

损失分布法下操作风险的尾部风险度量

宋坤

(西南财经大学中国金融研究中心, 成都, 610074)

摘要: 操作风险广泛存在于金融机构的经营环节, 事关金融机构的内部控制结构, 其发生机制和控制方法等均有与其它风险不同的鲜明特点。面临着损失加剧、危害日趋严重的操作风险, 金融监管部门和金融机构均愈加重视对操作风险的防范。文章在以收集的国内外公开报道的操作损失数据的基础上, 运用变点理论来确定 POT 模型的阈值, 以实现在损失分布法的框架下运用极值理论对操作风险监管资本的度量。对于国内的金融机构, 本文所采用的算法在方法上可以对度量和规避为由操作风险所带来的损失提供支持。

关键词: 损失分布 操作风险 风险度量

引言

巴塞尔委员会提出了三大类操作风险度量方法, 其中高级计量法中的损失分布法是基于操作风险历史损失数据之上的, 其且所采用的技术是开放式的, 因此它是目前国内外学者研究最频繁的方法之一。操作风险服从尖峰厚尾的分布, 也就是说分布的尾部更能反映出极端操作风险事件所引发的巨额损失, 若能处理好分布的尾部就能更真实地反映出操作风险的特征。而 POT 模型能够针对性地拟合损失的尾部分布, 是公认的能较好地衡量分布尾部及损失极值的方法。但在实际应用中, 如何客观、精确地确定 POT 模型的阈值(临界点)仍是一道关键的障碍。

在样本数据数学特征相同的情况下, 适合于商业银行的操作风险度量方法在理论上也同样适合于投资银行和保险公司。巴塞尔委员会建议的高级计量法中的损失分布法使用较为广泛, 本文就采用损失分布法来度量操作风险。由于金融机构需要额外关注发生次数少但带来损失巨大的操作风险损失事件, 鉴于极值理论法在处理小概率事件以及外部冲击所引起的大损失方面所具有的良好特性, 同时 POT 模型能够用简单的方法合并损失强度分布与损失频率以得到总损失的尾部风险, 从而极大地减少了计算量, 所以就采用 POT 模型来拟合操作风险的损失强度分布, 以实现在损失分布法下对操作风险尾部风险的度量。

一、损失分布法概述

损失分布法在量化操作风险资本的高级计量法中, 是一种重要和应用广泛的方法, 它是一种数据驱动方法。根据巴塞尔委员会的定义, 是指在对操作损失事件的损失频率和损失强度进行假设基础上, 对业务线/损失事件类型矩阵中的每一类操作损失的复合分布分别进行估计, 进而计算得到某一时期一定置信度下该类型操作风险价值的方法¹。可表述如下:

总的损失是 n 个随机事件所对应的操作风险 $X_1, \dots, X_n$ (假设 X_i 是独立同分布的随机变量) 的加总 (L):

$$L = X_1 + X_2 + \dots + X_n \quad (1-1)$$

其中 n 为损失频率, X_i 为损失强度。

一般而言, 总的损失分布是强度分布与损失频率分布的卷积, 而损失强度和损失频率之间的关系是未知的, 所以较难得到总损失分布的解析形式。但极值理论中的 POT 模型却能提供出性质良好的处理途径², 因为它能够以简明的解析方法合并处理损失强度和损失频率,

¹ Basel Committee on Banking Supervision, 2004: Consultative document: the New Basel Capital Accord.

² Reiss R, Thomas M, 2001: Statistical Analysis of Extreme Values from Insurance, Finance, Hydrology and

从而获得总的损失的尾部估计。这样不仅能够降低直接模拟带来的计算成本，还能减少模拟误差。通常，模拟误差一般来源于对两类分布的假设与真实分布之间的差异，差异常常在模拟中被放大。由于金融机构需要额外关注发生次数少但带来损失巨大的操作风险损失事件，所以借助于极值理论是种较好的选择。鉴于此方法的结果是依据一定置信度得到的，且随着时间的变化而动态变化，因此能客观、有效地反映出金融机构特有的风险特征。

二、损失强度分布—POT 模型

（一）极值理论

1、极值理论概述

极值理论是次序统计理论的一个分支，1923 年和 1928 年 Dodd 和 Fisher 分别开始研究极值理论，1943 年 Gendendo 建立了极值定理，1955 年 Jenkinson 在极值风险研究中采用了极值理论，接着极值理论被越来越广泛地应用于自然科学领域，并逐渐延伸至金融领域。它是分析物理过程的一种技术，它描述了大量同分布样本极限值的统计特征，可以有效处理小概率事件和外部冲击所引起的大损失。它的研究只针对分布的尾部而非整个分布，所以是一种在极端条件下描述尾部特征的方法。它在自然学科及工程研究领域有较为广泛的应用，如水文勘察和气象预测等³。

操作风险损失事件的发生频率低而发生损失大，对于这种具有厚尾特征的分布，若采用传统的标准方差模型来计算 VaR 值必然会导致低估损失。而极值理论法则有着得天独厚的优势，因为分布的尾部体现出的是潜在的灾难性风险事件所引发的极端损失，极值理论法以样本中极端数据的极限定理为基础，允许仅对损失分布的高分位点进行参数估计，并不要求像其它统计方法一样假设出整个损失的分布形状，即极值理论法能够根据样本的极端值在总体分布未知的情况下，有针对性地拟合出损失的尾部分布，所以将极值理论用于度量金融机构的操作风险是一种合适的方法。

2、次序统计量

定义 x 为操作风险损失的金额， x 是独立同分布的，都来自同一个总体分布函数 $F(x)$ 。

现将 x 进行由大到小的排序，得到 $x_1 \geq x_2 \geq \dots \geq x_n$ ， $x_1, x_2, \dots, x_n$ 是 n 个随机变量，则称

$(x_1, x_2, \dots, x_n)$ 为次序统计量， x_i 为第 i 个次序统计量，令 $W(x) = \max\{x_1 \geq x_2, \dots, x_n\}$ ，

