Combining internal data with scenario analysis

E. KARAM & F. PLANCHET

March, 2013

Université de Lyon, Université Lyon 1,
ISFA,
laboratoire SAF EA2429,
69366 Lyon France

Abstract

A Bayesian inference approach offers a methodical concept that combines internal data with experts' opinions. Joining these two elements with precision is certainly one of the challenges in operational risk. In this working paper, we are interested in applying a Bayesian inference technique in a robust manner to be able to estimate a capital requirement that best approaches the reality.

Keywords: Bayesian Inference, Operational Risk, MCMC.

1 Introduction

Under the new regulations of Basel II and Solvency II, to be able to estimate their aggregate operational risk capital charge, many financial institutions have adopted a Loss Distribution Approach (LDA), consisting of a frequency and a severity distribution, based on its own internal losses. Yet, basing our models on historical losses only might not be the perfect robust approach since no future attention is being taken into consideration which can generate a biased capital charge, defined as the 0.01 % quantile of the loss distribution, facing reality. On the other hand, adding scenario analysis given by the experts assume the adoption of future losses.

The main idea in this article is the following: A Bayesian inference approach offers a methodical concept that combines internal data with scenario analysis. We are searching first to integrate the information generated by the experts with our internal database; by working with conjugate family distributions, we determine a prior estimate. This estimate is then modified by integrating internal observations and experts' opinion leading to a posterior estimate; risk measures are then calculated from this posterior knowledge. See Shevchenko [2011] for more on the subject.

On the second half, we use Jeffreys non-informative prior and apply Monte Carlo Markov Chain with Metropolis Hastings algorithm, thus removing the conjugate family restrictions and developing, as the article shows, a generalized application to set up a capital allocation. For a good introduction to non-informative prior distributions and MCMC see Robert [2007].
Combining these different data sources for model estimation is certainly one of the main challenges in operational risk. More on Bayesian Inference techniques could be found in Berger [1985].

2 Bayesian techniques in combining two data sources: Conjugate prior

In our study, our data related to retail banking business line and external fraud event type is of size 279, collected in \$ over 4 years. The data follows the $Poisson(5.8)$ as a frequency distribution, and $\mathcal{LN}(\mu = 6.7, \sigma = 1.67)$ as the severity distribution.

Applying Monte Carlo simulation (cf. Frachot *et al.* [2001]), with $\lambda_{ID} = 5.8$, $\mu_{ID} = 6.7$, and $\sigma_{ID} = 1.67$, we obtained a Value-at-Risk of $VaR_{ID} = 1,162,215.00$ at 99.9%, using internal losses only.
On the other hand, working with the scenario analysis, our experts gave us their assumptions for the frequency parameter λ. As for the severity, our experts represent a histogram reflecting the probability that a loss is in an interval of losses (see table 1 below).

Losses Interval in \$	Expert Opinion
$[0, 5000[$	65%
$[5000, 20000[$	19%
$[20000, 50000[$	10%
$[50000, 100000[$	3.5%
$[100000, 250000[$	1.5%
$[250000, 400000[$	0.7%
≥ 400000	0.3%

Table 1: Scenario analysis

If we consider our severity distribution being Lognormal with paramters μ and σ^2, the objective is to find the parameters $(\mu_{exp}, \sigma_{exp})$ that adjust our histogram in a way to approach as much as possible the theoretical lognormal distribution. For this we can use chi-squared statistic that allows us to find (μ, σ) that minimize the chi-squared distance:

$$\tilde{T} = \sum_{i=1}^{n} \frac{(E_i - O_i)^2}{E_i},$$

where E_i and O_i are respectively the empirical and theoretical probability.

Our experts provided $\lambda = 2$, and by applying chi-squared, we obtained our lognormal parameters: $(\mu = 7.8, \sigma = 1.99)$ with the $VaR(99.9\%) = 6,592,086.00$.

We can easily check the high spread between the two values which can cause a problem in allocating the capital requirement. In the next sections, we will apply the Bayesian inference techniques, thus joining our internal observations with the experts opinion.

2.1 Modelling Frequency distribution: The Poisson Model

We are going to work with the Poisson and Lognormal distributions since they are the most used distributions in Operational Risk (cf. Shevchenko [2011]).
Consider the annual number of events N for a risk in a bank modelled as a random variable from the Poisson distribution $\mathcal{P}(\lambda)$, where Λ is considered as a random variable with the prior distribution $Gamma(a, b)$. So we have: $\mathbb{P}(N = n) = e^{-\lambda}\frac{\lambda^n}{n!}$, and λ has a prior density:

$$\Pi(\Lambda = \lambda) = \frac{(\frac{\lambda}{b})^{a-1}}{\Gamma(a)b} e^{\frac{-\lambda}{b}}, \quad \lambda > 0,\ a > 0,\ b > 0$$

As for the likelihood function, given the assumption that $n_1, n_2, ..., n_T$ are independent, for $N = n$:

$$h(n|\lambda) = \prod_{i=1}^{T} e^{-\lambda}\frac{\lambda^{n_i}}{n_i!},$$

where n is the number of historical losses and n_i is the number of losses in month i.

