

【编者按】量化市场风险对于期权定价、投资、金融监管和风险管理具有重要意义。而风险量化的关键在于选择科学的评估指标和工具。本期《财经研究》“海外归来”栏目翻译了西南财经大学经济与管理研究院杜在超教授和美国印第安纳大学 Juan Carlos Escanciano 教授在 *Management Science* 2017 年第 4 期合作发表的论文“*Backtesting Expected Shortfall: Accounting for tail risk*”。<sup>①</sup>该文首次提出了期望损失模型的后验分析方法,填补了金融风险领域的一个空白,也可为国内金融机构和学者提供借鉴和参考。

## 期望损失的后验分析

杜在超<sup>1</sup>, Juan Carlos Escanciano<sup>2</sup>

(1.西南财经大学 经济与管理研究院,四川 成都 610074;

2.美国印第安纳大学文理学院经济系,布鲁明顿 47405)

**摘要:**巴塞尔委员会已经批准用期望损失(*Expected Shortfall, ES*)作为市场风险指标对银行业进行监管,以替代现有的在险价值指标(*Value-at-Risk, VaR*)。这主要是因为期望损失满足风险度量应该满足的性质,而在险价值则不满足。在这个转变过程中,金融机构面临的主要困难是没有工具可以用来评估期望损失模型,即后验分析,而本文则提出了一套简单的 *ES* 后验分析工具。具体而言,我们基于累积碰撞序列(*Cumulative Violations*)构造了一套 *ES* 模型的检验方法,该方法是对 *VaR* 后验分析的自然推广。我们推导出了该检验的渐进分布性质,并且通过蒙特卡洛模拟研究了它们的有限样本表现。通过将该方法应用于全球三大股票指数,我们发现,在 2007 年金融危机期间,*VaR* 模型无法识别出金融危机时期与正常时期市场风险模型的差异,而 *ES* 模型则可以。

**关键词:**风险管理;期望损失;后验分析;尾部风险;在险价值

**中图分类号:**F830.91;F830.99 **文献标识码:**A **文章编号:**1001-9952(2017)12-0074-26

**DOI:**10.16538/j.cnki.jfe.2017.12.006

### 一、引言

市场风险的量化是学者和金融机构普遍感兴趣的问题,这对于期权定价、投资组合以及风险管理都非常重要。20 世纪 90 年代以来,度量市场风险的主要指标是在险价值(*VaR, Value-at-Risk*) (Jorion, 2006; Christoffersen, 2009)。*VaR* 描述了在给定置信水平的情况

收稿日期:2017-06-28

基金项目:国家自然科学基金(71401140);the Spanish Plan Nacional de I+D+I (ECO2014-55858-P)

作者简介:杜在超(1980—),男,山东烟台人,西南财经大学经济与管理研究院教授;

Juan Carlos Escanciano(1975—),男,美国印第安纳大学经济系教授。

<sup>①</sup>原文版权归属 INFORMS(<http://www.informs.org>),译文已获得原文的中文翻译版权。中文译文中出现的错误 INFORMS 概不负责,中文译文文责作者自负。我们感谢张栋浩在中文译稿中的大力帮助。*Management Science* 是管理类国际顶级期刊。杜在超教授于 2010 年获得印第安纳大学经济学博士学位,2010 年 8 月回国任教,2015 年被评为教育部首届青年长江学者,主要研究领域包括金融计量和应用计量。

下,一个投资组合可能遭受的最大损失。尽管  $VaR$  概念简单、计算方便,但是它却不是一个很好的风险度量指标。 $VaR$  没有考虑尾部风险,它仅仅告诉我们在没有发生尾部事件时的最大损失,但没有指出尾部事件发生时可能遭受的损失大小。 $VaR$  也存在其他的缺陷,包括不满足次可加性(Artzner 等,1997,1999; Acerbi 和 Tasche,2002)和凸性(Basak 和 Shapiro,2001)。期望损失( $ES$ , *Expected Shortfall*)克服了上述  $VaR$  的缺陷,是一种满足“一致性”的风险度量指标, $ES$  计算了超过给定置信水平的损失的平均值。<sup>①</sup> 在 2012 年 5 月 3 日的巴塞尔协议 III 中,巴塞尔委员会明确地提出要逐步淘汰  $VaR$ ,并采用  $ES$  作为替代指标。但是,使用  $ES$  作为度量市场风险的主要指标面临的一大挑战是,没有工具对  $ES$  进行评估(Yamai 和 Yoshida,2002,2005; Kerkhof 和 Melenberg,2004)。本文的主要目的就是提供这样的评估工具。

我们提出的方法主要是基于下面的观察。对每个置信水平而言,如果  $VaR$  模型是设定正确的,中心化的碰撞序列(投资组合损失超过  $VaR$  与否)应该是一个鞅差序列(Berkowitz, Christoffersen 和 Pelletier,2011)。如果在左侧尾部的置信区间内求取中心化碰撞序列的积分,也可以形成一个鞅差序列,我们把它叫做“累积碰撞序列”。就像  $ES$  是左侧尾部  $VaR$  的积分一样,累积碰撞序列是对左侧尾部所有碰撞序列的积分。因此,我们可以通过检验累积碰撞序列是否为鞅差序列来检验  $ES$  模型(Escanciano 和 Lobato,2009a)。具体而言,我们提出了一种 *Box-Pierce* 形式的检验(Box 和 Pierce,1970)。我们的 *Box-Pierce* 检验是对以往文献关于  $VaR$  条件后验分析的自然推广(Christoffersen,1998; Berkowitz, Christoffersen 和 Pelletier,2011)。同样地,我们基于累积碰撞序列的  $t$  检验是对  $VaR$  无条件后验分析的自然推广(Kupiec,1995)。McNeil 和 Frey(2000)、Berkowitz(2001)、Kerkhof 和 Melenberg(2004)、Wong(2008,2010)以及 Acerbi 和 Szekely(2014)曾提出了  $ES$  的无条件后验分析方法,而本文是第一个提出  $ES$  条件后验分析方法的文章。<sup>②</sup> 但是出于完整性考虑,我们也报告了无条件后验分析的结果。

我们对  $ES$  的后验分析考虑了有碍参数估计对检验的影响。已有文献对  $VaR$  或  $ES$  预测的估计误差进行了广泛研究,包括 Christoffersen 和 Goncalves(2005)、Gourieroux 和 Zakoian(2013)以及 Francq 和 Zakoian(2015)。Escanciano 和 Olmo(2010)研究了  $VaR$  后验分析存在的估计风险。针对  $ES$  后验分析的文献,只有 Kerkhof 和 Melenberg(2004)在对无条件后验分析进行历史模拟时考虑了估计风险。与 Escanciano 和 Olmo(2010)类似,我们从理论和模拟上证明了,当样本内观测没有比样本外观测大很多的时候,忽略估计风险将会给  $ES$  后验分析带来极大的 *Size* 扭曲(*size distortions*)。为了解决该局限性,我们提出了修正的后验分析,该方法充分考虑了估计风险,并通过蒙特卡洛模拟证明了稳健性。

本文想传达的主要信息是,与大多数学术和非学术文献相反, $ES$  后验分析并没有比  $VaR$  后验分析困难很多。<sup>③</sup> 我们提出的后验分析很容易执行,是对  $VaR$  后验分析方法的自然推广,也可以作为巴塞尔委员会建议的内部模型方法的工具箱,因此可谓是一个可以更好捕捉尾部事件的市场风险度量与评估。

<sup>①</sup>  $ES$  其他名称包括:条件  $VaR$ 、平均  $VaR$ 、尾部  $VaR$  或者平均尾部损失。

<sup>②</sup> 众所周知,对常用的模型(比如说,滤波历史模拟法)而言,条件后验分析比无条件后验分析更有检验力(Escanciano 和 Pei,2012)。

<sup>③</sup> 我们的观点与 Kerkhof and Melenberg(2005)、Acerbi 和 Szekely(2014)相同。并且,Acerbi 和 Szekely(2014)讨论了对  $ES$  做后验分析的可能性,以及阐述了“诱发性”(Elicitability)的概念。

本文余下内容安排如下。第二部分介绍本文常用的表达符号,以及后验分析所涉及的基本内容。第三部分我们提出新的无条件条件和条件后验分析,并求解它们的渐进性质。第四部分通过蒙特卡洛模拟研究后验分析的有限样本表现。第五部分采用金融危机后的日度数据,将本文提出的后验分析应用于三大主要股票指数,包括标普 500、德国 DAX 指数以及恒生指数。实证结果显示, VaR 对金融危机期间的极端事件的反应较为迟钝,而 ES 则提供了尾部风险的准确描述。第六部分总结全文。附录是本文的数学证明。

## 二、累积碰撞序列

$Y_t$  是银行在  $t$  时刻的收益,  $\Omega_{t-1}$  是风险管理者获得的  $t-1$  期的信息集, 包括  $Y_t$  的滞后值以及其他变量  $X_t$  的滞后值。也就是,  $\Omega_{t-1} = \{X_{t-1}, X_{t-2}, \dots; Y_{t-1}, Y_{t-2}, \dots\}$ 。我们假设,  $\{Y_t, X_t\}_{t=-\infty}^{\infty}$  是严格平稳、遍历性过程, 参考 Escanciano(2007) 的研究, 我们的结果也可以扩展到一些非严格平稳和非遍历性序列中。  $G(\cdot, \Omega_{t-1})$  是给定  $\Omega_{t-1}$  时  $Y_t$  的条件累积分布函数, 即  $G(\cdot, \Omega_{t-1}) = \Pr(Y_t \leq \cdot | \Omega_{t-1})$ 。为了表达的简便, 我们在所有包含随机变量的等式中删去了“almost surely”的说法。假设  $G(\cdot, \Omega_{t-1})$  是连续的。将  $\alpha \in [0, 1]$  记为置信水平。  $\alpha$  置信水平的 VaR 定义如下:

$$\Pr(Y_t \leq -VaR_t(\alpha) | \Omega_{t-1}) = \alpha \quad (1)$$

也就是说,  $-VaR_t(\alpha)$  是累积分布函数  $G$  在  $\alpha$  分位点的取值,  $VaR_t(\alpha) = -G^{-1}(\alpha, \Omega_{t-1}) = -\inf\{y: G(y, \Omega_{t-1}) \geq \alpha\}$ 。

定义时刻  $t$ , 置信水平  $\alpha$  时的碰撞序列是  $h_t(\alpha) = 1(Y_t \leq VaR_t(\alpha))$ , 其中,  $1(\cdot)$  是示性函数。当  $t$  时刻的损失大于等于  $VaR_t(\alpha)$  时, 示性函数取值 1, 否则取值 0。(1) 式的含义是碰撞序列是一个均值为  $\alpha$  的伯努利变量, 并且对于每一个  $\alpha \in [0, 1]$ , 中心化的碰撞序列 (*centered violations*) 都是鞅差序列, 即对每一个  $\alpha \in [0, 1]$ , 都有  $E[h_t(\alpha) - \alpha | \Omega_{t-1}] = 0$ 。

这个约束条件已经成为对 VaR 进行后验分析的基础。它的两个主要推论包括碰撞序列  $\{h_t(\alpha) - \alpha\}_{t=1}^{\infty}$  的零均值性质和不相关性, 分别是 Kupiec(1995) 的无条件后验分析以及 Christoffersen(1998) 的条件后验分析的基础。这也是最为广泛使用的后验分析方法。

VaR 由于没有捕捉尾部风险而被广泛批评。这也可以从碰撞序列  $\{h_t(\alpha) - \alpha\}_{t=1}^{\infty}$  中看出来, 它仅仅包含了损失是否超过 VaR 的信息, 但是并没有包含尾部事件所造成的损失额信息。VaR 没有捕捉尾部风险的缺陷加上其他方面的不足, 促使了 ES 概念的形成。ES 在度量投资组合风险时, 既考虑了损失超过一定置信水平的可能性, 也考虑了损失的规模大小。ES 的定义是超过  $VaR_t(\alpha)$  的条件期望损失, 也就是:

$$ES_t(\alpha) = E[-Y_t | \Omega_{t-1}, -Y_t > VaR_t(\alpha)] \quad (2)$$

通过条件概率以及变量转换, 可以将  $ES_t(\alpha)$  表达为  $VaR_t(\alpha)$  的一个函数:

$$ES_t(\alpha) = \frac{1}{\alpha} \int_0^\alpha VaR_t(u) du \quad (3)$$

不像  $VaR_t(\alpha)$  只包含了  $\alpha$  分位点的信息,  $ES_t(\alpha)$  通过对 VaRs 从 0 到  $\alpha$  积分, 包含了整个左侧尾部的信息。为了检验  $ES_t(\alpha)$  是否设定正确, 很自然想到要对  $h_t(\alpha)$  进行积分, 也就是累积碰撞序列  $H_t(\alpha) = \frac{1}{\alpha} \int_0^\alpha h_t(u) du$ 。

因为  $h_t(u)$  的均值是  $u$ , 通过 Fubini 定理可得,  $H_t(\alpha)$  的均值是  $1/\alpha \int_0^\alpha u = \alpha/2$ 。并且, 由

Fubini 定理可得,碰撞序列  $\{h_t(\alpha) - \alpha : \alpha \in [0, 1]\}_{t=1}^{\infty}$  的积分,即  $\{H_t(\alpha) - \alpha/2\}_{t=1}^{\infty}$  也是一个鞅差序列。

为了计算方便,定义  $u_t = G(Y_t, \Omega_{t-1})$ 。采用  $h_t(u) = 1(Y_t \leq -\text{VaR}_t(u)) = 1(u_t \leq u)$ , 我们可以得到:

$$H_t(\alpha) = \frac{1}{\alpha} \int_0^\alpha 1(u_t \leq u) du = \frac{1}{\alpha} (\alpha - u_t) 1(u_t \leq \alpha) \quad (4)$$

与碰撞序列相似,累积碰撞序列与  $Y_t$  的分布无关,因为  $\{u_t\}_{t=1}^{\infty}$  包括了服从独立同均匀分布的变量 (Rosenblatt, 1952; Berkowitz, 2001; Hong 和 Li, 2005)。Acerbi 和 Tasche (2002) 以及 Emmer 等 (2014) 采用 4 个式子的黎曼和 (Riemann sum) 对 (3) 式的积分进行了近似估计。<sup>①</sup> 累积碰撞序列可以不采用近似求解方法,因为 (4) 式的积分可以直接计算出来。但是与碰撞序列不同,累积碰撞序列还包括了尾部风险:当碰撞序列等于零时,累积碰撞序列也是零;当碰撞序列不等于零时,累积碰撞序列通过  $\alpha - u_t$  计算了  $Y_t$  的实际值与置信水平之间的差距。其中,  $\alpha - u_t = G(G^{-1}(\alpha, \Omega_{t-1}), \Omega_{t-1}) - G(Y_t, \Omega_{t-1})$ 。<sup>②</sup>

