AN EVOLUTIONARY CREDIBILITY MODEL FOR CLAIM NUMBERS

BY PETER ALBRECHT*

University of Mannheim, FR Germany

KEY WORDS

Credibility, doubly stochastic Poisson sequences, weakly stationary sequences, generalized Pólya sequence.

1. INTRODUCTION

This paper considers a particular credibility model for the claim numbers $N_1, N_2, \ldots, N_n, \ldots$ of a single risk within a collective in successive periods $1, 2, \ldots, n, \ldots$. In the terminology of JEWELL (1975) the model is an evolutionary credibility model, which means that the underlying risk parameter Λ is allowed to vary in successive periods (the structure function is allowed to be time dependent). Evolutionary credibility models for claim *amounts* have been studied by BUHLMANN (1969, pp. 164-165), GERBER and JONES (1975), JEWELL (1975, 1976), TAYLOR (1975), SUNDT (1979, 1981, 1983) and KREMER (1982). Again in Jewell's terminology the considered model is on the other hand stationary, in the sense that the conditional distribution of N_i given the underlying risk parameter does not vary with i.

The computation of the credibility estimate of N_{n+1} involves the considerable labor of inverting an $n \times n$ covariance matrix (n is the number of observations). The above mentioned papers have therefore typically looked for model structures for which this inversion is unnecessary and instead a recursive formula for the credibility forecast can be obtained. Typically nth order stationary *a priori* sequences (e.g., ARMA (p, q)-processes) lead to an nth order recursive scheme. In this paper we impose the restriction that the conditional distribution of N_i is Poisson (which by the way leads to a model identical to the so called "doubly stochastic Poisson sequences" considered in the theory of stochastic point processes). What we gain is a recursive formula for the coefficients of the credibility estimate (not for the estimate itself!) in case of an *arbitrary* weakly stationary *a priori* sequence. In addition to this central result the estimation of the structural parameters is considered in this case and some more special models are analyzed. Among them are EARMA-processes (which are positive-valued stationary sequences possessing exponentially distributed marginals and the same autocorrelation structure as ARMA-processes) as *a priori* sequence and models which can be considered as (discrete) generalizations of the Pólya process.

* I thank the editor and an anonymous referee for valuable suggestions

2. DEFINITION OF THE MODEL AND BASIC PROPERTIES

Let Λ_i denote the risk parameter in period i and let $U_n(\lambda_1, \ldots, \lambda_n)$—the structure function of the considered collective—denote the joint distribution function of $\Lambda_1, \ldots, \Lambda_n$. We make the following assumptions:

ASSUMPTION 1

$$P(N_1 = k_1, \ldots, N_n = k_n | \{\Lambda_i\}) = \prod_{i=1}^{n} P(N_i = k_i | \Lambda_i). \tag{1}$$

This means that the $\{N_i\}$ *are conditionally independent given the* $\{\Lambda_i\}$.

ASSUMPTION 2. *The conditional distribution of* N_i *given* $\Lambda_i = \lambda$ *is a Poisson distribution*

$$P(N_i = k_i | \Lambda_i = \lambda) = \frac{\lambda^{k_i}}{k_i!} e^{-\lambda}. \tag{2}$$

It is Assumption 2 which creates the difference to the other above mentioned evolutionary models. The price we have to pay is the specification of the conditional distribution—which, however, is very natural for claim number models—what we get on the other hand are more specific and useful results.

Combining (1) and (2) we obtain the multivariate distribution of the claim numbers

$$P(N_1 = k_1, \ldots, N_n = k_n) = \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^{n} \left\{ \frac{\lambda_i^{k_i}}{k_i!} e^{-\lambda_i} \right\} dU_n(\lambda_1, \ldots, \lambda_n). \tag{3}$$

This, however, means that the sequence $\{N_i\}_{i \in \mathbb{N}}$ is a "doubly stochastic Poisson sequence". Such sequences have been studied by GRANDELL (1971, 1972, 1976) as a special case of the doubly stochastic Poisson process, which itself can be considered as an evolutionary credibility model for claim numbers in continuous time. We will for practical purposes, however, consider only the discrete time model. A main implication of (3) is that it is possible to establish more properties of the model than just the form of the conditional linear forecast of N_{n+1} as in the usual credibility models. E.g., one can solve other statistical problems and one can give limit theorems for the process. For a lot of detailed results, cf. GRANDELL (1971, 1972, 1976) and SNYDER (1975).

If we denote

$$\begin{cases} E(\Lambda_i) = m_i, & \mathrm{Cov}(\Lambda_i, \Lambda_j) = r_{ij}, \\ \mathrm{Var}(\Lambda_i) = r_{ii} = r_i, \end{cases} \tag{4}$$

we obtain the corresponding moments of $\{N_i\}$ as

$$\begin{cases} E(N_i) = m_i, & \mathrm{Cov}(N_i, N_j) = r_{ij}, \qquad i \neq j \\ \mathrm{Var}(N_i) = r_i + m_i. \end{cases} \tag{5}$$

From (2) we see that the marginal distributions of the process $\{N_t\}$ are mixed Poisson distributions

$$P(N_t = k) = \int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} \, dU_{\Lambda_t}(\lambda). \tag{6}$$

This implies that $P(N_t = k)$ can be calculated for various mixing distributions $U_{\Lambda_t}(\lambda)$. For some recent results see ALBRECHT (1984). The multivariate counting distribution of the process is given by (3), but can alternatively be derived as follows.

Let $L_n^{\Lambda}(s_1, \ldots, s_n)$ denote the Laplace functional of $(\Lambda_1, \ldots, \Lambda_n)$ and let $\Phi_n^N(t_1, \ldots, t_n)$ denote the probability generating functional of $(N_1, \ldots, N_n)$.

