

文章编号: 2095-4980(2022)12-1277-08

## 基于 MES 频谱数据异常贡献度估计与后验分析

张继丹<sup>1</sup>, 肖 东<sup>2</sup>, 侯燕曦<sup>\*1</sup>

(1.复旦大学 大数据学院, 上海 200433; 2.哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150001)

**摘 要:** 近些年, 基于大数据分析模型的风险度量和控制方法研究变得越来越重要, 而风险度量模型的后验分析研究能够保障和检验所用分析技术在实际数据分析中的有效性。边际期望损失(MES)作为度量个体对系统性风险的边际贡献的重要工具, 其后验分析也是一个值得关注的问题。本文将 C. Acerbi 等提出的关于 ES 的后验分析方法进行二维变量下的延伸, 提出 2 个新的对于 MES 的统计量。模拟实验的结果表明, 在原假设分布和备择假设分布相差相对较小的情况下, 2 个统计量的统计功效均大于 D. Banulescu 等采用的统计量。实证分析的结果也表明, 对于同样的预测结果, 文中新提出的统计量在原假设的接受程度上相对更为谨慎。该方法对于大数据模型算法的后验分析具有一定的理论借鉴意义。

**关键词:** 后验算法分析; 统计功效; 边际期望损失; 系统性风险

中图分类号: TN914.42

文献标志码: A

doi: 10.11805/TKYDA2021178

## Estimation of anomaly contribution in spectrum data based on MES and backtesting

ZHANG Jidan<sup>1</sup>, XIAO Dong<sup>2</sup>, HOU Yanxi<sup>\*1</sup>

(1.School of Data Science, Fudan University, Shanghai 200433, China;

2.School of Information Engineering, Harbin Engineering University, Harbin Heilongjiang 150001, China)

**Abstract:** In recent years, the research on risk measurement and control methods based on big data analysis model has become more and more important, and the backtesting analysis on risk measurement tools can guarantee the effectiveness of the techniques used in actual data analysis. Marginal Expected Shortfall(MES) is an important tool to measure the marginal contribution of individual institutions to systemic risk, and the backtest methodologies for MES is also worthy to focus on. In this paper, the backtest method of ES in C. Acerb et al. is extended to the two-dimensional case and two backtest methodologies are proposed for MES. The results of simulation show that these two new statistics are more powerful than the statistics used in D.Banulescu et al. under situations that the difference between the null hypothesis and the alternative hypothesis is relatively small. The results of empirical analysis also support that these two new statistics proposed in this paper accept the null hypothesis more cautiously under the same prediction model. This method can give some reference for model algorithm backtesting under big data.

**Keywords:** algorithm backtesting; statistical power; Marginal Expected Shortfall; systemic risk

近年来, 随着数据量的急剧暴增以及计算性能的提高, 基于大数据的风险度量和分析模型开始盛行并推广应用于各行各业。系统性风险的概念最早起源于金融领域, 因为其极强的破坏性, 不少学者提出各种系统风险度量工具。边际期望损失(MES)则是其中的度量工具之一, 指的是在市场遭受比较极端损失情况下金融机构个体的期望损失, 衡量的是在危机发生时金融机构个体对整个系统风险的边际贡献<sup>[1]</sup>。

一般通信系统中传输的信号都具有一定的不确定性, 因此都属于随机信号(Random Signal), 它的幅度未可预

收稿日期: 2021-04-26; 修回日期: 2021-05-18

基金项目: 上海市自然科学基金面上项目资助项目(20ZR140390022)

\*通信作者: 侯燕曦 email:yxhou@fudan.edu.cn

知,任何一次观测只代表其在变动范围中可能产生的结果之一,但又服从一定的统计特性,又称不确定信号。

从信息论的观点,对接受者而言只有信号表现出某种不确定性才蕴含信息,关于更多随机信号的建模和预测详见赵淑清等<sup>[2]</sup>、王永德<sup>[3]</sup>等文献。而对随机信号的建模也有学者陆续提出各种方法,2001年,张海勇<sup>[4]</sup>对局域波法进行深入研究,研究了影响局域波法性能的一些因素,并提出改善该分析方法效能的具体指标及措施;2003年,张海勇<sup>[5]</sup>分别从自适应自回归模型谱分析法等几个方面综述了非平稳随机信号参数模型分析方法的发展现状,杨娟<sup>[6]</sup>建立了一种能够描述随机信号结构基本特征的双参数脉冲信号统计模型。

环境信号的异常是频谱感知中通常会关注的问题。随着无线电技术的迅猛发展,越来越多的通信设备利用电磁波传播信息,电磁空间环境日益复杂,对于电磁空间环境异常信号的检测也受到了越来越多的关注。近年来,频谱异常检测问题已得到了广泛研究,并取得了一定的成果。根据频谱异常检测所采用的方法,现有的研究可分为两类:第一类是利用信号传输时的各个特征进行频谱异常检测,包括接收信号强度空间分布、接收信号强度变化、物理信道属性、直接路径的幅度差、受扰路径的幅度差以及通话时间利用率等,相关成果见文献[7-10]。第二类方法是采用机器学习的方法进行频谱异常检测,夏伟等<sup>[11]</sup>提出了一种基于单类支持向量机(v-OCSVM)模型的频谱异常检测方法,通过实验检测和ROC曲线分析,表明v-OCSVM模型用于频谱异常检测是可行的。闫戈<sup>[12]</sup>提出了一种基于最小二乘法的动态逻辑回归学习模型,并设计了基于朴素贝叶斯分类算法的频谱异常检测和异常预测的实现方案;Li等<sup>[13]</sup>提出了一种新的基于长短期记忆(Long Short-Term Memory, LSTM)的数据驱动频谱异常检测模型,通过学习频谱数据的时间相关性来预测下一时刻的幅度值,根据预测值与测量值之间的误差来发现异常频谱。通常,当某个信号的能量偏离正常值较远时,会认为它有异常的可能,当偏离了正常值较远并持续了给定的一段时间时,则会认为该信号异常。李忠等<sup>[14]</sup>采用滑动均值 $\pm 2$ 倍均方差作为阈值提取异常,并阐明其合理性,表明其能够快速定量提取前兆数据异常信号。张小飞等<sup>[15]</sup>设计了一种基于能量上下界的缓存告警检测算法,对原有相关算法进行改进,解除了对固定阈值的依赖。由信号的傅里叶变化可知,信号由多个频率的正弦波叠加而成,不同频率分量出现异常都会反映在信号上。频谱上不同频点的电磁信号的能量波动对于电磁空间环境异常信号的影响是不同的,正如MES起初的含义是用来衡量个体风险对整体风险的边际贡献,本文中将其用于衡量每个分信号的异常波动对叠加后信号的异常程度的贡献。

