金融时间序列模型的贝叶斯单位根检验

李 勇／著

Bayesian unit rooting testing for financial times series

经济科学出版社
Economic Science Press

金融时间序列模型的贝叶斯单位根检验

李 勇 著

经济科学出版社

图书在版编目（CIP）数据

金融时间序列模型的贝叶斯单位根检验/李勇著．
—北京：经济科学出版社，2014.5
ISBN 978－7－5141－4612－7

Ⅰ.①金…　Ⅱ.①李…　Ⅲ.①金融－时间序列
分析　Ⅳ.①F830

中国版本图书馆 CIP 数据核字（2014）第 085083 号

责任编辑：柳　敏　周秀霞
责任校对：杨　海
版式设计：齐　杰
责任印制：李　鹏

金融时间序列模型的贝叶斯单位根检验
李　勇　著
经济科学出版社出版、发行　新华书店经销
社址：北京市海淀区阜成路甲 28 号　邮编：100142
总编部电话：010－88191217　发行部电话：010－88191522
网址：www.esp.com.cn
电子邮件：esp@esp.com.cn
天猫网店：经济科学出版社旗舰店
网址：http://jjkxcbs.tmall.com
北京京鲁创业科贸有限公司印装
710×1000　16 开　8 印张　120000 字
2014 年 4 月第 1 版　2014 年 4 月第 1 次印刷
ISBN 978－7－5141－4612－7　定价：26.00 元
（图书出现印装问题，本社负责调换。电话：010－88191502）

导　言

在计量经济学中，单位根检验是一个非常重要的研究问题，它是观察时间序列是否平稳的一个重要方法。从计量经济学建模理论上讲，单位根检验的理论目的在于：确定一个金融时间序列是否差分平稳。其中的平稳是指序列在不同时刻会保持稳定，更正规的说法是，平稳意味着一个序列的均值和协方差与时间无关。在金融研究的各种方法中，如回归、格兰杰因果检验，都要求相应的序列是平稳的。所以，我们必须进行单位根检验，以确定我们是否违背了这一假定。

从应用的角度讲，金融时间序列研究中的单位根检验有着非常重要的应用价值。比如说，它是观测收益和波动是否具有持续性的一个有效的计量方法。对于金融资产来说，收益或者波动持续性存在与否对金融资产定价（如期权定价）的影响是不同的。持续性现象表明当前的扰动（如当前的突发事件）对今后将有一个持续累积的影响。从宏观经济政策层面上看，政府对于股市调控的一系列方针、政策是否有效果，也可以通过观察资产收益或者波动的持续性，从而得到反映。举例来说，美联储什么时候宣布退出量化宽松的货币政策可能对金融市场有一个非常长期的影响，这个影响就可

以从计量的角度，通过构造单位根来检验。从微观层面讲，从金融风险角度看，持续性的存在增大了未来收益的风险，反之，如果不存在持续性，对于长期投资者来说，当前的扰动就可忽略不计。因此，显然地，对于厌恶风险的投资者，收益或者波动的持续性将是一个必须考虑的因素。

基于以上原因，开展基于不同模型的单位根检验方法研究，具有非常重要的理论和实际意义。因此，如何探测收益或者波动具有持续性也引起了研究者的广泛注意和兴趣，许多研究者都为此付出了努力。

对于单位根检验问题来说，传统的经典方法主要是经典单位根检验统计量（如 DF 和 ADF 检验统计量）及一些延伸研究。然而，这些检验统计量往往在实践中遭受一些困难，主要体现在以下几个方面：（1）对于一些比较复杂的模型，如随机波动模型，似然函数没有解析解，经典的估计方法 OLS 估计、ML 估计等估计方法相当困难。（2）基于频率统计的传统单位根检验统计量是建立在大样本理论基础上的，推导统计量的渐近分布异常复杂，即使是针对简单的 AR(p) 模型也是如此。（3）基于频率统计学派的传统计量经济学忽视研究者对所研究问题的“先验信息”，而主要关注“样本信息”。尤其在小样本的情况下，会使得统计推断存在严重的信息不充分。（4）传统单位根检验的一个较大的理论困难是：如果单位根检验不能拒绝单位根，并不能说明金融时间序列存在单位根。

