

Combining Internal Data with Scenario Analysis

Elias Karam, Frédéric Planchet

Laboratoire SAF EA2429, ISFA, Université Claude Bernard Lyon 1, Université Lyon, Lyon, France
Email: ek.eliaskaram@gmail.com, frederic.planchet@univ-lyon1.fr

Received 8 April 2015; accepted 18 May 2015; published 22 May 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

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 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. In addition, we illustrate the importance of a consistent scenario analysis in showing that the expert opinion coherence leads to a robust estimation of risk.

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 provide to some extent a future vision.

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 [1] 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 evaluation. For a good introduction to non-informative prior distributions and MCMC see [2].

Combining these different information sources for model estimation is certainly one of the main challenges in operational risk.

Modelling frequency and severity losses for estimating annual loss distribution, is known actuarial technique used to model, as well, solvency requirements in the insurance industry, see for e.g. [3] [4]. More literature on Operational Risk and Bayesian Inference techniques could be found in [5]-[8].

2. Combining Two Data Sources: The 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 fits 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 [9], with $\lambda_{ID} = 5.8$, $\mu_{ID} = 6.7$, and $\sigma_{ID} = 1.67$, we obtained a Value-at-Risk of $VaR_{ID} = 1162215.00$ 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).

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 $VaR(99.9\%) = 6592086.00$.

The high spread between the two VaR values, can cause a problem in allocating a non-biased 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

We are going to work with the Poisson and Lognormal distributions since they are the most used distributions in Operational Risk [10].

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

Table 1. Scenario analysis.

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%

we have: $\mathbb{P}(N = n) = e^{-\lambda} \frac{\lambda^n}{n!}$, and λ has a prior density:

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

As for the likelihood function, given the assumption that n_1, n_2, \dots, 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:

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

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:

$$\mathbb{E}[\lambda] = a \times b = 2 \tag{1}$$

$$\mathbb{P}(a_0 \leq \lambda \leq b_0) = \int_{a_0}^{b_0} \pi(\lambda|a, b) d\lambda = F_{a,b}^{(G)}(b_0) - F_{a,b}^{(G)}(a_0) = 0.7 \tag{2}$$

where $F_{a,b}^{(G)}(\cdot)$ is the *Gamma*(a, b) cumulative distribution function.

Solving the above equations would give us the prior distribution parameters $\lambda \mapsto \text{Gamma}(a = 0.79, b = 2.52)$, 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 $\text{VaR}(99.9\%) = 1117821.00$ using Monte Carlo simulation:

- Using the estimated Posterior *Gamma*(a_T, b_T) distribution, generate a value for λ ;
- Generate n number of monthly loss regarding the frequency of loss distribution *Poisson*(λ)
- Generate n losses $X_i, (i = 1, \dots, n)$ regarding the loss severity distribution $\mathcal{LN}(\mu, \sigma^2)$;
- Repeat steps b and c for $N = 12$. Summing all the generated X_i to obtain S which is the annual loss;
- Repeat steps a to d many times (in our case 10^5) to obtain the annual aggregate loss distribution.
- The VaR is calculated taking the 99.9th percentile of the aggregate loss distribution.

We notice that the 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: $\mathcal{LN}(\mu, \sigma)$ 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 $\mu \rightarrow \mathcal{N}(\mu_0, \sigma_0^2)$ as a prior distribution.

So we have,

$$\Pi(\mu) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left\{ -\frac{(\mu - \mu_0)^2}{2\sigma_0^2} \right\}.$$

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:

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

By identification we obtain:

$$\begin{cases} \frac{1}{2\sigma_1^2} = \frac{1}{2\sigma_0^2} + \frac{n}{2\sigma^2} \\ \frac{\mu_1}{\sigma_1^2} = \frac{\mu_0}{\sigma_0^2} + \frac{\sum_{i=1}^n y_i}{\sigma^2} \end{cases}$$

