Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: - (1983)

Heft: 1

Artikel: Credibility models allowing durational effects

Autor: Sundt, Bjørn

DOI: https://doi.org/10.5169/seals-967132

Bjørn Sundt, Oslo

Credibility models allowing durational effects

1 Introduction

1A. For a numerical study on parameter estimation in credibility models [Sundt (1983)] the present author used data from an automobile liability insurance portfolio observed for three consecutive years. From this portfolio we extracted the subportfolio consisting of all policies that had been in force for all three years. In Table 1.1 we have given the number of policies and claim frequencies both for the subportfolio and the whole portfolio. It is clearly seen that the claim frequencies of the subportfolio are significantly smaller than the frequencies of the whole portfolio. Hence, there must have been some selectional effect by our construction of the subportfolio; it seems that old policies have smaller claim frequencies than young policies.

One way to explain such a selectional effect would be by a learning effect; the drivers get better as time passes. However, if this were the right explanation, then the claim frequencies of the subportfolio should have been decreasing with time, and such an effect is not detected in the data.

It seems more appropriate to assume that there are individual differences between the policies, and that policies with high claim frequencies are more apt to leave the portfolio than policies with low claim frequencies. This aspect has so far not been encountered in credibility theory [except by Taylor (1975)], and it is the purpose of the present paper to examine it more closely.

Table 1.1

Claim frequencies in an automobile liability insurance portfolio (data from Storebrand Insurance Co. Ltd.)

Period	Subportfolio		Whole portfolio	
	Policy years	Frequency	Policy years	Frequency
1976	2697	0.023	11098	0.050
1977	2697	0.027	12103	0.057
1978	2697	0.026	11249	0.052
1976–78	8091	0.025	34450	0.053

1B. As usual in credibility theory, we assume that to each policy in an insurance portfolio there is connected an unknown random variable $\tilde{\theta}$, describing how this policy may differ from the other policies in the portfolio; to concretize, let us roughly say that the policy is more risky the greater the value taken by $\tilde{\theta}$ is.

Let $\tilde{t}$ be the total time the policy stays in the portfolio. Thus, for an n years old policy we know that $\tilde{t} \geqslant n$. We shall assume that $\tilde{t}$ is integer-valued.

The key idea of the present paper is that the fact that the policy has been in force for a certain number of years, tells something about its $\tilde{\theta}$. By this we do not mean that $\tilde{\theta}$ is changing over time, but that *the conditional distribution of $\tilde{\theta}$ given that the policy has been in force for this period, is different from the unconditional distribution of $\tilde{\theta}$.*

1C. Let $\tilde{x}_i$ denote the claim number or the claim amount of an insurance policy from the i-th year it is in force. We assume that $\tilde{x}_1, \tilde{x}_2, \ldots$ are conditionally independent and identically distributed given $\tilde{\theta}$.

It is important in what follows, to imagine that $\tilde{x}_i$ is also defined for $i > \tilde{t}$. Thus, when we want to estimate $\tilde{x}_{n+1}$ utilizing the past experience, the fact that $\tilde{t} > n$, may give some information in addition to the information lying in the values of $\tilde{x}_1, \ldots, \tilde{x}_n$.

1D. The present paper summarizes parts of the author's doctoral dissertation [Sundt (1982)], available from the author. For more details, we refer to the dissertation.

2 Conventions

In what follows, many expressions will be simplified by the concept that when $z_1, z_2, \ldots$ is some sequence, then by ${}_n\underline{\underline{z}}$ we mean the vector ${}_n\underline{\underline{z}} = (z_1, \ldots, z_n)'$.

Another simplifying notation is the following: Let $\tilde{y}$ and $\tilde{z}$ be two random variables. Then, if there is no possibility of misunderstanding, we write $\mathscr{E}(\tilde{y}|z)$ as an abbreviation for $\mathscr{E}(\tilde{y}|\tilde{z}=z)$. We do similarly for conditional variances, densities, etc.

For indexed quantities we skip the index when it does not give any information. For instance, if $\tilde{y}_1, \ldots, \tilde{y}_n$ are identically distributed random variables, we write $\mathscr{E}(\tilde{y})$ instead of $\mathscr{E}(\tilde{y}_i)$.

We shall often introduce conditional densities, expectations, etc. Let $\tilde{y}$ and $\tilde{z}$ have the joint density $f(y, z)$ and $\tilde{z}$ marginal density $g(z)$. Then, if $g(z) \neq 0$, the conditional density $f(y|z)$ of $\tilde{y}$ given $\tilde{z}=z$, is given by

$$f(y|z)=\frac{f(y,z)}{g(z)}$$

For $g(z)=0$ $f(y|z)$ may be defined in an arbitrary way. Because of this, we shall not bother about mentioning the case when denominators are equal to zero when we work with conditional densities, expectations, etc.
By $\underline{\underline{R}}_0$ we shall mean the non-negative real numbers; by $\underline{\underline{N}}_0$ the non-negative integers.

3 The general setup

3A. Let $\tilde{\theta}$ be a random variable with measurable density $u(\theta)$ with respect to a measure space $(\Theta, \mathscr{A}, a)$.
For given $\tilde{\theta}=\theta$ the random variables $\tilde{x}_1, \tilde{x}_2, \ldots$ are independent and identically distributed with measurable density $f(x|\theta)$ with respect to the measure space $(\underline{\underline{R}}_0, \mathscr{B}, b)$. In most applications b is either the counting measure or the Lebesgue measure, in most cases extended by an atom at zero.
The random variable $\tilde{t}$ is defined on $\underline{\underline{N}}_0$, and we assume that

$$Pr(\tilde{t}>n|{}_{\infty}\underline{\underline{\tilde{x}}}={}_{\infty}\underline{\underline{x}},\ \tilde{\theta}=\theta)=\begin{cases}1 & (n=0)\\ \prod_{i=1}^{n} g_i(\theta, {}_i\underline{\underline{x}}) & (n=1,2,\ldots)\end{cases}$$

This gives

$$Pr(\tilde{t}>n|\tilde{t}\geqslant n,\ {}_{\infty}\underline{\underline{\tilde{x}}}={}_{\infty}\underline{\underline{x}}, \tilde{\theta}=\theta)=g_n(\theta, {}_n\underline{\underline{x}})$$

Hence, if $\tilde{\theta}$ and the past claim amounts are given, the future claim amounts do not influence the probability of leaving the portfolio. We also assume that these probabilities are not influenced by the rating of the policy.
The joint density of $({}_n\underline{\underline{\tilde{x}}}', \tilde{t}, \tilde{\theta})'$ is now given by

$$\left(\prod_{i=1}^{n} f(x_i|\theta)\right)\left(\prod_{i=1}^{t-1} g_i(\theta, {}_i\underline{\underline{x}})\right)(1-g_t(\theta, {}_t\underline{\underline{x}}))\,u(\theta) \qquad (t\leqslant n)$$

We let $\tilde{\imath}_n$ denote the indicator

$$\tilde{\imath}_n=\begin{cases}1 & \text{if} \quad \tilde{t}>n\\ 0 & \text{if} \quad \tilde{t}\leqslant n\end{cases}$$

3B. After n years we want to estimate $\tilde{x}_{n+1}$. We assume that the past claim amounts ${}_n\underline{\underline{\tilde{x}}}$ are available, and that it is known whether the policy will be in force next year or not, that is, the available policy data are $({}_n\underline{\underline{\tilde{x}}}', \tilde{\imath}_n)'$. We use quadratic loss, that is, we choose the estimator minimizing the risk

$$\mathscr{R}_{n+1}(\hat{x}_{n+1}) = \mathscr{E}(\hat{x}_{n+1} - \tilde{x}_{n+1})^2$$

within some class $\mathscr{X}_{n+1}$ of estimators $\hat{x}_{n+1}$ based on the available data. The widest possible class is of course

$$\mathscr{X}^*_{n+1} = \{\hat{x}_{n+1} = h({}_n\underline{\underline{\tilde{x}}}, \tilde{\imath}_n) | h: \underline{\underline{R}}^n_0 \times \{0,1\} \to \underline{\underline{R}}\}$$

that is, the class of all estimators utilizing the available data. The optimal estimator in this class is

$$x^*_{n+1} = \mathscr{E}(\tilde{x}_{n+1} | {}_n\underline{\underline{\tilde{x}}}, \tilde{\imath}_n) = \mathscr{E}(\mathscr{E}(\tilde{x}|\tilde{\theta}) | {}_n\underline{\underline{\tilde{x}}}, \tilde{\imath}_n)$$

