BAYESIAN INFERENCE IN CREDIBILITY THEORY

L. D'Hooge and M. J. Goovaerts
Leuven (K.U.L.)

Abstract

In 1967, Bühlmann has shown that the credibility formula was the best linearized approximation to the exact Bayesian forecast.

His result for the credibility factor $z = V_\theta E(\xi/\theta)/V(1/n \sum_{t=1}^{n} \xi_t)$ can be found back by means of some Bayesian inference techniques. Introducing a uniform prior probability density function for the credibility factor provides us with a method for estimating z, a correction term to the Bühlmann's result is obtained. It is shown how prior boundary conditions can be introduced.

1. Introduction

In the present contribution the credibility factor z will be introduced by means of some adequate Bayesian inference technique. As has already been remarked several times [1] completely different methods leading to the same expression for z. We will show that also in a Bayesian framework for estimating parameters the same expressions can be obtained under certain general conditions. However we aim to suggest a method for deriving z as a function, being the mean value of a stochastical variable z imposing some inequality constraints. Let us first recall some elements of Bayesian inference [2]. Let $f_X(x, \lambda)$ be a notation for the distribution density function of a one dimensional stochastical variable X. This distribution depends on the parameter λ. Of course the mathematical admissable range of Λ, say $\lambda \in \Lambda$, can be determined by examining the given function $f_X(x, \lambda)$. Λ is supposed to be a continuous parameter space. It is clear that for a Gaussian distribution density:

$$f_X(x, \lambda) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\{-(x-\lambda)^2/2\sigma^2\} \qquad (1)$$

where σ is a number, Λ can be defined by the inequality: $\Lambda : \{\lambda \mid -\infty < \lambda < +\infty\}$. In fact $f_X(x, \lambda)$ can be interpreted as the distribution density of X, under the condition that λ is given, it is

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

$f_X(x, \lambda)$ is the joint posterior distribution density of X with given λ. In estimating λ in a Bayesian way one ought to consider the joint distribution density of λ, given $(x_1, \ldots, x_n)$, where $(x_1, \ldots, x_n)$ is constructed from some experimental data. The question that arises is what prior density of λ has to be taken, anyhow the economical meaning of the parameter itself often determines this prior density and certainly the range of the stochastical variable X. In case of "knowing nothing" [2] Jeffreys' rule becomes:

> "If $\lambda \in [-\infty, +\infty]$ then the prior density $\Phi_\lambda(\lambda) = c$, which is an improper density; if $\lambda \in [0, +\infty]$ then one has $\Phi_\lambda(\lambda) = c/\lambda$".

Given a vector $\bar{x}^0(x_1^0, \ldots, x_n^0)$ of n independent observations the likelihood-function of the sample is given by:

$$L(\bar{x}^0/\lambda) = \prod_{j=1}^{n} f_X(x_j^0, \lambda) \tag{2}$$

and the Bayesian estimator of λ is known to be given by:

$$\text{B.E.}(\lambda) = \frac{\int_\Lambda d\lambda L(\bar{x}^0/\lambda)\, \Phi_\lambda(\lambda) \cdot \lambda}{\int_\Lambda d\lambda L(\bar{x}^0/\lambda)\ \Phi_\lambda(\lambda)} \tag{3}$$

It can be shown that under some general conditions [3], B.E.(λ) can also be obtained by the least square method. Let us introduce next some notations for describing the credibility model.

A collective of heterogenous risks, in which each member is characterized by a risk parameter θ is considered. The claim experience for a certain time period t is a random variable with known distribution:

$$P_t(x/\theta) = \text{Prob}\ (\xi_t \leq x/\theta) \qquad (t = 1, 2, \ldots) \tag{4}$$

and with density $p_t(x/\theta)$.

We will assume the ξ_t to be mutually independent.