后验分析,也即检验真实值是否与估计值一致,是为了确保一个估计方法在实际数据分析中的有效性。对于比较主流、提出时间比较早的风险度量,譬如在险价值(Value at Risk, VaR),其后验分析方法相对来说比较多,相关理论发展得也比较完善。但是关于MES的后验分析方法目前文献中涉及得还不是太多。从其定义可以看到, MES可看成是ES(Expected Shortfall)(也叫期望损失,代表了一个随机变量在大于自身较高分位数下的条件期望,可表示为 $E(X)=E(X|X>VaR(X))$ 在二维变量下的延伸。由于ES不是可诱发的,因此曾经在很长一段时间内被认为是无法进行后验分析的。出于ES在实际应用的需要,许多学者也相继对其后验分析方法进行探究,这些研究成果既促进了ES在实际中的使用,也为对其他风险度量工具进行后验分析提供了思路和借鉴。其中, Acerbi等<sup>[16]</sup>较为详细地讨论了对ES做后验分析的可能性,并提出了3个不依赖于分布的、非参的ES后验分析方法; Escanciano and Du<sup>[17]</sup>借鉴了VaR的后验分析中通过碰撞序列进行假设检验分析的方式,基于累计碰撞序列对ES进行后验分析;此外,和系统风险度量相关的后验分析方法也相继有提出,但相对来说还是比较稀少。其中, Banulescu等<sup>[18]</sup>巧妙地借鉴了Escanciano and Du基于累积碰撞序列来对ES进行后验分析的方法,构造了一套关于MES及其相关系统性风险度量的检验方法。

后验分析相当于一个统计上的假设检验,对于一个假设检验,其统计功效是值得关注的问题。针对VaR的后验分析问题, Dumitrescu等<sup>[19]</sup>提到,后验分析的功效,尤其是在小样本下,对于评估检验结果起着非常关键的作用。而一般的检验通常有较低的功效,也即一般不会拒绝原假设的成立。Hurlin和Tokpavi<sup>[20]</sup>也阐述了相似的观点,该文中展示了大部分的后验分析都不会拒绝原假设,意味着VaR预测结果通常都被认为是有效的。Acerbi等在提出了3个新的ES后验分析方法的同时,也将其统计功效与VaR进行对比,并表明该方法相比于VaR的后验分析有更大的功效。而据了解,对于系统性风险度量的后验分析的功效问题暂时还没有相关文献涉及。

本文对Acerbi等提出的ES进行后验分析的方法进行维度上的延伸,从而提出2个新的关于MES后验分析的统计量。设计模拟实验来检验其统计功效,并将其与Banulescu等提出的用于检验MES的方法的统计功效进行对比。最后将这2个方法应用于电磁频谱数据的实证分析。

## 1 后验分析方法介绍

首先给出本文中MES的定义:令 $X_t, Y_t$ 分别代表个体和整个系统在 $t$ 时刻的取值,其中 $t=1, 2, \dots, T$ 。在 $(X_t, Y_t)$

有连续分布函数  $F_t$  的条件下，给定置信水平为  $1-\alpha$ ，则个体在  $t$  时刻的边际期望损失可表示为：

$$\text{MES}_\alpha(X_t) = E(X_t | Y_t \geq \text{VaR}_\alpha(Y_t)) \quad (1)$$

式中： $\text{VaR}_\alpha(Y_t) = \inf\{y_\alpha : 1-\alpha \leq F_{Y_t}(y_\alpha)\}$ ，相当于  $Y_t$  的  $1-\alpha$  分位数。如果真实的分布  $F_t$  是已知的，则可以根据定义来计算其对应的 MES 度量值。但在实际中， $F_t$  其实是未知的，往往需要通过一个预测的分布，暂且记作  $P_t$ ，来计算个体在  $t$  时刻的边际期望损失。这样一来自然就需要考虑一个问题：能否接受在  $P_t$  下计算得到的 MES 就是真实的 MES 这个假设，这正是后验分析需要解决的问题。这一节将介绍 2 个 MES 的后验分析方法。

