

# 小样本条件下操作风险度量中的参数估计研究

陈倩

(北京第二外国语学院 国际商学院, 北京 100024)

**摘要:** 针对操作风险损失数据匮乏的特点,引入贝叶斯推断理论来保证小样本条件下模型参数估计的质量,将专家经验以先验分布的形式引入,从贝叶斯推断的视角构建了融入专家经验的操作风险损失频率模型、损失强度模型和风险度量模型。其中,损失频率建模中假设损失频率服从 Poisson-Gamma 分布,损失强度建模中应用极值理论对损失分布的“厚尾”特性进行描述,利用广义帕累托分布(GPD)对损失强度分布进行拟合,并打破 GPD 的位置参数和形状参数常数化的限制,视其为随机变量并假定两者服从不同的 Gamma 分布。实证结果证明:与传统的 MLE 方法相比,在小样本的情况下,基于贝叶斯推断的 MCMC 方法对参数的估计更有效和稳定,较好地解决了在操作风险度量中由于损失数据不足给损失分布拟合带来的难题。

**关键词:** 贝叶斯推断; 参数估计; 操作风险度量; 极值理论

**中图分类号:** F831

**文献标识码:** A

**文章编号:** 1009-3370(2015)03-0092-08

伴随中国经济减速和信贷紧缩的压力,国内银行操作风险呈全面上升态势。2011 年的齐鲁银行票据诈骗案、温州银行信贷员伪造房地产证骗贷案、富滇银行倒券风波等以内部欺诈、银行卡诈骗、科技信息风险等为代表的操作风险事件,给银行乃至整个金融业的稳定发展造成了极大威胁,也将银行脆弱的内控体系曝光于公众面前。2013 年 1 月 1 日开始,《商业银行资本管理办法》在中国正式实施。除信用风险和市场风险外,资本监管框架中融入了操作风险,将对商业银行资本管理形成巨大的考验。如何正确地对操作风险展开合理的资本计提,是目前广大商业银行共同关注的焦点。

在操作风险的度量模型和方法中,高级度量法中的损失分布法是目前公认灵敏度最高、度量最准确、应用前景最广泛的方法。通过对操作风险损失数据的分析,de Fontnouvelle 等 (2006)<sup>[1]823</sup> 和 Moosa Imad A(2008)<sup>[2]</sup>认为损失分布法能用操作风险损失频率分布与损失强度分布来较好地描述各风险单元操作风险的大小。由于操作风险数据的匮乏性及其分布的“厚尾”性,在对于操作风险损失强度的拟合研究中面临着许多困难,因此目前的研究大多集中在在以损失分布法为基础的精算模型的改进和扩展上,特别是对损失强度分布的选择与拟合方面。Müller(2002)<sup>[3]</sup>指出 Exponential 分布、Weibull 分布、Gamma 分布和 Lognormal 分布等均能用来对操作

风险的损失强度分布进行拟合。但 de Fontnouvelle 等(2006)<sup>[1]840</sup>的实证研究证明上述几个分布并不能准确地反映出操作风险的尾部特征,往往会造成对风险的低估。为解决这些问题,Chernobai 等(2006)<sup>[4]</sup>提出了使用  $t$  分布、Mixed-Normal 分布模型、Generalized Error 分布和 Pareto 分布等厚尾分布来对操作风险事件的损失强度分布进行拟合。由于在度量风险时关注的重点往往是放在分布的尾部,针对操作损失数据量较少且具有“厚尾”“右偏”等特点,为了弥补用传统损失强度分布度量风险的一些缺陷,Chavez-Demoulin (2006)<sup>[5]</sup>、Cowell (2007)<sup>[6]</sup>和 Franklin(2008)<sup>[7]</sup>等学者相继提出运用极值理论来度量操作风险,与传统分布的拟合结果相比较。极值理论能较好地捕获操作风险的“厚尾”特性,对银行提取的监管资本比较准确,有利于保证经济资本的充分性。在国内,继全登华(2002)<sup>[8]</sup>从理论上分析了如何利用极值理论来计量银行操作风险后,高丽君等(2007)<sup>[9]</sup>、宋加山(2008)<sup>[10]</sup>、司马则茜等(2009)<sup>[11]</sup>、陈倩(2012)<sup>[12]</sup>相继将极值理论运用到操作风险的度量中,并分别对阈值的不同选取方法进行了研究。随后,针对操作风险损失分布的尾部特征,莫建明和周宗放(2008)<sup>[13]</sup>推导出厚尾性操作风险的风险价值和高置信度下厚尾性操作风险价值的标准不确定度及置信区间。司马则茜(2011)<sup>[14]</sup>根据采用具有厚尾特点的  $g-h$  分布度量银行的操作风险。丰吉闯等

收稿日期: 2013-07-02

基金项目: 教育部人文社科青年基金项目(12YJC790013);北京高等学校青年英才计划项目(YETP1522)

作者简介: 陈倩(1982—),女,副教授,管理学博士,E-mail:monicacq@163.com

(2011)<sup>[15]</sup>采用左截尾分布对损失强度分布进行拟合。

由上可知,统计精算模型是建立在损失数据的基础上的,因此损失数据的质量对于操作风险的度量至关重要。与国际银行操作风险管理水平相比,中国对操作风险的管理仍处于初级阶段,大部分银行存在信息披露制度不全、信息技术运用滞后、缺乏统一的数据收集机制和数据平台等问题,给操作风险的量化带来重重障碍:一方面,就“低频高损”损失数据而言,由于事件发生的频率低,样本个数比较少,特别是当确定损失阈值后,超过阈值的损失数据就更为匮乏,数据的不全面、不完整,可能影响到操作风险度量的正确性和准确性,进而影响操作风险管理的合理性和有效性;另一方面,仅基于历史数据的度量结果,过分依赖于数据,特别当数据质量不高时,其度量结果的准确性将大打折扣。因此操作风险数据不足成为操作风险度量瓶颈,银行传统的风险管理理念正面临着前所未有的挑战,如何在小样本条件下保证参数估计的稳定性,实现操作风险资本的合理计提,成为每个商业银行乃至整个金融界亟待解决的重要问题。