If the individual θ were not known a prior distribution $U(\theta)$ is introduced. As is explained f.e. in [4] the forecast density of the next year's risk would be the conditional density:

$$P_{n+1}(y/x_1, \ldots, x_n) = \frac{\int P_{n+1}(y/\theta) \prod_{t=1}^{n} P_t(x_t/\theta)\ dU(\theta)}{\int \prod_{t=1}^{n} P_t(x_t/\theta)\ dU(\theta)} \tag{5}$$

The fair premium for the year $n + 1$ would then be:

$$E(\xi_{n+1}/\xi_t = x_t\,(t = 1, 2, \ldots, n)) = \int y\, P_{n+1}(y/x_1, x_2, \ldots, x_n)\, dy \quad (6)$$

The collective distribution of the premium is given by:

$$P_t(x) = E_\theta(P_t(x/\theta)) = \int P_t(x/\theta)\, dU(\theta) \quad (7)$$

The premium, not taking into account the individual experience data, would be given by:

$$E(\xi_{n+1}) = \int x\, P_{n+1}(x)\, dx = E(\xi_n) = \ldots = E(\xi_1) \quad (8)$$

with:

$$P_t(x) = P_{t'}(x) \quad \text{for} \quad \forall t, t' \quad (9)$$

2. The Approximate Credibility Formula

The fair premium for the year $n + 1$ is usually written as a linear combination of the collective mean $E(\xi)$ and the sample mean $(1/n) \sum_{t=1}^{n} x_t$ of the individual experience data

$$E(\xi_{n+1}/\bar{x}) \cong (1 - z)\, E(\xi) + z \frac{1}{n} \sum_{t=1}^{n} x_t \quad (10)$$

The factor z is called the credibility factor, and was assumed to be of the form:

$$z = \frac{n}{n + N} \quad (11)$$

Bühlmann showed that the best approximation to $E(\xi_{n+1}/\bar{x})$ in the sense of minimizing [5] [6]

$$I = E\left\{\left[E(\xi_{n+1}/\bar{x}) - a - b\frac{1}{n}\sum_{t=1}^{n}\xi_t\right]^2\right\} \quad (12)$$

is given by:

$$a = (1 - b)\, E(\xi) \quad (13)$$

$$b = \frac{V_\theta E(\xi_{n+1}/\theta)}{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)}$$

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

If course one can also look for the best approximation to $E(\xi_{n+1}/\bar{x})$ in the sense of minimizing:

$$I = E \left\{ \left[E(\xi_{n+1}/\bar{x}) - (1 - b)\, E(\xi) - b\, \frac{1}{n} \sum_{t=1}^{n} \xi_t \right]^2 \right\} \tag{15}$$

The same result is obtained as in the previous model.

The credibility factor $b = \dfrac{V_\theta E(\xi_{n+1}/\theta)}{V\left(\frac{1}{n} \sum_{t=1}^{n} \xi_t\right)}$ can still be cast into the form:

$$b = \frac{n}{n + N} \tag{16}$$

with

$$N = \frac{E_\theta V(\xi/\theta)}{V_\theta E(\xi/\theta)} \tag{17}$$

The same result can still be obtained introducing Bayesian inference techniques. Indeed, in the present case the likelihood function becomes:

$$L(b) = c\, e^{-\frac{1}{2} E\{[E(\xi_{n+1}/\bar{x}) - E(\xi) - b(1/n \sum_{t=1}^{n} \xi_t - E(\xi))]^2\}} \tag{18}$$

Because no information of $E(\xi_{n+1}/\bar{x})$ is available in the sense that an experiment would give us some values for $E(\xi_{n+1}/\xi_n, \ldots, \xi_1)$, the usual summation over the number of experiments becomes the operator E, where the integrations have to be carried out over the $(n + 1)$-dimensional space generated by all prior possible $\{\xi_1, \xi_2, \ldots, \xi_n; \theta\}$ with measure

$$dU(\theta) \prod_{t=1}^{n} p(x_t/\theta)\, dx_t \tag{19}$$

Following Jeffreys the prior distribution density turns out to be a constant. So the posterior distribution function $pdf(b)$ turns out to be:

$$pdf(b) \cong e^{-\frac{1}{2} E\{[E(\xi_{n+1}/\bar{x}) - E(\xi) - b(1/n \sum_{t=1}^{n} \xi_t - E(\xi))]^2\}} \tag{20}$$

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

The Bayesian estimation of b turns out to be:

$$\text{B.E.}(b) = \frac{\int_{-\infty}^{+\infty} b\, db\, e^{-\frac{1}{2} E\{[E(\xi_{n+1}/\bar{x}) - E(\xi) - b(1/n \sum_{t=1}^{n} \xi_t - E(\xi))]^2\}}}{\int_{-\infty}^{+\infty} db\, e^{-\frac{1}{2} E\{[E(\xi_{n+1}/\bar{x}) - E(\xi) - b(1/n \sum_{t=1}^{n} \xi_t - E(\xi))]^2\}}} \qquad (21)$$