Thus, the posterior density would be: $\Pi(\lambda|N = n) = \frac{h(n|\lambda)\Pi(\lambda)}{h(n)}$, but since $h(n)$ plays the role of a normalizing constant, $\Pi(\lambda|N = n)$ could be rewritten as:

$$\Pi(\lambda|N = n) \propto h(n|\lambda)\Pi(\lambda) \propto \frac{(\frac{\lambda}{b})^{a-1}}{\Gamma(a)b} e^{\frac{-\lambda}{b}} \prod_{i=1}^{T} e^{-\lambda}\frac{\lambda^{n_i}}{n_i!} \propto \frac{\lambda}{b}^{\sum_{i=1}^{T} n_i + a - 1} e^{-\lambda(T + \frac{1}{b})} \propto \lambda^{a_T - 1} e^{-\frac{\lambda}{b_T}}.$$

Which is $Gamma(a_T, b_T)$, i.e. the same as the prior distribution with $a_T = \sum_{i=1}^{T} n_i + a$ and $b_T = \frac{b}{(1 + Tb)}$
So we have:

$$\mathbb{E}(\lambda|N = n) = a_T b_T = \omega\bar{N} + (1 - \omega)(ab) = \omega\bar{N} + (1 - \omega)\mathbb{E}(\Lambda), \text{ with } \omega = \frac{n}{n + \frac{1}{b}}$$

To apply this, and since the only unknown parameter is λ that is estimated by our experts with, $\mathbb{E}(\lambda) = 2$.

The experts may estimate the expected number of events, but cannot be certain of the estimate. Our experts specify $\mathbb{E}[\lambda]$ and an uncertainty that the "true" λ for next month is within the interval $[a_0, b_0] = [0.5, 8]$ with a probability $p = 0.7$ that $\mathbb{P}(a_0 \leq \lambda \leq b_0) = p$, then we obtain the below equations:

$$\begin{aligned}
\mathbb{E}[\lambda] &= a \times b = 2 \\
\mathbb{P}(a_0 \leq \lambda \leq b_0) &= \int_{a_0}^{b_0} \pi(\lambda|a,b)d\lambda = F_{a,b}^{(G)}(b_0) - F_{a,b}^{(G)}(a_0) = 0.7
\end{aligned}$$

Where $F_{a,b}^{(G)}(.)$ is the $Gamma(a, b)$ cumulative distribution function.

Solving the above equations would give us the prior distribution parameters $\lambda \hookrightarrow Gamma(a = 0.79, b = 2.52)$, and by using the formulas stated, we obtain: $a_T = 279.8$ and $b_T = 0.02$ as our posterior parameters distribution. At the end, we calculate a $VaR(99.9\%) = 1,117,821.00$ using Monte Carlo simulation:

1- Using the estimated Posterior $Gamma(a_T, b_T)$ distribution, generate a value for λ;

2- Generate n number of monthly loss regarding the frequency of loss distribution $Poisson(\lambda)$

3- Generate n losses X_i, $(i = 1, ..., n)$ regarding the loss severity distribution $\mathcal{LN}(\mu, \sigma^2)$;

4- Repeat steps 2 and 3 for $N = 12$. Summing all the generated X_i to obtain S which is the annual loss;

5- Repeat steps 1 to 4 many times (in our case 10^5) to obtain the annual aggregate loss distribution.

6- The VaR is calculated taking the $99.9th$ percentile of the aggregate loss distribution.

We notice that our Value-at-Risk is close to the VaR generated by the internal losses alone, since the only thing took as unknown was λ, both parameters μ and σ are equal to (μ_{ID}, σ_{ID}).

2.2 Modelling severity distribution: Lognormal $\mathcal{LN}(\mu, \sigma)$ distribution with unknown μ

Assume that the loss severity for a risk is modelled as a random variable from a lognormal distribution $\mathcal{LN}(\mu, \sigma)$ and we consider $\mu \hookrightarrow \mathcal{N}(\mu_0, \sigma_0^2)$ as a prior distribution.

So we have, $\Pi(\mu) = \dfrac{1}{\sigma_0\sqrt{2\pi}} exp\left\{-\dfrac{(\mu-\mu_0)^2}{2\sigma_0^2}\right\}$.

