

第1章 绪论

1.1 高频金融数据

近年来,受经济全球化、金融自由化及金融创新加速等因素的影响,金融市场规模迅速扩大、效率明显提高的同时,资产价格的波动性和风险也大为加剧.从亚洲金融危机到美国次贷危机引发的全球金融危机,金融市场剧烈震荡等极端事件频繁发生,既使金融机构和个人投资者遭受巨额损失,也给金融体系稳定及宏观经济政策设计带来了巨大挑战.在此背景下对我国金融市场价格运行规律、信息传导机制等市场微观结构进行深入研究,并在此基础上有效识别、测度与管理各种金融风险,对于有效防范与规避风险、控制恶性风险事件的冲击、维护金融系统安全具有重要意义.

20 世纪 90 年代以前,学者们通常只能利用日、周、月、季度或者年度等频率的数据对金融市场的上述问题进行研究,这类金融数据通常被称为低频数据.近年来,计算机技术的快速发展使得数据存储成本大幅降低,证券交易所和数据服务商开始提供高频数据供学者和金融从业人员使用.所谓高频金融数据(high frequency data)即金融资产的日内交易数据,是指在金融市场运行过程中以小时、分钟、秒或更高频率采集的数据.相比于传统的低频率数据,高频数据的一个显著特征就是数据量大.从理论上说,金融市场中的信息会连续不断地对资产价格的动态过程产生影响,采集频率越高,信息丢失则相对越少.当数据采集频率逐渐增高时,数据所包含的资产价格信息渐接近于理论上连续时间的资产价格模型.因此,高频金融数据包含金融资产价格更丰富的信息,能提供对金融市场更精细的分析,这正是研究高频数据的一大优势.此外,高频数据不仅包含丰富的资产价格信息,还包含众多其他维度的信息,例如交易的时间间隔、交易量、买卖价差,等等,这些不同的信息维度对于理解市场价格的形成、信息的传递机制等市场微观结构方面的特征具有相当重要的作用.因此,高频数据有助于对市场微观结构、波动性以及风险测度与管理等传统金融问题进行全新的审视,具有重要的研究价值.

然而另一方面,高频数据又具有与传统的日、周、月、季度或者年度等低频时间序列数据不同的数据特征.例如,受市场交易规则的影响,日内的高频交易价格一般都具有最小变动单位,资产价格数据的取值是离散化的.另外,交易的非同步性以及市场摩擦都会导致数据的实际观测值与理论值之间存在偏差,这类偏差通常称为市场微观结构噪声.现有的研究已经证实市场微观结构噪声的影响会导致具有良好理论性质的计量方法在高频数据的实际应用中失效.正是高频金融数据具备的这些独特特征,给高频数据的理论和实证研究带来了巨大的挑战,从而促使高频数据的研究成为近二十年来计量经济领域研究的热点问题之一.

1.2 应用领域

高频金融数据有助于对金融市场的微观结构、波动性、资产价格的动态规律以及风险度量等问题进行深入研究,具有广泛的应用背景.本节将对高频数据的这些理论和应用研究领域做简要介绍.

1.2.1 市场微观结构

金融市场微观结构的概念有狭义与广义之分.狭义的市场微观结构仅指市场价格发现机制,广义的市场微观结构则是各种交易制度的总称,包括价格发现机制、清算机制与信息传播机制等.关于市场微观结构的研究对象和内容,许多学者从不同的角度给出了不同定义.按 O'Hara(1995)的说法,市场微观结构研究“给定交易规则下,资产的交换过程及其结果”.主要包括有关交易机制的研究,在特定交易规则下价格在资产交易中的动态形成过程,交易机制对交易的结果,即资源配置和信息是否有效,以及相应的社会福利最大化等问题. Madhavan(2000)认为“市场微观结构研究投资者的潜在需求最终转化为价格和交易量的过程”. Engle(2002)从信息和高频金融数据研究角度出发,认为市场微观结构理论主要研究金融资产价格如何根据新信息进行调整,以及交易机制如何影响资产价格.由于市场微观结构的研究内容丰富、包罗万象,本节我们主要介绍基于高频数据的日内收益率序列相关性研究,以及交易机制,例如开收盘机制,对波动性的影响研究.

