More General Credibility Models

Virginia ATANASIU
Department of Mathematics, Academy of Economic Studies
e-mail: virginia_atanasiu@yahoo.com

This communication gives some extensions of the original Bühlmann model. The paper is devoted to semi-linear credibility, where one examines functions of the random variables representing claim amounts, rather than the claim amounts themselves. The main purpose of semi-linear credibility theory is the estimation of $\mu_0(\theta)=E[f_0(X_{t+1})|\theta]$ *(the net premium for a contract with risk parameter:* θ*) by a linear combination of given functions of the observable variables:* $\underline{X}^{'}=(X_1,X_2,...,X_t)$*. So the estimators mainly considered here are linear functions of several functions* $f_1,f_2,...,f_n$ *of the observable random variables. The approximation to* $\mu_0(\theta)$ *based on prescribed approximating functions* $f_1,f_2,...,f_n$ *leads to the optimal non-homogeneous linearized estimator for the semi-linear credibility model. Also we discuss the case when taking* $f_p=f$ *for all:* p*, try to find the optimal function* f*. It should be noted that the approximation to* $\mu_0(\theta)$ *based on a unique optimal approximating function* f *is always better than the one furnished in the semi-linear credibility model based on prescribed approximating functions:* $f_1,f_2,...,f_n$*. The usefulness of the latter approximation is that it is easy to apply, since it is sufficient to know estimates for the structural parameters appearing in the credibility factors. From this reason we give some unbiased estimators for the structure parameters. For this purpose we embed the contract in a collective of contracts, all providing independent information on the structure distribution. We close this paper by giving the semi-linear hierarchical model used in the applications chapter.*

Mathematics Subject Classification: 62P05.
Keywords*: contracts, unbiased estimators, structure parameters, several approximating functions, semi-linear credibility theory, unique optimal function, parameter estimation, hierarchical semi-linear credibility theory.*

Introduction

In this article we first give the semi-linear credibility model (see **Section 1**), which involves only one isolated contract. Our problem (from **Section1**) is the estimation of $\mu_0(\theta)=E[f_0(X_{t+1})|\theta]$ (the net premium for a contract with risk parameter: θ) by a linear combination of given functions $f_1,f_2,...,f_n$ of the observable variables: $\underline{X}^{'}=(X_1,X_2,...,X_t)$. So our problem (from **Section 1**) is the determination of the linear combination of 1 and the random variables: $f_p(X_r)$, $p=\overline{1,n}$, $r=\overline{1,t}$ closest to $\mu_0(\theta)=E[f_0(X_{t+1})|\theta]$ in the least squares sense, where θ is the structure variable. The solution of this problem:

$$\underset{\alpha_0,\alpha}{Min}\, E\left\{\left[\mu_0(\theta)-\alpha_0-\sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}f_p(X_r)\right]^2\right\}, \text{ where: } \alpha=(\alpha_{pr})_{p,r},$$

is the optimal non-homogeneous linearized estimator (namely the semi-linear credibility result). In **Section 2** we discuss the case when taking $f_p=f$ for all: p, try to find the unique optimal function f. It should be noted that the approximation to $\mu_0(\theta)$ based on a unique optimal approximating function f is always better than the one furnished in the semi-linear credibility model based on prescribed approximating functions: $f_1, f_2, ..., f_n$. The usefulness of the latter approximation is that it is easy to apply, since it is sufficient to know estimates for the structural parameters: a_{pq}, b_{pq} (with p, $q = \overline{0,n}$) appearing in the credibility factors z_p (where $p = \overline{1,n}$). To obtain estimates for these structure parameters from the semi-linear credibility model, in **Section 3** we embed the contract in a collective of contracts, all providing independent information on the structure distribution. We close this paper by giving the semi-linear hierarchical model used in the applications chapter (see **Section 4**).