针对以上两个问题,本文对小样本条件下度量操作风险过程中的参数估计问题展开研究。为了保证参数估计的稳定性,引入贝叶斯推断理论,在操作风险的度量中将专家的经验 and 意见以先验信息的形式引入,利用贝叶斯推断将专家的经验知识同银行内部实际的操作风险损失数据加以综合,对历史数据、专家经验和样本信息进行有效地整合,避免数据匮乏对参数估计带来的障碍。也有一些学者在操作风险的度量中采用贝叶斯估计方法,如 Dalla Valle (2008<sup>[16]</sup>, 2009<sup>[17]</sup>)、高丽君 (2012)<sup>[18]</sup>、冉丽宏 (2012)<sup>[19]</sup>应用贝叶斯来度量操作风险,对操作风险的损失频率和损失强度的分布函数进行了估计。但目前学者在操作风险损失强度的拟合中,其所选择的损失分布通常是具有共轭先验分布的,如正态分布、指数分布、对数正态分布等。而本文将利用极值理论对损失强度分布进行建模,其损失强度分布广义帕累托分布 (Generalized Pareto Distribution, GPD) 没有特定的先验共轭分布,这给在操作风险度量中利用极值理论和贝叶斯推断带来了困难和挑战。本文将打破 GPD 的位置参数和形状参数常数化的限制,视其为随机变量并假定两者服从不同的 Gamma 分布,结合贝叶斯推断理论分析和推导其后验分布,构建融入贝叶斯推断的损失频率模型和损失强度模型。在实证分析中,利用 MCMC 的 Gibbs 方法抽取操作风险总损失分布的样本,解决高维复杂积

分的问题,计算出基于内部数据和专家信息的操作风险大小,并将基于贝叶斯推断的 MCMC 方法的参数估计结果和传统 MLE 方法的结果进行对比,以解决由数据不足所引起的参数估计误差较大的问题。

## 一、贝叶斯推断的基本原理及在操作风险度量中的应用

贝叶斯推断统计方法的基本观点是<sup>[20]</sup>:首先,任何一个总体分布的未知参数  $\theta$  都可以看作是一个随机变量,应该用一个概率分布来描述对  $\theta$  的未知状况;其次,分布参数  $\theta$  具有先验分布,即在进行抽样之前就有关于  $\theta$  的先验信息的概率分布  $\pi(\theta)$ ;再次,对于分布参数的任何统计推断,必须依据  $\theta$  的后验分布来进行,这是因为在获得样本的观察值之后,后验分布包含了对分布参数  $\theta$  的全部信息。

在操作风险的度量中,设需要估计的参数为  $\theta$ , 具有密度为  $\pi(\theta)$  的先验分布,当参数  $\theta$  已知时,如果操作风险损失数据  $\mathbf{x}=(x_1, x_2, \dots, x_n)^T$  是由独立同分布的随机变量组成的,则此时损失的分布为

$$p(\mathbf{x}|\theta)=p(x_1|\theta)p(x_2|\theta)\cdots p(x_n|\theta) \quad (1)$$

在多元统计的概念下,则关于  $\mathbf{x}$  和参数  $\theta$  的联合分布概率密度函数  $p(\mathbf{x}, \theta)$  为

$$p(\mathbf{x}, \theta)=p(\mathbf{x}|\theta)\pi(\theta) \quad (2)$$

$\mathbf{x}$  的边缘分布概率密度  $p(\mathbf{x})$  为

$$p(\mathbf{x})=\int p(\mathbf{x}|\theta)\pi(\theta)d\theta \quad (3)$$

式(3)描述的是连续分布的情况,若为离散分布时,应该用求和号来替换积分号。

收集到操作风险的损失数据  $\mathbf{x}$  后,可以形成对  $\theta$  的新认识,则在  $\mathbf{x}$  条件下  $\theta$  的条件概率密度为

$$\pi(\theta|\mathbf{x})=\frac{p(\mathbf{x}, \theta)}{p(\mathbf{x})}=\frac{p(\mathbf{x}, \theta)\pi(\theta)}{\int p(\mathbf{x}|\theta)\pi(\theta)d\theta} \quad (4)$$

式(4)其为  $\theta$  的后验分布,它包含了总体、样本和先验三种信息中与  $\theta$  有关的一切信息,它是在给定数据后所得到的关于参数的条件概率分布。

式(4)中的分母为一个积分,故其最终为一个不依赖于  $\theta$  的常数,称为归一化常数。因此在应用中一般只关心式(4)的分子部分,将式(4)变形为

$$\pi(\theta|\mathbf{x})\propto p(\theta|\mathbf{x})\pi(\theta) \quad (5)$$

从式(5)可以看出,  $\theta$  的后验分布  $\pi(\theta|\mathbf{x})$  是与似然函数  $p(\mathbf{x}|\theta)$  和先验分布  $\pi(\theta)$  的乘积成比例,可以解释为当从操作风险损失数据总体获得样本  $\mathbf{x}$  后,利用贝叶斯推断可以把人们对  $\theta$  的认识从先验分布  $\pi(\theta)$  调整到后验分布  $\pi(\theta|\mathbf{x})$ 。

若后验分布  $\pi(\theta|\mathbf{x})$  与先验分布  $\pi(\theta)$  有相同的

分布类型(只是参数不同时),称先验分布  $\pi(\theta)$  为  $\theta$  因此共轭先验分布在贝叶斯推断中的应用较为广泛。常见的共轭先验分布如表 1 所示。