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

Assuming that the loss severity for a risk is modelled as a random variable from a lognormal distribution $X \rightarrow \mathcal{LN}(\mu, \sigma)$, $\Omega = \mathbb{E}[X | \mu, \sigma] = e^{\mu + \frac{1}{2}\sigma^2} \rightarrow \mathcal{LN}\left(\mu_0 + \frac{1}{2}\sigma^2, \sigma_0^2\right)$ and we consider $\mu \rightarrow \mathcal{N}(\mu_0, \sigma_0^2)$ 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:

$$\mathbb{E}[\Omega] = e^{\mu_0 + \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma_0^2} = 15825 \text{ \$} \tag{3}$$

$$\mathbb{P}(1 \leq \Omega \leq 250000) = \Phi\left(\frac{\ln 250000 - \frac{1}{2}\sigma^2 - \mu_0}{\sigma_0}\right) - \Phi\left(\frac{\ln 1 - \frac{1}{2}\sigma^2 - \mu_0}{\sigma_0}\right) = 99\% \tag{4}$$

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: $\mu \rightarrow \mathcal{N}(\mu_0 = 8.15, \sigma_0^2 = 0.25)$.

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

We find out that the posterior distribution: $\mu \rightarrow \mathcal{N}(\mu_1 = 6.72, \sigma_1 = 0.1)$.

At the end, using the posterior μ distribution and Monte Carlo method, we calculate the 99.9% Value-at-Risk: $VaR(99.9\%) = 1188079.00$.

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

2.3. Modelling Frequency and Severity Distributions

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, $\Pi(\mu|\mu_0, \sigma_0^2) \rightarrow \mathcal{N}(\mu_1, \sigma_1^2)$,

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:

$$\lambda \rightarrow Gamma(279.8, 0.02)$$

$$\mu \rightarrow \mathcal{N}(6.72, 0.1)$$

By simulating those two laws using Monte Carlo simulation (cf. Section 2.1), we obtain a Value-at Risk of 1199000.00 using the estimated posterior Gamma and Normal distributions.

The result is 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.

2.4. Sensitivity Analysis

Working with this method, is generally simple since conjugate prior is involved, yet one of the main questions is how to ensure that experts opinion are consistent, relevant, and capture well the situation, which might in a way, cause a model error. In this work we did not take into consideration this aspect and the experts opinion were treated as correct. To improve our results, we can do a sensitivity test regarding our prior parameters. On the other hand, we only have the mean of the number of losses, given by our expert. So it appears difficult to obtain the distribution of λ with this only information. So in a way, we are immensely relying on our prior parameters which in reality don't give us a banking sense and are not easily comprehensive.

We are going to test out prior parameters given by the experts and highlight the direct consequence on our Capital Required. To start with the first case, where we are working with unknown λ ; our experts gave us a value of $a_0 = 0.5$ and $b_0 = 8$ as seen in section 2.1, so taking into consideration a step of 0.035 for an interval $[0.1, 1.5]$ and 0.1 for [6] [10], respectively, we obtain the following VaR results (check [Figure 1](#)).

As for the cases of unknown μ and unknown σ , we took an interval for $\mu_0 \in [6, 10]$ and $\sigma_0 \in [0.3, 4.1]$ with a step of 0.1 (see [Figure 2](#)).

The following figures show the stability of our Value-at-Risk calculations regardless of all changes in our prior parameters $(a_0, b_0, \mu_0, \sigma_0)$. Relying on the choice of our intervals, we notice that the boundaries of our VaR are in an acceptable range.

3. MCMC-Metropolis Hastings Algorithm

In this section, we will use a noninformative prior and more particularly the Jeffreys prior [11], 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:

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

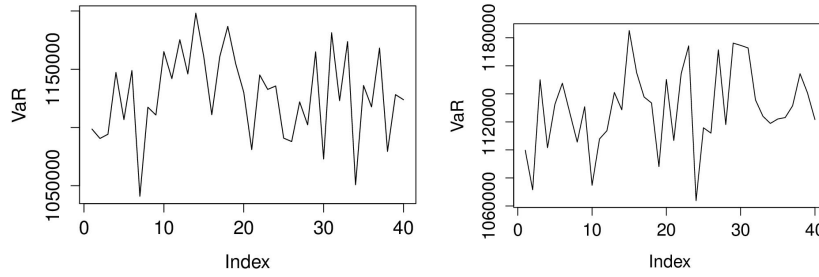


Figure 1. Sensitivity for a_0 and b_0 respectively.

