Optimal approximating function

Virginia Atanasiu

Abstract. In this paper we give the semi-linear credibility theory of some approximating functions and we try to demonstrate what kind of data is needed to apply the variational calculus in the semi-linear credibility model. The semi-linear credibility model is extension of the original Bühlmann model. So the communication is devoted to semi-linear credibility, where one examines functions of the random variables representing claim amounts, rather than the claim amounts themselves.

M.S.C. 2000: 62P05.
Key words: variational calculus, contracts, unbiased estimators, structure parameters, several approximating functions, semi-linear credibility theory, unique optimal function, 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 Section 1) 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 $f_1, f_2, \ldots, f_n$ of the observable variables $\underline{X}' = (X_1, X_2, \ldots, X_t)$. 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:

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

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, using the variational calculus, and try to find the unique optimal function f. It should be noted that the approximation to μ_0 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: $f_1, f_2, \ldots, f_n$. In Section 3 we will show the usefulness of the first approximation, giving an application of the unique optimal approximating function f.

Differential Geometry - Dynamical Systems, Vol.10, 2008, pp. 13-20.

1 Section 1 (The approximation to μ_0 based on prescribed approximating functions: $f_1, f_2, \ldots, f_n$)

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 $X_1, \ldots, X_t$. 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 $X_1, \ldots, X_t$ 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:

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

This expression does not depend on r. We define the following structure parameters:

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

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

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

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

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

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 $\mu_0(\theta) = E[f_0(X_{t+1})|\theta]$ 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 $z_1, z_2, \ldots, z_n$ is a solution to the linear system of equations:

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

or to the equivalent linear system of equations:

$$\sum_{p=1}^{n}(a_{pq} + tb_{pq})z_p = tb_{0q} \quad (q = \overline{1,n}). \tag{1.11}$$

Proof. We have to examine the solution of the problem:

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

Taking the derivative with respect to α_0 gives:

$$E[\mu_0(\theta)] - \sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}E[f_p(X_r)] = \alpha_0, \quad \text{or :} \quad \alpha_0 = m_0 - \sum_{p=1}^{n}\sum_{r=1}^{t}\alpha_{pr}m_p.$$

Inserting this expression for α_0 into (1.12) leads to the following problem:

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

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

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

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

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

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

2 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:

$$M = f(X_1) + \ldots + f(X_t), \tag{2.1}$$

where $f(x)=\frac{1}{t}\sum_{p=1}^{n}z_pf_p(x)+\frac{1}{t}m_0-\frac{1}{t}\sum_{p=1}^{n}z_pm_p$. Indeed, we have:

$$M=\sum_{p=1}^{n}z_p\sum_{r=1}^{t}f_p(X_r)+m_0-\sum_{p=1}^{n}z_pm_p=$$
$$=\frac{1}{t}\sum_{p=1}^{n}z_p\sum_{r=1}^{t}f_p(X_r)+m_0-\sum_{p=1}^{n}z_pm_p=$$
$$=\frac{1}{t}\sum_{p=1}^{n}z_p[f_p(X_1)+f_p(X_2)+\ldots+f_p(X_t)]+\frac{1}{t}m_0t-\frac{1}{t}\left(\sum_{p=1}^{n}z_pm_p\right)t=$$
$$=\left(\frac{1}{t}\sum_{p=1}^{n}z_pf_p(X_1)+\frac{1}{t}m_0-\frac{1}{t}\sum_{p=1}^{n}z_pm_p\right)+$$
$$+\left(\frac{1}{t}\sum_{p=1}^{n}z_pf_p(X_2)+\frac{1}{t}m_0-\frac{1}{t}\sum_{p=1}^{n}z_pm_p\right)+$$
$$+\ldots+\left(\frac{1}{t}\sum_{p=1}^{n}z_pf_p(X_t)+\frac{1}{t}m_0-\frac{1}{t}\sum_{p=1}^{n}z_pm_p\right)=f(X_1)+f(X_2)+\ldots+f(X_t),$$

as was to be proven.

