基于经验似然的 VaR 计算

郑浩，李金玉
（中国矿业大学理学院，江苏 徐州 221008）

摘要：本文介绍了非参数方法中的经验似然方法在估计样本分位点时的应用，最后将该方法应用到了金融资产 VaR 的计算上. 由于金融资产收益率的分布往往不能用参数方法准确的估计，经验似然方法作为非参数方法克服了这种局限，并改进了历史模拟方法. 文章对人民币／美元、人民币／英镑 两种汇率进行了 VaR 的计算的实证分析，计算了 VaR 的点估计和区间估计，并比较了几种计算方法，得到了一些有意义的结果。

关键词：VaR; 历史模拟法; 方差-协方差法; 经验似然; 分位数
中图分类号：O182

Empirical likelihood-based evaluations of Value at Risk

ZHENG Hao, LI Jinyu
(Depatment of Mathematics ,China University of Mining and Technology, JiangSu XuZhou 221008)

Abstract: In this paper the nonparametic empirical likelihood method is introduced and the VaR of the financial capital is evaluated. Because the financial asset yield is not easy to be estimated by parametic method，as a nonparametic method，Empirical likelihood method overcomes those defects，and also improves the history simulation method. In this paper VaR of American dollar and Pound sterling exchange rates are calculated. Both point estimator and interval estimator are given. By comparing those methods.some significative results are found.

Keywords: Value at risk; Historical Simulation approaches; analysis approaches; empirical likelihood; quantiles

0 引言

VaR(Value at Risk) 简称在险价值，是在一定置信水平下的头寸价值最大的变化范围，是近几年度量风险的重要的工具之一，可以被定义为“在一定的期间内，在一定的置信水平(例如 99%)下，一个金融头寸所面临的最大的潜在损失”。用数学公式来表示：

$$P\left(\Delta p \le -VaR\right) = \alpha ,$$

其中 Δp 为资产(或资产组合)在持有期内的收益，VaR 称为在置信水平 α 下的在险价值[1]。历史模拟方法、方差-协方差方法(分析方法)和蒙特卡罗模拟方法是计算 VaR 主要有 3 种方法。经验似然是 Owen(1988,1990)在完全样本下提出的一种统计推断方法，它是一种基于经验似然比函数的非参数的统计方法，具有无须假设样本的分布即可构造置信区间的统计特性和类似于 bootstrap 的抽样特性。本文应用经验似然方法，利用拉格朗日函数，对 VaR 分位点，置信区间进行了估计并且对人民币／美元、人民币／英镑两种汇率进行了 VaR 的计算，提出了基于经验似然的新的方法，它即克服了方差协-方差无法准确估计具有厚尾分布的金融数据的缺点，又不象历史模拟法需要过多的历史数据。本文将这种新方法得到的结果和其它方法进行了比较，得到一些有意义的结果。

作者简介：郑浩，（1983-），男，在读硕士，主要研究方向：数理统计
通信联系人：李金玉，（1961-），男，硕士生导师，主要研究方向：数理统计. E-mail: lijinyu@126.com

1 VaR

VaR 是在一定置信水平下的头寸价值最大的变化范围。头寸价值改变是指所持有的金融头寸的价值改变，也就是损益或回报。VaR 是面向金融工具的，因为金融头寸是指投资者在金融工具上所持有的头寸。头寸价值的改变石油时间性的，如果在无穷小的时间里，可以说任何头寸价值的改变都是零。所以，完整地讲，VaR 是在一段时间内，头寸价值改变在满足一定置信水平下的最大变化范围。计算 VaR 的三种方法：历史模拟法，方差-协方差法。

1.1 历史模拟法

历史模拟法假定投资组合的回报分布是独立同分布，市场因子的未来波动和历史波动完全一样，其核心是利用过去一段时间资产回报率数据，估算资产回报率的统计分布，再根据不同的分位数求得相应置信水平的 VaR。历史模拟法的步骤是:(l)将回报率按由小到大的顺序排列;(2)对于数据窗口宽度(样本区间长度)T，排序后的回报率分布的第 5 分位和第 1 分位数等对应为 95%VaR 和 99%VaR。[2]

