probabilities are obtainable under this assumption. To obtain them, the following probability-generating function is defined.

Definition 3.

$$P_n(z, \tau_1, \tau_2) = \sum_{k=0}^{\infty} p^*_{n,n+k}(\tau_1, \tau_2) z^k. \tag{11}$$

The solution to (8) under the assumption (10) is summarized by the following lemma.

LEMMA III.3.3. *The probability-generating function (11) satisfies the partial differential equation*

$$\frac{\partial}{\partial \tau_2} P_n(z,\tau_1,\tau_2) + z(1-z)b\frac{\partial}{\partial z} P_n(z,\tau_1,\tau_2) = (z-1)(a+bn)P_n(z,\tau_1,\tau_2) \tag{12}$$

subject to the initial condition

$$P_n(z, \tau_1, \tau_1) = 1. \tag{13}$$

Proof. The multiplication of (8) by z^{m-n} and summation over all $m \geqslant n$ yield equation (12). Since $p_0(\tau_1,\tau_1) = 1$, equation (13) follows.

Clearly $a > 0$ since $\lambda_0 > 0$. The solution to (12) depends upon the value of b. These solutions are summarized in the next few theorems.

THEOREM III.3.3. *If $b = 0$, the solution to (3) under the assumption (10) is the set of nonhomogeneous Poisson probabilities*

$$p_{n,n+k}(t, s) = \frac{1}{k!}\left[a\int_t^s \beta(\mu)d\mu\right]^k \exp\left[-a\int_t^s \beta(\mu)d\mu\right], \tag{14}$$
$$k = 0, 1, 2, \ldots.$$

Proof. If $b = 0$, then the subsidiary equations from (12) are

$$\frac{d\tau_2}{1}=\frac{dz}{0}=\frac{dP_n}{a(z-1)P_n}. \tag{15}$$

Thus,

$$z=c_1$$

and

$$P_n=c_2e^{a\tau_2(c_1-1)}$$

or

$$P_n(z,\tau_1,\tau_2)=e^{a\tau_2(z-1)}\phi(z).$$

Condition (13) implies

$$\phi(z)=e^{-a\tau_1(z-1)},$$

and so

$$P_n(z,\tau_1,\tau_2)=e^{a(\tau_2-\tau_1)(z-1)}. \tag{16}$$

Since $b = 0$, (14) is independent of n. If $b \neq 0$, there is a unique solution to (12) and (13).

LEMMA III.3.4. *If $b \neq 0$, the solution to (12) and (13) is*

$$P_n(z,\tau_1,\tau_2)=\left\{\frac{e^{-b(\tau_2-\tau_1)}}{1-z(1-e^{-b(\tau_2-\tau_1)})}\right\}^{a/b+n}. \tag{17}$$

Proof. The subsidiary equations to (12) are

$$\frac{d\tau_2}{1}=\frac{dz}{z(1-z)b}=\frac{dP_n}{(z-1)(a+bn)P_n}. \tag{18}$$

Thus,

$$b\tau_2=\ln\left(\frac{z}{1-z}\right)+k,$$

or

$$\left(\frac{z}{1-z}\right)e^{-b\tau_2}=c_1.$$

Also,

$$-\left(\frac{a}{b}+n\right)\frac{dz}{z}=\frac{dP_n}{P_n},$$

or

$$P_n z^{a/b+n}=c_2.$$

The general solution is

$$P_n(z,\tau_1,\tau_2)=z^{-a/b-n}\phi\left\{\frac{z}{1-z}e^{-b\tau_2}\right\}.$$

Condition (13) implies, for fixed τ_1,

$$\phi(w)=\left\{\frac{we^{b\tau_1}}{1+we^{b\tau_1}}\right\}^{a/b+n}.$$

Thus,

$$P_n(z,\tau_1,\tau_2)=z^{-a/b-n}\left\{\frac{ze^{-b(\tau_2-\tau_1)}}{1-z+ze^{-b(\tau_2-\tau_1)}}\right\}^{a/b+n}$$

$$=\left\{\frac{e^{-b(\tau_2-\tau_1)}}{1-z[1-e^{-b(\tau_2-\tau_1)}]}\right\}^{a/b+n}.$$

Two different sets of probabilities are obtained depending upon whether b is greater or less than 0. If $b>0$, the probabilities are given as follows.