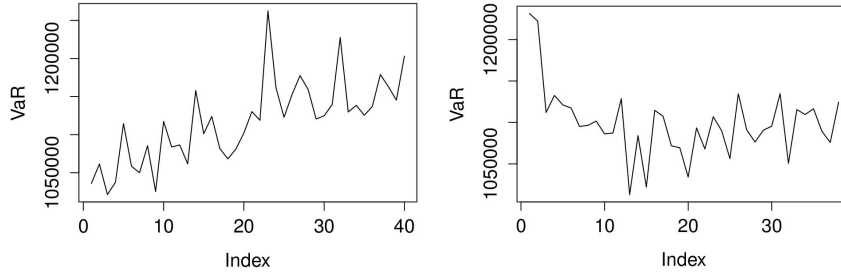


Figure 2. Sensitivity for μ_0 and σ_0 respectively.

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 [12].

3.1. MCMC with 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 5.1), thus finding the posterior distribution $f(\lambda|n_{SA}, n_{ID})$ with the use of experts Scenario Analysis and Internal Data would be:

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

So by applying Metropolis Hastings algorithm, (check Appendix 5.2.1 for full support on detailed algorithm), with the objective density:

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

and with a uniform proposal density: $U(\lambda_{SA}, \lambda_{ID})$, we obtain the parameter λ distribution see **Figure 3**.

We have removed the first 3000 iterations so that the chain is stationary (burn-in iterations effect), [13]. We

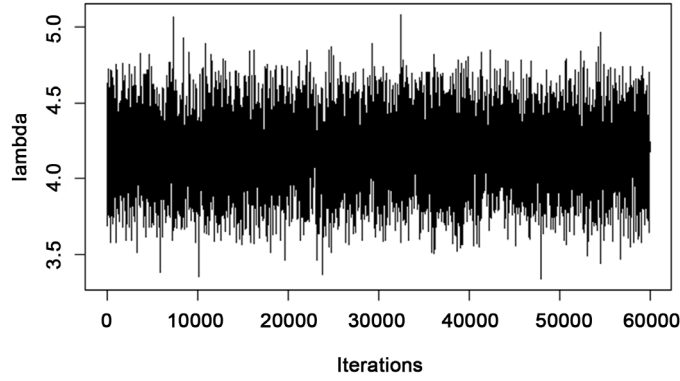


Figure 3. MCMC for the parameter λ .

obtain a 99.9% Value-at-Risk of 1000527.00

The result is close to the VaR considered with the use of conjugate family.

3.2. MCMC with $\mathcal{LN}(\mu, \sigma)$ Distribution with Unknown μ

Assuming that the parameters λ and μ are the only things unknown, we will treat them independently and since the $\text{Poisson}(\lambda)$ case has already been treated, the Jeffreys prior distribution for μ is: $\pi(\mu) \propto \frac{1}{\sigma} \propto 1$ (see Appendix 5.1), thus finding the posterior distribution $f(\mu|x, y)$ with the use of experts Scenario Analysis and Internal Data would be:

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

So by applying Metropolis Hastings algorithm, (check Appendix 5.2 for full support on detailed algorithm), with the objective density:

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

and with a uniform proposal density: $U(0,12)$, we obtain the parameter μ distribution see **Figure 4**.

We obtain a Value-at-Risk of 1167060.00.

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. MCMC: The General Case

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 $\text{Poisson}(\lambda)$ case has already been employed, the Jeffreys prior distribution for $\omega = (\mu, \sigma)$ is: $\pi(\omega) \propto \frac{1}{\sigma^3}$ (cf. Appendix 5.1), thus finding the posterior distribution $f(\omega|x, y)$ with the use of experts Scenario Analysis and Internal Data would be:

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

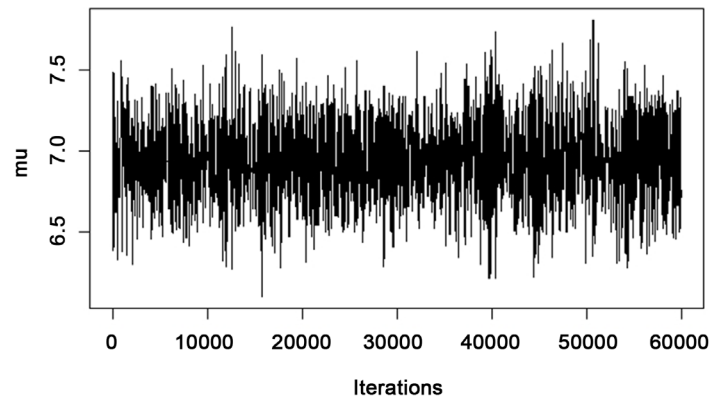


Figure 4. MCMC for the parameter μ .