As $e^{-\lambda(1-t)}$ is the probability generating function of a Poisson variable with parameter λ, we obtain from (3)

$$\begin{aligned} \Phi_n^N(t_1, \ldots, t_n) &= \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^n E[t_i^{N_i} | \Lambda_i = \lambda_i] \, dU_n(\lambda_1, \ldots, \lambda_n) \\ &= \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^n e^{-\lambda_i(1-t_i)} \, dU_n(\lambda_1, \ldots, \lambda_n) \\ &= L_n^{\Lambda}(1 - t_1, \ldots, 1 - t_n). \end{aligned} \tag{7}$$

The multivariate counting distribution then is given by the relation

$$P(N_1 = k_1, \ldots, N_n = k_n) = \left[\prod_{i=1}^n \frac{1}{k_i!}\right] \frac{\partial^{\Sigma k_i} \Phi_n^N(t_1, \ldots, t_n)}{\partial t_n^{k_n} \ldots \partial t_1^{k_1}} \bigg|_{t_i=0}. \tag{8}$$

We now come to the central problem of credibility, the calculation of the optimal linear forecast of N_{n+1} given the $N_1, \ldots, N_n$. If $f_n(N_1, \ldots, N_n) = a_0 + \sum_{i=1}^n a_i N_i$ denotes the linear forecast function, the parameters which make $E\{N_{n+1} - f_n(N_1, \ldots, N_n)\}^2$ a minimum are determined in the following way (this is easily established by straightforward calculation, or as a special case from the general result of JEWELL (1971, p. 15) or GRANDELL (1976, p. 128)).

a_0 is given by a single equation which makes the forecast unbiased

$$a_0 = E(N_{n+1}) - \sum_{i=1}^n a_i E(N_i) = m_{n+1} - \sum_{i=1}^n a_i m_i. \tag{9}$$

The remaining coefficients are given by the $n \times n$ system of linear equations

$$\sum_{j=1}^n \operatorname{Cov}(N_i, N_j) a_j = \operatorname{Cov}(N_i, N_{n+1}), \qquad i = 1, \ldots, n \tag{10}$$

or more specifically

$$a_i m_i + \sum_{j=1}^n r_{ij} a_j = r_{in+1}, \qquad i = 1, \ldots, n. \tag{11}$$

We note, that because of the identical expectation and covariance structure the optimal linear forecast of N_{n+1} given the $N_1, \ldots, N_n$ equals the optimal linear forecast of $\Lambda_{n+1} = E[N_{n+1}|\Lambda_{n+1}]$ given $N_1, \ldots, N_n$. In turn this means that it is also identical to the optimal linear forecast of $\mathrm{Var}(N_{n+1}|\Lambda_{n+1}) = \Lambda_{n+1}$ given $N_1, \ldots, N_n$.

We now consider in detail a rather general class of doubly stochastic Poisson sequences, which turns out to have nice properties with respect to the calculation of the credibility forecast and the estimation of the structural parameters.

3. WEAKLY STATIONARY A PRIORI SEQUENCES

We require that $\{\Lambda_n\}_{n\in\mathbb{N}}$ is a weakly stationary sequence characterized by the following moment structure:

$$E(\Lambda_i) = m \quad \text{for all } i \in \mathbb{N} \tag{12}$$

$$\mathrm{Cov}(\Lambda_i, \Lambda_j) = r_{|i-j|} \quad \text{for all } i, j \in \mathbb{N} \tag{13}$$

The main result in connection with this special model is that we are able to simplify the calculation of the credibility forecast. Whereas the general case only allows that the inverse of $C(n) = (\mathrm{Cov}(N_i, N_j))_{i=1,\ldots,n}$ can be calculated recursively we are able to give a recursive formula for the optimal coefficients a_i, however, *not* a recursive formula for the credibility forecast.

Let now

$$f_n^*(N_1, \ldots, N_n) = a_0(n) + \sum_{i=1}^{n} a_i(n) N_i \tag{14}$$

denote the optimal linear forecast of N_{n+1} given $N_1, \ldots, N_n$ and

$$C(n) = (\mathrm{Cov}(N_i, N_j))_{i,j=1,\ldots,n} = (c_{ij}) \tag{15}$$

denote the covariance matrix of $(N_1, \ldots, N_n)$.

We have

$$c_{ij} = \begin{cases} r_0 + m & i = j, \\ r_{|i-j|} & i \neq j. \end{cases} \tag{16}$$

Let

$$\boldsymbol{a}(n) = (a_1(n), \ldots, a_n(n))' \tag{17}$$

$$\tilde{\boldsymbol{a}}(n) = (a_n(n), \ldots, a_1(n))' \tag{18}$$

$$\boldsymbol{r}(n) = (r_1, \ldots, r_n)' \tag{19}$$

and

$$\tilde{\boldsymbol{r}}(n) = (r_n, \ldots, r_1)'. \tag{20}$$

From (10) we obtain that the optimal coefficients of the credibility forecast are given by

$$\boldsymbol{a}(n) = \boldsymbol{C}^{-1}(n)\tilde{\boldsymbol{r}}(n). \tag{21}$$

The following lemma gives the form of the inverse of a partitioned matrix.

LEMMA 1. *Let the symmetric (n, n) matrix $\boldsymbol{C}$ be decomposed to*

$$\boldsymbol{C}=\left(\begin{array}{c|c} c_{11} & \boldsymbol{u}' \\ \hline \boldsymbol{u} & \boldsymbol{D} \end{array}\right),$$

where $\boldsymbol{D}$ is of order $(n-1, n-1)$. Then we have

$$\boldsymbol{C}^{-1}=\left(\begin{array}{c|c} \dfrac{1}{s} & -\dfrac{1}{s}\boldsymbol{v}' \\ \hline -\dfrac{1}{s}\boldsymbol{v} & \boldsymbol{D}^{-1}+\dfrac{1}{s}\boldsymbol{v}\boldsymbol{v}' \end{array}\right), \tag{22}$$

where

$$\boldsymbol{v}=\boldsymbol{D}^{-1}\boldsymbol{u}$$

$$s=c_{11}-\boldsymbol{v}'\boldsymbol{u}=c_{11}-\boldsymbol{u}'\boldsymbol{D}^{-1}\boldsymbol{u}.$$

The following lemma gives some useful elementary properties of the covariance matrix $\boldsymbol{C}(n)$.