THEOREM III.3.4. *If $b>0$, the solution to (3) under the assumption (10) is the set of negative binomial transition probabilities*

$$p_{n,n+k}(t,s)=\binom{a/b+n+1}{k}\left\{\exp\left[-b\int_t^s\beta(r)dr\right]\right\}^{a/b+n}$$
$$\times\left\{1-\exp\left[-b\int_t^s\beta(r)dr\right]\right\}^k,\qquad k=0,1,2,\ldots. \tag{19}$$

Proof. If $b>0$, (17) is simply a negative binomial probability-generating function, which yields probabilities (19) after substituting back from Definition 2.

Finally, if $b<0$, the probabilities are as follows.

THEOREM III.3.5. *If $b<0$, and $-a/b$ a positive integer N, the solution to (3) under the assumption (10) is the set of binomial probabilities*

$$p_{n,n+k}(t, s) = \binom{N-n}{k}\left\{1 - \exp\left[b\int_t^s \beta(r)dr\right]\right\}^k$$
$$\times\left\{\exp\left[b\int_t^s \beta(r)dr\right]\right\}^{N-n-k}, \quad k = 0, 1, 2, \ldots, N-n, \tag{20}$$

for all $n \leq N$.

Proof. From (17),

$$P_n(z, \tau_1, \tau_2) = \{e^{b(\tau_2 - \tau_1)} + z(1 - e^{b(\tau_2 - \tau_1)})\}^{N-n},$$

which is a binomial generating function for $n \leq N$.

If $-a/b$ is not a positive integer, then the probabilities are given by a generalized binomial. See Bühlmann [6] for details of these "contagion" models.

Thus the Poisson, negative binomial, and binomial distributions arise from the nonhomogeneous birth process contagion models.

Many counting processes can be modeled as a nonhomogeneous birth process with intensities of frequency of the form (10). These include the Poisson process, the nonhomogeneous Poisson process, the Polya process, and the Yule process. For a definition and discussion of these processes, see Parzen [34], Feller [16], or Lundberg [25].

III.4. *A Branching Process Model*

The number of claims can be considered to have arisen from a branching process in certain instances. Each claim is directly responsible for the occurrence of a random number of claims, each of which in turn causes a random number of claims in accordance with the same probability distribution. This model might perhaps be useful in a situation where there is an infectious disease epidemic.

Definition 1. A single claim gives rise to n further claims directly (not including the original) with probability

$$q_n, \quad n = 0, 1, 2, \ldots, \tag{1}$$

where $q_0 > 0$,

$$Q(s) = \sum_{n=0}^{\infty} q_n s^n, \tag{2}$$

and $0 < Q'(1) < 1$.

It is assumed that claims occur initially in accordance with the distribution (1), each of which in turn causes n more in accordance with (1).

Definition 2. Let the probability-generating function of the total number of claims be given by

$$P(s) = \sum_{n=0}^{\infty} p_n s^n. \tag{3}$$

The relationship between $P(s)$ and $Q(s)$ is given by the following theorem.

THEOREM III.4.1.

$$P(s) = Q(sP(s)) \tag{4}$$

with $P(1) = 1$.

Proof. Feller ([16], p. 298) shows that the probability-generating function $P_1(s)$ of the total number of claims arising from 1 individual claim (including this claim) satisfies

$$P_1(s) = sQ(P_1(s)).$$

Thus if the number of these original claims has distribution (1), it follows that

$$P(s) = Q(P_1(s)) = \frac{P_1(s)}{s}$$

or

$$sP(s) = P_1(s).$$

This implies $sP(s) = sQ(sP(s))$. Feller ([16], p. 298) also shows that the distribution $\{p_n\}$ is not defective (i.e., $P(1) = 1$) when $Q'(1) < 1$.

The probability distribution $\{p_n\}$ has moments easily obtainable from (4). In particular, it is easily verified that the mean of the distribution may be expressed as

$$P'(1) = \frac{Q'(1)}{1 - Q'(1)}. \tag{5}$$

Further, the form of the distribution $P(1)$ is given in the following theorem.

THEOREM III.4.2. *The probability-generating function $P(s)$ given by (4) is Compound Poisson with generating function*

$$P(s) = e^{\lambda[A(s)-1]}, \tag{6}$$

where

$$\lambda = -\log q_0 \tag{7}$$

and

$$A(s) = \sum_{n=1}^{\infty} \frac{q_n^{n*}}{-n \log q_0} s^n \tag{8}$$

is a probability-generating function.

Proof: See Dwass [13].

Thus the solution to equation (4) is an infinitely divisible probability-generating function even if $Q(s)$ is not. Once again, the importance of the Compound Poisson distribution is highlighted. Here, the explicit form of the distribution is obtainable. The probabilities p_n could be calculated using the results of the above theorem, but a simpler form is given in the following theorem.

