B. Wissenschaftliche Mitteilungen

RAGNAR NORBERG, Oslo

A note on experience rating of large group life contracts

1 Introduction

Much of the existing literature on experience rating and credibility has been written with a primary view to group insurance, such as group life assurance and workmen's compensation insurance. This was the case with *Whitney's* (1918) pioneering paper, from which credibility methods originated, *Keffer's* (1929) paper which advanced matters related to the technique of conjugate prior distributions, and also the works of *Bailey* (1945, 1950).
One principal purpose of group insurance is to reduce administration expenses by recording only very summary data on the insurees. The point is that for the group as a whole the amount of risk exposure will be large, so that its risk level can be assessed from the claim files.
So far as this author knows, as yet only the most simple credibility formulas have been applied to group life insurance. This paper presents an attempt to work out more sophisticated tools for experience rating of group life treaties by making the age distribution and the mortality law explicit parts of the model framework. The basic model is presented in Paragraph 2. We consider groups which are not too small, so that the deaths can be assumed to conform with a Poisson process assumption.
In Paragraph 3 we work out an experience rating plan akin to the bonus schemes well known from ordinary life insurance. The plan is based on currently updated estimates of the total force of mortality for the whole group and the distribution of ages at death.
In Paragraph 4 we attack the problem of experience rating by credibility methods. Empirical credibility procedures involving estimation of structural parameters are included.
In Paragraph 5 some of the methods are applied to data from the group of municipal employees in the city of Oslo.

2 The basic model

We consider a group life assurance treaty which has been in force throughout the time period $[0, t]$, t being the present moment. The policy specifies

$s(y)$, the sum payable by death at age y.

By time t the insurer has observed

$n(\tau)$, the group size at time τ, $0 \leqslant \tau \leqslant t$,
$D(\tau)$, the number of deaths in $[0, \tau]$, $0 \leqslant \tau \leqslant t$,

and, if $D(t) > 0$,

Y_j, the age at death by death No. j, $j = 1, \ldots, D(t)$.

It is convenient to put

$Y_0 = 0$.

The elements governing the course of deaths are

A, the age distribution of the group,

and

μ, the force of mortality of the group,

which are both assumed to be independent of time. Thus, at time τ there are $n(\tau)\{A(x+dx) - A(x)\}$ group members at ages between x and $x + dx$, each with a probability $\mu(x)d\tau$ of dying before time $\tau + d\tau$. We introduce the weighted force of mortality,

$$\lambda = \int_0^\infty \mu(x)\, dA(x). \tag{2.1}$$

A member chosen at random from the group at time τ has probability $\lambda\, d\tau$ of dying within time $\tau + d\tau$.
Assuming that the group is not too small, we may adopt the point of view of collective risk theory and assume that the numbers of deaths $\{D(\tau);\ \tau \geqslant 0\}$ constitute a Poisson point process with intensity

$$n(\tau)\lambda \tag{2.2}$$

at time τ. Accordingly, $D(t)$ is Poisson distributed with parameter $N(t)\lambda$, where

$$N(t) = \int_0^t n(\tau)\, d\tau,$$

the amount of risk exposure up to time t, i.e. the total time at risk for all persons in the group in $[0, t]$.
The process $\{D(\tau);\ \tau \geqslant 0\}$ is independent of $Y_1, Y_2, \ldots$, which are mutually independent and identically distributed (i.i.d.) with cumulative distribution function (c.d.f.) given by

$$G(y) = \int_0^y \mu(x)\, dA(x)/\lambda. \tag{2.3}$$

It follows that the likelihood function of our observations $D(t), Y_0, \ldots, Y_{D(t)}$ is given by

$$P\{D(t) = d,\ \bigcap_{j=1}^{d} y_j < Y_j \leqslant y_j + dy_j\} = \frac{\{N(t)\lambda\}^d}{d!} e^{-N(t)\lambda} \prod_{j=1}^{d} dG(y_j), \tag{2.4}$$

with the obvious interpretations of $\bigcap_{j=1}^{0}$ and $\prod_{j=1}^{0}$ as the certain event and 1, respectively.
The risk process arising from these assumptions is nonhomogeneous generalized Poisson with intensity given by (2.2) and claim size distribution F given by

$$F(z) = P\{s(Y_1) \leqslant z\}.$$

In connection with tarrification a basic quantity is

$$\pi = \int_0^\infty s(x)\mu(x)\, dA(x), \tag{2.5}$$

the annual per capita risk premium. Since G and A are related by (2.3), π may also be expressed by G. By (2.3) we have

$$dA(y) = \lambda \frac{1}{\mu(y)} dG(y), \tag{2.6}$$

and, as A is a probability distribution,

$$\lambda = \left\{ \int_0^\infty \frac{1}{\mu(y)} dG(y) \right\}^{-1} \tag{2.7}$$

From (2.5), (2.6), and (2.7) we gather

$$\pi = \lambda \int_0^\infty s(y)\, dG(y) \tag{2.8}$$

$$= \int_0^\infty s(y)\, dG(y) \Big/ \int_0^\infty \frac{1}{\mu(y)} dG(y). \tag{2.9}$$

3 Experience rating; a non-Bayesian approach

We now turn to the problem of rating the group. Our calculations will be on a net basis, so that a loading for administration expenses will have to be added to the premiums proposed here. The correct premium is π given by (2.5). However, the age distribution A and (possibly) also the force of mortality μ are unknown, and so the insurer has to rely on some estimate of π. As time passes and risk experience is obtained, he will get steadily improved estimates of A and μ and, consequently, of π. Thus, he will actually be exercising some kind of experience rating. In our discussions of how to design a scheme for experience rating we shall distinguish between the case where both A and μ are unknown and the case where only A is unknown.