LEMMA 2.
1. *$\boldsymbol{C}(n+1)$ can for $n \geq 1$ be decomposed in the following way:*

$$\boldsymbol{C}(n+1)=\left(\begin{array}{c|c} r_0+m & \boldsymbol{r}(n)' \\ \hline \boldsymbol{r}(n) & \boldsymbol{C}(n) \end{array}\right). \tag{23}$$

2.

$$\boldsymbol{C}(n)\tilde{\boldsymbol{a}}(n)=\boldsymbol{r}(n). \tag{24}$$

This implies
3.

$$\boldsymbol{C}^{-1}(n)\boldsymbol{r}(n)=\tilde{\boldsymbol{a}}(n). \tag{25}$$

We now define (the $a_i(n)$ are the coefficients of the credibility forecast $f_n^*(N_1, \ldots, N_n)$)

$$s(n)=r_0+m-\boldsymbol{r}(n)'\tilde{\boldsymbol{a}}(n)=r_0+m-\sum_{i=1}^{n} r_i a_{n-i+1}(n), \qquad n \geq 1 \tag{26}$$

$$k(n)=r_{n+1}-\boldsymbol{r}(n)'\boldsymbol{a}(n)=r_{n+1}-\sum_{i=1}^{n} r_i a_i(n), \qquad n \geq 1. \tag{27}$$

REMARK. $s(n)=E\{N_{n+1}-f_n^*(N_1, \ldots, N_n)\}^2$, i.e., the minimum mean square error of a linear forecast of N_{n+1}, given $N_1, \ldots, N_n$.

We now come to the central result.

THEOREM. *For the coefficients* $a_0(n+1)$, $\boldsymbol{a}(n+1)$ *of the credibility forecast* $f_{n+1}(N_1, \ldots, N_{n+1})$ *the following relations are valid* $(n \geq 1)$:

$$a_0(n+1) = \left(1 - \frac{k(n)}{s(n)}\right) a_0(n) \tag{28}$$

$$a_1(n+1) = \frac{k(n)}{s(n)} \tag{29}$$

$$a_i(n+1) = a_{i-1}(n) - \frac{k(n)}{s(n)} a_{n-i+2}(n), \qquad 2 \leq i \leq n+1. \tag{30}$$

The starting values are $a_0(1) = m(1 - r_1/(r_0 + m))$ *and* $a_1(1) = r_1/(r_0 + m)$.

REMARK. (30) can alternatively be written as

$$(a_2(n+1), \ldots, a_{n+1}(n+1))' = \boldsymbol{a}(n) - \frac{k(n)}{s(n)} \tilde{\boldsymbol{a}}(n). \tag{31}$$

PROOF.

$$\boldsymbol{a}(n+1) = \boldsymbol{C}^{-1}(n+1)\tilde{\boldsymbol{r}}(n+1). \tag{32}$$

From the decomposition (23) of $C(n+1)$, we obtain in the notation of lemma 1, using (25):

$$\boldsymbol{v} = \boldsymbol{C}^{-1}(n)\boldsymbol{r}(n) = \tilde{\boldsymbol{a}}(n)$$

$$s = (r_0 + m) - \tilde{\boldsymbol{a}}(n)'\boldsymbol{r}(n) = s(n).$$

The following partitioned form of $\boldsymbol{C}^{-1}(n+1)$ results:

$$\boldsymbol{C}^{-1}(n+1) = \left(\begin{array}{c|c} \dfrac{1}{s(n)} & -\dfrac{1}{s(n)} \tilde{\boldsymbol{a}}(n)' \\ \hline -\dfrac{1}{s(n)} \tilde{\boldsymbol{a}}(n) & \boldsymbol{C}^{-1}(n) + \dfrac{1}{s(n)} \tilde{\boldsymbol{a}}(n)\tilde{\boldsymbol{a}}(n)' \end{array}\right).$$

From (32), the relations (29) and (30) easily follow. Then (28) is obtained from (9).

COROLLARY. *For the mean square error* $s(n)$ *of the credibility forecast the following recursive formula is valid*:

$$s(n+1) = s(n) - \frac{k(n)^2}{s(n)}, \quad n \geq 1; \qquad s(1) = r_0 + m - \frac{r_1^2}{r_0 + m}. \tag{33}$$

PROOF. From (26)

$$s(n+1) = r_0 + m - \sum_{i=1}^{n} r_i a_{n+2-i}(n+1) - r_{n+1} a_1(n+1);$$

using (29), (30) this simplifies to

$$r_0+m-\sum_{i=1}^{n} r_i\left\{a_{n+1-i}(n)-\frac{k(n)}{s(n)}a_i(n)\right\}-r_{n+1}\frac{k(n)}{s(n)}.$$

Using (26), (27) this in turn simplifies to (33).

The theorem allows recursive calculation of the credibility forecast of N_{n+1} in case of a known risk structure. To obtain an *empirical* credibility forecast, we have to estimate the unknown parameters, which here are: m, r_0, r_1, $r_2, \ldots$.

The estimation problem exhibits the second important property of the model considered in this section. If we assume that the *a priori* sequence $\{\Lambda_i\}$ is weakly stationary, then we obtain from (5), that the observable sequence $\{N_i\}$ is a weakly stationary one, too. We then have the possibility to apply results from the well-developed theory of the statistical analysis of weakly stationary time series, see e.g., HANNAN (1960, Chapters II-IV) or DOOB (1953, Chapter X). For example a *spectral analysis* of the sequence $\{N_i\}$ is possible. Some results in this direction can be found in GRANDELL (1976, Chapter 7.2). We will here, however, confine to the above mentioned estimation problem. Up to now we have only considered the claim number sequence of a single risk, observed for n years. We now assume that we observe a collective of K independent risks, each having the same probability law of its claim number sequence.

Let

(34) $$N_{ji} = \text{number of claims of risk } i \text{ in year } j$$

$$i=1,\ldots,K; j=1,\ldots,n.$$

From standard results of time series analysis, e.g., HANNAN (1960, pp. 30-33), we obtain the following natural estimators of the above mentioned parameters.

