TRANSACTIONS OF SOCIETY OF ACTUARIES
1990 VOL. 42

A QUEUEING THEORETIC APPROACH TO THE ANALYSIS OF THE CLAIMS PAYMENT PROCESS

GORDON E. WILLMOT

ABSTRACT

This paper represents an attempt to formulate a cohesive and consistent approach to the analysis of claim liabilities. Probabilistic tools from risk and queueing theory have been incorporated into a stochastic model that quantifies the variability inherent in such liabilities and at the same time reproduces intuitive results that may be arrived at from a deterministic standpoint. The model can be used to estimate various quantities of interest while providing a yardstick for measuring the accuracy of the estimates. Numerical examples illustrate the methodology.

Section 1 describes the problem and reviews some results from probability and risk theory. The liability of unreported claims is the subject of Section 2, where the first two parts outline an intuitive model that is well suited for practical implementation, as is illustrated by numerical examples. More general approaches that take into account seasonality of claims incurral, inflation, business growth, variations in risk levels, and other factors are considered in the final part.

The analysis of the liability of reported claims is considered in Section 3. This liability is shown to be statistically independent of the liability of unreported claims. A queueing theoretic approach to the modeling of the claim settlement process is proposed. In addition, models of various degrees of complexity are analyzed, and some numerical examples are provided. A recurring theme of this section is the approximate right tail behavior of the distribution of the liability of reported claims. Thus the amount needed to cover such liabilities can be estimated with a specified probability.

In Section 4 analysis of the delay in claims processing is discussed. An example illustrates how this information can be used to help analyze the efficiency of the claims administration system. Again, an approximation technique is developed for the distribution of the delay. Section 5 discusses areas for future research.

1. INTRODUCTION

1.1 The Claims Payment Process

The claims payment process is of considerable interest to insurers for various reasons. The process normally involves a time lag following incurral of the accident, death or other claim-causing event until final payment is made and the claim is settled. A consequence of the delay in payment of claims is the need to estimate outstanding claim liabilities as of a particular accounting date. This allows for the measurement of profit and loss within a particular accounting period to be made on a revenue basis. Estimation of these outstanding claim liabilities is a required component of any insurance company financial statement, whether it be annual statements required by regulatory authorities, in which case reporting is usually done on a statutory (conservative) basis, or an internal profit-and-loss statement on a more realistic basis. The import attached to the accuracy of such estimates is demonstrated in health insurance, for example, by the requirement that a retrospective test be performed to determine the accuracy of such liabilities in Schedule H of the NAIC statement in the U.S. Fundamental concepts involved in the analysis of these liabilities can be found in Bragg [3] or in Bowers et al. [2, chapters 5 and 9]. More advanced discussion of the philosophy of these liabilities and their intended purposes can be found in Koppel et al. [18]; see also Barnhart [1].

The time required to pay claims also reflects the efficiency of the insurers' claims processing system. Thus, a less efficient system will take longer, on average, to process a claim than a more efficient one. This may be a particularly important criterion in the selection of a carrier in group life and health insurance. The speed with which an insurer can efficiently process claims and remit payment may help determine whether new business can be obtained. Obviously, the need to monitor the whole claims payment process (as well as its constituent components) is crucial for insurers.

A mathematical model of the claims payment process can be quite useful. The model can normally provide numerical estimates of quantities of interest such as the liability at a point in time (needed for financial statements) or the time required to process a given type of claim. If a component of the system is unacceptable relative to expectation, the mathematical model can help predict the effect of a change. For example, if the time to approve a claim is unduly long due to too high a volume of claims, the effect of hiring additional staff can be assessed. Thus, a mathematical model that captures the salient physical features of the process can act as a "window" to the world that is being modeled. Considerable insight into the process can be obtained, a point that Bragg discusses in [3, p. 26], suggesting that such a model is of particular use for new blocks of business or where information is difficult or even impossible to obtain.

Holsten proposed one such mathematical model [13]. A major drawback of this model is its deterministic nature. Clearly, the problem is stochastic because the exact amount of outstanding claims cannot be ascertained. As a result, any deterministic formulation cannot capture the random variability inherent in the claims incurral process (the subject matter of risk theory, for example, Bowers et al. [2]) or the effect of an increased volume of claims in course of settlement, resulting in an increase in the total time to pay claims. Furthermore, the accuracy of the amount held for claim liabilities in various financial statements (the subject of the test discussed earlier) should more properly be assessed in light of its inherent variability before deciding whether the process used to set such amounts needs modification. This is particularly important because this variability can be quite large for some types of coverages, and such assessment cannot be made by using a deterministic model. Bragg [3, p. 36] suggests that the use of confidence intervals is appropriate in this connection; specifically, the amount held should have a "three-to-one likelihood of sufficiency." Such a requirement necessitates the use of a stochastic rather than a deterministic model.

A wide variety of stochastic models have been proposed in connection with "loss reserving," and many of these are described by Taylor [29] and Van Eeghen [32]. These models consider the "incurred but not reported," or IBNR, issue and, as Ruohonen notes [27], do not refer to the use of queueing theoretic techniques, nor do they attempt to integrate the methodology with standard risk theoretic models (for example, Bowers et al. [2]).

In this paper, the use of queueing techniques is shown to retain the advantages of other stochastic models by quantifying the inherent variability, while also allowing for modeling other important features such as the effect of congestion (due to large numbers of claims) on the claims payment process. Consequently, in addition to providing a stochastic model for the total time from incurral of a claim to payment (as well as the constituent parts), models for the number of outstanding claims at each stage of the payment process can be obtained (amounts held to cover the associated liabilities may need to be subdivided similarly for statement purposes; see O'Grady [21, p. 105]). In some situations, a model for the number of claims reported but unpaid may be unnecessary because the required claim counts can be obtained exactly. In many instances, however, such data may not be available in the required format (particularly if collected for another purpose), or they may be costly to obtain. Furthermore, one is often interested in forecasting profit-and-loss statements for several accounting periods into the future, and in these situations predictions of reported claims may need to be made.

Risk theoretic tools (for example, Bowers et al. [2]) can be employed to combine information on individual losses with the number of claims reported but unpaid, resulting in a stochastic model for the outstanding liability, and hence allowing variability to be quantified. Thus, the accuracy of an amount set aside to cover such liabilities may be assessed in light of the associated variability (which can be quite substantial).

An additional feature of the queueing theoretic approach is that, unlike many other models (compare Ruohonen [27]), the results are both consistent with and enhanced by the use of risk theory models. Consequently, the data required to use the models are the same as those needed for standard risk theoretic calculations. Thus, for weekly indemnity-type coverages, for example, a continuance table (for example, Bowers et al. [2, p. 377]) would be needed, whereas for life insurance the face amount and mortality rates are required (for example, Bowers et al. [2, section 13.3]). For other health- and casualty-type coverages, the data on individual losses are the same as those required for rate-setting purposes. A thorough discussion of modeling claim size distributions based on observed losses can be found in Hogg and Klugman [12].

The aim of this paper is to indicate various ways in which queueing theoretic tools can provide valuable insight into the claims payment process. Although some characteristics of practical situations are considered, it is not intended that the models or methods be used in any given situation. Consequently, only standard queueing methodology is used, but a knowledge of risk theory at the level of Bowers et al. [2] is sufficient background, as all other ideas are presented as needed. Furthermore, whereas the techniques may be applied to blocks of business in various lines of insurance (for example, group or individual, life or health), there can be specific coverages that are of a sufficiently long-term nature (for example, long-term disability) that the methods are not recommended.

1.2 Outline of the Paper

The remainder of the paper is devoted to the analysis of the claims payment process. Section 1.3 briefly reviews some of the important probabilistic and risk theoretic concepts, including generating functions, some parametric distributions, and compound distributions. Chapter 11 of Bowers et al. [2] covers many of these concepts. The claims incurral process is discussed in Section 1.4. The number of claims process is assumed to be a Poisson process, the usual risk theoretic assumption [2, chapter 12].

Section 2 deals with models for the claim liability due to unreported claims. The basic model is presented in Section 2.1 along with a numerical example involving life insurance, which helps to illustrate the methodology. A more general approach that can relate the reporting time to factors such as the size of a claim (a claimant may report a large claim more promptly than a relatively insignificant one) is proposed in Section 2.2. A numerical example is given. Other important subjects, such as inflation (clearly of interest in connection with various types of medical coverages), seasonality of claims incurral and reporting, growth of the business, and heterogeneity of risk levels, are treated in Section 2.3.

Reported claims are the subject matter of Section 3. Section 3.1 considers the reported claims process, and Section 3.2 presents the basic model with a numerical example. Section 3.3 utilizes queueing network theory in the simultaneous modeling of claims in various stages of the claims evaluation process. Such a breakdown is sometimes needed for statutory purposes (compare O'Grady [21, p. 105]). A more complicated model for the claim approval process is considered in Section 3.4. Section 3 also shows that relatively simple estimates of the claim liability can be obtained by using these models. Thus, for example, the amount needed to cover the liability with a specified confidence level can easily be estimated in the terminology of Bragg [3, p. 36].

Section 4 deals specifically with the analysis of the time that a claim is delayed in various processing stages. Thus, a policyholder or certificate holder may be interested in the total time from claim incurral until payment is actually received, because this determines the delay in receipt of funds. The insurer, on the other hand, may be interested in the time from receipt of notification of the claim until final approval or even actual disposal of the proceeds, because this time reflects the efficiency of the claims administration system.

Section 5 indicates areas for further research.

1.3 Concepts from Probability and Risk Theory

This section reviews the concepts useful in the stochastic modeling of the claims payment process.

Suppose that X is a random variable with probability density function (pdf) $f_X(x)$ if X is continuous or probability function (pf) $f_X(x)$ if X is discrete. The distribution function (df) is

$$F_X(x) = Pr(X \leq x) \tag{1.3.1}$$

and the moment generating function (mgf) of X is

$$M_X(s) = E(e^{sX}) = \int_{-\infty}^{\infty} e^{sx}\, dF_X(x). \tag{1.3.2}$$

If X is a discrete random variable defined on the nonnegative integers, it is often convenient to use the probability generating function (pgf)

$$P_X(s) = E(s^X) = \sum_{x=0}^{\infty} f_X(x) s^x \tag{1.3.3}$$

rather than (1.3.2). Evidently, $M_X(s) = P_X(e^s)$. The moments of X can be obtained from (1.3.2) or (1.3.3). Thus,

$$E(X) = M_X'(0) = P_X'(1), \tag{1.3.4}$$

whereas, for the variance,

$$Var(X) = M_X''(0) - \{M_X'(0)\}^2 = P_X''(1) + P_X'(1) - \{P_X'(1)\}^2. \tag{1.3.5}$$

If no ambiguity results, the subscript X can be dropped from (1.3.1), (1.3.2), or (1.3.3).

Various probability distributions will be used for modeling. A flexible family of distributions is the gamma family, with pdf

$$f(x) = \frac{\beta^{-\alpha} x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)}, \quad x > 0 \tag{1.3.6}$$

and mgf

$$M(s) = (1 - \beta s)^{-\alpha}, \quad s < \beta^{-1}, \tag{1.3.7}$$

where α and β are positive parameters. The exponential distribution is the special case $\alpha = 1$, and in this case the df is given by

$$F(x) = 1 - e^{-x/\beta}, \quad x > 0. \tag{1.3.8}$$

A second family of distributions that has slightly thicker tails than the gamma is the inverse Gaussian, with pdf

$$f(x) = \frac{\mu}{2}\left(\frac{\beta}{\pi x^3}\right)^{1/2} e^{-\frac{(2x-\mu\beta)^2}{4\beta x}}, \quad x > 0 \tag{1.3.9}$$

and mgf

$$M(s) = e^{-\mu\{(1-\beta s)^{1/2}-1\}}, \quad s \leq \beta^{-1}, \tag{1.3.10}$$

where μ and β are positive parameters. This latter family is discussed in detail by Chhikara and Folks [7]. Various other continuous pdf's are of use in various insurance contexts, and Hogg and Klugman [12] consider many of these in detail in connection with individual losses.

Of fundamental importance in connection with claim counts is the Poisson distribution with pf

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}; \quad x = 0, 1, 2, \ldots \tag{1.3.11}$$

and pgf

$$P(s) = e^{\lambda(s-1)}, \quad s < \infty. \tag{1.3.12}$$

Many important distributions in insurance may be obtained by mixing (compare Hogg and Klugman [12, section 2.7]). For example, if $f_i(x)$ is a pdf or pf for each $i \epsilon (1, 2, \ldots, k)$, then so is

$$f(x) = \sum_{i=1}^{k} q_i f_i(x) \tag{1.3.13}$$

where $\{q_i; i=1, 2, \ldots, k\}$ is a probability distribution. Mixtures of exponentials, for example, have been used in Bowers et al. [2, chapter 12] in connection with ruin theory. An important class of discrete distributions is obtained by letting the Poisson parameter be random, thus

$$f(x) = \int_0^\infty \frac{(\lambda y)^x e^{-\lambda y}}{x!} u(y)dy, \quad x = 0, 1, 2, \ldots \tag{1.3.14}$$

where $u(y)$ is itself the pdf of a positive random variable. The pgf associated with (1.3.14) is

$$P(s) = \int_0^\infty e^{\lambda y(s-1)} u(y)dy = M_1\{\lambda(s-1)\}, \tag{1.3.15}$$

where

$$M_1(s) = \int_0^\infty e^{sy}\, u(y)dy$$

is the mgf associated with the pdf $u(y)$. Mixed Poisson distributions are important in insurance modeling, as well as in a queueing context. The negative binomial distribution is the special case in which $u(y)$ is a gamma pdf, and in this case (1.3.14) becomes (with $\lambda = 1$)

$$f(x) = \binom{\alpha + x - 1}{x} \left(\frac{1}{1+\beta}\right)^\alpha \left(\frac{\beta}{1+\beta}\right)^x;$$

$$x = 0, 1, 2, \dots, \qquad (1.3.16)$$

and (1.3.15) is, using (1.3.7),

$$P(s) = \{1 - \beta(s - 1)\}^{-\alpha}, \quad s < 1 + \beta^{-1}. \qquad (1.3.17)$$

The geometric distribution is the special case $\alpha = 1$.