Case 1. Both A and μ are unknown.

In this case the intensity λ defined by (2.1) and the distribution G defined by (2.3) are functionally unrelated. From (2.4) it is seen that at time t the number of deaths, $D(t)$, is sufficient for λ and ancillary for G. It follows that inference about λ should be based on $D(t)$ and inference about G should be based on the ages at death Y_i, conditional on $D(t)$.
At time t the optimal (uniformly minimum variance unbiased and maximum likelihood) estimator of λ is

$$\hat{\lambda}_t = D(t)/N(t). \tag{3.1}$$

An interval estimator based on the normal approximation is given by the bounds

$$\lambda_t^{\pm} = \hat{\lambda}_t \pm \alpha(\varepsilon)\{\hat{\lambda}_t/N(t)\}^{1/2}, \tag{3.2}$$

where $\alpha(\varepsilon)$ is the upper $\varepsilon/2$ point in the $N(0, 1)$-distribution. The confidence level of this interval approaches $1-\varepsilon$ as $N(t)\to\infty$. More refined asymptotic bounds and also exact bounds for the case where $N(t)$ is small are available, see e.g. *Sverdrup* (1967).
Making no assumption concerning the shape of G, we estimate it by the empirical c.d.f. $\hat{G}_t$, defined for $D(t)\geqslant 1$ by

$$\hat{G}_t(y) = \frac{1}{D(t)} \sum_{i=1}^{D(t)} I(Y_i \leqslant y), \tag{3.3}$$

where $I(A)$ is the indicator function of the event A. An asymptotic $(1-\varepsilon)$-level confidence band for G is

$$G_t^-(y) \leqslant G(y) \leqslant G_t^+(y) \quad \text{for all } y, \tag{3.4}$$

with confines defined by

$$G_t^-(y) = \{\hat{G}_t(y) - \beta(\varepsilon)/D(t)^{1/2}\} \vee 0 \tag{3.5}$$

and

$$G_t^+(y) = \{\hat{G}_t(y) + \beta(\varepsilon)/D(t)^{1/2}\} \wedge 1. \tag{3.6}$$

(Here $a \wedge b$ and $a \vee b$ denote the smaller and the larger, respectively, of the numbers a and b.) The constant $\beta(\varepsilon)$ may be picked from statistical tables, e.g. the one in *Owen* (1962), which also provides exact bounds (3.4) for the case where $D(t)$ is small. *Billingsley* (1968) gives the asymptotic relation

$$\sum_{k=1}^{\infty} (-1)^{k+1} \exp\{-2k^2\beta^2(\varepsilon)\} = \varepsilon/2.$$

The product $(\lambda_t^-, \lambda_t^+) \times (G_t^-, G_t^+)$ forms a confidence region of (λ, G) with asymptotic (as $N(t)\to\infty$) confidence level $(1-\varepsilon)^2$. If (λ, G) belongs to this confidence region, π is between the bounds

$$\pi_t^- = \lambda_t^- \min_{G_t^- \leqslant G \leqslant G_t^+} \int s(y)\, dG(y) \tag{3.7}$$

and

$$\pi_t^+ = \lambda_t^+ \max_{G_t^- \leqslant G \leqslant G_t^+} \int s(y)\, dG(y). \tag{3.8}$$

To find $\pi_t^+(\pi_t^-)$, we have to maximize (minimize) the integral $\int s(y)\, dG(y)$ subject to the constraint (3.4). Roughly speaking, we seek the function G which allots as much mass as possible to those values of y where $s(y)$ is large (small). If s is monotone, the extreme values of $\int s(y)\, dG(y)$ are easily obtained by the following lemma.

Lemma 3.9 Let G_1 and G_2 be c.d.f.-s satisfying

$$G_1(y) \leqslant G_2(y) \quad \text{for all } y. \tag{3.10}$$

Then the inequality

$$\int s(y)\, dG_1(y) \leqslant \int s(y)\, dG_2(y) \tag{3.11}$$

is valid for each nonincreasing function s.