则 $W(x)$ 为极值分布函数。 $F(x)$ 与 $W(x)$ 之间的关系如下：

$$\begin{aligned} W(x) &= \Pr(x_1 \leq x, \dots, x_n \leq x) \\ &= \Pr(x_1 \leq x) \times \dots \times \Pr(x_n \leq x) \\ &= [F(x)]^n \end{aligned} \quad (2-1)$$

由公式 (2-1) 理论上可得到极值分布函数。但由于总体分布函数 $F(x)$ 是未知的，所以

可考虑通过样本中的极端数据得到当 $n \rightarrow \infty$ 时 $[F(x)]^n$ 的渐进分布。

极值理论法一般有两类，一类为 Fisher-Tippett 定理，指极大值序列的渐近分布收敛于广义极值分布 (GEV, Generalized Extreme-value Distribution)，包括 Gumbell、Frechet 和

Weibull 分布，这种方法考虑到损失事件的发生时间，被称为 BMM (Block Maxima Methods) 模型，针对块最大值建模。另一类指超过某一阈值后的样本服从广义帕累托分布，这种方法被称为 POT (Peaks over Threshold) 模型，专门把超过较大阈值的数据作为样本来建模，但忽略损失事件的发生时间。在实际应用中，由于 POT 模型能够较为有效地使用数量有限的极端样本值，而成为公认较好的度量操作风险损失极值的方法。下面就详细介绍本章要用到的 POT 模型。

(二) POT 模型

基于点过程法的 POT (Peaks Over Threshold) 模型选择位于某一较大阈值 (threshold) 之上的损失数据进行研究，是对数据进行广义帕累托分布 (GPD, Generalized Pareto Distribution) 拟和的技术，它忽略操作损失事件的发生时间，能够充分利用有限的极端数据建模。

x 为操作风险损失的金额， x 独立同分布，它们来自同一个总体分布函数 $F(x)$ 。对于置信度 p (即是大数，非显著性水平那个小数， $P(x \leq VaR) = p$)， p -th 分位数为：

$$x_p = F^{-1}(p) \quad (2-2)$$

即：

$$VaR_p = F^{-1}(p) \quad (2-3)$$

其中， F^{-1} 是分布函数 $F(x)$ 的反函数。

定义损失数据 x 超过阈值 x_M 的条件分布函数为 $F_{x_M}(y)$ (其中 $y = x - x_M$ ，表示超额损失； M 为超过阈值的损失数据的个数)，也将其称为阈值 x_M 的超额损失分布，表示为：

$$F_{x_M}(y) = P\{x - x_M \leq y | x > x_M\} = \frac{F(x_M + y) - F(x_M)}{1 - F(x_M)} \quad (2-4)$$

$$\Rightarrow F(x) = F_{x_M}(y)[1 - F(x_M)] + F(x_M) \quad x \geq x_M \quad (2-5)$$

根据 Pickands (1975)、Balkama 和 de Haan (1974) 定理，某一较高阈值超出值的极限分布可以用 GPD 来模拟，其分布具有厚尾特征。若存在大于零的常数 a_{x_M} 和 b_{x_M} ，使得当阈值 x_M 取较大值时， $F_{x_M}(a_{x_M} + b_{x_M})$ 具有连续的极限分布，那么：

$$\lim_{x_M \rightarrow x_+} \sup_{0 \leq y \leq x_+ - x_M} |F_{x_M}(y) - G_{\lambda, \eta(x_M)}(y)| = 0 \quad (2-6)$$

那么超额损失 y 的累积分布函数用 GPD 分布函数可表示为：

$$G_{\lambda, \eta}(y) = \begin{cases} 1 - \left(1 + \lambda \frac{y}{\eta}\right)^{-\frac{1}{\lambda}}, & \lambda \neq 0, \eta > 0, 1 + \lambda \frac{y}{\eta} > 0 \\ 1 - \exp\left(-\frac{y}{\eta}\right), & \lambda = 0, \eta > 0 \end{cases} \quad (2-7)$$

其中, λ 是 GPD 分布的尾部参数 (tail parameter), 决定尾部的消失速度。 λ 越大, 则尾部越厚, 反之越薄。当 $\lambda \geq 0$ 时, $y \in [0, \infty)$; 当 $\lambda < 0$ 时, $y \in (0, -\eta/\lambda]$ 。若一分布的尾部参数 $\lambda = 0$, 该分布为正态分布、对数正态分布或指数分布等; 若一分布的 $\lambda > 0$, 则表明该分布具有厚尾特征。

η 是尺度参数。当 $\eta = 1$ 时, 为广义帕累托分布的标准形式, 由三种互不相交的子类型分布构成:

指数分布:

$$G_0(y) = 1 - e^{-y} \quad x > 0 \quad (2-8)$$

Pareto 分布:

$$G_{1,\lambda}(y) = 1 - x^{-1/\lambda} \quad \lambda > 0, x \geq 1 \quad (2-9)$$

Weibull 分布:

$$G_{2,\lambda}(y) = 1 - (-x)^{-1/\lambda} \quad \lambda < 0, -1 \leq x \leq 0 \quad (2-10)$$

公式 (4-8) 中, $G_{\lambda,\eta}(y)$ 的密度函数为:

$$g_{\lambda,\eta}(y) = \frac{1}{\eta} \left(1 + \lambda \frac{y}{\eta} \right)^{-\frac{1}{\lambda}-1} \quad (2-11)$$