So by applying Metropolis Hastings algorithm, (check Appendix 5.2.3 for full support on detailed algorithm), with the objective density:

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

$$\propto \frac{1}{\sigma^3} \frac{1}{\sigma^{n_{SA}}} \exp\left\{-\sum_i \frac{(\ln x_i - \mu)^2}{2\sigma^2}\right\} \frac{1}{\sigma^{n_{ID}}} \exp\left\{-\sum_i \frac{(\ln y_i - \mu)^2}{2\sigma^2}\right\}$$

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 5.

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 3061151.00.

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

3.4. Confidence Interval Calculation

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 95% 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) , (see [14], pp. 183-186), for more details and demonstrations.

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 non-informative prior and MCMC-Metropolis Hastings algorithm, we will have a wider options and generate a good combination between internal data and experts' opinion.

In addition, we are going to use the Monte Carlo confidence intervals approach [15] in order to compare it with the previous approach and ensure the similarities.

Consider a parameter X with its consistent estimator Y , with cumulative distribution function $F(y|x)$ generated by some process which can be simulated. Number of simulations n and y values are independently generated and then ordered from largest to smallest. An approximate $100(1-2\alpha)$ confidence level for X is given by (y_j, y_k) where j and k represent respectfully the lower and upper bound of the interval and they are set as: $j = (n+1)\alpha$ and $k = (n+1)(1-\alpha)$. Usually j and k will not be integer; therefore we can simply round it to the nearest integer values or even use linear interpolation.

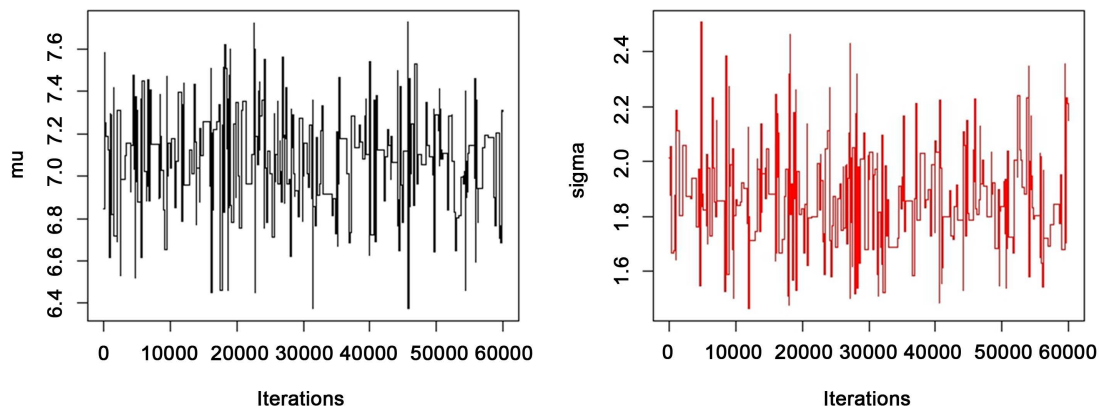


Figure 5. MCMC for the parameters μ and σ .

Table 2. Value at risk and confidence intervals for all cases treated.

Case	Confidence interval		VaR (99.9%)	Length
Aggregate	\$1040697.00	\$1230492.00	\$1162215.00	15.42%
Scenario analysis	\$6094853.00	\$7171522.00	\$6592086.00	15.01%
Bayesian unknown λ	\$1053861.00	\$1184129.00	\$1117821.00	11.00%
Bayesian unknown μ	\$1097195.00	\$1268136.00	\$1188079.00	13.48%
Bayesian unknown λ and μ	\$1141767.00	\$1318781.00	\$1199000.00	13.42%
MCMC λ	\$944793.10	\$1101274.00	\$1000527.00	14.21%
MCMC λ, μ	\$1098930.00	\$1244564.00	\$1167060.00	11.70%
MCMC λ, μ, σ	\$2839706.00	\$3310579.00	\$3061151.00	14.22%

We seek to calculate (y_j, y_k) , this may be found, using a conventional $100(1-2\alpha)\%$ confidence level, by solving: $F(Y|X = y_j) = 1-\alpha$ and $F(Y|X = y_k) = \alpha$.