Similarly, the optimal estimator of $\tilde{x}_{n+1}$ based on the past claim amounts, that is, the optimal estimator from the class

$$\mathscr{X}^{**}_{n+1} = \{\hat{x}_{n+1} = h({}_n\underline{\underline{\tilde{x}}}) | h: \underline{\underline{R}}^n_0 \to \underline{\underline{R}}\}$$

is

$$x^{**}_{n+1} = \mathscr{E}(\tilde{x}_{n+1} | {}_n\underline{\underline{\tilde{x}}}) = \mathscr{E}(\mathscr{E}(\tilde{x}|\tilde{\theta}) | {}_n\underline{\underline{\tilde{x}}})$$

The classical credibility estimator $\ddot{x}_{n+1}$ is the optimal estimator from the class

$$\ddot{\mathscr{X}}_{n+1} = \left\{\hat{x}_{n+1} = a_0 + \sum_{i=1}^{n} a_i \tilde{x}_i | a_0, a_1, \ldots, a_n \in \underline{\underline{R}}\right\}$$

It is well known [see e.g. Bühlmann (1967)] that

$$\ddot{x}_{n+1} = \frac{n}{n+\kappa} \bar{\tilde{x}}_n + \frac{\kappa}{n+\kappa} \mu$$

with

$$\bar{\tilde{x}}_n = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \quad \kappa = \frac{\phi}{\lambda} \quad \phi = \mathscr{E}\mathscr{V}(\tilde{x}|\tilde{\theta}) \quad \lambda = \mathscr{V}\mathscr{E}(\tilde{x}|\tilde{\theta}) \quad \mu = \mathscr{E}(\tilde{x})$$

The estimator $\ddot{x}_{n+1}$ was originally developed in a model in which variables like $\tilde{\imath}$ and $\tilde{\imath}_n$ were not incorporated. In the present model the information lying in $\tilde{\imath}_n$ may be utilized by the introduction of

$$\dot{x}_{n+1} = \alpha_0(\tilde{i}_n) + \sum_{i=1}^{n} \alpha_i(\tilde{i}_n)\tilde{x}_i$$

being the optimal estimator from the class

$$\dot{\mathscr{X}}_{n+1} = \left\{ \hat{x}_{n+1} = a_0(\tilde{i}_n) + \sum_{i=1}^{n} a_1(\tilde{i}_n)\tilde{x}_i \middle| a_0, \ldots, a_n : \{0, 1\} \to \underline{\underline{R}} \right\}$$

It is well known [see e.g. Jewell (1975a)] that the coefficients $\alpha_0(\tilde{i}_n), \ldots, \alpha_n(\tilde{i}_n)$ are determined by the normal equations

$$\mathscr{C}(\dot{x}_{n+1}, {}_n\underline{\underline{\tilde{x}}}'|\tilde{i}_n) = \mathscr{C}(\tilde{x}_{n+1}, {}_n\underline{\underline{\tilde{x}}}'|\tilde{i}_n)$$

$$\mathscr{E}(\dot{x}_{n+1}|\tilde{i}_n) = \mathscr{E}(\tilde{x}_{n+1}|\tilde{i}_n)$$

As

$$\ddot{\mathscr{X}}_{n+1} \subset \mathscr{X}^{**}_{n+1} \subset \mathscr{X}^{*}_{n+1}$$

$$\ddot{\mathscr{X}}_{n+1} \subset \dot{\mathscr{X}}_{n+1} \subset \mathscr{X}^{*}_{n+1}$$

we have

$$\mathscr{R}_{n+1}(x^{*}_{n+1}) \leqslant \mathscr{R}_{n+1}(x^{**}_{n+1}) \leqslant \mathscr{R}_{n+1}(\ddot{x}_{n+1})$$

$$\mathscr{R}_{n+1}(x^{*}_{n+1}) \leqslant \mathscr{R}_{n+1}(\dot{x}_{n+1}) \leqslant \mathscr{R}_{n+1}(\ddot{x}_{n+1}) \tag{3.1}$$

giving a partial ranking of the four estimators.

3.C In practice the estimators of $\tilde{x}_{n+1}$ are to be used for rating the policy for its $(n+1)$-th insurance year, and thus the expressions for the estimators x^{*}_{n+1} and $\dot{x}_{n+1}$ for the case $\tilde{i}_n = 0$ are of no practical interest. Hence, for the rest of the paper we let

$$x^{*}_{n+1} = \mathscr{E}(\tilde{x}_{n+1}|{}_n\underline{\underline{\tilde{x}}}, (\tilde{t} > n)) = \mathscr{E}(\mathscr{E}(\tilde{x}|\tilde{\theta})|{}_n\underline{\underline{\tilde{x}}}, (\tilde{t} > n))$$

and

$$\dot{x}_{n+1} = \alpha_0 + \sum_{i=1}^{n} \alpha_i \tilde{x}_i$$

where the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_n$ are determined by the normal equations

$$\mathscr{C}(\dot{x}_{n+1}, {}_n\underline{\underline{\tilde{x}}}'|\tilde{t} > n) = \mathscr{C}(\tilde{x}_{n+1}, {}_n\underline{\underline{\tilde{x}}}'|\tilde{t} > n) \tag{3.2}$$

$$\mathscr{E}(\dot{x}_{n+1}|\tilde{t} > n) = \mathscr{E}(\tilde{x}_{n+1}|\tilde{t} > n) \tag{3.3}$$

3D. We easily find

$$f({}_{n+1}\underline{\underline{x}}|\tilde{t}>n)=\frac{\mathscr{E}\left(\left(\prod_{i=1}^{n} g_i(\tilde{\theta},{}_{i}\underline{\underline{x}})f(x_i|\tilde{\theta})\right)f(x_{n+1}|\tilde{\theta})\right)}{\mathscr{E}\left(\prod_{i=1}^{n} g_i(\tilde{\theta},{}_{i}\underline{\underline{\tilde{x}}})\right)} \tag{3.4}$$

$$u(\theta|\tilde{t}>n,{}_{n}\underline{\underline{x}})=\frac{\prod_{i=1}^{n} g_i(\theta,{}_{i}\underline{\underline{x}})f(x_i|\theta)}{\mathscr{E}\left(\prod_{i=1}^{n} g_i(\tilde{\theta},{}_{i}\underline{\underline{x}})f(x_i|\tilde{\theta})\right)} \tag{3.5}$$

It is interesting to note that (3.4) remains unchanged if $g_i(\theta,{}_{i}\underline{\underline{x}})$ is multiplied by a constant c_i. As the coefficients $\alpha_0,\alpha_1,\ldots,\alpha_n$ of $\dot{x}_{n+1}$ are determined by the moments of the distribution $f({}_{n+1}\underline{\underline{x}}|\tilde{t}>n)$, we have that also these coefficients remain unchanged.
Furthermore, (3.5) remains unchanged when $g_i(\theta,{}_{i}\underline{\underline{x}})$ is multiplied by a function $d_i({}_{i}\underline{\underline{x}})$ of ${}_{i}\underline{\underline{x}}$. As x^*_{n+1} is an expectation with respect to the distribution $u(\theta|(\tilde{t}>n),{}_{n}\underline{\underline{x}})$, we have that also x^*_{n+1} remains unchanged.