正是由于这些实际限制，近些年来，贝叶斯单位根检验方法受到了研究者的广泛关注。在贝叶斯的方法框架下，通过后验抽样方法，得到后验分布的随机样本。而且贝叶斯统

计推断建立在后验分布的基础上，并不需要用到复杂的大样本理论。近十年来，随着蒙特卡罗模拟（MCMC）技术的进步，现在从后验分布模拟出随机样本已经非常简单，因而参数估计、标准误等统计量都是容易获得的。对于单位根检验问题，贝叶斯检验统计量能够告诉研究者："相对于备择假设，零假设成立的概率会更大吗?"另外，当先验信息有效时，通过贝叶斯建模，先验信息可以通过先验分布吸收到统计推断中。

在贝叶斯文献中，贝叶斯因子是一个非常重要的统计量，主要用于模型的检验和选择，在本书中用以单位根检验。然而，由于贝叶斯因子是边际似然的比值，对于随机波动模型来说，其边际似然涉及到一个高维积分（超过样本个数），因此计算贝叶斯因子是一个有相当挑战性的任务。本书的最主要理论贡献是：针对单位根检验问题，我们推导了一个贝叶斯因子的特殊形式，该形式可以看作是贝叶斯后验输出结果的副产品。因为计算上非常简单，我们可以用以单位根检验。

本书第一次系统地研究了基于不同金融时间序列模型的单位根检验问题。本书的主要思路分为三个方面：（1）金融时间序列的单位根检验问题。（2）基于随机波动模型的波动单位根检验问题。（3）随机误差为随机波动模型的单位根检验问题。我们考虑了多种常用模型，比如说，具有未知自由度的厚尾金融时间序列模型、随机波动模型、具有未知自由度的厚尾随机波动模型、基于跳跃的随机波动模型的贝叶斯单位根检验问题，等等。通过随机模拟方法，本书调查了检验统计量的检验功效，模拟结果表明，检验统计量能

够取得很好的有限样本行为。最后，通过实证分析论证了所发展的方法，得到了一些令人启发的结论。

本书主要由我几年以来和合作者（包括我的研究生）共同发表的论文组成。本书的第 2 章内容来源于我和孙瑞博、王贵银合作发表于《数理统计与管理》（2012 年第 31 卷第 1 期）的学术论文。第 3 章内容来源于我和余俊合作的工作文章——*A New Bayesian Unit Root Test in Stochastic Volatility Models*。第 4 章内容来源于我和王贵银发表于《中大管理研究》（2011 年第 6 卷第 4 期）的学术论文。第 5 章内容来源于我和刘晓斌待发表于 *Quantitative finance* 的学术论文。第 6 章内容来源于我和张瑾玉、陈珠明共同发表于 *Computational economics*（2013，41：89 – 100）的学术论文。第 7 章来源于我和庄太量、张杰共同发表于 *Economic modeling*（2012，41：2035 – 2038）的学术论文。在本书中使用这些研究成果，均征得了他们的同意，在这里对合作者们表示衷心的感谢。另外，本书的写作也得到了我的研究生李刚、刘晓斌、潘祺、吴文博、钱智俊等的大力帮助和支持。他们在翻译、校对和排版过程中都做出了很多贡献，在这里也一并表示感谢。

目　　录

第 1 章

导　论

1.1 频率单位根检验

在计量经济学中，单位根检验是一个非常重要的研究问题，它是观察时间序列是否稳定的一个重要方法，许多研究者为此付出了很多的努力。假设给定时间序列 $\{y_t\}_{t=0}^{T}$，考虑下列 AR(1) 模型：

$$y_t = \rho y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N[0, \sigma^2] \tag{1-1}$$

单位根检验问题可归为如下假设检验问题：

$$\begin{aligned} H_0 &: \rho = 1 \\ H_1 &: \rho < 1 \end{aligned} \tag{1-2}$$

对于该假设检验问题，迪基和富勒（Dickey & Fuller，1979）给出了著名的检验统计量：

$$DF_t = \frac{\hat{\rho} - 1}{s_{\hat{\rho}}} \tag{1-3}$$

这里，$\hat{\rho}\left(\hat{\rho} = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}\right)$为最小二乘估计量，$s_{\hat{\rho}}$是相应的标准误。

