

DEPENDENCE IN DYNAMIC CLAIM FREQUENCY CREDIBILITY MODELS

BY

OANA PURCARU AND MICHEL DENUIT

ABSTRACT

In nonlife insurance, actuaries usually resort to random effects to take unexplained heterogeneity into account (in the spirit of the Bühlmann-Straub model). This paper aims to study the kind of dependence induced by the introduction of correlated latent variables in the annual numbers of claims reported by policyholders. The effect of reporting claims on the *a posteriori* distribution of the random effects will be made precise. This will be done by establishing some stochastic monotonicity property of the *a posteriori* distribution with respect to the claims history.

KEYWORDS

Poisson mixture, credibility theory, experience rating, risk classification, copulas, time series for non-gaussian responses.

1. INTRODUCTION AND MOTIVATION

One of the main tasks of the actuary is to design a tariff structure that will fairly distribute the burden of claims among policyholders. If the risks in the portfolio are not all equal to each other (which means that the associated random financial losses have different distribution functions), it is fair to partition all policies into homogeneous classes with all policyholders belonging to the same class paying the same premium. In automobile third party liability insurance, examples of classification variables encountered in practice include the age, gender and occupation of the policyholders, the type and use of their car, the place where they reside and sometimes even the number of cars in the household, marital status, smoking behavior or the color of the vehicle. It is convenient to achieve *a priori* classification with the help of generalized linear models; see e.g. RENSHAW (1994).

However, many important factors cannot be taken into account at this stage; think for instance of swiftness of reflexes or aggressiveness behind the wheel in

automobile insurance. Consequently, tariff cells are still quite heterogeneous. This residual heterogeneity can be represented by a random effect in a statistical model. The amount of premium charged to all policyholders in a risk class is thus itself an average, so that some policyholders pay too much and subsidize the others. The claims histories can be used to restore fairness in the risk classes, increasing the premium for policyholders reporting claims and decreasing those of good drivers. The allowance for the history of the policyholder in a rating model thus derives from interpretation of serial correlation for longitudinal data resulting from hidden features in the risk distribution.

In a seminal paper, DIONNE and VANASSE (1989) proposed a credibility model which integrates *a priori* and *a posteriori* information on an individual basis. These authors introduced a regression component in the Poisson counting model in order to use all available information in the estimation of accident frequency. The unexplained heterogeneity was then modeled by the introduction of a latent variable representing the influence of hidden policy characteristics. Taking this random effect Gamma distributed yields the Negative Binomial model for the claim number. Of course, there is no particular reason to restrict ourselves to Gamma distributed random effects (except perhaps mathematical convenience).

The vast majority of the papers appeared in the actuarial literature considered time-independent (or static) heterogeneous models. Noticeable exceptions include the pioneering papers by GERBER & JONES (1975), SUNDT (1988) and PINQUET, GUILLÉN & BOLANCÉ (2001). The allowance for an unknown underlying random parameter that develops over time is justified since unobservable factors influencing the driving abilities are not constant. One might consider either shocks (induced by events like divorces or nervous breakdown, for instance) or continuous modifications (e.g. due to learning effect).

Another reason to allow for random effects that vary with time relates to moral hazard. Indeed, individual efforts to prevent accidents are unobserved and feature temporal dependence. The policyholders may adjust their efforts for loss prevention according to their experience with past claims, the amount of premium and awareness of future consequences of an accident (due to experience rating schemes).

The main technical interest of letting the random effects evolve over time is to take into account the date of claims. This reflects the fact that the predictive ability of a claim depends on its age: a recent claim is a worse sign to the insurer than a very old one. Contrarily to the static case, the total number of claims reported in the past is no more an exhaustive summary of policyholders' history. Rather, the sequence of annual claim numbers has now to be memorized to determine future premiums.

In this context, the present paper aims to examine the kind of dependence induced among annual claim numbers by the introduction of random effects taking unexplained heterogeneity into account. We will see that this dependence is often very strong, because of the total positivity of the Poisson kernel. We will also make precise the effect of reporting claims on the *a posteriori* distribution of the random effect. This will be done by establishing some stochastic

monotonicity property of the *a posteriori* distribution with respect to the claims history.

The main interest of this paper is to formalize intuitive ideas with the help of stochastic orderings. Every actuary intuitively feels that the *a posteriori* claim frequency distribution must become more dangerous as claims are reported. We make here precise the meaning of “more dangerous” and we prove that the *a posteriori* premium must increase with the total claim number in the mixed Poisson model. The result is known to be true in the linear credibility approach pioneered by BÜHLMANN (1967, 1970), which can also cope with regression components: the predictor is an increasing function of the number of claims regardless of the mixing distribution. This paper establishes an analogous result for the expected value principle.

To end with this introduction, let us say a few words about the notation used throughout this paper. As usual, the abbreviations iid, rv, cdf stand for independent and identically distributed, random variable and cumulative distribution function, respectively. Henceforth, bold symbols are used for multivariate quantities. Further, $\mathcal{N}(\mu, \sigma^2)$ denotes the Normal law with mean μ and variance σ^2 , $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate Normal law with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\mathcal{LN}(\mu, \sigma^2)$ stands for the LogNormal distribution with parameters μ and σ^2 . All the operations involving vectors have to be interpreted componentwise. Henceforth, given a rv (or random vector) X and an event A , we denote as $[X|A]$ a rv (or vector) with cdf $x \mapsto \Pr[X \leq x|A]$. The symbol “ $=_d$ ” is used to indicate that two rv’s (or two random vectors) have the same distribution, whereas the symbol “ \sim ” means “is distributed according to”.

2. POISSON CREDIBILITY MODELS INCORPORATING A PRIORI RISK CLASSIFICATION

During the observation period, n policies were in portfolio, each one observed during v_i periods. Let N_{it} be the number of claims reported by policyholder i during the period t , $i = 1, 2, \dots, n$, $t = 1, 2, \dots, v_i$. Let d_{it} be the length of this period (the risk exposure). Usually, $d_{it} = 1$, but there are a variety of situations where this is not the case. Indeed, a new period of observation starts as soon as some policy characteristics are modified (think for instance to a moving of the policyholder for a company using postcode as rating factor, policyholder’s wedding for a company using marital status, etc.).

At the beginning of each insurance period, the actuary has at his disposal some information about each policyholder, summarized in a vector \mathbf{x} . Resorting to standard Poisson regression machinery, this information is integrated into the prediction of the annual claim frequency λ . A random effect is superposed to the prediction λ to recognize the residual heterogeneity of the portfolio.