4 Some special cases

4A. In the present section we are going to look at some different additional assumptions on the probabilities $g_i(\theta,{}_{i}\underline{\underline{x}})$.
We first assume that

$$g_i(\theta,{}_{i}\underline{\underline{x}})=h_i(\theta)g(\theta,x_i) \tag{4.1}$$

Then (3.4) becomes

$$f({}_{n+1}\underline{\underline{x}}|\tilde{t}>n)=\frac{\mathscr{E}\left(\left(\prod_{i=1}^{n} h_i(\tilde{\theta})\right)\left(\prod_{i=1}^{n} g(\tilde{\theta},x_i)f(x_i|\tilde{\theta})\right)f(x_{n+1}|\tilde{\theta})\right)}{\mathscr{C}\left(\mathscr{E}^n(g(\tilde{\theta},\tilde{x})|\tilde{\theta})\prod_{i=1}^{n} h_i(\tilde{\theta})\right)} \tag{4.2}$$

From this we see that $\tilde{x}_1,\ldots,\tilde{x}_n$ are conditionally relatively exchangeable with respect to $\tilde{x}_{n+1}$ given $(\tilde{t}>n)$ [see Sundt (1979)]. Thus $\dot{x}_{n+1}$ may be written

$$\dot{x}_{n+1}=\gamma_n+\delta_n\bar{\tilde{x}}_n$$

and from (3.2) and (3.3) we get

$$\delta_n=\frac{\lambda_n}{\tau_n} \qquad \gamma_n=\nu_n-\delta_n\mu_n$$

with

$$\lambda_n = \mathscr{C}(\tilde{x}_{n+1}, \bar{\tilde{x}}_n | \tilde{t} > n) \qquad \tau_n = \mathscr{V}(\bar{\tilde{x}}_n | \tilde{t} > n)$$

$$\mu_n = \mathscr{E}(\bar{\tilde{x}}_n | \tilde{t} > n) \qquad \nu_n = \mathscr{E}(\tilde{x}_{n+1} | \tilde{t} > n)$$

that is,

$$\dot{x}_{n+1} = \delta_n(\bar{\tilde{x}}_n - \mu_n) + \nu_n = \frac{\lambda_n}{\tau_n}(\bar{\tilde{x}}_n - \mu_n) + \nu_n \tag{4.3}$$

4B. An interesting special case of (4.1) is the case

$$g(\theta, x) = 1 \tag{4.4}$$

This means that only the risk parameter $\tilde{\theta}$ has influence on the probabilities of departure, not the claim amounts. From (3.4) we get

$$f({}_{n+1}\underline{\underline{x}} | \tilde{t} > n) = \frac{\mathscr{E}\left(\left(\prod_{i=1}^{n} h_i(\tilde{\theta})\right) \prod_{i=1}^{n+1} f(x_i | \tilde{\theta})\right)}{\mathscr{E}\left(\prod_{i=1}^{n} h_i(\tilde{\theta})\right)}$$

and we see that $\tilde{x}_1, \ldots, \tilde{x}_{n+1}$ are now conditionally exchangeable given $(\tilde{t} > n)$. This brings some simplifications in $\dot{x}_{n+1}$. We have

$$\nu_n = \mu_n \qquad \lambda_n = \mathscr{V}(\mathscr{E}(\tilde{x}_1 | \tilde{\theta}, (\tilde{t} > n)) | \tilde{t} > n)$$

$$\tau_n = \frac{1}{n} \phi_n + \lambda_n$$

with

$$\phi_n = \mathscr{E}(\mathscr{V}(\tilde{x}_1 | \tilde{\theta}, (\tilde{t} > n)) | \tilde{t} > n)$$

With $\kappa_n = \phi_n / \lambda_n$ we have

$$\dot{x}_{n+1} = \frac{n}{n + \kappa_n} \bar{\tilde{x}}_n + \frac{\kappa_n}{n + \kappa_n} \mu_n$$

For practical applications the assumption (4.4) may be criticized; it seems natural to assume that $g(\theta, x)$ depends on x as the policy may leave the portfolio because of a claim. However, as will become more clear in Section 5, assumption (4.4) has the advantage that we avoid the rather unstable estimators of ν_n and λ_n that we get under the more general assumption (4.1). This means that in some cases the choice between dropping the assumption (4.4) or not, may be considered as the choice between an appropriate model with unstable estimation methods or a (slightly?) less appropriate model with stable methods.

4C. From (3.1) we see that $\ddot{x}_{n+1}$ cannot be a better estimator of $\tilde{x}_{n+1}$ than $\dot{x}_{n+1}$. From (3.4) we see that if

$$g_i(\theta, {}_i\underline{\underline{x}}) = c \tag{4.5}$$

then

$$f({}_{n+1}\underline{\underline{x}}|\tilde{t} > n) = \mathscr{E}\left(\prod_{i=1}^{n+1} f(x_i|\tilde{\theta})\right) = f({}_{n+1}\underline{\underline{x}})$$

where $f({}_{n+1}\underline{\underline{x}})$ denotes the unconditional density of ${}_{n+1}\underline{\underline{\tilde{x}}}$, that is, the fact that $\tilde{t} > n$, does not give any information about the claim amounts. In that case $\ddot{x}_{n+1} = \dot{x}_{n+1}$ and $x^{**}_{n+1} = x^{*}_{n+1}$.

Let us assume that we have a portfolio of independent policies that all satisfy the conditions of Section 3. Then, if (4.5) holds, the $\tilde{\theta}$-value of a policy drawn at random from the portfolio, is distributed according to the density $u(\theta)$. The structural parameters ϕ, λ, and μ can be estimated from portfolio data, e.g. by the estimators proposed by Bühlmann & Straub (1970).

It is tempting to argue that also in the general case where (4.5) does not necessarily hold, one should use $\ddot{x}_{n+1}$ instead of $\dot{x}_{n+1}$, even if $\dot{x}_{n+1}$ is better; in $\dot{x}_{n+1}$ one has to estimate the structural parameters $(\tau_n, \lambda_n, \mu_n, \nu_n)$ for each n, whereas in $\ddot{x}_{n+1}$ one could use the same two structural parameters (κ, μ) for all n. Besides, for the estimation of (κ, μ) one could use the simple Bühlmann-Straub estimators.

As to the second argument one has to be a bit careful. Assume that the age $\tilde{s}$ of a policy randomly drawn from the portfolio is distributed by a distribution with point probabilities

$$p(s) = Pr(\tilde{s} = s)$$

Then the structural distribution of the $\tilde{\theta}$'s in the portfolio has the density

$$v(\theta) = \sum_s u(\theta|\tilde{t} > s - 1)\, p(s)$$

which in general is not equal to $u(\theta)$. Hence, the Bühlmann-Straub estimators are not applicable when (4.5) is not satisfied.

4D. We now assume that

$$g_i(\theta, {}_i\underline{\underline{x}}) = g_i({}_i\underline{\underline{x}}) \tag{4.6}$$

This means that given the past experience, the value of $\tilde{\theta}$ would not give any new information as to whether the policy is going to continue or not.

From (3.5) we now see that

$$u(\theta|(\tilde{t}>n), {}_{n}\underline{x}) = \frac{\prod_{i=1}^{n} f(x_i|\theta)}{\mathscr{E}\left(\prod_{i=1}^{n} f(x_i|\tilde{\theta})\right)} u(\theta) = u(\theta|{}_{n}\underline{x})$$

where $u(\theta|{}_{n}\underline{x})$ denotes the conditional density of $\tilde{\theta}$ given ${}_{n}\underline{\tilde{x}} = {}_{n}\underline{x}$. From this we see that when ${}_{n}\underline{\tilde{x}}$ is known, the fact that $\tilde{t}>n$, does not give any additional information about $\tilde{\theta}$. In particular, we get that $x^{**}_{n+1} = x^{*}_{n+1}$. However, the conditional distribution

$$u(\theta|\tilde{t}>n) = \frac{\mathscr{E}\left(\prod_{i=1}^{n} g_i({}_{i}\underline{\tilde{x}})|\theta\right)}{\mathscr{E}\left(\prod_{i=1}^{n} g_i({}_{i}\underline{\tilde{x}})\right)} u(\theta)$$

still depends on n.