1. 日内收益率序列相关性

在市场微观结构领域,对资产收益率序列相关性进行研究具有重要的意义,因为资产收益率序列的相关性有助于揭示交易过程的基本特征. Fama(1970)的弱式市场有效性假说认为收益率序列应该不存在序列相关性.如若不然,则资产价格过程就在某种程度上可以预测,即理性交易者可以通过预测资产价格获得超额收益.国内外很多关于弱式市场有效性假说的实证研究显示个股和股指的日收益率序列不存在序列相关性,即支持股票市场弱式有效性的假说.但是也有一些实证结果拒绝了股票市场弱式有效性的假说. Roll(1984)的研究显示个股收益率序列存在负的序列相关性. Atchison(1987)的研究表明股指收益率序列存在正的序列相关性. Lo 和 MacKinlay(1988,1989)构建了方差比检验(variance ratio test)来研究个股和股指收益率序列的序列相关性,并构建理论模型解释了股指收益率序列存在正序列相关是因为异步交易现象(asynchronous trade)的存在. Goodhart 和 Figliuoli(1991)、McNish 和 Wood(1991)的实证结果显示收益率序列存在负的一阶序列相关.以上的实证研究虽然结果不同但都是基于日收益率序列、周收益率序列甚至是

月收益率序列等低频数据的。近年来,随着高频和超高频数据的数据质量和可获得性的提高,很多学者利用高频和超高频数据来研究资产收益率序列的序列相关问题。Low 和 Muthusamy(1996)最先利用高频数据研究收益率序列相关性,他们利用方差比检验研究了不同汇率收益率序列相关问题,并得到了存在负的序列相关性的结论。Thomas 和 Patnaik(2003)利用方差比检验研究了印度股市个股和股指收益率序列的序列相关问题,并研究了序列相关与流动性之间的关系。他们的研究显示个股收益率序列呈现出显著的负相关,而股指收益率序列则不存在显著的序列相关性。Bianco 和 Reno(2006)的研究显示由于买卖跳动效应的影响意大利股指期货收益率序列存在负的序列相关,并且序列相关与波动率之间存在着关系。Bianco 和 Reno(2009)研究了美国 S&P500 股指期货收益率的序列相关性,其实证结果显示收益率序列存在负相关,序列相关与总的日波动率呈正相关而与未预期到的波动率呈负相关。国内学者于亦文等(2005)利用方差比检验检验了上证综指 5 分钟高频数据收益率序列的序列相关问题,其结果显示收益率序列存在显著的正自相关性。吴斌哲等(2008)研究发现上证综指收益率序列大约具有 20 分钟的时间相关性且在 5~15 分钟存在负的序列相关。

2. 开收盘机制对波动性的影响

证券交易机制是指有组织的证券交易场所为履行其基本职能而制定的与证券交易有关的运作规则,它的重要功能之一是使潜在的投资需求转化为实际交易,发现市场的出清价格。这种证券价格的形成机制,能够将投资者的潜在需求和对证券未来走势的预期用实际的价格和成交量反映出来。在古典的经济学理论中,经济学家认为价格是供求关系平衡的结果,和交易机制没有太大的关系。但是现实的市场并非如此,现实市场中不仅存在交易成本,而且信息也各不相同,市场参与者不只是以达到供求平衡为目标,更多地,他们看中的是如何充分利用市场信息使自己的收益最大化。所以,在考虑证券价格发现过程和收益率序列波动性的时候,我们必须考虑交易机制的不同对它们的影响。