Proof: Let X be a random variable with a uniform distribution over $[0,1]$. For each $i = 1, 2$ let $Y_i = G_i^{-1}(X)$, where G_i^{-1} is the quasi-inverse of G_i defined by $G_i^{-1}(x) = \inf\{y;\, G_i(y) \geqslant x\}$. As is well known, Y_i has G_i as its c.d.f. From (3.10) it follows that $Y_1 \geqslant Y_2$ and, as s is nonincreasing, $s(Y_1) \leqslant s(Y_2)$. It follows that $\mathrm{Es}(Y_1) \leqslant \mathrm{Es}(Y_2)$, which is just the asserted inequality (3.11). □

On combining Lemma 3.9 with (3.7) and (3.8), we arrive at the following result.

Theorem 3.12 An interval estimator of π, with asymptotic confidence level not less than $(1-\varepsilon)^2$, is (π_t^-, π_t^+) defined by (3.7) and (3.8). If the sum function s is nonincreasing, the bounds are

$$\pi_t^{\pm} = \lambda_t^{\pm} \int s(y)\, dG_t^{\mp}(y), \tag{3.13}$$

where $\lambda_t^{\pm}$ are defined by (3.2) and $G_t^{\mp}$ by (3.5) and (3.6).

A great variety of experience rating procedures can be proposed on basis of the theory in this paragraph. We shall describe one which in a sense conforms with the principles of rating in ordinary life insurance, by which the technical bases include security loadings that serve as a buffer in case of unfavourable changes in the basic data. Under normal circumstances these loadings create a reserve fund which is to be allotted to the policyholders in accordance with some bonus scheme.

The experience rating plan we propose is as follows. As a part of the underwriting procedure, the parties agree on some initial per capita annual premium, π_0^+. This premium should be composed of an *a* priori assessment of π (i.e. of A and μ), made as realistic as possible, and a security loading. As time passes and risk experience is obtained, the premium is adjusted in the following manner. At time t a "safe" value of the annual per capita premium is the upper confidence bound, π_t^+, defined by (3.8). The corresponding premium for the period $(t, t+1)$ due at time $t = 0, 1, \ldots,$ is

$$\pi_t^+ \int_{\tau=t}^{t+1} (1+i)^{-(\tau-t)} n(\tau)\, d\tau, \tag{3.14}$$

where i is the annual rate of interest. If, moreover, the premiums paid in former periods are adjusted annually to the current safe level, the bonus payable to the group at time $t = 1, 2, \ldots,$ is

$$(\pi_{t-1}^+ - \pi_t^+) \int_{\tau=0}^{t} (1+i)^{t-\tau} n(\tau)\, d\tau \; I(\pi_{t-1}^+ \geqslant \pi_t^+). \tag{3.15}$$

The factor $I(\pi_{t-1}^+ \geqslant \pi_t^+)$ ensures that negative bonus allotments cannot occur.

The expression (3.15) needs not converge to zero as $t \to \infty$; it may even diverge to infinity. One should, therefore, modify the rule to make it behave well also for large values of t. One could, for instance, close the bonus allotments after some time T, which may be either fixed in advance or dependent on the development of the confidence bounds (3.7) and (3.8). One possibility is to let the bonus payments fall in as soon as π is sufficiently accurately estimated, say when $\pi_t^+ - \pi_t^- < c$.

Case 2. A is unknown, whereas μ is known

In this case λ and G are related by (2.7), and an estimator of π, which makes use of the fact that μ is known, may be obtained by inserting an estimator of G in (2.9). We readily get the following result.

Theorem 3.16 An interval estimator of π, with asymptotic confidence level not less than $(1-\varepsilon)^2$, is (π_t^-, π_t^+) defined by

$$\pi_t^- = \min_{G_t^- \leqslant G \leqslant G_t^+} \int s(y)\, dG(y) / \max_{G_t^- \leqslant G \leqslant G_t^+} \int \frac{1}{\mu(y)} dG(y) \tag{3.17}$$

and

$$\pi_t^+ = \max_{G_t^- \leqslant G \leqslant G_t^+} \int s(y)\, dG(y) / \min_{G_t^- \leqslant G \leqslant G_t^+} \int \frac{1}{\mu(y)} dG(y). \tag{3.18}$$

If s is nonincreasing and μ is nondecreasing, the confidence bounds are

$$\pi_t^{\pm} = \int s(y)\, dG_t^{\pm}(y) / \int \frac{1}{\mu(y)} dG_t^{\mp}(y). \tag{3.19}$$

The scheme for rating and bonus allotments defined by (3.14) and (3.15), may be applied without modifications to this case.

4 Experience rating by credibility methods

We now switch to the Bayes or empirical Bayes setting, by which the unknown parameters are viewed as random variables. As in the previous paragraph we also here distinguish between two cases.

Case 1. Both A and μ are unknown.

The unknown random risk parameter characterizing the group is (λ, G). (The parameter (μ, A) appears to be more "basic" than (λ, G) since the mapping $(\mu, A) \rightarrow (\lambda, G)$ defined by (2.1) and (2.3) is not one-to-one. However, being interested in the risk premium, which by (2.8) depends only on (λ, G), we may equally well take (λ, G) as risk parameter.)
Concerning the distribution of (λ, G), usually termed prior (distribution), we make the convenient assumption that λ and G are independent. Beyond this we do not want to be very specific as regards the shape of the prior and, therefore, resort to credibility estimators, which depend only on certain first and second order prior moments.