In this paper, we will assume that the unknown characteristics relating to policyholder i during year t are represented by a rv Θ_{it} . The annual numbers of claims $N_{i1}, N_{i2}, N_{i3}, \dots$ are assumed to be independent given the sequence $\Theta_i = (\Theta_{i1}, \Theta_{i2}, \Theta_{i3}, \dots)$ of random effects. The latent unobservable process Θ_i

characterizes the correlation structure of the N_{it} 's. Specifically, the model is based on the following assumptions:

A1 given $\Theta_i = \theta_i$, the rv's N_{it} , $t = 1, 2, \dots, v_i$, are independent and conform to the Poisson distribution with mean $\lambda_{it}\theta_{it}$, i.e.

$$\Pr[N_i = \mathbf{k}_i | \Theta_i = \theta_i] = \prod_{t=1}^{v_i} \Pr[N_{it} = k_{it} | \Theta_{it} = \theta_{it}] = \prod_{t=1}^{v_i} \exp(-\lambda_{it}\theta_{it}) \frac{(\lambda_{it}\theta_{it})^{k_{it}}}{k_{it}!},$$

$\mathbf{k}_i \in \mathbb{N}^{v_i}$;

A2 at the portfolio level, the sequences (Θ_i, N_i) , $i = 1, 2, \dots, n$, are assumed to be independent. Moreover, the Θ_{it} 's are non-negative rv's with unit mean ($\mathbb{E}[\Theta_{it}] = 1$ for all i, t , which means that the *a priori* ratemaking is correct on average). Defining $v_{\max} = \max_i v_i$, Θ_i has the same distribution as the first v_i components of some random vector $(\Theta_1, \dots, \Theta_{v_{\max}})$.

We denote by $G(\cdot)$ (resp. $g(\cdot)$) the common cumulative distribution function (resp. probability density function) of the Θ_{it} 's, for $t = 1, \dots, v_i$. We suppose also that the squared random effects are integrable.

Assuming that $\lambda_{it} = \lambda_{it+h}$, for some integer h , the stationarity of the random effects specified in the condition A2 implies that

$$\Pr[N_{it+h+1} = l_{t+1} | N_{ih+1} = l_1, \dots, N_{it+h} = l_t] = \Pr[N_{it+1} = l_{t+1} | N_{i1} = l_1, \dots, N_{it} = l_t]$$

whatever l_1, \dots, l_{t+1} . Therefore the predictive ability of claims will depend solely on the lag between the date of prediction and the date of occurrence, that is, the age of the claim.

In the model A1-A2, we intuitively feel that the following statements should be valid: provided that the Θ_{it} 's are "positively dependent",

- S1** the N_{it} 's are "positively dependent"
- S2** Θ_i "increases" in the claims N_i
- S3** Θ_{i,v_i+1} "increases" in the past claims N_i
- S4** N_{i,v_i+1} "increases" in the past claims N_i

The purpose of the remainder of this paper will be to formalize these ideas and to make precise the concepts of positive dependence and increasingness involved in statements S1-S4.

3. MODELLING HETEROGENEITY

Let us now complete the assumptions A1-A2 by various structures matching the constraints enumerated in A2. Model A3 is the classical static credibility model. Models A4-A6 come from PINQUET ET AL. (2001) and PINQUET (2000). Model A7 extends the specifications A3-A6 using copulas.

3.1. Model A3

As it is often the case in actuarial science, we could opt for static heterogeneity, i.e. $\Theta_t \equiv \Theta$. In this first case, statements S1-S4 should be true since the Θ_i 's have perfectly dependent components (these random vectors are comonotonic). This model was studied in details in PURCARU & DENUIT (2002).

3.2. Model A4

Let $W \sim \mathcal{N}(\mathbf{0}, \Sigma)$ where $\sigma_{st} = \sigma_W^2 \rho_W(|s-t|)$ with $|\rho_W(h)| \leq 1$, $\rho_W(0) = 1$. The correlation function ρ_W plays a central role in the analysis of the predictive ability of past claims, as shown in PINQUET ET AL. (2001). Now, define

$$\Theta_t = \frac{\exp W_t}{\mathbb{E}[\exp W_t]} \sim \mathcal{LN}\left(-\frac{\sigma_W^2}{2}, \sigma_W^2\right), \quad t = 1, \dots, v_{\max}.$$

Hence, all the Θ_i 's conform to multivariate LogNormal distributions.

Note that if the autocorrelation function $\rho_W(h) = 1$ for all h , we have in fact $\Theta_i = (\Theta_i, \dots, \Theta_i)$, with each Θ_i conforming to the LogNormal distribution; we thus find a special case of model A3.

3.3. Model A5

We could further specify model A4 by assuming in addition that W has an autoregressive structure of order one, that is,

$$W_t = \rho W_{t-1} + \epsilon_t, \quad t \geq 2,$$

where the errors $\epsilon_t \sim \mathcal{N}(0, \sigma^2(1-\rho^2))$ are independent, $|\rho| < 1$, and $W_1 \sim \mathcal{N}(0, \sigma^2)$. In this case, $\rho_W(|s-t|) = \rho^{|s-t|}$ so that the autocorrelation function decreases exponentially with the lag between observations. This specification is particularly interesting for ratemaking purposes.

3.4. Model A6

This model postulates that there is a static baseline heterogeneity which is perturbed by iid annual effects, so that the vector Θ_t is exchangeable. Specifically, $\Theta_t = RS_t$ where the S_t 's are iid and independent from R , $S_t \sim \mathcal{LN}(-\frac{\sigma_S^2}{2}, \sigma_S^2)$ and $R \sim \mathcal{LN}(-\frac{\sigma_R^2}{2}, \sigma_R^2)$. In this case, $\Theta_t \sim \mathcal{LN}(-\frac{\sigma_R^2 + \sigma_S^2}{2}, \sigma_R^2 + \sigma_S^2)$. Therefore in this model the autocorrelation function between the random effects is constant since

$$\begin{aligned} \text{Cov}[\Theta_t, \Theta_{t+h}] &= \mathbb{E}[\text{Cov}[\Theta_t, \Theta_{t+h} | R]] + \text{Cov}[\mathbb{E}[\Theta_t | R], \mathbb{E}[\Theta_{t+h} | R]] \\ &= \text{Cov}[R\mathbb{E}[S_t], R\mathbb{E}[S_{t+h}]] = \text{Var}[R]. \end{aligned}$$

3.5. Model A7

Models A4-A6 are based on LogNormal random effects. Now, we would also like to be able to specify other distributions for the random effects, like the Gamma law, for instance (which facilitates the Bayesian analysis of the data). In that respect, the copula construction is of prime interest.