表 1 常见的共轭先验分布

总体分布	参数	共轭先验分布
Normal 分布(方差已知)	均值	Normal 分布
Normal 分布(均值已知)	方差	Inverted Gamma 分布
Binomial 分布	成功概率	Beta 分布
Negative Binomial 分布	成功概率	Beta 分布
Poisson 分布	均值	Gamma 分布
Exponential 分布	均值的倒数	Gamma 分布

在后验分布的基础上,进一步得到当操作风险数据  $\mathbf{x}=(x_1, x_2, \dots, x_n)^T$  给定时,下一个新的损失数据  $y$  的条件概率分布  $p(y|\mathbf{x})$ ,也称为预测分布

$$p(y|\mathbf{x}) = \int p(y|\theta)\pi(\theta|\mathbf{x})d\theta \quad (6)$$

其中,  $p(y|\theta)$  是给定参数  $\theta$  后,一定新观测的概率密度函数。

为了达到预测的目的,通常关注的预测分布的均值,可以看作是给定  $n$  个观测和先验分布后对第  $n+1$  个观测的估计,即

$$E[Y|\mathbf{x}] = \int yp(y|\mathbf{x})dy = \int y \int p(y|\theta)\pi(\theta|\mathbf{x})d\theta dy = \int \pi(\theta|\mathbf{x}) \int yp(y|\theta)dy d\theta = \int E[Y|\theta]\pi(\theta|\mathbf{x})d\theta \quad (7)$$

## 二、基于贝叶斯推断的商业银行操作风险度量模型

### (一) 融入专家经验的商业银行操作风险损失频率模型

假设某银行每年的操作风险发生的次数服从参数为  $\lambda$  的 Poisson 分布,参数  $\lambda$  为连续随机变量,具体数值未知。专家根据经验指出  $\lambda$  服从参数分别为  $\alpha, \beta$  的 Gamma 分布,记作  $\text{Gamma}(\alpha, \beta)$ ,根据贝叶斯推断可知,此处的  $\text{Gamma}(\alpha, \beta)$  就是参数  $\lambda$  的先验分布,其密度函数为

$$\pi(\lambda) = \frac{\lambda^{\alpha-1}e^{-\lambda\beta}}{\Gamma(\alpha)\beta^\alpha} \quad (8)$$

那么在给定参数  $\lambda=\theta$  的条件下,2008 年该银行操作风险发生的次数为  $x$  的概率为

$$p(x|\theta) = \frac{e^{-\theta}\theta^x}{x!} \quad (9)$$

银行操作风险发生次数  $X$  与参数  $\theta$  的联合分布为

$$p(x, \theta) = p(x|\theta) \cdot \pi(\theta) = \frac{e^{-\theta}\theta^x}{x!} \cdot \frac{\theta^{\alpha-1}e^{-\theta\beta}}{\Gamma(\alpha)\beta^\alpha} = \frac{\theta^{x+\alpha-1}e^{-(1+\beta)\theta}}{x! \Gamma(\alpha)\beta^\alpha} \quad (10)$$

银行操作风险发生次数  $X$  的边缘分布为

$$p(x) = \int_{-\infty}^{+\infty} p(x|\theta) \cdot \pi(\theta) d\theta = \int_0^{+\infty} \frac{\theta^{x+\alpha-1}e^{-(1+\beta)\theta}}{x! \Gamma(\alpha)\beta^\alpha} d\theta = \frac{\Gamma(\alpha+x)}{x! \Gamma(\alpha)\beta^\alpha(1+\beta)^{x+\alpha}} = \binom{x+\alpha-1}{x} \left(\frac{1}{1+\beta}\right)^\alpha \left(\frac{\beta}{1+\beta}\right)^x \quad (11)$$

从式(11)可以看出,此时边缘分布为 Negative Binomial 分布,其参数分别为  $\alpha$  和  $\beta$ 。

根据贝叶斯推断公式,  $\lambda$  的后验密度函数为

$$\pi(\theta|\mathbf{x}) = \frac{p(\mathbf{x}, \theta)}{p(\mathbf{x})} = \frac{\theta^{x+\alpha-1}e^{-(1+\beta)\theta}}{x! \Gamma(\alpha)\beta^\alpha} / \frac{\Gamma(\alpha+x)}{x! \Gamma(\alpha)\beta^\alpha(1+\beta)^{x+\alpha}} = \frac{\theta^{x+\alpha-1}e^{-(1+\beta)\theta}(1+\beta)^{x+\alpha}}{\Gamma(\alpha+x)} \quad (12)$$

从后验的表达(12)可以看出,此时  $\lambda$  的后验分布为 Gamma 分布,其形状参数为  $x+\alpha$ ,尺度参数为  $(1+\beta)^{-1}=\beta/(1+\beta)$ 。

进一步,打破银行只有当年操作风险损失数据的限制,假设银行现有多年的操作风险损失数据,在 2008 年银行操作风险发生的次数为  $x_1, x_2, \dots, x_n$ ,沿用上面单观测样本的计算流程,在扩充样本下,即在拥有多个观测样本也就是多年的操作风险损失次数的条件下,可以对的后验分布和下一年的期望损失次数重新进行调整。

多观测样本与单观测样本的主要区别是在于在给定参数  $\lambda=\theta$  的条件下,银行每年操作风险发生的次数  $X$  服从参数为  $\theta$  的 Poisson 分布表达形式不同,即在多观测样本条件下,式(9)变形为

$$p(\mathbf{x}|\theta) = \prod_{i=1}^n \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{x_1+x_2+\dots+x_n}}{x_1! x_2! \dots x_n!} = \frac{e^{-n\theta}\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \quad (13)$$

银行操作风险发生次数  $X$  与参数  $\theta$  的联合分布为

$$p(\mathbf{x}, \theta) = p(\mathbf{x}|\theta)\pi(\theta) = \frac{e^{-n\theta}\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \cdot \frac{\theta^{\alpha-1}e^{-\theta\beta}}{\Gamma(\alpha)\beta^\alpha} \quad (14)$$



