Bayesian Approach to Measuring Parameter and Model Risk in Loss Ratio Estimation

Shingo SAITO

Faculty of Arts and Science, Kyushu University

7 November 2013

Joint work with Hiroki KONDO

(Nisshin Fire & Marine Insurance Company, Limited (till 31 Oct.)
American Home Assurance Company (since 1 Nov.)
Graduate School of Mathematics, Kyushu University)

The problem considered in this presentation

Given data

$\boldsymbol{x} = (x_1, \dots, x_n)$: annual **loss ratios** $\left(:= \dfrac{\text{total losses}}{\text{total premiums}} \right)$ in the past n years.

Example: $x_1 = 0.33$, $x_2 = 0.42$, $\dots$.

Aim

Estimate the **Value at Risk** (**VaR**) of the future annual loss ratio y.

For $0 < \alpha < 1$ (e.g. $\alpha = 0.99$), the $100\alpha\%$**VaR** is the α-quantile of y, i.e. the value y_0 for which

$$\mathrm{Prob}(y \le y_0) = \alpha.$$

[Assumption] $\boldsymbol{x_1, \dots, x_n, y}$ are i.i.d.

x: any one of $x_1, \dots, x_n$.

$f(z)$: the probability density function of a random variable z.

Conventional method

Typical conventional method

- Assume that x is **normally distributed**: $x \sim N(\mu, \sigma^2)$.
- Find the **maximum likelihood estimators** of μ, σ^2:

$$\widehat{\mu} = m_x = \frac{x_1 + \cdots + x_n}{n} \qquad \text{(sample mean)},$$

$$\widehat{\sigma}^2 = s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)^2 \qquad \text{([biased] sample variance)}.$$

- Estimate the $100\alpha\%$VaR of the future annual loss ratio y by

$$\widehat{\mu} + z_\alpha \widehat{\sigma} = \boldsymbol{m_x + z_\alpha s_x} \qquad \text{(Equation (1))}.$$

Here z_α is the α-quantile of the standard normal distribution:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_\alpha} \exp\left(-\frac{z^2}{2}\right) dz = \alpha.$$

Drawback of the conventional method: three types of risk

Risk being taken into account

(A) **Process risk**: caused by the stochastic nature of the model.

Risk NOT being taken into account

(B) **Parameter risk**: caused by the **parameter estimation error**.

(C) **Model risk**: caused by using a **wrong model** (distribution).

$\longrightarrow$ We employ **Bayesian inference** to take (B) and (C) into account.

Incorporating parameter risk

Likelihood: $x|\mu,\tau \sim N(\mu,\tau^{-1})$ $\big((\mu,\tau)\in\mathbb{R}\times\mathbb{R}_{>0}\big)$.

Prior distribution: $(\mu,\tau)\sim \mathrm{NG}(\alpha,\beta,\gamma,\delta)$ $(\alpha,\beta,\delta>0,\ \gamma\in\mathbb{R})$.
The **normal-gamma distribution** $\mathrm{NG}(\alpha,\beta,\gamma,\delta)$ is characterised by

- $\boldsymbol{\tau\sim\Gamma(\alpha,\beta)}$, i.e. $f(\tau)=\dfrac{\beta^\alpha}{\Gamma(\alpha)}\tau^{\alpha-1}\exp(-\beta\tau)$, and
- $\boldsymbol{\mu|\tau\sim N\left(\gamma,\frac{1}{\delta\tau}\right)}$, i.e. $f(\mu|\tau)=\sqrt{\dfrac{\delta\tau}{2\pi}}\exp\left(-\dfrac{\delta\tau(\mu-\gamma)^2}{2}\right)$.

$\longrightarrow$ **Posterior distribution**: $(\mu,\tau)|\boldsymbol{x}\sim\mathrm{NG}(\alpha',\beta',\gamma',\delta')$.
Here α', β', γ', δ' are functions of α, β, γ, δ and the data $\boldsymbol{x}$.
(The normal-gamma distribution is the **conjugate prior**.)

Incorporating parameter risk

Likelihood: $x|\mu,\tau \sim N(\mu,\tau^{-1})$.
Prior: $(\mu,\tau) \sim \mathrm{NG}(\alpha,\beta,\gamma,\delta)$ $(\alpha,\beta,\delta>0,\ \gamma\in\mathbb{R})$.
$\longrightarrow$ **Posterior**: $(\mu,\tau)|\boldsymbol{x} \sim \mathrm{NG}(\alpha',\beta',\gamma',\delta')$.