Section 1 (**The approximation to** $\mu_0(\theta)$ **based on prescribed approximating functions:** $f_1, f_2, ..., f_n$)

In this section, we consider one contract with unknown and fixed risk parameter: θ, during a period of t years. The yearly claim amounts are denoted by: $X_1, ..., X_t$. The risk parameter θ is supposed to be drawn from some structure distribution function: $U(\cdot)$. It is assumed that, for given: θ, the claims are conditionally independent and identically distributed (conditionally i.i.d.) with known common distribution function $F_{X|\theta}(x,\theta)$. The random variables $X_1, ..., X_t$ are observable, and the random variable X_{t+1} is considered as being not (yet) observable. We assume that: $f_p(X_r)$, $p = \overline{0,n}$, $r = \overline{1,t+1}$ have finite variance. For: f_0, we take the function of X_{t+1} we want to forecast.

We use the notation:

$$\mu_p(\theta) = E[f_p(X_r) \mid \theta] \quad (p = \overline{0,n}; r = \overline{1,t+1}) \tag{1.1}$$

This expression does not depend on r.

We define the following structure parameters:

$$m_p = E[\mu_p(\theta)] = E\{E[f_p(X_r) \mid \theta]\} = E[f_p(X_r)] \tag{1.2),}$$

$$a_{pq} = E\{Cov[f_p(X_r), f_q(X_r) \mid \theta]\} \tag{1.3),}$$

$$b_{pq} = Cov[\mu_p(\theta), \mu_q(\theta)] \tag{1.4),}$$

$$c_{pq} = Cov[f_p(X_r), f_q(X_r)] \tag{1.5),}$$

$$d_{pq} = Cov[f_p(X_r), \mu_q(\theta)] \tag{1.6),}$$

for: p, $q = \overline{0,n} \wedge r = \overline{1,t+1}$. These expressions do not depend on: $r = \overline{1,t+1}$. The structure parameters are connected by the following relations:

$$c_{pq} = a_{pq} + b_{pq} \tag{1.7),}$$

$$d_{pq} = b_{pq} \tag{1.8),}$$

for: $p, q = \overline{0,n}$. This follows from the covariance relations obtained in the probability theory where they are very well-known. Just as in the case of considering linear combinations of the observable variables themselves, we can also obtain non-homogeneous credibility estimates, taking as estimators the class of linear combinations of given functions of the observable variables, as shown in the following theorem:

Theorem 1.1 (*Optimal non-homogeneous linearized estimators*)

The linear combination of 1 and the random variables $f_p(X_r), p = \overline{1,n}; r = \overline{1,t}$ closest to $\mu_0(\theta) = E[f_0(X_{t+1}) \mid \theta]$ and to $f_0(X_{t+1})$ in the least squares sense equals:

$$M = \sum_{p=1}^{n} z_p \sum_{r=1}^{t} \frac{1}{t} f_p(X_r) + m_0 - \sum_{p=1}^{n} z_p m_p \tag{1.9),}$$

where $z_1, z_2, ..., z_n$ is a solution to the linear system of equations:

$$\sum_{p=1}^{n}\left[c_{pq}+(t-1)d_{pq}\right]z_p = td_{0q}\ (q=\overline{1,n}) \qquad (1.10)$$

or to the equivalent linear system of equations:

$$\sum_{p=1}^{n}\left(a_{pq}+tb_{pq}\right)z_p = tb_{0q}\ (q=\overline{1,n}) \qquad (1.11)$$

Proof: we have to examine the solution of the problem:

$$\underset{\alpha_0,\alpha}{Min}\, E\left\{\left[\mu_0(\theta)-\alpha_0-\sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}f_p(X_r)\right]^2\right\} \qquad (1.12)$$

Taking the derivative with respect to α_0 gives:

$E[\mu_0(\theta)]-\sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}E[f_p(X_r)]=\alpha_0$, or: $\alpha_0 = m_0-\sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}m_p$. Inserting this expression for α_0 into (1.12) leads to the following problem:

$$\underset{\alpha}{Min}\, E\left\{\left[\mu_0(\theta)-m_0-\sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}\left(f_p(X_r)-m_p\right)\right]^2\right\} \qquad (1.13)$$

On putting the derivatives with respect to $\alpha_{qr'}$ equal to zero, we get the following system of equations ($q=\overline{1,n}$; $r'=\overline{1,t}$):

$$Cov[\mu_0(\theta), f_q(X_{r'})]=\sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}Cov[f_p(X_r), f_q(X_{r'})] \qquad (1.14)$$

Because of the symmetry in time clearly: $\alpha_{p1}=\alpha_{p2}=...=\alpha_{pt}=\alpha_p$, so using the co-variance results, for $q=\overline{1,n}$ this system of equations can be written as:

$$b_{0q}=\sum_{p=1}^{n}\alpha_p\left[c_{pq}+(t-1)d_{pq}\right] \qquad (1.15)$$

Now (1.15) and (1.13) lead to (1.9) with: $\alpha_p=\frac{z_p}{t}$, $p=\overline{1,n}$.