银行操作风险发生次数  $X$  的边缘分布为

$$p(x) = \int_{-\infty}^{+\infty} p(x|\theta) \cdot \pi(\theta) d\theta = \int_0^{+\infty} \frac{e^{-n\theta} \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \cdot \frac{\theta^{\alpha-1} e^{-\theta\beta}}{\Gamma(\alpha)\beta^\alpha} d\theta = \frac{\Gamma(\alpha + \sum_{i=1}^n x_i)}{\prod_{i=1}^n x_i! \cdot \Gamma(\alpha)\beta^\alpha (1+1/\beta)^{\sum_{i=1}^n x_i + \alpha}} \quad (15)$$

于是根据贝叶斯推断公式,  $\lambda$  的后验密度函数为

$$\pi(\theta|x) = \frac{p(x, \theta)}{p(x)} = \frac{e^{-n\theta} \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \cdot \frac{\theta^{\alpha-1} e^{-\theta\beta}}{\Gamma(\alpha)\beta^\alpha} / \frac{\Gamma(\alpha + \sum_{i=1}^n x_i)}{\prod_{i=1}^n x_i! \cdot \Gamma(\alpha)\beta^\alpha (1+1/\beta)^{\sum_{i=1}^n x_i + \alpha}} = \frac{\theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+1)\beta\theta} (1+1/\beta)^{\sum_{i=1}^n x_i + \alpha}}{\Gamma(\alpha + \sum_{i=1}^n x_i)} \quad (16)$$

从后验分布的表达式式(16)可以看出, 此时  $\lambda$  的后验分布仍为 Gamma 分布, 其形状参数为  $\sum_{i=1}^n x_i + \alpha = n\bar{x} + \alpha$ , 尺度参数为  $(n+1/\beta) - 1 = \beta/(1+n\beta)$ , 其中  $\bar{x}$  为  $n$  年操作风险发生次数的均值, 即  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 。

则在过去  $n$  年中银行操作风险发生的次数为  $x_1, x_2, \dots, x_n$  的假设下, 利用贝叶斯推断可推理得出下一年的银行操作风险发生次数的预测分布为

$$p(y|x) = \int_{-\infty}^{+\infty} p(y|\theta) \cdot \pi(\theta|x) d\theta = \int_0^{+\infty} \frac{e^{-\theta} \theta^y}{y!} \cdot \frac{\theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+1)\beta\theta} (1+1/\beta)^{\sum_{i=1}^n x_i + \alpha}}{\Gamma(\alpha + \sum_{i=1}^n x_i)} d\theta = \frac{(1+1/\beta)^{\sum_{i=1}^n x_i + \alpha}}{y! \Gamma(\alpha + \sum_{i=1}^n x_i)} \int_0^{+\infty} \theta^{y + \sum_{i=1}^n x_i + \alpha - 1} e^{-(1+n+1/\beta)\theta} d\theta = \frac{(1+1/\beta)^{\sum_{i=1}^n x_i + \alpha} \Gamma(y + \alpha + \sum_{i=1}^n x_i)}{y! \Gamma(\alpha + \sum_{i=1}^n x_i) (1+n+1/\beta)^{y + \sum_{i=1}^n x_i + \alpha}} \quad (17)$$

此时银行操作风险发生次数的预测分布为 Negative Binomial 分布, 其参数为  $n\bar{x} + \alpha, \beta/(1+n\beta)$ , 因此可知下一年银行操作风险发生的期望次数为  $(n\bar{x} + \alpha)\beta/(1+n\beta)$ 。

通过对预测条件分布求期望能得出相同结论

$$E[Y|x] = \int_{-\infty}^{+\infty} yp(y|x) dy = \int_{-\infty}^{+\infty} E[y|\theta] \pi(\theta|x) d\theta = \int_{-\infty}^{+\infty} \theta$$

$$\frac{\theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+1)\beta\theta} (1+1/\beta)^{\sum_{i=1}^n x_i + \alpha}}{\Gamma(\alpha + \sum_{i=1}^n x_i)} d\theta = \frac{(\sum_{i=1}^n x_i + \alpha)\beta}{1+n\beta} \quad (18)$$

从式(18)可知, 银行在下一年发生操作风险的次数  $y$  的期望与损失频率服从参数为  $n\bar{x} + \alpha, \beta/(1+n\beta)$  的 Negative Binomial 分布的均值一致, 将式(18)整理后得到

$$E[Y|x] = \frac{(\sum_{i=1}^n x_i + \alpha)\beta}{1+n\beta} = \frac{\beta \sum_{i=1}^n x_i}{1+n\beta} + \frac{\alpha\beta}{1+n\beta} = \frac{n\beta}{1+n\beta} \bar{x} + \frac{1}{1+n\beta} \alpha\beta = \frac{n\beta}{1+n\beta} \bar{x} + (1 - \frac{n\beta}{1+n\beta}) \alpha\beta \quad (19)$$

分析式(19)可知,  $\bar{x}$  为银行前  $n$  年所发生风险事件次数的平均值,  $\alpha\beta$  为 Gamma( $\alpha, \beta$ ) 分布均值, 且 Gamma( $\alpha, \beta$ ) 分布为专家给出的先验分布; 故银行在下一年发生操作风险的期望次数可以表示为历史数据的均值与先验分布均值的加权平均。当银行的历史损失数据增加时, 更多的权重将放在历史损失数据上, 而先验分布的权重相应减小。同时, 随着  $\beta$  的增大, 先验分布的方差也随之增加, 即专家的主观经验的不确定性也随之增大, 此时历史数据对发生次数期望的影响大于专家的主观经验。