Taking $Y = \ln X$, we calculate the posterior distribution as previously by:

$$\Pi(\mu|\mu_0, \sigma_0^2) \propto \Pi(\mu)h(Y|\mu, \sigma) \propto \frac{e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}}}{\sigma_0\sqrt{(2\pi)}} \prod_{i=1}^{n} \frac{e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

since we are using a conjugate prior distribution, we know that the posterior distribution will follow a Normal distribution with parameters (μ_1, σ_1^2), where:

$$\Pi(\mu|\mu_0, \sigma_0^2) \propto e^{-\frac{(\mu-\mu_1)}{2\sigma_1^2}}$$

$$\text{By identification we obtain: } \begin{cases} \dfrac{1}{2\sigma_1^2} &= \dfrac{1}{2\sigma_0^2} + \dfrac{n}{2\sigma^2} \\ \dfrac{\mu_1}{\sigma_1^2} &= \dfrac{\mu_0}{\sigma_0^2} + \dfrac{\sum_{i=1}^n y_i}{\sigma^2} \end{cases}$$

So, $\mu_1 = \dfrac{\mu_0 + \omega_0 \sum_{i=1}^n y_i}{1 + n\omega_0} = \omega\bar{Y} + (1-\omega)\mu_0$, $\sigma_1^2 = \dfrac{\sigma_0^2}{1 + n\omega_0}$, with $\omega_0 = \dfrac{\sigma_0^2}{\sigma^2}$, and $\omega = \dfrac{n\omega_0}{1 + n\omega_0}$

Assuming that the loss severity for a risk is modelled as a random variable from a lognormal distribution $X \hookrightarrow \mathcal{LN}(\mu, \sigma)$, $\Omega = \mathbb{E}[X|\mu, \sigma] = e^{\mu + \frac{1}{2}\sigma^2} \hookrightarrow \mathcal{LN}(\mu_0 + \frac{1}{2}\sigma^2, \sigma_0^2)$ and we consider $\mu \hookrightarrow \mathcal{N}(\mu_0, \sigma_0^2)$ as a prior distribution.
Since the only thing unknown is μ, we already have $\sigma = 1.67$ and $\lambda = 5.8$, and the experts gave us:

$$\begin{aligned} \mathbb{E}[\Omega] &= e^{\mu_0 + \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma_0^2} = 15,825 \ \$ \\ \mathbb{P}(1 \le \Omega \le 250,000) &= \Phi\left(\frac{\ln 250,000 - \frac{1}{2}\sigma^2 - \mu_0}{\sigma_0}\right) - \Phi\left(\frac{\ln 1 - \frac{1}{2}\sigma^2 - \mu_0}{\sigma_0}\right) = 99\% \end{aligned}$$

Where Φ is the cumulative distribution function of the standard normal distribution.

Solving these two equations, we find that the prior distribution of μ is: $\mu \hookrightarrow \mathcal{N}(\mu_0 = 8.15, \sigma_0^2 = 0.25)$.

Hence using the formulas stated above where, $\mu_1 = \dfrac{\mu_0 + \omega_0 \sum_{i=1}^n y_i}{1 + n\omega_0} = 6.72$, $\sigma_1^2 = \dfrac{\sigma_0^2}{1 + n\omega_0} = 0.0096$, and $\omega_0 = \dfrac{\sigma_0^2}{\sigma^2} = 0.0898$, with $n = 279$ is the total number of historical losses.

We find out that the posterior distribution: $\mu \hookrightarrow \mathcal{N}(\mu_1 = 6.72, \sigma_1 = 0.1)$.
At the end, using the posterior μ distribution and Monte Carlo method, we calculate the 99.9% Value-at-Risk: $VaR(99.9\%) = 1,188,079.00$.

The same analysis goes here as well, since the only unknown parameter is μ, $(\lambda, \sigma) = (\lambda_{ID}, \sigma_{ID})$, the VaR calculated will be closer to our Internal Data Value-at-Risk.

2.3 Modelling frequency and severity distributions: Unknown Poisson(λ) parameter and Lognormal $\mathcal{LN}(\mu, \sigma)$ distribution with unknown μ

In the two previous subsections, we illustrated the case of modelling frequency and severity distributions with unknown λ that follows a $Gamma(a, b)$ distribution and μ that follows a $\mathcal{N}(\mu_0, \sigma_0^2)$ respectively.
Joining these two distributions is relatively simple since we have the hypothesis of independence between frequency and severity, which allows us to estimate independently the two posterior distributions and estimate the parameters.