由于分布函数  $G$  是未知的,所以构造  $\{H_t(\alpha)\}_{t=1}^{\infty}$  的变量  $\{u_t\}_{t=1}^{\infty}$  也是未知的。在实践中,学者和风险管理人员可以通过设定参数条件分布函数  $G(\cdot, \Omega_{t-1}, \theta_0)$  以及估计未知参数  $\theta_0$  的办法,来计算 VaR 或 ES。一般来说,选择服从 Student-t 分布的位置尺度模型 (Location-Scale Models) 来刻画分布函数  $G(\cdot, \Omega_{t-1}, \theta_0)$ 。在我们的模型设定中,当然也可以考虑其他的模型选择。<sup>③</sup> 在选定了参数模型之后,可以定义“广义误差”,  $u_t(\theta_0) = G(Y_t, \Omega_{t-1}, \theta_0)$ , 以及与之相关的累积碰撞序列  $H_t(\alpha) = \frac{1}{\alpha} (\alpha - u_t(\theta_0)) 1(u_t(\theta_0) \leq \alpha)$ 。

上述论据为 ES 的后验分析提供了理论基础。我们可以通过检验  $\{H_t(\alpha, \theta_0) - \alpha/2\}_{t=1}^{\infty}$  是否具有零均值,对 ES 做无条件后验分析;通过检验  $\{H_t(\alpha, \theta_0) - \alpha/2\}_{t=1}^{\infty}$  是否序列相关,对 ES 做条件后验分析。在下一节,我们将提出这些假设的检验统计量。

### 三、ES 后验分析

在本节,我们提出 ES 的后验分析方法。ES 的无条件后验分析是一个简单的  $t$  检验,原假设是  $E[H_t(\alpha, \theta_0)] = \alpha/2$ ,该检验是 Kupiec (1995) 提出的 VaR 无条件后验分析的推广。ES 条件后验分析是一个 Box-Pierce 检验,这也是 Christoffersen (1998) 和 Berkowitz 等 (2011) 提出的 VaR 条件后验分析的自然推广。我们首先研究了一般后验分析方法 (考虑了估计风险) 的渐进分布,并且证明了不考虑估计风险可能会导致极大的 size 扭曲。继而,我们又提出了对估计风险稳健的修正后验分析方法。

#### (一) 一般无条件和条件后验分析

在实践中,模型的参数  $\theta_0$  是未知的,为了构造 ES 的预测而需要估计  $\theta_0$ 。为了行文的简单,我们采用固定预测机制,但是本文理论可以一般化地推广到其他预测机制上 (比如,移动窗口和递归机制); 对此问题,可以参考 Escanciano 和 Olmo (2010) 及其他相关文献。具体

<sup>①</sup>同样可以参考在一个风险水平的子集上对 VaR 做后验分析的文献,包括 Hurlin 和 Tokpavi (2006)、Perignon 和 Smith (2008) 以及 Colletaz 等 (2013)。

<sup>②</sup>实际上,  $d(y, x) = |G(y, \Omega_{t-1}) - G(x, \Omega_{t-1})|$  是一个距离函数。

<sup>③</sup>我们可以将该方法扩展到半参数设定中,其中  $\theta_0$  包括一个无穷维的元素。我们将在以后的研究中对此问题进行扩展。

而言,样本内共  $T$  期  $\{Y_{-T+1}, \Omega_{-T}, \dots, Y_0, \Omega_{-1}\}$  用来估计参数  $\theta_0$ , 即  $\hat{\theta}_T$ 。其中,  $\hat{\theta}_T$  是  $\theta_0$  的一致估计量, 比如是条件极大似然估计量 (CMLE);  $\hat{\Omega}_{t-1}$  是可观测信息集, 用来近似不可完全观测的信息集  $\Omega_{t-1}$  (例如, 通过设定初始值来近似很久之前的数据)。通过  $\hat{\theta}_T$  我们构造残差项  $\hat{u}_t = G(Y_t, \hat{\Omega}_{t-1}, \hat{\theta}_T)$ , 以及累积碰撞序列的估计式  $\hat{H}_t(\alpha) = \frac{1}{\alpha}(\alpha - \hat{u}_t)1(\hat{u}_t \leq \alpha)$ 。

然后, 采用时间长度为  $n$  的样本外观测  $\{Y_1, \Omega_0, \dots, Y_n, \Omega_{n-1}\}$  来评估 ES 模型 (后验分析)。我们的后验分析是基于累积碰撞序列的估计值。

ES 的无条件后验分析是一个标准的  $t$  检验, 原假设是  $H_{0\alpha}: E(H_t(\alpha, \theta_0)) = \alpha/2$ 。

经过简单计算可知,  $E(H_t^2(\alpha)) = \alpha/3$ , 以及  $Var(H_t(\alpha)) = \alpha(1/3 - \alpha/4)$ 。因此, 一个简单的  $t$  检验统计量是  $U_{ES} = \frac{\sqrt{n}(\bar{H}(\alpha) - \alpha/2)}{\sqrt{\alpha(1/3 - \alpha/4)}}$ , 其中,  $\bar{H}(\alpha)$  是  $\{\hat{H}_t(\alpha)\}_{t=1}^n$  的样本均值, 也

就是说,  $\bar{H}(\alpha) = \frac{1}{n} \sum_{t=1}^n H_t(\alpha)$ 。

由于参数估计效应 (estimation effect) 的存在,  $U_{ES}$  的渐进分布一般不服从标准正态, 而是依赖于样本内时间  $T$  和样本外时间  $n$  的渐进相对大小。假设,  $T$  和  $n$  都趋近于无穷大, 并且  $n/T \rightarrow \lambda < \infty$ 。定理 1 给出了  $U_{ES}$  的极限分布, 其中  $F_t(\theta, x)$  是参数误差  $u_t(\theta)$  的条件累积分布函数,  $F_t(\theta, x) \equiv \Pr[u_t(\theta) \leq x | \Omega_{t-1}]$ ,  $l_t$  是估计值  $\hat{\theta}_T$  的影响函数 (influence function), 附录 A 中的假设 A2 给出了该影响函数的定义。符号  $\rightarrow^d$  表示依分布收敛,  $B'$  表示矩阵  $B$  的转置。

**定理 1** 在附录 A 中的假设 A0-A4 下,  $U_{ES} \rightarrow^d N(0, \sigma_\lambda^2(\alpha))$ , 其中,

$$\sigma_\lambda^2(\alpha) = \frac{\alpha(1/3 - \alpha/4) + \lambda R'_{ES} E[l_t l_t'] R_{ES}}{\alpha(1/3 - \alpha/4)},$$

这里  $R_{ES} = 1/\alpha \cdot E[\int_0^\alpha (\partial F_t(\theta_0, x) / \partial \theta) dx]$ 。

**注释 1:** 类似于 Escanciano 和 Olmo(2010), 我们在定理 1 的证明中量化了参数估计效应, 具体为:

$$\sqrt{n}(\bar{H}(\alpha) - \frac{\alpha}{2}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n [H_t(\alpha, \theta_0) - \frac{\alpha}{2}] + \underbrace{\frac{\sqrt{\lambda}}{\sqrt{n}} R'_{ES} \sum_{t=-T+1}^0 l_t}_{\text{Estimation Effect}} + o_p(1),$$

附录 B 中给出了  $R_{ES}$  的详细表达式, 包括在一般的位置尺度模型中的具体表达式。 $\lambda=0$  是定理 1 的特殊情况, 即样本内观测显著大于样本外观测。当  $\lambda=0$  时,  $U_{ES}$  服从标准正态的极限分布, 无条件后验分析的推论将会极大简化。

**推论 1** 在定理 1 的假设下以及  $\lambda=0$ ,  $U_{ES} \rightarrow^d N(0, 1)$ 。

接下来, 我们讨论 ES 的条件后验分析。原假设是  $H_{0\alpha}: E[H_t(\alpha, \theta_0), \alpha/2 | \Omega_{t-1}] = 0$ , 该原假设是对 VaR 条件后验分析原假设的自然推广 (Christoffersen, 1998)。我们需要如下的符号。定义  $H_t(\alpha)$  滞后  $j$  阶 ( $j \geq 0$ ) 的自协方差和自相关系数是  $r_j = Cov(H_t(\alpha), H_{t-j}(\alpha))$  和  $\rho_j = \frac{r_j}{r_0}$ , 基于样本  $\{H_t(\alpha)\}_{t=1}^n$  的  $\gamma_j$  和  $\rho_j$  的表达式是,  $r_{nj} = \frac{1}{n-j} \sum_{t=1+j}^n (H_t(\alpha) -$

$\alpha/2)(H_{t-j}(\alpha) - \alpha/2)$  和  $\rho_{nj} = \frac{r_{nj}}{r_{n0}}$ 。

在自相关系数的定义中,我们代进了均值等于  $\alpha/2$  的约束,这使得基于  $\gamma_{nj}$  的检验对原假设  $H_{oc}$  具有更强的检验效能 (power)<sup>①</sup>。但是,在我们目前的设定中,  $\{H_t(\alpha)\}_{t=1}^{\infty}$  不可观测、 $\theta_0$  未知、 $\Omega_{t-1}$  也不可完全观测,所以我们需要将  $H_t(\alpha)$  替换为  $\hat{H}_t(\alpha)$ , 则有

$$\hat{r}_{nj} = \frac{1}{n-j} \sum_{t=1+j}^n (\hat{H}_t(\alpha) - \alpha/2)(\hat{H}_{t-j}(\alpha) - \alpha/2) \text{ 和 } \hat{\rho}_{nj} = \frac{\hat{r}_{nj}}{\hat{r}_{n0}}$$

注意在原假设  $H_{oc}$  下,对于  $j \geq 1$  都有  $\rho_j = 0$ 。所以,可以通过  $\hat{\rho}_{nj}$  构造简单的条件检验,比如说 *Box-Pierce* 检验统计量,

$$C_{ES}(m) = n \sum_{j=1}^m \hat{\rho}_{nj}^2 \quad (6)$$

至于无条件后验分析,  $C_{ES}(m)$  的渐进分布不仅依赖于  $\lambda$ , 还依赖于其他的未知量。

**定理 2** 在附录 A 的假设 A0-A4 下,  $C_{ES}(m) \rightarrow^d \sum_{j=1}^m \pi_j Z_j^2$ , 其中,  $\{\pi_j\}_{j=1}^m$  是矩阵  $\Sigma$  的  $m$  个特征值, 矩阵  $\Sigma$  的第  $ij$  个元素是:

$$\sum_{ij} = \delta_{ij} + \lambda R'_i E(l_i l'_i) R_j \quad (7)$$

$$R_j = \frac{1}{\alpha(1/3 - \alpha/4)} E[(H_{t-j}(\alpha) - \alpha/2) \int_0^\alpha \frac{\partial F_t(\theta_0, x)}{\partial \theta} dx] \quad (8)$$

$\delta_{ij}$  是克罗内克函数 (Kronecker delta function), 该函数在  $i=j$  时取值 1, 否则取值 0。  $\{Z_j\}$  是服从独立标准正态分布的变量。

从定理 2 可知,  $C_{ES}(m)$  服从一个加权的卡方极限分布。该权重是  $R_j$  的函数, 因为与  $Y_t$  的分布有关。但当  $\lambda=0$  时, 也就是  $\pi_j=1$  (对于任意的  $j=1, \dots, m$ ),  $C_{ES}(m)$  就服从标准极限分布。

**推论 2** 在定理 2 的假设下以及  $\lambda=0$ ,  $C_{ES}(m) \rightarrow^d \chi_m^2$ , 其中,  $\chi_m^2$  是自由度为  $m$  的卡方分布。

综上所述, 当  $\lambda=0$  时, 本文提出的无条件后验分析在零假设下具有标准极限分布; 但当  $\lambda>0$  时, 它们不再具有标准极限分布。这说明, 除非  $T$  显著大于  $n$  (比如  $T=2500$ ,  $n=250$ ), 否则基于标准分布临界值的基本检验不能控制第一类错误 (Type-I error)。为了避免发生该问题, 我们提出对估计风险稳健的修正后验分析。

### (二) 修正的无条件后验分析

在本节, 我们提出修正的后验分析方法, 该方法明确考虑了估计效应 (estimation effect), 并且对于任意的  $\lambda (0 \leq \lambda < \infty)$  都有标准极限分布。定理 1 给出了无条件后验分析 (式 5) 的估计效应表达式, 据此我们提出如下的修正检验统计量:

$$MU_{ES} = \frac{\sqrt{n}(\bar{H}(\alpha) - \alpha/2)}{\sqrt{\alpha(1/3 - \alpha/4) + \frac{n}{T} R'_{ES} W_T R_{ES}}}$$

其中,  $\hat{R}_{ES} = \frac{1}{\alpha n} \sum_{t=1}^n \int_0^\alpha \frac{\partial \hat{F}_t(\theta_T, x)}{\partial \theta} dx$ ,  $\hat{F}_t(\theta, x)$  是  $F_t(\theta, x)$  的一致估计量, 从而  $\hat{r}_{ES} = R_{ES} + o_p(1)$

<sup>①</sup> 经过一些代数计算可得,  $\gamma_{nj}$  近似等于  $1/(n-j) \sum_{t=1+j}^n [(H_t(\alpha) - \tilde{H}_t)(H_{t-j}(\alpha) - \tilde{H}_t)] + (\tilde{H}_t - \alpha/2)$ , 其中  $\tilde{H}_t$  是  $H_t(\alpha)$  的样本均值, 第一项可以检验  $H_t(\alpha)$  的自相关, 第二项检验无条件后验分析的原假设  $H_{0u}$ 。或者, 在定义自相关系数时不采用无条件均值约束 (比如说采用累积违背的样本均值)。但是, 蒙特卡洛模拟结果显示, 这些检验都没有我们采用的方法更具有检验效能 (Power)。

( $op(1)$ 表示依概率收敛于0)。并且,  $W_T = \frac{1}{T} \sum_{t=-T+1}^0 \hat{l}_t \hat{l}_t'$ ,  $\hat{l}_t$  是  $l_t$  的一致估计量,  $W_T = E[l_t l_t'] + op(1)$ 。