The Gaussian integrals in the nominator and in the denominator are readely performed:

$$z(t) = \text{B. E. } (b) = \frac{V_\theta E(\xi_{n+1}/\theta)}{V\left(\frac{1}{n} \sum_{t=1}^{n} \xi_t\right)} \qquad (22)$$

3. The Approximate Credibility Formula, Introducing Uniform Priors

It is clear that the nature of the approximate formula has as a consequence:

$$0 < z(t) < 1 \qquad (23)$$

To take into account these constraints our $pdf(b)$ is constructed as a product of the likelihood function with the uniform prior $p(b)$ defined as:

$$p(b) = \begin{cases} 0 & \text{if } b < 0 \\ 1 & \text{if } 0 \leq b \leq 1 \\ 0 & \text{if } 1 < b \end{cases} \qquad (24)$$

So the Bayesian estimator of the credibility factor is given by:

$$z = \frac{\int_0^1 bdb \exp\left\{-\frac{1}{2}E\left\{\left[E(\xi_{n+1}/\bar{x}) - E(\xi) - b\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t - E(\xi)\right)\right]^2\right\}\right\}}{\int_0^1 d \exp\left\{-\frac{1}{2}E\left\{\left[E(\xi_{n+1}/\bar{x}) - E(\xi) - b\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t - E(\xi)\right)\right]^2\right\}\right\}} \qquad (25)$$

which can still be cast into the form:

$$z = \frac{\int_0^1 e^{-\frac{1}{2} b^2 V(1/n \sum_{t=1}^{n} \xi_t) + b V_\theta E(\xi_{n+1}/\theta)}\, bdb}{\int_0^1 e^{-\frac{1}{2} b^2 V(1/n \sum_{t=1}^{n} \xi_t) + b V_\theta E(\xi_{n+1}/\theta)}\, db} \qquad (26)$$

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

Making use of a result obtained by H. Bühlmann [6]

$$V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right) = \left(1-\frac{1}{n}\right) V_\theta E(\xi/\theta) + \frac{V(\xi)}{n} = V_\theta E(\xi/\theta) + E_\theta V(\xi/\theta)\cdot\frac{1}{n} \tag{27}$$

one is faced with:

$$z(n) = \frac{\int_0^1 b\,db \exp\left\{-\frac{1}{2}b^2\left[\left(1-\frac{1}{n}\right)V_\theta E(\xi/\theta) + \frac{V(\xi)}{n}\right] + bV_\theta E(\xi/\theta)\right\}}{\int_0^1 db \exp\left\{-\frac{1}{2}b^2\left[\left(1-\frac{1}{n}\right)V_\theta E(\xi/\theta) + \frac{V(\xi)}{n}\right] + bV_\theta E(\xi/\theta)\right\}} \tag{28}$$

By means of one partial integration one obtains:

$$z(n) = \frac{n}{n + \dfrac{E_\theta V(\xi/\theta)}{V_\theta E(\xi/\theta)}} - \frac{\exp\left\{\frac{1}{2} V_\theta E(\xi/\theta) - \dfrac{1}{2n} E_\theta V(\xi/\theta)\right\} - 1}{\left[V_\theta E(\xi/\theta) + E_\theta V(\xi/\theta)\dfrac{1}{n}\right] \int_0^1 e^{-\frac{1}{2}b^2[V_\theta E(\xi/\theta) + E_\theta V(\xi/\theta)1/n] + bV_\theta E(\xi_{n+1}/\theta)}\, db} \tag{29}$$

So if one neglects the correction term the result of (4) Bühlmann [6] is found back.

Of course it is possible to think about other prior densities for b, depending on n. In fact there is no mathematical argument taking an arbitrary function $z(t)$. Indeed:

$$E\left[E(\xi_{n+1}/\theta, \bar{x}) - (1 - z(t))\, E(\xi) - z(t)\,\frac{1}{n}\sum_{t=1}^{n}\xi_t\right] = 0 \tag{30}$$

for all $z(t)$.

4. The Approximate Credibility Formula, Introducing a Combination of Normal Priors

It is clear that in our expression for $z(t)$ the limit of the correction term for $n \to \infty$ doesn't approach to zero. To avoid this difficulty, if it is one, other prior densities can be introduced.