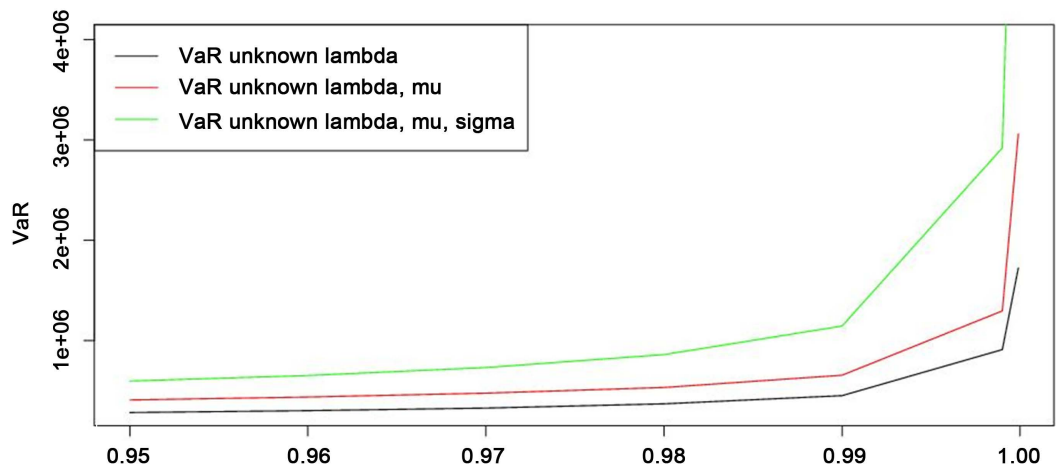
The actual confidence level has a beta distribution with parameters $k-j$ and $n-k+j+1$, this is concluded when percentiles of the distribution $F(Y|X = y)$ are estimated by simulation.

$$\text{Respecting that } B \text{ has a beta distribution, } E(B) = \frac{k-j}{n+1} \text{ and } \text{Var}(B) = \frac{(k-j)(n-k+j+1)}{(n+1)^2(n+2)}$$

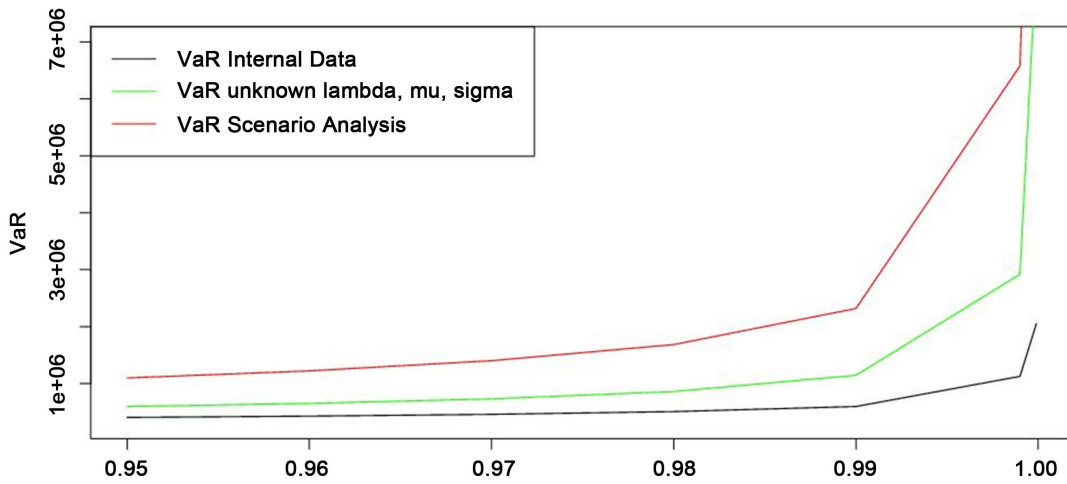
In our case, by using the confidence level of 99.9% and by applying the previous calculations we have obtained an approximation 95% interval for actual confidence level with $n = 10^5$, $p = 99.9\%$, $j = 50$, $k = 99951$, $\alpha = 0.05\%$, $\zeta = 2.5\%$, $\sigma = \sqrt{\text{Var}(B)}$ and by moving 1.96 standard errors in either direction from the estimate we obtain our confidence interval: $[p - \Phi^{-1}(\zeta)\sigma, p + \Phi^{-1}(\zeta)\sigma] = [99.88\%, 99.92\%]$, which is very close to the previous interval calculation in Table 2.