所以,  $W_T$  是估计量  $\hat{\theta}_T$  的渐进方差的一致估计, 这可以在许多统计软件中实现。在以 CMLE 方法估计的 GARCH 模型情形下, 本文第四部分提供了  $\hat{F}_t(\theta, x)$  和  $\hat{l}_t$  的表达式。

类似地, 通过定理 2 可以按照下式修正原有的条件后验分析(式子 6),  $MC_{ES}(m) = n \hat{\rho}_n^{(m)'} \hat{\Sigma}^{-1} \hat{\rho}_n^{(m)}$ , 其中,  $\hat{\rho}_n^{(m)} = (\hat{\rho}_{n1}, \hat{\rho}_{n2} \cdots \hat{\rho}_{nm})'$ ;  $\hat{\Sigma}$  是  $\Sigma$  的一致估计量, 即  $\hat{\Sigma} = \hat{\Sigma} + op(1)$ ,  $\hat{\Sigma}$  的第  $ij$  个元素是  $\hat{\Sigma}_{ij} = \delta_{ij} + \frac{n}{T} \hat{R}'_i W_T \hat{R}_j$ 。其中,  $\hat{R}_j = \frac{1}{\alpha(1/3 - \alpha/4)} \frac{1}{n-j} \sum_{t=j+1}^n \left\{ \left( \hat{H}_{t-j}(\alpha) - \frac{\alpha}{2} \right) \int_0^\alpha \frac{\partial \hat{F}_t(\hat{\theta}_T, x)}{\partial \theta} dx \right\}$ 。

在附录 B 中, 我们给出了位置尺度模型的  $\hat{r}_{ES}$  和  $\hat{r}_j$  的表达式。也可以从本文第四部分得知这些信息。

从定理 1 和定理 2 可知, 无论  $\lambda$  取何值, 修正的检验统计量  $MU_{ES}$  和  $MC_{ES}(m)$  都服从标准极限分布。

**推论 3** 在定理 1 的假设下以及  $\hat{r}_{ES}$  和  $W_T$  具有一致性,  $MU_{ES} \rightarrow^d N(0, 1)$ 。

**推论 4** 在定理 2 的假设下以及  $\hat{\Sigma}$  具有一致性,  $MC_{ES}(m) \rightarrow^d \chi_m^2$ 。

如推论 1 和推论 2 所示, 一般检验统计量只在  $\lambda=0$  时具有标准极限分布; 而推论 3 和推论 4 表明, 稳健的检验统计量对任意的  $0 \leq \lambda < \infty$  都具有标准极限分布。

#### 四、蒙特卡洛模拟

我们采用蒙特卡洛模拟评估本文后验分析方法的有限样本表现。为了便于比较, 我们同时汇报了 ES 和 VaR 的检验结果。参照 Kerkhof 和 Melenberg(2004) 及其他相关文献, 我们为 ES 选择了比 VaR 更大的置信水平  $\alpha$ 。具体而言, 根据经验准则, 我们为 ES 选择了比 VaR 大两倍(或接近于两倍)的置信水平, 从而使碰撞序列和累积碰撞序列的期望值保持一致(或接近一致)。在模拟中, 我们为 ES 设定的置信水平包括  $\alpha=0.1, 0.05$  以及  $0.025$ , 对应的 VaR 置信水平是  $\alpha=0.05, 0.025$  以及  $0.01$ 。我们比较 ES 和 VaR 的无条件后验分析和

条件后验分析。其中, VaR 的后验分析是  $U_{VaR} = \frac{\sqrt{n}(\bar{h}(\alpha) - \alpha)}{\sqrt{\alpha(1-\alpha)}}$ ,  $\bar{h}(\alpha)$  是

$\{\hat{h}_t(\alpha) = 1(\hat{u}_t \leq \alpha)\}_{t=1}^n$  的样本均值; VaR 的 Box-Pierce 检验是  $C_{VaR}(m) = n \sum_{j=1}^m \tilde{\rho}_{nj}^2$ , 其中,

$\tilde{\rho}_{nj} = \tilde{\gamma}_{nj} / \tilde{\gamma}_{no}$ ,  $\tilde{\gamma}_{nj} = 1/(n-j) \sum_{t=1+j}^n (\hat{h}_t(\alpha) - \alpha)(\hat{h}_{t-j}(\alpha) - \alpha)$ 。我们同样汇报了对估计效应保

持稳健性的修正的 VaR 后验分析检验结果。  $MU_{VaR} = \frac{\sqrt{n}(\bar{h}(\alpha) - \alpha)}{\sqrt{\alpha(1-\alpha) + \frac{n}{T} \tilde{R}'_{VaR} W_T \tilde{R}_{VaR}}}$ , 以及

$MC_{VaR}(m) = n \tilde{\rho}_n^{(m)'} \tilde{\Sigma}^{-1} \tilde{\rho}_n^{(m)}$ , 其中,  $\tilde{\rho}_n^{(m)} = (\tilde{\rho}_{n1}, \tilde{\rho}_{n2} \cdots \tilde{\rho}_{nm})'$ ;  $\tilde{\Sigma}_{ij} = \delta_{ij} + n/T \tilde{R}_i \tilde{W}_T \tilde{R}_j$ 。  $\tilde{R}_{VaR}$  和  $\tilde{R}_j$  的表达式可以从附录 B 中得知, 也可以参考 Escaniano 和 Olmo(2010)。

我们采用常用的 AR(1)-GARCH(1,1) 设定  $Y_t$  的真模型(原假设), 此时 VaR 和 ES

的表达式是:

$$\begin{aligned} VaR_t(\alpha) &= -a_0 Y_{t-1} - \sigma_t F_v^{-1}(\alpha), \sigma_t^2 = \omega_0 + \alpha_0 Y_{t-1} + \beta_0 \sigma_{t-1}^2 \\ ES_t(\alpha) &= -a_0 Y_{t-1} - \sigma_t m(\alpha), m(\alpha) = E[\epsilon_t | \epsilon_t \leq F_v^{-1}(\alpha)] \end{aligned} \quad (9)$$

其中,  $\epsilon_t$  服从自由度是  $v$  (未知) 的 *Student-t* 分布, 即  $\epsilon_t \sim t_v$ , 其中  $\alpha$  分位数记为  $F_v^{-1}(\alpha)$ 。我们将参数  $\theta_0$  的真实值设定为  $\theta_0 = (a_0, \omega_0, \alpha_0, \beta_0) = (0.05, 0.05, 0.1, 0.85)$  以及  $v = 5$ , 这也是实证文献中典型的参数值。

在每次模拟中, 我们都采用 *CMLE* 方法估计  $\theta_0$  和  $v$ , 从而得到  $\hat{u}_t = F_v(\epsilon_t(\hat{\theta}_T))$ ,  $F_v$  是自由度为  $v$  的 *Student-t* 分布的累积分布函数,  $f_v(\cdot)$  是其概率密度函数。  $\epsilon_t(\theta) = \epsilon_t(a, \omega, \alpha, \beta) = (Y_t - a Y_{t-1}) / \sigma_t(\theta)$ , 其中  $\sigma_t^2(\theta) = \omega + \alpha(Y_{t-1} - a Y_{t-2}) + \beta \sigma_{t-1}^2(\theta)$ 。这些表达式被用来计算如下修正的后验分析:  $\hat{F}_t(\theta, x) = F_v((a Y_{t-1} - \hat{a} Y_{t-1} + x \sigma_{t-1}(\theta)) / \sigma_t(\hat{\theta}_T))$ , 以及  $l_t = S^2 \partial(\ln f_v(\epsilon_t(\theta_T)) - \ln \sigma_t(\theta_T)) / \partial \theta$ , 其中  $S^2$  是  $\sqrt{T}(\hat{\theta}_T - \theta_0)$  渐进方差的一致估计量。我们通过设定  $m = 1, 3$  和  $5$  来计算未检验统计量  $U_{ES}, U_{VaR}, C_{ES}(m), C_{VaR}(m)$ , 以及修正后的检验统计量  $MU_{ES}, MU_{VaR}, MC_{ES}(m), MC_{VaR}(m)$ , 但在正文中只汇报了  $m = 5$  的结果。<sup>①</sup> 尽管只有当  $n/T$  很小的时候, 下述近似才具有理想的效果, 我们仍简单地采用标准正态分布近似  $U_{ES}$  和  $U_{VaR}$  的极限分布, 采用  $\chi_m^2$  分布计算  $C_{ES}(m)$  和  $C_{VaR}(m)$  的极限分布。通过设定样本内时间  $T = 250, 500, 2500$ , 以及样本外时间  $n = 250, 500$ , 我们进行了 1000 次蒙特卡洛模拟。

$Y_t$  数据生成过程的原假设和备择假设如下,

$H_0$ : *AR(1)-GARCH(1,1)* 模型,

$$\begin{aligned} Y_t &= 0.05 Y_{t-1} + v_t, v_t = \sigma_t \epsilon_t, \epsilon_t \sim t_5 \\ \sigma_t^2 &= 0.05 + 0.1 Y_{t-1}^2 + 0.85 \sigma_{t-1}^2 \end{aligned} \quad (10)$$

$A_1$ : *TAR* 模型,  $Y_t = a_t Y_{t-1} + v_t, v_t = \sigma_t \epsilon_t$

$$\begin{aligned} \sigma_t^2 &= 0.04 + 0.1 v_{t-1}^2 + 0.89 \sigma_{t-1}^2 \\ a_t &= 0.7 \times 1(v_{t-1} \leq -2). \end{aligned}$$

$A_2$ : 均值 *GARCH* 模型,  $Y_t = 2.5 \sigma_t^2 + v_t, v_t = \sigma_t \epsilon_t$

$$\sigma_t^2 = 0.01 + 0.29 v_{t-1}^2 + 0.7 \sigma_{t-1}^2.$$

$A_3$ : *AR(1)-ARCH(2)* 模型,  $Y_t = 0.05 Y_{t-1} + v_t, v_t = \sigma_t \epsilon_t$

$$\sigma_t^2 = 0.1 + 0.1 v_{t-1}^2 + 0.8 v_{t-2}^2.$$

$A_4$ : *AR(1)-EGARCH(1,1)* 模型,  $Y_t = 0.05 Y_{t-1} + v_t, v_t = h_t \epsilon_t$

$$\ln h_t^2 = 0.01 + 0.9 \ln h_{t-1}^2 + 0.3(|\epsilon_{t-1}| - \sqrt{2/\pi}) - 0.8 \epsilon_{t-1}.$$

$A_5$ : *AR(1)-Stochastic Volatility* 模型,  $Y_t = 0.05 Y_{t-1} + v_t, v_t = h_t \epsilon_t$

$$\begin{aligned} h_t^2 &= 0.1 Y_{t-1}^2 + \exp(0.98 \ln h_{t-1}^2 + e_t) \\ e_t &\sim iid N(0, 1). \end{aligned}$$

$A_6$ : *AR(1)-GARCH(1,1)* 并且具有混合正态残差项,  $Y_t$  与 (10) 式相同,

$$e_t \sim [0.6 \cdot N(1, \sqrt{2}) + 0.4 \cdot N(-1.5, \sqrt{0.75})] / \sqrt{3}.$$

除非被另外设定, 在这些模型中  $\{\epsilon_t\} \sim iid t_5$ 。Escanciano 和 Velasco (2010)、Escanciano 和 Olmo (2010) 采用了类似的模型。  $A_1$  和  $A_2$  中的条件均值没有被正确设定,  $A_3, A_4$  和

<sup>①</sup>关于  $m = 1$  和  $3$  的模拟结果, 读者如有兴趣, 可以向作者索取。



$A_5$  中的条件方差没有被正确设定,  $A_6$  中则是误差项  $\{\epsilon_t\}$  的分布没有被正确设定。

表 1-9 汇报了在 5% 显著性水平下各检验的实证 *size* 和 *size-corrected powers*。<sup>①</sup> 与前文理论一致, 当  $T$  较小时, 估计效应导致基本检验存在严重的 *size* 扭曲(犯第一类错误的可能性很高); 但当  $T$  很大时(比如,  $T=2\ 500$ ), *size* 扭曲的程度显著下降。除了  $VaR$  的条件后验分析  $MC_{VaR}(5)$  存在一些过度拒绝的情形外, 修正的后验分析在  $T$  很小时仍可以得到满意的 *size*。因为在尾部存在很少的观测值, 所以  $ES_t(0.025)$  和  $VaR_t(0.01)$  后验分析的 *size* 扭曲程度最大。对于给定的样本内观测  $T$ , 修正的  $ES$  后验分析  $MU_{ES}$  和  $MC_{ES}(5)$  的 *size* 表现, 随着样本外观测  $n$  从 250 增加到 500 而显著改善, 这与基本检验的结论相反。

由于小样本  $T$  存在 *size* 扭曲, 我们汇报了 *size-corrected powers*。其中,  $ES$  的无条件后验分析很好地识别了备择假设  $A_2$  和  $A_6$ , 条件后验分析对备择假设  $A_1$  和  $A_6$  具有最大的检验效能( $Power$  值最大), 对备择假设  $A_2-A_5$  也具有一定的检验效果。在绝大多数情况下,  $C_{ES}(5)$  和  $MC_{ES}(5)$  有着比  $C_{VaR}(5)$  和  $MC_{VaR}(5)$  更大的检验效能( $Power$  值更大)。最后,  $MU_{ES}$  和  $MC_{ES}(5)$  的检验效能随着样本外观测  $n$  的增加而变大, 说明了这些检验对备择假设而言是满足一致性的。当模型的残差项服从其他分布时, 比如说时变高阶距的 *Hansen-Skewed-t* 分布, 也可以得到一致的结论。<sup>②</sup>

蒙特卡洛模拟的主要结论如下。当  $T$  很小时( $T=250$ ), 基本后验分析存在显著的 *size* 扭曲, 但这些扭曲会随着  $T$  的增加而显著下降( $T=2\ 500$ )。与之相比, 修正的后验分析对  $T$  的取值具有稳健性。由于检验统计量需要根据模型和估计量去修正, *size* 的改善会导致计算成本的增加。所以, 我们建议大家样本内观测  $T$  很大时( $T=2\ 500, n=250$ ) 采用基本后验分析方法, 在样本内观测  $T$  较小时( $T=250, n=250$ ) 采用修正的后验分析方法。

模拟检验结果也说明了, 无条件后验分析和条件后验分析是互补的, 而不是替代的。因此, 我们建议在实践中同时使用两种后验分析方法。模拟结果也显示, 相比于  $VaR$  的检验, 基于  $ES$  的检验效能更好。在本文下一节, 我们将采用 2007 年金融危机期间三大股指的真实数据, 进一步验证我们的结论。