### (二) 融入专家经验的商业银行操作风险损失强度模型

操作风险的损失分布具有明显“厚尾”的特性, 为此我们选择极值理论来对操作风险进行度量, 并假定损失强度分布为 GPD 进行拟合。GPD 的分布函数为

$$G_{\xi, \sigma}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-1/\xi}, & \xi \neq 0 \\ 1 - e^{-x/\sigma}, & \xi = 0 \end{cases} \quad (20)$$

其中,  $\sigma$  为分布的尺度参数;  $\xi$  为形状参数, 用以描述分布尾部的消失速度。当  $\xi > 0$  时,  $G_{\xi, \sigma}(x)$  为厚尾分布, 如 Pareto 分布、Loggamma 分布和 Cauchy 分布等。在操作风险的实证研究中, 大部分操作风险的损失数据的分布都属于这类厚尾分布, 因此本文仅考虑 GPD 的形状参数  $\xi > 0$  的情形, 对其他两种情形也可按照下面思路进行分析, 此时 GPD 的密度函数为

$$f(x) = \frac{1}{\sigma} \left(1 + \frac{\xi x}{\sigma}\right)^{-1/\xi - 1}, \quad \xi > 0 \quad (21)$$

令  $\delta = \sigma/\xi$ , 当  $\xi > 0$  时, GPD 的累计分布函数和概率密度函数分别为

$$F(x|\xi, \delta) = 1 - (1 - x/\delta)^{1/\xi} \quad (22)$$

$$f(x|\xi, \delta) = \frac{1}{\xi\delta} \left(1 + \frac{x}{\xi}\right)^{-1/\xi - 1} \quad (23)$$

当有多个超额样本数据  $x_1, x_2, \dots, x_n$  时,可知样本的联合条件密度函数即似然函数为

$$p(\mathbf{x}|\xi, \delta) = p(x_1, x_2, \dots, x_n|\xi, \delta) = \frac{1}{\xi^{a_1}} \frac{1}{\delta^{a_2}} \prod_{i=1}^n \left(1 + \frac{x_i}{\delta}\right)^{-1/\xi-1} \quad (24)$$

关于对  $\xi$  和  $\sigma$  的先验分布的选取,国内、外学者提出许多有效、可行的分布,如  $\xi$  服从正态分布,  $\sigma$  服从逆 Gamma 分布或  $\xi$  服从 Pareto 分布、 $\sigma$  的倒数服从 Gamma 分布等<sup>[20]</sup>。在此,本文为 GPD 参数选用较广泛的先验分布形式<sup>[21]</sup>,即  $\xi$  服从参数为  $a_1$  和  $b_1$  的 Gamma 分布,记作  $\xi \sim \text{Gamma}(a_1, b_1)$ ,  $\delta$  服从参数为  $a_2$  和  $b_2$  的 Gamma 分布,记作  $\delta \sim \text{Gamma}(a_2, b_2)$ 。根据贝叶斯推断可知,此处的两个 Gamma 分布就分别为参数  $\xi$  和  $\delta$  的先验分布,其密度函数分别为

$$\pi(\xi) = \frac{\xi^{a_1-1} e^{-\xi/b_1}}{\Gamma(a_1)b_1^{a_1}}, \quad \pi(\delta) = \frac{\delta^{a_2-1} e^{-\delta/b_2}}{\Gamma(a_2)b_2^{a_2}} \quad (25)$$

根据贝叶斯法则,参数  $\xi$  和  $\delta$  的后验分布可以表示为

$$\pi(\xi, \delta|\mathbf{x}) = \frac{p(\mathbf{x}, \xi, \delta)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\xi, \delta)\pi(\xi)\pi(\delta)}{\iint p(\mathbf{x}|\xi, \delta)\pi(\xi)\pi(\delta)d\xi d\delta} \quad (26)$$

由于式(26)中分母为一固定常数项,将其忽略后,参数  $\xi$  和  $\delta$  的后验分布与损失大小的似然函数、参数  $\xi$  和  $\delta$  的先验分布成一定比例,如式(27)所示

$$\pi(\xi, \delta|\mathbf{x}) \propto p(\mathbf{x}|\xi, \delta)\pi(\xi)\pi(\delta) \quad (27)$$

将损失大小的似然函数、参数  $\xi$  和  $\delta$  的先验分布的概率密度表达式带入式(27)并整理得到

$$\pi(\xi, \delta|\mathbf{x}) \propto \xi^{-n+a_1-1} \delta^{-n+a_2-1} \exp[-b_1\xi - b_2\delta + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^n \ln\left(1 - \frac{x_i}{\delta}\right)] \quad (28)$$

据此,已经分别构建了融合专家信息的操作风险损失强度模型和损失频率模型。进一步对于操作风险的  $VaR_\alpha$  值可以表示为

$$VaR_\alpha = F^{-1}(1-\alpha) = u + \frac{\sigma}{\xi} \left\{ \left( \frac{1-\alpha}{\lambda} \right)^{-\xi} - 1 \right\} \quad (29)$$

在一定置信水平  $a$  下,操作风险损失超过 VaR 时的条件期望——期望短缺(Expect Shortfall, ES),可以表示为

$$ES_\alpha = E[X|X > VaR_\alpha] \quad (30)$$

将参数贝叶斯估计量带入操作风险度量模型式(28)和式(30)中,可分别得到操作风险事件尾部风险的 VaR 和 ES,进而可以得到在内部数据中融入专家经验后的银行操作风险和监管资本。