世界上现有的证券交易机制主要分为三种:连续竞价交易机制、集合竞价交易机制和做市商交易机制。但常见的只有连续竞价和集合竞价两种。连续竞价是指交易者在输入其买卖指令后,只要系统中有能与其匹配成交的存量指令,输入的指令即可立即执行,因而,在交易时间内,任何时点上交易都可发生,其成交价格完全由“即时”的市场供求订单决定。集合竞价是在规定时间内将订单数量累积到一定数量后再通过竞价最大限度地实现订单执行的一种价格生成机制,集合竞价的成交价为“单一”价格,所有执行成交的订单均按该价格成交。中国大陆证券交易市场的开收盘机制大致如此。每个交易日上午 9:15~9:25 为集合竞价阶段,由此产生的证券当日的第一笔成交价形成开盘价格。接下来从 9:30 开始为连续竞价阶段,

中午 11:30 到下午 1:00 休市,从 1:00 开始直接进入连续竞价阶段,下午 3:00 以后收市,并形成收盘价。收盘价是当日该证券最后一笔交易前一分钟所有交易的成交量加权平均得到的价格。当日没有成交的,以前收盘价作为当日收盘价。集合竞价这种价格决定方式有较好的价格发现功能,对市场的影响成本比较小。开盘价格是市场每个交易日产生的第一个价格,它包含了由于隔夜没有发生交易而累积的市场信息,以及很多未透露的“黑箱”信息,并且它对之后的证券价格变动有着重要影响。这对于弄清楚证券市场的微观结构有着十分重要的作用。收盘价格则包含了一整天的市场信息的累积,也有着不可忽视的地位。

国内外学者对开收盘机制以及它们如何影响证券价格进行了深入研究。Garman(1976)发现交易系统有一个演变模式,即随着交易量的增加,交易机制通常由集合竞价模式转向连续竞价模式。Amihud 和 Mendelson(1987)研究了开盘机制对证券价格变动的影响,他们对纽约证交所道琼斯指数的 30 种成分股的开盘价格收益率和收盘价格收益率作了检验,发现开盘价格收益率方差大于收盘收益率方差,显示了开盘相对于收盘波动性更大。Muscarella 和 Piwowar(2001)对巴黎交易所的股票进行研究发现,从连续竞价到集合竞价的转换降低了股票的成交量和交易价格。王春峰等(2005)采用上证 180 的 100 只典型个股样本对上海股市开收盘的波动性进行了实证研究,结果表明我国股市开盘回报的波动性显著大于收盘波动性。他们还进一步对影响波动性的因素进行考察,认为开盘方式和开盘交易过程中释放的私人信息是造成上述现象的重要原因。黄剑(2007)利用 EGARCH 模型,对 2000 年 1 月至 2007 年 4 月间沪深两市具有代表性的股票及指数的开收盘收益率的波动性进行实证分析,结果表明收益率序列有明显的 ARCH 效应,其波动性具有显著的非对称性的冲击的持续性;在样本期内,上交所的个股和指数未能观察到开盘波动性高于收盘波动性的现象,而深交所个股在 2006 年 7 月实施收盘集合竞价机制之后比较明显地观察到开盘波动性高于收盘波动性的现象。许香存等(2007)运用上海 A 股股票的交易数据,对开盘竞价透明度提高前后的价格变化进行了实证分析,结果表明开盘竞价的透明度提高之后,由开盘到连续竞价价格变化的幅度明显减小,连续竞价开始后 15 分钟内的价格波动率也减少。进一步,根据价格随着信息的变化而进行调整的过程,得到开盘竞价透明度增加能够提高开盘竞价收集信息的功能,稳定连续竞价市场的结论。

1.2.2 市场波动性

波动性是金融市场的重要特征,它与市场的不确定性和风险直接相关,是体现金融市场质量和效率的有效指标,同时也是证券组合理论、资产定价模型、套利定价模型及期权定价公式的核心变量。因此,如何有效地刻画金融市场波动的动态行