Use the **improper prior** $f(\mu,\tau)\propto\tau^{-1}$, i.e. $(\alpha,\beta,\gamma,\delta)=\left(-\frac{1}{2},0,0,0\right)$.

The posterior distribution is proper: $(\mu,\tau)\sim\mathrm{NG}\left(\frac{n-1}{2},\frac{ns_x^2}{2},m_x,n\right)$.

Estimator of VaR of $\boldsymbol{y}$: VaR of $y|\boldsymbol{x}$ (Equation (2)).
The distribution of $y|\boldsymbol{x}$ (**posterior predictive distribution**) turns out to be (a linear transformation of) Student's t-distribution.

Towards incorporating model risk

Want to incorporate model risk. $\longrightarrow$ Need another distribution.
$\longrightarrow$ The **log-normal distribution** turns out to be convenient.

Conventional method (corresponding to Equation (1))

- **Distribution**: Assume $x \sim LN(\mu, \sigma^2)$, i.e. $\log x|\mu,\tau \sim N(\mu, \sigma^2)$.
- **Parameter estimation**:
  Use MLE to get $\widehat{\mu} = m_{\log x}$ and $\widehat{\sigma}^2 = s^2_{\log x}$.
- **Estimator of VaR of $\boldsymbol{y}$**: VaR of $LN(\widehat{\mu}, \widehat{\sigma}^2)$ (Equation (3)).

Incorporating parameter risk (corresponding to Equation (2))

- **Likelihood**: $x|\mu,\tau \sim LN(\mu, \tau^{-1})$.
- **Prior**: $f(\mu,\tau) \propto \tau^{-1}$ (**same as before**).
- **Posterior**: $(\mu,\tau)|\boldsymbol{x} \sim \mathrm{NG}\left(\frac{n-1}{2}, \frac{n s^2_{\log x}}{2}, m_{\log x}, n\right)$.
- **Estimator of VaR of $\boldsymbol{y}$**: VaR of $y|\boldsymbol{x}$ (Equation (4)).

Incorporating model risk

- **Parameter space**: $\{\mathrm{N}, \mathrm{LN}\} \times \mathbb{R} \times \mathbb{R}_{>0} \left(\ni (M, \mu, \tau)\right)$ (N: normal, LN: log-normal).
- **Likelihood**: $x|M, \mu, \tau \sim \begin{cases} N(\mu, \tau^{-1}) & \text{if } M = \mathrm{N}; \\ LN(\mu, \tau^{-1}) & \text{if } M = \mathrm{LN}. \end{cases}$
- **Prior**: $f(M, \mu, \tau) \propto \tau^{-1}$ (**possible** because we use N and LN).
- **Posterior**:
  - $f(\mathrm{N}\,|\boldsymbol{x}) = p$ and $(\mu, \tau)|(\boldsymbol{x}, \mathrm{N}) \sim \mathrm{NG}(*, *, *, *)$;
  - $f(\mathrm{LN}\,|\boldsymbol{x}) = 1 - p$ and $(\mu, \tau)|(\boldsymbol{x}, \mathrm{LN}) \sim \mathrm{NG}(*, *, *, *)$,

  where p and the parameters $*$ are (unspecified) functions of $\boldsymbol{x}$.
- **Estimator of VaR of $\boldsymbol{y}$**: VaR of $y|\boldsymbol{x}$ (Equation (5)).

Numerical example and future work

Numerical example

$n = 10$, $\boldsymbol{x} = (0.33, 0.42, 0.37, 0.29, 0.31, 0.35, 0.42, 0.29, 0.23, 0.27)$.

	(1)	(2)	(3)	(4)	(5)
distribution	N	N	LN	LN	N/LN
process risk	✓	✓	✓	✓	✓
parameter risk	✗	✓	✗	✓	✓
model risk	✗	✗	✗	✗	✓
estimated 99% VaR	**0.466**	**0.513**	0.494	0.571	**0.558**

Future work

Extend the method to allow for other distributions.
$\longrightarrow$ Easy for parameter risk; difficult for model risk.