Is an original paper, which describes techniques for estimating premiums for risks, containing a fraction (a part) of the variance of the risk as a loading on the net risk premium. Mathematics Subject Classification: 62P05. Keywords: the linearized credibility formula, the best risk premium – in the sense of the minimal weighted mean squared error -, variance premium.

Introduction
One approach is to consider the best risk premium (in the sense of the minimal weighted mean squared error). In the first section it is shown that it can be used as an approximation to the variance loaded premium, by truncating a series expansion. In the second section this premium is derived as an optimal estimator minimizing a suitable loss function. In the credibility theory so far the credibility results described are intended to estimate pure net risk premiums. An important question arises if one is interested in estimating the variance loaded premiums. In a top-down approach, the collective premium can be distributed proportionally to this loaded risk premium. Therefore we also consider credibility estimates for loaded risk premiums.

Theory
Replacing the loss function \((y - x)^2\) when \(y\) is estimated instead of \(x\) by a slightly more general weighted loss function \((y - x)^2 e^{by}\), one gets the following best function of \(X\) to estimate a random variable \(Y\).

Minimizing weighted mean squared error
When \(X\) and \(Y\) are two random variables, and \(Y\) must be estimated using a function \(g(X)\) of \(X\), the choice yielding the minimal weighted mean squared error \(H(Y | X) = E[Y e^{by} | X] \big/ E[e^{by} | X] \) is the best risk premium (in the sense of the minimal weighted mean squared error) for \(Y\), given \(X\): \[H(Y | X) = E[Y e^{by} | X] \big/ E[e^{by} | X] \] (1.1)

Proof. Write:

\[
E[(Y - g(X))^2 e^{by}] = E[E[(Y - g(X))^2 e^{by} | X]] = \int E[(Y - g(x))^2 e^{by} | X = x] \cdot dF_X(x)
\]

For fixed \(x\), the integrand can be written as: \(E[(Z - p)^2 e^{bz}]\), with \(p = g(x)\) and \(Z\) distributed as \(Y\), given \(X = x\) \((Z \equiv Y \mid (X = x))\). This quadratic form in \(p\) is minimized taking \(p = E[Z e^{bz}] / E[e^{bz}]\) or what is the same \(g(x) = E[Y e^{by} | X = x] / E[e^{by} | X = x]\).

Indeed: \(\phi(p)_{\text{not}} = E[(Z - p)^2 e^{bz}] = E(Z^2 e^{hz}) + p^2 E(e^{hz}) - 2 p E(Z e^{hz})\).

We have to solve the following minimization problem:

\[
\min_p \phi(p) \quad (1.2)
\]

Since (1.2) is the minimum of a positive definite quadratic form, it is suffices to find a solution with the first derivative equal to zero. Taking the first derivative with respect to \(p\), we get the equation:

\[
2 p E(e^{hz}) - 2 E(Z e^{hz}) = 0.
\]

So: \(p = E(Z e^{hz}) / E(e^{hz})\), because:
\[ \varphi^*(p) = 2E(e^{hp}) > 0 . \]

If the integrand is chosen minimal for each \( x \), the integral over all \( x \) is minimized, too.

**Remark 1.1 (the best risk premium and variance premium)**

\[ H[Y | X] = \frac{E[Y | X] + hE[Y^2 | X] + O(h^2)}{1 + hE[Y | X] + O(h^2)} = (E[Y | X] + hE[Y^2 | X] + O(h^2)). \]

\[ (1 - hE[Y | X] + O(h^2)) = E[Y | X] + h\text{Var}[Y | X] + O(h^2). \]

Indeed, we have:

\[ e^{hy} = 1 + \frac{h}{2!} Y^2 + \frac{h^n}{n!} Y^n + \ldots, \]

and so: \( Y e^{hy} \approx Y + hY^2 + O(h^2) \). Therefore, from (1.1) we get:

\[ H[Y | X] = \frac{E[e^{hy} | X]}{E[e^{hy} | X]} = \frac{E[Y + hY^2 + O(h^2)]}{1 + hE[Y | X] + O(h^2)} = E[Y | X] + h\text{Var}[Y | X] + O(h^2). \]

Also: \( (1 + Z)^{-1} = 1 + \frac{(-1)}{2!} Z + \frac{(-1)(-1-1)}{3!} Z^2 + \ldots + \frac{(-1)(-1-1)...(-1-n+1)}{n!} Z^n + \ldots, \]

if: \( |Z| < 1 \). On taking \( Z = hE[Y | X] + O(h^2) \) one obtains:

\[ (1 + hE[Y | X] + O(h^2))^{-1} \approx 1 - h \cdot E[Y | X] + O(h^2), \]

and thus:

\[ H[Y | X] = E[Y | X] + h\text{Var}[Y | X] + O(h^2) \]  \( (1.3) \)