THEOREM III.4.3.

$$p_n = \frac{1}{n+1} q_n^{(n+1)*}, \qquad n = 0, 1, 2, \ldots . \tag{9}$$

Proof. Multiplying (4) by s and using Lagrange's expansion gives each side of (9) as the corresponding coefficient of s^{n+1}. See Steutel ([39], p. 13) for a discussion of this.

Thus the probabilities from this contagion model are easily calculated using the above theorem.

Example 1: $Q(s) = \exp\{\lambda(s - 1)\}$. In this Poisson case,

$$p_n = \frac{1}{n+1} \frac{[\lambda(n+1)]^n e^{-\lambda(n+1)}}{n!}, \tag{10}$$

$$P'(1) = \frac{\lambda}{1-\lambda},$$

and

$$P(s) = \exp\left\{\lambda\left[\sum_{n=1}^{\infty} \frac{(\lambda n)^n e^{-\lambda n}}{\lambda n(n!)} s^n - 1\right]\right\}. \tag{11}$$

Example 2: $Q(s) = [1 - a(s - 1)]^{-\alpha}$. For these negative binomial probabilities,

$$p_n = \frac{1}{n+1} \binom{\alpha(n+1)+n-1}{n} \left(\frac{1}{1+a}\right)^{\alpha(n+1)} \left(\frac{a}{1+a}\right)^n, \tag{12}$$

$$n = 0, 1, 2, \ldots .$$

If $\alpha = 1$, the geometric probabilities result in

$$p_n = \frac{1}{n+1}\binom{2n}{n}\left(\frac{1}{1+a}\right)^{n+1}\left(\frac{a}{1+a}\right)^n, \qquad n=0,\ 1,\ 2,\ \ldots . \tag{13}$$

Example 3: $Q(s) = (ps + q)^m$. The binomial probabilities lead to

$$p_n = \frac{1}{n+1}\binom{m(n+1)}{n}p^n q^{m(n+1)-n}, \qquad n=0,\ 1,\ 2,\ \ldots . \tag{14}$$

For the Bernoulli case when $m = 1$,

$$p_n = qp^n; \qquad n=0,\ 1,\ 2,\ \ldots , \tag{15}$$

which is a geometric distribution.

III.5. *Multivariate Distributions*

Many of the models considered assume that the individuals at risk have independent claim distributions (of each other). One possible method to account for a dependency of sorts is to assume that the number of claims from each individual at risk follows a multivariate distribution.

Definition 1. The k individual risks have a multivariate distribution with generating function

$$P_k(s_1,\ s_2,\ \ldots ,\ s_k). \tag{1}$$

The probability-generating function of the sum of the k risks is thus

$$P(s) = P_k(s,\ s,\ \ldots ,\ s). \tag{2}$$

Example 1: The Multivariate Poisson Distribution. This distribution, as defined by Teicher [43], has probability-generating function

$$P_k(s_1,\ s_2,\ \ldots ,\ s_k) = \exp\left\{\lambda\left[\sum_i \lambda_i(s_i - 1) + \sum_{i<j}\lambda_{ij}(s_is_j - 1) + \sum_{i<j<k}\lambda_{ijk}(s_is_js_k - 1) + \ldots + \lambda_{1,2,\ldots,k}(s_1 \ldots s_k - 1)\right]\right\}, \tag{3}$$

where the λ's are all positive. The marginal distribution of a subset of size r is of the same form as (3). Each individual risk has a marginal Poisson distribution. The correlation of any two risks is positive so that this distribution gives rise to a positive contagion model.

If μ_i = "the sum of all λ coefficients with i subscripted digits," for $i = 1, 2, \ldots, k$, then the generating function of the sum is

$$\begin{aligned} P(s) &= P_k(s, s, \ldots, s) \\ &= \exp\left\{\lambda \sum_{i=1}^{k} \mu_i(s^i - 1)\right. \\ &= \exp\{\lambda\mu[Q(s) - 1]\}, \end{aligned} \tag{4}$$

where

$$Q(s) = \sum_{i=1}^{k} \frac{\mu_i}{\mu} s^i \tag{5}$$

and

$$\mu = \sum_{i=1}^{k} \mu_i. \tag{6}$$

Thus $Q(s)$ is a probability-generating function and $P(s)$ a Compound Poisson.