Section 2 (The approximation to $\mu_0(\theta)$ based on a unique optimal approximating function: f)

The estimator M for $\mu_0(\theta)$ of *Theorem 1.1* can be displayed as:

$$M=f(X_1)+...+f(X_t) \qquad (2.1),$$

where:

$$f(x)=\frac{1}{t}\sum_{p=1}^{n}z_pf_p(x)+\frac{1}{t}m_0-\frac{1}{t}\sum_{p=1}^{n}z_pm_p.$$

Let us forget now about this structure of f and look for any function f such that (2.1) is closest to: $\mu_0(\theta)$. If are considered only functions f such that $f(X_1)$ has finite variance, then the optimal approximating function f results from the following theorem:

Theorem 2.1 (*Optimal approximating function*)

$f(X_1)+...+f(X_t)$ is closest to $\mu_0(\theta)$ and to $f_0(X_{t+1})$ in the least squares sense, if and only if f is a solution of the equation:

$$f(X_1)+(t-1)E[f(X_2)|X_1]-E[f_0(X_2)|X_1]\equiv 0 \qquad (2.2)$$

Proof: we have to solve the following minimization problem:

$$\underset{g}{Min}\, E\left\{[f_0(X_{t+1})-g(X_1)-...-g(X_t)]^2\right\} \qquad (2.3)$$

Suppose that f denotes the solution to this problem, then we consider: $g(X)=f(X)+\alpha h(X)$, with $h(\cdot)$ arbitrary, like in variational calculus. Let:

$$\varphi(\alpha)=E\left\{\left[f_0(X_{t+1})-f(X_1)-...-f(X_t)-\alpha h(X_1)-...-\alpha h(X_t)\right]^2\right\} \quad (2.4)$$

Clearly for f to be optimal, $\varphi'(0)=0$, so for every choice of h:

$$E\{[f_0(X_{t+1})-f(X_1)-...-f(X_t)][h(X_1)+...+h(X_t)]\}=0 \quad (2.5),$$

must hold. This can be rewritten as:

$$E[tf_0(X_2)h(X_1)-tf(X_1)h(X_1)-t(t-1)f(X_2)h(X_1)]=0 \quad (2.6),$$

or:

$$E\left[h(X_1)\left\{-f(X_1)-(t-1)E[f(X_2)|X_1]+E[f_0(X_2)|X_1]\right\}\right]=0 \quad (2.7)$$

Because this equation has to be satisfied for every choice of the function h one obtains, the expression in brackets in (2.7) must be identical to zero, which proves (2.2).

An application of Theorem 2.1:

If $X_1,...,X_{t+1}$ can only take the values $0,1,...,n$ and $p_{qr}=P[X_1=q,X_2=r]$ for: q, $r=\overline{0,n}$, then $f(X_1)+...+f(X_t)$ is closest to $\mu_0(\theta)$ and to $f_0(X_{t+1})$ in the least squares sense, if and only if for $q=\overline{0,n}$, $f(q)$ is a solution of the linear system:

$$f(q)\sum_{r=0}^{n}p_{qr}+(t-1)\sum_{r=0}^{n}f(r)p_{qr}=\sum_{r=0}^{n}f_0(r)p_{qr} \quad (2.8)$$

Indeed: $f(X_1):\begin{pmatrix} f(q) \\ P(X_1=q)\end{pmatrix}=\begin{pmatrix} f(q) \\ \sum_{r=0}^{n}p_{qr}\end{pmatrix}$, $q=\overline{0,n}$; $E[f(X_2)|X_1]=\sum_{r=0}^{n}f(r)P(X_2=r|X_1=q)=\sum_{r=0}^{n}f(r)\frac{p_{qr}}{\sum_{r=0}^{n}p_{qr}}$; $E[f_0(X_2)|X_1]=\sum_{r=0}^{n}f_0(r)P(X_2=r|X_1=q)=\sum_{r=0}^{n}f_0(r)\frac{p_{qr}}{\sum_{r=0}^{n}p_{qr}}$.