### 1.1 方法一

该方法是 Acerbi 等关于 ES 的后验分析方法在二维变量下的延伸。本文同样采用一般的假设检验框架来进行 MES 的无条件后验分析。

#### 1.1.1 提出假设与构造统计量

取原假设和备择假设分别为  $H_0: \text{MES}_\alpha^P(X_t) = \text{MES}_\alpha(X_t), \forall t \in T$ ,  $H_1: \text{MES}_\alpha^P(X_t) \leq \text{MES}_\alpha(X_t), \forall t \in T$ , 存在  $t \in T$ , 使  $\text{MES}_\alpha^P(X_t) < \text{MES}_\alpha(X_t)$ 。其中， $\text{MES}_\alpha^P$  表示在  $P_t$  下计算得到的 MES。本文进行的其实是单边的假设检验，也即检验的是在  $P_t$  下计算得到的 MES 是否低估其真实值。由式(1)可以定义如下的统计量：

$$Z_1(X) = \sum_{t=1}^T \frac{X_t I_{\{Y_t \geq \text{VaR}_\alpha^P(Y_t)\}}}{\text{MES}_\alpha^P(X_t) \sum_{t=1}^T I_{\{Y_t \geq \text{VaR}_\alpha^P(Y_t)\}}} - 1 \quad (2)$$

$$Z_2(X) = \sum_{t=1}^T \frac{X_t I_{\{Y_t \geq \text{VaR}_\alpha^P(Y_t)\}}}{\alpha \text{MES}_\alpha^P(X_t) T} - 1 \quad (3)$$

不难猜想，在  $\text{MES}_\alpha^P(X_t)$  与真实的 MES 非常接近的情况下，统计量  $Z_1(X)$  和  $Z_2(X)$  应该会非常接近零。如果  $\text{MES}_\alpha^P(X_t)$  低估了真实的 MES，则统计量的值会显著大于零；如果  $\text{MES}_\alpha^P(X_t)$  高估了真实的 MES，则统计量的值自然小于零。此外也可以看到，统计量  $Z_1$  和  $Z_2$  不同的地方在于分母，这体现了两者在出发点上略微的差异。 $Z_1$  在处理实际损失时是先对所有  $Y$  超过阈值的  $X$  求和，然后除以实际个数，相当于是对这部分的  $X$  取平均； $Z_2$  对所有  $Y$  超过阈值的  $X$  求和后除以的则是期望的个数，也即  $\alpha T$ 。同时注意到有  $E \sum_{t=1}^T I_{\{Y_t \geq \text{VaR}_\alpha^P(Y_t)\}} = \alpha T$  成立，也即如果  $Y_t \geq \text{VaR}_\alpha^P$  的估计是准确的， $Y$  超过阈值的实际个数应该会很接近于期望的个数。这个微小的区别使得这 2 个统计量之间也有所不同，其功效在下一节的模拟实验中进行展示。

#### 1.1.2 显著性水平

在通过某种方式计算得到了  $P_t$  后，将需要预测的样本数据代入式(2)和式(3)中，得到了统计量的值，也即  $Z_1$  和  $Z_2$  的一个实现，暂且记为  $Z_1(\bar{x})$  和  $Z_2(\bar{x})$ 。为了进行后验分析，需要进一步获得  $Z_1(\bar{x})$  和  $Z_2(\bar{x})$  各自所对应的  $p$  值，而如上定义的  $H_0$  并不足以获取  $p$  值，为此定义一个更强的原假设： $H_0: \bar{F}_{X_t}^{[a]} = \bar{F}_{X_t}^{[a]}, \forall t \in T$ 。其中， $\bar{F}_{X_t}^{[a]} = \bar{P}_{X_t}(x | Y_t \geq \text{VaR}_\alpha(Y_t)) = 1 - P(X_t \leq x | Y_t \geq \text{VaR}_\alpha(Y_t))$ ，表示预测得到的在  $Y_t \geq \text{VaR}_\alpha(Y_t)$  下  $X$  的条件分布。在  $H_0$  成立的条件下，有  $E_{H_0}(Z(X)) = 0$  以及  $E_{H_1}(Z(X)) > 0$ 。式(2)和式(3)中定义的统计量都是没有分布假定的，因此通过多次的模拟的方式获取  $Z_1(\bar{x})$  和  $Z_2(\bar{x})$  的  $p$  值。

具体来说，采取如下 3 步：a) 生成数据： $X_t^i \sim P_{X_t}, \forall t \in T, i=1, 2, \dots, M$ ；b) 计算统计量的值： $Z_j^i = Z_j(X^i), j=1, 2$ ；c) 根据式  $p = \sum_{i=1}^M Z_j^i > Z_j(\bar{x}) / M$  来获取  $p$  值。其中  $M$  是一个充分大的数，从而使得  $Z$  的经验分布非常接近真实分布。也即通过蒙特卡洛模拟的方式获得  $Z$  的近似分布，从而计算  $Z_1(\bar{x})$  和  $Z_2(\bar{x})$  的  $p$  值。

### 1.2 方法二

方法二是 Banulescu 等用来进行 MES 后验分析的方法，该方法通过巧妙地利用 CoVaR 和 MES 的关系，从而创造了一套关于 MES 的检验方法。

首先给出本文中 CoVaR 的定义，注意到  $X_t, Y_t$  均代表的是损失，本文将某个个体在  $t$  时刻的 CoVaR 表示为：

$$P(X_t \geq \text{CoVaR}_{\alpha, \beta}(X_t) | Y_t \geq \text{VaR}_\alpha(Y_t)) = \beta \quad (4)$$

式中  $\beta \in [0, 1]$ , 可看成是另一个置信水平。因此, MES 和 CoVaR 的关系为:

$$\text{MES}_\alpha(X_t) = \int_0^1 \text{CoVaR}_{\alpha, \beta}(X_t) d\beta \quad (5)$$

与 VaR 的后验分析类似, 可以通过一个碰撞序列对 CoVaR 进行后验分析。具体来说, 定义  $h_t(\alpha, \beta) = I\{X_t \geq \text{CoVaR}_{\alpha, \beta}^{P_t}(X_t) \cap Y_t \geq \text{VaR}_\alpha^{P_t}(Y_t)\}$ , 通过检验  $h_t(\alpha, \beta)$  是否是一个期望为  $\alpha\beta$  的伯努利变量, 或者等价地, 通过检验  $\{h_t(\alpha, \beta) - \alpha\beta\}$  是否为鞅差序列来说明对 CoVaR 的估计是否准确。基于式(5)可得

$$H_t(\alpha) = \int_0^1 h_t(\alpha, \beta) d\beta = u_{t|v} I_{\{v_t \geq \alpha\}} \quad (6)$$

