

Bayesian approach to measuring parameter and model risk in loss ratio estimation

Hiroki Kondo and Shingo Saito

Received on February 14, 2012 / Revised on March 4, 2012

Abstract. Using an approach based on Bayesian inference, we propose a method to compute an estimate for the Value-at-Risk of an insurance loss ratio, taking both parameter and model risk into account.

Keywords. Bayesian inference, parameter risk, model risk, loss ratio

1. INTRODUCTION

Let us consider the problem of finding the Value-at-Risk (VaR) of the insurance loss ratio for a line of business, given the data for previous years. Recall that for $\alpha \in (0, 1)$ the $100\alpha\%$ VaR, or α -quantile, of a random variable X is the least $q \in \mathbb{R}$ for which $P(X \leq q) \geq \alpha$; in this article, it can always be understood as the unique q for which $P(X \leq q) = \alpha$. The method currently used by many practitioners is to compute the VaR of the normal distribution (or sometimes other more heavy-tailed distributions) with mean and variance estimated from the given historical data. One of the drawbacks of this method is that it only takes process risk into consideration, ignoring parameter and model risk. Here process risk refers to the risk caused by the stochastic nature of the model; parameter risk by the parameter estimation error; model risk by using a wrong model. In this paper, we propose a Bayesian approach towards providing an estimate for the VaR with these three types of risk all taken into account; for exactly what is used as an estimate in our Bayesian framework, see the paragraph immediately before Proposition 2.

The rest of the paper is organized as follows. In Section 2, we describe without derivation our proposed formulae for VaR estimation, and provide a numerical example for comparison purposes. Sections 3–5 present how to apply Bayesian inference to derive our formulae. Section 6 offers some concluding remarks.

2. DESCRIPTION OF OUR FORMULAE

This section describes our main results and provides a numerical example. The derivation of our formulae will be given in Sections 3–5.

2.1. DESCRIPTION OF OUR FORMULAE

Given the data $\mathbf{x} = (x_1, \dots, x_n)$ of the loss ratio for the past n years, we aim to estimate the $100\alpha\%$ VaR of the loss ratio, where $0 < \alpha < 1$ and α is usually close to 1. We assume that x_1, \dots, x_n are positive and not all equal. For notational simplicity, set

$$m_x = \frac{1}{n} \sum_{i=1}^n x_i, \quad s_x = \left(\frac{1}{n} \sum_{i=1}^n (x_i - m_x)^2 \right)^{\frac{1}{2}},$$

$$m_{\log x} = \frac{1}{n} \sum_{i=1}^n \log x_i, \quad s_{\log x} = \left(\frac{1}{n} \sum_{i=1}^n (\log x_i - m_{\log x})^2 \right)^{\frac{1}{2}}.$$

(i) Normal without parameter risk This is the simple case where the loss ratio distribution is assumed to be normal and neither parameter nor model risk is taken into account. We estimate the VaR by

$$m_x + z_\alpha s_x,$$

where z_α is the α -quantile of the standard normal distribution.

(ii) Log-normal without parameter risk This is the same as Case (i) except that the loss ratio distribution is assumed to be log-normal. We estimate the VaR by

$$\exp(m_{\log x} + z_\alpha s_{\log x}).$$

(iii) Normal with parameter risk This is the case where the normal distribution is assumed and parameter risk is incorporated; model risk is not taken into consideration. We estimate the VaR by

$$m_x + \sqrt{\frac{n+1}{n-1}} t_\alpha(n-1) s_x,$$

where $t_\alpha(n-1)$ is the α -quantile of Student's t -distribution with $n-1$ degrees of freedom.