4E. We now assume that x^{**}_{n+1} is linear in $\tilde{x}_1, \ldots, \tilde{x}_n$, that is, that $x^{**}_{n+1} = \ddot{x}_{n+1}$. This is the case in some parametric classes [Jewell (1974, 1975b), Diaconis & Ylvisaker (1979)] and the nonparametric class introduced by Ferguson (1973) [see also Zehnwirth (1977, 1979)].
We also assume that (4.6) holds. This implies that $x^{**}_{n+1} = x^{*}_{n+1}$. But then x^{*}_{n+1} is linear, and thus $x^{*}_{n+1} = \dot{x}_{n+1}$. Hence, we have

$$x^{*}_{n+1} = x^{**}_{n+1} = \dot{x}_{n+1} = \ddot{x}_{n+1}$$

A natural question is: Are $\dot{x}_{n+1}$ and $\ddot{x}_{n+1}$ also equal under (4.6) if we do not make the assumption that $x^{**}_{n+1} = \ddot{x}_{n+1}$? From the following theorem follows that the answer is in general not yes.

Theorem 4.1. *If $\dot{x}_{n+1} = \ddot{x}_{n+1}$ for all choices of $g_i({}_{i}\underline{x})$, then $x^{**}_{n+1} = \ddot{x}_{n+1}$.*

For proof, we refer to Sundt (1982).
In most cases, x^{**}_{n+1} is not linear. From Theorem 4.1 follows that in such cases there exists at least one censoring mechanism of form (4.6) such that $\dot{x}_{n+1} \neq \ddot{x}_{n+1}$, that is, from (4.6) we cannot conclude that $\dot{x}_{n+1} = \ddot{x}_{n+1}$ if we are not willing to assume that x^{**}_{n+1} is linear.

4F. For the rest of Section 4 we assume that $g_i({}_{i}\underline{x}) = g(x_i)$. This is a special case of (4.1), and hence $\dot{x}_{n+1}$ is given by (4.3). Formula (4.2) may now be written

$$f({}_{n+1}\underline{x}|\tilde{t}>n) = \frac{\prod_{i=1}^{n} g(x_i)}{\mathscr{E}\mathscr{E}^{n}(g(\tilde{x})|\tilde{\theta})} f({}_{n+1}\underline{x})$$

The condition $\dot{x}_{n+1}=\ddot{x}_{n+1}$ may now be written

$$\frac{n}{n+\kappa}\,\bar{\bar{x}}_n+\frac{\kappa}{n+\kappa}\,\mu=\delta_n(\bar{\bar{x}}_n-\mu_n)+\nu_n$$

that is,

$$\frac{n}{n+\kappa}=\delta_n \qquad \frac{\kappa}{n+\kappa}\,\mu=\nu_n-\delta_n\mu_n$$

and by solving for κ and μ, we obtain

$$\kappa=n(\delta_n^{-1}-1)=n\left(\frac{\tau_n}{\lambda_n}-1\right)$$

$$\mu=\frac{n}{\kappa}(\nu_n-\mu_n)+\nu_n \tag{4.7}$$

We have the following analogue to Theorem 4.1.

Theorem 4.2. *If $\dot{x}_{n+1}=\ddot{x}_{n+1}$ for all choices of $g(x)$, and x^{**}_{n+1} depends on ${}_n\underline{\tilde{x}}$ only through $\bar{\bar{x}}_n$, then $x^{**}_{n+1}=\ddot{x}_{n+1}$.*

We close Section 4 by looking at a parametric example.

Example 4.1. We assume that

$$f(x|\theta)=Pr(\tilde{x}=x|\theta)=\frac{\theta^x}{x!}\,e^{-\theta}, \qquad (x=0,1,2,\ldots) \tag{4.8}$$

and that

$$g(x)=(1-\varrho)^x$$

for some $\varrho\in[0,1\rangle$. Let

$$\psi(s)=\mathscr{E}(e^{-s\tilde{\theta}})$$

be the Laplace transform of $\tilde{\theta}$. Then

$$\nu_n=-\frac{\psi'(n\varrho)}{\psi(n\varrho)} \qquad \mu_n=(1-\varrho)\,\nu_n \tag{4.9}$$

$$\lambda_n=(1-\varrho)\left[\frac{\psi''(n\varrho)}{\psi(n\varrho)}-\nu_n^2\right] \qquad \tau_n=(1-\varrho)\left[\lambda_n+\frac{\mu_n}{n}\right]$$

An interesting question is now: For which densities u of $\tilde{\theta}$ do we have $\dot{x}_{n+1}=\ddot{x}_{n+1}$ for all ϱ? By inserting (4.9) in (4.7) we obtain

$$\mu=-\left(\frac{s}{\kappa}+1\right)\frac{\psi'(s)}{\psi(s)}$$

with $s = n\varrho$. By solving this differential equation we get

$$\psi(s) = \left(\frac{\kappa}{\kappa + s}\right)^{\alpha}$$

with $\alpha = \mu\kappa$. But this is the Laplace transform of the Gamma density

$$u(\theta) = \frac{\kappa^{\alpha}}{\Gamma(\alpha)} \theta^{\alpha - 1} e^{-\kappa\theta}$$

If $\tilde{\theta}$ has this density, then

$$x^{*}_{n+1} = x^{**}_{n+1} = \dot{x}_{n+1} = \ddot{x}_{n+1} = \frac{\alpha + n\bar{\bar{x}}_n}{\kappa + n}$$

Combining this with the fact that if x^{**}_{n+1} is linear, then $\dot{x}_{n+1} = \ddot{x}_{n+1}$ for all choices of $g(x)$, we get that the Gamma densities are the only densities giving linear x^{**}_{n+1} when $f(x|\theta)$ satisfies (4.8). This result was proved more directly by Johnson (1957).

5 A numerical example

5A. In the following numerical example data from «Winterthur» Swiss Insurance Company were used. From the portfolio of compulsory automobile liability policies being in force per 31.12.80, policies satisfying the following criteria were extracted:

a) Passenger car for private use
b) Cylinder volume 1393–2963 cm^3
c) Final digits of the policy number: 12, 35, 58 or 79.

Criterion c) was introduced as a procedure for random sampling. Criterion b) may be criticized; it means cylinder volume per 31.12.80, but the cylinder volume may have been different earlier. This means that the claim numbers may have greater variances in earlier years. It should be mentioned that in «Winterthur» the policyholder usually keeps the same policy number when he changes his car. The extraction gave in all 11015 policies, but as the real age of the older policies was uncertain, we used for the investigations only policies originating not earlier than 1962.

As a simplification we assumed that for all policies the insurance year followed the calendar year. To obtain this, we defined the first insurance year to be the first whole calendar year the policy was in force, and we just neglected the data from the policy before it entered that year. We have not done anything to adjust for the fact that not all policies cancel at the end of a calendar year.

Table 5.1

Year of origin	Number of policies	Claims per policy year in					
		1976	1977	1978	1979	1980	1976–80
1963	470	0.05319	0.05319	0.07021	0.06596	0.08511	0.06553
1964	446	0.06951	0.07848	0.06726	0.05830	0.03140	0.06099
1965	520	0.07885	0.06731	0.06923	0.06346	0.06346	0.06846
1966	574	0.07317	0.07317	0.06969	0.05575	0.05749	0.06585
1967	533	0.06379	0.06379	0.07129	0.06567	0.08255	0.06942
1968	421	0.05463	0.08314	0.07126	0.05701	0.04038	0.06128
1969	378	0.07672	0.07143	0.08201	0.05820	0.07407	0.07248
1970	464	0.05819	0.04957	0.07112	0.07759	0.06897	0.06509
1971	528	0.07386	0.08902	0.07386	0.06250	0.05492	0.07083
1972	546	0.06272	0.05678	0.08242	0.06960	0.05678	0.06557
1973	523	0.07266	0.06119	0.08413	0.10134	0.07648	0.07916
1974	459	0.07190	0.07843	0.07190	0.05011	0.04357	0.06318
1975	496	0.07056	0.08266	0.09274	0.07863	0.06048	0.07702
1976	465	0.06882	0.06237	0.09892	0.08602	0.05591	0.07441
1977	503	—	0.08748	0.10537	0.06560	0.06163	0.08002
1978	609	—	—	0.07061	0.06404	0.07882	0.07115
1979	610	—	—	—	0.07869	0.10000	0.08934
1980	635	—	—	—	—	0.10709	0.10709
1963–80	9180	0.06786	0.07043	0.07813	0.06846	0.06808	0.07056

For each policy we registered the claim number from each year it had been in force in the period 1976–80. In Table 5.1 we show the observed claim frequencies.