(35) $$\hat{m}=\frac{1}{Kn}\sum_{i,j=1}^{n} N_{ji}$$

(36) $$\hat{r}_k=\frac{1}{K(n-k)-1}\sum_{i=1}^{K}\sum_{j=1}^{n-k}(N_{ji}-\hat{m})(N_{j+k,i}-\hat{m}), \quad \text{for } k\geq 1$$

(37) $$\widehat{\operatorname{Var}(N_{ji})}=\frac{1}{Kn-1}\sum_{i=1}^{K}\sum_{j=1}^{n}(N_{ji}-\hat{m})^2.$$

A natural estimate for r_0 then is

(38) $$\hat{r}_0=\frac{1}{Kn-1}\sum_{i=1}^{K}\sum_{j=1}^{n}(N_{ji}-\hat{m})^2-\hat{m}.$$

As pointed out by the referee the expected value of (37) is given by

$$\operatorname{Var}(N_{ji})-\frac{1}{Kn-1}\frac{1}{n}\sum_{j=1}^{n-1} r_j(n-j),$$

which implies a slight bias.

The theorem shows, how the coefficients of the credibility forecasts can be calculated recursively in the case of an arbitrary stationary *a priori* sequence. It is, however, not possible to develop a recursive formula for the credibility forecast itself for the general case. It would be interesting to examine special classes of stationary *a priori* sequences which give rise to recursive formulae for the credibility forecast itself. For a more general type of evolutionary models KREMER (1982) has considered ARMA (p, q) processes as a special class of stationary *a priori* sequences. In the model of this paper the *a priori* sequences have to be positive-valued to be admissible. Therefore the ARMA (p, q) processes are not admissible in general. However Lewis and a number of co-authors (see LAWRENCE and LEWIS (1980) for the most recent results) have developed models for positive-valued stationary time series $\{X_t\}_{t \in \mathbb{N}}$ which, being in general rather distinct from the ARMA-models, possess the same autocorrelation structure as the ARMA-processes. These processes are called EARMA (p, q)-processes, the E stemming from the additional feature of all these processes: they have an exponential marginal distribution!

The results of KREMER (1982) cannot be translated into the present context for several reasons, one being that the form of the linear regressions of the EARMA-processes have not yet been established. Another drawback of the EARMA-processes is that the statistical analysis of these processes is not yet well developed in general, contrary to the ARMA-processes. In the following, we consider some examples.

EXAMPLE 1. EAR (1)-process as *a priori* sequence. A stationary version of the first order autoregressive model with exponential marginals with "finite past" can be obtained as follows (cf. GAVER and LEWIS (1980, p. 732):

$$(39) \qquad \begin{cases} \Lambda_n = \rho \Lambda_{n-1} + I_n E_n, & n \geq 2 \\ \Lambda_1 = \rho E_0 + I_1 E_1, & \end{cases} \qquad (0 \leq \rho < 1)$$

where $\{I_n\}_{n \geq 1}$ is a sequence of i.i.d. Bernoulli-variables with $P(I_n = 0) = \rho$ and $\{E_n\}_{n \geq 0}$ is an independent sequence of i.i.d. exponentially distributed variables with parameter λ. The resulting sequence is a first order Markov process, the Λ_t are exponentially distributed with parameter λ and can alternatively be obtained in the usual first-order autoregressive form $\Lambda_n = \rho \Lambda_{n-1} + \varepsilon_n$ with a suitable $\{\varepsilon_n\}$. For the second order structure we obtain

$$(40) \qquad \begin{cases} m = E(\Lambda_n) = 1/\lambda, & r_0 = \operatorname{Var}(\Lambda_n) = 1/\lambda^2 \\ r_k = \operatorname{Cov}(\Lambda_n, \Lambda_{n+k}) = \rho^k/\lambda^2 = \rho^k r_0, & k \geq 1 \end{cases}$$

From (40) we see that the r_k fulfill the property (10) of SUNDT (1981, p. 7), which in our context reads:

$$(41) \qquad r_{i+1-j} = \rho_i \cdot r_{i-j} \quad \text{for all } i \geq j, \text{ for all } j \geq 1.$$

Clearly $\rho_i = \rho$ for all i and from Sundt's result (11) we obtain the following recursive formula for the credibility forecast.

As (notation as in SUNDT (1981)) $\varphi_i = E\{\text{Var}(N_i|\Lambda_i)\} = E(\Lambda_i) = 1/\lambda$, we define

$$\gamma_n = \frac{1}{\lambda s(n-1)}, \quad n \geq 2; \qquad \gamma_1 = \frac{1}{1+1/\lambda} \tag{42}$$

$$\chi = 2\rho^2 + (1-\rho^2)\left(1+\frac{1}{\lambda}\right) \tag{43}$$

and obtain

$$\gamma_{n+1} = (\chi - \gamma_n \rho^2)^{-1} \tag{44}$$

$$\begin{cases} f_n^*(N_1, \ldots, N_n) = \rho[(1-\gamma_n)N_n + \gamma_n f_{n-1}^*(N_1, \ldots, N_{n-1})] + (1-\rho)\lambda \\ f_0 = 1/\lambda. \end{cases} \tag{45}$$

This is the desired recursive formula for the credibility forecast.

It is interesting to note that the regressions of this *a priori* sequence are all linear, precisely

$$E(\Lambda_{n+1}|\Lambda_1, \ldots, \Lambda_n) = E(\Lambda_{n+1}|\Lambda_n) = \rho\Lambda_n + (1-\rho)/\lambda. \tag{46}$$

However, we have not been able to show that the regressions of the *a posteriori* process $\{N_i\}$ are linear too, i.e., the credibility forecast in the best forecast of N_{n+1} based on $N_1, \ldots, N_n$.

We now come to the estimation of the unknown parameters λ and ρ and consider again a *collective* of K independent risks each having the same law of its claim number sequence. Let N_{ji} be defined as in (34); noticing that $E(N_i) = 1/\lambda$ and $r_1 = \text{Cov}(N_i, N_{i+1}) = \rho/\lambda^2$ we obtain from (35) and (36) the following (consistent) natural estimators of λ and ρ:

$$\hat{\lambda} = 1\Big/\frac{1}{Kn}\sum_{i,j=1}^{n} N_{ji} \tag{47}$$

$$\hat{\rho} = \frac{\hat{\lambda}^2}{K(n-1)-1}\sum_{i=1}^{K}\sum_{j=1}^{n-1}(N_{ji} - \hat{\lambda}^{-1})(N_{j+1,i} - \hat{\lambda}^{-1}). \tag{48}$$

A drawback of the model is, that all correlations $\rho_k = \text{Corr}(\Lambda_i, \Lambda_{i+k})$ are positive. Indeed, one can show that there does not exist an autoregressive sequence $\Lambda_n = \rho\Lambda_{n-1} + \varepsilon_n$ possesssing exponentially distributed marginals *and* $\rho < 0$! However, GAVER and LEWIS (1980, p 741) present models of similar autocorrelation structure and negative correlation, which still possess the property of having an exponential marginal distribution. GAVER and LEWIS (1980, pp. 736-737) consider also an autoregressive process of first order with a gamma marginal distribution and a similar autocorrelation structure, the GAR (1)-process.