为一直是金融市场学研究的热点问题.尤其在金融经济波动频发,各国金融市场相互依存,全球金融风险不断扩大的经济环境下,深入探索波动的本质及其规律性对于金融风险防范与规避具有重大的理论与实际意义.在传统的低频数据研究领域,对波动性的研究主要通过建立时间序列模型,最有代表性的是 GARCH 族模型和 SV 族模型.这两类模型都能很好地刻画金融资产收益率序列波动的各种特性,在实际中有很广泛的应用.但由于高频金融数据所具有的一些独特特征,在低频数据领域具有很好拟合效果的 GARCH 族模型和 SV 族模型对于高频金融数据波动率建模效果不佳.近年来,在基于高频数据的波动率研究中,积分波动率这一类非参数波动率测度受到广泛关注.究其原因,主要是因为它不仅充分利用了日内的高频数据信息,而且在价格服从扩散过程的假设下,积分波动率正好是价格过程的二次变差,它的计算不依赖于模型具体形式,极为简便.

1. 波动率非参数估计

自 Andersen 和 Bollerslev(1998)提出利用日内高频数据计算已实现波动率用以评价 ARCH 类模型的预测能力之后,越来越多的学者对这种事后的积分波动率的非参数估计问题进行了研究. Barndorff-Nielsen 和 Shephard(2002)证明了当资产价格过程服从扩散模型时,已实现波动率是积分波动率的一致估计量,并给出了中心极限定理. Christensen 和 Posolskij(2005)提出利用日内价格极差信息构造已实现极差波动率来估计积分波动率,同样证明了扩散模型假设下估计量的一致性与中心极限定理.当资产价格过程包含跳跃行为时, Barndorff-Nielsen(2004, 2006, 2007)提出了二次幂变差与多次幂变差过程,得到了对跳跃稳健的积分波动率估计量. Christensen 和 Posolskij(2006)结合 Barndorff-Nielsen 的思想提出了对跳跃稳健的已实现极差二次变差过程. Mancini(2004, 2009)提出了阈值已实现波动率.积分波动率的估计是近二十年来高频数据领域研究的热点问题,它实际上刻画的是过去某段时间内波动的总量特征, Kristensen(2010)指出在更多场合下,瞬时波动率的估计对于投资者更实用、更有意义,尤其对于市场上出现的越来越多的高频交易者而言,他们需要比传统时间单位更为精细的日内风险测度,这就需要利用高频数据对日内的瞬时波动率进行估计.

对于瞬时波动率的估计,在低频领域常见的有 ARCH、GARCH 类模型、SV 模型,而在高频领域,这些模型都不再适用,为了避免设定任一模型所带来的误差,非参数估计提供了很好的解决方法.一般而言,瞬时波动率,即价格过程中的扩散函数是既随时间变化又与价格水平有关的,然而由于实际中我们只能观测到价格过程的一条路径,没有足够的信息去确定这样一个二元函数,因此目前的研究都只单独考虑与时间或者价格水平有关的一元扩散函数的非参数估计问题.当价格过程服从扩散过程时, Foster 和 Nelson(1996)最早提出了 rolling window 估计量;

Fan, Jiang 和 Zhang(2003)考虑了期限结构模型的时变扩散函数的非参数估计问题; 另外 Andreou 和 Ghysels(2002), Fan, Fan 和 Lv(2007), Mykland 和 Zhang(2003, 2006)都研究了时变扩散函数的非参数估计问题; Bandi 和 Phillips(2003)利用核方法估计了与价格水平有关的扩散函数. 当价格过程中包含跳跃时, Bandi 和 Nguyen(2003)以及 Johannes(2004)在跳跃过程具有有限活跃度且有界的假设下得到了与价格水平有关的扩散函数的核估计; Mancini 和 Renò(2006)分别在模型包含有限活跃度和无限活跃度跳跃行为的情况下给出了与价格水平有关的扩散函数的核方法估计. 而对于时变的扩散函数的研究却相对较少, 只有 Kristensen(2010)给出了当价格过程服从扩散模型时瞬时波动率的核估计方法, 将已实现波动率与瞬时波动率的估计联系了起来, 并指出当价格过程包含跳跃时可以结合幂变差过程与核估计方法来估计跳跃扩散模型里的扩散函数. 另外, Fan 和 Wang(2008)利用核方法给出了瞬时波动率矩阵的非参数估计.