Let us recall that a bivariate copula C is the joint distribution for a bivariate random couple with uniform marginals. By virtue of Sklar’s theorem, the joint distribution function of (Θ_{t-1}, Θ_t) can be represented as

$$\Pr[\Theta_{t-1} \leq \theta_{t-1}, \Theta_t \leq \theta_t] = C(G(\theta_{t-1}), G(\theta_t))$$

for some copula $C(.,.)$. Provided G is continuous then $C(.,.)$ is unique. For more details about copulas, we refer the interested reader e.g. to FREES and VALDEZ (1998). In a constructive approach, inserting any cdf $G(.)$ (gamma for instance) into some copula $C(.,.)$ (e.g Clayton, Frank or Gumbel) yields a correlated structure for Θ_t ’s.

Let us denote the conditional distribution of the copula by $C_{2|1}(v|u) = \frac{\partial}{\partial u} C(u, v)$. Then,

$$H(\theta_t | \theta_{t-1}) = \Pr[\Theta_t \leq \theta_t | \Theta_{t-1} = \theta_{t-1}] = C_{2|1}(G(\theta_t) | G(\theta_{t-1}))$$

and, if we denote by c the density of the copula C ,

$$h(\theta_t | \theta_{t-1}) = \frac{\partial}{\partial \theta_t} H(\theta_t | \theta_{t-1}) = c(G(\theta_{t-1}), G(\theta_t)) g(\theta_t).$$

Bivariate copulas offer a powerful tool to build autoregressive models for non-gaussian data (see e.g. JOE (1997)). Specifically, selecting some copula C and marginal cdf G , we compute H and the joint cdf of Θ is then given by

$$\Pr[\Theta \leq \theta] = G(\theta_1) H(\theta_2 | \theta_1) \dots H(\theta_n | \theta_{n-1}).$$

Note that taking $C(u, v) = \min\{u, v\}$ (i.e. the Fréchet upper bound copula) yields comonotonic random effects, which leads to the static credibility model A3.

4. STATEMENT S1

4.1. Stochastic order relations

This section gives the definitions of the stochastic orderings we will use, as well as some intuitive interpretations. For more details about stochastic orderings, we refer the reader e.g. to KAAS, VAN HEERWAARDEN and GOOVAERTS (1994).

Most positive dependence concepts aim to formalize the idea that large values of some component of a random vector tend to be associated with large values of the others. Therefore, we need to be able to decide whether a random vector is indeed “larger” than another one. To this end, we resort to stochastic

orderings. These probabilistic tools will also be used to formalize the increasingness mentioned in statements S2-S4. In this paper, we use two classical order relations, namely *stochastic dominance* and *likelihood ratio order*. Recall that a set $U \subseteq \mathbb{R}^n$ is called *upper* if $y \in U$ whenever $y \geq x$ and $x \in U$.

Definition 4.1. Let X and Y be two n -dimensional random vectors. Then X is said to be smaller than Y in the stochastic dominance, written as $X \preceq_{st} Y$, if

$$\Pr[X \in U] \leq \Pr[Y \in U], \text{ for all upper sets } U \text{ in } \mathbb{R}^n$$

Intuitively \preceq_{st} means that it is more likely that Y takes on “large” values (i.e. any value in an upper set U , for any upper set) than X . An univariate upper set is an interval of the form $(t, +\infty)$ so that Definition 4.1 gives for rv’s X and Y

$$X \preceq_{st} Y \Leftrightarrow \Pr[X > t] \leq \Pr[Y > t] \text{ for all } t \in \mathbb{R}.$$

Despite the intuitive nature of \preceq_{st} , we will also need a stronger order concept in the remainder of the paper, namely likelihood ratio order. This stochastic ordering is particularly useful in parametric models.

Definition 4.2. Let X and Y be two n -dimensional random vectors with continuous (or discrete) cdf’s and let f_X and f_Y denote their (continuous or discrete) pdf’s, respectively. Then X is said to be smaller than Y in the multivariate likelihood ratio order, written as $X \preceq_{lr} Y$, if:

$$f_X(x \wedge y) f_Y(x \vee y) \geq f_X(x) f_Y(y), \text{ for every } x \text{ and } y \text{ in } \mathbb{R}^n$$

where the lattice operators \wedge and \vee are defined as

$$x \vee y = (\max\{x_1, y_1\}, \dots, \max\{x_n, y_n\})$$

and

$$x \wedge y = (\min\{x_1, y_1\}, \dots, \min\{x_n, y_n\}).$$

If X and Y are rv’s then Definition 4.2 gives

$$X \preceq_{lr} Y \Leftrightarrow f_X(x) f_Y(y) \geq f_X(y) f_Y(x), \text{ whenever } x \leq y.$$

It can be shown that \preceq_{lr} is stronger than \preceq_{st} , that is if X and Y are two n -dimensional random vectors, $X \preceq_{lr} Y \Rightarrow X \preceq_{st} Y$ (the implication is strict). For a proof of the latter implication, see e.g. SHAKED and SHANTHIKUMAR (1994).

4.2. Dependence concepts

In order to formalize the positive dependence involved in statement S1, we will present several concepts of dependence for random vectors. For more details about these concepts we refer the reader e.g. to JOE (1997).

Definition 4.3. Let X be a n -dimensional random vector.

(i) X is Associated (A, in short) if

$$\text{Cov}[h_1(X), h_2(X)] \geq 0$$

for all componentwise non-decreasing functions $h_1, h_2: \mathbb{R}^n \rightarrow \mathbb{R}$;

(ii) X is Multivariate Positive Likelihood Ratio Dependent (MPLRD, in short) if its multivariate probability density function f_X is MTP_2 (Multivariate Totally Positive of Order 2), that is if

$$f_X(\mathbf{x} \vee \mathbf{y}) f_X(\mathbf{x} \wedge \mathbf{y}) \geq f_X(\mathbf{x}) f_X(\mathbf{y})$$

holds true for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (note that this amounts to $X \preceq_{lr} X$).