Example 2: The Multivariate Negative Binomial. If λ in (3) has a gamma density

$$g(\lambda) = \frac{\alpha(\alpha\lambda)^{n-1}e^{-\alpha\lambda}}{\Gamma(n)}, \qquad \lambda > 0, \tag{7}$$

then unconditionally, (3) becomes

$$\begin{aligned} P_k(s_1, s_2, \ldots, s_k) = \left\{1 - \frac{1}{\alpha}\left[\sum_i \lambda_i(s_i - 1)\right.\right. \\ \left.\left. + \ldots + \lambda_{1,2,\ldots,k}(s_1 s_2 \ldots s_k - 1)\right]\right\}^{-n}, \end{aligned} \tag{8}$$

which is a multivariate negative binomial. Again, the marginals are negative binomial and the generating function of the sum is (in the notation of Example 1)

$$P(s) = \left\{1 - \frac{\mu}{\alpha}[Q(s) - 1]\right\}^{-n}, \tag{9}$$

which is a Compound negative binomial and hence a Compound Poisson. As before, the correlations between the individual risks are positive.

Another interpretation of the individuals at risk is that they represent the number of claims in nonoverlapping time intervals, which are thus correlated in some way. Edwards and Gurland [14] applied the model (8) when $k = 2$ and found a better fit to some accident data than a univariate negative binomial.

Other multivariate discrete distributions may be found in Johnson and Kotz [21] or in Mardia [26].

III.6. *Catastrophe Models*

Another type of contagious process results from the occurrence of natural disasters. If an event such as a flood or an earthquake were to occur, the number of property or life insurance claims resulting from this single accident could quite easily be large. Thus, for each disaster that occurs, associated with it is a random number of claims resulting from it.

Definition 1. Let the stochastic process $\{K(t);\ t \geq 0\}$ represent the number of "claim causing" phenomena in the interval $(0, t]$ for $t > 0$, and $K(0) = 0$.

If the claims resulting from each single disaster are independent and identically distributed by number, it is clear that $\{N(t); t \geq 0\}$ is a compound counting process. (See [II.2] for a discussion of compound distributions.)

In particular, if the disasters occur randomly in time, it is quite likely that $\{K(t);\ t \geq 0\}$ is a (possibly nonhomogeneous) Poisson process. Thus the distribution of $N(t)$ would be Compound Poisson in this instance.

Furthermore, if $\{K(t); t \geq 0\}$ has stationary and independent increments, then from (II.4) both $K(t)$ and $N(t)$ are discrete Compound Poisson distributions.

It follows from these assumptions that $X(t)$ is also Compound Poisson (see Theorem II.4.6).

A likely candidate for the distribution of $N(t)$ therefore is any discrete distribution on the nonnegative integers that is infinitely divisible (or equivalently Compound Poisson). For examples of such distributions, see (II.4) or (III.1).

From (II.2) it is seen that the usual representation of the probability-generating function of a compound distribution is

$$P(z) = P_1(G(z)),$$

where $P_1(z)$ and $G(z)$ are also probability-generating functions.

Simple methods of computations of the probabilities are given by Panjer [31].

Alternatively, the probabilities may be expressed as a finite sum as a result of the following theorem, which generalizes a result given by Ord ([29], p. 129).

THEOREM III.6.1. *Suppose* $P(z)$ *and* $G(z) = \sum_{k=0}^{\infty} g_k z^k$ *are probability-generating functions and a parameterization exists such that*

$$P(z) = \sum_{r=0}^{\infty} p_r z^r = \sum_{n=0}^{\infty} f_n(\lambda)[G(z)]^n \tag{1}$$

may be written in the form

$$P(z) = B\{\lambda[G(z) - 1]\} \tag{2}$$

for some function B, then

$$p_r = \sum_{n=0}^{r} f_n(\lambda_1) h_{r-n}^{n*}, \tag{3}$$

where

$$\lambda_1 = \lambda(1 - g_0), \qquad h_k = g_{k+1}(1 - g_0)^{-1}.$$

Proof. From (2) it is seen that

$$P(z) = B\{\lambda_1[zH(z) - 1]\}, \tag{4}$$

where

$$H(z) = \frac{G(z) - g_0}{z(1 - g_0)} = \sum_{n=0}^{\infty} h_n z^n \tag{5}$$

is a probability-generating function.

Thus from (4), (2), and (1)

$$\begin{aligned} P(z) &= \sum_{n=0}^{\infty} f_n(\lambda_1)[zH(z)]^n \\ &= \sum_{n=0}^{\infty} f_n(\lambda_1)\left(\sum_{k=0}^{\infty} h_k^{n*} z^{n+k}\right) \\ &= \sum_{n=0}^{\infty} \sum_{r=n}^{\infty} f_n(\lambda_1) h_{r-n}^{n*} z^r \\ &= \sum_{r=0}^{\infty} \left[\sum_{n=0}^{r} f_n(\lambda_1) h_{r-n}^{n*}\right] z^r. \end{aligned} \tag{6}$$

From (5),

$$H(z) = \sum_{n=1}^{\infty} \frac{g_n z^{n-1}}{1 - g_0},$$

and so

$$h_k = g_{k+1}(1 - g_0)^{-1}. \tag{7}$$

Equating $P(z)$ in (1) and (6) and comparing the coefficients of z^r yields the result.