5B. In the present subsection we want to test whether the age of the policy does give any information about the value of $\tilde{\theta}$. Let

$$F_n(x) = Pr(\tilde{x}_n \leqslant x | \tilde{t} \geqslant n)$$

To test the hypothesis

$$(H) \qquad F_1 = F_2 = \ldots = F_{18}$$

against the general alternative that some of these distributions are different from each other, we used the homogeneity test described in Sverdrup (1967, Section XV.2.4) on the claim numbers from 1980. This gives a significance probability 0.003813, and thus we reject that the claim numbers from different years of origin have the same distribution.

It would also be natural to test (H) under the apriori condition

$$F_1 \geqslant F_2 \geqslant \ldots \geqslant F_{18}$$

[stochastic ordering, $F_i(x) \leqslant F_{i+1}(x)$ for all i and x]. This means that we test equality against the alternative that the more risky policies are more apt to leave the portfolio. For this situation we used the Jonckheere-Terpstra test described in Lehmann (1975, pp. 233–235) on the claim numbers from 1980. The normal approximation gives a significance probability 0.0003878.
It should be noted that the small significance probabilities do not necessarily mean that the durational effects are extreme; *statistical* significance is not always the same as *practical* significance. We have a very great sample, and even by small deviations from the hypothesis, we could have small significance probabilities. However, referring to Table 5.4 below, in the present case it seems that we also have practical significance.

5C. For the rest of the paper, we are going to discuss estimation of structural parameters. We assume that all the observed policies are independent and having the same structural density $u(\theta)$, sample density $f(x|\theta)$, and probabilities of remaining in the portfolio $g_i(\theta, {}_i\underline{x})$ satisfying (4.1). Then $\dot{x}_{n+1}$ is of the form (4.3), and we want to estimate the structural parameters μ_n, ν_n, τ_n, and λ_n.
For $n=0, 1, \ldots, 17$, let ${}_nN$ be the number of policies having been in force for exactly $n+1$ years, that is, the policies originating from calendar year $1980-n$. Let ${}_n\tilde{x}_{ij}$ be the claim number from year j of the i-th of these policies. We introduce

$$
{}_n\tilde{x}_{i.} = \begin{cases} \dfrac{1}{n} \sum\limits_{j=1}^{n} {}_n\tilde{x}_{ij} & (n \leqslant 4) \\[2ex] \dfrac{1}{4} \sum\limits_{j=n-3}^{n} {}_n\tilde{x}_{ij} & (n > 4) \end{cases}
$$

$$
{}_n\tilde{x}_{.j} = \frac{1}{{}_nN} \sum_{i=1}^{{}_nN} {}_n\tilde{x}_{ij} \qquad {}_n\tilde{x}_{..} = \frac{1}{{}_nN} \sum_{i=1}^{{}_nN} {}_n\tilde{x}_i
$$

For μ_n, ν_n, τ_n, and λ_n we introduce the unbiased estimators

$$
\hat{\mu}_n = {}_n\tilde{x}_{..} \qquad \hat{\nu}_n = {}_n\tilde{x}_{.n+1}
$$

$$
\hat{\tau}_n = \begin{cases} \dfrac{1}{{}_nN-1} \sum\limits_{i=1}^{{}_nN} ({}_n\tilde{x}_{i.} - {}_n\tilde{x}_{..})^2 & (n \leqslant 4) \\[2ex] \dfrac{1}{{}_nN-1} \sum\limits_{i=1}^{{}_nN} ({}_n\tilde{x}_{i.} - {}_n\tilde{x}_{..})^2 + \left(\dfrac{1}{4} - \dfrac{1}{n}\right) \dfrac{1}{3{}_nN} \sum\limits_{i=1}^{{}_nN} \sum\limits_{j=n-3}^{n} ({}_n\tilde{x}_{ij} - {}_n\tilde{x}_i)^2 & (n > 4) \end{cases}
$$

$$
\hat{\lambda}_n = \frac{1}{{}_nN-1} \sum_{i=1}^{{}_nN} ({}_n\tilde{x}_{i.} - {}_n\tilde{x}_{..})({}_n\tilde{x}_{i,n+1} - {}_n\tilde{x}_{.n+1})
$$

This procedure works for $n>0$. For $n=0$ $\hat{\mu}_n$, $\hat{\tau}_n$, and $\hat{\lambda}_n$ are not defined. However, as $\dot{x}_1=v_0$, they are not needed.
Even if the estimators $\hat{\mu}_n$, $\hat{v}_n$, $\hat{\tau}_n$, and $\hat{\lambda}_n$ are unbiased, the estimators

$$\hat{\delta}_n=\frac{\hat{\lambda}_n}{\hat{\tau}_n} \qquad \hat{\gamma}_n=\hat{v}_n-\hat{\delta}_n\hat{\mu}_n$$

of δ_n and γ_n are not necessarily unbiased as they are non-linear functions of the unbiased quantities. For small ${}_nN$ the bias may be considerable. However, the estimators are consistent as ${}_nN\uparrow\infty$.

Table 5.2

n	$\hat{\tau}_n$	$\hat{\lambda}_n$	$\hat{\mu}_n$	$\hat{v}_n$	$\hat{\delta}_n$	$\hat{\gamma}_n$
0	—	—	—	0.10709	0	0.10709
1	0.082468	0.001970	0.07869	0.10000	0.023893	0.09812
2	0.033289	0.002086	0.06732	0.07882	0.062672	0.07460
3	0.036388	0.000656	0.08615	0.06163	0.018030	0.06008
4	0.019737	0.003115	0.07903	0.05591	0.157802	0.04344
5	0.018090	0.004173	0.08115	0.06048	0.230665	0.04177
6	0.012016	−0.001881	0.06808	0.04357	−0.156564	0.05423
7	0.015594	0.003940	0.07983	0.07648	0.252680	0.05631
8	0.009636	0.002567	0.06777	0.05678	0.266457	0.03872
9	0.008539	0.002050	0.07481	0.05492	0.240112	0.03696
10	0.009669	0.000428	0.06412	0.06897	0.044289	0.06613
11	0.010031	−0.000049	0.07209	0.07407	−0.004897	0.07443
12	0.008191	0.000879	0.06651	0.04038	0.107362	0.03324
13	0.007748	0.002049	0.06614	0.08255	0.264444	0.06506
14	0.007924	0.001323	0.06794	0.05749	0.166910	0.04615
15	0.005292	0.001348	0.06971	0.06346	0.254657	0.04571
16	0.008193	−0.001028	0.06839	0.03139	−0.125464	0.03997
17	0.003781	0.005489	0.06064	0.08511	1.451663	−0.00292

In Table 5.2 we have given the values of $\hat{\mu}_n$, $\hat{v}_n$, $\hat{\tau}_n$, $\hat{\lambda}_n$, $\hat{\delta}_n$, and $\hat{\gamma}_n$ resulting from our numerical data. It is seen that when n varies, the values of $\hat{\delta}_n$ are very unstable. As this could be partly due to the above-mentioned bias, we also used the first order jackknife [see e.g. Zehnwirth (1981)] on $\hat{\delta}_n$ and $\hat{\gamma}_n$. The jackknife is known to have a bias-reducing effect. However, in the present case, this effect was negligible, and in Table 5.2 we have only given the unjackknifed estimates. For the jackknifed estimates we refer to Sundt (1982), where we have done most of the subsequent computations also on jackknifed estimates.