随着阈值 x_M 逐渐提高, $F_{x_M}(y)$ 逐渐收敛于 $G_{\lambda,\eta}(y)$, 即:

$$F_{x_M}(y) \approx G_{\lambda,\eta(x_M)}(y) \quad (2-13)$$

其中参数 λ 和 η 的取值依赖于 x_M 的大小。

可以得到超过阈值 x_M 的尾部分布:

$$F(x_M + y) = F(x_M) + (1 - F(x_M))G_{\lambda,\eta}(y) \quad (2-14)$$

三、损失频率分布的说明与操作风险监管资本的度量

操作风险的特征是“低频高损”, 对于发生概率较小的操作风险极端损失来说, 频率的估计往往较难处理。POT 模型的基本假设为, 大于阈值的超额部分是独立同分布的, 且此序列遵循泊松过程。有研究也表明当阈值足够大时, 超过阈值的超额损失会近似收敛于泊松分布⁴。这里设收集到的损失数据的样本总数为 n , 大于阈值 x_M 的操作风险损失次数为 M ,

损失频率分布 $F(x_M)$ 的估计能够通过 $(n - M)/n$ 得到。

因此, $\lambda \neq 0$ 时, 公式 (4-14) 可变为:

⁴ McNeil A., Frey R., Embrechts P.: 2005: Quantitative Risk Management: Concepts, Techniques and Tools, Princeton: Princeton University Press.

$$F(x_M + y) = \left(1 - \frac{M}{n}\right) + \frac{M}{n} \left[1 - \left(1 + \lambda \frac{y}{\eta}\right)^{-1/\lambda} \right] \quad (4-15)$$

尾部参数 λ 和尺度参数 η 的估计值可通过在给定阈值 x_M 下的极大似然函数 $\prod_{i=1}^n g(y_i)$ 来估计。

在 $\lambda \neq 0$ 时，由分布函数尾部数据的估计得到：

$$\hat{F}(x_M + y) = 1 - \frac{M}{n} \left(1 + \hat{\lambda} \frac{y}{\hat{\eta}} \right)^{-1/\hat{\lambda}} \quad (4-16)$$

则在 $\lambda \neq 0$ 时，，赋予指定的置信度 p ，则可得到 VaR 的估计值：

$$\hat{VaR}_p = x_M + \frac{\hat{\eta}}{\hat{\lambda}} \left\{ \left[\frac{n}{M} (1 - p) \right]^{-\hat{\lambda}} - 1 \right\} \quad (4-17)$$

巴塞尔委员会规定在计算银行操作风险监管资本时，要求涵盖操作风险的期望损失和非期望损失，除非银行有理由说明自己的期望损失已经得到很好的处理⁵。如果通过风险缓释手段已经规避了期望损失，那么操作风险的监管资本为非预期损失，即 VaR 和 EL 的差。

四、POT 模型阈值的确定

虽然 POT 模型能够直接处理操作风险损失数据的尾部，只依据真实的历史损失数据来选择分布函数。但金融机构运用 POT 模型来度量操作风险时，需要满足两个条件：一是需要有一定的历史损失数据，以便能精确估计出参数；二是需要设定合理的高阈值。也就是说，在 POT 模型下，度量结果很可能会出现良莠不齐的情况：因为如果阈值取得太高，那么能够被取入模型的损失数据样本点就会因为很少而无法建模，导致估计的偏差就会很大；而如果阈值取得太低，就会因为把不属于分布尾部的样本点当作是尾部的数据而进入模型，这就不能展现 POT 模型的优势，导致不相合的估计。因此，在实际的度量操作风险的应用中，阈值的确定是个非常关键问题，它决定了 POT 模型拟合操作风险损失分布的近似程度。国内外的很多学者（Hans, 2004; Stelios, 2005; Brooks, 2005; 田宏伟等, 2000）都在研究阈值的选取问题，但仍然没有就采用哪种方法选取的阈值结果最优而达成一致。目前使用较为频繁的阈值确定法有平均超额图法、Hill 图法、峰度法以及拟合优度法，下面逐一进行介绍。

（一）常见的阈值确定法

1、平均超额图法

由平均超额函数 $e(x_M) = E(x - x_M | x > x_M)$ ，可以得到：

$$e(x_M) = \frac{\eta + \lambda x_M}{1 - \lambda} \quad \eta + \lambda x_M > 0 \quad (4-1)$$

$e(x_M)$ 可通过样本的平均超额函数进行估计，得到：

⁵ Basel Committee on Banking Supervision, 2006: International Convergence of Capital Measurement and Capital Standards, A Revised Framework-Comprehensive Version, Switzerland: Bank for International Settlements.

$$e(x_M) = \sum_{i=1}^n (x_i - x_M)^+ / N_{x_M} \quad (4-2)$$

其中, N_{x_M} 为超过阈值 x_M 的样本个数。当 $x_i > x_M$ 时, $(x_i - x_M)^+ = x_i - x_M$; 当 $x_i \leq x_M$ 时, $(x_i - x_M)^+ = 0$ 。对于超过阈值 x_M 的样本, 函数曲线时会出现明显的线性变化: 当斜率为正时, 样本数据服从 GPD; 当成为水平线时, 样本数据服从指数分布; 当斜率为负时, 样本数据的尾部较薄。通过上述斜率的变化就能够确定样本的阈值。但对于函数曲线是否趋于线性, 则需要在观察图形的基础上借助于经验来判断。

2、Hill 图法

尾部指数的 Hill 统计量为:

$$\hat{\alpha} = \left(\frac{1}{M} \sum_{i=1}^M \ln \frac{x_i}{x_{M+1}} \right)^{-1}, \quad x_i > x_{M+1} \quad (4-3)$$

将 x_i 作为第 i 个降序样本统计量, 阈值为 x_M , 临界样本的序号为 M 。以 M 作为横轴、 $\hat{\alpha}$ 作为纵轴, 进行画图从而得到 Hill 图 (点 $\{(M, \alpha), 1 \leq M \leq n-1\}$ 的集合)。Hill 图中稳定区域的起点所对应的数值即可确定为阈值。当然, 何时进入稳定区域则需要借助于观察图形和经验判断, 具有一定的人为主观因素。

3、峰度法

峰度法是通过计算样本峰度来选取阈值的方法。样本峰度为:

$$K_n = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \mu_n)^4}{(S_n^2)^2} \quad (4-4)$$