表 1 显著性水平 5%,  $T=250, ES_t(0.1)$  和  $VaR_t(0.05)$  的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.169	0.150	0.118	0.103	0.043	0.039	0.053	0.075
$A_1$	0.080	0.096	0.218	0.164	0.077	0.086	0.207	0.235
$A_2$	0.549	0.578	0.229	0.197	0.478	0.209	0.212	0.052
$A_3$	0.035	0.052	0.130	0.096	0.006	0.011	0.155	0.137
$A_4$	0.063	0.076	0.206	0.142	0.012	0.022	0.199	0.169
$A_5$	0.072	0.084	0.101	0.076	0.075	0.050	0.114	0.085
$A_6$	0.351	0.509	0.230	0.227	0.441	0.557	0.236	0.203
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.230	0.214	0.125	0.107	0.044	0.043	0.057	0.064
$A_1$	0.086	0.088	0.356	0.298	0.079	0.098	0.249	0.393
$A_2$	0.651	0.596	0.283	0.216	0.564	0.229	0.219	0.056
$A_3$	0.044	0.043	0.190	0.154	0.004	0.004	0.244	0.237

① *Size*: 犯第一类错误的概率, 即在原假设为真的情况下, 拒绝了原假设; *Power*: 不犯第二类错误的概率, 即在原假设为假的情况下, 拒绝了原假设。

② 限于篇幅, 没有汇报这些模拟结果, 读者如有兴趣, 可以向作者索取。

续表1 显著性水平5%,  $T=250$ ,  $ES_i(0.1)$ 和 $VaR_i(0.05)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
$A_4$	0.063	0.062	0.284	0.197	0.019	0.020	0.248	0.237
$A_5$	0.076	0.076	0.114	0.110	0.091	0.062	0.130	0.086
$A_6$	0.409	0.531	0.294	0.314	0.518	0.667	0.295	0.295

表2 显著性水平5%,  $T=250$ ,  $ES_i(0.05)$ 和 $VaR_i(0.025)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.176	0.114	0.140	0.155	0.051	0.042	0.073	0.120
$A_1$	0.052	0.043	0.131	0.117	0.071	0.082	0.089	0.102
$A_2$	0.358	0.193	0.180	0.048	0.418	0.138	0.151	0.022
$A_3$	0.041	0.029	0.094	0.081	0.014	0.008	0.081	0.080
$A_4$	0.048	0.037	0.123	0.115	0.013	0.012	0.104	0.101
$A_5$	0.038	0.024	0.055	0.055	0.069	0.064	0.059	0.035
$A_6$	0.673	0.314	0.492	0.066	0.776	0.786	0.483	0.045
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.234	0.222	0.129	0.123	0.048	0.044	0.060	0.103
$A_1$	0.065	0.081	0.225	0.177	0.074	0.097	0.145	0.205
$A_2$	0.479	0.468	0.222	0.186	0.519	0.136	0.148	0.029
$A_3$	0.038	0.050	0.147	0.123	0.010	0.018	0.165	0.148
$A_4$	0.048	0.053	0.174	0.141	0.006	0.023	0.144	0.123
$A_5$	0.053	0.057	0.072	0.063	0.084	0.085	0.069	0.041
$A_6$	0.795	0.655	0.614	0.381	0.824	0.837	0.630	0.275

表3 显著性水平5%,  $T=250$ ,  $ES_i(0.025)$ 和 $VaR_i(0.01)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.179	0.034	0.140	0.090	0.071	0.078	0.064	0.075
$A_1$	0.028	0.086	0.078	0.107	0.040	0.053	0.077	0.064
$A_2$	0.120	0.001	0.130	0.046	0.180	0.116	0.115	0.009
$A_3$	0.028	0.030	0.067	0.093	0.010	0.008	0.076	0.065
$A_4$	0.032	0.102	0.076	0.102	0.006	0.003	0.088	0.095
$A_5$	0.018	0.031	0.082	0.101	0.024	0.028	0.073	0.079
$A_6$	0.265	0.000	0.275	0.067	0.418	0.812	0.294	0.004
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.239	0.258	0.140	0.133	0.068	0.071	0.059	0.096
$A_1$	0.048	0.073	0.136	0.093	0.059	0.058	0.097	0.098
$A_2$	0.312	0.254	0.216	0.003	0.421	0.068	0.140	0.003
$A_3$	0.036	0.045	0.112	0.083	0.011	0.015	0.115	0.098
$A_4$	0.030	0.057	0.127	0.095	0.017	0.008	0.117	0.119
$A_5$	0.019	0.037	0.051	0.047	0.059	0.065	0.051	0.036
$A_6$	0.640	0.139	0.544	0.007	0.897	0.864	0.562	0.008

表4 显著性水平5%,  $T=500$ ,  $ES_i(0.1)$ 和 $VaR_i(0.05)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.101	0.098	0.081	0.093	0.041	0.032	0.057	0.085
$A_1$	0.072	0.090	0.230	0.173	0.075	0.083	0.203	0.210

续表4 显著性水平5%,  $T=500$ ,  $ES_i(0.1)$ 和 $VaR_i(0.05)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
$A_2$	0.686	0.698	0.237	0.178	0.641	0.345	0.203	0.063
$A_3$	0.028	0.043	0.157	0.124	0.011	0.008	0.152	0.140
$A_4$	0.059	0.082	0.204	0.172	0.022	0.020	0.184	0.162
$A_5$	0.078	0.098	0.113	0.108	0.059	0.050	0.082	0.058
$A_6$	0.488	0.657	0.235	0.265	0.588	0.650	0.249	0.238
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.132	0.129	0.085	0.093	0.038	0.032	0.052	0.083
$A_1$	0.087	0.073	0.401	0.300	0.106	0.119	0.351	0.382
$A_2$	0.906	0.803	0.275	0.213	0.775	0.481	0.230	0.058
$A_3$	0.027	0.024	0.264	0.180	0.007	0.007	0.280	0.225
$A_4$	0.072	0.053	0.328	0.235	0.025	0.031	0.268	0.235
$A_5$	0.115	0.105	0.143	0.123	0.062	0.058	0.104	0.074
$A_6$	0.733	0.758	0.287	0.311	0.793	0.865	0.325	0.315

表5 显著性水平5%,  $T=500$ ,  $ES_i(0.05)$ 和 $VaR_i(0.025)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.112	0.082	0.098	0.109	0.067	0.065	0.071	0.091
$A_1$	0.042	0.041	0.145	0.113	0.052	0.045	0.104	0.125
$A_2$	0.490	0.356	0.216	0.067	0.493	0.207	0.175	0.010
$A_3$	0.027	0.022	0.120	0.097	0.012	0.006	0.099	0.091
$A_4$	0.031	0.020	0.128	0.114	0.007	0.003	0.109	0.101
$A_5$	0.038	0.052	0.072	0.070	0.043	0.047	0.052	0.043
$A_6$	0.837	0.492	0.623	0.083	0.900	0.901	0.628	0.024
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.151	0.143	0.100	0.098	0.054	0.055	0.063	0.094
$A_1$	0.058	0.057	0.253	0.165	0.064	0.070	0.201	0.191
$A_2$	0.721	0.694	0.257	0.167	0.693	0.313	0.182	0.035
$A_3$	0.024	0.033	0.173	0.127	0.013	0.012	0.140	0.114
$A_4$	0.036	0.039	0.195	0.146	0.018	0.021	0.185	0.141
$A_5$	0.081	0.080	0.102	0.088	0.061	0.063	0.065	0.038
$A_6$	0.956	0.970	0.789	0.467	0.959	0.964	0.804	0.476

表6 显著性水平5%,  $T=500$ ,  $ES_i(0.025)$ 和 $VaR_i(0.01)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.146	0.022	0.130	0.079	0.085	0.089	0.087	0.078
$A_1$	0.036	0.072	0.085	0.097	0.032	0.033	0.070	0.080
$A_2$	0.275	0.104	0.243	0.026	0.267	0.079	0.175	0.010
$A_3$	0.027	0.032	0.082	0.090	0.018	0.004	0.078	0.088
$A_4$	0.022	0.042	0.081	0.089	0.006	0.004	0.073	0.080
$A_5$	0.009	0.017	0.032	0.027	0.011	0.010	0.036	0.035
$A_6$	0.355	0.097	0.321	0.063	0.428	0.365	0.335	0.004
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.151	0.183	0.119	0.127	0.081	0.080	0.077	0.109
$A_1$	0.036	0.063	0.101	0.084	0.038	0.032	0.073	0.83

续表6 显著性水平5%,  $T=500$ ,  $ES_i(0.025)$ 和 $VaR_i(0.01)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
$A_2$	0.399	0.322	0.186	0.002	0.482	0.145	0.183	0.003
$A_3$	0.027	0.066	0.120	0.082	0.018	0.005	0.107	0.100
$A_4$	0.021	0.055	0.116	0.097	0.010	0.005	0.089	0.092
$A_5$	0.022	0.023	0.050	0.029	0.032	0.041	0.034	0.030
$A_6$	0.673	0.097	0.610	0.004	0.977	0.960	0.617	0.004

表7 显著性水平5%,  $T=2500$ ,  $ES_i(0.1)$ 和 $VaR_i(0.05)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.069	0.072	0.051	0.062	0.059	0.067	0.044	0.062
$A_1$	0.061	0.046	0.309	0.211	0.058	0.048	0.328	0.240
$A_2$	0.813	0.744	0.297	0.242	0.769	0.549	0.279	0.166
$A_3$	0.031	0.029	0.218	0.150	0.016	0.009	0.221	0.136
$A_4$	0.071	0.054	0.281	0.221	0.041	0.032	0.286	0.215
$A_5$	0.148	0.133	0.248	0.214	0.093	0.071	0.168	0.124
$A_6$	0.650	0.735	0.255	0.305	0.656	0.735	0.295	0.388
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.061	0.083	0.053	0.054	0.038	0.046	0.048	0.054
$A_1$	0.101	0.047	0.494	0.367	0.102	0.065	0.572	0.442
$A_2$	0.979	0.933	0.345	0.265	0.949	0.837	0.336	0.188
$A_3$	0.034	0.023	0.331	0.234	0.018	0.011	0.364	0.259
$A_4$	0.106	0.068	0.408	0.297	0.045	0.044	0.439	0.336
$A_5$	0.207	0.165	0.251	0.237	0.083	0.064	0.196	0.180
$A_6$	0.929	0.933	0.320	0.379	0.936	0.933	0.359	0.407

表8 显著性水平5%,  $T=2500$ ,  $ES_i(0.05)$ 和 $VaR_i(0.025)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								
$H_0$	0.083	0.047	0.081	0.095	0.078	0.047	0.068	0.093
$A_1$	0.036	0.047	0.187	0.131	0.036	0.035	0.182	0.136
$A_2$	0.618	0.505	0.251	0.063	0.598	0.346	0.252	0.021
$A_3$	0.029	0.041	0.136	0.110	0.023	0.015	0.141	0.100
$A_4$	0.043	0.053	0.180	0.136	0.028	0.023	0.173	0.121
$A_5$	0.105	0.100	0.168	0.148	0.072	0.056	0.119	0.087
$A_6$	0.946	0.515	0.670	0.082	0.953	0.946	0.697	0.042
<i>n=500, Size and Power(size-corrected)</i>								
$H_0$	0.069	0.071	0.072	0.063	0.049	0.041	0.061	0.069
$A_1$	0.056	0.048	0.301	0.208	0.050	0.053	0.276	0.212
$A_2$	0.931	0.862	0.325	0.286	0.901	0.691	0.291	0.176
$A_3$	0.031	0.039	0.221	0.149	0.018	0.014	0.214	0.147
$A_4$	0.065	0.057	0.236	0.193	0.030	0.029	0.238	0.200
$A_5$	0.195	0.162	0.191	0.189	0.120	0.089	0.158	0.146
$A_6$	1.000	0.999	0.850	0.604	1.000	0.999	0.875	0.584

表9 显著性水平5%,  $T=2500$ ,  $ES_i(0.025)$ 和 $VaR_i(0.01)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
<i>n=250, Size and Power(size-corrected)</i>								

续表9 显著性水平5%,  $T=2\ 500$ ,  $ES_t(0.025)$ 和 $VaR_t(0.01)$ 的实证拒绝率

	$U_{ES}$	$U_{VaR}$	$C_{ES}(5)$	$C_{VaR}(5)$	$MU_{ES}$	$MU_{VaR}$	$MC_{ES}(5)$	$MC_{VaR}(5)$
$H_0$	0.103	0.093	0.101	0.082	0.098	0.089	0.095	0.071
$A_1$	0.029	0.077	0.078	0.097	0.027	0.027	0.067	0.057
$A_2$	0.400	0.527	0.344	0.050	0.290	0.027	0.202	0.000
$A_3$	0.023	0.088	0.088	0.088	0.018	0.000	0.079	0.084
$A_4$	0.034	0.082	0.100	0.082	0.017	0.005	0.102	0.100
$A_5$	0.055	0.055	0.100	0.055	0.013	0.003	0.052	0.024
$A_6$	0.401	0.956	0.349	0.062	0.420	0.955	0.353	0.001
$n=500$ , Size and Power(size-corrected)								
$H_0$	0.075	0.124	0.090	0.095	0.062	0.103	0.082	0.092
$A_1$	0.030	0.040	0.123	0.083	0.028	0.029	0.107	0.090
$A_2$	0.732	0.606	0.360	0.051	0.720	0.292	0.249	0.001
$A_3$	0.033	0.058	0.133	0.092	0.022	0.005	0.124	0.114
$A_4$	0.036	0.054	0.156	0.120	0.021	0.009	0.141	0.124
$A_5$	0.117	0.058	0.133	0.101	0.071	0.040	0.100	0.082
$A_6$	1.000	0.967	0.648	0.002	1.000	0.995	0.657	0.002

### 五、实证应用

我们通过采用2007年金融危机期间全球三大股票指数,来检验ES作为风险度量指标的优点。实证结果显示,VaR后验分析方法在金融危机时期不能拒绝常规风险模型(残差项服从Student-t分布的AR(1)-GARCH(1,1)模型);但本文提出的ES后验分析方法能够在金融危机时期显著拒绝常规风险模型。我们的实证结果补充了Kourouma等(2011)及O'Brien和Szerszen(2014)的研究,他们分别以金融危机期间的股票指数和美国五大银行为研究对象,评估了传统后验分析的表现。为与之对比,我们将比较传统VaR后验分析和本文ES后验分析的表现。<sup>①</sup>