2. 波动率建模

大量研究结果表明, 高频金融数据存在很多低频数据所没有的特征和信息, 传统的低频波动率模型不再适用. 许多学者在研究已实现波动率序列特征的基础上, 提出了基于高频数据的波动率建模方法. Corsi(2004)基于异质市场假说, 将三类市场参与者的投资行为分别对应波动率的三种异质市场驱动因素, 提出了不同波动率驱动成分的简单叠加模型 HAR-RV. Barndorff-Nielsen 和 Shephard(2004)提出了已实现二次幂变差(realized bipower variation)将波动率分解为连续部分和跳跃部分. Andersen, Bollerslev 和 Diebold(2007)利用 Barndorff-Nielsen 和 Shephard(2004)提出的方法对 Corsi(2004)的模型进行了扩展, 构建了 HAR-RV-CJ 模型, 研究了波动率中跳跃部分和连续部分对波动率预测精度的影响. Brandt 和 Jones(2006)研究表明高低极差(high-low range)可以显著地提高波动率预测的精度. Ghysels, Santa-Clara 和 Valkanov(2006)提出了绝对幂变差(absolute power variation), 指出该统计量可以很好地对波动率进行预测. Engle 和 Gallo(2006)将已实现波动率与 GARCH 模型相结合, 以利用高频数据信息提高日波动率预测精度.

大量实证结果表明, 金融资产收益率序列的波动具有显著的非对称性. 目前国外学者将波动非对称性的产生机制主要归结为杠杆效应(leverage effect). Black(1976)和 Christie(1982)研究发现美国股票的当期收益率与未来波动存在负相关关系, 并试图解释个股收益的这种不对称波动特征, 他们将此现象解释为杠杆效应, 即在其他条件不变的条件下, 股票价格的下跌导致公司价值的降低, 增加债务/权益比率, 从而加剧波动性和持股风险. 反之股票价格的上升会降低债务/权益比率, 减少波动性和持股风险. 因此, 股票当前收益和未来波动之间负相关. Nelson

(1991), Engle 和 Ng(1993), Glosten, Jagannathan 和 Runkle(1993)将杠杆效应引入 GARCH 族模型的建模中,取得了良好的拟合和预测效果. 以上对波动非对称性的研究都是基于低频数据的,近年来,有少数学者利用高频数据研究了波动非对称性的建模问题. Barndorff-Nielsen, Kinnebrock 和 Shephard(2009)提出了已实现半方差(realized semi-variance)来度量高频金融数据的下方风险,并将已实现半方差引入波动率模型的建模中以研究波动的非对称性. Visser(2008)提出了向下绝对幂变差(downward absolute power variation)并将其引入 GARCH 族模型的建模中以提高对未来波动率的预测精确度. Corsi 和 Reno(2009)对 HAR-RV-CJ 模型进行了进一步的推广,加入了杠杆效应的影响,构建了 LHAR-RV-CJ 模型.