Except in special cases (such as the above), the integral in (1.3.14) is difficult to evaluate. An approximation can be given for large x, however. Using the notation $a(x) \sim b(x)$, $x \to \infty$ to mean $\lim_{x\to\infty} a(x)/b(x) = 1$, it can be shown (compare Willmot [36]) that if

$$u(x) \sim Cx^\phi e^{-\psi x}, \quad x \to \infty \qquad (1.3.18)$$

where $C > 0$, $-\infty < \phi < \infty$, and $\psi \geq 0$, then (1.3.14) satisfies

$$f(x) \sim \frac{Cx^\phi}{(\lambda + \psi)^{\phi+1}} \left(\frac{\lambda}{\lambda + \psi}\right)^x, \quad x \to \infty. \qquad (1.3.19)$$

Compound distributions play an important role in what follows. If N is a discrete random variable taking values on the nonnegative integers, and if $\{X_1, X_2, \dots\}$ is a sequence of independent and identically distributed random variables (also independent of N) with common mgf $M_X(s)$, then the random variable $Y = X_1 + X_2 + \dots + X_N$ (where $Y = 0$ if $N = 0$) has a compound distribution with mgf $M_Y(s) = P_N\{M_x(s)\}$. See Bowers et al. [2, chapter 11], for example. The distribution of Y is complicated, but Panjer [23] gives a recursive numerical algorithm for the evaluation of $f_Y(x)$ for various choices of $f_N(x)$. Also, suppose that

$$f_N(x) \sim Cx^\phi \theta^x, \quad x \to \infty, \qquad (1.3.20)$$

where $C>0, -\infty<\phi<\infty$, and $0<\theta<1$, and assume that there exists $\kappa>0$ satisfying $M_X(\kappa)=\theta^{-1}$. Then it can be shown (compare [8] and [35]) that

$$1 - F_Y(x) \sim C_1 x^{\phi} e^{-\kappa x}, \quad x \to \infty \tag{1.3.21}$$

where

$$C_1 = C/\{(e^{\kappa} - 1)(\theta M'_X(\kappa))^{\Theta+1}\}$$

if X is itself discrete on the nonnegative integers and

$$C_1 = C/\{\kappa[\theta M'_X(\kappa)]^{\phi+1}\}$$

if X is continuous. Clearly, (1.3.19) is itself of the form (1.3.20), and so tail estimates for the distribution of Y hold if N is negative binomial, for example.

1.4 The Claims Incurral Process

One of the main building blocks in the construction of a model for the claims payment process is a model for the claims incurral process. We assume that the number of incurred claims $\{K_t;\ t\geq 0\}$ is an ordinary Poisson process (that is, K_t is the number of claims incurred in $[0, t]$). This is the usual model in risk theory (compare Bowers et al. [2, chapter 12]). Thus, $\{K_t;\ t\geq 0\}$ has the following properties:

(i) $K_0=0$

(ii) $\{K_t;\ t\geq 0\}$ has stationary and independent increments

(iii) $Pr\{K_{t+h}-K_h=k\}=(\lambda t)^k e^{-\lambda t}/k!;\ k=0, 1, 2, \ldots$.

The parameter λ is called the rate of the process. A more detailed discussion of the assumptions leading to a Poisson process can be found in Bowers et al. [2, pp. 346–350].

A few other properties of the Poisson process are used subsequently, and they are recorded here for completeness. If a claim is classified upon incurral as being of type 1 with probability p and as type 2 with probability $1-p$, independently of other events, then the number of type 1 and the number of type 2 claims incurral processes are independent Poisson processes with rates λp and $\lambda(1-p)$, respectively. See Ross [26, pp. 203–206] for a proof of this statement. Thus, a Poisson process can be decomposed into independent Poisson processes, and the extension to more than two types of claims follows easily by induction. Similarly, if two independent Poisson processes with rates λ_1 and λ_2 are superimposed (that is, only the total process is observed), then the sum of the two processes is a Poisson process with rate $\lambda_1+\lambda_2$. The same property holds for more than two processes by induction.

Furthermore, the times of the k claims in $(0, t)$, given that k claims were incurred in $(0, t)$, are independent and identically distributed, each with the uniform density $f_t(x) = t^{-1}$, $0<x<t$. See Ross [26, pp. 209–211].

The total claims incurred process $\{Y_t;\ t \geq 0\}$ is then a compound Poisson process. Suppose that $\{X_1, X_2, \ldots\}$ is a sequence of independent and identically distributed random variables representing claim sizes (that is, X_i is the size of the i-th claim), also independent of $\{K_t;\ t \geq 0\}$. Then $Y_t = X_1 + X_2 + \ldots + X_{K_t}$ (with $Y_t = 0$ if $K_t = 0$). This process is the study of much of risk theory (for example, Bowers et al. [2, chapters 11–13]). Similar decomposition and superposition properties hold for $\{Y_t;\ t \geq 0\}$ as they do for the Poisson process (compare Karlin and Taylor [15, pp. 430–436]). In particular, the total of all claims of a certain size (that is, claims whose size is contained in a specified subset of the real line) is a compound Poisson process, independently of claims of other sizes.

In the remainder of the paper this model is assumed for the claims incurral process, and results quoted here are used freely in studying properties of the claims payment process.

2. UNREPORTED CLAIMS

2.1 The Basic Model

A main component of the claim liability is the portion attributable to the unreported claims. A wide variety of methods have been proposed (see Van Eeghen [32] and Taylor [29]), but, as noted in Ross [26], they do not use queueing theoretic techniques.

In this paper, the compound Poisson model for incurred claims (consistent with risk theory) is assumed, as discussed in Section 1.4. Suppose that the number of incurred claims $\{K_t;\ t \geq 0\}$ follows a Poisson process with rate λ. Let B be the random variable denoting the time from incurral of a claim to the time of reporting with distribution function $F_B(x)$. Furthermore, assume that reporting times are independent of each other. Then the distribution of N_t, the number of incurred but unreported claims at time t, can be determined. As shown by Ross [26, p. 212], in connection with the infinite server (M/G/∞) queue, the distribution of N_t is itself Poisson with mean

$$\lambda_t = \lambda \int_0^t \{1 - F_B(x)\}\, dx. \tag{2.1.1}$$

Under the risk theoretic model, the total unreported claims (U_t) is compound Poisson, that is, $U_t = X_1 + X_2 + \ldots + X_{N_t}$. As shown in Panjer [22], for example, if the single claim sizes (denoted generically by X) are discrete on the positive integers, then the distribution of U_t can be calculated recursively by using the formula

$$f_{U_t}(x) = \frac{\lambda_t}{x} \sum_{y=1}^{x} y f_X(y) f_{U_t}(x-y), \quad x > 0 \tag{2.1.2}$$

beginning with $f_{U_t}(0) = e^{-\lambda_t}$. A similar formula holds if X has a continuous distribution (compare Panjer [23]). Using (2.1.2), numerical values of the percentiles of the distribution of U_t can easily be obtained. The first two moments are

$$E(U_t) = \lambda_t E(X) \tag{2.1.3}$$

and

$$\mathrm{Var}(U_t) = \lambda_t E(X^2). \tag{2.1.4}$$

When statistical equilibrium has been reached, considerable simplification follows, and numerous intuitively appealing results can be obtained. From (2.1.1),

$$\lambda_\infty = \lim_{t\to\infty} \lambda_t = \lambda E(B), \tag{2.1.5}$$

and so evaluation of λ_∞ requires only knowledge of the mean reporting lag $E(B)$ rather than the distribution function $F_B(x)$, as is the case for λ_t when $t < \infty$. In particular, no distributional assumption need be made about B. Also, from (2.1.3) with $t \to \infty$,

$$E(U_\infty) = \lambda E(B)E(X) = E(Y_1)E(B), \tag{2.1.6}$$

that is,

expected liability = expected annual claims × expected reporting lag.

This result is very intuitive and might well be used in the absence of any formalized model. The model considered here may consequently be viewed as an aid to intuition, and not a replacement. Because $\lambda_t \le \lambda_\infty$, $E(U_t) \le E(Y_1)E(B)$ and so (2.1.6) provides a conservative bound on the mean claim liability.

These types of models have also been considered by Karlsson [16], Rantala [24], and Ruohonen [27]. A numerical example is now presented. The numerical values chosen are for illustration only and are not meant to be representative of a realistic situation.

Example 2.1.1

Vandebroek and DePril [31] considered a portfolio of lives insured under life insurance. Table 1 gives the number of lives n_{ij} in the portfolio for each insurance amount i and mortality rate q_j.

TABLE 1

NUMBER OF LIVES n_{ij}

Amount	100,000 q_j													
i	804	1,000	1,262	1,605	2,064	2,670	3,476	4,544	5,962	7,847	10,339	13,642	18,009	23,784
1	16	14	14	7	6	4	–	–	3	1	–	1	2	2
2	1	8	13	9	11	6	4	7	5	10	2	5	–	1
3	–	–	2	2	–	6	–	1	–	1	–	2	6	4
4	3	3	1	1	3	1	–	–	–	1	1	2	2	1
5	–	5	5	1	–	1	–	–	1	–	–	–	–	1
6	–	1	16	14	11	10	6	2	–	1	–	1	2	2
7	–	3	7	12	13	26	18	9	6	5	4	3	–	2
8	–	–	7	5	6	11	15	19	6	7	8	8	5	2
9	–	2	1	6	3	4	9	8	4	5	4	7	4	3
10	–	–	–	6	6	7	6	6	6	3	7	4	2	3
11	–	–	2	–	3	6	9	4	10	4	1	6	2	2
12	–	–	–	–	–	1	4	2	4	4	2	4	1	2
13	–	–	–	1	1	1	2	1	1	1	1	–	–	–
14	–	–	–	2	–	3	1	2	1	1	–	1	1	–
15	–	–	–	–	1	–	2	4	–	3	1	–	–	1
16	–	–	–	–	–	2	–	1	–	3	–	1	1	–
17	–	–	–	–	1	–	1	–	–	–	3	–	–	1
18	–	–	–	–	–	3	1	–	–	–	–	–	1	–
19	–	–	–	–	–	–	–	–	–	–	2	–	1	–
20	–	–	–	–	–	1	–	–	–	–	–	2	–	–
21	–	–	–	–	–	1	3	–	–	–	1	–	–	–
22	–	–	–	–	–	–	–	1	1	1	–	–	2	–
23	–	–	–	–	–	–	1	1	–	–	1	–	–	–
24	–	–	–	–	–	–	1	–	–	–	1	–	–	–
25	–	–	–	–	–	–	–	–	–	–	–	–	–	–
26	–	–	–	–	–	–	–	–	–	1	–	–	–	1
27	–	–	–	–	–	–	1	–	–	1	1	–	–	–
28	–	–	–	–	–	–	–	–	–	–	–	–	1	–

The compound Poisson model in Bowers et al. [2, pp. 381–382] can be used. Define

$$\lambda(i) = -\sum_j n_{ij} \log(1 - q_j) \tag{2.1.7}$$

and

$$\lambda = \sum_i \lambda(i). \tag{2.1.8}$$

According to the model, the total incurred claims process for the portfolio is a compound Poisson process with Poisson rate $\lambda = 4.27137$ and single claim amount distribution given by

$$f_X(x) = \lambda(x)/\lambda. \tag{2.1.9}$$

Suppose that previous studies indicate that the average reporting time for a claim is one month. Then, from (2.1.5), $\lambda_\infty = \lambda/12 = 0.355947$, and the total claim liability U_∞ is compound Poisson with parameter λ_∞ and single claim size distribution $f_X(x)$. In particular, the mean is 3.10424 from (2.1.3), and the variance is 36.7392 from (2.1.4). The distribution of U_∞ is easily obtained from (2.1.2), and the results are given in Table 2 together with the single claim size distribution $f_X(x)$ and df $F_X(x)$.

The mean could be used in choosing a numerical value to cover the liability. Alternatively, an amount that is adequate for a specified proportion of time could be chosen, as suggested by Bragg [3]. For example, an amount of 7 would be expected to cover the liability 80 percent of the time, as is evident from Table 2. The model yields simple quantitative estimates of the variability inherent in the liability, requiring only the mean reporting times as input. In fact, the entire distribution can be easily obtained numerically.

2.2 Individual Variations in Reporting Patterns

Although the model discussed in the previous section is sufficiently general for many applications, characteristics of particular situations may require refinements. One possible situation involves differences in reporting patterns for various segments of the portfolio. In particular, reporting patterns may be related to concomitant factors that are independent of the number of incurred claims process. The decomposition properties in Section 1.4 can be used to refine the model of Section 2.1.