式中:  $v_t = P_{t,2}(Y_t)$ ,  $u_{t|v} = P_{t,1|Y_t \geq \text{VaR}_\alpha^{P_t}(Y_t)}(X_t)$ ,  $P_{t,1}$  和  $P_{t,2}$  分别是在预测分布  $P_t$  下  $X_t$  和  $Y_t$  的边际分布。并且由于  $u_{t|v}$  和  $v_t$  相互独立, 故  $E(H_t(\alpha)) = \alpha/2$ , 从而检验 MES 估计是否准确就转化成了检验  $H_t(\alpha)$  的期望是否为  $\alpha/2$ , 也即可定义原假设为  $H_0: E(H_t(\alpha)) = \alpha/2, \forall t \in T$ 。

进一步计算可得到  $H_t(\alpha)$  方差为  $\alpha(1/3 - \alpha/4)$ , 因此可构造  $t$  统计量:

$$Z_t = \frac{\sqrt{T}(\bar{H}(\alpha) - \alpha/2)}{\sqrt{\alpha(1/3 - \alpha/4)}} \quad (7)$$

式中  $\bar{H}(\alpha) = \frac{1}{T} \sum_{t=1}^T H_t(\alpha)$ 。换言之, 如果在  $P_t$  下计算得到的 MES 是准确的, 则由式(7)定义的统计量应该服从自由度为  $T-1$  的  $t$  分布, 与方法一相统一。本文备择假设同样选取 MES 被低估的情况, 因为 MES 可看成是由 CoVaR 积分得到, 当 CoVaR 被低估时, 超过阈值的样本点会比预期多, 也即  $E(h_t(\alpha, \beta)) > \alpha\beta$ 。对应地, 当 MES 被低估时, 会有  $E(H_t(\alpha)) > \alpha/2$ , 即备择假设为  $H_1: E(H_t(\alpha)) \geq \alpha/2, \forall t \in T$ , 并且存在某个  $t \in T$ , 使不等号严格成立。

在实际操作中, 通过比较  $Z_t(\bar{x})$  与  $t(T-1)$  分布右侧的临界值, 就可对 MES 估计(预测)的准确性进行判断。

## 2 功效比较

由于样本的随机性, 当应用某个检验结果来做判断时, 有可能会犯弃真和取伪两类错误( $\alpha$  错误和  $\beta$  错误, 或者也叫第一类错误和第二类错误), 并且在样本量一定的条件下这两类错误不可能同时减小, 故通常会采取折中的方案, 即控制犯第一类错误的概率不大于某个给定的值。在统计学的假设检验中, 统计功效(Statistical Power)指的是由样本计算得到的统计量落入拒绝域的概率。当原假设不为真时, 这个概率越大, 意味着犯取伪错误的概率越小。这一节将通过模拟数据实验来对上述 2 种方法的统计量的功效进行比较。

采用二元的  $t$  分布生成模拟实验数据。首先给定 3 个原假设的分布, 每个原假设下分别设置 2 个备择假设, 第 1 个备择假设的自由度  $\nu$  与原假设保持一致,  $\rho$  比原假设大; 第 2 个备择假设在第一个备择假设的基础上减小了自由度  $\nu$ , 而  $\rho$  不变。不难看出, 第 2 个备择假设的分布与原假设分布的差别要比第 1 个备择假设与原假设的差别更大。此外,  $t$  分布中越小的自由度意味着越大的方差, 同时,  $\rho$  越大意味着变量间的相关关系越强, 而在这里每个原假设下的 2 个备择假设的  $\rho$  都大于原假设, 并且(或者)  $\nu$  要小于原假设, 这意味着每个备择假设的 MES 会高于原假设下的 MES。故该模拟实验探究的是当  $H_0$  不为真, 也即  $H_0$  下的 MES 确实低估了真实值时, 统计量取伪的概率大小。对每个备择假设, 都设置了 10% 和 5% 两个显著性水平, 每个模拟实验取样本量大小为 250 和 500 两种情况, MES 的置信水平分别取 90%、95% 和 97.5% 三个值。

表 1 汇总了当  $H_0$  不为真时, 方法一中的两个统计量在不同原假设和备择假设下的功效。从中可以看到一个有趣的结果: 对于每个  $H_0$ , 第一个  $H_1$  下的  $Z_1$  统计量的功效要优于  $Z_2$ , 第二个  $H_1$  下的情况则相反,  $Z_2$  统计量的功效要优于  $Z_1$ 。这说明当备择假设的分布与原假设的分布的差别相对较小时, 统计量  $Z_1$  的表现效果更好; 而当备择假设的分布与原假设的分布的差别相对较大时, 统计量  $Z_2$  的表现更佳。此外还注意到, 统计量  $Z_1$  和  $Z_2$  的功效均随 MES 置信水平的增大而减小, 这很有可能是因为这 2 个统计量在计算过程中需要对所有  $Y$  超过阈值的  $X$  进行某种程度的平均, 随着置信水平的增大,  $Y$  超过阈值所对应的  $X$  的个数也会减少, 取平均之后的可靠性也会降低, 从而影响最终统计量的效果。