近年来,国内不少学者也对基于高频数据的市场波动性问题进行了研究. 施红俊、马玉林和陈伟忠(2003)总结了关于已实现波动率的研究成果. 张伟、李平和曾勇(2008)对中国股票市场的个股进行了已实现波动率的估计. 徐正国和张世英(2004),唐勇和刘峰涛(2005)则比较了已实现波动率、GARCH 模型和 SV 模型对波动率的预测能力,通过实证研究表明已实现波动率优于其他波动率模型的预测结果,并阐述了后两种方法的维数灾难等问题. 在波动率的长记忆性研究方面,陈梦根(2003),王春峰、张庆翠和李刚(2003)对中国股市的长记忆性进行了实证研究. 徐正国和张世英(2004)提出了用调整已实现波动来降低其测量误差,并且用 ARFIMAX 模型来研究这个调整的已实现波动率的长记忆性. 李胜歌和张世英(2007)讨论了在估计波动率时二次幂变差和多次幂变差的有效性问题. 目前,国内也有大量学者对 HAR-RV 模型进行研究与推广. 西村友作和门明(2009)采用跳跃显著性检验方法、HAR-RV-J 及 HAR-RV-CJ 模型进行了深证成指已实现波动率的跳跃特征研究. 他们发现中国股市的已实现波动率发生显著性跳跃的频率高、幅度大,但是有些跳跃成分不影响未来波动率的预测. 同时,陈国进和王占海(2010)对沪深 300 指数进行了分析,分离出已实现波动率中的连续性波动和跳跃性波动序列,并检验了这两种波动成分的统计性质和杠杆效应. 他们发现,我国 A 股市场的连续性波动与跳跃性波动都具有显著自相关性,并且这种滞后相关性较美国股票市场要更为长久,而杠杆效应在考察期内则不显著.

1.2.3 资产价格跳跃行为

传统的资产价格模型以效率市场理论(EMH)为基础,认为资产价格主要反映经济体及企业自身基本面的变化. 然而,我们在现实的金融市场观察到,即使经济基本面或者企业基本面没有出现显著变化,资产价格也有可能出现极端波动;同时,未预期到的信息如企业特殊事件或者央行公告等事件的披露也会使资产价格出现极端波动. 近年兴起的行为金融学更是指出,即使没有任何实体经济面的变

化,投资者心理变化也会导致极端的资产价格变动.近年来基于高频数据的实证研究结果也印证了我们在现实金融市场中观察到的资产价格运动的跳跃行为.学界关于资产价格模型的研究主要分为连续时间模型和非连续时间模型.连续时间资产价格模型构建一直是金融领域的热点问题,大量学者对该问题进行了深入研究并取得了一系列重要的成果.粗略地说,金融资产价格连续时间模型构建经历了如下几个阶段.

(1) 扩散过程(连续路径)

扩散过程是最基本的资产价格模型.在扩散过程中,资产价格被假定来源于连续过程,即资产价格运动路径是连续的.例如,Black 和 Scholes(1973),Merton(1973)提出的著名的期权定价模型即假定资产价格服从隶属于扩散过程的几何Brown(布朗)运动.

(2) 跳扩散过程(扩散过程+大跳)

金融市场不时出现的极端资产价格变化使人们意识到用连续时间扩散过程构建资产价格模型存在一定的局限性,有学者开始利用跳跃扩散过程来构建资产价格模型.跳跃扩散过程包括用来解释价格过程中连续性波动的扩散过程及解释价格极端变化的跳跃过程,如Merton(1976)对扩散过程进行扩展加入了Poisson跳跃过程.Aït-Sahalia, Wang 和 Yared(2001), Andersen, Benzoni 和 Lund(2002), Ball 和 Torous(1985), Bates(1996), Duffie, Pan 和 Singleton(2000)也分别从理论和实证角度对跳跃扩散过程进行了研究.

(3) 纯跳过程(大跳+小跳)

纯跳过程仅使用跳跃过程来对资产价格进行建模,即资产价格运动过程仅由市场本身引起的微小跳跃过程和由企业特殊事件或者央行公告等事件的披露导致的资产价格极端波动组成.一些学者认为对资产价格建模时没有必要包含连续成分,如Geman, Madan 和 Yor(2001), Carr, Geman, Madan 和 Yor(2002)认为由市场出清条件引起的资产价格应该构建为一纯跳过程模型而不包含连续成分.