Suppose that there are m different classes of individuals in the portfolio with respect to reporting times, and the probability that a given incurred claim is of type i is q_i; $i = 1, 2, \ldots, m$. Then the incurred claims process for class i is compound Poisson with Poisson parameter λq_i and single claim size distribution denoted by $f_i(x)$. Let B_i denote the reporting time random variable for class i. By applying the results of Section 2.1 to each class, the total claim liability for class i can be modeled as a compound Poisson random

TABLE 2

x	$f_X(x)$	$F_X(x)$	$f_{U_\infty}(x)$	$F_{U_\infty}(x)$
0	0.000000	0.000000	0.700509	0.700509
1	0.047510	0.047510	0.011846	0.712356
2	0.081115	0.128626	0.020326	0.732682
3	0.062511	0.191137	0.015929	0.748611
4	0.028870	0.220007	0.007757	0.756368
5	0.010687	0.230694	0.003244	0.759612
6	0.053674	0.284368	0.013821	0.773433
7	0.102390	0.386758	0.026007	0.799440
8	0.145270	0.532028	0.037151	0.836591
9	0.103832	0.635860	0.027584	0.864175
10	0.090073	0.725933	0.024687	0.888862
11	0.080285	0.806219	0.022321	0.911183
12	0.052252	0.858471	0.015264	0.926447
13	0.009860	0.868330	0.004736	0.931183
14	0.016292	0.884623	0.006509	0.937692
15	0.019981	0.904604	0.007629	0.945321
16	0.015322	0.919925	0.006732	0.952053
17	0.014234	0.934159	0.006629	0.958682
18	0.006948	0.941107	0.004879	0.963561
19	0.009121	0.950228	0.005166	0.968727
20	0.007058	0.957285	0.004167	0.972894
21	0.005505	0.962790	0.003303	0.976197
22	0.012820	0.975610	0.004867	0.981064
23	0.004315	0.979924	0.002562	0.983626
24	0.003248	0.983173	0.002107	0.985733
25	0.000000	0.983173	0.001159	0.986893
26	0.007480	0.990653	0.002902	0.989795
27	0.005093	0.995745	0.002205	0.991999
28	0.004255	1.000000	0.001935	0.993935

variable with Poisson parameter $\lambda q_i E(B_i)$ and single claim size distribution $f_i(x)$, independently of other classes. Thus, by the additivity property of independent compound Poisson random variables (compare Bowers et al. [2, p. 327]), the total claim liability U_∞ is compound Poisson with Poisson parameter

$$\lambda_* = \lambda \sum_{i=1}^{m} q_i \, E(B_i) \tag{2.2.1}$$

and "single claim amount" distribution

$$f_*(x) = \frac{\sum_{i=1}^{m} q_i \, E(B_i) \, f_i(x)}{\sum_{i=1}^{m} q_i \, E(B_i)}. \tag{2.2.2}$$

Hence, the moments and probability distribution of U_∞ can be easily obtained by using the results of Section 2.1, but with λ_∞ and $f_X(x)$ replaced by λ_* and $f_*(x)$, respectively.

Using the more complicated model of this section requires knowing q_i and $f_i(x)$ for each class. In at least one important situation this is not difficult. Suppose that the time to report a claim depends on the size of the claim (for example, large claims may have a shorter mean reporting time than small claims). Then the total incurred claims process can be modeled as a compound Poisson process with parameter λ and claim size pdf or pf $f_X(x)$ and df $F_X(x)$ (as in Section 2.1). Partition the positive real line $[0, \infty)$ into the intervals $[c_{i-1}, c_i)$ for $i=1, 2, \dots, m$, where $c_0=0$ and $c_m=\infty$. Let a claim be of type i if the amount of the claim is in the interval $[c_{i-1}, c_i)$. Then

$$q_i = \int_{b[c_{i-1},c_i)} dF_X(x); \quad i = 1, 2, \dots, m \tag{2.2.3}$$

and

$$f_i(x) = \begin{cases} f_X(x)/q_i, & x \epsilon (c_{i-1}, c_i) \\ 0, & \text{otherwise.} \end{cases} \tag{2.2.4}$$

Hence, q_i and $f_i(x)$ are easily constructed from $f_X(x)$, and only the partition has to be determined. This should be done on the basis of observed variations in reporting time.

A numerical example follows. No significance should be attached to the numerical values, because they are purely for illustration.

Example 2.2.1

Consider the life portfolio of Example 2.1.1. Suppose that the average reporting time of claims in excess of 10 has been determined from previous studies to be one-half of a month, whereas claims of amount 10 or less are reported in one and a quarter months on average. This suggests the choice $m=2$ and the partition (0, 10.5) and (10.5, ∞). Using the distribution $f_X(x)$ as given in Example 2.1.1, $q_1=0.725933$ and $q_2=0.274067$. Because $E(B_1)=5/48$ and $E(B_2)=1/24$, (2.2.1) implies $\lambda_*=0.371769$. The distribution $f_1(x)$ and $f_2(x)$ can be obtained from (2.2.4), and from (2.2.2)

$$f_*(x) = 0.868799\, f_1(x) + 0.131201\, f_2(x). \tag{2.2.5}$$

Values of $f_1(x)$, $f_2(x)$, and $f_*(x)$ and the associated df $F_*(x)$ are given in Table 3.

TABLE 3

x	$f_1(x)$	$f_2(x)$	$f_*(x)$	$F_*(x)$
0	0.000000	0.000000	0.000000	0.000000
1	0.065447	0.000000	0.056860	0.056860
2	0.111740	0.000000	0.097079	0.153939
3	0.086111	0.000000	0.074813	0.228753
4	0.039770	0.000000	0.034552	0.263305
5	0.014722	0.000000	0.012790	0.276095
6	0.073938	0.000000	0.064237	0.340332
7	0.141046	0.000000	0.122541	0.462872
8	0.200115	0.000000	0.173860	0.636732
9	0.143033	0.000000	0.124267	0.760999
10	0.124079	0.000000	0.107799	0.868799
11	0.000000	0.292941	0.038434	0.907233
12	0.000000	0.190654	0.025014	0.932247
13	0.000000	0.035976	0.004720	0.936967
14	0.000000	0.059447	0.007800	0.944767
15	0.000000	0.072905	0.009565	0.954332
16	0.000000	0.055905	0.007335	0.961667
17	0.000000	0.051936	0.006814	0.968481
18	0.000000	0.025350	0.003326	0.971807
19	0.000000	0.033280	0.004366	0.976173
20	0.000000	0.025751	0.003379	0.979552
21	0.000000	0.020085	0.002635	0.982187
22	0.000000	0.046776	0.006137	0.988324
23	0.000000	0.015743	0.002065	0.990389
24	0.000000	0.011852	0.001555	0.991944
25	0.000000	0.000000	0.000000	0.991944
26	0.000000	0.027292	0.003581	0.995525
27	0.000000	0.018582	0.002438	0.997963
28	0.000000	0.015524	0.002037	1.000000

By using λ_* and $f_*(x)$ in place of λ_∞ and $f_X(x)$ in the results of Section 2.1, the mean claim liability is 2.78077 from (2.1.3). The variance is 27.8008 from (2.1.4). By using (2.1.2), the distribution of U_∞ is found, and this is given in Table 4. The third column can be used to select an adequate amount for covering the liability a specified proportion of the time.

2.3 Other Generalizations

In the previous two sections we proposed relatively simple models for the claim liability. In this section, we show how various realistic phenomena such as the effect of seasonality with respect to the incurral of claims, growth in the business, and heterogeneity of risks in the portfolio can be incorporated into the model by assuming a more general model of claims incurral than the Poisson. Other factors that can be modeled include inflation and seasonality of claims reporting.

TABLE 4

x	$f_{U_*}(x)$	$F_{U_*}(x)$	x	$f_{U_*}(x)$	$F_{U_*}(x)$
0	0.689513	0.689513	15	0.005955	0.961491
1	0.014576	0.704089	16	0.005948	0.967439
2	0.025039	0.729128	17	0.005819	0.973257
3	0.019705	0.748833	18	0.004525	0.977783
4	0.009717	0.758550	19	0.003903	0.981686
5	0.004172	0.762722	20	0.002906	0.984592
6	0.017144	0.779866	21	0.002066	0.986658
7	0.032151	0.812017	22	0.002683	0.989341
8	0.046001	0.858018	23	0.001490	0.990830
9	0.034467	0.892485	24	0.001310	0.992140
10	0.031073	0.923558	25	0.000829	0.992969
11	0.013414	0.936972	26	0.001628	0.994597
12	0.009567	0.946539	27	0.001240	0.995837
13	0.004007	0.950546	28	0.001068	0.996905
14	0.004990	0.955536			

2.3.1 *The Number of Claims Incurred*

The assumption that the number of incurred claims $\{K_t;\ t \geq 0\}$ is a Poisson process may be too restrictive in some situations. Because the rate of the process is a constant λ that does not change with time, the number of claims incurred in any period has the same distribution as the number incurred in any other period of the same length. It may be of interest to relax this assumption in various situations.

In this section $\{K_t;\ t \geq 0\}$ is assumed to be an order statistic process. This more general process has the property that, given $K_t = k \geq 1$, the times of the k claims are independent and identically distributed over $(0,t)$ with df

$$H_t(x) = \frac{E(K_x)}{E(K_t)}; \quad 0 < x < t. \tag{2.3.1}$$

This more general process can be used to accommodate the following factors.

(a) Incurred Claim Seasonality and Business Growth

There may be a seasonal pattern to claims incurral, such as a higher incidence of health-related claims during the winter months than in the summer. This can have a significant impact on the unreported claim liability at a given time. Another factor that can affect incurred claims is a change in the size of the portfolio over time. Growth in the business would result in an increase in the rate of claims incurral. These phenomena cannot be reflected by the ordinary Poisson process of claims incurral.

The nonhomogeneous Poisson process (for example, Ross [25, pp. 46–49, 53]) can be used in these situations. This process does not require that $E(K_x)$ be proportional to x as the ordinary Poisson process does. Thus the rate of the process

$$\lambda(x) = \frac{d}{dx} E(K_x)$$

is not restricted to a constant, but need only be nonnegative. Consequently, it can vary with time in such a manner as to describe these phenomena. Seasonality in claims incurral can be obtained by choosing $\lambda(x)$ to be a function both of the integer part of x in order to represent the year as well as the fractional part of x to represent the season. Similarly, growth in the business can be modeled by letting $\lambda(x)$ reflect the corresponding rate of change. For example, if the growth rate can be assumed to be exponential, this may be reflected by the choice $E(K_x) = ae^{bx}$, and thus the rate of the process is $\lambda(x) = abe^{bx}$. One could choose $\lambda(x)$ to reflect both seasonality of claims incurral and growth of the business.

(b) Heterogeneity of Risk Levels in the Portfolio

All risk classification schemes attempt to discriminate between different types of risk, with the intended result that all risks within a particular "cell" may be considered to be homogeneous with respect to the risk level. Unfortunately, this is not completely accomplished by even the most discriminating risk classification scheme, and some heterogeneity of risk levels (that is, some good and bad risks relative to the average) remains.

This characteristic may be reflected through the use of another fairly general type of process with the order statistic property, namely, the mixed Poisson process (for example, Willmot [34]). In this case

$$Pr\{K_{t+h} - K_h = k\} = \int_0^\infty \frac{(\lambda t)^k e^{-\lambda t}}{k!} dU(\lambda) \qquad (2.3.2)$$

where $U(\lambda)$ is the df of a nonnegative random variable (if $U(\lambda)$ is a gamma df, then the process is referred to as a Polya process). This model is common in automobile insurance; and in Buhlmann [4], Equation (2.3.2) is interpreted as the probability that one risk taken at random from the portfolio gives rise to k claims in $(h, h+t)$. The "structure function" $U(\lambda)$ represents the distribution of the levels of risk in the portfolio (as measured by the expected number of claims incurred), and thus provides a mechanism for dealing with the nonhomogeneity.

2.3.2 *A General Model*

Let $W(x,t)$ denote a random variable representing the liability at time t attributable to a claim that is incurred at time x. Then the total liability U_t at time t is the sum of the liabilities from all claims incurred before time t. The distribution of U_t is most easily characterized in terms of its mgf. By conditioning on both the number and times of the claims incurred,

$$M_{U_t}(s) = Pr\{K_t = 0\} + \sum_{k=1}^{\infty} Pr\{K_t = k\}$$

$$\times \int_0^t \int_0^t \cdots \int_0^t \prod_{i=1}^{k} [E\{e^{sW(x_i,t)}\} \, d_{x_i} H_t(x_i)].$$

Because the k-fold integral factors into the same integral repeated k times, the mgf of U_t is

$$M_{U_t}(s) = P_{K_t}\{M_{W_t}(s)\} \tag{2.3.3}$$

where $P_{K_t}(s)$ is the pgf of K_t and

$$M_{W_t}(s) = \int_0^t E\{e^{sW(x,t)}\} \, d_x H_t(x) \tag{2.3.4}$$

is the mgf of a random variable obtained by mixing the distribution of $W(x,t)$ over the interval $(0,t)$ by the mixing distribution $H_t(x)$.

It is evident from the discussion in the paragraph following (1.3.21) that the representation (2.3.3) implies that U_t has a compound distribution. Thus, if $K_t=0$, then $U_t=0$, and if $K_t>0$, then U_t is the sum of K_t independent random variables, each with mgf (2.3.4).

2.3.3 *Inflation and Seasonality of Reporting*

The relationship between the liability $W(x,t)$ at time t for the claim incurred at time x and both the amount of the claim and the reporting time can be quite complex when the effects of inflation and seasonality are considered in the reporting time of the claim.

(a) Inflation

To allow for inflation, assume that X is a random variable representing the amount of a single claim at some time point in the past (that is, before t), such as at time 0 or at time x. Then, as in Hogg and Klugman [12, section 5.2], the effects of inflation are such that the value of the claim at time t is a scalar multiple of X, namely, $a(x,t)X$. Suppose, for example, that X represents the amount payable on a claim incurred at time 0. If claims inflation is characterized by a force of inflation $\delta_1(y)$, the amount payable on a claim incurred at time x is

$$X e^{\int_0^x \delta_1(y)dy}.$$

If the time value of money involves a force of interest $\delta_2(y)$, the value at time t of a claim incurred at time x is

$$X e^{\int_0^x \delta_1(y)dy + \int_x^t \delta_2(y)dy}$$

if interest is payable on claim amounts. This suggests that one could choose

$$a(x,t) = e^{\int_0^x \delta_1(y)dy + \int_x^t \delta_2(y)dy}. \tag{2.3.5}$$

Because $a(x,t)$ can be an arbitrary function, however, other inflationary or trending patterns could be used.

(b) Seasonality in Claims Reporting

Seasonality in reporting can also be modeled by assuming that the reporting time B_x of a claim incurred at time x depends on the time of incurral x, perhaps through the integral and fractional part of x. With these assumptions,

$$W(x,t) = \begin{cases} 0, & B_x \leq t - x \\ a(x,t)X, & B_x > t - x \end{cases} \tag{2.3.6}$$

because there is no liability if the claim is reported by time t (that is, $B_x \leq t-x$). Thus, from (2.3.6), the mgf of $W(x,t)$ is

$$E\{e^{sW(x,t)}\} = F_{B_x}(t-x) + \{1 - F_{B_x}(t-x)\} M_X\{sa(x,t)\} \tag{2.3.7}$$

where $M_X(s)$ is the mgf of X. The expression (2.3.7) can be substituted into (2.3.4).