表1 方法一统计量的功效  
Table1 Statistical power of method one

H <sub>0</sub>		H <sub>1</sub>		significance level/%	power/%											
ρ	ν	ρ	ν		T=250						T=500					
					MES <sub>90%</sub>		MES <sub>95%</sub>		MES <sub>97.5%</sub>		MES <sub>90%</sub>		MES <sub>95%</sub>		MES <sub>97.5%</sub>	
				Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>1</sub>	Z <sub>2</sub>	
0.4	6	0.6	6	10	58.3	47.8	44.1	36.8	31.2	27.7	83.2	70.7	67.8	53.7	50.0	39.7
				5	41.3	34.1	27.7	24.6	17.2	18.0	70.6	56.1	51.0	39.3	32.9	26.8
		0.6	3	10	80.3	86.9	68.8	85.9	53.7	83.8	95.1	97.6	88.7	96.9	76.3	96.0
				5	71.4	80.0	55.6	79.0	38.7	77.2	91.5	95.8	81.4	94.9	64.7	93.3
0.4	30	0.6	30	10	71.8	56.3	58.2	42.8	42.8	32.5	92.3	75.6	81.8	58.1	64.3	42.3
				5	56.9	41.2	39.8	30.5	26.8	21.0	83.7	63.6	68.0	44.9	47.1	30.4
		0.6	10	10	80.0	76.0	68.0	71.1	52.1	65.4	95.4	92.0	88.8	87.9	75.2	83.4
				5	68.2	63.9	51.7	58.8	34.9	53.4	90.1	86.3	79.0	80.8	60.9	74.7
0.2	6	0.4	6	10	57.7	54.0	44.1	41.1	32.2	31.7	79.2	75.0	63.9	58.6	47.4	43.9
				5	41.9	39.5	29.4	27.7	17.8	20.7	66.4	61.3	47.9	44.3	33.1	28.5
		0.4	3	10	71.3	79.8	59.6	77.2	46.4	74.9	88.2	93.6	78.7	91.2	65.4	89.0
				5	60.4	72.7	48.0	69.6	32.2	67.2	82.0	90.2	69.0	87.5	53.8	84.5

从表2中可以看到，整体而言，Z<sub>i</sub>的功效要低于方法一的统计量。具体来说，对于每个H<sub>0</sub>，在第一个H<sub>1</sub>下，Z<sub>i</sub>的功效要明显低于Z<sub>1</sub>，整体低于Z<sub>2</sub>；在第二个H<sub>1</sub>下，Z<sub>i</sub>的功效整体来看要低于Z<sub>2</sub>。说明当备择假设的分布与原假设的分布的差别相对来说比较小时，统计量Z<sub>i</sub>的表现远比不上Z<sub>1</sub>和Z<sub>2</sub>；当备择假设的分布与原假设的分布的差别相对来说比较大时，统计量Z<sub>i</sub>的表现整体要劣于Z<sub>2</sub>。

表2 方法二统计量的功效  
Table2 Statistical power of method two

H <sub>0</sub>		H <sub>1</sub>		significance level/%	power/%					
ρ	ν	ρ	ν		T=250			T=500		
					MES <sub>90%</sub>	MES <sub>95%</sub>	MES <sub>97.5%</sub>	MES <sub>90%</sub>	MES <sub>95%</sub>	MES <sub>97.5%</sub>
0.4	6	0.6	6	10	34.6	27.6	23.6	48.0	38.5	30.6
				5	23.5	19.6	15.5	36.1	27.3	21.4
		0.6	3	10	77.6	83.2	84.6	94.2	96.4	96.7
				5	68.3	75.0	77.5	90.1	93.7	94.3
0.4	30	0.6	30	10	39.0	32.6	27.5	52.1	41.6	33.8
				5	27.3	23.1	19.2	39.4	30.9	24.2
		0.6	10	10	57.5	59.0	58.6	78.1	78.1	77.1
				5	45.6	47.3	47.9	68.4	69.1	68.0
0.2	6	0.4	6	10	31.6	27.1	22.4	45.3	36.1	29.8
				5	22.0	18.4	15.2	33.3	25.5	19.9
		0.4	3	10	73.7	79.1	81.4	91.2	94.5	95.4
				5	63.2	70.7	73.6	85.5	91.4	92.2

但Z<sub>i</sub>在某些方面也有着略微的优势。可以注意到的一点是，随着置信水平的增加，在第2个H<sub>1</sub>下，Z<sub>i</sub>的表现略优于Z<sub>1</sub>和Z<sub>2</sub>。将MES置信水平为97.5%下每个原假设对应的第2个备择假设的数据取出来，共12组对比数据，其中有8组数据Z<sub>i</sub>的功效值是大于Z<sub>1</sub>和Z<sub>2</sub>，说明在较高的置信水平并且备择假设的分布与原假设的分布相差较大的情况下(虽然实际中可能性较少)，统计量Z<sub>i</sub>的表现稍佳。可以猜想，在更高的置信水平下以及同样是备择假设的分布与原假设的分布相差较大的情况下，Z<sub>i</sub>的表现有可能会相对更优。并且还可以注意到，与Z<sub>1</sub>和Z<sub>2</sub>不同，Z<sub>i</sub>的功效并不一定随着MES置信水平的增大而减小。而且，由于在原假设下Z<sub>i</sub>的分布是已知的，故统计量的p值可直接获得，而不需要像Z<sub>1</sub>和Z<sub>2</sub>一样需要通过大量的模拟计算。

总的来说，在原假设分布和备择假设分布相差相对较小时，统计量Z<sub>1</sub>的统计功效最大，表现最好，其次是Z<sub>2</sub>，而Z<sub>i</sub>的统计功效最低，效果最差；在原假设分布和备择假设分布相差相对较大时，整体看来，统计量Z<sub>2</sub>的表现最好，对于置信水平较高的MES，Z<sub>i</sub>的统计功效相对稍大些。但这种情况在实际中发生的概率相对比较小，因为估计的分布一般不会和真实的分布相差特别远。因此可以认为在实际中，统计量Z<sub>1</sub>的结果可能相对会更可靠。