1.2 方差-协方差方法

假定组合回报服从正态分布，于是利用正态分布的良好特性——置信度与分位数的对应性计算的组合的 VaR 等于组合收益率①的标准差与相应置信度下分位数的乘积[3]：

$$VaR = Z_\alpha \sigma \sqrt{\Delta t}$$

其中：Z_α –标准正态分布下置信度α对应的分位数

σ –组合收益率的标准差

Δt-持有期

很显然，正如以上所述，VaR 取决于两个重要的参数：持有期和置信度。针对不同的投资对象和风险管理者，这两个值的选择有所差异。具体而言，选择一个适当的持有期主要考虑以下因素：头寸的波动性、交易发生的频率、市场数据的可获性、监管者的要求等。通常情况下，银行等金融机构倾向于按日计算 VaR；但对于一般投资者而言，可按周或月计算 VaR。国际清算银行规定的作为计算银行监管资本 VaR 持有期为 10 天。置信度水平通常选择 95%～99%之间。95%的置信度意味着预期 100 天里只有 5 天所发生的损失会超过相应的 VaR 值；而 99%的置信度意味着预期 100 天里只有 1 天所发生的损失会超过相应的 VaR 值。

2 经验似然

设 $X_1,\ldots,X_n \in \mathbb{R}^d$ 独立有共同的累积分布 F，则 F 的非参数似然是：

$$L(F) = \prod_{i=1}^{n} F(\{X_i\})$$

这里 $F(\{X_i\})$ 是分布 F 在 X_i 处的概率质量，其中 $i = 1,2,\ldots,n$．大家知道 $X_1,\ldots,X_n$ 的经验累积分布函数 $F_n = n^{-1}\sum_{i=1}^{n}\delta_{X_i}$ 使上式达到极大，其中 $\delta_x(A) = I(x \in A)$ 也就是 是 F 的非参数极大似然估计．在参数推断中人们利用参数似然比进行假设检验与置信区间估计。

类似地，在分布完全未知的情况下非参数似然比 $R(F)=\dfrac{L(F)}{L(F_n)}$ 也可以用于统计推断。不象参数似然比，非参数似然比中不包含未知参数，一个自然的问题是如何使用它对参数作统计推断。注意到一些参数 是总体分布的泛函，即 $\theta=T(F)\in\mathbb{R}^p$，其中 $T(F)$ 是分布 F 的某泛函，F 属于某分布类 ，如总体均值及分位点等就是有上述形式泛函的例子。为了对 $T(F)=\theta$ 作检验， Owen(1988)[4][5]定义如下经验似然比检验统计量：

$$R(\theta)=\sup\left\{R(F)\,|\,T(F)=\theta,F\in F\right\}$$

很显然，经验似然比实际上是一种截面非参数似然比函数，它要求 F 在满足约束条件 $T(F)=\theta$ 下使非参数似然比达到极大(在无约束条件时，极大非参数似然比是 1)，而参数 由这一约束条件引入这一极大似然比中，从而得到关于参数 的极大截面非参数似然比函数，用这一非参数似然比作假设检验、区间估计或进行其它统计推断，这一方法就是所谓的经验似然方法。

2.1 分位数

设 $X_1,\ldots,X_n$ 定义一个具有分布函数 F 的随机变量序列，并假设 p 分位数 $q=F^{-1}(p)$ 唯一地被定义。由于 $E\left(1_{X\le q}-p\right)=0$，其中 $0<p<1,-\infty<q<+\infty$，我们考虑构造经验似然函数。令 $Z_i(p,q)=1_{X_i\le q}-p$，我们得到如下的经验似然函数：

$$R(p,q)=\max\left\{\prod_{i=1}^{n}np_i\,|\,\sum_{i=1}^{n}p_iZ_i(p,q)=0,p_i\ge 0,\sum_{i=1}^{n}p_i=1\right\}$$

由拉格朗日乘数方法可得

$$R(p,q)=\left(\frac{p}{\hat{p}}\right)^{n\hat{p}}\left(\frac{1-p}{1-\hat{p}}\right)^{n\left(1-\hat{p}\right)}$$

其中 $\hat{p}=\hat{p}(q)=\dfrac{\#\{X_i\le q\}}{n}$

我们定义 $0\log 0=0$，上试可改写为

$$R(p,q)=n\left[\hat{p}\log\left(\frac{\hat{p}}{p}\right)+\left(1-\hat{p}\right)\log\left(\frac{1-\hat{p}}{1-p}\right)\right]$$

$R(p,q)$ 分布是渐近自由度为 1 的标准卡方分布。-2 $R(p,q)$ 分布是渐近自由度为 1 的标准卡方分布[6],下面我们将利用上述理论对 VaR 的分位数和置信区间进行估计。

3 基于经验似然 VaR 的计算及实证比较

本文采用的两种汇率数据（人民币 / 美元、人民币 / 英镑）全部来源于中国银行官方网站的金融历史数据库，采用的时间为 2009 年 2 月到 2010 年 12 月，由于金融数据的研究一般采用对数回报，所以本文也采用对数回报。

假设 R_t 表示金融资产第 t 天的价格，用 r_t 表示其对数回报，则 r_t 可以表示为如下形式：

$$r_t = \log\left(\frac{R_t}{R_{t-1}}\right)$$

令 $x_t = -r_t$，如果的上分位点的点估计是 δ_P，那么金融资产的在险价值 VaR=$-\delta_P$.