其中, $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_n)^2$, $\mu_n = \frac{1}{n} \sum_{i=1}^n x_i$ 。那么当 $K_n \geq 3$ 时, 把令 $(x_i - \mu_n)^2$ 最大

的 x_i 删除, 重复直至 $K_n < 3$ 为止。然后从剩余的样本中选取最大的 x_i , 此值便可作为阈值。

峰度法虽然计算简便, 但并没有严格的理论支持。

4、拟合优度法

拟合优度法的原理为: 由于超过阈值的样本的超额值的条件分布服从广义帕累托分布, 所以最优的阈值便是使超额值的条件分布最接近 GPD 的值。所以, 先按一定标准确定多个阈值, 把超过阈值的超额值看作一个新的序列, 构造 Pearson- χ^2 统计量:

$$\chi^2 = \sum_{i=1}^N \frac{(N_i - np_i)^2}{np_i} \quad (4-5)$$

公式 (4-5) 表示了理论频数与实际频数间的差别。此统计量的值越小, 分布的置信度就会越高, 拟合效果就会越好, 对应的阈值也越佳。

(二) 基于变点理论的阈值确定法

鉴于阈值确定的方法仍处于探索研究过程, 下面利用变点理论, 把阈值所在的位置精确

地定位出来,进而把阈值定量地计算出来,从而在确定阈值这个非常重要的环节时避免了仅靠肉眼和经验来判断的弊端。具体思路为:先确定损失数据是否具有厚尾分布,然后给出的形状参数的 Hill 估计 $\hat{\alpha}$ 画出 Hill 图,接着采用变点理论,寻找 Hill 图中 $\hat{\alpha}$ 出现相对稳定区域的起始位置 d^* , 则 d^* 位置所对应的损失金额,就是阈值 x_M 。

1、变点理论的适用性

在自然界、社会、经济等问题的研究中,常常出现系统的输出序列在某未知时刻(或位置)发生突然变化,这点即称为变点(Chang Point)。变点统计分析的目的是判断和检验变点的存在、位置、个数,并估计出变点的跃度,可以将前、后数据的均值、概率分布或模型参数发生显著改变的时点称为均值、概率或模型参数的变点⁶。

Hill 图曲线的表现形态通常具有如下特征⁷:首先曲线剧烈变化,快速拉升或急骤下降,或反复性的剧烈变化,然后进入某区域或跨过某点后,突然变化减缓,并最终逐渐趋于稳定。曲线开始由非稳定变化转为稳定变化会出现在某一点(或区域)的前后,该突然转变的点就是变点,而变点的出现往往说明,该点(或区域)前后的数据具有不同的结构和不同的特性。Hill 图曲线中的非稳定区域与稳定区域的分界点就是所谓的阈值。因此,对确定阈值而言,至关重要的就是准确地定位出变点所属区域和位置。在变点位置附近,其斜率变化会表现得比较显著,然而斜率之差反映的只是该点两侧斜率的变化幅度,分析的是斜率的变化情况,所以必须结合斜率局部加(减)速变化最大来进行分析。因为形状参数 α 在其斜率加速度达到最大值之前,往往是沿着某类直线或曲线移动;而 α 在其斜率加速度越过最大值之后,会开始转向,变为沿着另一类直线或曲线运动。因此,通过确认曲线斜率变化的加(减)速度局部极大值,就能够对变点(即阈值)所在的位置进行确定。在等间距的情况下,斜率变化幅度的二阶差分反映了斜率的变化率。因此,如果结合着一阶差分最大值找到最接近曲线稳定区域的二阶差分最大值所处的区域,那么就找到了变点所在的区间,就可以确定 α 从何处开始进入稳定状态,再用类似于求分组数据众数的方法,便能够对变点的位置进行准确定位,将其换算为所对应的损失金额,这个金额对应的就是要确定的阈值。

2、基于变点理论确定阈值的算法

用变点对阈值进行确定的具体算法:

(1) 由于超过阈值的样本数的间距是相等的,将阈值探索点序列定义为 $d_j, (j = 1, 2, \dots, M - 1)$, 每个点所对应的 Hill 估计值为 α_j 。

(2) 计算探索点前、后周围曲线的斜率(即探索点前、后各若干数据点的线性回归系数)时,需考虑以各探索点为中心的滑动窗口。回归系数的取值会受到参加回归的数据点数 l 的多少的影响,所以在构造滑动窗口时,需要在探索点前、后各取 l 个数据点。同时,在较小的距离内,曲线会近似成直线,所以 l 的取值不能太大,因此这里分别设置 $l = 2, 3, 4$, 从而构成三套滑动窗口。在以 d_j 为滑动窗口中,线性回归探索点 d_j 前(或左)的 l 个数据点,得到回归系数 $\beta_{before}^{(l)}(d_j)$; 用同样的方法,线性回归探索点 d_j 后(或右)的 l 个数据点,

⁶ 叶五一、缪柏其、谭常春:《基于分位点回归模型变点检测的金融传染分析》,《数量经济技术经济研究》,2007年第10期。

⁷ S Resnick, C starica, 1997: Smoothing the Hill estimator, Advance in Applied Probability, Vol.29, pp.271-293. C Sousa, 2002: A contribution to the estimation of the tail index of heavy-tailed distributions, Ph.D. thesis (in The University of Michigan), www.utstat.toronto.edu/desousa.