The credibility estimator of λ based on $D(t)$ is (see e.g. *Norberg* (1979))

$$\tilde{\lambda}_t = \frac{N(t)}{N(t)+\varkappa}\hat{\lambda}_t + \frac{\varkappa}{N(t)+\varkappa}\lambda_0, \tag{4.1}$$

where

$$\lambda_0 = E\lambda$$

and

$$\varkappa = E\lambda / \operatorname{Var}\lambda.$$

The credibility estimator of G based on $Y_1, \ldots, Y_{D(t)}$, conditional on the value of $D(t) > 0$, is

$$\tilde{G}_t = \frac{D(t)}{D(t)+\alpha}\hat{G}_t + \frac{\alpha}{D(t)+\alpha}G_0, \tag{4.2}$$

where

$$G_0 = EG,$$

$$\alpha = E\operatorname{Var}(Y/G)/\operatorname{Var}E(Y/G),$$

and $\hat{G}_t$ is the empirical c.d.f. defined by (3.3). From (4.2) we obtain the credibility estimator of $\int s(y)\,dG(y)$,

$$\begin{aligned} \int s(y)\,d\tilde{G}_t(y) &= \frac{D(t)}{D(t)+\alpha}\int s(y)\,d\hat{G}_t(y) + \frac{\alpha}{D(t)+\alpha}\int s(y)\,dG_0(y) \\ &= \frac{D(t)}{D(t)+\alpha}\bar{s}_t + \frac{\alpha}{D(t)+\alpha}\int s(y)\,dG_0(y), \end{aligned} \tag{4.3}$$

where $\bar{s}_t$ is defined for $D(t) > 0$ by

$$\bar{s}_t = \frac{1}{D(t)}\sum_{j=1}^{D(t)} s(Y_j).$$

Credibility estimators are Bayes solutions, with respect to quadratic loss, in the restricted class of linear estimators. The estimator $\tilde{\lambda}$ is unrestricted Bayes

in the particular case where λ has a gamma prior (see e.g. *De Groot* (1970), Ch. 9). *Ferguson* (1972) introduced the so-called Dirichlet process and proved that if G has a Dirichlet prior, the credibility estimators $\tilde{G}_t$ and $\int s(y)\, d\tilde{G}_t(y)$ are unrestricted Bayes. See also *Zehnwirth* (1977, 1978).
On inserting the estimators defined in (4.1) and (4.3) into (2.8), we obtain an estimator of π,

$$\tilde{\pi}_t = \tilde{\lambda}_t \int s(y)\, d\tilde{G}_t(y). \tag{4.4}$$

Note that even if $\tilde{\lambda}_t$ and $\tilde{G}_t$ are exactly Bayes, $\tilde{\pi}_t$ will not in general be a Bayes estimator of π_t. It is, however, consistent as $N(t) \to \infty$ since both $\tilde{\lambda}_t$ and $\tilde{G}_t$ are consistent.
Two interpretations are possible for the prior distribution. Which to choose depends on the situation at hand.
In the first place the prior may be viewed as a summary expression of our subjective, prior to data beliefs concerning the value of (λ, G). Then the constants $\varkappa$ and α in (4.1) and (4.2) measure our faith in the prior values λ_0 and G_0 as compared to the sample values $\hat{\lambda}_t$ and $\hat{G}_t$. For instance, if we choose $\alpha = 50$, it means that after 50 deaths we consider G_0 and $\hat{G}_t$ as equally trustworthy estimates of G. This purely Bayesian approach is appropriate when the group is unique in the sense that we have no statistical information from similar group assurance treaties.
In the second place the prior may be given a frequency interpretation, in which case we shall speak of it as the structural distribution. This point of view is appropriate if the treaty may be regarded as picked at random from a population of similar treaties. This situation is referred to as the empirical Bayes case since it allows for estimation of the structural parameters λ_0, $\varkappa$, G_0, and α from the claims records of a sample of treaties.
Suppose we have data from I independent treaties, and let us equip with subscript i all quantities originating from the i-th treaty. We adopt the convention that an estimator is denoted by the parameter symbol marked with an asterisk. A wide class of estimators of structural parameters is given by the defining relations (4.5)–(4.9) below. The $w_{\lambda i}$ and w_{Gi} are weights normed such that $\sum_{i=1}^{I} w_{\lambda i} = \sum_{i=1}^{I} w_{Gi} = 1$. For the sake of simplicity we drop the indices t_i, writing $\hat{\lambda}_i$, $\hat{G}_i$, ... etc. instead of $\hat{\lambda}_{t_i i}$, $\hat{G}_{t_i i}$, ... The estimators are of the form

$$\lambda_0^* = \sum_{i=1}^{I} w_{\lambda i}\hat{\lambda}_i, \tag{4.5}$$

$$(\operatorname{Var}\lambda)^* = (1 - \sum_{i=1}^{I} w_{\lambda i}^2)^{-1}\left[\sum_{i=1}^{I} w_{\lambda i}\{\hat{\lambda}_i - \lambda_0^*\}^2 - \lambda_0^* \sum_{i=1}^{I} w_{\lambda i}(1 - w_{\lambda i})/N_i\right], \tag{4.6}$$