5D. It is clear that the present estimates cannot be used directly in a rating scheme; they have properties that seem most unfair to the policy-holders. For instance, for a six years old policy the premium would be a decreasing function of the average claim number $\bar{\bar{x}}_6$. We are now going to introduce some properties that we want δ_n and γ_n to possess.
From the above example we clearly want

$$\text{i)} \quad \delta_n \geqslant 0$$

Furthermore,

$$\text{ii)} \quad \gamma_n \geqslant 0$$

Even with no claims in the past the policy-holder should not have a negative premium.
Let us now assume that the policy has no claims in year $n+1$. Then it would clearly seem unfair if the policy-holder should get a premium increase in year $n+2$, that is, we want

$$\delta_{n+1} \frac{n}{n+1} \bar{\bar{x}}_n + \gamma_{n+1} \leqslant \delta_n \bar{\bar{x}}_n + \gamma_n$$

A sufficient condition for this inequality to hold, is that

$$\text{iii)} \quad \frac{\delta_{n+1}}{n+1} \leqslant \frac{\delta_n}{n}$$

$$iv) \quad \gamma_{n+1} \leqslant \gamma_n$$

It should be noted that all of these conditions are violated by the estimates in Table 5.2.

5E. The conditions given in the previous subsection motivate graduation of the estimates $(\hat{\gamma}_n, \hat{\delta}_n)'$ against some parametric functions. Another reason for doing such a graduation is that we may then get estimates also for $n > 17$.
In the next two subsections we are going to discuss some graduation procedures. As it seems difficult to say anything about variances and covariances of the ungraduated quantities, it seems to be too ambitious to try to search for optimal graduation methods. Thus we are going to propose simple and intuitively reasonable methods without claiming any optimality properties. When we need weights, we simply use the number of policies.

5F. In this subsection we assume that

$$\delta_n = \frac{n}{n+\kappa'} \qquad \gamma_n = \frac{\kappa'}{n+\kappa'} \mu' \tag{5.1}$$

This is in particular the case when $\dot{x}_{n+1} = \ddot{x}_{n+1}$. However, the prime is used to indicate that (5.1) may be satisfied under more general conditions, and that we do not necessarily have $\kappa' = \kappa$ and $\mu' = \mu$ (κ and μ being defined in subsection 3B). From (5.1) and (4.3) we obtain

$$\kappa' = \frac{n(\tau_n - \lambda_n)}{\lambda_n} \qquad \mu' = \frac{n}{\kappa'}(v_n - \mu_n) + v_n$$

and we propose

$$\hat{\kappa}' = \frac{\sum_n {}_nNn(\hat{\tau}_n - \hat{\lambda}_n)}{\sum_n {}_nN\hat{\lambda}_n} \tag{5.2}$$

$$\hat{\mu}' = \frac{1}{\sum_n {}_nN} \sum_n {}_nN\left[\frac{n}{\hat{\kappa}'}(\hat{v}_n - \hat{\mu}_n) + \hat{v}_n\right] \tag{5.3}$$

as estimators for κ' and μ'.

Table 5.3

n	Present procedure		Generalized Bühlmann-Straub	
	δ_n	γ_n	$n/(n+\kappa)$	$\mu\kappa/(n+\kappa)$
0	0	0.06676	0	0.07073
1	0.02302	0.06522	0.03758	0.06807
2	0.04500	0.06376	0.07243	0.06561
3	0.06601	0.06235	0.10485	0.06331
4	0.08612	0.06101	0.13508	0.06117
5	0.10538	0.05973	0.16334	0.05918
6	0.12385	0.05849	0.18980	0.05730
7	0.14157	0.05731	0.21465	0.05555
8	0.15858	0.05617	0.23801	0.05389
9	0.17494	0.05508	0.26002	0.05234
10	0.19067	0.05403	0.28081	0.05087
11	0.20581	0.05302	0.30044	0.04948
12	0.22040	0.05205	0.31905	0.04816
13	0.23446	0.05111	0.33668	0.04692
14	0.24802	0.05020	0.35343	0.04573
15	0.26111	0.04933	0.36935	0.04460
16	0.27376	0.04848	0.38451	0.04353
17	0.28597	0.04769	0.39895	0.04251

From the «Winterthur» data we get $\hat{\kappa}'=42.4462$ and $\hat{\mu}'=0.06676$. As a comparison we also estimated κ and μ by the generalized Bühlmann-Straub procedure described by Sundt (1983, Section 3). This procedure does not take into account that the duration may give information about the risk parameter. We get 25.6118 and 0.07073 as estimates for κ and μ. It is interesting to note the great differences in the estimates of the two procedures, a strong indication that the duration should be taken into account. In Table 5.3 we have computed δ_n and γ_n given by (5.1) with κ' and μ' replaced by the estimates above. As a comparison we have also computed the coefficients of $\ddot{x}_{n+1}$ with κ and μ replaced by the Bühlmann-Straub estimates above.
To further illustrate the importance of including the durational effects in the model, we have in Table 5.4 for different values of n and $\sum_{i=1}^{n} \tilde{x}_i$ computed the premium increase in percent by using the Bühlmann-Straub procedure instead of the present procedure.

Table 5.4

n \ $\sum_{i=1}^{n} \tilde{x}_i$	0	1	2	3	4	5	6	7	8	9	10	∞
0	5.9	—	—	—	—	—	—	—	—	—	—	—
1	4.4	19.7	28.7	34.7	38.8	42.0	44.4	46.3	47.9	49.2	50.3	63.3
2	2.9	18.0	26.9	32.8	36.9	40.0	42.3	44.2	45.8	47.1	48.1	61.0
3	1.5	16.5	25.2	31.0	35.1	38.1	40.5	42.3	43.8	45.1	46.2	58.8
4	0.3	15.0	23.9	29.4	33.4	36.4	38.7	40.5	42.1	43.3	44.4	56.9
5	−0.9	13.7	22.2	27.8	31.8	34.8	37.1	38.9	40.4	41.6	42.6	55.0
6	−2.0	12.4	20.8	26.4	30.3	33.3	35.5	37.3	38.8	40.0	41.0	53.3
7	−3.1	11.2	19.6	25.1	28.9	31.8	34.1	35.9	37.3	38.5	39.5	51.6
8	−4.1	10.1	18.3	23.8	27.6	30.5	32.7	34.5	35.9	37.1	38.1	50.1
9	−5.0	9.0	17.2	22.6	26.4	29.2	31.4	33.2	34.6	35.8	36.8	48.6
10	−5.9	8.0	16.1	21.5	25.2	28.1	30.2	32.0	33.4	34.6	35.5	47.3
∞	−36.1	−26.7	−21.1	−17.5	−15.0	−13.0	−11.6	−10.4	−9.4	−8.6	−8.0	0

5G. Let us now assume that

$$\delta_n=\frac{n}{\alpha+n\sigma} \qquad \gamma_n=\frac{\beta+n\varrho}{\alpha+n\sigma} \tag{5.4}$$

This assumption is fulfilled in some parametric classes discussed in Sundt (1982). It seems to be very difficult to find good estimators for α, β, σ, and ϱ without making further assumptions. We just mention some attempts for this estimation and discuss their weaknesses.