### 三、实证分析

#### (一)数据来源及整体特征分析

本文仿照国内外学者的做法<sup>[4-10]</sup>,通过报纸、网络等媒体的公开报道、监管部门文件、法院检察院等审理的公开文件等渠道来搜集中国商业银行操作风险的损失案例,搜集到 1994—2012 年共 1 065 个案例,考虑到操作风险自身具有一定的隐藏性,且从发生到被发现往往具有一定的滞后期,由于未及时发现或延迟报道等事宜的存在,往往会低估操作风险损失强度和频率,为了降低估计误差,放弃 2010—2012 年的损失案例,最终搜集 1994—2009 年共 576 个操作风险损失案例,其最大损失为 39.94 亿,最小为 2 500 元。考虑到真实情况下银行收集数据存在一个特定的门槛值,我们将所收集到的小于 10 万元的数据舍去,最后得到 246 个损失数据。

为了对操作风险的损失大小有整体上的了解,用一般描述性统计来对操作风险的损失大小的统计特征进行了分析,结果如表 2 所示。

表 2 操作风险损失强度统计特征

案件数/件	均值/万元	标准差/万元	中位数/万元	偏度	峰度
246	20 828.84	52 788.28	2 283.5	4.159	19.911

从描述统计结果可知,这 246 个风险案件所造成的平均损失为 20 828.84 万元,与正态分布相比,实际操作风险损失数据的分布具有更尖的峰(19.911>3)。从偏度系数可知实际损失分布不服从对称分布,而是明显右偏态分布(4.159>0),这也印证了操作风险损失具有“右偏、厚尾”的特性。

#### (二)损失强度和损失频率分布参数估计

运用上文所构建的融入专家信息后的随机模拟模型,结合贝叶斯推断理论和 MCMC 方法,本文借助 WinBUGS 软件,对商业银行操作风险的损失

频率分布和基于极值理论的 GPD 分布的参数进行估计,为最终操作风险的度量打下基础。由于 WinBUGS 软件包中并未定义 GPD 的分布函数,故在估计损失强度分布的参数前,我们使用“one tricks”技巧来解决此问题。在对后验分布的参数估计过程中,Markov Chain 具有收敛性是使用 MCMC 的前提和关键。为了验证 Markov Chain 的收敛性,本文设置了 2 条具有不同初始值的 Markov Chain,分别对上述所构建的融入专家信息后的操作风险度量模型进行抽样模拟。图 1 分别

展示估计参数  $l$ 、 $x$  和  $s$  时的 Markov Chain 抽样路径。

从图 1 中 3 个参数的 Markov Chain 抽样路径图可以看出,3 个参数的两条 Markov Chain 重合在

一起,参数样本围绕中心线上下波动,说明参数样本值较为平稳,因此可以认为 Markov Chain 是收敛的,可以运用抽样的办法来得到参数的后验分布,继而实现对参数的估计。

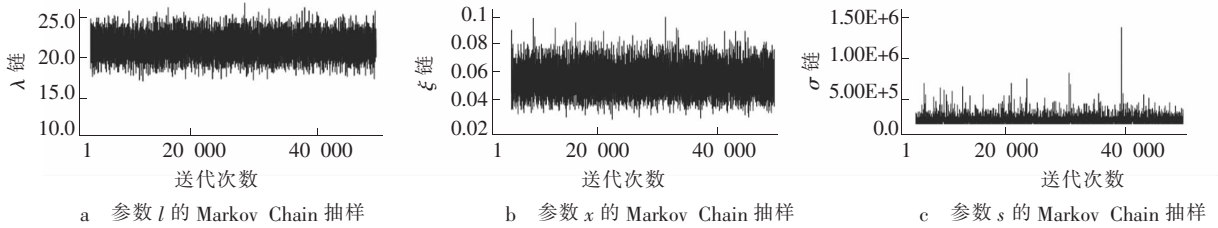


图 1 不同参数的 Markov Chain 抽样路径图

对所有的参数都进行了 50 000 连续抽样,并将初始的 2 000 个样本设置为燃烧期,即舍去前 2 000 个样本。同时为了减少参数的自相关性的影响,保证模拟的结果具有随机分布的特性,在进行随机模拟时,设定每个 5 个模拟数据中抽取 1 个,因此共抽取了 9 600 个样本来对分布的参数进行估计,参数的核密度图如图 2 所示,表 3 列出了所估计参数的统计结果。

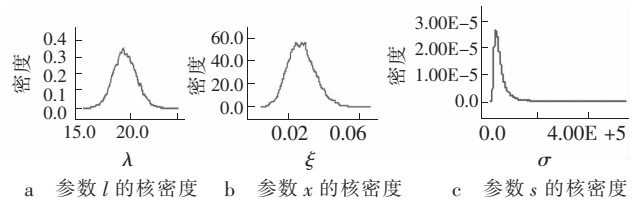


图 2 不同参数的核密度图

从表 3 中可以看出,MC Error 比较小,说明 MCMC 方法对参数估计的结果较好。

表 3 MCMC 法参数估计结果

参数名称	均值	标准差	MC error	2.5%分位数	中位数	97.5%分位数
$l$	18.920 000	1.216 000	0.007 317	16.600 000	18.870 000	21.370 000
$x$	0.04 7390	0.007 773	0.000 069	0.033 660	0.046 960	0.064 100
$s$	43 383	32 423	292	14 264	34 965	122 208

(三)基于贝叶斯推断的操作风险大小

将表 3 中的参数估计结果分别代入式 (29)和

式(30),即可得到融合专家意见和内部数据的操作风险的大小,结果如表 4 所示。

表 4 融入专家意见后的操作风险的尾部风险值

VaR 95%	VaR 99%	VaR 99.9%	ES 95%	ES 99%	ES 99.9%
281 768.83	452 909.02	725 121.92	389 360.28	570 269.57	85 021.06

(四)小样本条件下 MCMC 方法与传统 MLE 方法估计结果稳定性对比

为了更好地比较 MCMC 方法与 MLE 方法,将 MLE 方法所估计的参数值及 95%的置信区间(以 2.5%和 97.5%分位数表示)在表中 5 列出。

从表 5 中可看出,用 MCMC 方法和 MLE 方法所估计的参数值大小相差不大,但从参数估计的 95%的置信区间来看,较多情况下 MCMC 方法的置信区间比 MLE 方法估计的置信区间窄,说明在小样本条件下,考虑参数的先验信息能提高参数估计的精确度。图 3 也直观地展示了 POT 模型中两个参数