$$G_0^* = \sum_{i=1}^{I} w_{Gi}\hat{G}_i, \tag{4.7}$$

$$\{E\operatorname{Var}(Y/G)\}^* = \sum_{i;\ D_i>1} w_{Gi} \sum_{j=1}^{D_i} (Y_{ji} - \overline{Y}_{\cdot i})^2/(D_i - 1), \tag{4.8}$$

$$\begin{aligned}\{\operatorname{Var}E(Y/G)\}^* = \left(1 - \sum_{i=1}^{I} w_{Gi}^2\right)^{-1}\Bigg[&\sum_{i=1}^{I} w_{Gi}(\overline{Y}_{\cdot i} - \overline{Y}_{\cdot\cdot})^2 \\ &- \{E\operatorname{Var}(Y/G)\}^* \sum_{i=1}^{I} w_{Gi}(1 - w_{Gi})/D_i\Bigg],\end{aligned} \tag{4.9}$$

with $\overline{Y}_{\cdot i}$ defined for all i with $D_i > 0$ by

$$\overline{Y}_{\cdot i} = \sum_{j=i}^{D_i} Y_{ji}/D_i$$

and $\overline{Y}_{\cdot\cdot}$ defined when $\sum_{i=1}^{I} D_i > 0$ by

$$\overline{Y}_{\cdot\cdot} = \sum_{i=1}^{I} w_{Gi}\overline{Y}_{\cdot i}\,.$$

The weights should be measures of amounts of statistical information, and as such we propose to take the relative amounts of risk exposure

$$w_{\lambda i} = N_i/\sum_{k=1}^{I} N_k$$

and

$$w_{Gi} = D_i/\sum_{k=1}^{I} D_k,$$

$i = 1, \ldots, I$. In order that (4.9) be well defined we should always take $w_{Gi} = 0$ if $D_i = 0$. The estimators defined in (4.5)–(4.9) are easily shown to be unbiased.

Case 2. A is unknown, whereas μ is known

Following the ideas of Paragraph 3, we may now, as an alternative to the above approach, base our estimate of π on relation (2.9). Enter the credibility estimator $\widetilde{G}_t$ defined in (4.2) into (2.9). The resulting estimator is not Bayes, but consistent under weak assumptions. We do not elaborate on this method.

5 An application to real data

In this closing paragraph we apply our methods to the data presented in Table 1, which stems from the group of employees in the municipality of Oslo. A group life contract for this group has been in force since May 1, 1978, and the table includes all deaths incurred in 1978, 1979 and 1980. The table also specifies the sum function s in units of the basic amount of the National insurance scheme. The number $n(\tau)$ of insured persons at risk was constantly equal to 31 500 throughout the period.

Entering the data into the formulas (3.1), (3.2), (3.3), (3.5), and (3.6), we obtain the following point estimates and accompanying upper and lower asymptotic 95 % confidence bounds.

For the $N(1) = (8/12)\,31500 = 21000$ risk years exposed during 1978 we find $\hat{\lambda}_1 = 0.00205$, $\lambda_1^- = 0.00144$, $\lambda_1^+ = 0.00266$, and $\hat{G}_1$ and $G_1^\pm$ as shown in Fig. 1, which gives $\int s(y)\,dG_1^\mp(y) = 1.849$ and 2.459. Inserting these values into (3.13), we obtain $\pi_1^- = 0.00266$ and $\pi_1^+ = 0.00654$.

For $N(2) = (20/12)\,31500 = 52500$ risk years exposed during 1978 and 1979 we find $\hat{\lambda}_2 = 0.00232$, $\lambda_2^- = 0.00191$, $\lambda_2^+ = 0.00273$, and $\hat{G}_2$ and $G_2^\pm$ as shown in Fig. 2, which gives $\int s(y)\,dG_2^\mp(y) = 1.893$ and 2.262. Hence we obtain the bounds $\pi_2^- = 0.00362$ and $\pi_2^+ = 0.00618$, which are narrower than those obtained after the first year.

For $N(3) = (32/12)\,31500 = 84000$ risk years exposed during the years 1978–1980 we find $\hat{\lambda}_3 = 0.00237$, $\lambda_3^- = 0.00204$, $\lambda_3^+ = 0.00270$, and $\hat{G}_3$ and $G_3^\pm$ as shown in Fig. 3, which gives $\int s(y)\,dG_3^\mp(y) = 1.974$ and 2.264. This gives $\pi_3^- = 0.00403$ and $\pi_3^+ = 0.00611$.

Finally we employ the results in Paragraph 4 to accomplish a credibility analysis of the data in Table 1. Let us take the position that we have no prior information concerning the composition of the group beyond the fact that it is a sample from the work-force. Then it seems reasonable to base the a priori risk assessment on group life mortality statistics, if available.