具体而言,本文采用S&P 500指数、DAX指数以及恒生指数的日度数据展开分析。数据通过finance.yahoo.com获得,时间期限是1997.1.1-2009.6.30。表10从样本内和样本外两个时间段对三大股指进行了描述统计,样本内时间是1991.1.1-2007.6.30,样本外时间是金融危机期间的2007.7.1-2009.6.30。概括地讲,股票指数呈现了尖峰后尾和波动率聚集的特点,并且这些特点在金融危机期间尤为明显。表10显示,三大股指的超额峰度取值都很大,特别是1997年亚洲金融危机导致了恒生指数的超额峰度最大。图1显示,三大股指的对数收益率存在波动率聚集现象,金融危机期间股指收益的波动更为明显,尤其是标普500指数在2008.7-2009.3期间几乎下降了一半。在2001-2003年期间,由于911恐怖袭击和股市处于整体下行的阶段,S&P 500指数和DAX指数具有很大的波动。香港恒生指数则在1997年末经历了明显的波动。

表10 三大股指对数收益(%)的描述性统计

	样本内时间(1997-2007年)			样本外时间(2007-2009年)		
	S&P500	DAX	Hangseng	S&P500	DAX	Hangseng
观测值	2 639	2 658	2 596	504	509	503
均值	0.027	0.039	0.019	-0.098	-0.100	-0.034
中位数	0.059	0.106	0.042	0.036	0.016	0.052

<sup>①</sup>重复本文实证应用的R code,可以通过Run My Code网站获得(<http://www.runmycode.org/>)。

续表10 三大股指对数收益(%)的描述性统计

	样本内时间(1997—2007年)			样本外时间(2007—2009年)		
	S&P500	DAX	Hangseng	S&P500	DAX	Hangseng
标准差	1.131	1.577	1.677	2.218	2.045	2.762
偏度	-0.089	-0.152	0.146	-0.065	0.351	0.167
超额峰度	3.165	2.348	11.355	4.035	5.273	3.319
最大值	5.574	7.553	17.250	10.960	10.800	13.410
10分位点	-1.324	-2.132	-1.712	-2.478	-2.132	-3.238
5分位点	-1.815	-2.567	-2.552	-3.506	-3.371	-4.381
1分位点	-2.881	-4.547	-4.306	-6.310	-6.061	-7.305
最小值	-7.113	-6.652	-14.73	-9.470	-7.433	-13.580

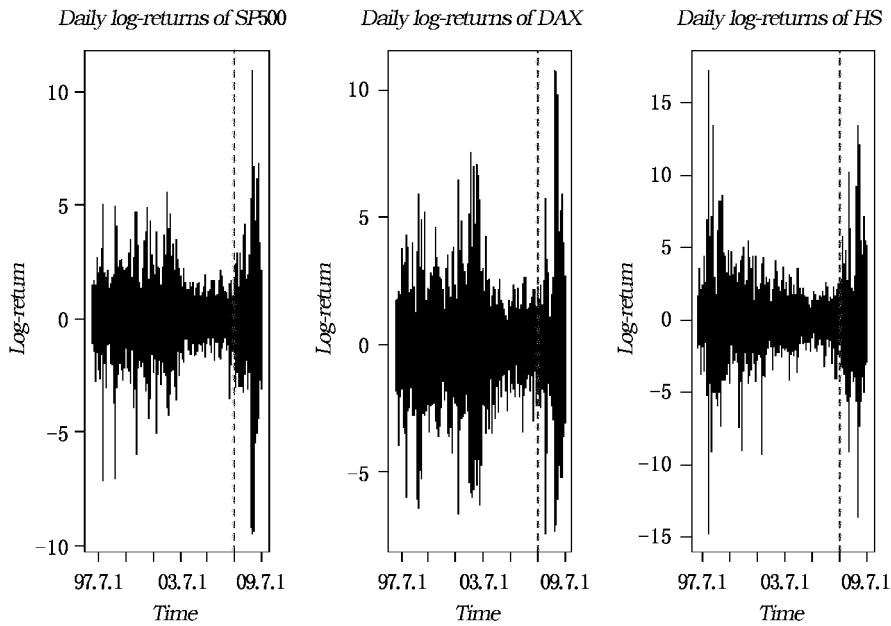


图1 三大股指的日均对数收益率

我们通过  $AR(1)-GARCH(1,1)-t$  模型拟合对数收益  $Y_t$ 。在给定置信水平  $\alpha$ ,  $VaR$  和  $ES$  等于:  $VaR_t(\alpha) = -a_0 Y_{t-1} - \sigma_t F_v^{-1}(\alpha)$ ,  $\sigma_t^2 = \omega_0 + \alpha_0 \sigma_{t-1}^2 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2$ ,  $ES_t(\alpha) = -a_0 Y_{t-1} - \sigma_t m(\alpha)$ ,  $m(\alpha) = E[\epsilon_t | \epsilon_t \leq F_v^{-1}(\alpha)]$ 。

我们采用样本内数据通过  $CMLE$  方法估计未知参数, 并采用样本外数据做后验分析。表 11 汇报了位置参数的估计结果, 包括  $F_v^{-1}(\alpha)$  和  $m(\alpha)$ 。<sup>①</sup> 表 11 显示, 三大股指存在明显的波动率惯性现象。恒生指数的自由度参数最小, 所以它的厚尾现象也最为明显, 这与表 10 中恒生指数有最大的峰度相一致。

表 11 条件极大似然估计(CMLE)

	S&P500	DAX	Hangseng
$a_0$	-0.027	0.004	0.034
$\omega_0$	0.007	0.016	0.010
$\alpha_0$	0.059	0.088	0.058
$\beta_0$	0.937	0.910	0.948

<sup>①</sup>本文中, 我们将自由度参数处理成离散、未知的。那么, 估计效应仅仅来源于位置参数和尺度参数的不确定性。Escanciano和 Olmo(2010)有相似地发现。

续表11 条件极大似然估计(CMLE)

	S&P 500	DAX	Hangseng
$v$	9	10	4
$F_v^{-1}(0.05)$	-1.617	-1.621	-1.507
$F_v^{-1}(0.01)$	-2.488	-2.472	-2.649
$m(0.1)$	-1.781	-1.779	-1.767
$m(0.25)$	-2.544	-2.521	-2.824

图2-4分别绘制了S&P 500指数、DAX指数以及恒生指数的VaR(0.05)和ES(0.1)估计结果。从中可知,  $-ES(0.1)$ 要小于 $-VaR(0.05)$ 。当 $Y_t$ 低于 $-VaR(0.05)$ 时,  $-ES(0.1)$ 更加接近真实的 $Y_t$ 。以2008年9月15日为例,当天雷曼兄弟破产,S&P 500指数下降4.83%,模型估计出的VaR(0.05)是1.82%,ES(0.1)则是2.65%,所以ES估计值更加接近真实损失。S&P 500指数收益在2007.7.1-2009.6.30期间共有504个观测值中,其中41个观测值小于 $-VaR(0.05)$ 。在这41个观测值中,S&P 500指数的平均损失是3.82%,VaR(0.05)的平均值是2.79%,而ES(0.1)的平均值则是3.07%。另外,还有11个观测值小于 $-ES(0.1)$ ,S&P 500指数的平均损失是3.76%,VaR(0.05)的平均值是3.13%,而ES(0.1)的平均值则是3.20%。因此,ES比VaR更好地描述了极端损失,通过DAX和恒生指数也可以得到类似结论。

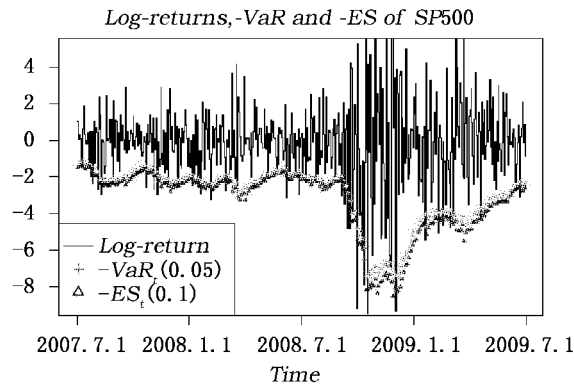


图2 S&P 500指数的对数收益以及VaR和ES估计值

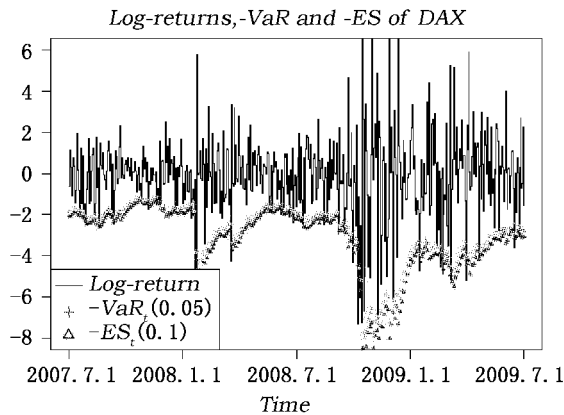


图3 DAX指数的对数收益以及VaR和ES估计值

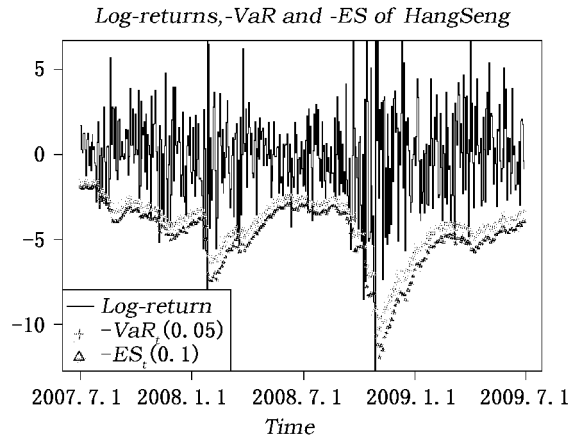


图 4 恒生指数的对数收益以及 VaR 和 ES 估计值

表 12 汇报了 2007 年金融危机前后三大股票指数的碰撞序列取值 ( $V(\alpha) = \sum_{i=1}^n \hat{h}_i(\alpha)$ )、累积碰撞序列取值 ( $CV(\alpha) = \sum_{i=1}^n \hat{H}_i(\alpha)$ ) 以及碰撞序列的预期值 ( $n\alpha$ )。危机前后时间段分别为 2005.07.01—2007.06.30 和 2007.07.01—2009.06.30。通过比较危机前后的  $V(\alpha)$  和  $CV(\alpha)$ ，可以看到危机发生后风险显著增加。一个例外是 DAX 指数的  $V(0.01)$  在危机发生后下降了，但是  $CV(0.025)$  在危机发生后上升了。进一步看，尽管 DAX 指数在危机发生后的碰撞序列取值  $V(0.01)$  下降了，但是危机发生后的损失要大于危机之前，这也就解释了危机发生后  $CV(0.025)$  的上升和  $V(0.01)$  的下降。表 12 的结果说明，在巴塞尔委员会建议的置信水平上(对 VaR 采取  $\alpha=0.01$ )，碰撞序列和累积碰撞序列存在显著的差异。

表 12 违背的描述统计:危机前与危机后

	金融危机前(2005—2007年)			金融危机后(2007—2009年)		
	S&P500	DAX	Hangseng	S&P500	DAX	Hangseng
$V(0.05)$	20	20	24	41	35	29
$CV(0.1)$	20.309	22.434	24.714	40.026	34.862	30.612
$n \times 0.05$	25.2	25.45	25.15	25.2	25.45	25.15
$V(0.01)$	5	8	2	11	5	5
$CV(0.025)$	6.110	6.360	4.063	13.702	9.101	6.145
$n \times 0.01$	5.04	5.09	5.03	5.04	5.09	5.03

图 5 描述了金融危机期间三大股票指数的累积碰撞序列  $\{\hat{H}_t(0.1)\}$ 。当  $\hat{H}_t(0.1)$  的取值较大，意味着当天股票指数产生了较大的损失。图 5 显示，S&P500 指数的累积碰撞序列取值大于 DAX 指数和恒生指数；三大股指的累积碰撞序列也都存在明显的聚集现象，反映了股指满足非鞅差序列的性质。我们将采用条件后验分析方法对此假设进行验证。

表 13 汇报了三大股票指数无条件后验分析 ( $U_{ES}, U_{VaR}, MU_{ES}, MU_{VaR}$ ) 以及条件后验分析 ( $C_{ES}(5), C_{VaR}(5), MC_{ES}(5), MC_{VaR}(5)$ ) 的  $p$  值。ES 的条件后验分析  $C_{ES}(5)$  和  $MC_{ES}(5)$  都拒绝了原假设中的风险模型，而 VaR 的条件后验分析  $C_{VaR}(5)$  和  $MC_{VaR}(5)$  则不能拒绝。图 6 和图 7 分别绘制了  $\hat{H}_t(0.025)$  和  $\hat{h}_t(0.01)$  的样本自相关图，从中也可以清楚地看到  $ES_t(0.025)$  在 5% 显著性水平上显著拒绝了原假设，而  $VaR_t(0.01)$  则没有拒绝。