下面利用 2009 年 2 月到 2010 年 8 月中的金融历史数据库，分别用历史模拟法（表 1），经验似然方法（表 2），方差-协方差方法（表 3）对的分位数及置信区间进行估计。

表 1 历史模拟法的 VaR 估计及置信区间
Table 1 Historical simulation-based evaluations and confidence intervals of Value at risk

	n	δ_p	VaR置信区间
人民币/美元	50	1.048%	(0.849%, 1.192%)
	500	1.012%	(0.864%, 1.173%)
人民币/英镑	50	1.021%	(0.823%, 1.163%)
	500	1.006%	(0.841%, 1.147%)

表 2 经验似然方法的 VaR 估计及置信区间
Table 2 Empirical likelihood-based evaluations and confidence intervals of Value at risk

	n	δ_p	VaR置信区间
人民币/美元	50	1.057%	(0.861%, 1.194%)
	500	1.023%	(0.871%, 1.169%)
人民币/英镑	50	1.044%	(0.832%, 1.161%)
	500	1.013%	(0, 853%, 1.141%)

表 3 方差-协方差方法的 VaR 估计及置信区间
Table 3 Analysis-based evaluations confidence and intervals of Value at risk

	均值μ	标准差σ	n	δ_p	VaR置信区间
人民币/美元	0.00074	0.00574	50	1.051%	(0.831%, 1.197%)
	0.00081	0.00531	500	1.029%	(0.857%, 1.184%)
人民币/英镑	0.00031	0.00421	50	1.017%	(0.819%, 1.154%)
	0.00039	0.00394	500	1.016%	(0.824%, 1.149%)

通过表 1 和表 2 的对比，我们发现经验似然方法和历史模拟法都和 n 有一定的关系，随着 n 的不同取值(n=50,n=500) δ_p 相差很大，这说明经验似然方法和历史模拟法都不是稳健的方法，它们估计值的精度依赖于样本 n 的个数。我们也发现经验似然方法和历史模拟法的置信区间也存在一定的差异，这主要是因为历史模拟法是利用渐近正态性来构造置信区间而经验似然方法是利用渐近卡方分布来构造置信区间，下面我们利用 2010 年 7 月到 2010 年 9 月的 94 组数据来检测经验似然方法和历史模拟法的置信区间覆盖率（表 4）。

表 4 经验似然方法和历史模拟法的置信区间覆盖率
Table 4 Coverage of empirical likelihood and Historical simulation

		T_1	T_2	α_1	α_2
人民币/美元	人名币/美元	79	85	84.04%	90.03%
	人名币/英磅	76	81	80.85%	86.17%
人民币/英镑	人名币/美元	86	89	91.46%	94.68%
	人名币/英磅	81	88	86.17%	93.62%

从表 4 可以看出经验似然方法具有更好的置信区间覆盖率，下面我们利用经验似然方法具有更好的置信区间覆盖率和具有与 bootstrap 类似的抽样特性，设技了如下的方法，具体的步骤为：

1) 设随机样本 $X^{'}=(X_1^{'},\ldots,X_n^{'})$ 来自总体分布 $F^{'}$，由观测值 $x^{'}=\left(x_1^{'},\ldots,x_n^{'}\right)$ 构造经验分布函数 $F_n^{'}$ 和置信区间（基于经验似然）；

2) 从 $F_n^{'}$ 中抽取简单样本 $x^{*}=\left(x_1^{*},\ldots,x_m^{*}\right)$，计算分位数 $\delta_p^{'}$，如果分位数在置信区间内则留下分位数作为估计值否则重新抽取简单样本估计分位数。

3) 重复步骤（2）的抽样 n 次得到样本分位数 $\delta_{p,i}^{'},i=1,\ldots,n$

4) 计算统计量 $\delta_p^{'}=\dfrac{1}{n}\sum_{i=1}^{n}\delta_{p,i}^{'}$ 作为分位数的估计值。

表 5 给出了基于经验似然方法的点估计和区间估计。

表 5 基于经验似然方法的 VaR 估计及置信区间

Table 5 Improved empirical likelihood-based evaluations and confidence intervals of Value at risk

	n	δ_p	VaR置信区间
人民币/美元	50	1.031%	(0.869%, 1.191%)
	500	1.024%	(0.873%, 1.172%)
人民币/英镑	50	1.015%	(0.841%, 1.163%)
	500	1.011%	(0, 847%, 1.153%)