(iii) X is Conditionally Increasing in Sequence (CIS, in short) if

$$[X_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}] \preceq_{st} [X_i | X_1 = x'_1, \dots, X_{i-1} = x'_{i-1}],$$

holds whenever $x_j \leq x'_j, j \in \{1, \dots, i-1\}$, for $i = 2, \dots, n$;

(iv) X is Comonotonic (C, in short) if there exists a rv Z and non-decreasing functions $\varphi_1, \dots, \varphi_n$ such that $X =_d (\varphi_1(Z), \dots, \varphi_n(Z))$.

The dependence concepts defined before are linked by the following chain of implications:

$$X C \Rightarrow X \text{ MPLRD} \Rightarrow X \text{ CIS} \Rightarrow X A. \tag{4.1}$$

The second implication follows from Property 4.4 and from Theorem 4.14 in BARLOW and PROSCHAN (1975); the third implication follows from Theorem 2.4 in JOE (1997).

Moreover, they enjoy some functional invariance property, i.e. for all increasing functions $\varphi_1, \dots, \varphi_n$, we have that:

$$X C \text{ (resp. MPLRD, CIS, A)} \Rightarrow (\varphi_1(X_1), \dots, \varphi_n(X_n)) C \text{ (resp. MPLRD, CIS, A)}.$$

4.3. MTP_2 functions

These functions will be often used in the remainder of the paper. Therefore, we recall hereafter some basic features of such functions. In the bivariate case, the MTP_2 concept for functions involved in the definition of MPLRD reduces to the classical TP_2 . Let us recall that a function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ is said to be TP_2 if the inequality

$$f(s_1, t_1) f(s_2, t_2) \geq f(s_1, t_2) f(s_2, t_1) \tag{4.2}$$

holds true for any $s_1 \leq s_2$ and $t_1 \leq t_2$.

The bivariate TP_2 is of great interest for studying MTP_2 as shown in the next result (taken from KEMPERMAN (1977)).

Property 4.4. A function f is MTP_2 if, and only if, it is TP_2 in pairs, that is for all $i \neq j$

$$(x_i, x_j) \mapsto f(\mathbf{x}) \text{ is } TP_2 \text{ for fixed } x_k, k \neq i, j.$$

The following fundamental property of MTP_2 known as the *basic composition formula* will play a central role in the proofs of the results stated in the next sections.

Property 4.5. Given some functions $h_1 : S \times \mathcal{T} \rightarrow \mathbb{R}$ and $h_2 : \mathcal{T} \times \mathcal{U} \rightarrow \mathbb{R}$, let us define the function h_3 as

$$h_3(\mathbf{s}, \mathbf{u}) = \int_{\mathbf{t} \in \mathcal{T}} h_1(\mathbf{s}, \mathbf{t})h_2(\mathbf{t}, \mathbf{u})d\sigma(\mathbf{t}).$$

If h_1 is MTP_2 on $S \times \mathcal{T}$ and h_2 is MTP_2 on $\mathcal{T} \times \mathcal{U}$ then h_3 is MTP_2 on $S \times \mathcal{U}$.

4.4. Proof of statement S1

Before formalizing statement S1, we need to establish the following technical property.

Property 4.6. In the model A1-A2, given $\theta \leq \theta' \in \mathbb{R}^+$, $[N_{it} | \Theta_{it} = \theta] \preceq_{lr} [N_{it} | \Theta_{it} = \theta']$.

Proof. For $k \leq k' \in \mathbb{N}$ and $\theta \leq \theta' \in \mathbb{R}^+$ we have that

$$\frac{\Pr[N_{it} = k | \Theta_{it} = \theta'] \Pr[N_{it} = k' | \Theta_{it} = \theta]}{\Pr[N_{it} = k | \Theta_{it} = \theta] \Pr[N_{it} = k' | \Theta_{it} = \theta']} = \left\{ \frac{\theta}{\theta'} \right\}^{k' - k}$$

so that

$$\Pr[N_{it} = k | \Theta_{it} = \theta] \Pr[N_{it} = k' | \Theta_{it} = \theta'] \geq \Pr[N_{it} = k | \Theta_{it} = \theta'] \Pr[N_{it} = k' | \Theta_{it} = \theta],$$

from which follows the conclusion. □

We are now ready to state the main result of this section, namely that positive dependence for Θ_i is transmitted to N_i . This result is closely related to previous studies by WHITT (1979), FAHMY ET AL. (1982) and SHAKED & SPIZZICHINO (1998).

Proposition 4.7. In the model A1-A2,

- (i) $\Theta_i \text{ MPLRD} \Rightarrow N_i \text{ MPLRD}$.
- (ii) $\Theta_i \text{ A} \Rightarrow N_i \text{ A}$.

Proof. (i) The joint probability function of N_i can be expressed as

$$\Pr[N_i = \mathbf{k}] = \int_{\theta_i \geq \mathbf{0}} \left\{ \prod_{t=1}^{v_i} \Pr[N_{it} = k_t | \Theta_{it} = \theta_{it}] \right\} f_{\Theta_i}(\boldsymbol{\theta}_i) d\boldsymbol{\theta}_i,$$

where f_{Θ_i} stands for the joint pdf of Θ_i . Denote as

$$\bar{f}(k_1, \dots, k_{v_i}, \theta_{i1}, \dots, \theta_{iv_i}) \equiv \prod_{t=1}^{v_i} \Pr[N_{it} = k_t | \Theta_{it} = \theta_{it}].$$

From Property 4.6 it follows that \bar{f} is a product of TP_2 functions and is hence MTP_2 in $(\mathbf{k}, \boldsymbol{\theta})$ by virtue of Property 4.4. Now, invoking the basic composition formula of Property 4.5 (with $h_1 = \bar{f}$ and $h_2 = f_{\Theta_i}$, the latter function being MTP_2 since Θ_i has been assumed MPLRD) we obtain that N_i is MPLRD.