As so, we have already demonstrated the fact that our posterior density $\Pi(\lambda | N = n)$ follows the $Gamma(a_T, b_T)$ distribution, with $a_T = \sum_{i=1}^{T} n_i + a$ and $b_T = \frac{b}{(1 + nb)}$
and, $\Pi(\mu | \mu_0, \sigma_0^2) \hookrightarrow \mathcal{N}(\mu_1, \sigma_1^2)$, with $\mu_1 = \frac{\mu_0 + \omega_0 \sum_{i=1}^{n} y_i}{1 + n\omega_0}$, $\sigma_1^2 = \frac{\sigma_0^2}{1 + n\omega_0}$, with $\omega_0 = \frac{\sigma_0^2}{\sigma^2}$
Since we have the hypothesis of independence between frequency and severity, which allows us to estimate independently the two posterior distributions, which have been already calculated for the parameter λ we took the gamma distribution and for the μ parameter, the posterior distribution was normal with:

$$\lambda \hookrightarrow Gamma(279.8, 0.02)$$
$$\mu \hookrightarrow \mathcal{N}(6.72, 0.1)$$

By simulating those two laws using Monte Carlo simulation (cf. section 2.1), we obtain a Value-at Risk of $1,199,000.00$ using the estimated posterior Gamma and Normal distributions.

This result is highly interesting, since with two unknown parameters λ and μ, the VaR is still closer to VaR_{ID}. This states that the parameter σ is the key parameter in this application, as we are going to see throughout this article.

The general case where all parameters are unknown will not be treated in this section since it is more complex to tackle it with the use of conjugate prior distributions (cf. Shevchenko [2011] pp. 129-131) for details.

3 Bayesian techniques in combining two data sources: MCMC-Metropolis Hastings algorithm

In this section, we will use a noninformative prior and more particularly the Jeffreys prior, (cf. Jeffreys [1946]), that attempts to represent a near-total absence of prior knowledge that is proportional to the square root of the determinant of the Fisher information:

$$\pi(\omega) \propto \sqrt{|I(\omega)|},$$

where $I(\omega) = -\mathbb{E}\left(\frac{\partial^2 \ln \mathcal{L}(X|\omega)}{\partial \omega^2}\right)$.
Then we are going to apply an MCMC model to obtain a distribution for the parameters and generate our capital required at 99.9%. This will allow us to compare both methods' results and develop a generalized application to set up our capital allocation, since no restrictions is made regarding the distributions. As for the parameter σ, it will no longer be fixed as in the previous sections. For more details on the Jeffreys prior and MCMC-Metropolis Hastings algorithm check Robert [2007].

3.1 MCMC with the Poisson(λ) distribution

Assuming that the parameter λ is the only thing unknown, the Jeffreys prior distribution is: $\pi(\lambda) \propto \frac{\sqrt{\lambda}}{\lambda}$ (see Appendix A), thus finding the posterior distribution $f(\lambda|n_{SA}, n_{ID})$ with the use of experts Scenario Analysis and Internal Data would be:

$$f(\lambda|n_{SA}, n_{ID}) \propto \overbrace{\pi(\lambda)}^{Jeffreys\ prior} \underbrace{\mathcal{L}(n_{SA}, \lambda)\mathcal{L}(n_{ID}, \lambda)}_{Likelihood\ functions}.$$

So by applying Metropolis Hastings algorithm, (check appendix $B.1$ for full support on detailed algorithm), with the objective density:

$$\begin{aligned} f(\lambda|n_{SA}, n_{ID}) &\propto \frac{1}{\sqrt{\lambda}} \prod_{k=1}^{n_{SA}} \frac{e^{-\lambda}\lambda^k}{k!} \prod_{k=1}^{n_{ID}} \frac{e^{-\lambda}\lambda^k}{k!} \\ &\propto \frac{1}{\sqrt{\lambda}} \prod_{k=1}^{n_{SA}} e^{-\lambda}\lambda^k \prod_{k=1}^{n_{ID}} e^{-\lambda}\lambda^k \\ &\propto \frac{1}{\sqrt{\lambda}} e^{n_{SA}\lambda} \lambda^{\sum_k^{n_{SA}} k} e^{n_{ID}\lambda} \lambda^{\sum_k^{n_{ID}} k} \end{aligned}$$

and with a uniform proposal density: $U(\lambda_{SA}, \lambda_{n_{ID}})$, we obtain the parameter λ distribution see Figure 1.

We have removed the first 3000 iterations so that the chain is stationary (burn-in iterations effect), (cf. Gilks *et al.* [1996] pp. 5-6). We obtain a 99.9 % Value-at-Risk of $1,000,527.00$.
The result is close to the VaR considered with the use of conjugate family.