得到回归系数 $\beta_{after}^{(l)}(d_j)$ 。可以发现， $l = 2$ 时，得到的回归系数能够较好地反映出局部的斜率性态，但却由于随机性较大而不能较完整地反映出整体的斜率性态； $l = 4$ 时，虽然得到的回归系数能够较完整地反映出整体的斜率性态，统计意义较强，但却不能较好地反映出局部的斜率性态。 $l = 3$ 时算出的回归系数性态则介乎二者之间，故应当更注重 $l = 4$ 时回归系数的结果。因此，将 $l = 2, 3, 4$ 时计算出的 $\beta_{before}^{(l)}(d_j)$ 和 $\beta_{after}^{(l)}(d_j)$ 进行加权平均时，权重取 l^2 ，即：

$$\bar{\beta}_{before}(d_j) = \frac{\sum_{l=2}^4 l^2 \cdot \beta_{before}^{(l)}(d_j)}{\sum_{l=2}^4 l^2} \quad (4-6)$$

$$\bar{\beta}_{after}(d_j) = \frac{\sum_{l=2}^4 l^2 \cdot \beta_{after}^{(l)}(d_j)}{\sum_{l=2}^4 l^2} \quad (4-7)$$

(3) 对每个探索点 d_j ，将 d_j 点前后曲线斜率的改变量记为 $\Delta S(d_j)$ ，即一阶差分：

$$\Delta S(d_j) = \bar{\beta}_{after}(d_j) - \bar{\beta}_{before}(d_j) \quad (4-8)$$

(4) 将得到 $\Delta S(d_j)$ 序列再进行计算，得到其二阶差分，即：

$$\Delta^2 S(d_j) = \Delta S(d_{j-1}) - \Delta S(d_j) \quad (4-9)$$

(5) 找到序列的 $\Delta^2 S(d_j)$ 的最大值所对应的区间 (d_{j-1}, d_j) ，这即为变点所在的区域。

根据其前后相邻区间 (d_{j-2}, d_{j-1}) 、 (d_j, d_{j+1}) 和对应的值 $\Delta^2 S(d_{j-1})$ 、 $\Delta^2 S(d_{j+1})$ ，用与分组数据众数类似的方法进行线性内插，则可精确定位到变点 d^* 的位置。

(6) 再由变点 d^* 的精确位置，以 d_{j-1} 、 d_j 与 d^* 的距离为权重，便得到阈值的精确值。

通过 Visual C++ 编程实现以上算法。

五、 POT 模型参数的估计

在用变点理论确定好 POT 模型的阈值后，下一步需要对参数 λ 和 η 进行估计。参数估

计的方法中，极大似然估计法 (MLE) 得到广泛的认可，特别是当 $\lambda > 0$ 时，估计效果更佳。但在数据出现异常值的情况下，MLE 法通常又是缺乏定量的稳健性的。由于在极值分布尾指数估计中出现的小误差就足以导致结果发生大的偏差⁸，为得到更稳健的估计量，采用平方误差积分 (ISE, Integrated Squared Error) 法来对 POT 模型中的参数进行估计，它与极大似然估计相比具有更良好的稳健性。下面分别对 MLE 法和 ISE 法进行介绍。

(一) 用 MLE 法估计参数

⁸ Brazauskas V., Serfling R., 2000: Robust and efficient estimation of the tail index of a single-parameter Pareto distribution, NorthAmer Actuar, Vol.4, pp.12-27.

把 $G_{\lambda,\eta}(y)$ 的密度函数 $g_{\lambda,\eta}(y)$ 可扩展为:

$$g_{\lambda,\eta}(y) = \begin{cases} \frac{1}{\eta} \left(1 + \lambda \frac{y}{\eta}\right)^{-\frac{1}{\lambda}-1}, & \lambda \neq 0 \\ \frac{1}{\eta} e^{-\frac{y}{\eta}}, & \lambda = 0 \end{cases} \quad (5-1)$$

超额损失 y 的对数似然函数为:

$$L(\lambda, \eta | y) = \begin{cases} -n \ln \eta - (1 + \frac{1}{\lambda}) \sum_{i=1}^n \ln(1 + \frac{\lambda}{\eta} y_i), & \lambda \neq 0 \\ -n \ln \eta - \frac{1}{\eta} \sum_{i=1}^n y_i, & \lambda = 0 \end{cases} \quad (5-2)$$

下面公式 (5-2) 求极大值, 就可得到参数 λ 和 η 的估计值。

(二) 用 ISE 法估计参数

基于平方误差积分的最小距离估计准则最早由 Terrel (1990) 提出的⁹。经证明此准则具有一致性和渐近正态性¹⁰。此准则直观地来说是取尽量使最多的数据与假定的参数模型相匹配的情况下所得到参数估计值, 因此其准确性损失不大, 得到的参数误差更加小, 结果更为稳定。

使用以平方误差积分为基础的最小距离估计准则, 能够得到适合的参数 $\hat{\theta}$ 使得密度函数 $f(\cdot | \theta)$ 最接近于真实未知的密度函数 f , 公式表述如下:

$$\hat{\theta} = \arg \min_{\theta} \left[\int (f(y|\theta) - f(y))^2 dy \right] \quad (5-3)$$

对一个具有参数 θ 的分布函数 $F(\cdot | \theta)$, 也可以表述为:

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum [F(y|\theta) - F(y)]^2 \right\} \quad (5-4)$$

针对所采用的 POT 模型, 在阈值 x_M 已确定的情况下, 在得出实际超额损失数据分布的条件概率后, 就可运用上述准则来拟出参数估计值。 $F_{x_M}(y)$ 为超额损失 y 的分布、 $G_{\lambda,\eta}(y)$ 为近似累积分布函数, 公式 (5-4) 可变为:

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum [G_{\lambda,\eta}(y) - F_{x_M}(y)]^2 \right\}$$

⁹ Terrel G, 1994: Linear density estimates, Proceedings of the Statistical Computing Section. American Statistical Association, Vol.1990, pp.297-302.

¹⁰ Hjort N.L., 1994: Minimum L_2 and robust Kullback-Leibler estimation, Proceedings of the 12th Prague Conference on Information Theory, Vol.3, pp.102-105.