EXAMPLE 2. EMA (1)-process as a priori sequence. A first-order moving average model with exponential marginal distribution, was considered by LAWRANCE and LEWIS (1977) and can be obtained as follows (forward formulation)

$$\Lambda_n = \beta\varepsilon_n + I_n\varepsilon_{n+1}, \qquad (0 \leq \beta \leq 1), \tag{49}$$

where the $\{I_n\}_{n \geq 1}$ are i.i.d. Bernoulli variables with $P(I_n = 1) = 1 - \beta$ and $\{\varepsilon_n\}_{n \geq 1}$ is an independent sequence of i.i.d. exponentially distributed (parameter λ) random variables. The process is not Markovian and the second order structure is given by

$$(50) \qquad \begin{cases} m = E(\Lambda_n) = 1/\lambda, \qquad r_0 = \mathrm{Var}\,(\Lambda_n) = 1/\lambda^2 \\ r_1 = \mathrm{Cov}\,(\Lambda_n, \Lambda_{n+1}) = \beta(1-\beta) r_0 \\ r_k = 0 \quad \text{for } k \geq 2 \end{cases}$$

To obtain a recursive formula for the credibility forecast we can use Theorem 2 in SUNDT (1981, p. 6).

We obtain the following recursive relation for the estimation error $s(n)$:

$$(51) \qquad \begin{cases} s(n) = \dfrac{1}{\lambda}\left(1 + \dfrac{1}{\lambda}\right) - \dfrac{\beta^2(1-\beta)^2}{\lambda^4 s(n-1)}, \qquad n \geq 2 \\ s(1) = \dfrac{1}{\lambda}\left(1 + \dfrac{1}{\lambda}\right) - \dfrac{\beta^2(1-\beta)^2}{\lambda^2 + \lambda^3} \end{cases}$$

and the following recursive formula for the credibility forecast

$$(52) \qquad \begin{cases} f_n^*(N_1, \ldots, N_n) = \dfrac{\beta(1-\beta)}{\lambda^2 s(n-1)} N_n - \dfrac{\beta(1-\beta)}{\lambda^2 s(n-1)} f_{n-1}^*(N_1, \ldots, N_{n-1}) + \dfrac{1}{\lambda}, \\ f_1^*(N_1) = \dfrac{1}{\lambda} - \dfrac{\beta(1-\beta)}{(1+\lambda)\lambda} + \dfrac{\beta(1-\beta)}{1+\lambda} N_1. \end{cases} \qquad n \geq 2$$

A natural estimator of the unknown parameter β (λ is estimated as under (47)) is given by

$$(53) \qquad \hat{\beta} = \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\hat{\lambda}^2 \frac{1}{K(n-1)-1} \sum_{i=1}^{K} \sum_{j=1}^{n-1} (N_{ji} - \hat{\lambda}^{-1})(N_{j+1,i} - \hat{\lambda}^{-1})}.$$

A drawback of the model is that the first-order autocorrelation $\rho_1 = \beta(1-\beta)$ is always nonnegative (one can show in addition, that it is always bounded from above by 1/4).

The regressions of the *a priori* process are given by

$$(54) \qquad E(\Lambda_{n+1} | \Lambda_n) = \frac{1}{\lambda}\left[\beta\lambda\Lambda_n + \frac{1-2\beta}{1-\beta} + \frac{\beta}{1-\beta} e^{-\lambda(1-\beta)\Lambda_n/\beta}\right]$$

and are therefore not linear.

EXAMPLE 3. EARMA (1, 1) process as a priori sequence. A first order mixed autoregressive-moving average process with an exponential marginal distribution was considered by JACOBS and LEWIS (1977) and can be obtained as follows ("backward formulation").

$$
(55)\qquad \begin{cases} \Lambda_n = \beta\varepsilon_n + U_n A_{n-1} & (0 \leq \beta \leq 1) \\ A_n = \rho A_{n-1} + V_n \varepsilon_n & (0 \leq \rho \leq 1) \\ & A_0 = \varepsilon_0 \end{cases} \quad (n \geq 1),
$$

where $\{U_n\}$ and $\{V_n\}$ are independent sequences of independent Bernoulli variables with $P(U_n = 0) = \beta$, $P(V_n = 0) = \rho$ and $\{\varepsilon_n\}$ is an independent sequence of i.i.d. exponentially distributed (parameter λ) random variables. The resulting process $\{\Lambda_n\}$ is stationary and in general non Markovian. The second order structure of the process is given by

$$
(56)\qquad \begin{cases} m = E(\Lambda_n) = 1/\lambda, \qquad r_0 \mathrm{Var}\,(\Lambda_n) = 1/\lambda^2 \\ r_1 = \mathrm{Cov}\,(\Lambda_n, \Lambda_{n+1}) = r_0(1-\beta)[\beta + \rho(1-2\beta)] \\ r_k = \rho^{k-1} r_1. \end{cases}
$$

Again we can apply Theorem 2 of SUNDT (1981) to obtain a recursive formula for the credibility forecast. The result is as follows:

$$
(57)\qquad \begin{cases} s(n) = (r_0 + m) + \rho^2(r_0 + m) - 2\rho r_1 - \dfrac{[\rho(r_0 + m) - r_1]^2}{s(n-1)}, & n \geq 2 \\ s(1) = (r_0 + m) - \dfrac{r_1^2}{r_0 + m} \end{cases}
$$

$$
(58)\qquad \begin{cases} f_n^*(N_1, \ldots, N_n) = \left(\rho - \dfrac{\rho(r_0+m) - r_1}{s(n-1)}\right) N_n \\ \qquad + \dfrac{\rho(r_0+m) - r_1}{s(n-1)} f_{n-1}^*(N_1, \ldots, N_{n-1}) + m(1-\rho), & n \geq 2 \\ f_1^*(N_1) = m\left(1 - \dfrac{r_1}{r_0+m}\right) + \dfrac{r_1}{r_0+m} N_1. \end{cases}
$$

4. SOME SPECIAL MODELS

We first treat two models which can be considered as generalizations of the Pólya-process in discrete time. The Pólya-process is a mixed Poisson process with the gamma distribution as mixing distribution.