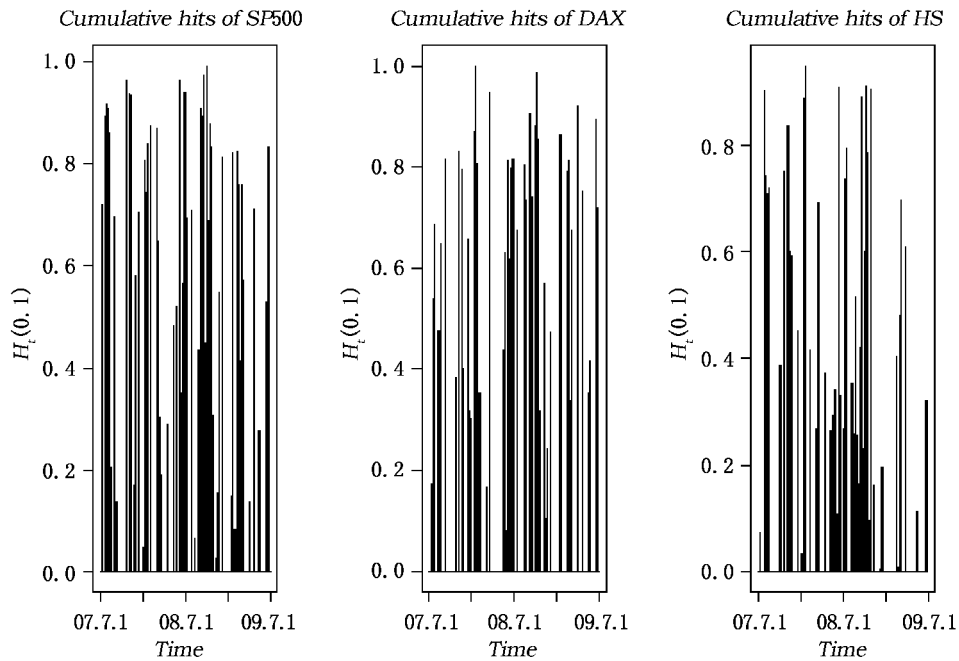


图5 金融危机期间三大股指的累积违背

图7显示,三个股指序列的碰撞序列 $\hat{h}_i(0.01)$ 都不存在显著的自相关性。*DAX*指数和恒生指数在滞后12阶时的 $\hat{h}_i(0.01)$ 自相关性都非常接近于0,因为这两个股指的 $Y_t$ 只有5次观测值小于 $-VaR_t(0.01)$ ,而*S&P500*指数有11次观测值要小于 $-VaR_t(0.01)$ 。2007年金融危机源于美国,并且对美国有更大的影响,所以导致了美国股市的极端损失要多于德国股市和香港股市。这也解释了为何表13中*S&P500*指数的无条件检验 $U_{VaR}$ 只有很小的 $p$ 值,而另外两个股票指数的 $p$ 值较大。

另一方面,三个股指序列的累积碰撞序列 $\hat{H}_i(0.025)$ 都具有显著的自相关性。发生极端损失的次数可能不是很多,但是平均损失额很大,且高度自相关。碰撞序列 $\hat{h}_i(0.01)$ 仅仅考虑了尾部事件的聚集特点,没有考虑损失规模。但是,累积碰撞序列 $\hat{H}_i(0.025)$ 则同时考虑了极端损失的发生次数和每次产生的损失额。所以,正如表13所示,基于累积碰撞序列 $\hat{H}_i(0.025)$ 的条件后验分析能更好地识别金融危机期间的风险模型问题。

总之,基于*VaR*的后验分析,我们会认为正常时期建立的风险模型在金融危机发生后也是适用的,并且据此计算的存款准备金也是足够的。但是,如果采用*ES*度量风险,本文提出的后验分析方法可以明确地拒绝正常时期的风险模型。我们的实证分析证实了*ES*相比于*VaR*更具备理论优势(Artzer等,1997,1999),并且该理论优势也可以在市场风险的后验分析中获得经验证据。本文基于累积碰撞序列提出的一套后验分析方法,不仅有助于评估金融损失发生的可能性,也有助于发现金融损失的规模大小。

表 13 ES 和 VaR 后验分析的 p 值

	ES(0.025)	VaR(0.01)	ES(0.1)	VaR(0.05)
<i>S&amp;P500</i>				
<i>U</i>	0.011	0.070	0.004	0.010
<i>C(5)</i>	0.007	0.270	0.009	0.052
<i>MU</i>	0.019	0.073	0.006	0.013
<i>MC(5)</i>	0.017	0.271	0.010	0.053
<i>DAX</i>				
<i>U</i>	0.224	0.968	0.045	0.095
<i>C(5)</i>	0.002	0.998	0.091	0.768
<i>MU</i>	0.253	0.968	0.052	0.102
<i>MC(5)</i>	0.015	0.998	0.095	0.769
<i>Hangseng</i>				
<i>U</i>	0.939	0.989	0.194	0.462
<i>C(5)</i>	0.002	0.998	0.002	0.002
<i>MU</i>	0.945	0.990	0.310	0.509
<i>MC(5)</i>	0.003	0.998	0.004	0.002

注: *U* 是基本的无条件后验分析, *MU* 是稳健的无条件后验分析, *C(5)* 是  $m=5$  的基本条件后验分析, *MC(5)* 是  $m=5$  的稳健条件后验分析。

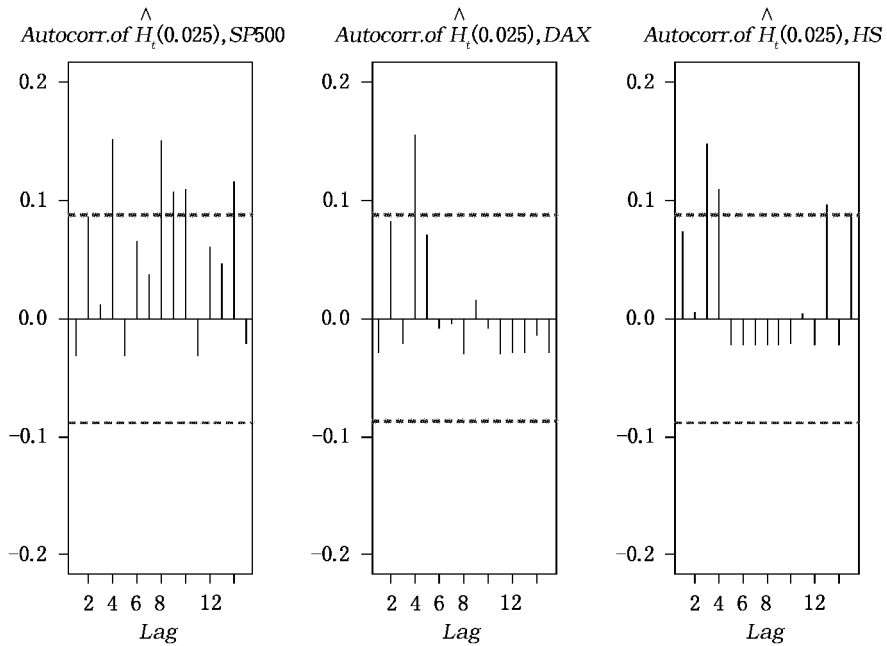


图 6 三大股指的  $\hat{H}_t(0.025)$  样本自相关

## 六、结 论

尽管大量的理论证据表明, *ES* 作为风险度量指标比 *VaR* 更具有优越性, 但是直到最近金融机构和监管者才开始采用 *ES* 替换 *VaR*。在这个转变过程中, 一大挑战是没有工具可以用来评估 *ES* 模型(即, 后验分析)。本文通过累积碰撞序列构建 *ES* 的无条件条件和后

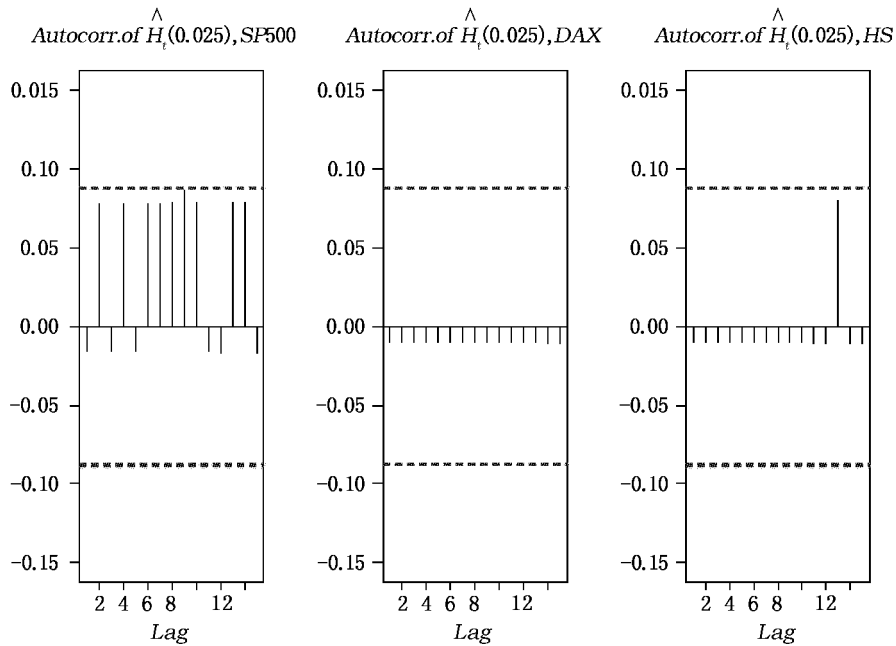


图 7 三大股指的  $\hat{h}_t(0.01)$  样本自相关

验分析,这也是基于碰撞序列对 VaR 进行后验分析的自然推广。但与碰撞序列不同,累积碰撞序列包括了尾部风险的信息,为风险提供了更加完整的描述。

本文提出了基本的无条件后验分析方法和对估计风险保持稳健性的经修正的后验分析方法。无条件后验分析采用  $t$  检验对累积碰撞序列的均值进行假设检验,条件后验分析采用基于累积碰撞序列的 *Box-Pierce* 检验。同时,我们也建议大家采用累积碰撞序列的图形和自相关系数作为补充工具。在具体应用中,实践者需要对条件后验分析做出两个选择,分别是置信水平  $\alpha$  和自相关阶数  $m$ 。对于 ES,我们建议选择  $\alpha$  等于 0.1、0.05 或 0.025,不建议取更小的值,因为此时需要很大的样本外观测才能对有限样本得到满意的渐进分布性质。关于自相关阶数的选择,我们选取了  $m=5$ 。当然,一个合理的做法是通过数据驱动方法选择  $m$  (Escanciano 和 Labato, 2009b)。该方法在其他环境下被证明可以获得可靠的推论,它也可以用于本文情形中,从而以较小的计算成本获得数据驱动的 ES 后验分析。

参考文献:

[1] Acerbi C, Tasche D. On the coherence of expected shortfall[J]. *Journal of Banking & Finance*, 2002, 26(7):1487-1503.  
 [2] Acerbi C, Szekely B. Backtesting expected shortfall [N]. *Risk Magazine*, 2014.  
 [3] Artzner P, Delbaen F, Eber M., et al. Thinking coherently [J]. *Risk*, 1997, 10:68-71.  
 [4] Artzner P, Delbaen F, Eber M., et al. Coherent measures of risk [J]. *Mathematical Finance*, 1999, 9(3): 203-228.  
 [5] Bai J. Testing parametric conditional distributions of dynamic models[J]. *Review of Economics and Statistics*, 2003, 85(3): 531-549.  
 [6] Basak S, Shapiro A. Value-at-risk based risk management: Optimal policies and asset prices [J]. *Review*

- of Financial Studies, 2001, 14(2):371–405.
- [7] Basel Committee on Banking Supervision. Consultative document: Fundamental review of the trading book [R]. Basel, Switzerland, 2012.
- [8] Berkowitz J. Testing density forecasts with applications to risk management [J]. Journal of Business and Economic Statistics, 2001, 19(4):465–474.
- [9] Berkowitz J, Christoffersen P, Pelletier D. Evaluating value-at-risk models with desk-level data [J]. Management Science, 2011, 57(12): 2213–2227.
- [10] Bose A. Bahadur representation of mestimates [J]. Annals of Statistics, 1998, 26(2):771–777.
- [11] Box G, Pierce D. Distribution of residual autocorrelations in autoregressive integrated moving average time series models [J]. Journal of American Statistical Association, 1970, 65(332):1509–1527.
- [12] Christoffersen P. Evaluating interval forecasts [J]. International Economic Review, 1998, 39(4):841–862.
- [13] Christoffersen P, Goncalves P. Estimation risk in financial risk management [J]. Journal of Risk, 2005, 7(3):1–28.
- [14] Christoffersen P. Value-at-risk models [M]. Handbook of Financial Time Series, 2009.
- [15] Colletaz G, Hurlin C, Perignon C. The risk map: A new tool for validating risk models [J]. Journal of Banking and Finance, 2013, 37(10):3843–3854.
- [16] Du Z. Nonparametric bootstrap tests for the independence of generalized errors [R]. CAEPR Working Paper No. 023–2009, 2015.
- [17] Emmer S, Kratz M, Tasche D. What is the best risk measure in practice? A comparison of standard measures [R]. ESSEC Business School: Paris, France, 2014.
- [18] Escanciano J. Weak convergence of non-stationary multivariate marked processes with applications to martingale testing [J]. Journal of Multivariate Analysis, 2007, 98(7):1321–1336.
- [19] Escanciano J, Lobato I. Testing the martingale hypothesis [M]. Palgrave Handbook of Econometrics, 2009a.
- [20] Escanciano J, Lobato I. An automatic portmanteau test for serial correlation [J]. Journal of Econometrics, 2009b, 151(2):140–149.
- [21] Escanciano J, Olmo J. Backtesting parametric value-at-risk with estimation risk [J]. Journal of Business and Economic Statistics, 2010, 28(1):36–51.
- [22] Escanciano J, Velasco C. Specification tests of parametric dynamic conditional quantiles [J]. Journal of Econometrics, 2010, 159(1):209–221.
- [23] Escanciano J, Pei P. Pitfalls in backtesting historical simulation models [J]. Journal of Banking and Finance, 2012, 36(8):2233–2244.
- [24] Francq C, Zakoian J. Risk-parameter estimation in volatility models [J]. Journal of Econometrics, 2015, 184(1):158–173.
- [25] Gouriéroux C, Zakoian J. Estimation-adjusted VaR [J]. Econometric Theory, 2013, 29(4):735–770.
- [26] Hansen B. Autoregressive conditional density estimation [J]. International Economic Review, 1994, 35(3):705–730.
- [27] Hong Y, Lee T. Diagnostic checking for adequacy of linear and nonlinear time series models [J]. Econometric Theory, 2003, 19(6):1065–1121.
- [28] Hong Y, Li H. Nonparametric specification testing for continuous-time models with applications to term structure of interest rates [J]. The Review of Financial Studies, 2005, 18(1):37–84.
- [29] Hurlin C, Tokpavi S. Backtesting value-at-risk accuracy: A simple new test [J]. Journal of Risk, 2007, 9(9):19–37.
- [30] Jorion P. Value at risk: The new benchmark for managing financial risk [M]. New York: McGraw-Hill,

- 2006.
- [31] Kerkhof J, Melenberg B. Backtesting for risk-based regulatory capital [J]. *Journal of Banking and Finance*, 2004, 28(8):1845–1865.
- [32] Kourouma L, Dupre D, Sanfilippo G, et al. Extreme value at risk and expected shortfall during financial crisis [R]. Working Paper, 2011.
- [33] Kupiec P. Techniques for verifying the accuracy of risk measurement models [J]. *The Journal of Derivatives*, 1995, 3(2):73–84.
- [34] McNeil A, Frey R. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach [J]. *Journal of Empirical Finance*, 2000, 7(3/4):271–300.
- [35] O’Brien J, Szerszen P. An evaluation of bank VaR measures for market risk during and before the financial crisis [R]. FEDS Working Paper No. 2014–21, 2014.
- [36] Perignon C, Smith D. A new approach to comparing VaR estimation methods [J]. *Journal of Derivatives*, 2008, 16(2):54–66.
- [37] Rosenblatt M. Remarks on a multivariate transformation [J]. *Annals of Mathematical Statistics*, 1952, 23:470–472.
- [38] Shiryaev A. *Probability* [M]. New York: Springer, 1996.
- [39] van der Vaart A, Wellner J. *Weak convergence and empirical processes* [M]. New York: Springer, 1996.
- [40] Wong W. Backtesting trading risk of commercial banks using expected shortfall [J]. *Journal of Banking & Finance*, 2008, 32(7):1404–1415.
- [41] Wong W. Backtesting value-at-risk based on tail losses [J]. *Journal of Empirical Finance*, 2010, 17(3): 526–538.
- [42] Wu W. M-estimation of linear models with dependent errors [J]. *Annals of Statistics*, 2007, 35(2):495–521.
- [43] Yamai Y, Yoshida T. Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization [J]. *Monetary and Economic Studies*, 2002, 20(3):87–122.
- [44] Yamai Y, Yoshida T. Value-at-risk versus expected shortfall: A practical perspective [J]. *Journal of Banking and Finance*, 2005, 29(4):99–115.