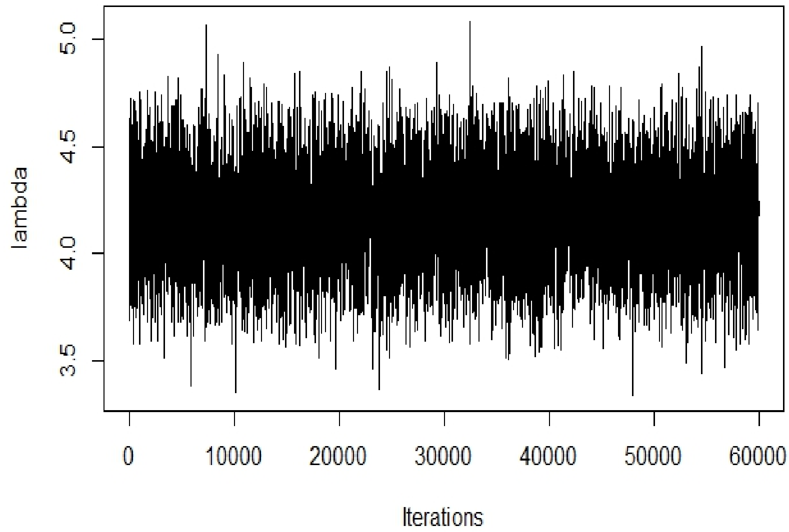

Figure 1: MCMC for the parameter λ

3.2 MCMC with Unknown Poisson(λ) parameter and Lognormal $\mathcal{LN}(\mu, \sigma)$ distribution with unknown μ

Assuming that the parameters λ and μ are the only things unknown, we will treat them independently and since the Poisson(λ) case has already been treated, the Jeffreys prior distribution for μ is: $\pi(\mu) \propto \frac{1}{\sigma} \propto 1$ (see Appendix A), thus finding the posterior distribution $f(\mu|x, y)$ with the use of experts Scenario Analysis and Internal Data would be:

$$f(\mu|x, y) \propto \overbrace{\pi(\mu)}^{Jeffreys\ prior} \underbrace{\mathcal{L}(x_1, x_2, ..., x_{n_{SA}}|\mu, \sigma_{SA}))\mathcal{L}(y_1, y_2, ..., y_{n_{ID}}|\mu, \sigma_{ID})}_{Likelihood\ functions}.$$

So by applying Metropolis Hastings algorithm, (check Appendix $B.2$ for full support on detailed algorithm), with the objective density:

$$\begin{aligned} f(\mu|x, y) &\propto \prod_{i=1}^{n_{SA}} \frac{1}{x_i\sqrt{2\pi\sigma_{SA}^2}} \exp\{-\frac{(\ln x_i - \mu)^2}{2\sigma_{SA}^2}\} \prod_{i=1}^{n_{ID}} \frac{1}{y_i\sqrt{2\pi\sigma_{ID}^2}} \exp\{-\frac{(\ln y_i - \mu)^2}{2\sigma_{ID}^2}\} \\ &\propto \exp\{-\sum_i \frac{(\ln x_i - \mu)^2}{2\sigma_{SA}^2}\} \exp\{-\sum_i \frac{(\ln y_i - \mu)^2}{2\sigma_{ID}^2}\} \end{aligned}$$

and with a uniform proposal density: $U(0, 12)$, we obtain the parameter μ distribution see Figure 2.

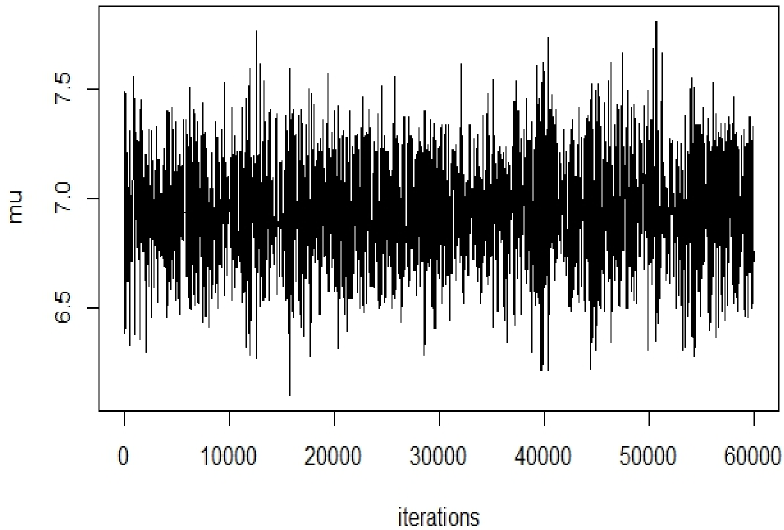

Figure 2: MCMC for the parameter μ

We obtain a Value-at-Risk of $1,167,060.00$.
Comparing this to the same case generated with conjugate prior, we can check the closeness of both values.