Indeed, some conditions like

$$z(0) = 0; \qquad z(\infty) = 1 \tag{31}$$

can be introduced. It is sufficient to take $\phi_b(b)$ defined by:

$$\phi_b(b) = \exp \left\{ - (b - 1)^2 \frac{n}{2} - \frac{b^2}{2n} \right\} \qquad N_b : [0, 1] \tag{32}$$

Indeed:

$$\lim_{n \to \infty} \phi_b(b) = \delta(b - 1) \qquad \lim_{n \to 0} \phi_b(bb) = \delta(b - 0) \tag{33}$$

In the present case our Bayesian estimator for $z(t)$ becomes

$$z(n) = \frac{n}{n + \frac{E_\theta V(\xi/\theta) + 1}{V_\theta E(\xi/\theta) + n}} - \frac{e^{\frac{1}{2} V_\theta E(\xi/\theta) + \frac{n}{2} - \frac{1}{2n}[E_\theta V(\xi/\theta) + 1]}}{\left[V_\theta E(\xi/\theta) + n + \frac{E_\theta V(\xi/\theta) + 1}{n} \right] \times \int_0^1 e^{-\frac{1}{2} b^2 [V_\theta E(\xi/\theta) + n] + nbV_\theta E(\xi/\theta)} \, db} \tag{34}$$

However in our opinion the question arises whether it is necessary to have $\lim_{n \to \infty} z(n) = 1$. Indeed, let us take a person who has a zero past, it is $\sum_{t=1}^{n} x_t = 0$. In that case his fair premium would be zero. It is clear that a priori b has 0 and 1 as constraints, but it is not clear that b necessary has to approach to one as $n \to \infty$. For our present $z(n)$ one has $\lim_{n \to \infty} z(n) = 1$.

In both cases one obtains:

$$z(n) = \frac{n}{n + N} - C(n) \tag{35}$$

But

$$E(\xi_{n+1}/\bar{x}) = \left(1 - \frac{n}{n + N}\right) E(\xi) + \frac{n}{n + N} \frac{1}{n} \sum_{t=1}^{n} x_t + C(n) \frac{1}{n} \sum_{t=1}^{n} x_t - C(n) E(\xi) \tag{36}$$

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

It is clear that

$$E\left\{C(n)\frac{1}{n}\sum_{t=1}^{n}\xi_t - C(n)\,E(\xi)\right\} = 0 \qquad (37)$$

So that in fact the correction term can be omitted, no differences will occur in the collective. Although taking into account the correction term one avoids a priori zero premiums. Of course up to know we have introduced in our Bayesian estimation problem an a priori variance $\sigma = 1$, in fact σ is unknown and has to be considered as a parameter. So we have to construct a new model.

5. Estimation of the Credibility Factor in Case of Unknown Variance

In fact one has to consider the problem of finding b in the linear regression model

$$E(\xi_{n+1}/\bar{x}) = E(\xi) + b\left(\frac{1}{n}\sum_{t=1}^{n} x_t - E(\xi)\right) + \varepsilon_n \qquad (38)$$

where ε_n is an error term, supposed to be normaly distributed with mean value 0 and unknown dispersion σ. So the likelihood function becomes:

$$L \cong \frac{1}{\sigma^N}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} n_i\left(E(\xi_{n+1}/x_1^{(i)},\ldots,x_n^{(i)}) - E(\xi) - b\left(\frac{\sum_{t=1}^{n} x_t^{(i)}}{n} - E(\xi)\right)\right)^2} \qquad (39)$$

Of course the n_i will be proportional (statistically) to the probability for finding $x_1^{(i)} \ldots x_n^{(i)}$. Supposing the number of experiments large (it is the number of elements in the port folio is large) one has:

$$\sum_{i=1}^{n} n_i\left(E\left(\xi_{n+1}/\bar{x}^{(i)} - E(\xi) - b\left(\frac{1}{n}\sum_{t=1}^{n} x_t^{(i)} - E(\xi)\right)\right)\right)^2 \approx$$
$$E\left[\left(E\,(\xi_{n+1}/x_1 \ldots x_n) - E(\xi) - b\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t - E(\xi)\right)\right)^2\right] \qquad (40)$$

The proportionality factor is not important, because in performing the integration he cancels.