A comparison of (2) with (1.18) implies that three distributions that satisfy (2) are the Compound Poisson, negative binomial, and binomial. Hence any of these three compound distributions have probabilities that may be expressed as a finite sum in the form given by (3). (See also Example 1.9.)

Any discrete compound distribution could thus serve as a distribution for the number of claims in a catastrophe situation, although in most instances a Compound Poisson model seems an appropriate choice for physical reasons. For a description of the many possibilities, see Kupper [23].

REFERENCES

1. BEEKMAN, J. *Two Stochastic Processes*. New York: Halsted Press, 1974.
2. BONDESSON, L. "On the Infinite Divisibility of Products of Powers of Gamma Variables," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, XLIX (1979), 171–75.
3. ———. "A General Result on Infinite Divisibility," *Annals of Probability*, VII (1979), 965–79.
4. BOWERS, N., GERBER, H., HICKMAN, J., JONES, D., and NESBITT, C. Part 5A Study Note on Risk Theory. Chicago: Society of Actuaries, 1982.
5. BÜHLMANN, H. "Note on the Collective Theory of Risk," *Skandinavisk Aktuarietidskrift*, LI (1968), 174–77.
6. ———. *Mathematical Methods in Risk Theory*. New York: Springer-Verlag, 1970.
7. BÜHLMANN, H., and BUZZI, R. "On a Transformation of the Weighted Compound Poisson Process," *ASTIN Bulletin*, VI (1971), 42–46.
8. COX, D. *Renewal Theory*. London: Methuen, 1962.
9. COX, D., and MILLER, H. *The Theory of Stochastic Processes*. London: Methuen, 1965.
10. CRAMÉR, H. *On the Mathematical Theory of Risk*. Skandinavian Jubilee Volume. Stockholm, 1930.
11. ———. *Collective Risk Theory*. Skandinavian Jubilee Volume. Stockholm, 1955.

12. DELBAEN, F., and HAEZENDONCK, J. *A Generalization of a Result of H. Bühlmann on Compound Processes* (to be published, 1983).
13. DWASS, M. "A Theorem about Infinitely Divisible Distributions," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, IX (1968), 287–89.
14. EDWARDS, C., and GURLAND, J. "A Class of Distributions Applicable to Accidents." *Journal of the American Statistical Association*, LVI (1961), 503–17.
15. FELLER, W. "On a General Class of Contagious Distributions," *Annals of Mathematical Statistics*, XIV (1943), 389–400.
16. ———. *An Introduction to Probability Theory and Its Applications*, Vol. I. 3d ed. New York: John Wiley, 1968.
17. ———. *An Introduction to Probability Theory and Its Applications*, Vol. II. 2d ed. New York: John Wiley, 1971.
18. GERBER, H. *An Introduction to Mathematical Risk Theory*, Homewood, Ill.: Irwin, 1979.
19. ———. "On the Numerical Evaluation of the Distribution of Aggregate Claims and Its Stop-Loss Premiums," *Insurance: Mathematics and Economics*, I, No. 1 (1982), 13–18.
20. GOOVAERTS, U., and VAN WOUWE, M. "The Generalized Waring Distribution as a Mixed Poisson with a Generalized Gamma Mixing Distribution," *Bulletin KVBA* (1981), 95–98.
21. JOHNSON, N. and KOTZ, S. *Discrete Distributions*. Boston: Houghton Mifflin, 1969.
22. KATTI, S. "Infinite Divisibility of Integer Valued Random Variables," *Annals of Mathematical Statistics*, XXXVIII (1967), 1306–8.
23. KUPPER, J. "Wahrscheinlichkeitstheoretische Modelle in der Schadenversicherung," *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, V (1960), 451–503; VI (1962), 95–130.
24. LUKACS, E. *Characteristic Functions*. 2d ed. London: Charles Griffin, 1970.
25. LUNDBERG, O. *On Random Processes and Their Application to Sickness and Accident Statistics*. Uppsala: Almqvist & Wiksells, 1940.
26. MARDIA, K. *Families of Bivariate Distributions*. London: Charles Griffin, 1970.
27. MCFADDEN, J. "On the Lengths of Intervals in a Stationary Point Process," *Journal of the Royal Statistical Society, B*, XXIV (1962), 364–82.
28. NEYMAN, J. "On a New Class of 'Contagious' Distributions, Applicable in Entomology and Bacteriology," *Annals of Mathematical Statistics*, X (1939), 35–57.
29. ORD, J. *Families of Frequency Distributions*. London: Charles Griffin, 1972.
30. PANJER, H. "The Aggregate Claims Distribution and Stop-Loss Reinsurance," *TSA*, XXXII (1980), 523–35.
31. ———. "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin*, XII (1981), 22–26.
32. PANJER, H., and WILLMOT, G. "Finite Sum Evaluation of the Negative Binomial-Exponential Model," *ASTIN Bulletin*, XII (1981), 133–37.