Scott D.W, 1998: Parametric modeling by minimum L_2 error, Rice University, Houston.

$$= \arg \min_{\lambda, \eta} \left\{ \sum \left[1 - \left(1 + \lambda \frac{y}{\eta} \right)^{-\frac{1}{\lambda}} - F_{x_M}(y) \right]^2 \right\} \quad (5-5)$$

涉及到权重，则为：

$$\hat{\theta} = \arg \min_{\lambda, \eta} \left\{ \sum \omega \cdot \left[1 - \left(1 + \lambda \frac{y}{\eta} \right)^{-\frac{1}{\lambda}} - F_{x_M}(y) \right]^2 \right\} \quad (5-6)$$

在其它信息匮乏的情况下，选取权函数时采用了经验估计的方法。损失数据的顺序统计量为 $x_1 \geq x_2 \geq \dots \geq x_n$ ，超额损失的顺序统计量为 $y_1 \geq y_2 \geq \dots \geq y_{M-1}$ ，则 y_i 的分布的经验估计为：

$$\hat{F}(y_i) = \frac{M-i}{M-1}, \quad y_{i-1} < y < y_i \quad (5-7)$$

公式（4-32）可化为：

$$\hat{\theta} = \arg \min_{\lambda, \eta} \left\{ \sum_{i=1}^{M-1} \frac{M-i}{M-1} \cdot \left[1 - \left(1 + \lambda \frac{y_i}{\eta} \right)^{-\frac{1}{\lambda}} - F_{x_M}(y) \right]^2 \right\} \quad (5-8)$$

对上式分别求 λ 和 η 的偏导，并使之等于零就可以得到上式的最小值。

以上方法的实现任通过 Visual C++编程来实现。

六、实证分析

下面用第三章中收集的 1987 年至 2011 年期间，监管部门公布以及国内外媒体公开报导的 279 件商业银行操作风险损失事件，来验证本章提出的基于变点理论的 POT 模型阈值确定方法度量操作风险的有效性。若将这 279 个损失数据看作是整体来对操作风险进行度量时，那么可将估计结果视为国内商业银行业应对操作风险的监管资本。原始损失数据的频数图如图 4-1 所示。

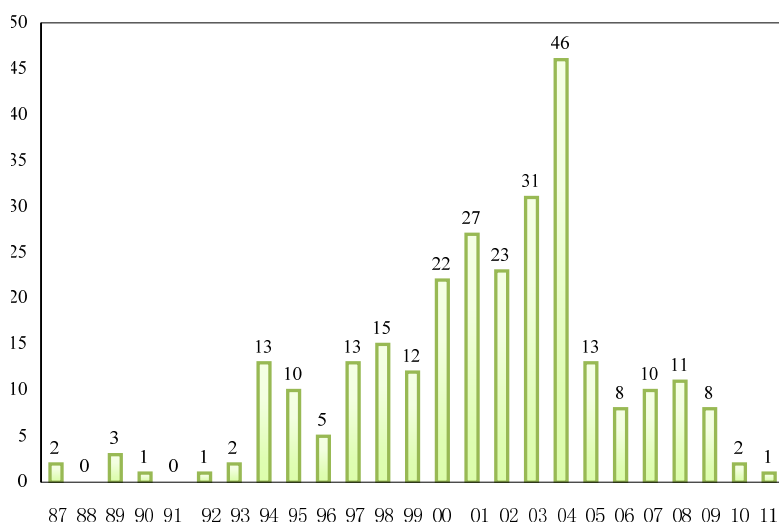


图6-1 1987年至2011年操作风险损失频数图

（一）阈值的确定

Hill图的稳定区域要具有优良的性质的前提是损失数据需要具有厚尾特征。因此必须首先检验损失数据是否服从厚尾分布。损失数据的描述性统计结果如表6-1所示。

表 6-1 操作风险损失数据的描述性统计

N	Valid	279
	Missing	0
Mean		16445.18
Std. Deviation		38926.35
Skewness		3.771
Std. Error of Skewness		.146
Kurtosis		16.250
Std. Error of Kurtosis		.291
Minimum		.578
Maximum		270000.0

由表6-1中损失数据可以看出，偏度为3.771，峰度为16.250，其特征有偏且尖峰。通过绘制QQ图来进一步检验操作风险的损失数据是否服从厚尾分布,其结果如下：

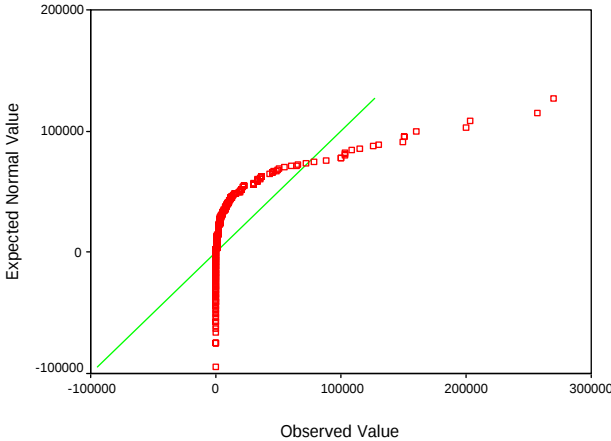


图6-2 操作风险损失数据的QQ图

在图6-2中选取正态分布为参考分布，其图形走势明显偏离直线并且向下弯曲，可以说明操作风险损失数据分布的尾部比正态分布的尾部厚，样本确实来源于厚尾性分布。通过 Visual C++编程处理得到的 $\hat{\alpha}$ 作为纵轴，以 M 作为横轴，绘制的Hill图如下：