### 3 实证分析

#### 3.1 数据描述

通过实际的数据分析来验证前文中后验分析方法的功效。其中的数据来自一个公开的射频频谱数据集，由伊利诺伊理工学院联合图尔库大学和弗吉尼亚理工大学共同收集，以用于对射频环境的时空变化进行建模，从

而实现对频谱利用率更精确的估计, 提高频谱利用率。观测点有 3 个在美国, 2 个在芬兰, 采集的时间区间是 2015 年到 2018 年间每年的 1 月 20 日至 1 月 27 日连续 8 天时间, 采集的频率范围为 30~6 000 Hz, 分为 30~130 Hz、130~800 Hz、650~1 200 Hz、1 200~3 000 Hz 和 3 000~6 000 Hz 五个频段。本小节从其中 650~1 200 Hz 的频段中随机选取 20 个频点, 暂且将每个频点对应的电磁信号表示为  $X_1, X_2, \dots, X_{20}$ , 同时假定在同一个时间点有一个信号是由这 20 个频点的电磁信号叠加得到, 暂且记为  $Y$ 。本文选取了其在 2015 年 1 月 22 日这一天, 在经度为 22.27, 纬度为 66.44 处的数据, 由于数据在采集过程中是每隔 3 s 左右记录 1 次, 因此在这 24 h 内每个频率共有 28 272 个采集样本。取训练和测试的样本数据比例大致为 3:1, 也即前面的 21 000 个样本数据作为第一部分用来估计模型参数, 剩下的作为第二部分用来做后验分析以检验估计结果的有效性。这里先展示每个频点的电磁信号的统计特征, 限于篇幅, 仅展示其中前 4 个频点上电磁信号的情况, 如表 3 所示。

从表 3 中可以看到, 每个频率的信号都有其各自的分布特点, 均值和波动情况也不一样, 但都有轻微的左偏以及比标准正态稍高一些的峰度。最后的两列是关于 ARCH 效应和自相关性的检验, 表格中展示的数值对应的  $p$  值, 可以看到, 在 5% 的显著性水平下, 不拒绝信号 2 和信号 3 不存在 ARCH 效应和自相关性的假定, 在 1% 的显著性水平下, 只有信号 1 会拒绝不存在 ARCH 效应和自相关性的假定。

表 3 四个频率信号的描述性统计  
Table3 Description statistic of four signals

signal	mean	SD	skewness	kurtosis	ARCH	Ljung_Box
signal 1	-93.105 9	6.811 7	-1.260 4	2.204 4	0	0
signal 2	-112.657 0	3.767 3	-0.788 1	1.921 1	0.103 6	0.083 8
signal 3	-114.452 0	5.614 1	-1.086 1	2.069 1	0.139 7	0.156 4
signal 4	-107.250 0	3.968 6	-0.538 9	1.318 3	0.019 4	0.020 1

### 3.2 分布拟合

在对  $(X, Y)$  分布的估计上, 为了增加一定的灵活性, 本文并不直接对其二元联合分布进行建模, 而是采用先估计边际分布, 再估计相依结构(Join Structure)的二步拟合的方式。从表 3 的描述性统计中可以看到, 每个信号的分布都存在左偏的形态, 故采取有偏的正态分布来进行拟合, 其密度函数为  $f=2f(x)\Phi(\lambda x)$ , 其中,  $f, \Phi$  分别表示标准正态分布的密度函数和分布函数,  $\lambda$  为偏态系数。在相依结构上, 采用正态 Copula 进行拟合。采用极大似然估计(Maximum Likelihood Estimation, MLE)来获取最优参数。

表 4 汇总了 4 个频点的电磁信号的边际分布和与叠加信号  $Y$  的相依结构的极大似然估计结果, 第 1 列  $\mu$  为位置参数, 衡量了每个频点电磁信号功率的平均情况, 第 2 列  $\sigma$  为尺度参数, 衡量了每个频点电磁信号功率的波动情况, 第 3 列  $\lambda$  衡量了偏移程度的大小, 最后一列为正态 Copula 的系数。从中可以看到, 信号 1 的平均功率最大, 波动的程度也最大, 信号 1 和信号 3 的非对称性相对来说最高, 也即左偏最明显, 并与  $Y$  的相关程度也相对高些。

表 4 分布拟合结果  
Table4 Distribution fitting

signal number	$\mu$	$\sigma$	$\lambda$	$\rho$
signal 1	-85.145 6	10.978 4	-5.652 7	0.427 9
signal 2	-108.462 1	5.574 6	-2.528 6	0.207 5
signal 3	-107.639 1	8.769 8	-4.157 7	0.285 6
signal 4	-103.335 5	5.501 8	-1.772 0	0.206 3

图 1 为 4 个频点上电磁信号的直方图, 以及表 4 中的估计参数值所对应的密度函数曲线, 从中可以看出, 对边际分布的估计结果较好地拟合了真实数据的分布情况。