该统计量的渐近分布可以推导为

$$\frac{\int_0^1 W\mathrm{d}W}{\left(\int_0^1 W^2\right)^{\frac{1}{2}}}$$

其中，W 是标准布朗运动，该检验统计量的临界值可以通过模拟的方法得到。对于单位根检验研究，国内研究者李子奈、王少平、张晓峒、白仲林、靳庭良等在经典频率单位根检验方法上做出了一定的贡献。

1.2 贝叶斯单位根检验

与经典频率统计学派不同，贝叶斯学派将概率理解为主观概率，概率不再是所谓的频率次数决定的，而是由主观决定的可能性大小，当这些主观概率满足一定的定理时，可以证明其满足概率的标准定义。贝叶斯学派因而由此出发，将事物的不确定性用概率进行概括，换句话说，所要研究的事物具有不确定性，而这不确定性是我们的重点所在，而且这些不确定性可以通过概率模型进行数量化。这为后续的统计推断建立了理论基础。

贝叶斯统计分析遵循同样的思路，明确所要研究的对象，针对对象的特点提出概率模型，从主观上将所要研究的事物的不确定性都概括在模型中。该模型必须与所要研究的对象的客观特征相一致。结合模型和实际样本数据，推导后验分布，后验分布必须具有可以理解性，并且具有实际意义。在后验分布的基础上，通过随机模拟抽样，得出后验的随机样本，在这些样本的基础上，进行统计推断，诸如假设检验、模型选择等。格尔曼等（Gelman et al.，2013）和罗伯特（Robert，2001）对贝叶斯学派的这一思想进行了详细解释，有兴趣的读者可以参考这两本经典教科书。

近些年来，随着计算机技术和统计技术的发展，特别是蒙特卡洛马尔科夫（MCMC）的方法提出来后，带来了贝叶斯统计分析的繁荣研究。MCMC 主要用两种方法：Gibbs 抽样和 Metropolis – Hastings 抽样，前者是后者的特例，算法的本质都是利用可逆马尔可夫链的性质通过大量迭代得到后验分布的随机抽样样本，从而可以运用这些样本进行统计推断，具体可以看卡塞拉和乔治（Casella & George，1992）以及契博和格林伯格（Chib & Greenberg，1995）的研究。以下对这两种方法进行简要的介绍。

1.2.1 Gibbs 抽样

假设参数向量为 $\theta = \{\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(k)}\}$，联合后验分布为 $\pi(\theta) = \pi(\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(k)})$，为了得到联合后验分布的样本，先对参数向量进行初始化，$\theta_0 = \{\theta_0^{(1)}, \theta_0^{(2)}, \cdots, \theta_0^{(k)}\}$，并逐次对各个参数的后验分布进行抽样如下：

初始化参数向量 $\theta_0 = \{\theta_0^{(1)}, \theta_0^{(2)}, \cdots, \theta_0^{(k)}\}$

从 $j = 1, \cdots, n$，运行以下步骤：

对 $\theta_j^{(1)}$ 进行抽样，$\theta_j^{(1)} \sim \pi(\theta_j^{(1)} \mid \theta_{j-1}^{(2)}, \cdots, \theta_{j-1}^{(k)})$

对 $\theta_j^{(2)}$ 进行抽样，$\theta_j^{(2)} \sim \pi(\theta_j^{(2)} \mid \theta_j^{(1)}, \theta_{j-1}^{(3)}, \cdots, \theta_{j-1}^{(k)})$

……

对 $\theta_j^{(k)}$ 进行抽样，$\theta_j^{(k)} \sim \pi(\theta_j^{(k)} \mid \theta_j^{(1)}, \cdots, \theta_j^{(k-1)})$

通过上述步骤大量迭代，可以使得所抽取的值逐步收敛于目标后验联合分布，从而可以得到所需要的样本。关于该算法的具体介绍可以参考杰曼 · S. 和杰曼 · D.（Geman，S. & Geman，D.，1984）。