Model A. A natural generalization, which was already considered by BATES and NEYMAN (1952), is to assume

$$
(59)\qquad \Lambda_j = a_j \cdot \Lambda,
$$

where Λ follows a gamma distribution with parameters b and p. The *a priori* moments are given by

$$
(60)\qquad E(\Lambda_i) = a_i \frac{p}{b}, \qquad \mathrm{Var}\,(\Lambda_i) = a_i^2 \frac{p}{b^2}, \qquad \mathrm{Cov}\,(\Lambda_i, \Lambda_j) = a_i a_j \frac{p}{b^2}.
$$

SNYDER (1975, p. 288) considers the continuous time analogue, which he terms "inhomogeneous Pólya process".

Model A seems to be the only known double stochastic Poisson sequence for which the multivariate counting distribution can be given explicitly. Bates and Neyman showed that

$$(61)\quad P(N_1=n_1,\ldots,N_k=n_k)=\left(1+\frac{a}{b}\right)^{-p} n!\binom{n+p-1}{n}\prod_{i=1}^{k}\frac{1}{n_i!}\left[\frac{a_i/b}{1+a/b}\right]^{n_i},$$

where $a=\sum_{i=1}^{k} a_i$ and $n=\sum_{i=1}^{k} n_i$.

Comparing (61) with JOHNSON and KOTZ (1969, p. 292, (32)) shows, that the multivariate counting distribution of the "discrete inhomogeneous Pólya process" is just a *multivariate negative binomial distribution* ($N=p$, $P_i=a_i/b$ in their notation).

JOHNSON and KOTZ (1969, p. 295) show also, that in case of a multivariate negative binomial distribution the regressions are always linear. Especially we obtain

$$(62)\quad E(N_{n+1}N_1,\ldots,N_n)=p\frac{a_{n+1}}{b+a}+\frac{a_{n+1}}{b+a}\sum N_i$$

$$=\frac{b}{b+a}E(N_{n+1})+\frac{a_{n+1}}{b+a}\sum N_i.$$

This implies that in case of the "discrete inhomogeneous Pólya process" the optimum forecast function (with respect to the mean square error) is identical to the best *linear* forecast function (the credibility forecast).

If we want to calculate the credibility forecast with the method of chapter 1 (equations (9) and (10)), we can apply a result of JEWELL (1976, pp. 16–17), because Cov (N_i, N_j) can be factored into $a_i\cdot((p/b^2)a_j)$.

It is interesting to note that already BUHLMANN (1969, pp. 164–165) considered a similar model. He considered a sequence of conditionally Poisson distributed claim variables $\{X_n\}$ with the property

$$(63)\quad E(X_n|\theta)=a_n\cdot\theta,$$

where $a_n=n+c$, c is a constant independent of n and θ follows a gamma distribution.

In addition to Buhlmann's results we show in the following how the structural parameters (especially c) can be estimated.

Assume that we have given a sample of size m of observations of $(N_1,\ldots,N_k)$. Let

$$n_{ij}=i\text{th observation of } N_j,\qquad i=1,\ldots,m;\ j=1,\ldots,k.$$

Let

$$n_i=\sum_{j=1}^{k} n_{ij},\qquad \tilde n_j=\sum_{i=1}^{m} n_{ij},\qquad n=\sum_{i=1}^{m} n_i,\qquad r=\tfrac{1}{2}k(k+1).$$

Then the log-likelihood-function of the observations is given by

$$(64)\qquad \log L = -(n+pm)\log\left(1+\frac{r+kc}{b}\right) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\log(p-1+j) + \sum_{j=1}^{k}\tilde{n}_j \log\left(\frac{j+c}{p}\right).$$

The likelihood equations then are given by

$$(65)\qquad \frac{\partial \log L}{\partial p} = -m\log\left(1+\frac{r+kc}{b}\right) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{1}{p-1+j} = 0$$

$$(66)\qquad \frac{\partial \log L}{\partial c} = \frac{-k(n+pm)}{b+r+kc} + \sum_{j=1}^{k}\frac{\tilde{n}_j}{j+c} = 0$$

$$(67)\qquad \frac{\partial \log L}{\partial b} = \frac{(n+pm)(r+kc)}{b^2+b(r+kc)} - \frac{n}{b} = 0.$$

If $\hat{p}$, $\hat{c}$, $\hat{b}$ denote the maximum likelihood estimators of p, c, b, then from (67) we obtain

$$(68)\qquad \hat{b} = \frac{m}{n}(r+k\hat{c})\hat{p}$$

Substituting (68) in (66), we obtain that $\hat{c}$ is the solution of

$$(69)\qquad \sum_{j=1}^{k}\frac{\tilde{n}_j}{j+\hat{c}} = \frac{nk}{(r+k\hat{c})}$$

Substituting (69) in (65), we obtain that $\hat{p}$ is the solution of

$$(70)\qquad \sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{1}{\hat{p}-1+j} = m\log\left(1+\frac{n}{m\hat{p}}\right).$$

Model B Another way to obtain a generalization of the Pólya process is to replace the gamma mixing distribution by a multivariate analogue, a multivariate gamma distribution for $(\Lambda_1, \ldots, \Lambda_n)$.

A natural way to obtain a multivariate gamma distribution, more precisely a multivariate χ^2-distribution is the following, cf. also JOHNSON and KOTZ (1972, chapter 40.3) or KRISHNAIAH and RAO (1961). The χ^2-distribution with n degrees of freedom is a special gamma distribution and is the distribution of $\sum_{i=1}^{n} X_i^2$, where the X_i are independent and identically $N(0, 1)$-distributed (normal distribution with mean 0 and variance 1). A natural multivariate analogue is obtained by starting with m independent and identically multivariate normal distributed random vectors $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{in})$, $i = 1, \ldots, m$. Precisely $\boldsymbol{Y}_i$ follows a $N(\boldsymbol{0}, \boldsymbol{\Sigma})$ distribution, where $\boldsymbol{\Sigma} = (\Sigma_{ij})$ is the variance covariance matrix of $(Y_{i1}, \ldots, Y_{in})$ and we assume that $\Sigma_{ii} = 1$.