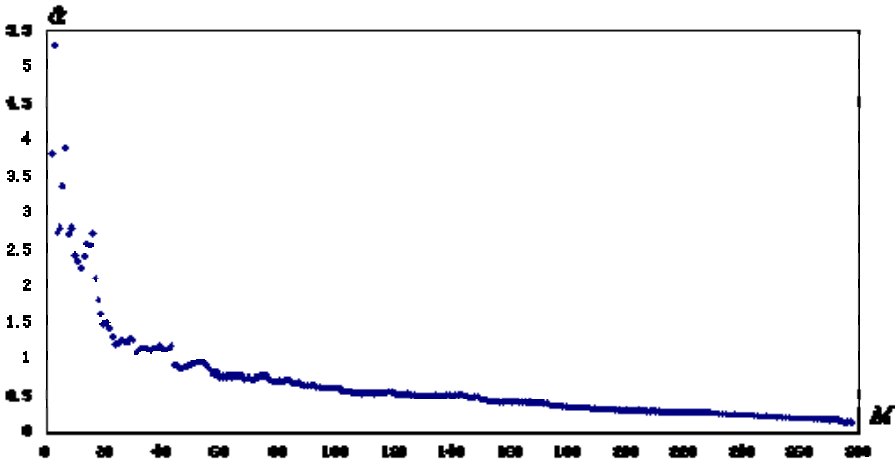


图6-3 Hill图

(由于计算出来第1个点的值为20.59，而其余点的值均小于5.5，为了图形显示效果，此图仅显示第2点至第278点)

由图 6-3 可以直观看出，从第 200 点后 $\hat{\alpha}$ 有进入稳定区域的趋势且出现了跳跃点，这可能就是变点所处的位置。下面对第 200 点之后的一、二阶差分值进行分析，如下图所示：

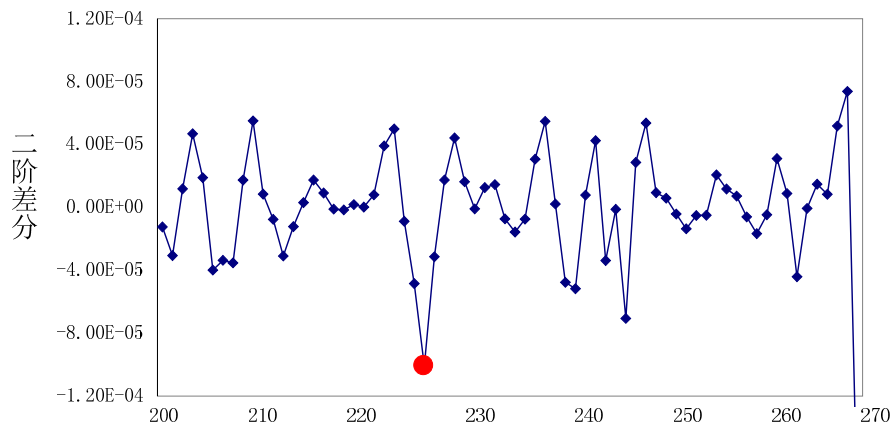


图 6-4 二阶差分值(第 200 点至第 269 点)

图 6-4 显示出距离稳定区域最近且二阶差分最大的点是第 226 个点。

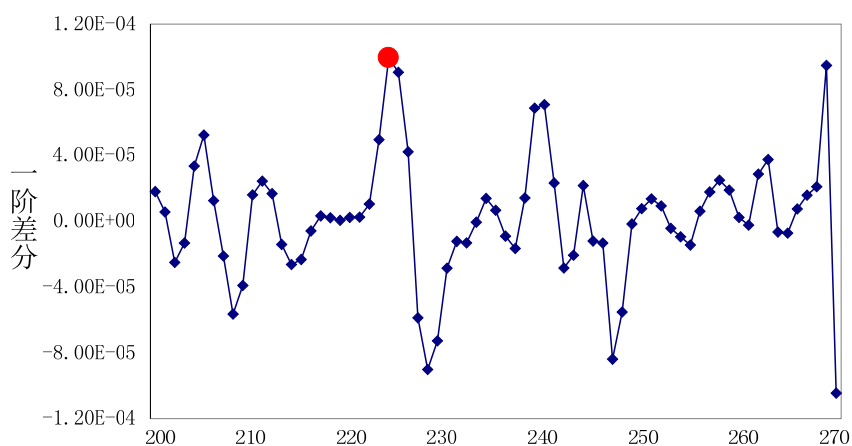


图 6-5 一阶差分值(第 200 点至第 270 点)

图 6-5 显示出距离稳定区域最近且一阶差分最大的点是第 224 个点。二阶差分与一阶差分最大值出现的位置意味着第 225 点前后数据的结构存在着显著差异，是区分 $\hat{\alpha}$ 值稳定与非稳定的变点，稳定区域的起始点就是阈值。

由于考虑到变点前后数据对该点的影响，若是只取第 225 点对应的损失金额做为阈值是不合适的。为保持数据的连续性，下面运用线性插值法精确定位变点 d^* 的位置，即：

$$d^* = d_{j-1} + \frac{\Delta^2 S(d_j) - \Delta^2 S(d_{j-1})}{[\Delta^2 S(d_j) - \Delta^2 S(d_{j-1})] + [\Delta^2 S(d_j) - \Delta^2 S(d_{j+1})]} \quad (6-1)$$

得 $d^* = 225.430$ 。以 d_{224} 、 d_{225} 与 d^* 的距离为权重，则阈值 x_M 为 19179.765 万元。

(二) 参数的估计

在确定阈值是 1.92 亿元的情况下，接下来对两个参数进行估计，超额损失 y 的分布见表 4-3。

表 6-2 超过阈值的损失数据的分布

Losses(万元)	Cumulative Percent(%)
20000	83.2
30000	85.3
37000	88.9
50000	91.4
79000	93.5
126000	96.8
160000	98.6
202900	99.3
270000	100

注：超额损失数据的总数为 54 个，由于篇幅限制只罗列部分结果。

对条件分布进行计算，得到的结果参见表 4-4。

$$F_{19179.765}(y) = \frac{F(19179.765 + y) - F(19179.765)}{1 - F(19179.765)} \quad (6-2)$$

表 6-3 超过阈值的损失数据的条件分布

Excess Losses(万元)	Conditional Probability
370.2	0.012
14320.2	0.329
29820.2	0.532
52820.2	0.646
89920.2	0.797
130520.2	0.870
183720.2	0.964
250820.2	1.000