Let m be the force of mortality and l the corresponding decrement function for insurees with group life coverage. Further, let y' be the (average or "typical") age of entrance into the work-force and y'' the age of retirement. Then, according to this mortality law, the proportion of people at age y or less is

$$\int_{y'}^{(y'' \wedge y) \vee y'} l(x)\,dx \Big/ \int_{y'}^{y''} l(x)\,dx. \tag{5.1}$$

On inserting m for μ and the expression (5.1) for $A(y)$ in (2.1) and (2.3), we obtain the prior mean values

$$\lambda_0 = \int_{y'}^{y''} m(x)\,l(x)\,dx \Big/ \int_{y'}^{y''} l(x)\,dx = \{l(y') - l(y'')\} \Big/ \int_{y'}^{y''} l(x)\,dx \tag{5.2}$$

and

$$G_0(y) = \int_{y'}^{(y \wedge y'') \vee y'} m(x)\,l(x)\,dx / \{l(y') - l(y'')\}$$

$$= [l(y') - l\{(y \wedge y'') \vee y'\}] / \{l(y') - l(y'')\}. \tag{5.3}$$

In our example we take m to be the Gompertz-Makeham function fitted to mortality data for Norwegian group life treaties by The Statistical Bureau of the Norwegian Life Insurance Companies. It is given by $10^3 m(y) = 0.2897 + 0.0204 \cdot 10^{0.04445y}$. Further we put $y' = 18$ and $y'' = 72$. The corresponding prior mean values in (5.2) and (5.3) are $\lambda_0 = 0.00552$ and G_0 as shown in Fig. 4, the latter giving $\int_{y'}^{y''} s(y)\,dG_0(y) = 1.862$. On entering these values together with the sample estimates $\hat{\lambda}_t$ and $\hat{G}_t$, $t = 1, 2, 3$, into formulas (4.1), (4.2), and (4.3) with $\varkappa = 50.000$ and $\alpha = 50$, we obtain the premiums $\tilde{\pi}_0 = \lambda_0 \int_{y'}^{y''} s(y)\,dG_0(y) = 0.01028$, $\tilde{\pi}_1 = 0.00897$, $\tilde{\pi}_2 = 00782$, and $\tilde{\pi}_3 = 0.00733$.
A comment is in order concerning the choice of the prior means given by (5.1)–(5.3). Actually the age distribution will be given by formula (5.1) only if the pattern of entries and decrements of the group conforms with the scheme of a stationary closed population. From Figures 3 and 4 it is readily seen that G_0 does not agree with the mortality experience for the group. The observed ages at death are on the average significantly lower than could be expected under the stationary population hypothesis.This may be explained by the strong growth in the staff of municipal employees during the post-war period, which has effected a shift to the left in the age distribution. Prior information of this kind might, of course, have been taken into account to produce a more judicious choice of G_0.

Table 1

Claims record for the years 1978, 1979, and 1980 for 31 500 insured employees in the municipality of Oslo

Age at death y	Sum $s(y)$	Number of deaths at age y by time t, $t=1$ (Dec. 31, 1978)	$t=2$ (Dec. 31, 1979)	$t=3$ (Dec. 31, 1980)
18				1
.				.
.				.
24			1	2
5			1	1
6				3
7				
8				1
9		1	2	2
30		2	2	2
1				
2		1	1	1
3				
4				1
5				2
6	3.00			
7				
8			1	2
9		1	1	1
40			2	2
1				1
2		1	2	4
3			1	2
4			3	3
5		1	3	3
6		1	2	2
7			1	2
8		1	1	1
9		2	3	4
50		1	3	5
1	2.85		2	8
2	2.70	2	3	5
3	2.55	1	1	5
4	2.40	1	3	7
5	2.25	2	7	9
6	2.10	2	9	12
7	1.95	4	6	10
8	1.80	3	5	9
9	1.65	1	3	10
60			2	7
1		2	11	14
2		2	7	11
3			6	6
4		3	6	11
5	1.50	3	6	8
6			1	3
7		1	6	7
8			2	3
9		1	3	3
70		2	2	2
1		1	1	1
		$D(1)=43$	$D(2)=122$	$D(3)=199$