Furthermore, Figure 6 and Figure 7 illustrate the calculation of our Value-at-Risk for different confidence level. It clearly shows the presence of the general case, between both Internal and Scenario Analysis curves. On the other hand, the conjugate prior Figure 7, regarding all 3 unknown variables, point out the closeness of the curves which add to our previous analysis that σ is our key parameter.

We have to note as well, that the concept of Scenario Analysis with the expert opinion should deserve more clarification. Roughly speaking, when we refer to experts judgments, we express the idea that banks' experts and experienced managers have some reliable intuitions on the riskiness of their business and that these intuitions are not entirely reflected in the bank's historical, internal data. In our case, experts' intuitions were directly plugged into severity and frequency estimations through building a loss histogram.



(a)



(b)

Figure 6. Different VaR calculation for all MCMC cases, internal data and scenario analysis.

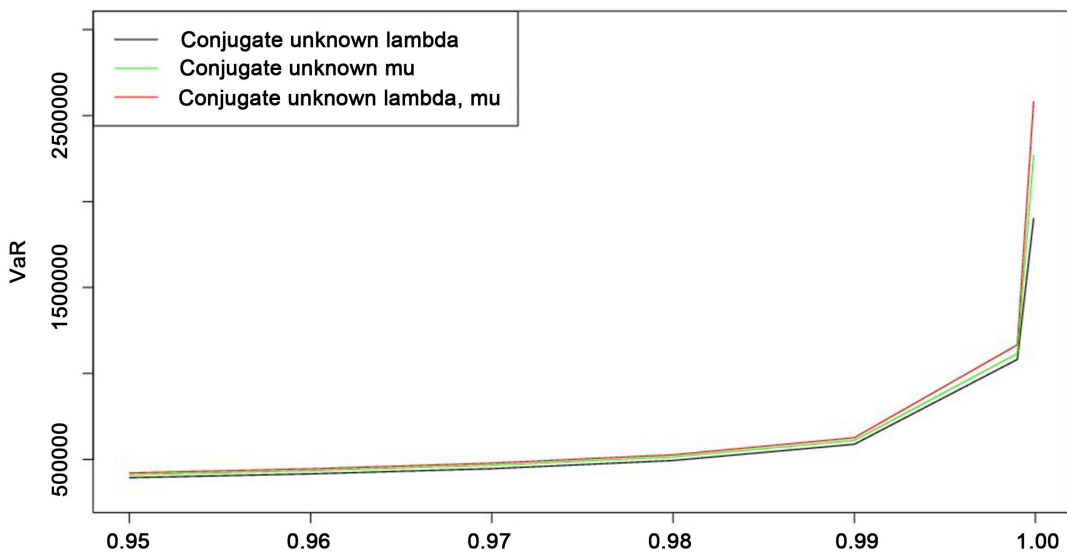


Figure 7. Different VaR calculation for all conjugate prior cases.

3.5. Bayesian Approach Reviewed

In this part, we are going to replace the experts opinions, by assuming that the experts' parameters are set using the Basel II standardized approach calculation. Hence, the experts opinion is questionable in the meaning of when it's used, we shift into the Markovian process which can cause problems.

3.6. Standardized Approach Reviewed

In the Standardized Approach (SA), banks' activities are divided into 8 business lines [16] [17]: corporate finance, trading & sales, retail banking, commercial banking, payment & settlements, agency services, asset management, and retail brokerage. Within each business line, there is a specified general indicator that reflects the size of the banks' activities in that area. The capital charge for each business line is calculated by multiplying gross income by a factor β assigned to a particular business line (Table 3).

The total capital charge is calculated as a three year average over all positive gross income (GI) as follows:

$$K_{SA} = \frac{\sum_{i=1}^3 \max\left(\sum_{j=1}^8 \beta_j GI^i, 0\right)}{3}$$

Hence, the application of the Standardized Approach generates a capital requirement of $K_{SA} = 2760780$.

Numerical Results for Expert Opinion Treated as SA

Setting the parameters to give us the same Standardized approach capital requirement and treating them as the expert parameters gave us: $\lambda = 6.8$, $\mu = 7.3$ and $\sigma = 1.72$.