(iv) **Log-normal with parameter risk** This is the same as Case (iii) except that the log-normal distribution is assumed. We estimate the VaR by

$$\exp\left(m_{\log x} + \sqrt{\frac{n+1}{n-1}} t_\alpha(n-1) s_{\log x}\right).$$

(v) **Normal and log-normal with parameter and model risk** This is the case where both possibilities of normal and log-normal distributions are considered, and model and parameter risk are both incorporated. Setting

$$p = \frac{s_{\log x}^{n-1} \prod_{i=1}^n x_i}{s_{\log x}^{n-1} \prod_{i=1}^n x_i + s_x^{n-1}},$$

we estimate the VaR by the solution $q > 0$ to the equation

$$pF_{t(n-1)}\left(\frac{q - m_x}{\sqrt{(n+1)/(n-1)} s_x}\right) + (1-p)F_{t(n-1)}\left(\frac{\log q - m_{\log x}}{\sqrt{(n+1)/(n-1)} s_{\log x}}\right) = \alpha,$$

where $F_{t(n-1)}$ is the cumulative distribution function of Student's t -distribution with $n-1$ degrees of freedom. Although the equation cannot be solved analytically, numerical approximation is not too difficult as the solution is always between the estimated values for Cases (iii) and (iv).

Remark 1. The above equation in q may not have a positive solution as the estimated VaR for Case (iii) may be negative. The estimated VaR for Case (i) may also be negative. No such anomalies occur when $\alpha \geq 1/2$.

2.2. NUMERICAL EXAMPLE

If our data is

$$\mathbf{x} = (0.33, 0.42, 0.37, 0.29, 0.31, 0.35, 0.42, 0.29, 0.23, 0.27)$$

with $n = 10$, then the estimated 99% VaR is the following:

	(i)	(ii)	(iii)	(iv)	(v)
99% VaR	0.466	0.494	0.513	0.571	0.558

Note that Case (iii) has a larger value than Case (i) and that Case (iv) than Case (ii); the differences can be thought of as the result of parameter risk. The estimated VaR for Case (v) lies between the values for Cases (iii) and (iv).

3. BASIC ESTIMATION: CASES (i) AND (ii)

In Case (i), we assume that the loss distribution is normal with, say, mean μ and standard deviation σ . Then the future loss ratio, a random variable with distribution $N(\mu, \sigma^2)$, has $100\alpha\%$ VaR $\mu + z_\alpha s$. Since the maximum likelihood estimators for μ and σ are m_x and s_x respectively, we estimate the $100\alpha\%$ VaR by

$$m_x + z_\alpha s_x.$$

Similarly, in Case (ii), since the maximum likelihood estimators for the log-mean and the log-standard deviation are $m_{\log x}$ and $s_{\log x}$ respectively, we estimate the $100\alpha\%$ VaR by

$$\exp(m_{\log x} + z_\alpha s_{\log x}).$$

4. INCORPORATING PARAMETER RISK: CASES (iii) AND (iv)

In what follows, with a slight abuse of notation, we will not explicitly distinguish between random variables and their realizations; also, we will always use the same letter f to denote probability density functions and probability mass functions of different random variables. For example, by $f(y)$ we can mean the value of the probability density function of a random variable y at a real number y .

In order to incorporate parameter risk, we adopt a Bayesian approach and place a distribution on the parameter space. For both normal and log-normal distributions, we use the parameter space $\Theta = \mathbb{R} \times \mathbb{R}_{>0}$ of which each element (μ, τ) represents the (log-)mean and (log-)precision, so that the likelihood functions are given by

$$f(x|\mu, \tau) = \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau(x-\mu)^2}{2}\right), \quad x \in \mathbb{R},$$

$$f(x|\mu, \tau) = \sqrt{\frac{\tau}{2\pi}} x^{-1} \exp\left(-\frac{\tau(\log x - \mu)^2}{2}\right), \quad x > 0.$$

In Bayesian inference, we first place a distribution, called the *prior distribution*, on the parameter space Θ before observing the data \mathbf{x} , and then take into account the information of \mathbf{x} to obtain the *posterior distribution* on Θ . See the proof of Proposition 1 for how to apply Bayes' theorem to obtain the posterior distribution; see [2, Section 1.3] for further details on the basics of Bayesian inference.

We will use the *normal-gamma* distribution as the prior distribution on Θ , since it is known to be a conjugate prior for both the normal and log-normal distributions. For the reader's convenience, we include the basics of the normal-gamma distribution (see also [2, Section 3.3]).