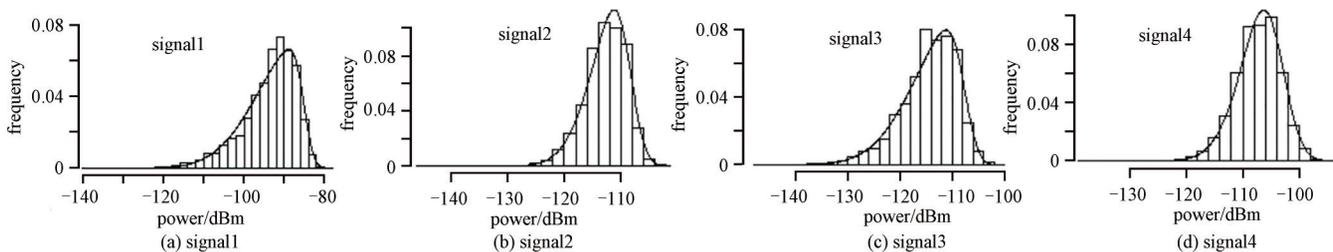


Fig.1 Boxplots of power for each signal

图 1 每个信号的功率箱线图

3.3 后验分析结果

用第二部分的数据对MES 的估计结果进行后验分析，即通过计算对应的统计量来判断模型对MES 的估计是否正确。由于每个统计量在计算过程中均需涉及到Y 与阈值的比较，故先对VaR 估计的准确性进行衡量，结果如表5 所示。从中可以看到，虽然Y 超过VaR 的实际个数与期望值相比略有出入，但从括号中的p 值可以得到，在10% 的显著性水平下，都可认为模型对VaR 的估计是正确的。

表5 实际数据超过VaR 的次数

Table5 Numbers exceeding VaR

VaR <sub>90%</sub> (Expectation:773)	VaR <sub>95%</sub> (Expectation:386)	VaR <sub>97.5%</sub> (Expectation:193)
789(0.539 0)	362(0.202 8)	173(0.141 1)

表6 为MES 后验分析的结果，括号中的数值为统计量对应的p 值。从中可以看到，在10% 的显著性水平下，除了信号3 的MES<sub>90%</sub> 和MES<sub>95%</sub> 在统计量Z<sub>1</sub> 下没有接受原假设和信号1 的MES<sub>90%</sub> 在统计量Z<sub>i</sub> 下没有接受原假设外，其他的都不拒绝原假设，在5% 的显著性水平下，所有的结果均不拒绝原假设，也即可以选择接受MES 没有被低估这一结论。此外，对表6 中的数据进行简单的统计发现，在12 组数据中，Z<sub>2</sub> 的p 值大于Z<sub>1</sub> 的有10 组，Z<sub>i</sub> 的p 值大于Z<sub>1</sub> 和Z<sub>2</sub> 的分别有10 组和8 组，表明对于同样的估计结果，Z<sub>i</sub> 要比Z<sub>1</sub> 和Z<sub>2</sub> 对原假设的接受程度更高，而Z<sub>2</sub> 又比Z<sub>1</sub> 对原假设的接受程度高。可以得出，在原假设和备择假设的分布差别相对较小时，Z<sub>2</sub> 和Z<sub>i</sub> 在H<sub>0</sub> 不为真时的统计功效比Z<sub>1</sub> 小，即取伪概率比Z<sub>1</sub> 大。由于在实际中原假设和备择假设的分布往往不会非常大，即估计的分布一般不会和真实的分布相差特远，故表6 中呈现的结果可认为与第2 小节中得到的结论比较一致。

表6 三个统计量的值

Table6 Values of the three statistics

statistics	signal	MES <sub>90%</sub>	MES <sub>95%</sub>	MES <sub>97.5%</sub>
Z <sub>1</sub>	signal 1	-0.006 072(0.990 1)	-0.002 618(0.781 6)	-0.000 472(0.579 5)
	signal 2	0.000 301(0.392 4)	-0.000 274(0.569 8)	-0.000 313(0.566 5)
	signal 3	0.003 176(0.050 7)	0.004 042(0.062 7)	0.002 439(0.285 9)
	signal 4	0.000 718(0.306 0)	-0.001 941(0.822 2)	-0.001 416(0.695 0)
Z <sub>2</sub>	signal 1	0.014 763(0.361 8)	-0.065 599(0.843 1)	-0.104 978(0.878 6)
	signal 2	0.021 270(0.311 4)	-0.063 403(0.839 0)	-0.104 835(0.879 2)
	signal 3	0.024 205(0.288 2)	-0.059 360(0.823 6)	-0.102 370(0.872 2)
	signal 4	0.021 695(0.308 8)	-0.064 965(0.845 3)	-0.105 823(0.882 1)
Z <sub>i</sub>	signal 1	1.328 517(0.092 0)	-0.945 105(0.827 7)	-1.356 476(0.912 5)
	signal 2	0.007 291(0.497 1)	-1.251 312(0.894 6)	-1.494 165(0.932 4)
	signal 3	-1.246 605(0.893 7)	-2.435 229(0.992 5)	-2.052 488(0.979 9)
	signal 4	-0.134 229(0.553 4)	-0.757 781(0.775 7)	-1.113 731(0.867 3)

4 结论

后验分析作为检验风险预测模型准确程度的重要步骤，直接关系到最终结论的可靠性。本文将Acerbi 等关于ES 的后验分析方法进行二维变量下的延伸，提出了2 个新的对于MES 后验分析的统计量。这2 个新的统计量基于MES 的定义式得到，同样也是无分布假定的、非参的，且比较直观，从概念上也易于理解。本文将其与Banulescu 等采用的统计量进行比较，发现在原假设分布和备择假设分布相差相对较小的情况下，2 个统计量的统计功效均大于Banulescu 等采用的统计量。由于这种情形一般情况下与实际更相符，因此可以认为在实际中，统计量Z<sub>i</sub> 的结果可能相对会更可靠。实证分析的结果也表明，对于同样的预测结果，文中新提出的统计量对原假设的接受相对更为谨慎。

本文的主要贡献在于：一方面提出了2 个新的对于MES 后验分析的统计量，丰富了对系统性风险度量进行后验分析的方法；另一方面也丰富了对系统性风险度量后验分析的统计功效问题的讨论，对于其他的后验分析方法，也可以类似的方式进行统计功效的对比讨论，衡量其在原假设不成立下的取伪概率，探讨在不同情况下各统计量的优劣。本文将该后验方法用于电磁频谱数据分析中，为单个频点的电磁信号对整体异常的贡献程度的衡量提供了思路，对于大数据模型算法的后验分析具有一定的理论借鉴意义。

参考文献：

[ 1 ] ACHARYA V V,PEDERSEN L H,PHILIPPON T,et al. Measuring systemic risk[J]. CEPR Discussion Papers, 2017,29(1002):85-119.  
 [ 2 ] 赵淑清,郑薇. 随机信号分析[M]. 哈尔滨:哈尔滨工业大学出版社, 1999. (ZHAO Shuqing,ZHENG Wei. Stochastic signal analysis[M]. Harbin,China:Harbin Institute of Technology Press, 1999.)