We note that the Standardized approach from Basel II is the one to rely on when it comes to calculate the VaR with its confidence interval. It is interesting to compare both results in Table 2 and Table 4 where we notice

Table 3. Business lines and the beta factors.

Business line (j)	Beta factors (β_j)
$j = 1$, corporate finance	18%
$j = 2$, trading & sales	18%
$j = 3$, retail banking	12%
$j = 4$, commercial banking	15%
$j = 5$, payment & settlement	18%
$j = 6$, agency services	15%
$j = 7$, asset management	12%
$j = 8$, retail brokerage	12%

Table 4. Value at risk and confidence intervals for all cases treated.

Case	Confidence interval		VaR (99.9%)	Length
Aggregate	\$1040697.00	\$1230492.00	\$1162215.00	15%
Scenario analysis	\$2570297.00	\$2876469.00	\$2759640.00	12%
Bayesian unknown λ	\$1084427.00	\$1257414.00	\$1172801.00	16%
Bayesian unknown μ	\$1045412.00	\$1183887.00	\$1118045.00	13%
Bayesian unknown λ and μ	\$1114267.00	\$1249999.00	\$1175326.00	12%
MCMC λ	\$1025132.00	\$1188600.00	\$1083511.00	16%
MCMC λ, μ	\$1169519.00	\$1347836.00	\$1253938.00	15%
MCMC λ, μ, σ	\$1678124.00	\$1912897.00	\$1769198.00	14%

that the VaR results in the cases of bayesian unknown λ , unknown μ and unknown λ, μ are very close to the result in the Aggregate case. As for the MCMC approach where experts opinion are respected, in the case of unknown λ, μ and σ the results are close to the Scenario Analysis VaR. We conclude that, the expert opinion used parameters can be uncertain and cause an issue because it can lead to a disruption in the Markovian process. Combining these different data sources, highlights the importance of experts opinion coherence which generate an estimation risk that affects our capital required calculation.

4. Conclusions

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.

Yet, treating the experts outputs, the same as Basel's operational risk Standardized approach, illustrated the necessity of calling attention to the judgments given. In our application, judgments were needed to make sensible choices but these choices will influence the results. Understanding this influence, should be an important aspect of capital calculations, since it created an estimation risk that has highly influenced our capital requirement [18], for the judgment under uncertainty. Moreover, we did not take into consideration external data, which might be interesting to elaborate and apply in practice, more on this subject could be found in [19].

References

- [1] Shevchenko, P.V. (2011) Modelling Operational Risk Using Bayesian Inference. Springer, Berlin. <http://dx.doi.org/10.1007/978-3-642-15923-7>
- [2] Robert, C.P. (2007) The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation second edition. Springer, Berlin.
- [3] Solvency II (2010) Quantitative Impact Study 5 (Final Technical Specifications). <http://www.eiopa.europa.eu>
- [4] Gamonet, J. (2006) Modélisation du risque opérationnel dans l'assurance. Mémoire d'actuaire, CEA.
- [5] Berger, J.O. (1985) Statistical Decision Theory and Bayesian Analysis. 2nd Edition, Springer, New York (NY). <http://dx.doi.org/10.1007/978-1-4757-4286-2>
- [6] Moosa, I.A. (2007) Operational Risk Management. Palgrave Macmillan, Basingstoke. <http://dx.doi.org/10.1057/9780230591486>
- [7] Moosa, I.A. (2008) Quantification of Operational Risk under Basel II: The Good, Bad and Ugly. Palgrave Macmillan, Basingstoke. <http://dx.doi.org/10.1057/9780230595149>
- [8] Santos, C.H., Krats, M. and Munoz, F.V. (2012) Modelling Macroeconomics Effects and Expert Judgements in Operational Risk: A Bayesian Approach. Research Center ESSEC Working Paper 1206.
- [9] Frachot, A., Georges, P. and Roncalli, T. (2001) Loss Distribution Approach for Operational Risk. Groupe de Recherche Opérationnelle, Crédit Lyonnais, France.
- [10] Shevchenko, P.V. and Wüthrich, M.V. (2006) The Structural Modelling of Operational Risk via Bayesian Inference: Combining Loss Data with Expert Opinions. *The Journal of Operational Risk*, **1**, 3-26.
- [11] 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*, **186**, 453-461. <http://dx.doi.org/10.1098/rspa.1946.0056>
- [12] Hastings, W.K. (1970) Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, **57**, 97-109. <http://dx.doi.org/10.1093/biomet/57.1.97>
- [13] Gilks, W.R., Richardson, S. and Spiegelhalter, D. (1996) Markov Chain Monte Carlo in Practice. Chapman & Hall, London.
- [14] David, H.A. and Nagaraja, H.N. (2003) Order Statistics. 3th Edition. John Wiley & Sons Inc., New York. <http://dx.doi.org/10.1002/0471722162>