33. ———. "Recursions for Compound Distributions," *ASTIN Bulletin*, XIII (1982), 1–11.
34. PARZEN, E. *Stochastic Processes*. San Francisco: Holden-Day, 1962.
35. ROSS, S. *Applied Probability Models with Optimization Applications*. San Francisco: Holden-Day, 1962.
36. SEAL, H. *Stochastic Theory of a Risk Business*. New York: John Wiley, 1969.
37. ———. *Survival Probabilities: The Goal of Risk Theory*. New York: John Wiley, 1978.
38. SPARRE-ANDERSEN, E. "On the Collective Theory of Risk in the Case of Contagion between the Claims," *Transactions of the Fifteenth International Congress of Actuaries*, New York, II (1957), 219–27.
39. STEUTEL, F. *Preservation of Infinite Divisibility under Mixing and Related Topics*. Math. Centre Tracts 33. Amsterdam: Math. Centre, 1970.
40. ———. "Some Recent Results on Infinite Divisibility," *Stochastic Processes and Their Applications*, I (1973), 125–43.
41. ———. "Infinite Divisibility in Theory and Practice," *Scandinavian Journal of Statistics*, VI (1979), 57–64.
42. SUNDT, B., and JEWELL, W. "Further Results on Recursive Evaluation of Compound Distributions," *ASTIN Bulletin*, XII (1981), 1, 27–39.
43. TEICHER, H. "On the Multivariate Poisson Distribution." *Skandinavisk Aktuarietidskrift*, XXXVII (1954), 1–9.
44. TELLENBACH U. "Berechnung der Verteilung der Schadenzahlen bei bekannter Verteilung der Wartezeisten," *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker*, LXXII (1977), 35–46.
45. THORIN, O. "On the Infinite Divisibility of the Pareto Distribution," *Scandinavian Actuarial Journal*, 1977, pp. 31–40.
46. ———. "On the Infinite Divisibility of the Log-Normal Distribution," *Scandinavian Actuarial Journal*, 1977, pp. 121–48.
47. THYRION, P. "Extension of the Collective Risk Theory," *Skandinavisk Aktuarietidskrift*, LII (1969) (Suppl.), 84–98.
48. VAN HARN, K. *Classifying Infinitely Divisible Distributions by Functional Equations*. Math. Centre Tracts 103. Amsterdam: Math. Centre, 1978.

DISCUSSION OF PRECEDING PAPER

ELIAS S.W. SHIU:

The authors are to be complimented for this comprehensive survey of the Compound Poisson process, infinite divisibility, and number-of-claims distributions. I wish to supplement this excellent paper with the following comments.

The Sparre-Andersen risk process is discussed in section III.2 of the paper. It is assumed that the claim interoccurrence times $\{W_i\}$ are independent and identically distributed with probability distribution function S, with $S(0) = 0$. However, $T_1 = W_1$ is the time between the origin and the first claim epoch. It may not be reasonable to assume that the distribution of T_1 is also S unless S is a memoryless distribution (i.e., $N(t)$ is Poisson). Thorin ([8], [9]) has pointed out that if the probability density function of T_1 is assumed to be

$$[1 - S(t)] \Big/ \int_0^\infty [1 - S(y)]dy, \tag{1}$$

then $N(t)$ has stationary increments. An elegant discussion of equation (1) can be found in McFadden ([6], p. 366).

It may be possible to simplify section III.3 of the paper. Bühlmann ([2], p. 50) has proved that an operational time for the claim number process exists if and only if the intensities of frequencies are of the form

$$c_n(t) = \lambda_n \cdot \beta(t). \tag{III.3.5}$$

Thus, it seems natural to define

$$p^*_{n,m}(\tau) = p_{n,m}(t,s), \tag{III.3.6'}$$

where

$$\tau = \int_t^s \beta(\mu)\, d\mu. \tag{III.3.7'}$$

Comparing equations (III.3.4) with the top two equations on page 52 of Bühlmann [2], we see that Theorems III.3.3, III.3.4 and III.3.5 follow from the results in section 2.2.4 of [2] because of the new definition.