的置信区间。

图 3 中间的粗实线表示模拟过程中参数的均值,上下 2 条虚线分别代表 2.5%和 97.5%的分位数。3 条线都基本为直线,说明所采用的抽样方法比较稳定,所模拟产生的参数值的波动性不大。

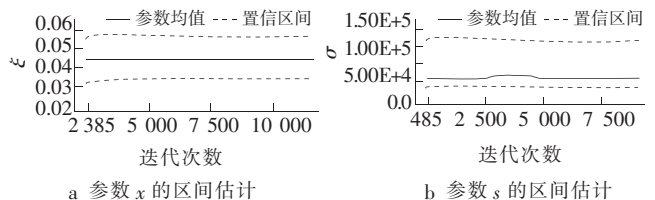


图 3 不同参数的区间估计图

表 5 MLE 法和 MCMC 法估计结果的比较

参数名称	MLE			MCMC		
	均值	2.5%分位数	97.5%分位数	均值	2.5%分位数	97.5%分位数
$l$	18.92	16.50	21.37	18.92	16.56	21.29
$x$	0.054 000	-0.378 000	0.388 800	0.047 390	0.030 366	0.064 100
$s$	87 646	53 534	143 500	43 383	14 264	122 208



在使用 Gibbs 抽样时,为了保证所抽取的参数具有随机分布的特性,我们采取每 5 个样本选取 1 个的措施。为了验证所选取的参数是否确实具备随机性,对其进行了自相关性检验,检验结果如图 4 所示。

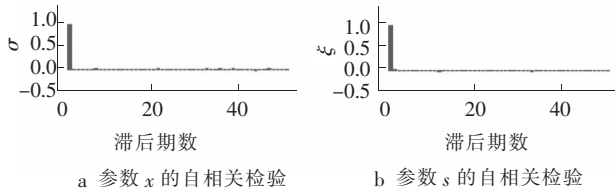


图 4 不同参数的自相关检验图

从图 4 中可以看出,除时滞为 0 处,其他各时滞状态下参数的自相关值均接近于 0,说明所模拟的参数的自相关性已经基本消除,保证了模拟结果的随机性。

随后,为了检验 MCMC 方法在小样本数下对参数估计的效果,逐步提升极值理论的阈值,并在各新的阈值和超额样本数下,分别使用 MLE 和 MCMC 方法来对 POT 模型的参数和 95% 的置信区间进行估计,其结果如表 6 所示。

在表 6 中,除了选取 2 亿元作为阈值外,分别

表 6 不同阈值下 MLE 和 MCMC 对 POT 模型的参数估计结果

阈值	超额样本数	参数	MLE			MCMC		
			均值	2.5%分位数	97.5%分位数	均值	2.5%分位数	97.5%分位数
2 亿	40	x	0.054 00	-0.378 00	0.388 80	0.047 39	0.033 66	0.064 10
		s	87 646	53 534	143 500	43 383	14 264	122 208
2.5 亿	39	x	0.061 00	-0.362 60	0.484 60	0.047 80	0.033 56	0.064 78
		s	80 191	47 350	135 810	44 224	14 815	126 389
3 亿 (15%)	35	x	-0.015 60	-0.432 30	0.401 20	0.047 50	0.032 67	0.065 29
		s	91 021	53 450	155 000	46 820	14 769	133 204
5 亿 (10%)	26	x	-0.153 80	-0.589 40	0.281 90	0.046 58	0.030 02	0.066 39
		s	11 218	62 940	199 950	57 954	16 935	178 603
11.5 亿 (5%)	13	x	-0.343 40	-0.986 50	0.299 70	0.046 24	0.024 45	0.076 17
		s	132 550	5 832	301 290	79 876	18 752	55 896

选取了 2.5 亿、3 亿、5 亿和 11.5 亿元作为新的阈值,其括号内的百分数代表大于阈值的损失个数占总损失个数的百分比。第 2 列是不同阈值下的超额样本数,随着阈值的增大,超额样本的个数也随之减少。从参数估计的结果可知,使用 MLE 方法对参数进行估计时,随之阈值的增大,GPD 分布的形状参数从正变为负,说明操作风险损失分布的尾部不再具有“厚尾”特性,这显然与现实不相吻合。究其原因是由于样本个数太少,从而在参数估计中产生较大的误差,因此当研究样本为小样本数据时,使用 MLE 方法进行估计其结果波动很大,传统方法对参数的估计变得很不稳定。而相比较而言,MCMC 方法中参数估计的结果波动较小,其参数估计结果的稳定性不受数据个数多少的影响,这也说明 MCMC 方法能有效解决样本数据不足的问题,其对参数的估计比 MLE 方法更稳定和可靠。

#### 四、结论

在操作风险建模过程中,各种模型的适用性和精确度往往受操作风险损失数据量的影响,操作风险损失数据的小样本特性制约着统计模型的建立的应用。在小样本损失数据的条件下,忽略数据匮乏的特性,盲目将国外先进的建模技术直接套用到