- [15] Buckland, S.T. (1985) Monte Carlo Confidence Intervals. *Journal of Royal statistical society, Series C (Applied statistics)*, **34**, 296-301.
- [16] Basel Committee on Banking Supervision (2006) International Convergence of Capital Measurement and Capital Standards. BCBS, 144-146. <http://www.bis.org>
- [17] Basel Committee on Banking Supervision (2002) Quantitative Impact Study 3 Technical Guidance. <http://www.bis.org/bcbs/qis/qis3tech.pdf>
- [18] Tversky, A. and Kahneman, D. (1974) Judgment under Uncertainty: Heuristics and Biases. *New Series*, **185**, 1124-1131. <http://dx.doi.org/10.1126/science.185.4157.1124>
- [19] Lambrigger, D.D., Shevchenko, P.V. and Wüthrich, M.V. (2008) Data Combination under Basel II and Solvency 2: Operational Risk Goes Bayesian. *Bulletin Français' Actuariat*, **8**, 4-13.

5. Appendix

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

5.1.1. Jeffreys Prior for Poisson(λ) and Lognormal(μ, σ) Distributions

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

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

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

Let $X \rightarrow \mathcal{L}(\mu, \sigma^2)$, with $f_X(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}$.

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

5.2. MCMC Metropolis-Hastings Algorithm

5.2.1. Applying MCMC with Metropolis Hastings Algorithm for λ

- Initialize $\lambda_0 = \frac{\lambda_{ID} + \lambda_{SA}}{2}$
- Update from λ_i to λ_{i+1} ($i=1, \dots, n$) by
 - Generating $\lambda \rightarrow U(\lambda_{SA}, \lambda_{ID})$
 - Define $\zeta = \min\left(\frac{f(\lambda|n_{SA}, n_{ID})}{f(\lambda_i|n_{SA}, n_{ID})}, 1\right)$
 - Generate $Rnd \rightarrow U(0,1)$
 - If $Rnd \leq \zeta$, $\lambda_{i+1} = \lambda$, else $\lambda_{i+1} = \lambda_i$
- Remove the first 3000 iterations, so that the chain is stationary (burn-in effect).

5.2.2. Applying MCMC with Metropolis-Hastings Algorithm for μ

- Initialize $\mu_0 = \mu_{ID}$
- Update from μ_i to μ_{i+1} ($i=1, \dots, n$) by
 - Generating $\mu \rightarrow U(0,12)$
 - Define $\zeta = \min\left(\frac{f(\mu|x, y)}{f(\mu_i|x, y)}, 1\right)$

- Generate $Rnd \rightarrow U(0,1)$
- If $Rnd \leq \zeta$, $\mu_{i+1} = \mu$, else $\mu_{i+1} = \mu_i$
- Remove the first 3000 iterations, so that the chain is stationary (burn-in effect).

5.2.3. Applying MCMC with Metropolis-Hastings Algorithm for $\omega = (\mu, \sigma)$

- Initialize $\mu_0 = \mu_{ID}$ and $\sigma_0 = \sigma_{ID}$
- Update from μ_i to μ_{i+1} and σ_i to σ_{i+1} , ($i = 1, \dots, n$) by
 - Generating $\mu \rightarrow U(0,12)$ and $\sigma \rightarrow U(0,7)$
 - Define $\zeta = \min\left(\frac{f(\mu, \sigma | x, y)}{f(\mu_i, \sigma_i | x, y)}, 1\right)$
 - Generate $Rnd \rightarrow U(0,1)$
 - If $Rnd \leq \zeta$, $\mu_{i+1} = \mu$ and $\sigma_{i+1} = \sigma$ else $\mu_{i+1} = \mu_i$ and $\sigma_{i+1} = \sigma_i$
- Remove the first 3000 iterations from both distributions, so that the chains is stationary (burn-in effect).