Note that (2.3.4) holds regardless of the manner in which $W(x,t)$ is dependent on the amount of the claim at time x and the ensuing reporting time.

Hence, although (2.3.6) seems reasonable, there may be other formulations that could be used.

2.3.4 *Further Remarks*

The model of Section 2.1 is a special case of the current model. Because $E(K_t) = \lambda t$, (2.3.1) yields the uniform distribution on $(0, t)$. With $a(x, t) = 1$ and $F_{B_x}(y) = F_B(y)$, (2.3.7) implies that (2.3.4) becomes

$$M_{W_t}(s) = \frac{1}{t}\int_0^t \{F_B(t-x) + [1 - F_B(t-x)]M_X(s)\}\,dx$$

$$= \frac{1}{t}\int_0^t \{F_B(x) + [1 - F_B(x)]M_X(s)\}\,dx.$$

Hence,

$$\lambda t\{M_{W_t}(s) - 1\} = \lambda \int_0^t \{F_B(x) + [1 - F_B(x)]\,M_X(s)\}\,dx - \lambda t$$

$$= \lambda \int_0^t \{F_B(x) + [1 - F_B(x)]M_X(s) - 1\}\,dx$$

$$= \lambda \int_0^t \{1 - F_B(x)\}\,\{M_X(s) - 1\}\,dx$$

$$= \lambda_t\,\{M_X(s) - 1\}$$

by using (2.1.1). Because $P_{K_t}(s) = \exp\{\lambda t(s-1)\}$, (2.3.3) is the mgf of the compound Poisson random variable U_t of Section 2.1.

This model is quite general, and the main difficulty in employing it lies in the evaluation of the distribution with mgf (2.3.4) (if it may be obtained, the recursive techniques in Panjer [23] often allow for the numerical evaluation of the distribution of U_t). In general, (2.3.4) and (2.3.7) yield

$$M_{W_t}(s) = \int_0^t \{F_{B_x}(t-x) + [1 - F_{B_x}(t-x)]M_X[sa(x, t)]\}\,d_xH_t(x).$$

By using this result and the properties of conditional expectation, the associated df $F_{W_t}(y)$ satisfies

$$F_{W_t}(y) = 1 - \int_0^t \{1 - F_{B_x}(t - x)\} \{1 - F_X[y/a(x, t)]\} \, d_x H_t(x). \quad (2.3.9)$$

If X is a pdf $f_X(\cdot)$, then (2.3.9) can be differentiated to give the pdf

$$f_{W_t}(y) = \int_0^t \frac{f_X\{y/a(x, t)\}}{a(x, t)} \{1 - F_{B_x}(t - x)\} \, d_x H_t(x). \quad (2.3.10)$$

Numerical integration could be used to evaluate (2.3.9) or (2.3.10). Panjer [23] describes how the pdf of the compound distribution of U_t with mgf (2.3.3) can be evaluated numerically if $\{K_t;\ t \geq 0\}$ is a (nonhomogeneous) Poisson or Polya process.

This approach has other uses as well. It shows how the model is modified if more complicated assumptions of phenomena such as inflation are incorporated. It also provides insight into the behavior of the liability. In particular, the compound Poisson form of the distribution of U_t holds quite generally as long as $\{K_t;\ t \geq 0\}$ is a (nonhomogeneous) Poisson process. Similarly, if $\{K_t;\ t \geq 0\}$ is assumed to be a Polya process, then K_t has a negative binomial distribution and the distribution of U_t remains of compound negative binomial form (compare Bowers et al. [2, pp. 323–325]), as is evident from (2.3.3).

3. REPORTED CLAIMS

3.1 The Reported Claims Process

A second major category of the claim liability is that portion attributable to claims for which notification has reached the insurer but for which no payment has been made. As discussed in Section 1.1, one may be able to obtain the required claim amounts exactly, and hence they need not be estimated by using a model. If the data are not readily available, however, or if estimates of future reported claims are needed for forecasting profit-and-loss statements, the use of a model may be worthwhile. Furthermore, this portion of the claim liability can be influenced by the insurer through modifications to the claims settlement process. A model can often be used to predict the effect of these changes without actually implementing them.

The number of reported claims is of central importance in the analysis of the reported claim liability. Recalling from Section 1 that claims are incurred according to a Poisson process with rate λ (see Section 1.4) and that each of these is reported to the insurer at random time B with df $F_B(x)$ later, independently of all other claims, then from Ross [25, p. 39] the number of reported claims in $(0, t]$ is both Poisson distributed with mean

$$\lambda \int_0^t F_B(x)dx$$

and independent of the number of unreported claims N_t in $(0, t]$. The independence of the number of reported and unreported claims at a point in time is a useful feature of the model, because it implies that the unreported claim liability U_t and the reported claim liability R_t are independent. This follows from the fact that U_t is assumed to be the sum of N_t independent individual claim amounts, whereas R_t is the sum of A_t independent individual claim amounts, where A_t is the number of claims reported but unpaid at time t. Because A_t depends on the number of reported claims (which is independent of N_t) and the claim settlement process (which is independent of unreported claims), the independence of U_t and R_t follows. As a result, the unreported and reported claim liabilities can be analyzed separately and without regard for each other, clearly a simplifying feature of the model.

A second important property of this approach is that the number of reported claims is a nonhomogeneous Poisson process with rate $\lambda F_B(t)$, as shown in Ross [25, p. 48], where it is pointed out that as $t \to \infty$, the process becomes an ordinary Poisson process. This implies that the input process to the claims payment discipline can be assumed to be a Poisson process in equilibrium (that is, for large values of t). This result is heavily relied upon in the remainder of the paper.

3.2 The Basic Model

The analysis of the reported claim liability is fundamentally different than the unreported liability due to the interaction of claims. One can normally assume that the time it takes to report a claim does not depend on other claims in a similar incurred but unreported state. The same cannot be said for the reported claims in general, however, because the presence of too many claims waiting for approval at one time can cause a backlog and hence a delay in the time until payment is made.

This congestion can be incorporated into a stochastic framework through a queueing formulation of the problem. One imagines that claims are reported to the insurer and that they "queue up" in the claims area waiting to be processed. Once approved, payment is made. The process of approving claims for payment can then be visualized in terms of a particular queueing discipline. The number of reported claims process is the input process to this "queue," and this is a Poisson process once equilibrium has been reached (see Section 3.1), an assumption that will henceforth be made. Let A represent the number of claims that are reported but unpaid, that is, the number in the queueing system. In keeping with risk theoretic methodology, the total liability for reported but unpaid claims R is given by $R=X_1+X_2+\ \dots\ +X_A$ (with $R=0$ if $A=0$). As before $\{X_1, X_2, \dots\}$ is an independent and identically distributed sequence of claim amounts, and in this case X_i represents the amount of the i-th claim in the system.

Assume, in the simplest case, that claims are approved in the order that they are reported by a single claims evaluator (examiner) and that, once approved, they are paid immediately. Suppose that the time to approve a claim T is exponentially distributed with mean $E(T)=\rho/\lambda$ where λ is the Poisson claim rate and $\rho \epsilon (0,1)$ is a parameter. Then (for example, Kleinrock [17, p. 96])

$$Pr\,(A = n) = (1 - \rho)\rho^n,\ n = 0, 1, 2, \dots, \qquad (3.2.1)$$

that is, A is geometrically distributed. Then R has a compound geometric distribution (for example, Bowers et al. [2, p. 319]) with mgf

$$M_R(s) = \frac{1 - \rho}{1 - \rho M_X(s)}. \qquad (3.2.2)$$

From (1.3.4), the mean reported liability is

$$E(R) = \frac{\rho}{1 - \rho} E(X), \qquad (3.2.3)$$

and by using (1.3.5) the variance is

$$\text{Var}(R) = \frac{\rho}{1 - \rho} E(X^2) + \left\{\frac{\rho}{1 - \rho} E(X)\right\}^2. \qquad (3.2.4)$$

The distribution of R can be computed recursively (compare Panjer [23]). For example, if the single claim size distribution is discrete on the positive integers, then

$$f_R(x) = \rho \sum_{y=1}^{x} f_X(y) f_R(x - y), \tag{3.2.5}$$

which can be used to compute the distribution of R recursively, beginning with $f_R(0) = 1 - \rho$. In addition, (3.2.1) is of the form (1.3.20), implying that if there exists $\kappa > 0$ satisfying $M_X(\kappa) = \rho^{-1}$, then (1.3.21) yields

$$1 - F_R(x) \sim Ce^{-\kappa x}, \quad x \to \infty \tag{3.2.6}$$

where $C = (1 - \rho)/\{\rho(e^{\kappa} - 1)M'_X(\kappa)\}$ if X is discrete on the positive integers and $C = (1 - \rho)/\{\rho\kappa M'_X(\kappa)\}$ if X is continuous. Thus, under fairly general conditions, the distribution of R is asymptotically exponential. Numerical evaluation of κ and further discussion of this type of asymptotic result can be found in Willmot [35]. Numerical investigations indicate that the right side of (3.2.6) is an extremely good approximation to $1 - F_R(x)$ in a wide variety of situations. This suggests that a simple approximation of the amount needed to be adequate to cover the liability R a proportion α of the time can be obtained. Simply set $F_R(x) = \alpha$ in (3.2.6) and solve for x, yielding

$$\frac{1}{\kappa} \log \{C/(1 - \alpha)\} \tag{3.2.7}$$

as an approximation to the required value. The formula (3.2.7) can be used as a simple approximation to the exact procedure based on the recursive formula (3.2.5). An example is now presented, in which the numbers chosen are for illustration only.

Example 3.2.1

Consider the life portfolio of Example 2.1.1 where $\lambda = 4.27137$ and the single claim amount distribution is given by the first column in Table 2. Suppose studies indicate that the time from which notification of the claim reaches the insurer until payment is made (denoted by S) has an average of 1.5 months. It is known (for example, Kleinrock [17, p. 202]) that for this queueing system S is exponentially distributed with mean $E(S) = \rho/\{\lambda(1 - \rho)\}$. Hence $\rho = \lambda E(S)/\{1 + \lambda E(S)\}$. In this case $E(S) = 1/8$ and so $\rho = 0.348076$. From (3.2.3) and (3.2.4), the mean and variance of R are 4.65636 and 76.7905, respectively. Table 5 lists the exact distribution and corresponding df [obtained by using (3.2.5)], as well as the approximate df [denoted by $\bar{F}_R(x)$] from (3.2.6). In this case κ is easily found from $M_X(\kappa)=\rho^{-1}$ to be 0.101337.

TABLE 5

x	$f_R(x)$	$F_R(x)$	$\bar{F}_R(x)$	x	$f_R(x)$	$F_R(x)$	$\bar{F}_R(x)$
0	0.651924	0.651924	0.000000	36	0.001410	0.987232	0.987238
1	0.010781	0.662705	0.557139	37	0.001277	0.988509	0.988468
2	0.018585	0.681290	0.599818	38	0.001152	0.989661	0.989579
3	0.014797	0.696087	0.638384	39	0.000997	0.990658	0.990584
4	0.007555	0.703642	0.673234	40	0.000870	0.991528	0.991491
5	0.003480	0.707122	0.704725	41	0.000787	0.992315	0.992311
6	0.012999	0.720122	0.733181	42	0.000724	0.993039	0.993052
7	0.024131	0.744253	0.758894	43	0.000665	0.993705	0.993722
8	0.034669	0.778922	0.782130	44	0.000608	0.994313	0.994327
9	0.026646	0.805568	0.803126	45	0.000553	0.994866	0.994873
10	0.024525	0.830093	0.822099	46	0.000499	0.995364	0.995367
11	0.022514	0.852606	0.839244	47	0.000445	0.995809	0.995814
12	0.016107	0.868714	0.854736	48	0.000408	0.996218	0.996217
13	0.006612	0.875326	0.868735	49	0.000366	0.996584	0.996582
14	0.008413	0.883739	0.881386	50	0.000331	0.996914	0.996911
15	0.009666	0.893404	0.892817	51	0.000292	0.997206	0.997209
16	0.009170	0.902574	0.903146	52	0.000268	0.997474	0.997478
17	0.009339	0.911913	0.912480	53	0.000244	0.997717	0.997721
18	0.007924	0.919837	0.920914	54	0.000223	0.997940	0.997941
19	0.008072	0.927909	0.928536	55	0.000200	0.998140	0.998139
20	0.006829	0.934737	0.935423	56	0.000180	0.998320	0.998318
21	0.005687	0.940424	0.941646	57	0.000162	0.998482	0.998480
22	0.006920	0.947344	0.947270	58	0.000146	0.998628	0.998627
23	0.004696	0.952040	0.952352	59	0.000132	0.998759	0.998759
24	0.004163	0.956203	0.956944	60	0.000119	0.998879	0.998879
25	0.003226	0.959429	0.961093	61	0.000108	0.998987	0.998987
26	0.004735	0.964163	0.964843	62	0.000098	0.999084	0.999085
27	0.004008	0.968171	0.968231	63	0.000088	0.999173	0.999173
28	0.003679	0.971850	0.971292	64	0.000080	0.999253	0.999252
29	0.002581	0.974431	0.974059	65	0.000072	0.999325	0.999324
30	0.002407	0.976838	0.976559	66	0.000065	0.999390	0.999390
31	0.002112	0.978950	0.978818	67	0.000059	0.999449	0.999448
32	0.001893	0.980843	0.980859	68	0.000053	0.999502	0.999502
33	0.001759	0.982602	0.982704	69	0.000048	0.999550	0.999550
34	0.001692	0.984295	0.984371	70	0.000043	0.999593	0.999593
35	0.001528	0.985823	0.985877				

It is apparent from Table 5 that $\bar{F}_R(x)$ is an extremely good approximation to $F_R(x)$ even for small values of x, and (3.2.7) should provide a good approximation to the exact amount required to cover the liability a proportion α of the time, even for α as low as 0.75.