If one wishes to use the partial differential equation method in the paper, one may define

$$P_n(z,\tau) = \sum_{k=0}^{\infty} p^*_{n,n+k}(\tau)\, z^k. \qquad \text{(III.3.11')}$$

Then one has the partial differential equation

$$\frac{\partial}{\partial\tau} P_n(z,\tau) + z(1-z)b\frac{\partial}{\partial z} P_n(z,\tau) = (z-1)(a+bn)P_n(z,\tau) \qquad \text{(III.3.12')}$$

with the initial condition

$$P_n(z, 0) = 1. \qquad \text{(III.3.13')}$$

Since this paper is concerned with the distribution of aggregate claims, it might be of interest to review how Filip Lundberg derived the probability density functions of aggregate claims, $g(x,t)$ ([3]; [7], Chapter 1). Assume that

$$E[N(t)] = \lambda t.$$

Consider the infinitesimal time interval from t to $t + dt$. The probability that a claim occurs in this interval is λdt. The probability that this claim amounts to y is $f(y)dy$, where f is the claim amount density function. Hence,

$$g(x, t+dt) = (\lambda dt) \int_0^x g(x-y, t)f(y)dy + (1 - \lambda dt)g(x,t).$$

Rearranging the above, we obtain Lundberg's integro-differential equation for $g(x,t)$,

$$\frac{\partial}{\partial t} g(x,t) = \lambda\Big[\int_0^x g(x-y, t)f(y)dy - g(x, t)\Big]. \qquad (2)$$

Equation (2) can be solved by means of Laplace transformations. It follows from (2) that

$$\frac{\partial}{\partial t} \hat{g}(s,t) = \lambda[\hat{g}(s,t)\hat{f}(s) - \hat{g}(s, t)].$$

Thus,

$$\frac{\partial}{\partial t} \log \hat{g}(s, t) = \lambda(\hat{f}(s) - 1),$$

or

$$\hat{g}(s,t) = \hat{g}(s,0)e^{\lambda t(\hat{f}(s)-1)}$$
$$= e^{-\lambda t}\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} [\hat{f}(s)]^n \hat{g}(s,0).$$

Inverting the Laplace transforms, we have

$$g(x,t) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} f^{n*}(x) * g(x,0).$$

By hypothesis, $g(x,0)$ is a probability density function with whole mass 1 at $x = 0$. Thus,

$$g(x,t) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} f^{n*}(x).$$

Note that since $f^{0*}(x)$ is a delta function, the probability distribution function of aggregate claims has a jump of height $e^{-\lambda t}$ at $x = 0$. This fact also follows from Theorem II.2.4.

I find the statement of Theorem III.6.1 somewhat difficult to understand. The problem here is the computation of the composition of power series. Suppose that

$$P(z) = f(g(z)),$$

where

$$f(z) = \sum_{n=0}^{\infty} f_n z^n$$

and

$$g(z) = \sum_{n=0}^{\infty} g_n z^n.$$

Then

$$P(z) = \sum_{n=0}^{\infty} f_n (\sum_{i=0}^{\infty} g_i z^i)^n$$
$$= \sum_{n=0}^{\infty} f_n (\sum_{j=0}^{\infty} g_j^{n*} z^j)$$
$$= \sum_{j=0}^{\infty} (\sum_{n=0}^{\infty} f_n g_j^{n*}) z^j.$$

Thus, in general, the coefficients of the power series of P are infinite series, which can be difficult to evaluate.

On the other hand, suppose that we can find power series

$$k(z) = \sum_{n=0}^{\infty} k_n z^n$$

and

$$h(z) = \sum_{n=0}^{\infty} h_n z^n$$

such that

$$P(z) = k(zh(z)).$$

Then

$$\begin{aligned} P(z) &= \sum_{n=0}^{\infty} k_n \, (z \sum_{i=0}^{\infty} h_i z^i)^n \\ &= \sum_{n=0}^{\infty} k_n \, (\sum_{j=0}^{\infty} h_j^{n*} z^{n+j}) \\ &= \sum_{n=0}^{\infty} k_n \, (\sum_{r=n}^{\infty} h_{r-n}^{n*} \, z^r) \\ &= \sum_{r=0}^{\infty} (\sum_{n=0}^{r} k_n h_{r-n}^{n*}) \, z^r . \end{aligned}$$

Hence, in this case the coefficients of the power series of P are finite sums. Efficient algorithms for computing these coefficients can be found in the answer to exercise 4.7.11 on page 657 of Knuth [5] and also in section 1.6 of Henrici [4].