1.2.2 Metropolis – Hastings 算法

上述 Gibbs 抽样适用于对常见的分布进行抽样，但在实际应用

中，经常遇到的目标后验分布不是标准的后验分布，这给抽样带来了很大的困难，对此，学者们提出了另外一种更广泛的方法——Metropolis - Hastings 算法。Metropolis 算法的核心是引入较为容易抽样的函数 $q(\theta \mid g)$，在已拥有抽样 θ 的情况下对该函数进行抽样 $\tilde{\theta} \sim q(\tilde{\theta} \mid \theta)$，并建立接受/拒绝概率：

$$\alpha(\tilde{\theta}, \theta) = \min\left\{1, \frac{\pi(\tilde{\theta})q(\tilde{\theta} \mid \theta)}{\pi(\theta)q(\theta \mid \tilde{\theta})}\right\}$$

如果被拒绝了，那么 θ 将保持原来的值，如果接受了，那么 θ 将取新值 $\tilde{\theta}$。具体而言如下：

初始化参数向量 $\theta_0 = \{\theta_0^{(1)}, \theta_0^{(2)}, \cdots, \theta_0^{(k)}\}$

从 $j=1, \cdots, n$，运行以下步骤：

从抽样函数抽取样本 $\tilde{\theta}_j \sim q(\tilde{\theta}_j \mid \theta_{j-1})$

计算 $\alpha(\tilde{\theta}_j, \theta_{j-1})$，并抽取均匀分布 $U \sim U(0, 1)$

如果 $U < \alpha(\tilde{\theta}_j, \theta_{j-1})$，接受新值，$\theta_j = \tilde{\theta}_j$，如果 $U > \alpha(\tilde{\theta}_j, \theta_{j-1})$，那将保持原值，$\theta_j = \theta_{j-1}$。

Gibbs 抽样方法是 Metropolis - Hastings 算法的特例，但二者本质上是一致的，都是利用构造可逆马尔科夫链，通过大量重复的迭代抽样，使得抽取的样本渐渐趋向于实际后验分布。当进行了足够多的迭代次数后，去掉前面的样本使初始值的影响可以忽略，剩下的抽样值可以当作是从实际后验分布中抽取出来的，可用于统计推断。关于 Metropolis - Hastings 算法的具体介绍，可以参考黑斯廷斯（Hastings, 1970）、契博和格林伯格（1995）。

在贝叶斯方法框架下，通过 MCMC 抽样，得到后验分布的随机样本，统计推断是建立在后验分布的基础上，并不需要用到复杂的大样本理论。近十年来，随着蒙特卡罗模拟技术的进步，现在从后验分布模拟出随机样本已经非常简单，因而参数估计、标准误等统计量均容易获得。对于单位根检验问题，贝叶斯检验统计量能够告诉研究者“相对于备择假设，零假设成立的概率会更大吗?”另外，当先验信

息有效时，通过贝叶斯建模，先验信息可以通过先验分布吸收到统计推断中。

在贝叶斯文献中，贝叶斯因子是一个非常重要的统计量，用以模型的检验和选择，可参看卡斯和拉夫特里（Kass & Raftery，1995）、格韦克（Geweke，2007）。具体细节表达如下：记模型 M_0 为单位根零假设成立时的模型，称之为零模型，模型 M_1 为单位根备择假设成立时的模型，称之为备择模型。在贝叶斯分析中，检验单位根假设就转化为比较零模型 M_0 和备择模型 M_1 哪一个更好，这可归纳为一个非嵌套（nonnested）模型选择问题。贝叶斯因子定义可简单概括如下：对于 k =0，1，$p(y \mid M_k)$ 为模型 M_k 的密度函数，贝叶斯因子可表达为：

$$B_{01} = \frac{p(y \mid M_0)}{p(y \mid M_1)}$$

这里 $p(y \mid M_k) = \int p(y, \theta \mid M_k) d\theta$ 被称为边际似然。

1.3　基于随机波动模型的贝叶斯单位根检验

近三十年来，随着金融产品的多样化，尤其是金融衍生产品的不断涌现，估计和分析波动率无论是在业界或是学界均受到广泛的重视。在业界，波动率被作为一个重要考虑因素用于金融产品定价、金融风险规避、投资决策等领域；在学界，分析波动率现象、构建波动率模型、预测波动率等均是金融计量研究领域的热点问题。

对金融波动率进行建模的模型可分为两大类：一类是自回归条件异方差模型（ARCH），是由恩格尔（Engel，1982）在研究英国通货膨胀指数时提出的，后经波勒斯列夫（Bollerslev，1986）发展，提出了 GARCH 模型。另一类是由泰勒（Taylor，1986）提出的随机波动模型（stochastic volatility models）。随机波动模型在波动模型中引入