In fact any loss function \( (y - x)^2 w(y) \) with small \( h = w'(0)/w(0) \) leads to the expression (1.3). From this remark we may conclude that to derive credibility estimates for premiums loaded with a fraction of the variance, as well as for the best risk premium, one may consider a weighted loss function. To be able to compute the loaded credibility estimates, we will need (co-) variances of squares of the observations.

1. **Credibility for the best risk premium**

Consider the original Bühlmann model. Applying (1.1) - of the previous section - to \( Y = X_{t+1} \) and \( X = X = (X_1, \ldots, X_t) \), we see that the best risk premium - in the sense of weighted mean squared error- to charge for period \( t+1 \) is:

\[ H[X_{t+1} | X] = E[X_{t+1} e^{HY_{t+1}}] / E[e^{HY_{t+1}}] \]  \( (2.1) \)

For small \( h \), the expansion (1.3) for (2.1) can be rewritten as follows:

\[ H[X_{t+1} | X] = E[X_{t+1} | X] + h\text{Var}[X_{t+1} | X] + O(h^2) \approx E[X_{t+1} | X] + h\text{Var}[X_{t+1} | X] = E[\mu(\theta) | X] + h|E[\sigma^2(\theta) | X] + \text{Var}[\mu(\theta) | X]] \]  \( (2.2) \)

Indeed, we have:

\[ E[\mu(\theta) | X] = E[E(X_{t+1} | \theta) | X] = E[E(X_{t+1} | \theta, X) | X] = E(X_{t+1} | X) \]  \( (2.3) \),

and also:

\[ \text{Var}(X_{t+1} | X) = E[X_{t+1}^2 | X] - E^2[X_{t+1} | X] \]  \( (2.4) \)

But: \[ E[\sigma^2(\theta) | X] = E[\text{Var}(X_{t+1} | \theta) | X] = E[E(X_{t+1}^2 | \theta) - E^2(X_{t+1} | \theta) | X] = E[E(X_{t+1}^2 | \theta) | X] - E[E^2(X_{t+1} | \theta) | X] \]

\[ = E[E(X_{t+1}^2 | \theta) | X] - E\left[ E^2(X_{t+1} | \theta) | X \right] = E[E(X_{t+1}^2 | \theta, X) | X] \]

\[ - E\left[ E^2(X_{t+1} | \theta) | X \right] = E(X_{t+1}^2 | X) - E\left[ E^2(X_{t+1} | \theta) | X \right] \]  \( (2.5) \),
and:

\[
\text{Var}[\mu(\theta) | X] = E[\mu^2(\theta) | X] - E^2[\mu(\theta) | X] = E[E^2(X_{r+1} | \theta) | X] - E^2(X_{r+1} | X) \tag{2.6},
\]

because from (2.3) we have: \(E[\mu(\theta) | X] = E(X_{r+1} | X)\) and so: \(E^2[\mu(\theta) | X] = E^2(X_{r+1} | X)\). Furthermore:

\[
E[\sigma^2(\theta) | X] + \text{Var}[\mu(\theta) | X] = E(X_{r+1}^2 | X) - E^2(X_{r+1} | X) + E[E^2(X_{r+1} | \theta) | X] - E^2(X_{r+1} | X) = E(X_{r+1}^2 | X) - E^2(X_{r+1} | X) = \text{Var}(X_{r+1} | X) \text{ (see (2.4))}.
\]

Therefore,

\[
H[X_{r+1} | X] = E[\mu(\theta) | X] + h \left[ E[\sigma^2(\theta) | X] + \text{Var}[\mu(\theta) | X] \right].
\]

. Thus we have (2.2).

**Remark 2.1** The first term in (2.2) denotes the expected value part, the second term gives the variance part, the last term the fluctuation part.

**Remark 2.2** Apart from the optimal credibility result (2.1) for this situation, we are interested in obtaining a linearized credibility formula for estimating \(X_{t+1}\). Therefore, we prove the following application of the formula (1.1).

**The main results of this paper**

In the following we present the main results leaving the detailed computations to the reader.