3.3 Several Claims Evaluators and Network Liability Models

In this section a more general model for the reported claim liability is proposed, whereby a more complex claims evaluation process is considered. In practice the assumption that there is a single claims evaluator who approves claims for payment may be inappropriate. For example, several individuals may be involved at various stages in the process. Also, it may be of interest to subdivide the reported claim liability for purposes of monitoring the process or even for financial reporting purposes. Exhibit 11 of the U.S. Annual Statement requires reported health claim liabilities to be subdivided into "Due and Unpaid" and "In Course of Settlement"; see O'Grady [21, p. 105] for further details.

To begin, the assumption in Section 3.2 that there is one claims evaluator is relaxed. Hence claims, which are reported according to a Poisson process with rate λ, are immediately evaluated by any one of c evaluators (if not busy) in the order in which they are reported. The time T required for one evaluator to process a claim is assumed to be exponentially distributed with mean $E(T)=\rho c/\lambda$, with $\rho\epsilon(0,1)$ a parameter. The claim is then paid immediately.

Note that this model can be used to help monitor the efficiency of the claims evaluation process. A parameter ρ, which represents the expected proportion of evaluators who are busy at one time, is of interest in this connection (compare Kleinrock [17, p. 18]). If this number is too large or too small, the amount of time available to perform other tasks may not be appropriate relative to the needs of the claims department. Assuming that the mean processing time $E(T)=\rho c/\lambda$ is constant, ρ varies inversely with c. The effect of a change in ρ of the number of evaluators c can therefore be ascertained. A second quantity of interest is the total time S from reporting until payment (that is, the total system time). Because S is the sum of the time spent waiting to begin evaluation plus the actual evaluation time, it follows from Tijms [30, p. 333] and the fact that $\rho c=\lambda E(T)$ that

$$E(S) = E(T) + \frac{E(T)\,\{\lambda E(T)\}^c}{(c-1)!\{c-\lambda E(T)\}^2} \times \left\{\frac{\{\lambda E(T)\}^c}{(c-1)!\{c-\lambda E(T)\}} + \sum_{k=0}^{c-1}\frac{\{\lambda E(T)\}^k}{k!}\right\}^{-1}. \quad (3.3.1)$$

Thus, if λ and $E(T)$ are assumed to be fixed, (3.3.1) can be viewed as a function of c, and the effect of a change in the number of evaluators c on the average processing time can be studied. A more detailed study of the quantity S is in Section 4.

To analyze the claim liability R, note that (for example, Tijms [30, p. 332])

$$f_A(0) = Pr(A = 0) = \left\{\frac{(\rho c)^c}{c!(1-\rho)} + \sum_{k=0}^{c-1} \frac{(\rho c)^k}{k!}\right\}^{-1} \tag{3.3.2}$$

and

$$f_A(n) = Pr\,(A = n) = \begin{cases} \dfrac{(pc)^n}{n!} f_A(0); & n = 0, 1, ..., c-1 \\ \dfrac{\rho^n c^c}{c!} f_A(0); & n = c, c+1, \ldots . \end{cases} \tag{3.3.3}$$

Thus the reported claim liability R has mgf

$$\begin{aligned} M_R(s) &= \sum_{n=0}^{\infty} f_A(n)\{M_X(s)\}^n \\ &= f_A(0)\left\{\sum_{n=0}^{c-1} \frac{(\rho c)^n}{n!}\{M_X(s)\}^n + \frac{(\rho c)^c}{c!(1-\rho)} M_*(s)\right\} \end{aligned} \tag{3.3.4}$$

where

$$M_*(s) = \left\{\frac{1-\rho}{1-\rho M_X(s)}\right\} \{M_X(s)\}^c. \tag{3.3.5}$$

Moments of R can be found from (3.3.4). For example, the mean is, using (1.3.4),

$$E(R) = E(X) f_A(0)\left\{\sum_{n=1}^{c-1} \frac{(\rho c)^n}{(n-1)!} + \frac{(\rho c)^c}{c!(1-\rho)}\left(c + \frac{\rho}{1-\rho}\right)\right\} \tag{3.3.6}$$

where the summation is 0 if $c = 1$.

To obtain the distribution of R, assume that $Pr(X=0)=0$ and so $f_R(0)=f_A(0)$. Supposing that $\{M_X(s)\}^n$ and $M_*(s)$ are the moment generating functions of the distributions $f_X^{*n}(x)$ and $f_*(x)$, respectively, from (3.3.4), for $x>0$,

$$f_R(x) = f_A(0)\left\{\sum_{n=1}^{c-1}\frac{(\rho c)^n}{n!}f_X^{*n}(x) + \frac{(\rho c)^c}{c!(1-\rho)}f_*(x)\right\}. \qquad (3.3.7)$$

Clearly, $f_X^{*n}(x)$ is the n-fold convolution of $f_X(x)$ with itself and can be obtained by using techniques described in Panjer [23, section 2.3], for example. The distribution $f_*(x)$ can be found recursively. In the case when $f_X(x)$ is discrete on the positive integers (a similar formula holds in the continuous case), for $x = 1, 2, 3, \ldots$,

$$f_*(x) = (1-\rho)f_X^{*c}(x) + \rho\sum_{y=1}^{x} f_X(y)f_*(x-y), \qquad (3.3.8)$$

beginning with $f_*(0)=0$. To see (3.3.8), note that (3.3.5) implies that

$$M_*(s) = (1-\rho)\{M_X(s)\}^c + \rho M_X(s)M_*(s).$$

One may equate coefficients of e^{sx} on both sides of this equation to give (3.3.8).

Consequently, it is straightforward to obtain $f_R(x)$ numerically. The convolutions $f_X^{*n}(x)$ for $n=1, 2, \ldots, c$ can be obtained successively. Then $f_*(x)$ can be obtained by using (3.3.8) and $f_R(x)$ from (3.3.7).

In addition, a simple asymptotic formula holds. From (3.3.3),

$$f_A(n) \sim \frac{c^c}{c!}f_A(0)\rho^n,\ n \to \infty, \qquad (3.3.9)$$

which is of the form (1.3.20). Thus, from (1.3.21), if there exists $\kappa>0$ satisfying $M_X(\kappa)=\rho^{-1}$, then (3.2.6) holds with

$$C = c^c f_A(0)/\{c!\rho(e^\kappa - 1)M_X'(\kappa)\}$$

if X is discrete on the positive integers and

$$C = c^c f_A(0)/\{c!\rho\kappa M_X'(\kappa)\}$$

if X is continuous. Thus, the asymptotic exponentiality of R holds for this more general model. This implies that the simple approximation (3.2.7) to the quantity that is adequate to cover R a proportion α of the time still holds, but with the above definition of C. A numerical example is presented to illustrate these techniques.

Example 3.2.2

The life portfolio of Example 2.1.1 is used, where $\lambda = 4.27137$ and the single claim amount distribution is given by the first column in Table 2. Suppose that there are three evaluators ($c = 3$) and the average processing time is 1¼ months (that is, $E(S) = 5/48$). Note that (3.3.1) can be rewritten as

$$\rho c + \frac{c^{c-1}\rho^{c+1}}{(c-1)!(1-\rho)^2}\left\{\sum_{k=0}^{c-1}\frac{(\rho c)^k}{k!} + \frac{(\rho c)^c}{c!(1-\rho)}\right\}^{-1} - \lambda E(S) = 0.$$

With $\lambda E(S)$ known, this is an implicit function of ρ, which is easily solved numerically by using a Newton-Raphson procedure (for example, Burden and Faires [5, section 2.3]). In this case, one finds easily that $\rho = 0.147681$. The mean and variance of R are 3.88030 and 46.3413, respectively. The distribution $f_R(x)$ obtained from (3.3.7) is given in Table 6, together with the df $F_R(x)$. With $\kappa = 0.162247$, the approximate df $\tilde{F}_R(x)$ obtained from (3.2.6) is also given with C as above.

The models of this and the previous section can be used in the more general setting of the claims evaluation process. This process may involve several functions such as verification of coverage, claim validation, and actual payment (compare O'Grady [21, chapter 7]). These functions can be done separately or in conjunction with one another. If done separately, models of this sort can often be used independently at each stage of the process.

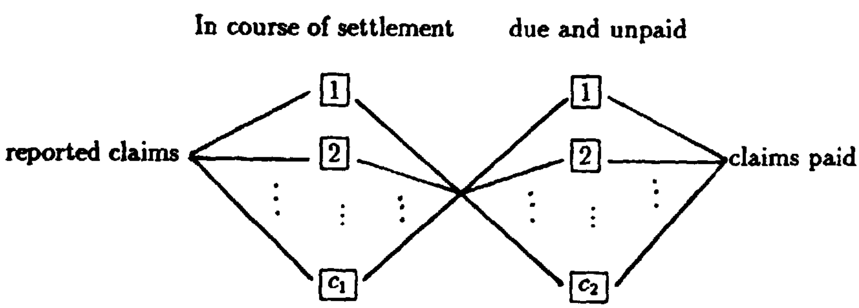
Suppose, for example, that there are two basic components of the claims evaluation process. Claims are reported to the insurer as before and queue up for evaluation and approval for payment by any one of c_1 available evaluators. Once approved, the claims are then routed to a second queue to await payment, and any one of c_2 individuals processes the claim for payment. The two stages are referred to as claims "In Course of Settlement" and "Due and Unpaid" and can be represented diagramatically as shown in Figure 1.

If it is assumed that the reported claims follow a Poisson process as before and processing time is exponential at each stage, then both the claim liability at each stage and the total time spent at each stage are independent of the corresponding quantity at the other stage (compare Kleinrock [17, section 4.8] and Burke [6]). That is, the model described earlier in this section for the reported claim liability can be applied to each stage independently. This provides a natural mechanism for the analysis of separate liabilities at each stage, because these are required to be reported separately for health claims

TABLE 6

x	$f_R(x)$	$F_R(x)$	$\hat{F}_R(x)$	x	$f_R(x)$	$F_R(x)$	$\hat{F}_R(x)$
0	0.641769	0.641769	0.000000	36	0.000560	0.997006	0.996981
1	0.013509	0.655278	0.116881	37	0.000479	0.997485	0.997433
2	0.023206	0.678484	0.249145	38	0.000409	0.997894	0.997818
3	0.018260	0.696744	0.361600	39	0.000317	0.998211	0.998145
4	0.009002	0.705746	0.457212	40	0.000246	0.998457	0.998423
5	0.003863	0.709609	0.538505	41	0.000209	0.998666	0.998659
6	0.015887	0.725496	0.607623	42	0.000188	0.998853	0.998860
7	0.029794	0.755290	0.666388	43	0.000167	0.999021	0.999030
8	0.042628	0.797918	0.716353	44	0.000145	0.999166	0.999176
9	0.031934	0.829852	0.758835	45	0.000124	0.999290	0.999299
10	0.028785	0.858637	0.794954	46	0.000105	0.999395	0.999404
11	0.026115	0.884752	0.825663	47	0.000087	0.999483	0.999493
12	0.018057	0.902809	0.851773	48	0.000081	0.999563	0.999569
13	0.006068	0.908877	0.873973	49	0.000067	0.999631	0.999634
14	0.008141	0.917017	0.892848	50	0.000058	0.999689	0.999689
15	0.009480	0.926497	0.908896	51	0.000044	0.999732	0.999735
16	0.008539	0.935036	0.922541	52	0.000039	0.999772	0.999775
17	0.008480	0.943516	0.934142	53	0.000035	0.999806	0.999809
18	0.006515	0.950031	0.944005	54	0.000031	0.999838	0.999837
19	0.006782	0.956814	0.952391	55	0.000025	0.999863	0.999862
20	0.005515	0.962329	0.959522	56	0.000021	0.999884	0.999882
21	0.004402	0.966731	0.965584	57	0.000017	0.999901	0.999000
22	0.006116	0.972847	0.970739	58	0.000015	0.999915	0.999915
23	0.003440	0.976286	0.975121	59	0.000012	0.999928	0.999928
24	0.002872	0.979158	0.978847	60	0.000011	0.999939	0.999939
25	0.001756	0.980914	0.982015	61	0.000009	0.999948	0.999948
26	0.003711	0.984625	0.984709	62	0.000008	0.999956	0.999956
27	0.002886	0.987511	0.986999	63	0.000007	0.999962	0.999962
28	0.002558	0.990070	0.988946	64	0.000006	0.999968	0.999968
29	0.001275	0.991344	0.990602	65	0.000005	0.999973	0.999973
30	0.001172	0.992516	0.992009	66	0.000004	0.999977	0.999977
31	0.000973	0.993489	0.993206	67	0.000003	0.999980	0.999980
32	0.000833	0.994322	0.994223	68	0.000003	0.999983	0.999983
33	0.000759	0.995081	0.995089	69	0.000002	0.999986	0.999986
34	0.000735	0.995816	0.995824	70	0.000002	0.999988	0.999988
35	0.000630	0.996446	0.996450	71	0.000002	0.999990	0.999990

FIGURE 1



in Exhibit 11 of the U.S. Annual Statement. Note that if there is no congestion at either stage, this may be accommodated by setting either c_1 or c_2 equal to infinity, and the corresponding liability model for that stage becomes the same model as that given in Section 2.1. The independence of the two stages still holds. For example, the payment stage may involve little or no congestion.

More general models can be employed for the reported claims process in which claims may be routed back and forth between various stages (as may occur if claims are resisted). If there is one evaluator at each stage, the liability attributable to each stage may be modeled by using the approach of Section 3.2, and the liabilities at each stage are independent of those at other stages. These network models are described, for example, in Kleinrock [17, section 4.8]. Note, however, that this independence does not hold in general for the total time spent in each stage, except in a few special cases such as that given in Figure 1. See Burke [6] for more details.

3.4 Arbitrary Processing Time

The assumption of an exponential distribution of processing time may not be reasonable in some situations. The processing time may not be exponential, or the mode of the processing time distribution may be greater than zero. Various tools are available even when this distribution is not exponential.