Definition 1. The *normal-gamma* distribution with parameters $\alpha, \beta, \delta > 0$ and $\gamma \in \mathbb{R}$, written $\text{NG}(\alpha, \beta, \gamma, \delta)$, is the continuous probability distribution on Θ whose density is given by

$$f(\mu, \tau) = \sqrt{\frac{\delta}{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1/2} \exp\left(-\beta\tau - \frac{\delta\tau(\mu-\gamma)^2}{2}\right).$$

Remark 2. We list the basic properties of the normal-gamma distribution.

The marginal distribution of τ is the Gamma distribution $\Gamma(\alpha, \beta)$, i.e.

$$f(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} \exp(-\beta\tau).$$

The conditional distribution of μ given τ is the normal distribution $N(\gamma, (\delta\tau)^{-1})$, i.e.

$$f(\mu|\tau) = \sqrt{\frac{\delta\tau}{2\pi}} \exp\left(-\frac{\delta\tau(\mu-\gamma)^2}{2}\right).$$

The marginal probability density function of μ is

$$\begin{aligned} f(\mu) &= \int_0^\infty f(\mu, \tau) d\tau \\ &= \sqrt{\frac{\delta}{2\pi\beta}} \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\alpha)} \left(1 + \frac{\delta(\mu - \gamma)^2}{2\beta}\right)^{-(\alpha + \frac{1}{2})}; \end{aligned}$$

if we put $\mu' = (\mu - \gamma)/\sqrt{\beta/\alpha\delta}$, we have

$$\begin{aligned} f(\mu') &= \sqrt{\frac{\beta}{\alpha\delta}} f(\mu) \\ &= \sqrt{\frac{1}{2\pi\alpha}} \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\alpha)} \left(1 + \frac{\delta(\mu - \gamma)^2}{2\beta}\right)^{-(\alpha + \frac{1}{2})} \\ &= \sqrt{\frac{1}{2\pi\alpha}} \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\alpha)} \left(1 + \frac{\mu'^2}{2\alpha}\right)^{-(\alpha + \frac{1}{2})}, \end{aligned}$$

which means that μ' has Student's t -distribution with 2α degrees of freedom. We therefore say that the marginal distribution of τ is $\gamma + \sqrt{\beta/\alpha\delta} t(2\alpha)$.

Suppose for now that we use $\text{NG}(\alpha, \beta, \gamma, \delta)$ as the prior distribution on Θ in Case (iii); the parameters will be chosen later. The following proposition shows that the normal-gamma distribution is indeed a conjugate prior for the normal distribution:

Proposition 1. *The posterior distribution on Θ is the normal-gamma distribution $\text{NG}(\alpha', \beta', \gamma', \delta')$, where*

$$\begin{aligned} \alpha' &= \alpha + \frac{n}{2}, & \beta' &= \beta + \frac{ns_x^2}{2} + \frac{\delta n(m_x - \gamma)^2}{2(\delta + n)}, \\ \gamma' &= \frac{\gamma\delta + nm_x}{\delta + n}, & \delta' &= \delta + n. \end{aligned}$$

Proof. Bayes' theorem and a simple computation show that, with proportionality constants independent of (μ, τ) , we have

$$\begin{aligned} f(\mu, \tau | \mathbf{x}) &\propto f(\mu, \tau) f(\mathbf{x} | \mu, \tau) = f(\mu, \tau) \prod_{i=1}^n f(x_i | \mu, \tau) \\ &\propto \tau^{\alpha - \frac{1}{2}} \exp\left(-\beta\tau - \frac{\delta\tau(\mu - \gamma)^2}{2}\right) \\ &\quad \times \prod_{i=1}^n \sqrt{\tau} \exp\left(-\frac{\tau(x_i - \mu)^2}{2}\right) \\ &= \tau^{\alpha' - \frac{1}{2}} \exp\left(-\beta'\tau - \frac{\delta'\tau(\mu - \gamma')^2}{2}\right), \end{aligned}$$

as desired. \square

Consider a random variable y with the property that x_1, \dots, x_n, y are conditionally independent and identically distributed given (μ, τ) . The random variable y can be thought of as the future loss ratio. The distribution of y given \mathbf{x} is called the *posterior predictive distribution*. In our Bayesian framework, we estimate the VaR of the loss ratio by the VaR of the posterior predictive distribution.