## 附录 A: 假设

本节介绍检验方法所需的假设和一些公式。我们首先介绍一些符号。 $\|\cdot\|$  表示欧几里德范数,  $C$  表示各个表达式中的一般常数, 可以随表达式而变化。

为了完整性, 我们将展示考虑了转换函数  $\varphi(u_t)$  的一般化结果。 $\varphi \in \bar{\Psi}$ ,  $\bar{\Psi}$  是测度函数集合,  $\varphi: [0, 1] \rightarrow R$ ,  $\bar{\Psi}$  满足左极限右连续 (cadlag)、有限变化或者非递减。累积碰撞序列是满足这些条件的特殊情况:

$$\varphi(u_t) = \frac{1}{\alpha}(\alpha - u_t)1(u_t \leq u_t) \quad (\text{A.1})$$

正文是在满足下述假设下对累计碰撞序列进行了分析, 但在此附录中我们以一般化的  $\bar{\Psi}$  展开分析。简化下述符号:  $u_t(\theta) = G(Y_t, \Omega_{t-1}, \theta)$ ,  $u_t \equiv u_t(\theta_0)$ ,  $c_\varphi = E[\varphi(u_t)]$  以及  $v_\varphi = \text{var}(\varphi(u_t))$ 。具体而言, 对 (A.1) 式中的  $\varphi$ , 我们以  $c_{ES} = \alpha/2$ ,  $v_{ES} = \alpha(1/3 - \alpha/4)$  分别对应它的均值和方差; 将  $\Theta_0$  记为  $\theta_0 \in \Theta$  的任一邻域。考虑下述假设:

**假设 A0:** 给定信息集  $\Omega_{t-1}$ ,  $Y_t$  的条件分布是  $G(\cdot, \Omega_{t-1}, \theta)$ 。

**假设 A1:**  $\{Y_t, X_t\}_{t=-T+1}^n$  满足严平稳和遍历性。

**假设 A2:**  $\hat{\theta}_T$  是  $\theta_0$  的  $\sqrt{T}$  一致估计量,  $\theta_0$  是  $\Theta$  的内点。并且,  $\hat{\theta}_T$  满足如下的渐进 Bahadur

扩展:  $\sqrt{T}(\theta_T - \theta_0) = \frac{1}{\sqrt{T}} \sum_{t=-T+1}^0 l_t + o_p(1)$ ,  $l_t$  满足  $E(l_t | \Omega_{t-1}) = 0$ ,  $E(l_t l_t')$  存在且是正定矩阵。

假设 A3: 信息截尾 (*information truncation*) 的影响要满足,

$$\sup_{\theta \in \Theta_n} \frac{1}{\sqrt{n}} \sum_{t=1}^n |\varphi(G(Y_t, \hat{\Omega}_{t-1}, \theta)) - \varphi(G(Y_t, \Omega_{t-1}, \theta))| = O_p(1).$$

假设 A4:  $F_t(\theta, x)$  关于  $\theta$  和  $x \in [0, 1]$  连续可微, 并且关于  $v_\varphi$  也几乎肯定连续可微,

$$E \left[ \sup_{\theta \in \Theta_n, 0 \leq x \leq 1} \left\| \frac{\partial F_t(\theta, x)}{\partial x} \right\| \right] < C \text{ and } E \left[ \int_0^1 \sup_{\theta \in \Theta_n} \left\| \frac{\partial F_t(\theta, x)}{\partial \theta} \right\| d\varphi(x) \right] < C.$$

其中, A0 在文献中常用, 它假设模型是正确设定的。该假设也可以放松到如下形式, 但不会改变本文的理论。  $P(Y_t \leq y | \Omega_{t-1}) = G(y, \Omega_{t-1}, \theta_0)$  for all  $y \leq G^{-1}(\alpha, \Omega_{t-1}, \theta_0)$ 。

假设 A1 是为了阐述的方便。我们的结果, 可以扩展到一些非平稳和非遍历性的序列中 (Escanciano, 2007)。大多数常用的估计量可以满足假设 A2, 比如 (拟) 极大似然估计量和广义距估计量 (Bose, 1998; Wu, 2007)。假设 A3 是信息截取的影响, 由于很久之前的观测数据不可获得, 所以产生了信息截取问题。很多有限二阶距的时间序列模型都可以满足假设 A3, 包括平稳不可逆的 ARMA 过程、GARCH 过程等等 (Bai, 2003; Hong 和 Li, 2003)。马尔科夫过程则不需要该假设。假设 A4 是为了满足某些实证过程和大数定律的渐进同等连续性。

#### 附录 B: 估计效应 (*Estimation Effect*) 的表达式

本节给出一般规模尺度模型中的  $\hat{r}_{ES}$ ,  $\hat{r}_j$ ,  $\hat{R}_{Var}$  以及  $\hat{R}_j$  详细表达式:

$$Y_t = \mu_t + \sigma_t \varepsilon_t \quad (\text{B.1})$$

其中,  $\mu_t = \mu(\Omega_{t-1}, \theta_0) = E[Y_t | \Omega_{t-1}]$ ,  $\sigma_t^2 = \sigma^2(\Omega_{t-1}, \theta_0) = \text{Var}[Y_t | \Omega_{t-1}]$ ; 并且  $\varepsilon_t$  的累积分布函数是  $G_\varepsilon(\cdot)$ , 概率密度函数是  $g_\varepsilon(\cdot)$ 。

$$\text{推导可得: } \frac{\partial F_t(\theta_0, x)}{\partial \theta} = g_\varepsilon(G_\varepsilon^{-1}(x)) \frac{\dot{u}_t + G_\varepsilon^{-1}(x) \dot{\sigma}_t}{\sigma_t} \quad (\text{B.2})$$

其中,  $\dot{\mu}_t = \partial \mu(\Omega_{t-1}, \theta_0) / \partial \theta$ , 以及  $\dot{\sigma}_t = \partial \sigma(\Omega_{t-1}, \theta_0) / \partial \theta$ 。因此,

$$\begin{aligned} R_{ES} &= \frac{1}{\alpha} E \left\{ \int_0^\alpha \frac{\partial F_t(\theta_0, x)}{\partial \theta} dx \right\} = \frac{1}{\alpha} E \left\{ \int_0^\alpha g_\varepsilon(G_\varepsilon^{-1}(x)) \frac{\dot{u}_t + G_\varepsilon^{-1}(x) \dot{\sigma}_t}{\sigma_t} dx \right\} \\ &= \frac{1}{\alpha} E \left\{ \int_{G_\varepsilon^{-1}(0)}^{G_\varepsilon^{-1}(\alpha)} g_\varepsilon(z) \frac{\dot{u}_t + z \dot{\sigma}_t}{\sigma_t} dG_\varepsilon(z) \right\} = \frac{1}{\alpha} E \left\{ g_\varepsilon(\varepsilon_t) 1(\varepsilon_t \leq G_\varepsilon^{-1}(\alpha)) \frac{\dot{u}_t + \varepsilon_t \dot{\sigma}_t}{\sigma_t} \right\}, \end{aligned}$$

$$\text{以及 } \hat{R}_{ES} = \frac{1}{\alpha n^{t-1}} \sum_{t=1}^n g_\varepsilon(\hat{\varepsilon}_t) 1(\hat{\varepsilon}_t \leq G_\varepsilon^{-1}(\alpha)) \frac{\dot{\mu}_t + \hat{\varepsilon}_t \dot{\sigma}_t}{\hat{\sigma}_t},$$

其中,  $\dot{\mu}_t = \partial \mu(\Omega_{t-1}, \theta_0) / \partial \theta$ , 以及  $\dot{\sigma}_t = \partial \sigma(\Omega_{t-1}, \theta_0) / \partial \theta$ 。并且,

$$\begin{aligned} R_j &= \frac{1}{v_{ES}(\alpha)} E \left\{ (H_{t-j}(\alpha) - \alpha/2) \int_0^\alpha g_\varepsilon(G_\varepsilon^{-1}(x)) \frac{\dot{u}_t + G_\varepsilon^{-1}(x) \dot{\sigma}_t}{\sigma_t} dx \right\} \\ &= \frac{1}{v_{ES}(\alpha)} E \left\{ (H_{t-j}(\alpha) - \alpha/2) \int_{G_\varepsilon^{-1}(0)}^{G_\varepsilon^{-1}(\alpha)} g_\varepsilon(z) \frac{\dot{u}_t + z \dot{\sigma}_t}{\sigma_t} dG_\varepsilon(z) \right\} \\ &= \frac{1}{v_{ES}(\alpha)} E \left\{ (H_{t-j}(\alpha) - \alpha/2) g_\varepsilon(\varepsilon_t) 1(\varepsilon_t \leq G_\varepsilon^{-1}(\alpha)) \frac{\dot{u}_t + \varepsilon_t \dot{\sigma}_t}{\sigma_t} \right\}, \end{aligned}$$

其对应的可行估计是,

$$\hat{R}_j = \frac{1}{v_{ES}(\alpha)} \frac{1}{n-j} \sum_{i=j+1}^n (\hat{H}_{i-j}(\alpha) - \alpha/2) g_\varepsilon(\hat{\varepsilon}_i) 1(\hat{\varepsilon}_i \leq G_\varepsilon^{-1}(\alpha)) \frac{\hat{\mu}_i + \hat{\varepsilon}_i \hat{\sigma}_i}{\hat{\sigma}_i}.$$

类似地,对修正的 VaR 后验分析而言,我们有:

$$\tilde{R}_{VaR} = \frac{g_\varepsilon(G_\varepsilon^{-1}(\alpha))}{n} \sum_{i=1}^n \frac{\hat{\mu}_i + G_\varepsilon^{-1}(\alpha) \hat{\sigma}_i}{\hat{\sigma}_i},$$

以及,  $\tilde{R}_j = \frac{-g_\varepsilon(G_\varepsilon^{-1}(\alpha))}{(n-j)\alpha(1-\alpha)} \sum_{i=j+1}^n \frac{\hat{\mu}_i + G_\varepsilon^{-1}(\alpha) \hat{\sigma}_i}{\hat{\sigma}_i} (1(\hat{u}_{i-j} \leq \alpha) - \alpha).$

附录 C:证明

对于一般化的定理 1 和定理 2,我们将再给予证明。这种一般性体现在  $\varphi(u_i), \varphi \in \bar{\Psi}$ , 而  $H(\cdot)$  和  $h(\cdot)$  只是  $\varphi(u_i)$  的特殊形式。这种一般化的结果,可以用于构建不同于 ES 的一致性风险度量的后验分析。

**定理 A1:**在假设 A0-A4 下,  $\frac{1}{\sqrt{n}} \sum_{i=1}^n (\varphi(\hat{u}_i) - c_\varphi) \rightarrow^d N(0, \sigma_\varphi^2),$

其中,  $\sigma_\varphi^2 = v_\varphi + \lambda R'_\varphi E[l_i l'_i] R_\varphi, R_\varphi = -E\left\{ \int_0^1 \frac{\partial F_t(\theta_0, x)}{\theta} d\varphi(x) \right\}.$

证明定理 A1:首先考虑没有信息截尾的情况。

如果  $G$  仅仅依赖于有限阶滞后的  $Y_i$  和  $X_i$ ,则不存在信息截尾的情形。可以从 Escanciano 和 Olmo(2010)定理 1 的证明中得到类似的论据。在假设 A0-A2 以及 A4 下,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(\hat{u}_i) - E[\varphi(\hat{u}_i) | \Omega_{i-1}]] - \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(u_i) - E[\varphi(u_i) | \Omega_{i-1}]] = o_P(1).$$

因此,

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(\hat{u}_i) - c_\varphi] &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(u_i) - c_\varphi] + \frac{1}{\sqrt{n}} \sum_{i=1}^n (\hat{\theta}_T - \theta_0)' \frac{\partial E[\varphi(u_i(\tilde{\theta}_T)) | \Omega_{i-1}]}{\partial \theta} + o_P(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(u_i) - c_\varphi] + \sqrt{\lambda} \sqrt{T} (\hat{\theta}_T - \theta_0)' \frac{1}{n} \sum_{i=1}^n \frac{\partial E[\varphi(u_i(\tilde{\theta}_T)) | \Omega_{i-1}]}{\partial \theta} + o_P(1), \end{aligned}$$

其中,  $\tilde{\theta}_T$  是介于  $\hat{\theta}_T$  和  $\theta_0$  之间的中值点。注意:

$$\frac{\partial E[\varphi(u_i(\theta)) | \Omega_{i-1}]}{\partial \theta} = \frac{\partial (\int_0^1 \varphi(x) dF_i(\theta, x))}{\partial \theta} = - \frac{\partial \int_0^1 F_i(\theta, x) d\varphi(x)}{\partial \theta}.$$

因此,

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(\hat{u}_i) - c_\varphi] &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(u_i) - c_\varphi] + \sqrt{\lambda} \sqrt{T} (\hat{\theta}_T - \theta_0)' R_\varphi + o_P(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [\varphi(u_i) - c_\varphi] + \sqrt{\lambda} R'_\varphi \frac{1}{\sqrt{T}} \sum_{i=-T+1}^0 l_i + o_P(1) \end{aligned}$$