As in previous sections, the number of reported claims is assumed to be a Poisson process with rate λ, and claims are immediately processed by any one of c claims evaluators (if free), but the processing time has an arbitrary distribution. As before, let A denote the number of claims reported but unpaid (the number in the system) and S the total time from reporting until payment (the total processing time). Then the means of A and S are related by Little's formula (for example, Tijms [30, p. 262]), namely, $E(A) = \lambda E(S)$. Thus, because the mean reported but unpaid claims liabilities $E(R) = E(A)E(X)$ where $E(X)$ is the mean claim size, $E(R) = \lambda E(S)E(X)$. Because the expected annual incurred claims is $E(Y_1) = \lambda E(X)$,

$$E(R) = E(Y_1)E(S) \tag{3.4.1}$$

In words,

expected reported liability =

expected annual claims × *expected processing time.*

This intuitive result is analogous to that for the unreported claims in Section 2.1 and does not depend on the distribution of processing time. Evidently, the queueing theoretic approach provides an aid to intuition by generating a distribution about the mean.

Suppose now that the processing time for one claim is denoted by T with df $F_T(x)$ and mean $E(T)$. Define the df

$$F_1(x) = \int_0^x \left\{\frac{1 - F_T(t)}{E(T)}\right\} dt. \tag{3.4.2}$$

See Bowers et al. [2, Section 12.5] for a discussion of (3.4.2). Define the df's

$$F_k(x) = 1 - \{1 - F_1(x)\}^k \tag{3.4.3}$$

for $k=1$, c, and associated mixed Poisson pgf's

$$Q_k(S) = \sum_{m=0}^{\infty} q_m(k) s^m = \int_0^{\infty} e^{\lambda k c^{-1} t(s-1)} dF_k(t). \tag{3.4.4}$$

Then an approximation to the distribution of A is given in terms of its pgf as

$$\begin{aligned} P_A(s) &= \sum_{n=0}^{\infty} f_A(n) s^n \\ &= \sum_{n=0}^{c-1} f_A(n) s^n + \frac{\rho f_A(c-1)}{1-\rho} s^c Q_c(s) \left\{\frac{1-\rho}{1-\rho Q_1(s)}\right\} \end{aligned} \tag{3.4.5}$$

In (3.4.5), $\rho = \lambda E(T)/c$ and $f_A(n)$ is given by (3.3.3) for $n=0, 1, 2, \ldots$, $c-1$. This approximation for the equilibrium distribution of A is exact when T is exponential and when $c=1$ or $c=\infty$. It is derived in Section 4.4.3 of Tijms [30]. Also, various reasons for the high degree of accuracy are given in Miyazawa [20] in connection with equivalent mathematical problems.

The reported claim liability has mgf $M_R(s) = P_A\{M_X(s)\}$, and so from (3.4.5)

$$M_R(s) = \sum_{n=0}^{c-1} f_A(n) \{M_X(s)\}^n + \frac{\rho f_A(c-1)}{1-\rho} M_{\bullet}(s) \tag{3.4.6}$$

where

$$M_{\bullet}(s) = \left\{\frac{1-\rho}{1-\rho M_{\bullet 1}(s)}\right\} \{M_X(s)\}^c M_{\bullet c}(s) \tag{3.4.7}$$

and

$$M_{\bullet k}(s) = Q_k\{M_X(s)\} \tag{3.4.8}$$

for $k=1, c$. The analysis of the moments and distribution of R proceeds in the same manner as for the model in Section 3.3, because Equation (3.3.4) is similar in structure to (3.4.6). A complicating factor is the presence of the compound mgf $M_{\bullet k}(s)$ and the associated distribution $f_{\bullet k}(s)$, both of which are often awkward to deal with. An important exception to this observation is given in the following example.

Example 3.4.1

Suppose that $c=1$ and the processing time T has a distribution that is a mixture of gammas (Section 1.3) with integral index parameters, that is, has pdf

$$\frac{d}{dx}F_T(x) = \sum_{i=1}^{k} q_i \left\{\frac{\beta^{-i} x^{i-1} e^{-x/\beta}}{(i-1)!}\right\} \tag{3.4.9}$$

where $\{q_1, q_2, \ldots, q_k\}$ is itself a probability distribution. The density (3.4.9) is referred to as a generalized Erlangian distribution and is frequently used in queueing applications because of its flexibility of shape and convenient mathematical properties (Tijms [30, pp. 271–272, 397–400]). The mean is

$$E(T) = \beta \sum_{i=1}^{k} iq_i.$$

Also, by using formula 1.22 of Tijms [30, p. 18], the df is

$$F_T(x) = 1 - \sum_{i=1}^{k} q_i \sum_{j=1}^{i} \frac{(x/\beta)^{j-1} e^{-x/\beta}}{(j-1)!}.$$

Interchanging the order of summation, the density corresponding to (3.4.2) is

$$\frac{d}{dx}F_1(x) = \frac{1-F_T(x)}{E(T)} = \sum_{j=1}^{k} q_j^{\bullet} \left\{\frac{\beta^{-j} x^{j-1} e^{-x/\beta}}{(j-1)!}\right\} \tag{3.4.10}$$

where

$$q_j^* = \left\{\sum_{i=j}^{k} q_i\right\} \Big/ \left\{\sum_{i=1}^{k} iq_i\right\};\ j = 1, 2, \ldots, k. \qquad (3.4.11)$$

Because $\Sigma_{j=1}^{k} q_j^* = 1$, (3.4.10) is of the same form as (3.4.9), but with different weights, that is, also a mixture of gamma distributions. Then, by using (3.4.4) with $k=1$ and $c=1$, (1.3.15), and (1.3.7),

$$Q_1(s) = \sum_{j=1}^{k} q_j^* \{1 - \lambda\beta(s - 1)\}^{-j}, \qquad (3.4.12)$$

a mixture of negative binomial pgf's. Thus, from (3.4.8),

$$M_{\cdot 1}(s) = \sum_{j=1}^{k} q_j^* \{1 - \lambda\beta[M_X(s) - 1]\}^{-j}, \qquad (3.4.13)$$

a mixture of compound negative binomials. Evaluation of the moments is straightforward by using (3.4.13), and the distribution $f_{\cdot 1}(x)$ can be evaluated recursively by using the techniques in Panjer [23]. Analysis of the distribution and moments of R follows easily by using (3.4.6), (3.4.7), and (3.4.8) with $c=1$.

Although the computational difficulties associated with the evaluation of the distribution of R may be overwhelming for arbitrary c and processing time distribution $F_T(x)$, some asymptotic help is available. From Tijms [30, p. 351], if there exists $\tau>1$ satisfying $Q_1(\tau)=\rho^{-1}$, then

$$f_A(n) \sim \frac{\tau^{c-1} f_A(c - 1)\, Q_c(\tau)}{Q_1'(\tau)}\, \tau^{-n},\ n \to \infty. \qquad (3.4.14)$$

This is clearly of the form (1.3.20), and so if there exists $\kappa>0$ satisfying $M_X(\kappa)=\tau$, then one obtains from (1.3.21) an asymptotic approximation of the form

$$1 - F_R(x) \sim Ce^{-\kappa x},\ x \to \infty. \qquad (3.4.15)$$

Thus the tail of the distribution of the reported claim liability is asymptotically exponential even for this fairly general model. As mentioned previously, this yields a simple approximation for the amount needed to be adequate for a proportion α of the time, namely,

$$\kappa^{-1} \log\{C/(1 - \alpha)\}. \qquad (3.4.16)$$

In (3.4.15) and (3.4.16), the constant C is given by

$$\tau^c f_A(c - 1)Q_c(\tau)/\{(e^\kappa - 1)Q_1'(\tau)M_X'(k)\}$$

if X is discrete and

$$\tau^c f_A(c - 1)Q_c(\tau)/\{\kappa Q_1'(\tau)M_X'(\kappa)\}$$

if X is continuous. The assumption that there exists $\tau > 1$ satisfying $Q_1(\tau) = \rho^{-1}$ is essentially the assumption that there exists an adjustment coefficient using a ruin theoretic interpretation. This issue is discussed in some detail in Bowers et al. [2, section 12.3], who point out that there usually does exist such a quantity. To see this interpretation, note that from (3.4.4) and (3.4.2),

$$Q_1(s) = \int_0^\infty e^{\lambda c^{-1} t(s-1)} \left\{\frac{1 - F_T(t)}{E(T)}\right\} dt,$$

and (1.3.15) together with formula (12.5.4) of Bowers et al. [2, p. 360] implies that

$$Q_1(s) = \frac{M_T\{\lambda c^{-1}(s - 1)\} - 1}{\rho(s - 1)} \tag{3.4.17}$$

where $M_T(s)$ is the mgf of the processing time T. Thus, $Q_1(\tau) = \rho^{-1}$ is equivalent to $M_T\{\lambda c^{-1}(\tau - 1)\} = \tau$. In other words, one needs to find $\phi > 0$ satisfying

$$M_T(\phi) = 1 + \rho^{-1}E(T)\phi, \tag{3.4.18}$$

and then $\tau = 1 + c\phi/\lambda$. Examination of (3.4.18) and section (12.3) of Bowers et al. [2] reveals that ϕ is simply the adjustment coefficient in a ruin theoretic context with "single claim size" random variable T and relative security loading $(1 - \rho)/\rho$.

Thus, in most instances there will exist $\tau > 1$ satisfying $Q_1(\tau) = \rho^{-1}$, and so (3.4.14), (3.4.15), and (3.4.16) will be applicable in general. In particular, τ will always exist if τ has a gamma distribution, or more generally, the pdf (3.4.9). Occasionally, however, this will not be the case. Consider, for example, the inverse Gaussian distribution. If $M_T(s)$ is given by (1.3.10), then $M_T(s) \leq e^\mu$, and because $E(T) = \mu\beta/2$ in this case, it is evident from (3.4.23) that no such τ will exist if $e^\mu < 1 + \mu/(2\rho)$, that is, if $\rho < \mu/\{2(e^\mu - 1)\}$. In this case and in some other situations, an alternative asymptotic formula to (3.4.14) and (3.4.15) is available. This is stated as a theorem.

Theorem 3.4.1

Suppose that the processing time df satisfies

$$1 - F_T(x) \sim Kx^{\alpha} e^{-\beta x}, \quad x \to \infty, \tag{3.4.19}$$

where $K>0$, $\alpha<-1$, and $\beta>0$, and $Q_1 \left\{\frac{\lambda+\beta c}{\lambda}\right\}<\rho^{-1}$. Then if $c=1$,

$$f_A(n) \sim \frac{K(1-\rho)}{(\lambda+\beta)^{\alpha}\left\{1-\rho Q_1\left(\frac{\lambda+\beta}{\lambda}\right)\right\}^2} n^{\alpha}\left(\frac{\lambda}{\lambda+\beta}\right)^n, \; n \to \infty, \tag{3.4.20}$$

whereas if $c>1$,

$$f_A(n) \sim \frac{K\rho c^{\alpha} f_A(c-1)\, Q_c\left(\frac{\lambda+\beta c}{\lambda}\right)}{\lambda^{c-1}(\lambda+\beta c)^{\alpha-c+1}\left\{1-\rho Q_1\left(\frac{\lambda+\beta c}{\lambda}\right)\right\}^2} n^{\alpha} \times \left(\frac{\lambda}{\lambda+\beta c}\right)^n, \; n \to \infty. \tag{3.4.21}$$

Proof: The case (3.4.20) with $c=1$ is proved in Willmot [35]. Hence assume $c>1$ and it is of interest to prove (3.4.21). Note that the density corresponding to (3.4.3) is

$$\frac{d}{dx} F_k(x) = \frac{k}{E(T)}\{1 - F_T(x)\}\{1 - F_1(x)\}^{k-1}. \tag{3.4.22}$$

L'Hospital's rule yields

$$\begin{aligned}
\lim_{x\to\infty} \frac{x^{\alpha} e^{-\beta x}}{1 - F_1(x)} &= \lim_{x\to\infty} \frac{\alpha x^{\alpha-1} e^{-\beta x} - \beta x^{\alpha} e^{-\beta x}}{-\{1 - F_T(x)\}/E(T)} \\
&= E(T) \lim_{x\to\infty}\left\{\frac{x^{\alpha} e^{-\beta x}}{1 - F_T(x)}\right\}\left\{\beta - \frac{\alpha}{x}\right\} \\
&= \frac{\beta E(T)}{K}.
\end{aligned}$$

In other words

$$1 - F_1(x) \sim \frac{K}{\beta E(T)} x^{\alpha} e^{-\beta x}, \; x \to \infty,$$

and so

$$\frac{d}{dx} F_k(x) \sim k\beta \left\{\frac{K}{\beta E(T)}\right\}^k x^{k\alpha} e^{-k\beta x}, \; x \to \infty.$$

Thus, (3.4.4) and (1.3.19) yield

$$q_n(k) \sim \frac{\beta c}{\lambda + \beta c} \left\{\frac{K c^{\alpha - 1} \lambda}{\rho\beta k^{\alpha} (\lambda + \beta c)^{\alpha}}\right\}^k n^{k\alpha} \left(\frac{\lambda}{\lambda + \beta c}\right)^n, \quad n \to \infty. \quad (3.4.23)$$

Define

$$J(s) = \sum_{n=0}^{\infty} j_n s^n = (1 - \rho)/\{1 - \rho Q_1(s)\}.$$

Then, because

$$\rho^{-1} > Q_1 \left\{\frac{\lambda + \beta c}{\lambda}\right\},$$

lemma 2 in Willmot [35] yields

$$j_n \sim \rho(1 - \rho) \left\{1 - \rho Q_1 \left(\frac{\lambda + \beta c}{\lambda}\right)\right\}^{-2} q_n(1), \; n \to \infty.$$

In other words,

$$j_n \sim \frac{K c^{\alpha} \lambda(1 - \rho)}{(\lambda + \beta c)^{\alpha+1} \left\{1 - \rho Q_1 \left(\frac{\lambda + \beta c}{\lambda}\right)\right\}^2} n^{\alpha} \times \left(\frac{\lambda}{\lambda + \beta c}\right)^n, \; n \to \infty. \quad (3.4.24)$$

Now define

$$H(s) = \sum_{n=0}^{\infty} h_n s^n = J(s) Q_c(s).$$

Because $c>1$ and $\alpha<-1$, it is clear from (3.4.23) and (3.4.24) that $\lim_{n\to\infty} q_n(c)/j_n=0$. Corollary 6.1 of Meir and Moon [19] then yields that

$$h_n \sim Q_c\left\{\frac{\lambda+\beta c}{\lambda}\right\} j_n,\ n\to\infty.$$

Thus, using (3.4.24),

$$h_n \sim \frac{K c^\alpha \lambda(1-\rho)Q_c\left(\frac{\lambda+\beta c}{\lambda}\right)}{(\lambda+\beta c)^{\alpha+1}\left\{1-\rho Q_1\left(\frac{\lambda+\beta c}{\lambda}\right)\right\}^2} n^\alpha \left(\frac{\lambda}{\lambda+\beta c}\right)^n,\ n\to\infty.$$

But from (3.4.5),

$$f_A(n) = \rho(1-\rho)^{-1} f_A(c-1)h_{n-c} \text{ for } n>c$$

and so (3.4.21) results.□

Note that Theorem 3.4.1 yields an asymptotic expression of the form (1.3.22), namely, $f_A(n)\sim K_1 n^\alpha\,[\lambda/(\lambda+\beta c)]^n$, $n\to\infty$, where K_1 varies depending on whether c is greater than or equal to 1. As a simple corollary to the theorem, from (1.3.21) is obtained the asymptotic expression for the reported liability df

$$1-F_R(x)\sim C_2 x^\alpha e^{-\kappa x},\ x\to\infty. \tag{3.4.25}$$

In this expression C_2 varies both as the claim size distribution is discrete or continuous and as c is greater than or equal to 1. In any event, C_2 is easily obtained from the theorem and the discussion immediately following (1.3.21). Also, κ in (3.4.25) satisfies $M_X(\kappa)=(\lambda+\beta c)/\lambda$.