**An application of the formula (1.1)** - Linearized credibility formula for exponentially weighted squared error loss function -

The solution to the following problem:

\[
\begin{aligned}
\text{Min} \quad & E\left[ X_{r+1} - c_0 - \sum_{r=1}^{t} c_r X_r \right]^2 e^{kX_{r+1}} \\
\text{subject to} \quad & M^* = z^* X + \left\{ \frac{E^*_\theta[H(X | \theta)]}{E^*_\mu(\theta)} - z^* \right\} E^*_\mu(\theta) \tag{2.8}
\end{aligned}
\]

Here \(H[X | \theta]\) is the best risk premium for the conditional distribution of \(X\) given \(\theta\), and:

\[
dU^*(\theta) = m_h(\theta)dU(\theta)/m_h \tag{2.10}, \text{ with:}
\]

\[
m_h(\theta) = E[e^{kX} \mid \theta] \quad \text{(2.11), and}
\]

\[
m_h = E[m_h(\theta)] = \int m_h(\theta)dU(\theta) \tag{2.12}
\]

**Proof.** To have a minimum in (2.7), the derivative with respect to \(c_0\) must equal to zero, so:

\[
E[X_{r+1}e^{kX_{r+1}}] = c_0E[e^{kX_{r+1}}] + \sum_{r=1}^{t} c_r E[X_r e^{kX_{r+1}}], \text{ or: } c_0 = \frac{E[X_{r+1}e^{kX_{r+1}}]}{E[e^{kX_{r+1}}]} = \sum_{r=1}^{t} \frac{c_r E[X_r e^{kX_{r+1}}]}{E[e^{kX_{r+1}}]}, \text{ that is:}
\]

\[
c_0 = E^*_\theta[H(X | \theta)m_h(\theta)/m_h] - \sum_{r=1}^{t} c_r E^*_\mu(\theta)m_h(\theta)/m_h \tag{2.13},
\]

where \(m_h(\theta)\) and \(m_h\) are as in (2.11) and (2.12), because:

\[
E[X_{r+1}e^{kX_{r+1}}] = E[E[X_{r+1}e^{kX_{r+1}} | \theta] = E\left[ \frac{E[X_{r+1}e^{kX_{r+1}} | \theta]}{E[e^{kX_{r+1}} | \theta]} \right] E[e^{kX_{r+1}} | \theta]
\]

\[
= E[H(X | \theta)m_h(\theta)] = E^*_\theta[H(X | \theta)m_h(\theta)] \text{ (see (1.1)).}
\]

\[
E[e^{kX_{r+1}}] = E[e^{kX}] = E[E[e^{kX} | \theta] = E[m_h(\theta)] = m_h
\]

\[
E[X_r e^{kX_{r+1}}] = E[E[X_r e^{kX_{r+1}} | \theta] = E[E[X_r | \theta]E[e^{kX_{r+1}} | \theta]
\]


\[ E[E(X \mid \theta)E(e^{X_1} \mid \theta)] = E[\mu(\theta)m_h(\theta)] = E_o[\mu(\theta)m_h(\theta)]. \]

Therefore, the verification of formula (2.13) is readily performed. Using the notation:
\[ E_o^*[f(\theta)] = \int f(\theta)dU^*(\theta) = \int f(\theta)\frac{m'(\theta)}{m_h}dU(\theta) = E_o[f(\theta)m_h(\theta)/m_h] \]

(2.13) can be written as: \[ c_0 = E_o^*[H(X \mid \theta)] - \sum_{r=1}^{c} c_r E_o^*[\mu(\theta)] \]  
(2.14)

Inserting (2.14) into (2.7) the problem is reduced to:
\[ \min E \left[ \left\{ X_{r+1} - E_o^*[H(X \mid \theta)] - \sum_{r=1}^{c} c_r (X_r - E_o^*[\mu(\theta)]) \right\}^2 e^{X_{r+1}} \right] \]  
(2.15),

where: \( c = (c_1, \ldots, c_\ell) \). On taking the derivative of (2.15) with respect to \( c_q, q = 1, \ell \) and putting the result equal to zero, one obtains:
\[ E\left[ (X_{r+1} - E_o^*[H(X \mid \theta)])(X_q - E_o^*[\mu(\theta)])e^{X_{r+1}} \right] = \sum_{r=1}^{c} c_r \left( (X_r - E_o^*[\mu(\theta)])(X_q - E_o^*[\mu(\theta)])e^{X_{r+1}} \right) \]  
(2.16)