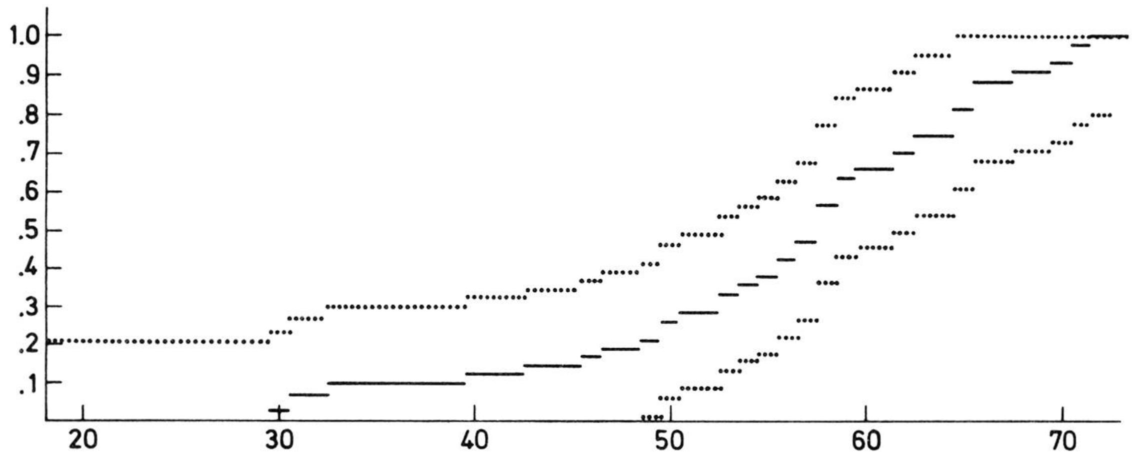

Figure 1

Empirical c.d.f. $\hat{G}_1$ and asymptotic 95% confidence band for G based on 43 deaths incurred during 1979 amongst employees in the municipality of Oslo

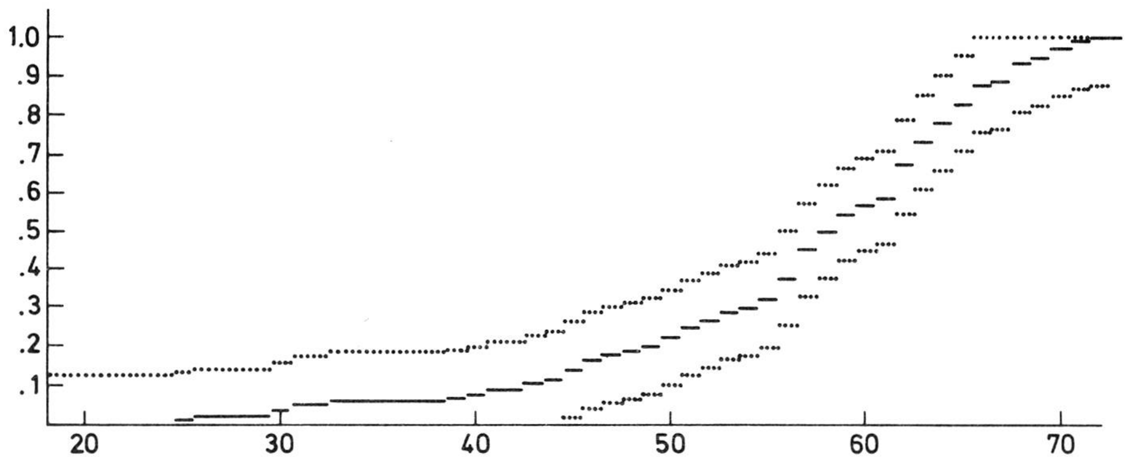

Figure 2

Empirical c.d.f. $\hat{G}_2$ and asymptotic 95% confidence band for G based on 122 deaths incurred during 1978 and 1979 amongst employees in the municipality of Oslo

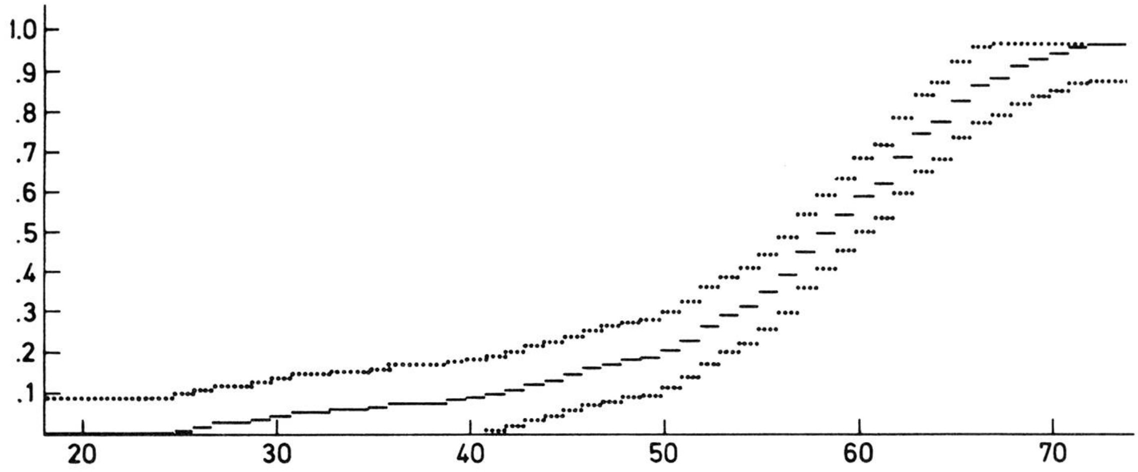

Figure 3

Empirical c.d.f. $\hat{G}_3$ and asymptotic 95 % confidence band for G based on 199 deaths incurred during 1978, 1979, and 1980 amongst employees in the municipality of Oslo