Consider the class of distributions satisfying (3.4.19). Now, $M_T(s)<\infty$ for $s\leq\beta$, and $M_T(\beta)<\infty$. Thus from (3.4.17), $Q_1(s)<\infty$ for $s\leq(\lambda+\beta c)/\lambda$ and $Q_1\,[(\lambda+\beta c)/\lambda]<\infty$. There will exist $\tau>1$ satisfying $Q_1(\tau)=\rho^{-1}$ if $Q_1\,[(\lambda+\beta c)/\lambda]\geq\rho^{-1}$ (that is, if $M_T(\beta)\geq1+\rho^{-1}\,E(T)\beta$). But Theorem 3.4.1 holds if $Q_1[(\lambda+\beta c)/\lambda]<\rho^{-1}$, and so one of the two asymptotic results will hold, namely, (3.4.14) or one of (3.4.20) or (3.4.21).

The inverse Gaussian pdf (1.3.9) satisfies

$$f(x)\sim\left\{\frac{\mu e^\mu}{2}\left(\frac{\beta}{\pi}\right)^{1/2}\right\} x^{-3/2}\,e^{-x/\beta},\quad x\to\infty,$$

and L'Hospital's rule yields

$$1 - F(x) \sim \left\{ \frac{\beta\mu e^{\mu}}{2} \left(\frac{\beta}{\pi}\right)^{1/2} \right\} x^{-3/2} e^{-x/\beta}, \quad x \to \infty. \qquad (3.4.26)$$

The relation (3.4.26) is clearly of the form (3.4.19) with $\alpha = -3/2$, β replaced by β^{-1}, and $K = \beta\mu e^{\mu}(\beta/\pi)^{1/2}/2$. Thus, if T has the inverse Gaussian pdf, (3.4.14) will hold if $e^{\mu} \geq 1 + \mu/(2\rho)$, but if $e^{\mu} < 1 + \mu/(2\rho)$, Theorem 3.4.1 applies.

Although the model of this section is more complex, it nevertheless provides some insight into the distributional behavior of R in a more general situation.

4. THE ANALYSIS OF DELAYS

4.1 Introduction

A quantity of interest to both the insured and the insurer is the length of time for processing and approving a claim for payment. We will ignore partial payments made prior to final settlement. The insured normally is interested in the total delay between the time of incurral of the claim and the time of receipt of payment, whereas the insurer is concerned with the time from receipt of notification of the claim until approval or payment. Because the time from incurral to receipt of notification is outside the insurer's control, this quantity is not of interest for analysis of the system's efficiency. In group insurance, this efficiency is one of the more important parameters involved in the decision of policyholders to place their business with a particular insurer. Hence, the time to process a claim is clearly a quantity of interest to the insurer.

Although the average processing time is certainly important, it is not sufficient for proper evaluation of the system's efficiency, because it does not allow for variability. For example, it does not account for variations in the time for processing a particular claim or in the delay due to an increased volume of incurred claims. A queueing approach can incorporate these quantities into the model. It is important to be able to assess whether a long delay in payment of a claim is reasonable in light of this variability. Clearly, improvement of the system's efficiency might be deemed appropriate if delays are too long.

4.2 Exponential Processing Models

Consider first the time S between receipt of notification of the claim and final approval of the claim for payment. For the single claims evaluator model of Section 3.2, Example 3.2.1 pointed out that S has an exponential distribution with mean $\rho/\{\lambda(1-\rho)\}$. Hence, because $\rho=\lambda E(T)$ where T is the time required to approve one claim, the distribution of S is given explicitly by

$$F_S(x) = 1 - e^{-\left\{\frac{1}{E(T)} - \lambda\right\}x}, \quad x > 0. \tag{4.2.1}$$

It follows at once that

$$\mathrm{Var}(S) = \{E(S)\}^2 = \left\{\frac{E(T)}{1 - \lambda\, E(T)}\right\}^2. \tag{4.2.2}$$

Thus, (4.2.1) and (4.2.2) give two simple measures of the variability in S.

As pointed out in Section 3.3, the model with c claims evaluators may be of more interest to the insured because the distribution of S and its moments can be modified by a change in c, a parameter that is under the control of the insurer. In this case (in the notation of Section 3.3) the distribution of S is given by (compare Gross and Harris [11], p. 91)

$$F_S(x) = 1 - (1 - \theta)e^{-\frac{x}{E(T)}} - \theta e^{-\left\{\frac{c}{E(T)} - \lambda\right\}x}, \quad x > 0, \tag{4.2.3}$$

where

$$\theta = \frac{1 - \sum_{n=0}^{c-1} f_A(n)}{1 - c + \lambda E(T)} \tag{4.2.4}$$

and for $n=1, 2, \ldots, c-1$,

$$f_A(n) = \frac{\{\lambda E(T)\}^n}{n!}\left(\frac{\{\lambda E(T)\}^c}{(c-1)!\,\{c - \lambda E(T)\}} + \sum_{k=0}^{c-1}\frac{\{\lambda E(T)\}^k}{k!}\right)^{-1}. \tag{4.2.5}$$

From (4.2.3),

$$E(S) = (1 - \theta)E(T) + \theta\left\{\frac{E(T)}{c - \lambda E(T)}\right\} \tag{4.2.6}$$

and

$$E(S^2) = 2(1 - \theta)\{E(T)\}^2 + 2\theta\left\{\frac{E(T)}{c - \lambda E(T)}\right\}^2, \qquad (4.2.7)$$

with $\text{Var}(S) = E(S^2) - \{E(S)\}^2$.

For analysis, it is convenient to express (4.2.3) through (4.2.7) in terms of λ, $E(T)$, and c. This is because the claims incurral rate λ and the mean processing time $E(T)$ would normally be beyond the control of the insurer, but the number of evaluators c is under the control of the insurer. Thus, as discussed in Section 3.3, the effect on the distribution of S of a change in the value of c may be ascertained. The following example illustrates this point.

Example 4.2.1

Consider the situation of Example 3.2.2 with $c = 3$ claims evaluators, $\lambda = 4.27137$, and $\rho = 0.147681$. Then the mean processing time of one claim is $E(T) = c\rho/\lambda = 0.103724$. It is a simple matter to evaluate $f_A(n)$ for $n = 1, 2, \ldots, c - 1$ by using (4.2.5). Then, from (4.2.4), $\theta = -0.00700951$. The mean and variance of S are 0.104167 (= 5/48, see Example 3.3.2) and 0.0109797, respectively, obtained by using (4.2.6) and (4.2.7). The df $F_S(x)$ from (4.2.3) is

$$F_S(x) = 1 - 1.00701e^{-9.64097x} + 0.00701e^{-24.6516x}.$$

Thus, for example $F_S(0.145) \approx 0.75$, implying that about 75 percent of the claims could be expected to take no more than 0.145 of a year to be approved (and 25 percent would take more than this length of time).

The effect on S of hiring or releasing claims evaluators can be evaluated by varying c but keeping λ and $E(T)$ constant. In this situation, for example, the effect of releasing one evaluator can be determined by reworking the calculation with $c = 2$. Then $\theta = -0.144258$, $E(S) = 0.109077$, and $\text{Var}(S) = 0.014050$. The fact that $E(S)$ increases only slightly for the case when $c = 3$ reflects the fact that ρ is quite small, and so there is little congestion. Note that the variability has increased, probably reflecting the fact that increased congestion has a greater effect with fewer evaluators. Finally, in this case

$$F_S(x) = 1 - 1.14426^{-9.64097x} + 0.14426e^{-15.0106x}.$$

One finds that $F_S\,(0.155) \approx 0.75$, that is, an increase from 0.145 to 0.155 of the 75th percentile from the case $c = 3$, again agreeing with intuition.

Note that the processing times in each stage of the two-stage network model described in Section 3.3 are independent of each other, and each is distributed as described above. The total processing time has a distribution that is the convolution of two distributions, each with df of the form (4.2.3). This independence does not hold for the more general network models (compare Burke [6]). Similarly, the total delay from the claimant's standpoint is simply the convolution of the distribution of S described above with that of B, the time from incurral to reporting, as described in Section 2. In these and other models, the distribution of interest involves convolutions of exponentials with different means. Rather than enumerate all possibilities, it suffices to point out that the sum of k independent exponentials with different means has mgf of the form

$$M(s) = \prod_{i=1}^{k} \left\{ \frac{\mu_i}{\mu_i - s} \right\} \tag{4.2.8}$$

and pdf

$$f(x) = \sum_{i=1}^{k} q_i \, \mu_i \, e^{-\mu_i x} \tag{4.2.9}$$

where, for $i = 1, 2, \ldots, k$,

$$q_i = \prod_{\substack{j=1 \\ j \neq i}}^{k} \{\mu_j / (\mu_j - \mu_i)\}. \tag{4.2.10}$$

The μ_i's are all assumed to be distinct in this formula. See Everitt and Hand [9, p. 79], for further references. In this situation as well as others, this result allows for a simple derivation of the distribution of interest and associated moments.

4.3 More General Delay Models

For situations not involving exponential processing models, the total delay distributions are more complex. However, a common underlying mathematical structure can be exploited to provide a unified treatment of the various delay distributions of interest to the insurer and the policyholder.

To begin, consider the model of Section 3.4 with c claims evaluators processing claims, with the time required to process one claim given by a generic variable T and with distribution $F_T(x)$. The notation of Section 3.4

will be used. Recall (Bowers et al. [2, p. 360]) that the random variable with pdf $F_1'(x)=\{1-F_T(x)\}/E(T)$ has mgf

$$M_1(s) = \frac{M_T(s) - 1}{sE(T)}. \tag{4.3.1}$$

Because

$$M_T(s) = \sum_{k=0}^{\infty} \frac{E(T^k)}{k!} s^k,$$

the moments of the distribution with df $F_1(x)$ are given by

$$M_1^{(k)}(0) = E(T^{k+1})/\{(k + 1)E(T)\}. \tag{4.3.2}$$

If we denote the delay random variable of interest to be W, the distribution of W is most easily characterized by its mgf, which is of the mixture form

$$M_W(s) = \theta M_{W_1}(s) + (1 - \theta)M_{W_2}(s) \tag{4.3.3}$$

where $\theta = 1 - \sum_{n=0}^{c-1} f_A(n)$,

$$M_{W_1}(s) = \frac{1 - \rho}{1 - \rho M_1(s/c)} M_{W_3}(s), \tag{4.3.4}$$

and the mgf's $M_{W_2}(s)$ and $M_{W_3}(s)$ are selected so that W represents the desired quantity.

To identify W_1 and W_2, suppose first that W is the time S between receipt of notification of the claim and approval for payment. Van Hoorn [33, p.37] shows that the time from receipt of notification until actual processing of the claim begins has mgf of the form $1-\theta+\theta M_{W_1}(s)$, where $M_{W_3}(s)$ is the mgf $M_c(s)$ of the random variable with df $F_c(x)$ given by (3.4.3). Thus S is obtained by convolving this distribution with that of T, the processing time. In other words, S has mgf of the form (4.3.3) with $M_{W_2}(s)=M_T(s)$ and $M_{W_3}(s)=M_c(s)M_T(s)$. From the policyholder's standpoint, $W=S+B$ where B is the time from incurral to reporting. Hence in this case W is still of the form (4.3.3) with $M_{W_2}(s)=M_T(s)M_B(s)$ and $M_{W_3}(s)=M_c(s)M_T(s)M_B(s)$.

The representation (4.3.3) allows for evaluation of the moments of W by differentiation. In general, the moments $M_c^{(k)}(0)$ may be difficult to evaluate, but if $c=1$, then (4.3.2) may be used so there is no difficulty, as long as the moments of T (and perhaps B) can be obtained.