了波动误差，相比较于 ARCH 类模型，更加贴近金融理论和符合金融实践，因此在金融计量经济学研究中被当作是 ARCH 类模型一个有效的替代，引起了国内外学术界的广泛兴趣和重视。在理论方法与应用研究方面，张世英（2004）的专著介绍了各种波动性模型的原理、方法和实际应用，谢泼德（Shephard，2005）详细介绍了随机波动模型的理论及其发展。

对于随机波动模型，之所以也要检验单位根检验，一是单位根被拒绝是模型稳定的一个重要条件，另外也是观察波动持续性的一个重要方法。由于似然函数没有解析解，经典单位根检验统计量应用是非常困难的。对于最简单的随机波动模型，如果回归模型存在协变量，经典单位根检验统计量如 DF 检验统计量是不能应用的。对于简单随机波动模型的一些推广，如具有厚尾，杠杆效应的随机波动模型等，经典单位根检验统计量更是无能为力。

对于金融资产来说，持续性存在与否对金融资产诸如期权定价的影响是不同的。波动持续性现象表明当前的扰动对今后的波动性有一个积聚的影响。从微观层面，金融风险的角度看，持续性的存在增大了未来收益的风险。反之，如果不存在持续性，那么对长期投资者来说，当前的扰动就可忽略不计。显然，对厌恶风险的投资者来说，波动持续性是一个必须考虑的因素。另外，从宏观经济政策层面上讲，政府对于股市调控的一系列方针、政策是否有效果，也可以通过观察资产波动持续性来反映。因此，如何探测波动具有持续性引起了研究者的广泛注意和兴趣。

现代资产定价理论已经表明，波动持续性可以通过检验波动模型是否有单位根来反映。因而对于随机波动模型来说，检查波动是否具有持续性也就是在随机波动模型中检验单位根是否存在。从计量经济学建模理论上讲，单位根检验的理论目的在于，确定一个金融时间序列是否差分平稳。平稳是指序列在不同时刻会保持稳定。更正规的说法是平稳意味着一个序列的均值和协方差与时间无关。在金融研究的

各种方法中，如回归、Granger因果检验中都要求相应的序列是平稳的，所以我们必须进行单位根检验，以确定我们是否违背了这一假定。综上所述，无论是从宏观经济政策效果检验，微观层面上的投资应用，还是计量经济学的理论建模，对于随机波动模型，检验其单位根是否存在都具有十分重要的价值和意义。

对于随机波动模型，由于潜在波动的引入，使得模型的似然函数涉及到一个高维积分，并且这个积分通常超过了样本的个数，从而导致了建立在似然函数基础上的参数估计和假设检验均非常困难。对于参数估计，研究者们提出了很多估计方法来克服这些困难，包括广义矩估计（GMM）方法、伪极大似然估计（QML）、模拟极大似然估计（SML）、贝叶斯估计（Bayesian estimation）。安德森等（Andersen et al.，1999）通过蒙特卡罗学习比较了这些估计方法，指出了贝叶斯估计方法是最有效的一种。因此，贝叶斯估计方法引起了研究者的广泛注意和兴趣，如雅基耶等（Jacquier et al.，2004）、余（Yu，2005）、李等（Li et al.，2008）等。国内的研究者如朱慧明等（2008）、王春峰等（2003）、孟利锋等（2004）等在随机波动模型的基础上，运用贝叶斯估计方法，对中国股市的波动率变化情况进行了研究和探讨。

对于随机波动模型单位根检验问题来说，经典单位根检验统计量如ADF检验统计量应用非常困难，甚至是不能应用，见索奥和李（So & Li，1999）。这主要体现在以下几个方面：（1）由于随机波动模型似然函数没有解析解，经典的估计方法OLS估计、ML估计等估计方法相当困难。（2）基于频率统计的传统单位根检验统计量是建立在大样本理论基础上的，推导统计量的渐近分布异常复杂，即使是简单的AR(p)模型，也是如此。（3）基于频率统计学派的传统计量经济学忽视研究者对所研究问题的“先验信息”，主要关注“样本信息”。尤其在小样本情况下，使得统计推断存在严重的信息不充分。（4）传统单位根检验的一个较大理论困难是：如果单位根检验不能