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

Applying Jeffreys' rules one gets for the posterior probability density function:

$$pdf(b) \cong \frac{1}{\sigma^{N+1}} e^{-\frac{1}{2\sigma^2} E((E(\xi_{n+1}/\bar{x}) - E(\xi) - b(1/n \sum_{t=1}^{n} x_t - E(\xi)))^2)} \tag{41}$$

which can still be cast into the form

$$pdf(b) = \frac{1}{\sigma^{N+1}} e^{-\frac{1}{2\sigma^2}[T - 2bV_\theta E(\xi/\theta) + b^2 V(1/n \sum_{t=1}^{n} \xi_t)]} \tag{42}$$

with:

$$T = E[(E(\xi_{n+1}/\bar{\xi}) - E(\xi))^2] \tag{43}$$

Performing the integration over σ gives:

$$pdf(b) \cong \frac{1}{\left(T - 2bV_0 E(\xi/\theta) + b^2 V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)\right)^{N/2}}$$

$$\cong \frac{1}{\left(1 - 2b\frac{V_\theta E(\xi/\theta)}{T} - b^2 \frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)}{T}\right)^{N/2}} \tag{44}$$

Such that

$$B.E.(b) = \frac{\int_0^1 bdb\left[1 - 2b\frac{V_\theta E(\xi/\theta)}{T} + b^2\frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)}{T}\right]^{-N/2}}{\int_0^1 db\left[1 - 2b\frac{V_\theta E(\xi/\theta)}{T} + b^2\frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)}{T}\right]^{-N/2}} \tag{45}$$

Introducing the new variable $b = (t/N)$ one gets:

$$B.E.(b) = \frac{1}{N}\frac{\int_0^N tdt\left(1 - 2t\frac{V_\theta E(\xi/\theta)}{NT} + t^2\frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)/N}{N}\right)^{-N/2}}{\int^N dt\left(1 - 2t\frac{V_\theta E(\xi/\theta)}{NT} + t^2\frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)/N}{N}\right)^{-N/2}} \tag{46}$$

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

For large values of N one gets:

$$\text{B.E.}(b) = \frac{1}{N} \frac{\int_0^N t\,dt\; e^{t \frac{V_\theta\, E(\xi/\theta)}{T} - \frac{1}{2}t^2 \frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)/N}{T}}}{\int_0^N dt\; e^{t \frac{V_\theta\, E(\xi/\theta)}{T} - \frac{1}{2}t^2 \frac{V\left(\frac{1}{N}\sum_{t=1}^{n}\xi_t\right)/N}{T}}} \tag{47}$$

By means of one partial integration one gets:

$$\text{B.E.}(b) = \frac{NT}{N4V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)} \cdot \frac{e^{N\frac{V_\theta\, E(\xi/\theta)}{T} - \frac{1}{2}N \frac{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)}{T}} - 1}{\int_0^N dt\; e^{t\frac{V_\theta\, E(\xi/\theta)}{T} - \frac{1}{2}t^2 V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right) N}} + \frac{T}{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)} \cdot \frac{V_\theta E(\xi/\theta)}{T} \tag{48}$$

And so, in the limit for N

$$\text{B.E.}(b) = \frac{V_\theta E(\xi/\theta)}{V\left(\frac{1}{n}\sum_{t=1}^{n}\xi_t\right)} \tag{49}$$

which gives Buhlmanns' result, having introduced as a prior condition that our estimation for $z(t)$ has to satisfy some inequality constraints.

6. References

[1] van Klinken Dr. J. and Drs. G. W. Euverman, *Overzicht van de theorie en toepassingen van credibility indices,* (unpublished).

[2] Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics,* (eds) John Wiley and Sons, Inc., New York.

[3] Broeckx, F., Goovaerts, M. and J. van den Broeck, *On the prior density functions proposed by Jeffreys and Haldane, in a Bayesian framework*, Sogesci, to be published.

https://doi.org/10.1017/S0515036100009314 Published online by Cambridge University Press

[4] Jewell, W. S., "The credible distribution", *The ASTIN Bull.*, Vol. VII, Part 3, pp. 237-269.
[5] For other derivations of the credibility factor see references given by W. S. Jewell in [4].
[6] Bühlmann, H. (1970), *Mathematical methods in risk theory*, (eds) Springer-Verlag, New York.
[7] Bühlmann, H. (1971), *Credibility Procedures*, Sixth Berkeley Symposium, pp. 515-525.