中国商业银行操作风险的度量中,在风险的度量中将可能产生由于模型应用不当而产生的模型风险。因此,针对在操作风险度量中用小样本进行分布拟合所存在的参数估计的困难,在银行内部损失数据的基础上,本文首先利用贝叶斯推断来实现内部数据与专家的经验 and 意见等信息的有效融合,建立了融入专家先验信息的度量操作风险的贝叶斯模型;其次,为解决度量中的高维复杂积分问题,在实证分析利用 MCMC 的 Gibbs 方法抽取操作风险总损失分布的样本,计算出基于内部数据和专家信息的操作风险大小。最后,为判断 MCMC 估计的效果和精度,通过实证分析将参数估计的结果与 MLE 方法的结果进行对比。特别需要指出的是,为了验证数据的小样本特性对 POT 模型参数估计的影响,将在不同阈值下使用 MLE 和 MCMC 方法估计 GPD 分布参数的结果进行了对比。实证结果表明:当研究样本为小样本数据时,使用传统方法 MLE 对参数估计结果波动很大,稳定性较差。而相比较而言,基于贝叶斯推断的 MCMC 方法中参数估计的结果波动较小,其参数估计结果的稳定性不受数据个数多少的影响,在小样本条件下仍能得到比较准确和稳定的估计结果,比传统方法更稳定、可靠。因此,基于贝叶斯推断理论的操作风险度量模型,可以在有限

的损失样本容量下,结合损失分布的先验信息,在小样本条件下实现对损失分布参数的有效估计,有效地解决了操作风险度量中的损失数据匮乏对模

型参数估计影响,在一定程度上弥补了操作风险损失数据匮乏对参数估计影响较大的不足。

#### 参考文献:

- [1] De Fontnouvelle Patrick, Dejesus-Rueff Virginia, Jordan John S, Rosengren Eric S. Capital and risk; new evidence on implications of large operational losses[J]. Journal of Money, Credit & Banking, 2006, 38(7): 1819-1846.
- [2] Moosa Imad A. A critique of the advanced measurement approach to regulatory capital against operational risk[J]. Journal of Banking Regulation, 2008, 9(3): 151-164.
- [3] Müller H. Quantifying operational risk in a financial institution[D]. Universität Karlsruhe, 2002.
- [4] Chernobai A, Rachev S. Applying robust methods to operational risk modeling[J]. Journal of Operational Risk, 2006(1): 27-41.
- [5] Chavez-Demoulin V, Embrechts P, Neslehova J. Quantitative models for operational risks: extremes, dependence and aggregation[J]. Journal of Banking and Finance, 2006(30): 2635-2658.
- [6] Cowell R G, Verrall R J, Yoon Y K. Modeling operational risk with bayesian networks[J]. Journal of Risk & Insurance, 2007, 74(4): 795-827.
- [7] Franklin James. Operational risk under Basel II: a model for extreme risk evaluation[J]. Banking & Financial Services Policy Report, 2008, 27(10): 10-16.
- [8] 全登华. 利用极值理论计量银行操作风险[J]. 统计与决策, 2002(3): 46.
- [9] 高丽君, 李建平, 徐伟宣, 王书平. 基于 POT 方法的商业银行操作风险极端值估计[J]. 运筹与管理, 2007(1): 1-6.
- [10] 宋加山. 基于极值理论的我国商业银行操作风险度量[D]. 北京: 中国科技大学, 2008.
- [11] 司马则茜, 蔡晨, 李建平. 度量银行操作风险的 POT 幂律模型及其应用[J]. 中国管理科学, 2009, 17(1): 36-41.
- [12] 陈倩. 基于极值理论的商业银行操作风险度量研究[J]. 中国管理科学, 2012, 20: 332-339.
- [13] 莫建明, 周宗放. 重尾性操作风险监控参数识别[J]. 系统工程, 2008, 26(8): 65-70.
- [14] 司马则茜, 蔡晨, 李建平. 基于 g-h 分布度量银行操作风险[J]. 系统工程理论与实践, 2011, 3(12): 2321-2327.
- [15] 丰吉闯, 李建平, 陈建明. 基于左截尾数据的损失分布法度量操作风险: 以中国商业银行为例[J]. 管理评论, 2011, 23(7): 171-176.
- [16] Dalla Valle L, Giudici P. A Bayesian approach to estimate the marginal loss distributions in operational risk management[J]. Computational Statistics & Data Analysis, 2008(52): 3107-3127.
- [17] Dalla Valle. Bayesian copula distributions, with application to operational risk management[J]. Methodology and Computing in Applied Probability, 2009, 11(1): 95-115.
- [18] 高丽君, 丰吉闯. 基于变位置参数贝叶斯预测银行内部欺诈研究[J]. 中国管理科学, 2012, 20(4): 20-25.
- [19] 冉丽宏. 基于 Bayese-Buhlmann 方法的操作风险度量研究[D]. 山西: 山西财经大学, 2012.
- [20] 钟波, 汪青松. 基于 Bayes 估计的金融风险值——VaR 计算[J]. 数理统计与管理, 2007, 26(5): 881-886.
- [21] Claudia Czado. Einführung zu markov chain monte carlo verfahren mit anwendung auf gesamtschadenmodelle[R]. Germany: Technical University of Munich, 2006.

## Research on Parameter Estimation for Operational Risk Measurement based on Small Sampling Data

CHEN Qian

(International Business School, Beijing International Studies University, Beijing 100024, China)

**Abstract:** To ensure the parameter estimation quality based on small sampling data, Bayesian inference theory was applied. Taking experts' experience and opinion as the prior distributions of parameters, the way to assess the frequency distribution and severity distribution was developed from a Bayesian perspective. Poisson distribution was chosen as the loss frequency distribution, and Gamma Distribution was set as the prior distribution. Extreme Value Theory (EVT) was applied in loss severity modeling, and GPD was used to fit the loss distribution. Both the location parameter and scale were assumed to have the prior Gamma distribution separately. Methodology for estimating parameters of the loss distribution was extended from Maximum Likelihoods Estimation (MLE) to Markov Chain Monte Carlo (MCMC) method. Results reveal that compared with the MLE method, MCMC method, which is based on Bayesian Inference, provides more effective and stable estimation for parameters in circumstances of small sampling data.

**Key words:** Bayesian inference; parameter estimation; operational risk measurement; extreme value theory

[责任编辑: 宋宏]