We remark that if f is a "nice" function, then the coefficients of the power series of $f(g(z))$ can be computed using the Faà di Bruno formula, which is equation (II.2.15).

It is pointed out in section II.3 of the paper that the equation

$$\frac{1}{n!} \int_0^t s^n \, e^{-s} \, ds = 1 - \sum_{j=0}^{n} \frac{t^j e^{-t}}{j!} \qquad (3)$$

can be verified easily by integration by parts. As actuarial students have learned about operators in finite differences, they may be amused by the following proof of (3).

Let D denote the differentiation operator $\frac{d}{dx}$. By the product rule

$$D[e^{ax} f(x)] = a\, e^{ax} f(x) + e^{ax} Df(x)$$
$$= e^{ax} [a + D]f(x). \quad (4)$$

From equation (4) one derives the exponential shift formula ([1], Section 36)

$$q(D)\,[e^{ax} f(x)] = e^{ax} q(a + D)f(x).$$

Thus,

$$D^{-1}\,(e^{-x} x^n) = e^{-x}\,(-1 + D)^{-1} x^n$$
$$= e^{-x}\,(-1 - D - D^2 - D^3 - \ldots)x^n$$
$$= e^{-x}\,(-x^n - nx^{n-1} - \ldots - n!).$$

Since

$$\int_0^t e^{-x} x^n\, dx = D^{-1}\,(e^{-x} x^n) \Big|_{x=0}^{x=t},$$

we obtain (3).

REFERENCES

1. BRAND, L. *Differential and Difference Equations*, New York: Wiley, 1966.
2. BÜHLMANN, H. *Mathematical Methods in Risk Theory*, New York: Springer-Verlag, 1970.
3. CRAMÉR, H. "Historical Review of Filip Lundberg's Works on Risk Theory," *Skandinavisk Aktuarietidskrift*, LII (1969), Supplement, 6–12.
4. HENRICI, P. *Applied and Computational Complex Analysis*, Vol. I, New York: Wiley, 1974.
5. KNUTH, D.E. *The Art of Computer Programming*, Vol. II, 2d ed., Reading, Mass.: Addison-Wesley, 1981.
6. McFADDEN, J.A. "On the Lengths of Intervals in a Stationary Point Process," *Journal of the Royal Statistical Society*, B, XXIV (1962), 364–82; Corrigenda, XXV (1963), 500.
7. SEAL, H.L. *Survival Probabilities: The Goal of Risk Theory*, New York: Wiley, 1978.
8. THORIN, O. "Stationarity Aspects of the Sparre-Andersen Risk Process and the Corresponding Ruin Probabilities," *Scandinavian Actuarial Journal*, 1975, 87–98.
9. ——————. "Stationarity and the Start of a Renewal Process. A Comment to

Hiliary L. Seal's Note: 'When Does a Renewal, or Other Stationary Point Process, Start?'," *Scandinavian Actuarial Journal*, 1976, 235–39.

(AUTHORS' REVIEW OF DISCUSSION)

HARRY H. PANJER AND GORDON E. WILLMOT:

Dr. Shiu's discussion is a valuable addition to the paper. The derivation of the Compound Poisson distribution using the method of infinitesimals, which was first obtained by Filip Lundberg, is important in motivating the use of this distribution. The derivation of formula II.3.2 using finite difference techniques is most interesting.

The point is raised with respect to the renewal (or Sparre-Andersen) process that it may not be reasonable to assume that the distribution of $T_1 = W_1$ be the same as that of the other W_i's. This reflects the fact that the origin may not be at the time of a claim occurrence. If the distribution of T_1 is assumed to be different, the resulting process is called a "modified renewal process." The special case when T_1 has the probability density function given by

$$\frac{1 - S(t)}{\int_0^\infty [1 - S(y)]dy}$$

is called an "equilibrium renewal process." The motivation for this latter density arises since it is the long-run distribution of future time until a claim occurs, conditional on the time since the last claim (often called the forward recurrence time). An excellent reference for renewal processes is given by [8] in the paper.

Dr. Shiu is correct in pointing out that the probabilities defined by III.3.6 depend only on the difference $\tau_2 - \tau_1$ and not on τ_1 and τ_2 separately. This leads to a somewhat simpler notational treatment of the model, although the final results are unaffected.

The point of Theorem III.6.1 is that the probability generating function of the number of claims can indeed be written in the form

$$P(z) = k(zh(z)),$$

as Dr. Shiu discusses, if $k(z)$ itself is of a particular form. This means that the probabilities can always be computed as finite sums. We thank Dr. Shiu for his additional references on power series.