We concentrate on estimating α and σ by the $\hat{\delta}_n$'s. When we have found estimates $\hat{\alpha}$ and $\hat{\sigma}$, we propose to estimate β and ϱ by the estimators $\hat{\beta}$ and $\hat{\varrho}$ minimizing

$$Q_2(\beta, \varrho) = \sum_n {}_nN \left(\hat{\gamma}_n - \frac{\beta + n\varrho}{\hat{\alpha} + n\hat{\sigma}} \right)^2$$

i) The perhaps most natural procedure would be to minimize

$$Q_1(\alpha, \sigma) = \sum_n {}_nN \left(\hat{\delta}_n - \frac{n}{\alpha + n\sigma} \right)^2 \tag{5.5}$$

By putting the partial derivatives of Q_1 with respect to α and σ equal to zero, we get a system of two equations in α and σ. These equations have to be solved numerically. Unfortunately they have several solutions, and it thus seems to be quite a lot of work to find the values minimizing Q_1.

ii) If we instead of minimizing (5.5) minimize

$$Q_1(\alpha, \sigma) = \sum_n {}_nN \left(\frac{1}{\hat{\delta}_n} - \frac{\alpha}{n} - \sigma \right)^2$$

we have reduced the problem to an ordinary linear regression. However, in many practical applications we risk that some of the δ_n's are close to zero; they may even be negative. Such values will of course make the estimators of α and σ extremely unstable, and we may get completely wild estimates.

iii) Minimize

$$Q_1(\alpha, \sigma) = \sum_n {}_nN((\alpha + n\sigma)\hat{\delta}_n - n)^2$$

This expression can be rewritten

$$Q_1(\alpha, \sigma) = \sum_n {}_nN(\alpha + n\sigma)^2 \left(\hat{\delta}_n - \frac{n}{\alpha + n\sigma} \right)^2 \tag{5.6}$$

In practical applications we may risk that for some value of n, for which we have given a $\hat{\delta}_n$, the estimated $\alpha + n\sigma$ is close to zero. Then $n/(\alpha + n\sigma)$ could be a very bad estimate of δ_n. But as in (5.6) such estimates are given very little weight, they do not influence very much the quantity to be minimized.

In Table 5.5 we have estimated the parameters α, σ, β, and ϱ by Procedures ii) and iii) above using the $(\hat{\delta}_n, \hat{\gamma}_n)$'s of Table 5.2, and in Table 5.6 these estimates have

Table 5.5

	Procedure ii)	Procedure iii)
α	63.2760	42.6411
σ	−12.2680	− 1.6949
β	3.5465	3.1563
ϱ	− 0.6902	− 0.1719

Table 5.6

	Procedure ii)		Procedure iii)	
n	δ_n	γ_n	δ_n	γ_n
0	0	0.05605	0	0.07402
1	0.01960	0.05600	0.02442	0.07289
2	0.05163	0.05591	0.05095	0.07165
3	0.11333	0.05575	0.07988	0.07031
4	0.28161	0.05530	0.11154	0.06884
5	2.58250	0.04922	0.14634	0.06722
6	−0.58073	0.05758	0.18478	0.06543
7	−0.30974	0.05687	0.22744	0.06345
8	−0.22944	0.05665	0.27508	0.06124
9	−0.19094	0.05655	0.32862	0.05875
10	−0.16834	0.05649	0.38922	0.05593
11	−0.15348	0.05645	0.45838	0.05272
12	−0.14296	0.05643	0.53805	0.04902
13	−0.13512	0.05641	0.63083	0.04471
14	−0.12906	0.05639	0.74024	0.03962
15	−0.12423	0.05638	0.87118	0.03354
16	−0.12029	0.05637	1.03072	0.02613
17	−0.11702	0.05636	1.22937	0.01690

been inserted in (5.4). None of these sets of estimated credibility coefficients satisfy all the conditions of subsection 5D.

The set found by Procedure ii) does not seem to have very much to do with the ungraduated set of Table 5.2. It seems that the ungraduated $\hat{\delta}_{11}$ must have had a disastrous effect. What we really do in our graduation, is linear regression on $\hat{\delta}_n^{-1}$; we would have expected $\hat{\delta}_{11}^{-1}$ to be something positive, but we actually have $\hat{\delta}_{11}^{-1} = -204.2$!

The set found by Procedure iii) looks much more reasonable.

5H. For the rest of the paper we make the assumption (4.4) in addition to (4.1). We have to estimate the three structural parameters μ_n, ϕ_n, and λ_n, and propose the unbiased estimators $\hat{\mu}_n = {}_n\tilde{x}_{..}$

$$\hat{\phi}_n = \frac{1}{{}_nNk_n} \sum_{i=1}^{{}_nN} \sum_{j=n-k_n+1}^{n+1} ({}_n\tilde{x}_{ij} - {}_n\tilde{x}_{i\cdot})^2$$

$$\hat{\lambda}_n = \frac{1}{{}_nN-1} \sum_{i=1}^{{}_nN} ({}_n\tilde{x}_{i\cdot} - {}_n\tilde{x}_{..})^2 - \frac{\hat{\phi}_n}{k_n+1}$$

with

$${}_n\tilde{x}_{i\cdot} = \frac{1}{k_n+1} \sum_{j=n-k_n+1}^{n+1} {}_n\tilde{x}_{ij}$$

$${}_n\tilde{x}_{..} = \frac{1}{{}_nN} \sum_{i=1}^{{}_nN} {}_n\tilde{x}_{i\cdot}$$

$$k_n = \min(n, 4)$$

(note that the notations have other meanings than in subsection 5C). If we are going to use the ungraduated estimates in the credibility estimators, one should replace $\hat{\lambda}_n$ by max $(\hat{\lambda}_n, 0)$ as λ_n is non-negative. The estimators introduced are in

Table 5.7

n	$\hat{\phi}_n$	$\hat{\lambda}_n$	$\hat{\tau}_n$	$\hat{\mu}_n$	$\hat{\delta}_n$	$\hat{\gamma}_n$
0	—	—	—	0.10709	0	0.10709
1	0.09262	0.001933	0.094556	0.08934	0.02044	0.08752
2	0.06623	0.002092	0.035207	0.07115	0.05943	0.06693
3	0.08565	0.002923	0.031474	0.08002	0.09287	0.07259
4	0.07505	0.000734	0.019497	0.07441	0.03764	0.07161
5	0.07923	0.002371	0.018218	0.07702	0.13017	0.06699
6	0.06797	−0.000914	0.010415	0.06318	−0.08777	0.06873
7	0.08050	0.003906	0.015406	0.07916	0.25354	0.05909
8	0.06630	0.001589	0.009876	0.06557	0.16088	0.05502
9	0.07102	0.000882	0.008774	0.07083	0.10057	0.06371
10	0.06379	0.002262	0.008641	0.06509	0.26174	0.04805
11	0.07143	0.002197	0.008690	0.07249	0.25278	0.05416
12	0.06033	0.001979	0.007006	0.06128	0.28239	0.04398
13	0.06811	0.002528	0.007767	0.06942	0.32553	0.04682
14	0.06463	0.002485	0.007101	0.06585	0.34988	0.04281
15	0.06904	0.000918	0.005521	0.06846	0.16633	0.05707
16	0.06166	0.001918	0.005772	0.06099	0.33231	0.04072
17	0.06234	0.002333	0.006000	0.06553	0.38881	0.04005

full accordance with the estimators proposed by Bühlmann and Straub (1970). As estimators for τ_n, δ_n, and γ_n we use

$$\hat{\tau}_n = \frac{1}{n}\,\hat{\phi}_n + \hat{\lambda}_n \qquad \hat{\delta}_n = \frac{\hat{\lambda}_n}{\hat{\tau}_n} \qquad \hat{\gamma}_n = (1-\hat{\delta}_n)\,\hat{\mu}_n$$

In Table 5.7 we have given the values of $\hat{\phi}_n$, $\hat{\lambda}_n$, $\hat{\tau}_n$, $\hat{\mu}_n$, $\hat{\delta}_n$, and $\hat{\gamma}_n$ resulting from our numerical data. The estimates are seen to be much more stable than the ones in Table 5.2. In particular, it is astonishing to see how different the estimates of δ_n are from the ones in Table 5.2. The reason seems to be the great unstability of the estimates of λ_n in that table.

5I. For the present subsection we make the assumptions (5.1). Then, analogously to (5.2) and (5.3), we get the following estimators for κ' and μ'

$$\hat{\kappa}' = \frac{\sum_n {}_nN\hat{\phi}_n}{\sum_n {}_nN\hat{\lambda}_n} \qquad \hat{\mu}' = \frac{\sum_n {}_nN\hat{\mu}_n}{\sum_n {}_nN}$$

Table 5.8

n	δ_n	γ_n
0	0	0.07345
1	0.02608	0.07154
2	0.05083	0.06972
3	0.07436	0.06799
4	0.09675	0.06635
5	0.11808	0.06478
6	0.13842	0.06328
7	0.15785	0.06186
8	0.17643	0.06049
9	0.19420	0.05919
10	0.21122	0.05794
11	0.22753	0.05674
12	0.23419	0.05559
13	0.25822	0.05448
14	0.27267	0.05342
15	0.28656	0.05240
16	0.29994	0.05142
17	0.31282	0.05047

When using these estimators on our numerical data, we obtain $\hat{\kappa}'=37.3448$ and $\hat{\mu}'=0.07345$. In Table 5.8 we have computed δ_n and γ_n given by (5.1) with κ' and μ' replaced by these estimates.