In the next subsection, we will tackle the general case, where all parameters are unknown, this case was not treated with conjugate prior distributions since it would be more complicated.

3.3 General case: MCMC with Unknown Poisson(λ) parameter and Lognormal $\mathcal{LN}(\mu, \sigma)$ distribution with unknown μ and σ

We are going to assume the general case, where all the parameters are unknown λ, μ and σ, we will treat them independently and since the Poisson(λ) case has already been employed, the Jeffreys prior distribution for $\omega = (\mu, \sigma)$ is: $\pi(\omega) \propto \frac{1}{\sigma^3}$ (cf. Appendix A), thus finding the posterior distribution $f(\omega|x, y)$ with the use of experts Scenario Analysis and Internal Data would be:

$$f(\omega|x,y) \propto \overbrace{\pi(\omega)}^{Jeffreys\ prior} \underbrace{\mathcal{L}(x_1, x_2, ..., x_{n_{SA}}|\mu,\sigma))\mathcal{L}(y_1, y_2, ..., y_{n_{ID}}|\mu,\sigma)}_{Likelihood\ functions}.$$

So by applying Metropolis Hastings algorithm, (check appendix $B.3$ for full support on detailed algorithm), with the objective density:

$$f(\omega|x,y) \propto \frac{1}{\sigma^3}\prod_{i=1}^{n_{SA}}\frac{1}{x_i\sqrt{2\pi\sigma^2}}\exp\{-\frac{(\ln x_i-\mu)^2}{2\sigma^2}\}\prod_{i=1}^{n_{ID}}\frac{1}{y_i\sqrt{2\pi\sigma^2}}\exp\{-\frac{(\ln y_i-\mu)^2}{2\sigma^2}\}$$
$$\propto \frac{1}{\sigma^3}\frac{1}{\sigma^{n_{SA}}}\exp\{-\sum_i\frac{(\ln x_i-\mu)^2}{2\sigma^2}\}\frac{1}{\sigma^{n_{ID}}}\exp\{-\sum_i\frac{(\ln y_i-\mu)^2}{2\sigma^2}\}$$

and with a uniform proposal density: $U(0,12)$ and $U(0,7)$ for μ and σ respectively, we obtain the parameters μ and σ distributions, illustrated in Figure 3.

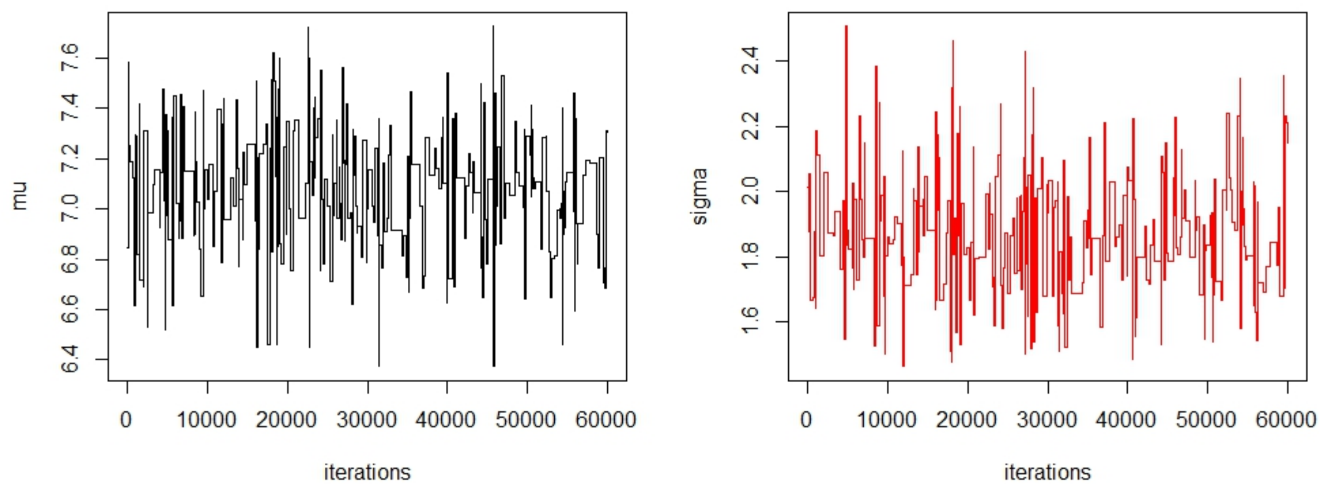

Figure 3: MCMC for the parameters μ and σ

We have removed as well, the first 3000 iterations so that the chain is stationary (burn-in iteration effect). We obtain a Value-at-Risk of $3,061,151.00$.