Evaluation of the distribution of W is also complicated in general, primarily because of the presence of the distribution of W_1 with mgf (4.3.4). The tail of the distribution may be asymptotically exponential, however. Notice that (4.3.4) can be expressed as

$$M_{W_1}(s) = \rho M_1(s/c)M_{W_1}(s)+(1-\rho)M_{W_3}(s).$$

Assuming that W_3 is continuous, $\rho=\lambda E(T)/c$ implies

$$f_{W_1}(x) = \lambda \int_0^x \{1 - F_T(cy)\} f_{W_1}(x-y)dy + (1-\rho)f_{W_3}(x). \quad (4.3.5)$$

This relation is useful because it may sometimes be solved numerically for $f_{W_1}(x)$, being a Volterra integral equation (compare Ströter [28]). Also, it is a defective renewal equation (for example, Gerber [10, chapter 8]). To see this, note that the mgf $M_1(s/c)$ is associated with the pdf

$$f_1(x) = \frac{c}{E(T)}\{1 - F_T(cx)\}, \quad (4.3.6)$$

and (4.3.5) can be expressed as

$$f_{W_1}(x) = \rho \int_0^x f_1(y)f_{W_1}(x-y)dy + (1-\rho)f_{W_3}(x). \quad (4.3.7)$$

Thus, if there exists $\kappa>0$ satisfying

$$M_1(\kappa/c) = \rho^{-1}, \quad (4.3.8)$$

then (4.3.7) satisfies

$$e^{\kappa x} f_{W_1}(x) = \int_0^\infty \{\rho e^{\kappa y} f_1(y)\}\{e^{\kappa(x-y)} f_{W_1}(x-y)\}dy + (1-\rho)e^{\kappa x} f_{W_3}(x).$$

By (4.3.8), this is an ordinary renewal equation, and by the renewal theorem (compare Karlin and Taylor [14, p. 191]),

$$\lim_{x\to\infty} e^{\kappa x} f_{W_1}(x) = \frac{(1-\rho)\int_0^\infty e^{\kappa x} f_{W_3}(x)dx}{\rho\int_0^\infty x e^{\kappa x} f_1^*(x)dx}.$$

In other words,

$$f_{W_1}(x) \sim \frac{c(1-\rho)M_{W_3}(\kappa)}{\rho M_1'(\kappa/c)} e^{-\kappa x}, \quad x \to \infty,$$

and because asymptotic expressions can be integrated,

$$1 - F_{W_1}(x) \sim \frac{c(1-\rho)M_{W_3}(\kappa)}{\rho\kappa M_1'(\kappa/c)} e^{-\kappa x}, x \to \infty.$$

Finally, if $M_{W_2}(\kappa) < \infty$, then $e^{\kappa x}\{1-F_{W_2}(x)\}\to 0$ as $x\to\infty$, and so (4.3.3) yields

$$1 - F_W(x) \sim Ke^{\kappa x}, \quad x \to \infty, \tag{4.3.9}$$

where the constant K is given by

$$K = \frac{c\theta(1-\rho)M_{W_3}(\kappa)}{\rho\kappa M_1'(\kappa/c)}. \tag{4.3.10}$$

Evidently, (4.3.9) demonstrates that the distribution of W is asymptotically exponential under these conditions, which provides qualitative insight into its behavior. Note that (4.3.1) and (4.3.8) combine to yield an alternative definition of the decay parameter κ in (4.3.9), namely,

$$M_T(\kappa/c) = 1 + \rho^{-1}\left\{\frac{E(T)}{c}\right\}\kappa. \tag{4.3.11}$$

Equation (4.3.11) reveals, upon examination of Section 12.3 of Bowers et al. [2], that κ is the adjustment coefficient in a ruin theoretic context with "single claim size" mgf $M_T(s/c)$ and relative security loading $(1-\rho)/\rho$. This is analogous to the condition for the asymptotic exponentiality of the reported claim liability distribution of Section 3.4, as is discussed following (3.4.18).

5. CONCLUSIONS AND AREAS FOR FURTHER RESEARCH

This paper presents a cohesive and comprehensive modeling approach to the analysis of the claims payment process, relying heavily on risk and queueing theoretic techniques to account for the effects of statistical variation in both claims incurral and processing. A unique feature of the approach is the attempt to incorporate the effects of increased congestion of claims on the reported claims process.

Section 1 describes both the nature of the problem and the relevant risk theoretic background. In particular, the number of incurred claims is assumed to be a Poisson process, and some of its properties are described.

The unreported claim liability is the topic of Section 2. First, a compound Poisson model is proposed that requires knowledge only of the average reporting delay as well as the usual incurred claims information. The mean unreported liability is consistent with intuition, higher moments such as the variance are easily obtained, and the entire distribution can be calculated recursively with a computer. This allows one to choose the amount adequate to cover the liabilities with a specified probability, an approach suggested by Bragg [3]. Second, a generalization is presented that reflects differences in reporting patterns, while retaining the advantages of the compound Poisson form. Finally, a much more general model is discussed, which allows for a great deal of flexibility with respect to realistic phenomena, such as seasonality of incurred claims, business growth, heterogeneity of risk levels in the portfolio, inflation, and variability of reporting patterns. The added expense of the generalizations is more complicated mathematics, but the compound form of the unreported liability distribution is retained. Many of the various desirable ramifications are discussed.

The reported claim liability is considered in Section 3. First, it is shown quite generally that the reported and unreported claim liabilities are statistically independent of each other, implying that they can be analyzed separately. Also, the number of reported claims is shown to be approximately a Poisson process, which facilitates the use of queueing techniques. A compound geometric model for the unreported liability is proposed in Section 3.2 under the assumption that one claims evaluator processes claims in the order in which they are reported and that the processing times required for each claim are independent and exponentially distributed. The reported claim liability distribution can be evaluated recursively on a computer, and a simple exponential approximation for the right tail allows for a simple estimate of the amount needed to cover the liability for a fixed proportion of the time.

A somewhat more complex model involving several claims evaluators is described in Section 3.3. Although the computational details are slightly more onerous, there is little difficulty calculating the moments and the distribution (the latter recursively), and an exponential tail approximation for the right tail of the reported claim liability distribution is still available. These models can be combined to describe more complex evaluation systems through the use of networks, and a two-stage model representing claims "In Course of Settlement" and "Due and Unpaid" is outlined. Finally, an arbitrary processing time distribution and several evaluators are assumed. This general model reproduces an intuitively appealing mean reported claim liability. Although this model is more mathematically complex, the right tail of the reported claim liability distribution is still approximately exponential under fairly general conditions. In fact, these conditions are shown to be essentially those for the existence of the adjustment coefficient in ruin theory. An alternative asymptotic formula is given for some situations in which the exponential form does not hold.

The analysis of the delays in processing claims for payment is the subject of Section 4, in which it is argued that this is an important tool in the analysis of the efficiency of the claims evaluation system. The situation involving exponential processing times yields relatively simple moments and distributions of the delays, as shown in Section 4.2. In the more general formulation of Section 4.3, an expression is given for the moment generating function of the delay distribution, and it is shown that this formulation may represent different time periods of interest to the policyholder and the insurer. An exponential tail approximation for the delay is then derived, again under essentially the same conditions as those underlying the existence of the adjustment coefficient of ruin theory.

The paper describes a general approach to modeling the claims payment process and provides a basic set of quantitative tools for a variety of situations. Although the use of network models discussed in Section 3 provides an important framework within which quite complicated claims processing systems can be modeled, certain situations require more complicated models.

One such situation involves the possibility of resisted claims, which can often be dealt with through a redefinition of the single claim size distribution. Suppose, for example, that a proportion p of claims are assumed to be ultimately not paid. Then the single claim amount distribution $f_X(x)$ could be replaced by one of the form $p+(1-p)\, f_X(x)$, where $f_X(x)$ is now interpreted as the distribution of the amount payable given that something is payable. Even in the more complicated situation with partial payments, past experience data could possibly be used to construct a distribution of the amount actually paid (if the data available do not already reflect this). A more difficult problem involves situations in which the size of the claim cannot be ascertained at the time of incurral and is not independent of the processing time, and the approach of this paper may be unsatisfactory. Note that these may be the same situations in which the standard model of risk theory is also unsuitable, however.

One other possible feature involves the queueing mechanism assumed in the liability of reported claims. Rather than working on one particular claim until it is approved for payment, an evaluator may work on other claims (or even other types of insurance) while other work is done on the original claim or other information is obtained. Thus several claims may be processed simultaneously by a single evaluator. The use of network models may be appropriate here because the claims could be routed to another queue and then returned after additional information is obtained. A second possibility is to formulate a model in which the time of the evaluator is "shared" by several claims in the course of being approved. A simple method for incorporating this feature would be to assume that the evaluator acts like several evaluators, one for each claim.

There may be other features that one may wish to incorporate into the model for the liability of reported claims, and a queueing approach provides a systematic and unified methodology that can be utilized in a wide variety of situations.

As with other models such as Holsten's [13], the model for reported claims assumes equilibrium has been reached, and removal of this assumption may be both desirable and difficult. Nevertheless, we hope that this approach provides valuable insight into the claims payment process and, in particular, to claim liabilities and the delays involved.

REFERENCES

1. BARNHART, E.P. "A New Approach to Premium, Policy and Claim Reserves for Health Insurance," *TSA* 37 (1985): 13–95.
2. BOWERS, N., GERBER, H., HICKMAN, J., JONES, D., AND NESBITT, C. *Actuarial Mathematics*. Itasca, Ill.: Society of Actuaries, 1986.
3. BRAGG, J. "Health Insurance Claim Reserves and Liabilities," *TSA* 16 (1964): 17–54.
4. BÜHLMANN, H. *Mathematical Methods in Risk Theory*. New York: Springer Verlag, 1970.
5. BURDEN, R., AND FAIRES, J. *Numerical Analysis*, 3rd ed. Boston: Prindle, Weber, and Schmidt, 1985.

6. BURKE, P. "The Dependence of Sojourn Times in Tandem M/M/s Queues," *Operations Research* 17 (1969): 754–55.
7. CHHIKARA, R., AND FOLKS J. *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*. New York: Marcel Dekker, 1989.
8. EMBRECHTS, P., MAEJIMA, M., AND TEUGELS, J. "Asymptotic Behaviour of Compound Distributions," *ASTIN Bulletin* 15 (1985): 45–48.
9. EVERITT, B., AND HAND, D. *Finite Mixture Distributions*. London: Chapman and Hall, 1981.
10. GERBER, H. *An Introduction to Mathematical Risk Theory*. Philadelphia: S.S. Huebner Foundation, University of Pennsylvania, 1979.
11. GROSS, D., AND HARRIS, C. *Fundamentals of Queueing Theory*, 2nd ed. New York: John Wiley, 1985.
12. HOGG, R., AND KLUGMAN, S. *Loss Distributions*. New York: John Wiley, 1984.
13. HOLSTEN, F. Discussion of Pike, B. "Some Considerations in Determining Incurred Claims Used in the Computation of Dividends under Group Accident and Health Insurance," *TSA* 10 (1958): 643–49.
14. KARLIN, S., AND TAYLOR, H. *A First Course in Stochastic Processes*, 2nd ed. New York: Academic Press, 1975.
15. KARLIN, S., AND TAYLOR, H. *A Second Course in Stochastic Processes*. New York: Academic Press, 1981.
16. KARLSSON, J. "A Stochastic Model for Time Lag in Reporting of Claims," *Journal of Applied Probability* 11 (1974): 382–87.
17. KLEINROCK, L. *Queueing Systems–Volume 1, Theory*. New York: John Wiley, 1975.
18. KOPPEL, S., O'GRADY, F., SEE, G., AND SHAPLAND, R. "Reserve Principles for Individual Health," *TSA* 37 (1985): 201–40.
19. MEIR, A., AND MOON, J. "Some Asymptotic Results Useful in Enumeration Problems," *Aequationes Mathematicae* 33 (1987): 260–68.
20. MIYAZAWA, M. "Approximation of the Queue-Length Distribution of an M/GI/s Queue by the Basic Equations," *Journal of Applied Probability* 23 (1986): 443–58.
21. O'GRADY, F., ed. *Individual Health Insurance*. Itasca, Ill.: Society of Actuaries, 1988.
22. PANJER, H. "The Aggregate Claims Distribution and Stop-Loss Reinsurance," *TSA* 32 (1980): 523–45.
23. PANJER, H. "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin* 12 (1981): 22–26.
24. RANTALA, J. "An Application of Stochastic Control Theory to Insurance Business," *Acta University Tamp. Series A*. 164 (1984).
25. ROSS, S. *Stochastic Processes*. New York: John Wiley, 1983.
26. ROSS, S. *Introduction to Probability Models*, 3rd ed. Orlando: Academic Press, 1985.

27. RUOHONEN, M. "The Claims Occurrence Process and the I.B.N.R. Problem," *Proceedings of the 23rd International Congress of Actuaries*, Helsinki, Finland, (1988): 113–23.
28. STRÖTER, B. "The Numerical Evaluation of the Aggregate Claim Density Function Via Integral Equations," *Blätter der Deutschen Gesellschaft für Versicherungs Mathematik* 17 (1985): 1–14.
29. TAYLOR, G. *Claims Reserving in Non-Life Insurance*. Amsterdam: North Holland, 1986.
30. TIJMS, H. *Stochastic Modelling and Analysis: A Computational Approach*. Chichester: John Wiley, 1986.
31. VANDEBROEK, M., AND DEPRIL, N. "Recursions for the Distribution of a Life Portfolio: A Numerical Comparison," *Bulletin of the Royal Association of Belgian Actuaries*, 1987, in press.
32. VAN EEGHEN, J. *Loss Reserving Methods*. Vol. 1 in *Surveys of Actuarial Studies*. Holland: Nationale-Nederlanden, 1981.
33. VAN HOORN, M. *Algorithms and Approximations for Queueing Systems*. CWI Tract No. 8. Amsterdam: CWI, 1984.
34. WILLMOT, G. "The Total Claims Distribution under Inflationary Conditions." University of Waterloo, Institute of Insurance and Pension Research, Report 88-02; *Scandinavian Actuarial Journal*, no. 1 (1989): 1–12.
35. WILLMOT, G. "Limiting Tail Behaviour of Some Discrete Compound Distributions." University of Waterloo, Institute of Insurance and Pension Research, Report 88-16; *Insurance: Mathematics and Economics* 8 (1989): 175–85.
36. WILLMOT, G. "Asymptotic Tail Behaviour of Poisson Mixtures with Applications." University of Waterloo, Institute of Insurance and Pension Research, Report 88-07; *Advances in Applied Probability* 22 (1990): 147–59.