注：类似于表 6-2，这里仅列出部分结果。

根据 $G_{\lambda,\eta}(y)$ 来进行参数估计。为了计算方便，现对 $G_{\lambda,\eta}(y)$ 分布进行相应的变化，如下：

设 $1 - G_{\lambda,\eta}(y) = Z$ ， $\lambda/\eta = \beta_1$ ， $1/\lambda = \beta_2$ ，则 $G_{\lambda,\eta}(y)$ 可化简为：

$$Z = (1 + \beta_1 y_i)^{-\beta_2} \quad (6-3)$$

通过 Matlab 软件拟合这条曲线来获得参数的估计值。为比较此方法与极大似然法的估计结果，采用计算机模拟产生 279 个随机数，进行随机排序，从而得到一组随机损失数据，利用极大似然法计算出参数估计值，重复 100 次上述过程就能够获得估计参数的均值。两种方法估计出的参数值的比较结果参见表 6-4。

表 6-4 两种方法得到的参数估计值

	$\hat{\lambda}$	$\hat{\eta}$
平方误差积分法	0.558(0.164)	186654 (4124.210)

极大似然估计法	0.582 (0.387)	1289956(9754.074)
---------	---------------	-------------------

括号内值为标准差。由表 6-4 可以看出，基于平方误差积分的最小距离估计的标准差，是小于 MLE 估计的标准差的。

设横坐标是有效超额损失数据（扣除阈值后的数据），纵坐标是大于阈值的条件累积概率，图 6-6 就是拟合效果。图 6-6 表明拟合效果良好，模型适宜。

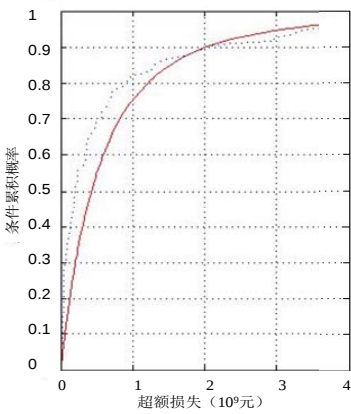


图 6-6 超额损失数据的拟合效果

注：连续曲线为帕累托的拟合损失分布，虚线为超额损失的损失分布

（三）操作风险监管资本的度量结果

表 6-5 VaR 值与监管资本 （单位:亿元）

	平方误差积分法		极大似然估计法	
置信度	95%	99.9%	95%	99.9%
VaR值	40.401	607.492	272.770	4579.835
监管资本	38.7565	605.8475	271.1255	4578.1905

巴塞尔委员会规定在计算操作风险监管资本时取 99.9%的置信度。在 99.9%的置信水平上，监管资本计算所得为 606 亿元。由于以所收集到的所有商业银行的损失事件为样本，因此这 606 亿元在 99.9%的置信水平下可近似看成对整个商业银行的操作风险监管资本。

参考文献

[1]Antoine Frachot, Thierry Roncalli, 2002: Mixing internal and external data for managing operational risk, Working Paper.

[2]Frachot A., Georges P., Roncalli T., 2001: Loss distribution approach for operational risk, Groupe de Recherche Operationnelle, Credit Lyonnais, France.

[3]Frachot A., Roncalli T., 2002: Mixing internal and external data for managing operational risk, Groupe de Recherche Operationnelle, Credit Lyonnais, France.

[4]A. M. Best, 2007: Draft risk management and the rating process for insurance companies, Vol.3, pp.9-12.

[5]Aue F., Kalkbrener M., 2007: LDA at work, www.gloriamundi.org/picsresources/famk_ida_v2.pdf.

[6]Basel Committee on Banking Supervision, 1999: The New Basel Capital Accord (CP1), June, 1-55.

[7]Jorion Philippe, 2001: The New Benchmark for Managing Financial Risk, 2nd edition, Mc - Graw - Hill.

[8]樊欣、杨晓光：《我国银行业操作风险的蒙特卡罗模拟估计》，《系统工程理论与实践》，2005 年第 5 期。

[9]高丽君、李建平、徐伟宣、陈建明：《基于 HKKP 估计的商业银行操作风险估计》，《系统工程》，2006

年第 6 期。

[10]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》，《运筹与管理》，2007 年第 1 期。

[11]高丽君：《商业银行操作风险外部数据的内生偏差研究》，《管理评论》，2011 年第 7 期。

[12]郭丹：《基于新巴塞尔资本协议看我国商业银行操作风险管理架构》，吉林大学硕士论文，2007。

[13]扈倩倩：《商业银行操作风险度量方法研究》，青岛大学硕士论文，2010。

[14]嵇尚洲、陈方正：《金融风险中的新领域—操作风险的度量与管理》，《上海金融》，2003 年第 1 期。

[15]刘张发：《商业银行操作风险度量方法比较研究》，云南大学硕士论文，2010。

[16]侣传振：《商业银行操作风险识别与管理》，山东大学硕士论文，2010。

Loss distribution method under the tail risk measure of operational risk

Song Kun

(Chinese Financial Research Centre of Southwestern University of Finance and
Economics, Chengdu, 610074)

Abstract: operational risk exists widely in the daily operation of financial institutions, and its characteristic and controlling method is significantly different from other risks. So facing over-growing operational risk, financial supervision department and financial institutions pay more and more attention on the measures against it. Based on the collection of data loss at home and abroad, this paper uses changing point theory to determine the threshold of POT model, in order to realize the loss distribution method by means of extremum theory under the framework of operational risk regulatory capital measurement. For domestic financial institutions, the algorithm adopted in this paper is on the way to measure and avoid operation risk brought about by the loss of support.

Key words: Operational Risk; Loss Distribution Approach; Risk measurement

收稿日期：2013-12-26

作者简介：宋坤，西南财经大学中国金融研究中心金融学博士生，研究方向：资本市场与证券投资