The general case clearly generates a good combination between internal data and experts' opinion with a capital requirement of $3,061,151$ \$.

3.4 Bayesian approach reviewed

To recapitulate on all the calculations, table 2 summarizes all Value-at-Risk generated. As for the calculation of the confidence interval, since we are working with order statistics, the interval (x_l, x_u) would cover our quantile x_p with a 99.5% probability that depends on the lower bound l, upper bound u, number of steps n and confidence level p.

In our calculations, we took $n = 10^5$, $p = 99.9\%$ and our integers (l, u), were constructed using the normal approximation $\mathcal{N}(np, np(1-p))$ to the binomial distribution $\mathcal{B}(n, p)$,(since n is large). Then a simple linear interpolation has been made to obtain the values of (x_l, x_u), (cf. David & Nagaraga [2003] pp. 183-186), for more details and demonstrations.

Case	***Confidence Interval***		***VaR (99.9%)***	**Length**
Aggregate	*$ 1,040,697.00*	*$ 1,230,492.00*	*$ 1,162,215.00*	*15.42%*
Scenario Analysis	*$ 6,094,853.00*	*$ 7,171,522.00*	*$ 6,592,086.00*	*15.01%*
Bayesian unknowm λ	*$ 1,053,861.00*	*$ 1,184,129.00*	*$ 1,117,821.00*	*11.00%*
Bayesian unknown μ	*$ 1,097,195.00*	*$ 1,268,136.00*	*$ 1,188,079.00*	*13.48%*
Bayesian unknowm λ and μ	*$ 1,141,767.00*	*$ 1,318,781.00*	*$ 1,199,000.00*	*13.42%*
MCMC λ	***$ 944,793.10***	***$1,101,274.00***	***$1,000,527.00***	*14.21%*
MCMC λ, μ	***$1,098,930.00***	***$1,244,564.00***	***$1,167,060.00***	*11.70%*
MCMC λ, μ, σ	***$2,839,706.00***	***$3,310,579.00***	***$3,061,151.00***	*14.22%*

Table 2: Value at Risk and Confidence intervals for all cases treated

Table 2 clearly shows the helpful use of the Bayesian inference techniques. The results of both methods are close and comparable; though conjugate prior is simple but the distributions are restricted to the conjugate family, yet with the Jeffreys noninformative prior and MCMC-Metropolis Hastings algorithm, we will have a wider options and generate a good combination between internal data and experts' opinion.

4 Conclusion

Using the information given by the experts, we were able to determine all the parameters of our prior distribution, leading to the posterior distributions with the use of internal data, which allowed us to compute our own capital requirement. This approach offers a major simplicity in its application through the employment of the conjugate distributions. Therefore, allowing us to obtain explicit formulas to calculate our posterior parameters. Yet, the appliance of this approach could not be perfected since it's restricted to the conjugate family.

On the other hand, Jeffreys prior with MCMC-Metropolis Hastings algorithm provided us with wider options and generated a satisfactory result regarding all three unknown variables λ, μ and σ, with the only difference of using complex methods. Taking σ unknown as well, was very essential in reflecting the credibility of estimating our capital requirement.

In our study, we displayed a particular application of the Bayesian inference methods showing a more robust capital allocation that better approaches the reality. The only thing not taken into consideration is external data, which might be interesting to elaborate and apply in practice, more on this subject could be found in Lambrigger *et al.* [2011].

References

Basel Committee on Banking Supervision [2002], *Quantitative Impact Study 3 Technical guidance*, www.bis.org/bcbs/qis/qis3tech.pdf

Berger J.O. [1985], *Statistical Decision Theory and Bayesian Analysis*, 2nd edition, Springer, New York (NY)

David H.A. & Nagaraja H.N. [2003], *Order Statistics third edition*, New York: John Wiley & Sons inc.