几种方法的比较

VaR 是基于历史数据得出的未来风险价值，这样就需要我们对其预测结果的有效性与准确性进行检验，Kupiec P 在 Techniques for verifying the accuracy of risk measurement models 中提出 Kupiec 检验法，也称为失败检验法。Kupiec 在原假设 $H_0: p=p^{'}$ 条件下对失败率运用似然比检验，得到了似然比统计量：

$$LR=\ln\frac{\left[\left(1-n/t\right)^{t-n}\left(n/t\right)n\right]^2}{\left[\left(1-p^{'}\right)^{t-n}\left(p^{'}\right)^n\right]^2}$$

其中t－实际考察天数

n−失败天数

p−失败的概率, 其频率估计值为n/t

$p^{'}$－给定的值

在原假设 $H_0: p=p^{'}$ 条件下，LR 是服从于自由度为 1 的卡方分布，利用 LR 检验法，用 2010 年 9 月到 12 月的 92 个数据来检验预测效果，结果见下表：

表 6 三种方法的效果检验表

Table 6 Comparison of empirical likelihood,historical simulation and analysis

		n	LR
历史模拟法	人民币/美元	9	2.1692
	人民币/英镑	4	0.0597
方差-协方差法	人民币/美元	9	2.3472
	人民币/英镑	4	0.4971
基于EL方法	人民币/美元	4	0.0537
	人民币/英镑	3	0.3869

统计量 LR 越小，表示预测效果越好，从表 6 可以看出，EL 方法要优于历史模拟法和方差-协方差法。

大量的金融研究表明，金融市场收益率是非正态的，是厚尾的，所谓厚尾分布的定义如下：

如果分布函数$G(x)$的尾函数$G(x)^{'}=1-G(x)$以幂函数形式衰减，既满足：

$$1-G(x)=x^{\frac{1}{\eta}}l_G(x),\eta>1$$

这里，l_G是无穷远处缓变函数，满足对所有的$\alpha>0,\dfrac{l_G(\alpha x)}{l_G(x)}\to 1,x\to\infty$。

Pareto 分布，Burr 分布，学生 t 分布，对数伽马分布以及 Frechet 分布都属于厚尾分布。

下面我们通过自由度为 5 的 t 分布来检验计算结果，首先通过随机抽样得到样本，然后分别利用历史模拟法，方差-协方差，EL 方法进行了计算，具体结果见表 7：

表 7 模拟结果比较表

Table 7 Simulation of empirical likelihood,historical simulation and analysis

	t分布真实值	历史模拟法	方差-协方差法	基于EL方法
95%VaR	2.105	2.247	1.812	2.196
99%VaR	3.365	3.648	2.769	3.347

从表 7 可以看出 EL 最接近真实的 VaR 值，由统计理论可知，在给定固定的样本后，由经验似然方法得到的分位点会比历史模拟方法更接近实际情况，尤其是由于本文中的数据量不是很大，历史模拟方法得到的 VaR 会有较大的偏差. 方差-协方差方法假设了收益率分布服从正态，当其与正态相差较大时，出现厚尾分布时，作为一种参数方法不如经验似然方法准确。

4 结论

本文提出的基于经验似然方法的 VaR 计算是一种计算方法的改进，将这种比较新的计算方法用到了金融领域中，该方法具有很好的稳定性，克服了历史模拟法过度依赖于样本，样本的大小会对 VaR 值造成较大的影响，产生一个较大的方差的缺点，也克服了方差-协方差法由于正态假设条件，对于“厚尾”现象广泛存在，许多金融资产的收益率分布并不符合正态分布，这样，基于正态近似的模型往往会低估实际的风险值的缺点，可以给出一定置信水平下的置信区间，经验似然在其它方面，如估计波动率，收益率均值等也可以得到应用。

5 致谢

感谢导师给予的指导和审稿老师的仔细审稿。

[参考文献] (References)

[1] 陈忠阳 模型与金融风险管理[J]. 金融论坛，2001，（5）
[2] 刘玲，赵娇 风险测度和管理的 VaR 方法及其优缺点[J].北方金贸，2003
[3] Jorion P.Value at Risk.The New Benchmark for Controlling Market Risk[M].McGraw-Hill,New York,1997.
[4] Owen A .Empirical likelihood ratio confidence intervals[J].Biometrika,1988,75:237-149.
[5] Owen A.Empirical likelihood for linear models[J]. The Annals of Statistics,1990,18:90-120.
[6] Chen S X,Hall P.Smoothed empirical likelihood confidence intervals for quantiles.Ann Statist,1993,21:1166-1181.