Proposition 2. *The posterior predictive distribution is*

$$\gamma' + \sqrt{\frac{\beta'(\delta' + 1)}{\alpha'\delta'}} t(2\alpha').$$

Proof. We have

$$\begin{aligned} f(y | \mathbf{x}) &= \iint_{\Theta} f(y | \mu, \tau) f(\mu, \tau | \mathbf{x}) d\mu d\tau \\ &= \iint_{\Theta} \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau(y - \mu)^2}{2}\right) \sqrt{\frac{\delta'}{2\pi}} \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \tau^{\alpha' - \frac{1}{2}} \\ &\quad \times \exp\left(-\beta'\tau - \frac{\delta'\tau(\mu - \gamma')^2}{2}\right) d\mu d\tau \\ &= \sqrt{\frac{\delta'}{2\pi\beta'(\delta' + 1)}} \frac{\Gamma(\alpha' + \frac{1}{2})}{\Gamma(\alpha')} \\ &\quad \times \left(1 + \frac{\delta'(y - \gamma')^2}{2\beta'(\delta' + 1)}\right)^{-(\alpha' + \frac{1}{2})}, \end{aligned}$$

from which the proposition follows as in Remark 2. \square

We now set the parameters as $\alpha = -1/2$ and $\beta = \gamma = \delta = 0$. This means that we use the improper prior

$$f(\mu, \tau) \propto \tau^{-1} \quad (*)$$

as a non-informative prior on Θ as suggested in [2, Section 3.2], where the prior is given in terms of the variance τ^{-1} rather than the precision τ . Then Proposition 1 shows that the posterior distribution on Θ is

$$\text{NG}\left(\frac{n-1}{2}, \frac{ns_x^2}{2}, m_x, n\right);$$

Proposition 2 shows that the posterior predictive distribution is

$$m_x + \sqrt{\frac{n+1}{n-1}} t(n-1)s_x.$$

It follows that the estimated $100\alpha\%$ VaR is

$$m_x + \sqrt{\frac{n+1}{n-1}} t_\alpha(n-1)s_x.$$

Case (iv) can be dealt with in exactly the same manner. The posterior distribution on Θ is

$$\text{NG}\left(\frac{n-1}{2}, \frac{ns_{\log x}^2}{2}, m_{\log x}, n\right),$$

and the posterior predictive distribution is

$$\exp\left(m_{\log x} + \sqrt{\frac{n+1}{n-1}} t(n-1)s_{\log x}\right),$$

which is the distribution of a positive random variable whose logarithm has the distribution

$$m_{\log x} + \sqrt{\frac{n+1}{n-1}} t(n-1)s_{\log x}.$$

It follows that the estimated $100\alpha\%$ VaR is

$$\exp\left(m_{\log x} + \sqrt{\frac{n+1}{n-1}} t_\alpha(n-1)s_{\log x}\right).$$

5. INCORPORATING MODEL RISK: CASE (v) and

Our method of incorporating model risk was inspired by Cairns [1], though the context is rather different. Incorporating model risk requires that we place a distribution on the model space, as well as on the parameter space Θ . In our setting, the model space is $\mathcal{M} = \{N, LN\}$, where N and LN respectively denote the models in which the loss ratio distribution is normal and log-normal. We choose to adopt the uniform distribution $f(N) = f(LN) = 1/2$ as the prior distribution on \mathcal{M} . It is also possible to use any non-uniform distribution, and the following computation will not be any more difficult; see Remark 3.