The *a priori* vector

$$(71)\qquad (\Lambda_1, \dots, \Lambda_n) = \left(\sum_{i=1}^{m} Y_{i1}^2, \dots, \sum_{i=1}^{m} Y_{in}^2 \right)$$

then follows a distribution, which can be considered as a multivariate χ^2-distribution. Especially each Λ_i is χ^2-distributed with n degrees of freedom. The Laplace functional $L_n^{\Lambda}(s_1, \dots, s_n) = E[e^{-\Sigma s_i \Lambda_i}]$ is given by

$$(72)\qquad L_n^{\Lambda}(s_1, \dots, s_n) = |\boldsymbol{I} + 2 s_\Delta \boldsymbol{\Sigma}|^{-m/2},$$

where s_Δ is a diagonal matrix with diagonal elements $s_1, \dots, s_n$. From (7) it follows that the probability generating functional of $N_1, \dots, N_n$ is given by

$$(73)\qquad \Phi_n^N(t_1, \dots, t_n) = |\boldsymbol{I} + 2(\boldsymbol{I} - \boldsymbol{t}_\Delta)\boldsymbol{\Sigma}|$$

where $\boldsymbol{I} - \boldsymbol{t}_\Delta$ is a diagonal matrix with diagonal elements $(1 - t_1, \dots, 1 - t_n)$.

A simple special case is obtained when we assume a first-order correlation for the Y_{ij}, i.e., $\boldsymbol{\Sigma}$ is of the form

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & r & 0 & \cdots & \cdots & 0 \\ r & 1 & r & \ddots & & \vdots \\ 0 & r & 1 & r & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & r & 1 & r \\ 0 & \cdots & \cdots & 0 & r & 1 \end{pmatrix}$$

We then obtain the following second order recursive relations for $\varphi_n(t_1, \dots, t_n) = |\boldsymbol{I} + 2(\boldsymbol{I} - \boldsymbol{t}_\Delta)\boldsymbol{\Sigma}|$:

$$(74)\qquad \begin{cases} \varphi_{n+2}(t_1, \dots, t_{n+2}) = (3 - 2t_{n+2})\varphi_{n+1}(t_1, \dots, t_{n+1}) \\ \qquad\qquad -4(t_{n+2} - 1)(t_{n+1} - 1) r^2 \varphi_n(t_1, \dots, t_n) \quad \text{for } n \geqslant 0 \\ \varphi_1(t_1) = (3 - 2t_1), \qquad \varphi_0(t_0) = 1. \end{cases}$$

The probability generating functional in this special case then is given by $\Phi_n^N(t_1, \dots, t_n) = \varphi(t_1, \dots, t_n)^{-m/2}$.

We obtain that

$$(75)\qquad \frac{\partial^k \Phi_1(t_1)}{\partial t_1^k} = 2^k \frac{\Gamma((m/2)+k)}{\Gamma(m/2)} \varphi_1(t_1)^{-((m/2)+k)}.$$

From (8) we obtain

$$(76)\qquad P(N_1 = k) = \frac{1}{k!} \frac{\partial^k \Phi_1(t_1)}{\partial t_1^k}\bigg|_{t_1 = 0} = \frac{2^k}{k! 3^{(m/2)+1}} \frac{\Gamma((m/2)+k)}{\Gamma(m/2)}.$$

This result is identical (for $t = 1$) with a result of ALBRECHT (1984), who calulated $P(N(t) = n)$ for a mixed Poisson process $N(t)$ with a χ^2-mixing distribution.

In addition we obtain after some calculation that

$$(77)\quad \frac{\partial^{k_1+k_2}\Phi_2(t_1, t_2)}{\partial t_1^{k_1}\partial t_2^{k_2}} = (-1)^{k_1}\frac{\Gamma((m/2)+k_1)}{\Gamma(m/2)}\sum_{k=0}^{k_2}\binom{k_2}{k}\{\varphi_2(t_1, t_2)^{-((m/2)+k_1)}\}^{(k)}$$
$$\times\{[(4-4r^2)t_2+4r^2-6]^{k_1}\}^{(k_2-k)}$$

where (k) denotes the kth derivative with respect to t_2.

We obtain after some calculation

$$(78)\quad P(N_1 = k_1, N_2 = k_2) = \frac{1}{k_1!}\frac{1}{k_2!}\frac{\partial^{k_1+k_2}\Phi_2(t_1, t_2)}{\partial t_1^{k_1}\partial t_2^{k_2}}\bigg|_{t_1=t_2=0} = \frac{(-1)^k}{k_1!k_2!\Gamma(m/2)}$$
$$\times\sum_{k=0}^{k_2}\binom{k_2}{k}(-1)^k\frac{\Gamma((m/2)+k_1+k)\Gamma(k_1+1)}{\Gamma(k_1-k_2+k+1)}$$
$$\times(9-4r^2)^{-((m/2)+k_1+k)}(4r^2-6)^{2k+k_1-k_2}(4-4r^2)^{k_2-k}.$$

Even in the simple first-order case we have not been able to develop an expression for $E(N_{n+1}|N_1, \ldots, N_n)$, the "best" estimate of N_{n+1} given $N_1, \ldots, N_n$.