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 results from the following theorem:

Theorem 2.1 (Optimal approximating function) $f(X_1)+\ldots+f(X_t)$ *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*:

$$f(X_1)+(t-1)E[f(X_2)|X_1]-E[f_0(X_2)|X_1]\equiv 0. \tag{2.2}$$

Proof. We have to solve the following minimization problem:

$$\operatorname*{Min}_{g} E\{[f_0(X_{t+1})-g(X_1)-\ldots-g(X_t)]^2\}. \tag{2.3}$$

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:

$$\varphi(\alpha)=E\{[f_0(X_{t+1})-g(X_1)-\ldots-g(X_t)]^2\},$$

or

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

So the problem (2.3) reads:

$$\operatorname*{Min}_{\alpha}\varphi(\alpha). \tag{2.5}$$

Since (2.5) is the minimum of a positive definite quadratic form, it suffices to find a solution with the first derivative equal to zero. For this reason we may conclude that: clearly for f to be optimal, $\varphi'(0) = 0$, so for every choice of h:

$$E\{[f_0(X_{t+1}) - f(X_1) - \ldots - f(X_t)][h(X_1) + \ldots + h(X_t)]\} = 0, \tag{2.6}$$

must hold. To obtain the relation (2.6), we proceed as follows: We have:

$$\begin{aligned}
&\varphi(\alpha) = E[f_0^2(X_{t+1}) + f^2(X_1) + \ldots + f^2(X_t) + \alpha^2 h^2(X_1) + \ldots + \alpha^2 h^2(X_t) - \\
&-2f_0(X_{t+1})f(X_1) - \ldots - 2f_0(X_{t+1})f(X_t) - 2f_0(X_{t+1})\alpha h(X_1) - \ldots - \\
&-2f_0(X_{t+1})\alpha h(X_t) + 2f(X_1)f(X_2) + \ldots + 2f(X_1)f(X_t) + \\
&+2f(X_1)\alpha h(X_1) + \ldots + 2f(X_1)\alpha h(X_t) + \ldots + 2f(X_t)\alpha h(X_1) + \ldots + \\
&+2f(X_t)\alpha h(X_t) + 2\alpha^2 h(X_1)h(X_2) + \ldots + 2\alpha^2 h(X_1)h(X_t) + 2\alpha^2 h(X_2)h(X_3) + \\
&+ \ldots + 2\alpha^2 h(X_2)h(X_t) + \ldots + 2\alpha^2 h(X_{t-1})h(X_t)] = E[f_0^2(X_{t+1})] + E[f^2(X_1)] + \\
&+ \ldots + E[f^2(X_t)] + \alpha^2 E[h^2(X_1)] + \ldots + \alpha^2 E[h^2(X_t)] - 2E[f_0(X_{t+1})f(X_1)] - \\
&- \ldots - 2E[f_0(X_{t+1})f(X_t)] - 2E[f_0(X_{t+1})h(X_1)]\alpha - \ldots - 2E[f_0(X_{t+1})h(X_t)]\alpha + \\
&+2E[f(X_1)f(X_2)] + \ldots + 2E[f(X_1)f(X_t)] + 2E[f(X_1)h(X_1)]\alpha + \ldots + \\
&+2E[f(X_1)h(X_t)]\alpha + \ldots + 2E[f(X_t)h(X_1)]\alpha + \ldots + 2E[f(X_t)h(X_t)]\alpha + \\
&+2\alpha^2 E[h(X_1)h(X_2)] + \ldots + 2\alpha^2 E[h(X_1)h(X_t)]\alpha + 2\alpha^2 E[h(X_2)h(X_3)] + \\
&+ \ldots + 2\alpha^2 E[h(X_2)h(X_t)] + \ldots + 2\alpha^2 E[f(X_{t-1})h(X_t)].
\end{aligned} \tag{2.7}$$

Taking the first derivative of (2.7) with respect to α, gives:

$$\begin{aligned}
&\varphi'(\alpha) = 2\alpha E[h^2(X_1)] + \ldots + 2\alpha E[h^2(X_t)] - 2E[f_0(X_{t+1})h(X_1)] - \\
&- \ldots - 2E[f_0(X_{t+1})h(X_t)] + 2E[f(X_1)h(X_1)] + \ldots + 2E[f(X_1)h(X_t)] + \\
&+ \ldots + 2E[f(X_t)h(X_1)] + \ldots + 2E[f(X_t)h(X_t)] + 4\alpha E[h(X_1)h(X_2)] + \\
&+ \ldots + 4\alpha E[h(X_1)h(X_t)] + 4\alpha E[h(X_2)h(X_3)] + \\
&+ \ldots + 4\alpha E[h(X_2)h(X_t)] + \ldots + 4\alpha E[h(X_{t-1})h(X_t)].
\end{aligned} \tag{2.8}$$