- [3] 王永德,王军. 随机信号分析基础[M]. 北京:电子工业出版社, 2013. (WANG Yongde,WANG Jun. Fundamentals of stochastic signal analysis[M]. Beijing:China Publishing House, 2013.)
- [4] 张海勇. 基于局域波法的非平稳随机信号分析中若干问题的研究[D]. 大连理工大学, 2001. (ZHANG Haiyong. Research on some problems in non-stationary random signal analysis based on local wave method[D]. Dalian,China:Dalian University of Technology, 2001.)
- [5] 张海勇,李勘. 非平稳随机信号参数模型分析方法[J]. 系统工程与电子技术, 2003,25(3):386-390. (ZHANG Haiyong,LI Kan. Parametric model analysis method of non-stationary random signal[J]. Systems Engineering and Electronics, 2003,25(3):386-390.)
- [6] 杨娟,卞保民,闫振纲,等. 典型随机信号特征参数统计分布的分形特性[J]. 物理学报, 2011,60(10):86-92. (YANG Juan,BIAN Baomin,YAN Zhengang,et al. Fractal characteristics of statistical distribution of typical random signal characteristic parameters[J]. Acta Physica Sinica, 2011,60(10):86-92.)
- [7] SONG L,CHEN Y,TRAPPE W,et al. ALDO:an anomaly detection framework for dynamic spectrum access networks[C]// IEEE INFOCOM 2009. Rio de Janeiro,Brazil:IEEE, 2009:675-683.
- [8] YIN S,LI S,YIN J. Temporal-spectral data mining in anomaly detection for spectrum monitoring[C]// The 5th International Conference on Wireless Communications, Networking and Mobile Computing. Beijing:IEEE, 2009:1-5.
- [9] LIU S,GREENSTEIN L J,TRAPPE W,et al. Detecting anomalous spectrum usage in dynamic spectrum access networks[J]. Ad Hoc Networks, 2012,10(5):831-844.
- [10] ZHANG L,DING G,WU Q,et al. Spectrum sensing under spectrum misuse behaviors:a multi-hypothesis test perspective[J]. IEEE Transactions on Information Forensics and Security, 2018,13(4):993-1007.
- [11] 夏伟,王琳娜,沈小青,等. 基于单类支持向量机模型的频谱异常检测方法[J]. 微处理器, 2012,33(4):73-75. (XIA Wei,WANG Linna,SHEN Xiaoqing,et al. Spectral anomaly detection method based on single-class support vector machine model[J]. Microprocessor, 2012,33(4):73-75.)
- [12] 闫戈. 基于大数据的频谱分析算法的研究[D]. 北京:北京邮电大学, 2017. (YAN Ge. Research on spectrum analysis algorithm based on big data[D]. Beijing:Beijing University of Posts and Telecommunications, 2017.)
- [13] LI Zhijing,XIAO Zhujun,WANG Bolun,et al. Scaling deep learning models for spectrum anomaly detection[C]// Proceedings of the Twentieth ACM International Symposium on Mobile Ad Hoc Networking and Computing. New York,NY,USA:ACM, 2019:291-300.
- [14] 李忠,安建琴,宋奕瑶,等. 一种滑动阈值的自适应随机信号异常提取方法:CN108898117A[P]. 2018. (LI Zhong,AN Jianqin,SONG Yiyao,et al. Abnormal extraction method of adaptive random signal with sliding threshold:CN108898117A[P]. 2018.)
- [15] 张小飞,李恒帅,石志鑫,等. 一种电磁频谱异常告警分析算法及实现[J]. 中国无线电, 2020(6):34-36,41. (ZHANG Xiaofei,LI Hengshuai,SHI Zhixin,et al. An electromagnetic spectrum anomaly alarm analysis algorithm and its implementation[J]. China Radio, 2020(6):34-36,41.)
- [16] ACERBI C,SZEKELY B. Backtesting expected shortfall[R]. New York,NY,USA:MSCI Inc., 2014:1-33.
- [17] DU Z,ESCANCIANO J C. Backtesting expected shortfall:accounting for tail risk[J]. Management Science, 2017,63(4):940-958.
- [18] BANULESCU D,HURLIN C,LEYMARIE J,et al. Backtesting marginal expected shortfall and related systemic risk measures[J]. Management Science, 2021,67(9):5730-5754.
- [19] DUMITRESCU E I,HURLIN C,PHAM V. Backtesting value-at-risk:from dynamic quantile to dynamic binary tests[J]. Finance, 2012,33(1):79-112.
- [20] HURLIN C,TOKPAVI S. Une evaluation des procédures de backtesting:tout va pour le mieux dans le meilleur des mondes[J]. Finance, 2008,29(1):53-80.

#### 作者简介:

张继丹(1994-),女,硕士,主要研究方向为统计推断与风险管理.email:2375707879@qq.com.

肖东(1999-),男,在读硕士研究生,主要研究方向为电磁大数据以及压缩感知.

侯燕曦(1978-),男,博士,副教授,博士生导师,主要研究方向为极值理论、统计推断在风险管理中的应用.