When finding the posterior distribution on \mathcal{M} , we need to calculate

$$f(\mathbf{x}|N) = \iint_{\Theta} f(\mathbf{x}|N, \mu, \tau) f(\mu, \tau|N) d\mu d\tau,$$

$$f(\mathbf{x}|LN) = \iint_{\Theta} f(\mathbf{x}|LN, \mu, \tau) f(\mu, \tau|LN) d\mu d\tau.$$

However, since we have decided to use an improper prior distribution on Θ as given in Equation (*), the density functions $f(\mu, \tau|N)$ and $f(\mu, \tau|LN)$ are not well defined, and so neither are $f(\mathbf{x}|N)$ and $f(\mathbf{x}|LN)$. We can overcome this problem by making the most of the fact that we are using the same prior distribution.

Suppose for the moment that we use the same *proper* prior distribution $NG(\alpha, \beta, \gamma, \delta)$ for the normal and log-normal distributions. Then if we put

$$\alpha' = \alpha + \frac{n}{2},$$

$$\beta' = \beta + \frac{ns_x^2}{2} + \frac{\delta n(m_x - \gamma)^2}{2(\delta + n)},$$

$$\beta'' = \beta + \frac{ns_{\log x}^2}{2} + \frac{\delta n(m_{\log x} - \gamma)^2}{2(\delta + n)},$$

$$\gamma' = \frac{\gamma\delta + nm_x}{\delta + n}, \quad \gamma'' = \frac{\gamma\delta + nm_{\log x}}{\delta + n}, \quad \delta' = \delta + n,$$

we have

$$f(\mathbf{x}|N)$$

$$= \iint_{\Theta} f(\mathbf{x}|N, \mu, \tau) f(\mu, \tau|N) d\mu d\tau$$

$$= \iint_{\Theta} \prod_{i=1}^n \left(\sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau(x_i - \mu)^2}{2}\right) \right)$$

$$\times \sqrt{\frac{\delta}{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha - \frac{1}{2}} \exp\left(-\beta\tau - \frac{\delta\tau(\mu - \gamma)^2}{2}\right) d\mu d\tau$$

$$= (2\pi)^{-\frac{n}{2}} \sqrt{\frac{\delta}{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)}$$

$$\times \iint_{\Theta} \tau^{\alpha' - \frac{1}{2}} \exp\left(-\beta'\tau - \frac{\delta'\tau(\mu - \gamma')^2}{2}\right) d\mu d\tau$$

$$= (2\pi)^{-\frac{n}{2}} \sqrt{\frac{\delta}{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)} \sqrt{\frac{2\pi}{\delta'}} \frac{\Gamma(\alpha')}{\beta'^{\alpha'}}$$

$$f(\mathbf{x}|LN)$$

$$= \iint_{\Theta} f(\mathbf{x}|LN, \mu, \tau) f(\mu, \tau|LN) d\mu d\tau$$

$$= \iint_{\Theta} \prod_{i=1}^n \left(\sqrt{\frac{\tau}{2\pi}} x_i^{-1} \exp\left(-\frac{\tau(\log x_i - \mu)^2}{2}\right) \right)$$

$$\times \sqrt{\frac{\delta}{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha - \frac{1}{2}} \exp\left(-\beta\tau - \frac{\delta\tau(\mu - \gamma)^2}{2}\right) d\mu d\tau$$

$$= (2\pi)^{-\frac{n}{2}} \sqrt{\frac{\delta}{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^n x_i^{-1} \sqrt{\frac{2\pi}{\delta'}} \frac{\Gamma(\alpha')}{\beta'^{\alpha'}}.$$

It follows that

$$\frac{f(N|\mathbf{x})}{f(LN|\mathbf{x})} = \frac{f(N)f(\mathbf{x}|N)}{f(LN)f(\mathbf{x}|LN)} = \frac{\beta''^{\alpha'}}{\beta'^{\alpha'}} \prod_{i=1}^n x_i; \quad (**)$$

we can substitute $\alpha = -1/2$ and $\beta = \gamma = \delta = 0$ in the most right-hand side to obtain

$$\frac{\beta''^{\alpha'}}{\beta'^{\alpha'}} \prod_{i=1}^n x_i = \frac{s_{\log x}^{n-1}}{s_x^{n-1}} \prod_{i=1}^n x_i.$$