注意,等式右侧第一项依分布收敛于  $N(0, v_\varphi)$ , 并且前两项的协方差为零,因为第一项是针对样本外观测的式子,第二项是针对样本内观测的式子。从而完成了定理 A1 的证明。

其次,考虑信息存在截尾的情形。

定义  $\tilde{u}_t = G(Y_t, \Omega_{t-1}, \hat{\theta}_T)$ , 我们可得,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n [\varphi(\hat{u}_t) - c_\varphi] = \frac{1}{\sqrt{n}} \sum_{t=1}^n [\varphi(\hat{u}_t) - \varphi(\tilde{u}_t)] + \frac{1}{\sqrt{n}} \sum_{t=1}^n [\varphi(\tilde{u}_t) - c_\varphi].$$

假设 A3 隐含了上式的第一项是  $o_P(1)$ 。然后, 针对没有信息截尾的  $1/\sqrt{n} \sum_{t=1}^n (\varphi(\hat{u}_t) - c_\varphi)$  的证明可以直接应用到  $1/\sqrt{n} \sum_{t=1}^n (\varphi(\tilde{u}_t) - c_\varphi)$ , 从而完成了定理 A1 的证明。

定理 1、推论 1 以及推论 3 的证明, 可以直接从定理 A1 中得出。

接下来, 我们证明一般化的定理 2, 为此我们需要如下引理。定义,

$$R_{nj}(x, y) = \frac{1}{n-j} \sum_{t=1+j}^n \{1(u_t \leq x) - x\} \{1(u_{t-j} \leq y) - y\}$$

$$\hat{R}_{nj}(x, y) = \frac{1}{n-j} \sum_{t=1+j}^n \{1(\hat{u}_t \leq x) - x\} \{1(\hat{u}_{t-j} \leq y) - y\}.$$

引理 A1: 在假设 A0-A4 下, 我们有:

$$\sup_{0 \leq x \leq 1, 0 \leq y \leq 1} \left| \sqrt{n-j} [\hat{R}_{nj}(x, y) - R_{nj}(x, y)] - \sqrt{\lambda} \sqrt{T} (\hat{\theta}_T - \theta)' E_j(x, y) \right| = o_P(1),$$

其中,  $E_j(x, y) = E \left\{ \frac{\partial F_t(\theta_0, x)}{\partial \theta} [1(u_{t-j} \leq y) - y] \right\}$ 。

引理 A1 是 Du(2015) 定理 1 的一个特殊情形, 所以省略此处的证明。

引理 A2: 定义  $R(x, y)$  是  $[0, 1]^2$  上的函数。对于任意的  $0 \leq y \leq 1$  都有  $R(\cdot, y) \in \bar{\Psi}$ , 对于任意的  $0 \leq x \leq 1$  都有  $R(x, \cdot) \in \bar{\Psi}$ , 以及在边界上  $R=0$ 。定义  $l([0, 1]^2)$  是所有  $R(x, y)$  函数的矩阵空间, 并且拥有上确界函数。然后, 如下映射在  $R$  上对任意的  $\varphi \in \bar{\Psi}$  都是连续的。

$$R \rightarrow \int_0^1 \int_0^1 \varphi(x) \varphi(y) R(dx, dy)$$

证明引理 A2: 根据分部积分定理 (Shiryayev, 1996, pp206, 定理 11) 以及  $R$  的定义, 我们有  $\int_0^1 \int_0^1 \varphi(x) \varphi(y) R(dx, dy) = \int_0^1 \int_0^1 R(x, y) \varphi(dx) \varphi(dy)$ 。

注意, 下式对任意的  $R_1, R_2 \in l([0, 1]^2)$ , 以及  $\int_0^1 |\varphi(dx)| < \infty$  都成立。从而, 完成了证明。

$$\left| \int_0^1 \int_0^1 R_1(x, y) \varphi(dx) \varphi(dy) - \int_0^1 \int_0^1 R_2(x, y) \varphi(dx) \varphi(dy) \right|$$

$$\leq \sup |R_1(x, y) - R_2(x, y)| \int_0^1 \int_0^1 |\varphi(dx) \varphi(dy)|,$$

在表述了两个引理之后, 我们将证明一般化的定理 2。定义  $\varphi(u_t)$  滞后  $j$  阶 ( $j > 0$ ) 的自协方差和自相关系数是  $\gamma_j = Cov(\varphi(u_t), \varphi(u_{t-j}))$  和  $\rho_j = \frac{\gamma_j}{\gamma_0}$ , 样本  $\{u_t\}_{t=1}^n$  对应的  $\gamma_j$  和  $\rho_j$  是,

$\gamma_{nj} = \frac{1}{n-j} \sum_{t=1+j}^n (\varphi(u_t) - c_t)(\varphi(u_{t-j}) - c_t)$  和  $\rho_{nj} = \frac{\gamma_{nj}}{\gamma_{n0}}$ , 但是由于  $\{u_t\}_{t=1}^n$  不可观测, 所以我们采用  $\hat{u}_t$  替换  $u_t$ ,

$$\hat{\gamma}_{nj} = \frac{1}{n-j} \sum_{t=1+j}^n (\varphi(\hat{u}_t) - c_t)(\varphi(\hat{u}_{t-j}) - c_t) \text{ 和 } \hat{\rho}_{nj} = \frac{\hat{\gamma}_{nj}}{\hat{\gamma}_{n0}}.$$

定理 A2: 在假设 A0-A4 下,  $\sqrt{n} \hat{\rho}_n^{(m)} \rightarrow^d N(0, \Sigma)$ 。  $\Sigma$  的第  $ij$  个元素是  $\sum_{ij} = \delta_{ij} + \lambda R'_{ij} E[l_i l'_i] R_j$ ,



其中,  $R_j$  见下式,  $\delta_{ij}$  是克罗内克函数,  $\delta_{ij}$  在  $i=j$  时取值 1, 否则取值 0。

$$R_j = \frac{-1}{v_\varphi} E \left\{ (\varphi(u_{t-j}) - c_\varphi) \int_0^1 \frac{\partial F_t(\theta_0, x)}{\partial \theta} d\varphi(x) \right\} \quad (C.1)$$

证明定理 A2: 首先考虑没有信息截尾的情形。

如果  $G$  仅仅依赖于有限阶滞后的  $Y_t$  和  $X_t$ , 则不存在信息截尾的情形。注意,

$$\begin{aligned} \sqrt{n-j} \hat{\gamma}_{nj} &= \sqrt{n-j} \int_0^1 \int_0^1 \varphi(x) \varphi(y) \hat{R}_{nj}(dx, dy) \\ \sqrt{n-j} \gamma_{nj} &= \sqrt{n-j} \int_0^1 \int_0^1 \varphi(x) \varphi(y) R_{nj}(dx, dy). \end{aligned}$$

通过引理 1 和引理 2 可得,

$$\sqrt{n-j} \hat{\gamma}_{nj} = \sqrt{n-j} \gamma_{nj} + \sqrt{\lambda} \sqrt{T} (\hat{\theta}_T - \theta_0)' \int_0^1 \int_0^1 \varphi(x) \varphi(y) E_j(dx, dy) + o_p(1) \quad (C.2)$$

其中,

$$\begin{aligned} \int_0^1 \int_0^1 \varphi(x) \varphi(y) E_j(dx, dy) &= \int_0^1 \int_0^1 E \left\{ \varphi(x) \frac{\partial^2 F_t(\theta_0, x)}{\partial x \partial \theta} dx \varphi(y) [1(u_{t-j} \leq dy) - dy] \right\} \\ &= \int_0^1 E \left\{ \varphi(x) \frac{\partial^2 F_t(\theta_0, x)}{\partial x \partial \theta} dx (\varphi(u_{t-j}) - c_\varphi) \right\} \\ &= E \left\{ \int_0^1 \varphi(x) \frac{\partial^2 F_t(\theta_0, x)}{\partial x \partial \theta} dx (\varphi(u_{t-j}) - c_\varphi) \right\} \\ &= E \left\{ \left[ \varphi(x) \frac{\partial F_t(\theta_0, x)}{\partial \theta} \right]_0^1 - \int_0^1 \frac{\partial F_t(\theta_0, x)}{\partial \theta} dx \right\} (\varphi(u_{t-j}) - c_\varphi) \\ &= -E \left\{ \int_0^1 \frac{\partial F_t(\theta_0, x)}{\partial \theta} d\varphi(x) (\varphi(u_{t-j}) - c_\varphi) \right\} \\ &= v_\varphi R_j \end{aligned}$$

(C.1) 式定义了  $R_j$ 。由假设 A4 可对上式进行期望与积分互换, 分部积分原理参见 Shiryaev (1996, pp.206)。

因此, 我们证明了  $\sqrt{n-j} (\hat{\gamma}_{nj} - \gamma_{nj}) = \sqrt{\lambda} \sqrt{T} (\hat{\theta}_T - \theta_0)' v_\varphi R_j + o_p(1)$ 。

然后, 我们可得出,

$$\sqrt{n-j} (\hat{\rho}_{nj} - \rho_{nj}) = \frac{\sqrt{n-j} (\hat{\gamma}_{nj} - \gamma_{nj})}{v_\varphi} + o_p(1) = R'_j \sqrt{\lambda} \sqrt{T} (\theta_T - \theta_0) + o_p(1)。$$

因此,  $\sqrt{n-j} \hat{\rho}_{nj} = \sqrt{n-j} \rho_{nj} + R'_j \sqrt{\lambda} \sqrt{T} (\theta_T - \theta_0) + o_p(1)$

$$= \frac{1}{\sqrt{n-j} v_\varphi} \sum_{t=1+j}^n (\varphi(u_t) - c_\varphi) (\varphi(u_{t-j}) - c_\varphi) + R'_j \sqrt{\lambda} \frac{1}{\sqrt{T}} \sum_{t=-T+1}^0 l_t + o_p(1)。$$

注意,  $\sqrt{n} (\rho_{n1}, \rho_{n2}, \dots, \rho_{nm})' \rightarrow^d N(0, I_m)$ , 等式右侧前两项的协方差为 0, 因为第一项是样本外观测的式子, 第二项是样本内观测的式子。综上, 我们完成定理 A2 的证明。

接下来, 我们考虑存在信息截尾的情形。

定义  $\tilde{u}_t = G(Y_t, \Omega_{t-1}, \hat{\theta}_T)$ , 以及  $\tilde{\gamma}_{nj} = 1/(n-j) \sum_{t=1+j}^n (\varphi(\tilde{u}_t - c_\varphi)) (\varphi(\tilde{u}_{t-j} - c_\varphi))$ 。从而,

$$\sqrt{n-j} (\hat{\gamma}_{nj} - \gamma_{nj}) = \sqrt{n-j} (\hat{\gamma}_{nj} - \tilde{\gamma}_{nj}) + \sqrt{n-j} (\tilde{\gamma}_{nj} - \gamma_{nj})。$$

通过证明可知, 上式中的第一项是  $o_p(1)$ 。因为,

$$\begin{aligned}
\sqrt{n-j}(\hat{\gamma}_{nj} - \tilde{\gamma}_{nj}) &= \frac{1}{\sqrt{n-j}} \sum_{t=1+j}^n [(\varphi(\hat{u}_t) - c_\varphi)(\varphi(\hat{u}_{t-j}) - c_\varphi) - (\varphi(\tilde{u}_t) - c_\varphi)(\varphi(\tilde{u}_{t-j}) - c_\varphi)] \\
&= \frac{1}{\sqrt{n-j}} \sum_{t=1+j}^n (\varphi(\hat{u}_t) - \varphi(\tilde{u}_t))(\varphi(\hat{u}_{t-j}) - c_\varphi) \\
&\quad + \frac{1}{\sqrt{n-j}} \sum_{t=1+j}^n (\varphi(\tilde{u}_t) - c_\varphi)(\varphi(\hat{u}_{t-j}) - \varphi(\tilde{u}_{t-j})).
\end{aligned}$$

以及从假设 A3 可知,  $\varphi(\tilde{u}_t) = O_p(1)$ 。不存在信息截尾的  $\sqrt{n-j}(\hat{\gamma}_{nj} - \tilde{\gamma}_{nj})$  论证可以直接应用到  $\sqrt{n-j}(\tilde{\gamma}_{nj} - \gamma_{nj})$ , 从而完成了定理 A2 的证明。

定理 2 的证明: 定义  $\Sigma = Q\Lambda Q'$ , 其中  $Q$  是正交矩阵,  $\Lambda$  是对角矩阵, 对角线元素是  $\{\pi\}_{j=1}^m$ 。从定理 A2 可知,  $Q'\sqrt{n}\rho_n^{(m)} \rightarrow^d N(0, Q\Sigma Q') = N(0, \Lambda)$ 。然后, 可以通过  $C_{ES}(m) = (Q'\sqrt{n}\rho_n^{(m)})'(Q'\sqrt{n}\rho_n^{(m)})$  完成定理 2 的证明。

推论 2 和推论 4 可从定理 2 得出。

## Backtesting Expected Shortfall: Accounting for Tail Risk

Du Zaichao<sup>1</sup>, Juan Carlos Escanciano<sup>2</sup>

(1. *Research Institute of Economics and Management, Southwestern  
University of Finance and Economics, Chengdu 610074, China;*

2. *Department of Economics, Indiana University, Bloomington IN 47405, USA)*

**Abstract:** The Basel Committee on Banking Supervision (BIS) has recently sanctioned Expected Shortfall (ES) as the market risk measure to be used for banking regulatory purposes, replacing the well-known Value-at-Risk (VaR). This change is motivated by the appealing theoretical properties of ES as a measure of risk and the poor properties of VaR. In particular, VaR fails to control for “tail risk”. In this transition, the major challenge faced by financial institutions is the unavailability of simple tools for evaluation of ES forecasts (i.e. backtesting ES). The main purpose of this article is to propose such tools. Specifically, we propose backtests for ES based on cumulative violations, which are the natural analogue of the commonly used backtests for VaR. We establish the asymptotic properties of the tests, and investigate their finite sample performance through some Monte Carlo simulations. An empirical application to three major stock indexes shows that VaR is generally unresponsive to extreme events such as those experienced during the recent financial crisis, while ES provides a more accurate description of the risk involved.

**Key words:** risk management; expected shortfall; backtesting; tail risk; Value-at-Risk.

(责任编辑 石头)