Frachot A. & Georges P. & Roncalli T. [2001], *Loss distribution approach for Operational Risk*, Groupe de Recherche Opérationelle, Crédit Lyonnais, France

Hastings W.K. [1970], *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, Vol. 57, No. 1., pp. 97-109

Jeffreys H. [1946], *An invariant form for the prior probability in estimation problems*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, Vol. 186, No. 1007, pp. 453-461

Lambrigger D. D. & Shevchenko P. V. & Wüthrich M. V. [2008], *Data combination under Basel II and Solvency 2: Operational Risk goes Bayesian*, Bulletin Français d'Actuariat, vol. 8, $n°16$

Robert C.P. [2007], *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation second edition*, Springer

Shevchenko P.V. [2011], *Modelling Operational risk Using Bayesian Inference*, Springer

Solvency II [2010], *Quantitative Impact Study 5 (Final Technical Specifications)*, www.eiopa.europa.eu

A Jeffreys prior distribution

Jeffreys prior attempts to represent a near-total absence of prior knowledge that is proportional to the square root of the determinant of the Fisher information:

$$\pi(\omega) \propto \sqrt{|I(\omega)|},$$

where $I(\omega) = -\mathbb{E}\left(\dfrac{\partial^2 \ln \mathcal{L}(X|\omega)}{\partial \omega^2}\right)$.

A.1 Jeffreys prior for Poisson(λ) and Lognormal(μ, σ) distributions

Let $N \hookrightarrow \mathcal{P}(\lambda)$, the poisson density function is: $f(k|\lambda) = \mathbb{P}(N = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$ with,

$$\frac{\partial^2 \ln f(k|\lambda)}{\partial \lambda^2} = -\frac{k}{\lambda^2}$$

and consequently, $\pi(\lambda) \propto \dfrac{\sqrt{\lambda}}{\lambda}$.

Let $X \hookrightarrow \mathcal{LN}(\mu, \sigma^2)$, with $f_X(x) = \dfrac{1}{x\sqrt{2\pi\sigma^2}} \exp\{-\dfrac{(\ln x - \mu)^2}{2\sigma^2}\}$.

Hence, by letting $\omega = (\mu, \sigma)$ and calculating the corresponding partial derivatives to $\ln f_X(x)$ we obtain:

$$I(\omega) = \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 1/2\,\sigma^4 \end{bmatrix}$$

As a consequence, $\pi(\mu) \propto \dfrac{1}{\sigma^2} \propto 1$ and $\pi(\omega) = \dfrac{1}{\sqrt{2\sigma^6}} \propto \dfrac{1}{\sigma^3}$

B MCMC Metropolis-Hastings algorithm

B.1 Applying MCMC with Metropolis Hastings algorithm for λ

1- Initialize $\lambda_0 = \dfrac{\lambda_{ID} + \lambda_{SA}}{2}$

2- Update from λ_i to λ_{i+1} $(i = 1, ..., n)$ by

- Generating $\lambda \hookrightarrow U(\lambda_{SA}, \lambda_{ID})$
- Define $\zeta = \min\left(\dfrac{f(\lambda|n_{SA}, n_{ID})}{f(\lambda_i|n_{SA}, n_{ID})}, 1\right)$
- Generate $Rnd \hookrightarrow U(0, 1)$
- If $Rnd \leq \zeta$, $\lambda_{i+1} = \lambda$, else $\lambda_{i+1} = \lambda_i$

3- Remove the first 3000 iterations, so that the chain is stationary (burn-in effect).

B.2 Applying MCMC with Metropolis-Hastings algorithm for μ

1- Initialize $\mu_0 = \mu_{ID}$

2- Update from μ_i to μ_{i+1} $(i = 1, ..., n)$ by

- Generating $\mu \hookrightarrow U(0, 12)$
- Define $\zeta = \min\left(\dfrac{f(\mu|x, y)}{f(\mu_i|x, y)}, 1\right)$
- Generate $Rnd \hookrightarrow U(0, 1)$
- If $Rnd \leq \zeta$, $\mu_{i+1} = \mu$, else $\mu_{i+1} = \mu_i$

3- Remove the first 3000 iterations, so that the chain is stationary (burn-in effect).

B.3 Applying MCMC with Metropolis-Hastings algorithm for $\omega = (\mu, \sigma)$

1- Initialize $\mu_0 = \mu_{ID}$ and $\sigma_0 = \sigma_{ID}$

2- Update from μ_i to μ_{i+1} and σ_i to σ_{i+1}, $(i = 1, ..., n)$ by

- Generating $\mu \hookrightarrow U(0, 12)$ and $\sigma \hookrightarrow U(0, 7)$
- Define $\zeta = \min\left(\frac{f(\mu, \sigma|x, y)}{f(\mu_i, \sigma_i|x, y)}, 1\right)$
- Generate $Rnd \hookrightarrow U(0, 1)$
- If $Rnd \leq \zeta$, $\mu_{i+1} = \mu$ and $\sigma_{i+1} = \sigma$ else $\mu_{i+1} = \mu_i$ and $\sigma_{i+1} = \sigma_i$

3- Remove the first 3000 iterations from both distributions, so that the chains is stationary (burn-in effect).