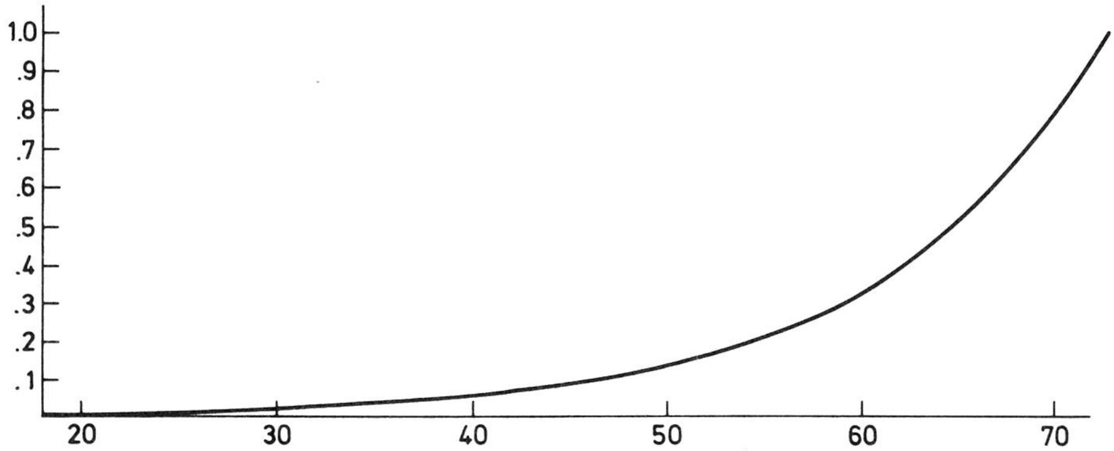

Figure 4

The c.d.f. G_0 given by (5.3) with mortality given by $10^3 m(y) = 0.2897 + 0.0204 \cdot 10^{0.04445\,y}$

Acknowledgements

This paper was prepared for the Storebrand Life Insurance Company while I was affiliated to Storebrand in 1980–81. My thanks are due to Swiss actuaries who were the first to take interest in the project. Table 1 is compiled from data kindly put to my disposal by The Statistical Bureau of the Norwegian Life Insurance Companies.

Ragnar Norberg
Matematisk institutt
P.b. 1053 Blindern
N–0316 Oslo 3
Norway

References

Bailey, A.L. (1945): A generalized theory of credibility. Proceedings of The Casualty Actuarial Society *32*, 13–20.

Bailey, A.L. (1950): Credibility procedures, La Place's generalization of Bayes' rule, and the combination of collateral knowledge with observed data. Proceedings of The Casualty Actuarial Society *37*, 7–23.

Billingsley, P. (1968): Convergence of probability measures. John Wiley & Sons, Inc., New York.

De Groot, M. (1970): Optimal statistical decisions. McGraw-Hill Company, New York.

Ferguson, T.S. (1972): A Bayesian analysis of some nonparametric problems. Annals of Statistics, *1*, 209–230.

Keffer, R. (1929): An Experience rating formula. Transactions of the Actuarial Society of America *30*, 130–139.

Norberg, R. (1979): The credibility approach to experience rating. Scandinavian Actuarial J., 181–221.

Owen, D.B. (1962): Handbook of Statistical Tables. Addison Wesley Publ. Co. Inc., Reading, Massachussets.

Sverdrup, E. (1967): Laws and chance variations. Vol. 1. North Holland Publishing Co., Amsterdam.

Whitney, A.W. (1918): The theory of experience rating. Proceedings of The Casualty Actuarial Society *4*, 274–292.

Zehnwirth, B. (1977): The mean credibility formula is a Bayes rule. Scandinavian Actuarial J., 212–216.

Zehnwirth, B. (1978): The credible distribution function is an admissible Bayes rule. Scandinavian Actuarial J., 121–127.

Abstract

The risk process of a group life treaty is assumed to be of generalized Poisson type. Experience rating is discussed within a non-Bayesian as well as in an (empirical) Bayes framework. Two situations are considered: first the one where both the age distribution and the mortality law of the group are unknown and, second, the one where only the age distribution is unknown.

Zusammenfassung

Unter der Annahme, dass der Risikoprozess einer Gruppen-Lebensversicherung vom Typ zusammengesetzt-Poisson ist, wird die Erfahrungstarifierung diskutiert, und zwar in einem nicht-Bayesschen als auch in einem (empirischen) Bayesschen Rahmen. Es werden zwei Situationen untersucht: die eine, in der Altersverteilung und Sterblichkeitsgesetz unbekannt sind, die andere, in der lediglich die Altersverteilung unbekannt ist.

Résumé

L'évolution du risque d'un contrat d'assurance-vie de groupe est supposé être de type poissonnien composé. L'auteur traite de la tarification expérimentale dans un cadre non-Bayesien, également dans un cadre Bayesien empirique. Il considère deux situations: lorsque sont inconnues premièrement la distribution des âges et la loi de mortalité du groupe, secondement seulement la loi de distribution des âges.