Inserting these expressions for: $f(X_1)$, $E[f(X_2)|X_1]$ and $E[f_0(X_2)|X_1]$ into (2.2) leads to (2.8).

Section 3 (**Parameter estimation**)

It should be noted that the approximation to $\mu_0(\theta)$ based on a unique optimal approximating function f is always better than the one furnished in **Section 1** based on prescribed approximating functions: $f_1,f_2,...,f_n$. The usefulness of the latter approximation is that it is easy to apply, since it is sufficient to know estimates for the structural parameters a_{pq}, b_{pq} (with p, $q=\overline{0,n}$) appearing in the credibility factors z_p (where $p=\overline{1,n}$). From this reason we give some unbiased estimators for the structure parameters. For this purpose we consider k contracts, $j=\overline{1,k}$, and k (≥ 2) independent and identically distributed vectors $(\theta_j,\underline{X}'_j)=(\theta_j,X_{j1},...,X_{jt})$, for $j=\overline{1,k}$. The contract indexed j is a random vector consisting of a random structure parameter θ_j and observations: $X_{j1},...,X_{jt}$, where $j=\overline{1,k}$. For every contract $j=\overline{1,k}$ and for θ_j fixed, the variables: $X_{j1},...,X_{jt}$ are conditionally independent and identically distributed.

Theorem 3.1 (Unbiased estimators for the structure parameters)

Let:

$$\hat{m}_p=\frac{1}{kt}X^p_{..}=\frac{1}{kt}\sum_{j=1}^{k}\sum_{r=1}^{t}f_p(X_{jr}) \quad (3.1)$$

$$\hat{a}_{pq}=\frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}\left(X^p_{jr}-\frac{1}{t}X^p_{j.}\right)\left(X^q_{jr}-\frac{1}{t}X^q_{j.}\right) \quad (3.2)$$

$$\hat{b}_{pq} = \frac{1}{k-1}\sum_{j=1}^{k}\left(\frac{1}{t}X_{j.}^{p} - \frac{1}{kt}X_{..}^{p}\right)\left(\frac{1}{t}X_{j.}^{q} - \frac{1}{kt}X_{..}^{q}\right) - \frac{\hat{a}_{pq}}{t} \qquad (3.3),$$

, then: $E\left(\hat{m}_p\right) = m_p$, $E\left(\hat{a}_{pq}\right) = a_{pq}$, $E\left(\hat{b}_{pq}\right) = b_{pq}$, where: $X_{j.}^{p} = \sum_{r=1}^{t} X_{jr}^{p}$, $X_{j.}^{q} = \sum_{r=1}^{t} X_{jr}^{q}$, $X_{..}^{p} = \sum_{j=1}^{k}\sum_{r=1}^{t} X_{jr}^{p}$, $X_{..}^{q} = \sum_{j=1}^{k}\sum_{r=1}^{t} X_{jr}^{q}$, with $X_{jr}^{p} = f_p\left(X_{jr}\right)$, ($j = \overline{1,k}$ and $r = \overline{1,t}$), $X_{jr}^{q} = f_q\left(X_{jr}\right)$, ($j = \overline{1,k}$ and $r = \overline{1,t}$), for p, $q = \overline{0,n}$, such that $p < q$.

Proof: note that the usual definitions of the structure parameters apply, with θ_j replacing θ and X_{jr} replacing X_r so: $E\left(\hat{m}_p\right) = \frac{1}{kt}\sum_{j,r} E\left[f_p\left(X_{jr}\right)\right] = \frac{1}{kt}\sum_{j,r} m_p = \frac{kt}{kt} m_p = m_p$;