(ii) By reformulating the covariance we obtain the expression:

$$\text{Cov}[h_1(N_i), h_2(N_i)] = \mathbb{E} \left[\text{Cov}[h_1(N_i), h_2(N_i) | \Theta_i] \right] + \text{Cov} \left[\mathbb{E}[h_1(N_i) | \Theta_i], \mathbb{E}[h_2(N_i) | \Theta_i] \right].$$

Given Θ_i , the components of N_i are independent by virtue of A2. From BARLOW & PROSCHAN (1975, Theorem 2.2, p31) independent rv's are associated and hence the conditional covariance in the first term is almost surely positive and so is the whole first term. Let us denote by $\bar{h}_1(\boldsymbol{\theta}_i) \equiv \mathbb{E}[h_1(N_i) | \Theta_i = \boldsymbol{\theta}_i]$ and by $\bar{h}_2(\boldsymbol{\theta}_i) \equiv \mathbb{E}[h_2(N_i) | \Theta_i = \boldsymbol{\theta}_i]$. Since h_1 and h_2 are increasing functions, Property 4.6 ensures that $\bar{h}_1(\cdot)$ and $\bar{h}_2(\cdot)$ are both increasing functions. In the hypothesis Θ_i was supposed to be associated, thus the second term in the right is also positive, and hence the conclusion follows. □

4.5. Statement S1 in the models A3-A7

Since the type of dependence existing between the N_{it} 's is induced by the type of dependence existing between the Θ_{it} 's, as presented in Proposition 4.7, it remains to study the dependence structure of the random effects Θ_{it} in the models mentioned at the beginning of the article, i.e. A3-A7.

4.5.1. Model A3

Since Θ_i is C, it is *a fortiori* MPLRD, so by virtue of Proposition 4.7(i) N_i is MPLRD. Note that N_i is obviously not C.

4.5.2. Model A4

This model is rather general. Not surprisingly, the type of dependence between the Θ_{it} 's is induced by the form of the covariance matrix Σ of W .

Property 4.8. *In the model A1-A2-A4, the following properties hold:*

- (i) *If the inverse of the covariance matrix Σ of W has all the off-diagonal components nonpositive then N_i is MPLRD.*
- (ii) *If all the components of the covariance matrix Σ of W are nonnegative then N_i is A.*

- Proof.* (i) From TONG (1990, Theorem 4.3.2.), we know that W is MPLRD if, and only if, all the off-diagonal components of the inverse of its covariance matrix are non-positive. Since MPLRD is closed under increasing transformation, each Θ_i is then MPLRD. The announced result then follows from Proposition 4.7(i).
- (ii) The reasoning is similar to the one in (i), since we know from PITT (1982) that W is A if, and only if, all the elements of its covariance matrix are non-negative. □

4.5.3. Model A5

In this third model, the covariance structure of W is expressed in terms of the autoregressive parameter ρ . Provided ρ is non-negative, we get strong positive dependence for the components of N_i , as shown in the following result.

Property 4.9. *In the model A1-A2-A5, $\rho \geq 0 \Rightarrow N_i$ MPLRD.*

Proof. In the model A5, the elements of Σ are given by:

$$\sigma_{it} = \sigma_W^2, \quad \sigma_{st} = \sigma_{ts} = \rho^{|s-t|} \sigma_W^2 \text{ for } |s-t| \geq 1.$$

As it can be seen from JOE (1997, Example 8.1 on page 253), the off-diagonal elements of the matrix $R = \Sigma^{-1}$ are as follows:

$$r_{t,t+1} = r_{t+1,t} = -\frac{\rho}{\sigma_W^2(1-\rho^2)} \text{ for } t = 1, \dots, v_{\max}-1 \text{ and } r_{st} = 0 \text{ for } |s-t| \geq 2$$

and are all non positive when $\rho \geq 0$. Hence W is MPLRD. Since the MPLRD property is functionally invariant, Θ_i is also MPLRD and hence N_i by Proposition 4.7(i). □

4.5.4. Model A6

Let us now turn to the exchangeable random effects. When the LogNormal specification is retained for R_i and the S_{it} 's, no further conditions are needed for MPLRD, as shown in the following result.

Property 4.10. *In the model A1-A2-A6, N_i is MPLRD.*

Proof. The joint density of Θ_i is given by

$$f_{\Theta_i}(\theta_{i1}, \dots, \theta_{iv_i}) = \int_{r \geq 0} \left\{ \prod_{t=1}^{v_i} f_{S_{it}}\left(\frac{\theta_{it}}{r}\right) \right\} f_{R_i}(r) dr. \tag{4.3}$$

Let us show that $f_{S_{it}}\left(\frac{\theta_{it}}{r}\right)$ is TP₂ in (θ_{it}, r) , i.e. that for all $\theta_1 < \theta_2$ and all $r_1 < r_2$ the inequality

$$f_{S_{it}}\left(\frac{\theta_1}{r_1}\right)f_{S_{it}}\left(\frac{\theta_2}{r_2}\right) \geq f_{S_{it}}\left(\frac{\theta_1}{r_2}\right)f_{S_{it}}\left(\frac{\theta_2}{r_1}\right)$$

holds true. This amounts to prove that

$$\begin{aligned} & \exp\left\{-\frac{1}{2\sigma_S^2}\left[\left(\ln\frac{\theta_1}{r_1}-\frac{\sigma_S^2}{2}\right)^2+\left(\ln\frac{\theta_2}{r_2}-\frac{\sigma_S^2}{2}\right)^2\right]\right\} \\ & \geq \exp\left\{-\frac{1}{2\sigma_S^2}\left[\left(\ln\frac{\theta_1}{r_2}-\frac{\sigma_S^2}{2}\right)^2+\left(\ln\frac{\theta_2}{r_1}-\frac{\sigma_S^2}{2}\right)^2\right]\right\}, \end{aligned}$$

or, equivalently, that

$$-\frac{1}{2\sigma_S^2}\{(\ln\theta_1-\ln r_1)^2+(\ln\theta_2-\ln r_2)^2-(\ln\theta_1-\ln r_2)^2-(\ln\theta_2-\ln r_1)^2\} \geq 0$$

which reduces to

$$\frac{1}{\sigma_S^2}(\ln\theta_2-\ln\theta_1)(\ln r_2-\ln r_1) \geq 0.$$

The latter relation is obviously true. Thus $f_{S_{it}}\left(\frac{\theta_{it}}{r}\right)$ is TP_2 in (θ_{it}, r) . The integrand of (4.3) is then MTP_2 in (θ_i, r) whence it follows that f_{θ_i} is MTP_2 . \square

4.5.5. Model A7

In the general autoregressive model induced by bivariate copulas C , the dependence structure of N_i is induced by the properties of C , as expected.