Using conditional expectations over \( \theta \) an dividing by \( m_h \), this equation can be written as:
\[ E_o^*[H(X \mid \theta)] - E_o^*[H(X \mid \theta)][\mu(\theta) - E_o^*[\mu(\theta)]m_h(\theta)/m_h] = \]
\[ \sum_{r=1}^{c} c_r E_o^*[\mu(\theta) - E_o^*[\mu(\theta)]m_h(\theta)/m_h] + c_q E_o^*[Var[X \mid \theta]]m_h(\theta)/m_h \]  
(2.17),

since:
\[ E[X_{r+1}X_qe^{X_{r+1}}|\theta] = E[E[X_{r+1}X_qe^{X_{r+1}}|\theta]e^{X_q}|\theta] = \]
\[ = E[H(X \mid \theta)m_h(\theta)\mu(\theta)] = E_o[H(X \mid \theta)m_h(\theta)\mu(\theta)], \]

from which:
\[ E[X_{r+1}X_qe^{X_{r+1}}/m_h] = E_o[H(X \mid \theta)m_h(\theta)\mu(\theta)/m_h]; \]
\[ E_o^*[\mu(\theta)]e^{X_{r+1}} = E[E[X_{r+1}E_o^*[\mu(\theta)]e^{X_{r+1}}|\theta]] = E[E_o^*[\mu(\theta)]e^{X_{r+1}}|\theta] = E_o^*[\mu(\theta)]E_o[H(X \mid \theta)m_h(\theta)] = E_o^*[\mu(\theta)]; \]
\[ H(X \mid \theta)m_h(\theta), \text{ from which: } E[X_{r+1}E_o^*[\mu(\theta)]e^{X_{r+1}}/m_h] = E_o^*[\mu(\theta)]H(X \mid \theta) \cdot \]
\[ \cdot m_h(\theta)/m_h; \]  
\[ E_o^*[H(X \mid \theta)]X_qe^{X_{r+1}} = E_o^*[H(X \mid \theta)]E[X_qe^{X_{r+1}}] = \ldots = E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta) \cdot m_h(\theta)], \]

from which: \[ E_o^*[H(X \mid \theta)]X_qe^{X_{r+1}}/m_h = \]
\[ E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta)]m_h(\theta)/m_h; \]
\[ E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta)]e^{X_{r+1}} = E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta)]e^{X_{r+1}} = \ldots = E_o^*[H(X \mid \theta) \cdot \]
\[ \cdot E_o^*[\mu(\theta)]E[m_h(\theta)] = E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta)]E[m_h(\theta)] = E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta)] \cdot \]
\[ \cdot m_h(\theta), \text{ from which: } E_o^*[H(X \mid \theta)]E_o^*[\mu(\theta)]e^{X_{r+1}}/m_h = \]
\[ E_o^*[H(X \mid \theta)]e^{X_{r+1}}; \]  
if \( r \neq q \) then we obtain:
\[ E(X_rX_qe^{X_{r+1}}) = E[E(X_rX_qe^{X_{r+1}}|\theta)] = E[E(X_r \mid \theta)E(X_q \mid \theta)e^{X_{r+1}}|\theta] = \]
\[ = E[\mu(\theta)\mu(\theta)m_h(\theta)] = E_o[\mu(\theta)\mu(\theta)m_h(\theta)], \text{ from which: } E(X_rX_qe^{X_{r+1}}/m_h) = \]
\[ = E_o[\mu(\theta)\mu(\theta)\mu(\theta)/m_h]; \]  
if \( r = q \) then we obtain:
\[ E(X_q^2e^{X_{r+1}}) = E[E(X_q^2e^{X_{r+1}}|\theta)] = E[E(X^2 \mid \theta)e^{X_{r+1}}|\theta] = E[E(X^2 \mid \theta)m_h(\theta)] = \]
\[ E_0\left[ E\left(X^2 \mid \theta \right) m_\theta(\theta) \right] \]

from which: \( E\left(X^2 e^{kX_{\text{E},t}} / m_\theta \right) = E_0\left[ E\left(X^2 \mid \theta \right) m_\theta(\theta) / m_\theta \right] \); also, we have

\[ E\left[X, E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} \right] = E\left[E\left[X, E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} \right] \mid \theta \right] = E\left[E_0^*\left[ \mu(\theta) \right] E(X, \theta) e^{kX_{\text{E},t}} \mid \theta \right] = E\left[E_0^*\left[ \mu(\theta) \right] \mu(\theta) m_\theta(\theta) \right], \]

from which: \( E\left\{ X, \cdot \right\} \)