$$E\left(\hat{a}_{pq}\right) = \frac{1}{k(t-1)}\sum_{j,r}\left[Cov\left(X_{jr}^{p}, X_{jr}^{q}\right) + E\left(X_{jr}^{p}\right)E\left(X_{jr}^{q}\right) - Cov\left(X_{jr}^{p}, \frac{1}{t}X_{j.}^{q}\right) - E\left(X_{jr}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right) - \right.$$
$$\left. - Cov\left(\frac{1}{t}X_{j.}^{p}, X_{jr}^{q}\right) - E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(X_{jr}^{q}\right) + Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{t}X_{j.}^{q}\right) + E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right)\right] = \frac{1}{k(t-1)}\cdot$$
$$\cdot\sum_{j,r}\left[\left(a_{pq} + b_{pq}\right) + m_p m_q - \left(\frac{1}{t}a_{pq} + b_{pq}\right) - m_p m_q - \left(\frac{1}{t}a_{pq} + b_{pq}\right) - m_p m_q + \left(\frac{1}{t}a_{pq} + b_{pq}\right) + m_p m_q\right]$$
$$= \frac{1}{k(t-1)}\sum_{j,r}\left(a_{pq} + b_{pq} - \frac{1}{t}a_{pq} - b_{pq}\right) = \frac{1}{k(t-1)}kt\frac{(t-1)}{t}a_{pq} = a_{pq};\; E\left(\hat{b}_{pq}\right) = \frac{1}{k-1}\sum_{j}\left[Cov\left(\frac{1}{t}\cdot\right.\right.$$
$$\left. \cdot X_{j.}^{p}, \frac{1}{t}X_{j.}^{q}\right) + E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right) - Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{kt}X_{..}^{q}\right) - E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(\frac{1}{kt}X_{..}^{q}\right) - Cov\left(\frac{1}{kt}X_{..}^{p},\right.$$
$$\left. \left.,\frac{1}{t}X_{j.}^{q}\right) - E\left(\frac{1}{kt}X_{..}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right) + Cov\left(\frac{1}{kt}X_{..}^{p}, \frac{1}{kt}X_{..}^{q}\right) + E\left(\frac{1}{kt}X_{..}^{p}\right)E\left(\frac{1}{kt}X_{..}^{q}\right)\right] - \frac{a_{pq}}{t} = \frac{1}{k-1}\cdot$$
$$\cdot\sum_{j}\left[\left(\frac{1}{t}a_{pq} + b_{pq}\right) + m_p m_q - \left(\frac{1}{kt}a_{pq} + \frac{1}{k}b_{pq}\right) - m_p m_q - \left(\frac{1}{kt}a_{pq} + \frac{1}{k}b_{pq}\right) - m_p m_q + \left(\frac{1}{kt}a_{pq} + \right.\right.$$
$$\left.\left. + \frac{1}{k}b_{pq}\right) + m_p m_q\right] - \frac{a_{pq}}{t} = \frac{1}{k-1}\sum_{j}\left(\frac{1}{t}a_{pq} + b_{pq} - \frac{1}{kt}a_{pq} - \frac{1}{k}b_{pq}\right) - \frac{a_{pq}}{t} = \frac{1}{k-1}k\frac{k-1}{k}b_{pq} +$$
$$+ \frac{1}{k-1}k\frac{k-1}{kt}a_{pq} - \frac{a_{pq}}{t} = b_{pq} + \frac{a_{pq}}{t} - \frac{a_{pq}}{t} = b_{pq}.$$

Section 4 (**Applications of semi-linear credibility theory**)

We close this paper by giving **the semi-linear hierarchical model** used in the applications chapter. Like in Jewell's hierarchical model we consider a portfolio of contracts, which can be broken up into P sectors each sector p consisting of k_p groups of contracts. Instead of estimating: $X_{p,j,t+1}$, $\mu\left(\theta_p, \theta_{p_j}\right) = E\left[X_{p,j,t+1}\middle|\theta_p, \theta_{p_j}\right]$ (the pure net risk premium of the contract (p, j)), $\nu\left(\theta_p\right) = E\left[X_{p,j,t+1}\middle|\theta_p\right]$ (the pure net risk premium of the sector p), we now estimate: $f_0\left(X_{p,j,t+1}\right)$, $\mu_0\left(\theta_p, \theta_{p_j}\right) = E\left[f_0\left(X_{p,j,t+1}\right)\middle|\theta_p, \theta_{p_j}\right]$ (the pure net risk premium of the contract (p, j)), $\nu_0\left(\theta_p\right) = E\left[f_0\left(X_{p,j,t+1}\right)\middle|\theta_p\right]$ (the pure net risk premium of the sector p), where $p = \overline{1,P}$ and $j = \overline{1,k_p}$. In semi-linear credibility theory the following class of estimators is con-