Property 4.11. *In the model A1-A2-A7, N_i is MPLRD provided C is TP_2 , that is*

$$c(u, v)c(u', v') \geq c(u', v)c(u, v') \text{ whenever } u \leq u', v \leq v'.$$

Proof. Let us write the joint pdf of N_i as

$$\Pr[N_i = \mathbf{k}] = \int_{\theta_i \geq 0} \left\{ \prod_{t=1}^{v_i} \Pr[N_{it} = k_t | \Theta_{it} = \theta_{it}] \right\} f_{\theta_i}(\theta_i) d\theta_i.$$

Now, exploiting the autoregressive of order 1 structure yields

$$\begin{aligned} f_{\theta_i}(\theta_i) &= g(\theta_{i1})h(\theta_{i2}|\theta_{i1}) \dots h(\theta_{iv_i}|\theta_{iv_i-1}) \\ &= \left\{ \prod_{t=1}^{v_i} g(\theta_{it}) \right\} c(G(\theta_{i1}), G(\theta_{i2})) \dots c(G(\theta_{iv_i-1}), G(\theta_{iv_i})). \end{aligned}$$

If $(u, v) \mapsto c(u, v)$ is TP_2 then $(\theta_{it}, \theta_{it+1}) \mapsto c(G(\theta_{it}), G(\theta_{it+1}))$ is also TP_2 so that f_{Θ_i} is MTP_2 by Property 4.4. Then, invoking Property 4.5 shows that $\mathbf{k} \mapsto \Pr[N_i = \mathbf{k}]$ is MTP_2 , which ends the proof. \square

Let us give now some prominent examples of copulas, which enjoy the TP_2 property (at least for some values of their parameters), resulting in MPLRD claim numbers N_i :

(i) Frank copula:

$$C_\alpha(u, v) = -\frac{1}{\alpha} \ln \left(\frac{[\eta - (1 - e^{-\alpha u})(1 - e^{-\alpha v})] / \eta}{\eta} \right), \text{ where } \eta = 1 - e^{-\alpha}$$

is TP_2 for $\alpha \geq 0$.

(ii) Normal copula:

$$C_\alpha(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\alpha^2}} \int_{\xi_1=-\infty}^{\Phi^{-1}(u_1)} \int_{\xi_2=-\infty}^{\Phi^{-1}(u_2)} \exp \left\{ \frac{-\left(\xi_1^2 - 2\alpha\xi_1\xi_2 + \xi_2^2\right)}{2(1-\alpha^2)} \right\} d\xi_1 d\xi_2$$

where Φ is the cdf of the $\mathcal{N}(0,1)$ distribution, is TP_2 for $\alpha \geq 0$.

(iii) Gumbel copula:

$$C_\alpha(u, v) = \exp \left\{ -\left(\tilde{u}^\alpha + \tilde{v}^\alpha\right)^{-1/\alpha} \right\}$$

where $\tilde{u} = -\ln u$ and $\tilde{v} = -\ln v$, is TP_2 for $\alpha \geq 1$.

(iv) Clayton copula:

$$C_\alpha(u, v) = \left(u^{-\alpha} + v^{-\alpha} - 1\right)^{-1/\alpha}$$

is TP_2 for $\alpha > 0$.

5. STATEMENT S2

Let us now establish the following result inspired from Theorems 3-4 in FAHMY ET AL. (1982). We purpose to prove that policyholders reporting more claims in the past become more dangerous on unobservable characteristics.

Proposition 5.1. *In the model A1-A2, $[\Theta_i | N_i = \mathbf{k}] \preceq_{lr} [\Theta_i | N_i = \mathbf{k}']$ whenever $\mathbf{k} \leq \mathbf{k}'$ provided Θ_i is MPLRD.*

Proof. The conditional pdf of Θ_i given $[N_i = \mathbf{k}]$ can be expressed as

$$f_{\Theta_i}(\theta | \mathbf{k}) = c(\mathbf{k}) \Pr[N_i = \mathbf{k} | \Theta_i = \theta] f_{\Theta_i}(\theta)$$

where $c(\mathbf{k})$ is a normalizing constant given by

$$[c(\mathbf{k})]^{-1} = \Pr[N_i = \mathbf{k}] = \int_{\lambda \geq 0} \Pr[N_i = \mathbf{k} | \Theta_i = \lambda] f_{\Theta_i}(\lambda) d\lambda.$$

Let us now consider two claim histories $\mathbf{k} \leq \mathbf{k}'$ as well as two possible values for the random effects θ and θ' . Starting from

$$\begin{aligned} f_{\Theta_i}(\theta|\mathbf{k})f_{\Theta_i}(\theta'|\mathbf{k}') &= c(\mathbf{k})c(\mathbf{k}')\Pr[N_i = \mathbf{k} | \Theta_i = \theta] \Pr[N_i = \mathbf{k}' | \Theta_i = \theta'] f_{\Theta_i}(\theta)f_{\Theta_i}(\theta') \\ &= c(\mathbf{k})c(\mathbf{k}') \left\{ \prod_{t=1}^{v_i} \Pr[N_{it} = k_t | \Theta_{it} = \theta_{it}] \right\} \\ &\quad \left\{ \prod_{t=1}^{v_i} \Pr[N_{it} = k'_t | \Theta_{it} = \theta'_{it}] \right\} f_{\Theta_i}(\theta)f_{\Theta_i}(\theta') \end{aligned}$$

we get from the MPLRD nature of Θ_i together with Property 4.6 that

$$\begin{aligned} f_{\Theta_i}(\theta|\mathbf{k})f_{\Theta_i}(\theta'|\mathbf{k}') &\leq c(\mathbf{k})c(\mathbf{k}') \left\{ \prod_{t=1}^{v_i} \Pr[N_{it} = k_t \wedge k'_t | \Theta_{it} = \theta_{it} \wedge \theta'_{it}] \right\} \\ &\quad \left\{ \prod_{t=1}^{v_i} \Pr[N_{it} = k_t \vee k'_t | \Theta_{it} = \theta_{it} \vee \theta'_{it}] \right\} f_{\Theta_i}(\theta \wedge \theta')f_{\Theta_i}(\theta \vee \theta') \\ &= c(\mathbf{k})c(\mathbf{k}')\Pr[N_i = \mathbf{k} \wedge \mathbf{k}' | \Theta_i = \theta \wedge \theta'] \Pr[N_i = \mathbf{k} \vee \mathbf{k}' | \Theta_i = \theta \vee \theta'] \\ &\quad f_{\Theta_i}(\theta \wedge \theta')f_{\Theta_i}(\theta \vee \theta') \\ &= \left\{ c(\mathbf{k})\Pr[N_i = \mathbf{k} | \Theta_i = \theta \wedge \theta'] f_{\Theta_i}(\theta \wedge \theta') \right\} \\ &\quad \left\{ c(\mathbf{k}')\Pr[N_i = \mathbf{k}' | \Theta_i = \theta \vee \theta'] f_{\Theta_i}(\theta \vee \theta') \right\} \\ &= f_{\Theta_i}(\theta \wedge \theta'|\mathbf{k}) f_{\Theta_i}(\theta \vee \theta'|\mathbf{k}') \end{aligned}$$

whence the announced result follows. □

The conditions under which Θ_i is MPLRD have been studied in Section 4.4 for the models A3-A6. It is interesting to contrast Proposition 5.1 to the corresponding result for the static credibility model (Proposition 3.4 in PURCARU & DENUIT (2002)); see Section 7.