5J. For the present subsection we make the assumptions (5.4). In Table 5.9 we have estimated the parameters α, σ, β, and ϱ by Procedures ii) and iii) of subsection 5G using the $(\hat{\delta}_n, \hat{\gamma}_n)$'s of Table 5.7, and in Table 5.10 these estimates have been inserted in (5.4). The graduated credibility coefficients of Table 5.10 seem much more reasonable than the corresponding ones in Table 5.6, and the ones obtained by Procedure iii) satisfy all the conditions of subsection 5D.

Table 5.9

	Procedure ii)	Procedure iii)
α	46.9086	24.3015
σ	− 0.7681	1.4327
β	3.8972	2.2538
ϱ	− 0.1614	− 0.0178

Table 5.10

	Procedure ii)		Procedure iii)	
n	δ_n	γ_n	δ_n	γ_n
0	0	0.08308	0	0.09274
1	0.02167	0.08097	0.03886	0.08689
2	0.04408	0.07878	0.07362	0.08166
3	0.06726	0.07651	0.10490	0.07694
4	0.09125	0.07417	0.13319	0.07268
5	0.11609	0.07175	0.15891	0.06881
6	0.14184	0.06923	0.18238	0.06527
7	0.16854	0.06663	0.20390	0.06203
8	0.19625	0.06392	0.22369	0.05905
9	0.22502	0.06111	0.24196	0.05630
10	0.25492	0.05819	0.25888	0.05375
11	0.28601	0.05516	0.27458	0.05139
12	0.31837	0.05200	0.28920	0.04918
13	0.35207	0.04871	0.30284	0.04713
14	0.38721	0.04528	0.31560	0.04521
15	0.42387	0.04170	0.32757	0.04340
16	0.46216	0.03796	0.33881	0.04171
17	0.50219	0.03405	0.34938	0.04012

5K. In Section 5 we have used several assumptions and several estimation procedures on the «Winterthur» data, and we now try to make some conclusions.
i) It seems clear that the durational effects should be taken into account. Strong indications for this are the test of subsection 5B and the comparison with the Bühlmann-Straub estimators in Subsection 5F.
ii) The estimators are very unstable. To get reasonable results one should, preferably, have a very great quantity of data. It also seems necessary to make additional assumptions that make smoothing or similar procedures possible.
iii) The gain by jackknifing the estimators is negligible.

Acknowledgements

The present research was financed by the Association of Norwegian Insurance Companies and the Norwegian Research Council for Science and the Humanities. Data were provided by «Winterthur» Swiss Insurance Company; I would in particular like to thank Fritz Bichsel in this connection. I am grateful to Ragnar Norberg, Hans Bühlmann, Andre Dubey, and Nils Lid Hjort for useful suggestions and discussions.

Bjørn Sundt
Forsikringsrådet
Postboks 4301, Torshov
Oslo 4
Norway

References

Bühlmann, H. (1967), Experience rating and credibility. *Astin Bulletin 4*, 199–207.

Bühlmann, H. and *Straub, E.* (1970), Glaubwürdigkeit für Schadensätze. *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker 70*, 111–133.

Diaconis, P. and *Ylvisaker, D.* (1979), Conjugate priors for exponential classes. *Annals of Statistics 7*, 269–281.

Ferguson, T. S. (1973), A Bayesian analysis of some nonparametric models. *Annals of Statistics 1*, 209–230.

Jewell, W. S. (1974), Credibility means are exact Bayesian for exponential families. *Astin Bulletin 8*, 77–90.

Jewell, W. S. (1975a), Model variations in credibility theory. In Kahn (1975), 193–244.

Jewell, W. S. (1975b), Regularity conditions for exact credibility. *Astin Bulletin 8*, 336 – 341.

Johnson, N. L. (1957), Uniqueness in a result in the theory of accident proneness. *Biometrika 44*, 530–531.

Kahn, P. M. (ed.) (1975), *Credibility: Theory and applications*. Academic Press, New York.

Lehmann, E. L. (1975), *Nonparametrics: Statistical methods based on ranks*. Holden-Day, San Francisco.

Sverdrup, E. (1967), *Laws and chance variations*. Vol. II. North-Holland, Amsterdam.

Sundt, B. (1979), Credibility and relative exchangeability. *Scandinavian Actuarial Journal*, 45–51.

Sundt, B. (1982), Credibility models allowing durational effects. *Doctoral dissertation* No 7141, ETH Zürich.

Sundt, B. (1983), Parameter estimation in some credibility models. To be published in *Scandinavian Actuarial Journal.*

Taylor, G. C. (1975), Credibility under conditions of imperfect persistency. In Kahn (1975), 391–400.

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

Zehnwirth, B. (1979), Credibility and the Dirichlet process. *Scandinavian Actuarial Journal*, 13–23.

Summary

In a classical credibility model it is assumed that the total claim amounts from different years are conditionally independent and identically distributed, given an unknown random risk parameter $\tilde{\theta}$. In the present paper, we introduce an additional random variable $\tilde{t}$, denoting the total time the policy stays in the portfolio. It is assumed that information about $\tilde{t}$ may say something about the value of $\tilde{\theta}$, and it should therefore be used in the rating scheme. In such models we deduce and discuss credibility estimators. In connection to a numerical example with real automobile liability data, we first test under different prior assumptions the hypothesis that there are no durational effects. The hypothesis is rejected. Then we propose estimators for structural parameters under various assumptions.

Zusammenfassung

In den klassischen Kredibilitätsmodellen wird davon ausgegangen, dass die Schadentotale von verschiedenen Jahren für einen gegebenen Wert des unbekannten Zufalls-Risikoparameters $\tilde{\theta}$ voneinander unabhängig und gleichverteilt sind. In der vorliegenden Arbeit wird eine zusätzliche Zufallsvariable $\tilde{t}$ eingeführt, welche die Gesamtzeit misst, welche eine Police im Portefeuille bleibt. Es wird angenommen, dass Informationen über $\tilde{t}$ Aussagen über den Wert von $\tilde{\theta}$ gestatten, $\tilde{t}$ sollte daher im Tarifierungssystem benutzt werden. In solchen Modellen werden Kredibilitätsschätzer hergeleitet und diskutiert. In einem numerischen Beispiel aus der Praxis der Automobil-Haftpflichtversicherung wird unter verschiedenen a priori Bedingungen die Hypothese getestet, dass keine dauerabhängigen Effekte vorliegen; die Hypothese wird verworfen. Schliesslich werden unter verschiedenen Bedingungen Schätzer für die Strukturparameter vorgeschlagen.

Résumé

Les modèles classiques de crédibilité supposent les charges de sinistres provenant de différentes années conditionellement indépendantes et identiquement distribuées, pour une valeur donnée du paramètre inconnue de risque $\tilde{\theta}$. Le présent article introduit une variable aléatoire supplémentaire $\tilde{t}$, qui mesure le temps total que la police a passé dans le portefeuille. On y suppose que l'information sur $\tilde{t}$ peut fournir des renseignements sur la valeur de $\tilde{\theta}$ et que $\tilde{t}$ peut ainsi être utilisé dans le schéma du calcul des primes. L'auteur utilise les modèles de ce type pour déduire puis discuter des estimateurs de crédibilité. Dans un exemple numérique provenant d'observations réelles d'assurance RC-automobile, l'auteur teste l'hypothèse – sous diverses conditions a priori – qu'il n'y a pas d'effet chronique. L'hypothèses est rejetée. Enfin l'auteur propose des estimateurs pour les paramètres de structure sous des conditions diverses.