Hence, for $\alpha = 0$ we have:

$$\begin{aligned}
&\varphi'(0) = -2E[f_0(X_{t+1})h(X_1)] - \ldots - 2E[f_0(X_{t+1})h(X_t)] + \\
&+2E[f(X_1)h(X_1)] + \ldots + 2E[f(X_1)h(X_t)] + \\
&+ \ldots + 2E[f(X_t)h(X_1)] + \ldots + 2E[f(X_t)h(X_t)]
\end{aligned} \tag{2.9}$$

So the equation $\varphi'(0) = 0$ reads:

$$\begin{aligned}
E\{f_0(X_{t+1})[h(X_1) + \ldots + h(X_t)] - f(X_1)[h(X_1) + \ldots + h(X_t)] - \\
- \ldots - f(X_t)[h(X_1) + \ldots + h(X_t)]\} = 0,
\end{aligned}$$

or $E\{[f_0(X_{t+1}) - f(X_1) - \ldots - f(X_t)][h(X_1) + \ldots + h(X_t)]\} = 0$, as was to be proven (see (2.6)). This can be rewritten as:

$$E[tf_0(X_2)h(X_1) - tf(X_1)h(X_1) - t(t-1)f(X_2)h(X_1)] = 0, \tag{2.10}$$

(because $X_{t+1} \overset{(P)}{\equiv} X_2$; $X_i \overset{(P)}{\equiv} X_1$, $\forall i = \overline{2,t}$; $X_i \overset{(P)}{\equiv} X_2$, $\forall i = \overline{3,t}$ and so: $f_0(X_{t+1}) \overset{(P)}{\equiv} f_0(X_2)$; $h(X_i) \overset{(P)}{\equiv} h(X_1)$, $\forall i = \overline{2,t}$; $f(X_i) \overset{(P)}{\equiv} f(X_2)$, $\forall i = \overline{3,t}$, that is: $E\{[f_0(X_2) - f(X_1) - (t-1)f(X_2)]th(X_1)\} = 0$, so

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

that is

$$tE[f_0(X_2)h(X_1)] - tE[f(X_1)h(X_1)] - t(t-1)E[f(X_2)h(X_1)] = 0,$$

or

$$E[f_0(X_2)h(X_1)] - E[f(X_1)h(X_1)] - (t-1)E[f(X_2)h(X_1)] = 0. \tag{2.11}$$

But

$$E[f_0(X_2)h(X_1)] = E\{E[f_0(X_2)h(X_1)|X_1]\} = E\{h(X_1)E[f_0(X_2)|X_1]\} \tag{2.12}$$

and

$$E[f(X_2)h(X_1)] = E\{E[f(X_2)h(X_1)|X_1]\} = E\{h(X_1)E[f(X_2)|X_1]\}. \tag{2.13}$$

Inserting (2.12) and (2.13) in (2.11) one obtains:

$$E\{h(X_1)E[f_0(X_2)|X_1]\} - E[h(X_1)f(X_1)] - (t-1)E\{h(X_1)E[f(X_2)|X_1]\} = 0,$$

or

$$E[h(X_1)\{-f(X_1) - (t-1)E[f(X_2)|X_1] + E[f_0(X_2)|X_1]\}] = 0. \tag{2.14}$$

Because this equation has to be satisfied for every choice of the function h one obtains, the expression in brackets in (2.14) must be identical to zero, which proves (2.2).

3 Section 3 (An application of the unique optimal approximating function f)

If $X_1, \ldots, X_{t+1}$ can only take the values $0, 1, \ldots, n$ and $p_{qr} = P[X_1 = q,\ X_2 = r]$ for: $q, r = \overline{0,n}$, then $f(X_1) + \ldots + f(X_t)$ 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:

$$f(q)\sum_{r=0}^{n} p_{qr} + (t-1)\sum_{r=0}^{n} f(r)p_{qr} = \sum_{r=0}^{n} f_0(r)p_{qr}. \tag{3.1}$$