As the second-order structure of the sequence N_n is given by

$$(79)\quad \begin{cases} E(N_i) = m, \qquad \mathrm{Var}\,(N_i) = 3m, \\ r_1 = \mathrm{Cov}\,(N_i, N_{i+1}) = 2mr^2 \\ r_k = \mathrm{Cov}\,(N_i, N_{i+k}) = 0, \qquad k \geq 2 \end{cases}$$

we can apply Theorem 2 of SUNDT (1981) to obtain a recursive formula for the credibility forecast. The result is as follows:

$$(80)\quad \begin{cases} s(n) = 3m - \dfrac{4m^2r^4}{s(n-1)}, \qquad n \geq 2 \\[2ex] s(1) = 3m - \dfrac{4}{3}mr^4 \end{cases}$$

$$(81)\quad \begin{cases} f_n^*(N_1, \ldots, N_n) = \dfrac{2mr^2}{s(n-1)}(N_n - f_{n-1}^*(N_1, \ldots, N_{n-1})) + m, \qquad n \geq 2 \\[2ex] f_1^*(N_1) = m(1-\frac{2}{3}r^2) + \frac{2}{3}r^2N_1 \end{cases}$$

Model C (*a priori* sequence with independent increments). If we assume that the *a priori* sequence $\{\Lambda_0 \equiv 0, \Lambda_1, \Lambda_2, \ldots\}$ possesses independent increments, this means—cf. DOOB (1953, p. 96)—that for all $n \geq 3$ and $i_1 < i_2 < \cdots < i_n$ the random variables $\Lambda_{i_2} - \Lambda_{i_1}, \ldots, \Lambda_{i_n} - \Lambda_{i_{n-1}}$ are mutually independent. An additional assumption is that $E(\Lambda_i) = m$; let $V_i = \mathrm{Var}\,(\Lambda_i)$, then we obtain for $i < j$

$$\begin{aligned} \mathrm{Cov}\,(\Lambda_i, \Lambda_j) &= \mathrm{Cov}\,(\Lambda_i - \Lambda_0, \Lambda_j - \Lambda_i + \Lambda_i) \\ &= \mathrm{Var}\,(\Lambda_i) + \mathrm{Cov}\,(\Lambda_i - \Lambda_0, \Lambda_j - \Lambda_i) \\ &= \mathrm{Var}\,(\Lambda_i), \end{aligned}$$

i.e., in general

$$\mathrm{Cov}\,(\Lambda_i, \Lambda_j) = \mathrm{Var}\,(\Lambda_{\min(i,j)}). \tag{82}$$

A credibility model with the above moment structure for the *a priori* variables was already considered by GERBER and JONES (1975, pp. 98–99), they show that the credibility forecast $f_n^*(N_1, \ldots, N_n)$ of N_{n+1} is of the "updating type"

$$f_n^*(N_1, \ldots, N_n) = (1 - Z_n) f_{n-1}^*(N_1, \ldots, N_{n-1}) + Z_n N_n. \tag{83}$$

The weights can be calculated recursively, we have

$$Z_1 = \frac{V_1}{m + V_1}$$

$$Z_n = \frac{V_n - V_{n-1} + mZ_{n-1}}{V_n - V_{n-1} + mZ_{n-1} + m}.$$

Additional models for the *a priori* sequence are considered in GRANDELL (1972) (e.g., $\{\Lambda_j\}$ is in the form of a linear regression model, pp. 106–108) and GRANDELL (1976) (e.g. $\{\Lambda_j\}$ is a stationary alternating Markov chain, pp 153–157).

REFERENCES

ALBRECHT, P (1984) Laplace Transforms, Mellin Transforms and Mixed Poisson Processes *Scandinavian Actuarial Journal*, 58–64

BATES, G E and NEYMAN, J (1952) Contributions to the Theory of Accident Proneness *University of California Publications in Statistics* **1**, 215–276

BUHLMANN, H (1969) Experience Rating and Credibility *ASTIN Bulletin* **5**, 157–165

DOOB, J L (1953) *Stochastic Processes* Wiley New York

GAVER, D G and LEWIS, P A W (1980) First-Order Autoregressive Gamma Sequences and Point Processes *Advances in Applied Probability* **12**, 727–745

GERBER, H U and JONES, D A (1975) Credibility Formulas of the Updating Type In KAHN (1975), pp 89–105

GRANDELL, J (1971) On Stochastic Processes Generated by a Stochastic Intensity Function *Skandinavisk Aktuarietidskrift* **54**, 204–240

GRANDELL, J (1972) Statistical Inference for Doubly Stochastic Poisson Processes In *Stochastic Point Processes Statistical Analysis, Theory and Applications*, edited by P A W LEWIS Wiley New York, pp 90–121

GRANDELL, J (1976) *Doubly Stochastic Poisson Processes Lecture Notes in Mathematics, No 529* Springer Berlin

HANNAN, E J (1960) *Time Series Analysis* Methuen London.

JACOBS, P A and LEWIS, P. A W (1977) A Mixed Autoregressive-Moving Average Exponential Sequence and Point Process, EARMA(1, 1) *Advances in Applied Probability* **9**, 87–104

JEWELL, W S (1975) Model Variations in Credibility Theory In KAHN (1975), pp 193–244

JEWELL, W S (1976) Two Classes of Covariance Matrices Giving Simple Linear Forecasts *Scandinavian Actuarial Journal*, 15–29

JOHNSON, N L and KOTZ, S (1969) *Distributions in Statistics Discrete Distributions* Wiley New York

JOHNSON, N L and KOTZ, S (1969) *Distributions in Statistics Continuous Multivariate Distributions* Wiley New York

KAHN, P M (ed) (1975) *Credibility* New York Academic Press

KREMER, E (1982) Credibility Theory for some Evolutionary Models *Scandinavian Actuarial Journal*, 129–142

KRISHNAIAH, P R and RAO, M M (1961) Remarks on a Multivariate Gamma Distribution *American Mathematical Monthly* **68**, 342–346

LAWRANCE, A J and LEWIS, P A W (1977) An Exponential Moving Average Sequence and Point Process, EMA (1) *Journal of Applied Probability* **14**, 98–113

LAWRANCE, A J and LEWIS, P A W (1980) The Exponential Autoregressive-Moving Average EARMA (p, q) Process *Journal of the Royal Statistical Society* B **42**, 150–161

SNYDER, D L (1975) Random point processes Wiley, New York

SUNDT, B (1979) An Insurance Model with Collective Seasonal Random Factors *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker*, 57–64

SUNDT, B (1981) Recursive Credibility Estimation *Scandinavian Actuarial Journal*, 3–21

SUNDT, B (1983) Finite Credibility Formulae in Evolutionary Models *Scandinavian Actuarial Journal*, 106–116

TAYLOR, G C (1975) Credibility for Time-Heterogeneous Loss Ratio In KAHN (1975), pp 363-389