sidered: $\alpha_0 + \sum_{p=1}^{n}\sum_{q=1}^{P}\sum_{i=1}^{k_q}\sum_{r=1}^{t}\alpha_{pqir} f_p\left(X_{qir}\right)$,
where $f_1(\cdot),...,f_n(\cdot)$ are functions given in advance. Let us consider the case of one given function f_1 in order to approximate $f_0\left(X_{p,j,t+1}\right)$ or $\nu_0\left(\theta_p\right)$ and $\mu_0\left(\theta_p,\theta_{p_j}\right)$. We formulate the following theorem:

Theorem 4.1 (*Hierarchical semi-linear credibility*)

Using the same notations as introduced for the hierarchical model of Jewell and denoting $X_{pjs}^0 = f_0\left(X_{pjs}\right)$ and $X_{pjs}^1 = f_1\left(X_{pjs}\right)$ one obtains the following least squares estimates for the pure net risk premiums:

$$\hat{\nu}_0\left(\theta_p\right) = \left(m_0 - z_p m_1\right) + z_p X_{pzw}^1,$$

$$\hat{\mu}_0\left(\theta_p,\theta_{pj}\right) = \left(m_0 - z_{pj} m_1\right) + z_{pj} X_{pjw}^1 \quad (3.1)$$

where: $X_{pjw}^1 = \sum_{r=1}^{t}\frac{w_{pjr}}{w_{pj.}}X_{pjr}^1$,

$$X_{pzw}^1 = \sum_{j=1}^{k_p}\frac{z_{pj}}{z_{p.}}X_{pjw}^1,$$

$$z_{pj} = w_{pj.}d_{01} / \left[c_{11} + \left(w_{pj.} - 1\right)d_{11}\right]$$

(the credibility factor on contract level), with: $d_{01} = Cov\left(X_{pjr}^0, X_{pjr'}^1\right)$, $d_{11} = Cov\left(X_{pjr}^1, X_{pjr'}^1\right)$, $r \neq r'$, $c_{11} = Cov\left(X_{pjr}^1, X_{pjr}^1\right) = Var\left(X_{pjr}^1\right)$,
and: $z_p = z_{p.}D_{01} / \left[C_{11} + \left(z_{p.} - 1\right)D_{11}\right]$ (the credibility factor at sector level), with: $D_{01} = Cov\left(X_{pjw}^0, X_{pj'w}^1\right)$,
$D_{11} = Cov\left(X_{pjw}^1, X_{pj'w}^1\right)$, $j \neq j'$, $C_{11} = Cov\left(X_{pjw}^1, X_{pjw}^1\right) = Var\left(X_{pjw}^1\right)$.

Remark 4.1: the linear combination of 1 and the random variables X_{pjr}^1 ($p = \overline{1,P}$, $j = \overline{1,k_p}$, $r = \overline{1,t}$) closest to $f_0\left(X_{p,j,t+1}\right)$ and to $\nu_0\left(\theta_p\right)$ in the least squares sense equals $\hat{\nu}_0\left(\theta_p\right)$, and the linear combination of 1 and the random variables X_{pjr}^1 ($p = \overline{1,P}$, $j = \overline{1,k_p}$, $r = \overline{1,t}$) closest to $\mu_0\left(\theta_p,\theta_{p_j}\right)$ in the least squares sense equals $\hat{\mu}_0\left(\theta_p,\theta_{pj}\right)$.

References

[1] Goovaerts, M.J., Kaas, R., Van Herwaarden, A.E., Bauwelinckx,T.: *Insurance Series, volume 3, Efective Actuarial Methods,* University of Amsterdam, The Netherlands, 1991.

[2] Pentikäinen, T., Daykin, C.D., and Pesonen, M.: *Practical Risk Theory for Actuaries***,** Université Pierré et Marie Curie, 1990.

[3] Sundt, B.: *An Introduction to Non-Life Insurance Mathematics***,** Veroffentlichungen des Instituts für Versicherungswissenschaft der Universität Mannheim Band 28 , VVW Karlsruhe, 1984.