Indeed, Theorem 2.1 affirms that: "$f(X_1) + \ldots + f(X_t)$ is closest to $f_0(X_{t+1})$ (and to $\mu_0(\theta)$ in the least squares sense, if and only if f is a solution of the equation:

$$f(X_1) + (t-1)E[f(X_2)|X_1] - E[f_0(X_2)|X_1] \equiv 0.\text{''} \tag{3.2}$$

Here we have:

$$f(X_1):\begin{pmatrix} f(q) \\ P[f(X_1)=f(q)] \end{pmatrix} = \begin{pmatrix} f(q) \\ P(X_1=q) \end{pmatrix}, \quad q=\overline{0,n} \tag{3.3}$$

and

$$[f(X_2)|X_1]:\begin{pmatrix} f(r) \\ P[f(X_2)=f(r)|X_1=q] \end{pmatrix} = \begin{pmatrix} f(r) \\ P(X_2=r|X_1=q) \end{pmatrix} =$$

$$= \begin{pmatrix} f(r) \\ \dfrac{P(X_1=q, X_2=r)}{P(X_1=q)} \end{pmatrix}, \quad r=\overline{0,n} \tag{3.4}$$

But

$$p_{qr} = P[X_1=q,\ X_2=r], \quad q,r=\overline{0,n} \tag{3.5}$$

and

$$P(X_1=q) = P(X_1=q,\ \Omega) = P\left[X_1=q,\ \bigcup_{r=0}^{n}(X_2=r)\right] =$$

$$= P\left[\bigcup_{r=0}^{n}(X_1=q,\ X_2=r)\right] = \sum_{r=0}^{n} P(X_1=q,\ X_2=r) = \sum_{r=0}^{n} p_{qr}; \quad q=\overline{0,n} \tag{3.6}$$

(see the hypothesis of the application). The relations (3.5) and (3.6) imply:

$$[f(X_2)|X_1]:\begin{pmatrix} f(r) \\ \dfrac{p_{qr}}{\sum_{r=0}^{n} p_{qr}} \end{pmatrix}, \quad r=\overline{0,n} \tag{3.7}$$

(see (3.4)). So

$$E[f(X_2)|X_1] = \sum_{r=0}^{n} f(r)P(X_2=r|X_1=q) = \sum_{r=0}^{n} f(r)\frac{p_{qr}}{\sum_{r=0}^{n} p_{qr}} \tag{3.8}$$

Also, we have:

$$[f_0(X_2)|X_1]:\begin{pmatrix} f_0(r) \\ P(f_0(X_2)=f_0(r)|X_1=q] \end{pmatrix} = \begin{pmatrix} f_0(r) \\ P(X_2=r|X_1=q) \end{pmatrix} =$$

$$= \begin{pmatrix} f_0(r) \\ \dfrac{P(X_1=q, X_2=r)}{P(X_1=q)} \end{pmatrix} = \begin{pmatrix} f_0(r) \\ \dfrac{p_{qr}}{\sum_{r=0}^{n} p_{qr}} \end{pmatrix}, \quad r=\overline{0,n} \tag{3.9}$$

(see (3.5) and (3.6)). So:

$$E[f_0(X_2)|X_1] = \sum_{r=0}^{n} f_0(r)P(X_2 = r|X_1 = q) = \sum_{r=0}^{n} f_0(r)\frac{p_{qr}}{\sum_{r=0}^{n} p_{qr}}. \tag{3.10}$$

Inserting (3.3), (3.8) and (3.10) in (3.2) one obtains:

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

or

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

as was to be proven (see (3.1)).

References

[1] M.J. Goovaerts, R. Kaas, A.E. Van Herwaarden and T. Bauwelinckx, *Insurance Series*, volume 3, Efective Actuarial Methods, University of Amsterdam, The Netherlands, 1991, 187–211.

[2] T. Pentikäinen, C.D. Daykin, and M. Pesonen, *Practical Risk Theory for Actuaries*, Université Pierré et Marie Curie, 1990, 22–54.

[3] B. Sundt, *An Introduction to Non-Life Insurance Mathematics*, Veroffentlichungen des Instituts für Versicherungswissenschaft der Universität Mannheim Band 28 , VVW Karlsruhe, 1984, 189–194.

Author's address:

Virginia Atanasiu
Department of Mathematics, Academy of Economic Studies,
15-17 Calea Dorobanţilor st., Bucharest, 010552, Romania.
E-mail: virginia_atanasiu@yahoo.com