In the light of the above computation, we set the posterior distribution on \mathcal{M} to be $f(N|\mathbf{x}) = p$ and $f(LN|\mathbf{x}) = 1 - p$, where

$$p = \frac{s_{\log x}^{n-1} \prod_{i=1}^n x_i}{s_{\log x}^{n-1} \prod_{i=1}^n x_i + s_x^{n-1}}.$$

Then if y is a random variable such that x_1, \dots, x_n are conditionally independent and identically distributed given a model in \mathcal{M} and $(\mu, \tau) \in \Theta$, we have for $q > 0$

$$P(y \leq q | \mathbf{x})$$

$$= P(y \leq q | \mathbf{x}, N)P(N|\mathbf{x}) + P(y \leq q | \mathbf{x}, LN)P(LN|\mathbf{x})$$

$$= pP(y \leq q | \mathbf{x}, N) + (1 - p)P(y \leq q | \mathbf{x}, LN).$$

It follows from the arguments in Section 4 that the posterior predictive distribution is the mixture of

$$m_x + \sqrt{\frac{n+1}{n-1}} t(n-1)s_x$$

and

$$\exp\left(m_{\log x} + \sqrt{\frac{n+1}{n-1}} t(n-1)s_{\log x}\right)$$

with weights p and $1 - p$ respectively. Hence the estimated 100% VaR is the solution $q > 0$ to the equation

$$pF_{t(n-1)}\left(\frac{q - m_x}{\sqrt{(n+1)/(n-1)}s_x}\right)$$

$$+ (1 - p)F_{t(n-1)}\left(\frac{\log q - m_{\log x}}{\sqrt{(n+1)/(n-1)}s_{\log x}}\right) = \alpha.$$

Remark 3. If we decide to use a non-uniform prior $f(N) = p_0$ and $f(LN) = 1 - p_0$, all we need to do is replace the definition of p by

$$p = \frac{p_0 s_{\log x}^{n-1} \prod_{i=1}^n x_i}{p_0 s_{\log x}^{n-1} \prod_{i=1}^n x_i + (1 - p_0) s_x^{n-1}}.$$

If $p_0 = 1$, then $p = 1$ and the estimated VaR is the same as the one for Case (iii); similarly, the case $p_0 = 0$ yields $p = 0$ and corresponds to Case (iv).

6. CONCLUDING REMARKS

We have proposed a Bayesian approach to measuring model and parameter risk in a typical loss ratio estimation problem based on historical data of previous years.

Our method applies to general population distribution estimation, provided that the population distribution may be assumed to be either normal or log-normal. In cases where other distributions should be taken into consideration, parameter risk can often be measured in the same manner, but model risk cannot; this is because our method relies essentially on the fact that we can cancel out the infinite parts produced by the use of improper priors for the normal and log-normal distributions (see Equation (**)). We may need to find appropriate proper priors to handle other distributions.

It should also be noted that in Cases (iv) and (v), the posterior predictive distribution has infinite mean, which prevents us from using Tail VaR instead of VaR in these cases.

ACKNOWLEDGEMENTS

The authors obtained these results while conducting joint research between Nisshin Fire & Marine Insurance Co., Ltd. and Kyushu University. They wish to express their sincere gratitude to the other group members for useful discussions and helpful comments. They are also grateful to Hirokazu Iwasawa and Koji Aihara for their invaluable suggestions that significantly improved the present paper.

REFERENCES

- [1] Cairns, A. J. G.: A discussion of parameter and model uncertainty in insurance, *Insurance Math. Econom.* **27** (2000) 313–330.
- [2] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B.: *Bayesian Data Analysis*, second edition, Texts in Statistical Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2004.

Hiroki Kondo
Nisshin Fire & Marine Insurance Company, Limited, 2–3
Kanda-Surugadai, Chiyoda-ku, Tokyo, 101-8329, Japan
E-mail: hiroki.kondou(at)nisshinfire.co.jp

Shingo Saito
Institute of Mathematics for Industry, Kyushu University,
744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
E-mail: ssaito(at)imi.kyushu-u.ac.jp