(4) Itô(伊藤)半鞅(扩散过程+大跳+小跳)

Itô 半鞅是描述资产价格过程最一般化的模型,它包括连续成分、大跳和小跳,即资产价格运动过程由资产价格的连续性运动、市场本身引起的微小跳跃过程和由企业特殊事件或者央行公告等事件的披露导致的资产价格极端波动三部分组成.实际上,样本总是在离散时间上观测到的.所以,通过样本决定一个金融时间序列是否能由扩散过程、跳跃过程或两者的组合构建模型就变得十分重要.在大跳发生的情况下,对于数据集直观的判断就足以解决这一问题.但是对大多数实际时间序列的直观判断并不能提供小跳是否存在的清晰证据.由于频繁发生的小跳无疑

应该包含在资产价格模型中,而且模型中是否含跳跃会使模型具有不同的数学性质并导致期权套期保值、资产组合优化等金融实践活动产生不同的结果,所以解决这一问题的统计方法就变得极为重要。

连续时间扩散过程模型已经成为金融研究的一个简便工具,它提供了市场完备性和无套利衍生定价公式,这些假设在 Black-Scholes 期权定价模型及相关的动态对冲策略的有效性中十分关键.但大量的理论和实证研究表明连续时间扩散过程模型难以准确地描述现实中的资产价格运动.自 Merton 的开创性论文以来,金融计量经济学界对资产价格跳跃行为的研究进行了深入的研究,相关文献回顾可见 Cont 和 Tankov(2004).随着高频数据的可获得,有部分学者研究了通过样本确定资产价格过程是否存在跳跃的统计检验问题. Aït-Sahalia(2004)使用转移函数构造了判断跳跃是否存在的参数检验方法. Carr 和 Wu(2003)基于高频数据研究了短期期权的微分性质并以此为基础构造了跳跃是否存在的检验方法. Barndorff-Nielsen 和 Shephard(2004,2006)提出了跳跃稳健波动率估计量——已实现二次幂变差,并基于已实现方差和二次幂变差之间的差值构建了跳跃检验统计量. Cont 和 Mancini(2011)提出了已实现门限积分波动率并根据该波动率估计量构造了跳跃检验方法. Fan 和 Wang(2007)利用小波技术对资产价格过程中的跳跃成分进行剥离并据此提出了跳跃检验方法. Jiang 和 Oomen(2008)基于对方差掉期复制策略应用 Itô 公式构造了跳跃检验统计量. Lee 和 Mykland(2008)通过对已实现收益率进行标准化得到了非参数的跳跃检验方法. Aït-Sahalia 和 Jacod(2009)根据已实现幂变差的渐近理论性质提出了确定资产价格过程中跳跃是否存在的非参数检验方法。

近年来,国内学者对基于高频数据的资产价格过程跳跃行为相关问题进行了研究.李胜歌和张世英(2007)研究表明跳跃行为对于股票市场波动过程具有重要影响.王春峰、姚宁、房振明和李晔(2011)以上证综指数据为研究对象,研究了已实现波动率中的跳跃行为,他们发现已实现波动率中的连续成分占据了波动率的主要部分,但是跳跃行为对波动率也有显著的影响.沈根祥(2010)构造了跳跃行为的 Hausman 检验统计量并对沪深 300 指数高频数据进行了分析,研究发现沪深 300 指数样本区间内有三分之一以上的交易日存在跳跃行为.陈国进和王占海(2010)根据 Barndorff-Nielsen 和 Shephard(2004,2006)的思路研究了中国股票市场的跳跃行为,他们发现中国股票市场的跳跃行为比美国股票市场具有更长期的滞后相关性.王春峰、郝鹏和房振明(2011)基于高频数据特征构建了跳跃检验及估计跳跃时刻和规模的统计方法,并研究了跳跃的特征及跳跃与信息融入效率之间的关系.杨科和陈浪南(2011)基于中国股票市场数据运用核估计量及修正的已实现门限多次幂变差估计跳跃行为及相应的跳跃特征。