6. STATEMENTS S3 AND S4 IN MODELS A5-A7

The whole credibility theory aims to predict as accurately as possible the future claims given the past observed. This section focusses on predictive distribution, that is the distribution of the future number of claims N_{i_v+1} given past claims N_i . Model A4 is hardly used to perform prediction on longitudinal basis. This is due to the fact that the correlation function ρ_W has to be continued beyond the observation period before evaluating experience premiums (see PINQUET

ET AL. (2001) for a complete treatment of this subject). In this section, we concentrate on models A5-A7 and assume that the conditions of Properties 4.9, 4.10 and 4.11 are fulfilled. This ensures that the vector $(\Theta_{i1}, \dots, \Theta_{iv_i}, \Theta_{iv_i+1})$ is also MPLRD.

Property 6.1. *In the model A1-A2 completed by either A5, A6 or A7, $[\Theta_{iv_i+1} | N_i = \mathbf{k}] \preceq_{lr} [\Theta_{iv_i+1} | N_i = \mathbf{k}']$ for all $\mathbf{k} \leq \mathbf{k}'$ under the conditions of Property 4.9, 4.10 or 4.11.*

Proof. The aim is to prove that the inequality

$$f(\theta_{iv_i+1} | \mathbf{k}) f(\theta'_{iv_i+1} | \mathbf{k}') \geq f(\theta'_{iv_i+1} | \mathbf{k}) f(\theta_{iv_i+1} | \mathbf{k}') \tag{6.1}$$

holds true for all $\mathbf{k} \leq \mathbf{k}'$ and $\theta_{iv_i+1} \leq \theta'_{iv_i+1}$, where $f(\cdot | \mathbf{k})$ denotes the conditional density of Θ_{iv_i+1} given $N_i = \mathbf{k}$. Bayes formula gives

$$f(\theta_{iv_i+1} | \mathbf{k}) = \frac{\Pr[N_i = \mathbf{k} | \Theta_{iv_i+1} = \theta_{iv_i+1}] g(\theta_{iv_i+1})}{\int_{\lambda \geq 0} \Pr[N_i = \mathbf{k} | \Theta_{iv_i+1} = \lambda] g(\lambda) d\lambda}$$

and thus showing (6.1) is equivalent to verify that the inequality

$$\begin{aligned} & \Pr[N_i = \mathbf{k} | \Theta_{iv_i+1} = \theta_{iv_i+1}] \Pr[N_i = \mathbf{k}' | \Theta_{iv_i+1} = \theta'_{iv_i+1}] \\ & \geq \Pr[N_i = \mathbf{k}' | \Theta_{iv_i+1} = \theta_{iv_i+1}] \Pr[N_i = \mathbf{k} | \Theta_{iv_i+1} = \theta'_{iv_i+1}], \end{aligned}$$

holds for all $\mathbf{k} \leq \mathbf{k}'$ and $\theta_{iv_i+1} \leq \theta'_{iv_i+1}$. Let us denote as $h(\mathbf{k}, \theta)$ the joint density of (N_i, Θ_{iv_i+1}) evaluated at (\mathbf{k}, θ) . It then remains to prove that

$$h(\mathbf{k}, \theta_{iv_i+1}) h(\mathbf{k}', \theta'_{iv_i+1}) \geq h(\mathbf{k}', \theta_{iv_i+1}) h(\mathbf{k}, \theta'_{iv_i+1}).$$

To establish the latter inequality, we use an interesting property of the MTP_2 concept, which stipulates that if the joint distribution of a random vector is MTP_2 , then the same property still holds for any of the marginals. Since the joint density of $(N_{i1}, \dots, N_{iv_i+1}, \Theta_{i1}, \dots, \Theta_{iv_i+1})$ given by

$$h(k_1, \dots, k_{v_i+1}, \theta_{i1}, \dots, \theta_{iv_i+1}) \equiv \left\{ \prod_{t=1}^{v_i+1} \Pr[N_{it} = k_t | \Theta_{it} = \theta_{it}] \right\} f_{\Theta_i}(\theta_{i1}, \dots, \theta_{iv_i+1})$$

is obviously MTP_2 when Θ_i is MPLRD, we get the announced result. □

Now that we have shown that a policyholder having reported more claims in the past will be more dangerous on the unobservable characteristics relating to the year $v_i + 1$, we would like to establish the same result for the future number of claims. This is precisely done next.

Property 6.2. *Under the assumptions of Property 6.1, $[N_{i_v+1} | N_i = \mathbf{k}] \preceq_{lr} [N_{i_v+1} | N_i = \mathbf{k}']$, for all $\mathbf{k} \leq \mathbf{k}'$.*

Proof. We have to show that for all $n \leq n'$ and $\mathbf{k} \leq \mathbf{k}'$ the following inequality

$$\begin{aligned} & \Pr[N_{i_v+1} = n | N_i = \mathbf{k}] \Pr[N_{i_v+1} = n' | N_i = \mathbf{k}'] \\ & \geq \Pr[N_{i_v+1} = n | N_i = \mathbf{k}'] \Pr[N_{i_v+1} = n' | N_i = \mathbf{k}] \end{aligned}$$

holds true. The conditional probability could also be expressed in the form:

$$\Pr[N_{i_v+1} = n | N_i = \mathbf{k}] = \int_{\theta \geq 0} \Pr[N_{i_v+1} = n | \Theta_{i_v+1} = \theta] f(\theta | \mathbf{k}) d\theta$$

where $f(\cdot | \mathbf{k})$ denotes the conditional density of Θ_{i_v+1} given $N_i = \mathbf{k}$. The first factor of the integrand is TP_2 in (n, θ) , based on Property 4.6. Moreover, we have seen in Property 6.1 that $f(\theta | \mathbf{k})$ is MTP_2 in $(k_1, \dots, k_{v_i}, \theta)$. By applying the basic composition formula for

$$h_1(n, \theta) \equiv \Pr[N_{i_v+1} = n | \Theta_{i_v+1} = \theta] \text{ and } h_2(\theta, \mathbf{k}) \equiv f(\theta | \mathbf{k})$$

it follows that $(n, \mathbf{k}) \mapsto \Pr[N_{i_v+1} = n | N_i = \mathbf{k}]$ is MTP_2 , which ends our proof. \square