\[ E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} / m_\theta = E_0^*\left[ \mu(\theta) \right] \mu(\theta) m_\theta(\theta) / m_\theta \), where \( r = 1 \), \( t \); similarly, we have:

\[ E\left[ E_0^*\left[ \mu(\theta) \right] X e^{kX_{\text{E},t}} / m_\theta = E_0^*\left[ \mu(\theta) \right] \mu(\theta) m_\theta(\theta) / m_\theta \right], \]

where \( q = 1 \), \( t \); we write:

\[ E\left[ E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} / m_\theta = E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} \right] = E_0^*\left[ \mu(\theta) \right] E(X, \theta) e^{kX_{\text{E},t}} \]

\[ m_\theta(\theta) = E_0^*\left[ \mu(\theta) \right] m_\theta(\theta), \]

from which: \( E\left[E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} / m_\theta \right] = E_0^*\left[ \mu(\theta) \right] m_\theta(\theta) \)

\[ \cdot \]

finally, we observe that:

\[ E\left[ X, -E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} \right] = E_0^*\left[ \mu(\theta) \right] m_\theta(\theta) \]

\[ E\left[ E_0^*\left[ \mu(\theta) \right] e^{kX_{\text{E},t}} \right] = E_0^*\left[ \mu(\theta) \right] E(X, \theta) e^{kX_{\text{E},t}} \]

\[ m_\theta(\theta) = E_0^*\left[ \mu(\theta) \right] m_\theta(\theta) \]

Defining \( \text{Cov}^* \) and \( \text{Var}^* \) by using \( E_0^* \) instead of \( E_0 \) (2.17) becomes:

\[ \text{Cov}^*\left[ H(\theta) \right] E_0^*\left[ H(\theta) \right] E_0^*\left[ H(\theta) \right] = \sum_{r=1}^{\infty} c_r \text{Cov}^* \left[ \mu(\theta) \right] + c_0 E_0^*\left[ \mu(\theta) \right] E_0^*\left[ \mu(\theta) \right] = \sum_{r=1}^{\infty} c_r \text{Var}^* \left[ \mu(\theta) \right] + c_0 E_0^*\left[ \mu(\theta) \right] E_0^*\left[ \mu(\theta) \right], \]

where \( q = 1 \).

Because of the symmetry of this system of equations in the variables: \( c_1, c_2, \ldots, c_i \), one obtains

\[ c_1 = c_2 = \ldots = c_i = 1 \]

and therefore:

\[ c = \text{Cov}^*\left[ H(\theta) \right] \text{Var}^* \left[ \mu(\theta) \right] + c_0 \text{Var}^* \left[ \mu(\theta) \right] \]

(2.18)

Inserting (2.18) into (2.14) and taking \( \bar{z}^* = ct \) as in (2.9) one obtains:

\[ c_0 = E_0^*\left[ H(\theta) \right] - c \sum_{r=1}^{\infty} E_0^*\left[ \mu(\theta) \right] = E_0^*\left[ H(\theta) \right] - \frac{\bar{z}^*}{t} E_0^*\left[ \mu(\theta) \right] = E_0^*\left[ H(\theta) \right] \]

Consequently:

\[ M^* = c_0 + \sum_{r=1}^{\infty} c_r X_r = c_0 E_0^*\left[ H(\theta) \right] - \frac{\bar{z}^*}{t} E_0^*\left[ \mu(\theta) \right] + \frac{\bar{z}^*}{t} \sum_{r=1}^{\infty} X_r = \bar{z}^* \bar{X} + \]

\[ \left( \frac{E_0^*\left[ H(\theta) \right] - \bar{z}^*}{E_0^*\left[ \mu(\theta) \right]} \right) E_0^*\left[ \mu(\theta) \right], \]

as was to be proven. Taking the limit for: \( h \to 0 \), the original Bühlmann credibility formula results (see (1.1)).
Conclusions
In this paper we have obtained the best risk premium - in the sense of weighted mean squared error - to charge for period \((t + 1)\), by truncating a series expansion. To be able to compute the loaded credibility estimates, we demonstrated the relevant (co-) variances of squares of the observations. The fact that it is based on complicated mathematics, involving conditional expectations, needs not bother the user more than it does when he applies statistical tools like SAS, GLIM, discriminant analysis, and scoring models.
Apart from the optimal credibility result (2.1) for this situation, we have presented a linearized credibility formula for exponentially weighted squared error loss function (a linearized credibility formula for estimating \(X_{t+1}\)), using the greatest accuracy theory.

References