7. SOME PARTICULARITIES OF THE STATIC MODEL A3

In the static credibility models, which are investigated in details in PURCARU & DENUIT (2002), conditions over the N_i in the dynamic case, reduce to conditions over

$$N_{i\bullet} = \sum_{t=1}^{v_i} N_{it},$$

where $N_{i\bullet}$ is the total claim number reported by policyholder i during the v_i observation periods. The statistic $N_{i\bullet}$ is a convenient summary of past claims history but it neglects the age of the claims. Therefore we have the following result, which formalizes statements S2 and S3; see PURCARU & DENUIT (2002) for formal proofs.

Proposition 7.1. *In the model A3, the following statements are true:*

- (i) $[\Theta_i | N_{i\bullet} = n] \preceq_{lr} [\Theta_i | N_{i\bullet} = n']$ for $n \leq n'$;
- (ii) $[N_{i_v+1} | N_{i\bullet} = n] \preceq_{lr} [N_{i_v+1} | N_{i\bullet} = n']$ for $n \leq n'$.

In the static case (model A3), the condition thus involves $N_{i\bullet}$ (and not the whole vector N_i), the total number of claims reported by policyholder i in the past: increasing the value of $N_{i\bullet}$ makes the unique random effect Θ_i larger in the univariate \preceq_{lr} -sense. In that framework, the distribution of the claims during the driving career of the policyholder does not matter.

8. CONCLUSION

The present paper aimed to investigate the kind of dependence generated by actuarial credibility models with time-dependent random effects. To the best of the authors' knowledge, this aspect of actuarial modelling has never been investigated in the literature so far. As expected, the structure of dependence between the observable claim numbers is inherited from the type of correlation existing between the unobservable random effects modelling residual heterogeneity. It turns out that the kind of dependence induced by these models is often very strong, namely MPLRD.

It is worth mentioning that most of the reasonings only use the fact that the Poisson distribution is monotone in its mean in the \leq_{lr} -sense (as shown in Property 4.6). So the results are readily extended to any other claim frequency distribution possessing this property. See also SHAKED and SPIZZICHINO (1998) for similar results involving absolutely continuous conditional distributions.

ACKNOWLEDGEMENTS

The support of the Belgian Government under contract "Project d'Action de Recherche Concertées" ARC 98/03-217 is gratefully acknowledged.

The authors warmly thank two anonymous referees and an editor for careful reading and for numerous remarks that have substantially improved the paper.

REFERENCES

- BARLOW, R.E. and PROSCHAN, F. (1975) *Statistical Theory of Reliability and Life Testing – Probability Models*. Holt, Rinehart and Winston.
- BÜHLMANN, H. (1967) Experience rating and credibility. *ASTIN Bulletin* **4**, 199-207.
- BÜHLMANN, H. (1970) *Mathematical Methods in Risk Theory*. Springer Verlag, New York.
- DIONNE, G. and VANASSE, C. (1989) A generalization of actuarial automobile insurance rating models: the Negative Binomial distribution with a regression component. *ASTIN Bulletin* **19**, 199-212.
- FREES, E.W. and VALDEZ, E.A. (1998) Understanding relationships using copulas. *North American Actuarial Journal* **2**, 1-15.
- FAHMY, S., PEREIRA, C.A., PROSCHAN, F. and SHAKED, M. (1982) The influence of the sample on the posterior distribution. *Communications in Statistics – Theory & Methods* **11**, 1157-1168.
- GERBER, H.U. and JONES, D. (1975) Credibility formulas of the updating type. *Transactions of the Society of Actuaries* **27**, 31-52.
- JOE, H. (1997) *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- KAAS, R., VAN HEERWAARDEN, A.E. and GOOVAERTS, M.J. (1994) *Ordering of Actuarial Risks*. CAIRE. Brussels.
- KEMPERMAN, J.H.B. (1977) On the FKG-inequality for measures on a partially ordered space. *Indagationes Mathematicae* **39**, 313-331.
- PINQUET, J. (2000) Experience rating through heterogeneous models. In *Handbook of Insurance*, edited by G. Dionne. Kluwer Academic Publishers.
- PINQUET, J., GUILLÉN, M. and BOLANÇÉ, C. (2001) Allowance for the age of claims in bonus-malus systems. *ASTIN Bulletin* **31**, 337-348.

- PITT, L.D. (1982) Positively correlated normal variables are associated. *The Annals of Probability* **10**, 496-499.
- PURCARU, O. and DENUIT, M. (2002) On the dependence induced by frequency credibility models. *Belgian Actuarial Bulletin* **2**, 74-80.
- RENSHAW, A.E. (1994) Modelling the claim process in the presence of covariates. *ASTIN Bulletin* **24**, 265-285.
- SHAKED, M. and SHANTHIKUMAR, J.G. (1994) *Stochastic Orders and Their Applications*. Academic Press, New York.
- SHAKED, M. and SPIZZICHINO, F. (1998) Positive dependence properties of conditionally independent random lifetimes. *Mathematics of Operations Research* **23**, 944-959.
- SUNDT, B. (1988) Credibility estimators with geometric weights. *Insurance: Mathematics and Economics* **7**, 113-122.
- TONG, Y.L. (1990) *The Multivariate Normal Distribution*. Springer-Verlag, New York.
- WHITT, W. (1979) A note on the influence of the sample on the posterior distribution. *Journal of the American Statistical Association* **74**, 424-426.

OANA PURCARU & MICHEL DENUIT
Institut de Statistique & Institut des Sciences Actuarielles
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve
Belgium
E-mail: purcaru@stat.ucl.ac.be & denuit@stat.ucl.ac.be