

6.2 日内跳跃试验综述

6.2.1 已实现波动率的度量和 BNS 试验

在金融经济学中，无套利原理要求资产价格动力学遵循一个半鞅过程。具体地，有效的日志价格过程过滤概率空间的定义 $(\Omega, \mathcal{F}, \mathbb{P})$ ， $(X_t)_{t \geq 0}$ 可以分解为：

$$X_t = X_0 + \int_0^t \sigma_s dW_s + M_t, \quad t \geq 0, \tag{6.1}$$

¹⁰[37, 70].

在有限的变异和 MT 是一种局部鞅[15]。11 是一个特殊的情况下
^ITO 带 Poisson 跳价格半鞅：

$$X_t = X_0 + \int_0^t a_u du + \int_0^t \sigma_u dW_u + \sum_{k=1}^{N(t)} J_k,$$

T

凡 $= () T \geq 0$ 是局部有界可预见的漂移功能， $= (T T) \geq 0$ 改编的 CA ` DLA ` G 波动过程， $W = (WT) T \geq 0$ 标准布朗运动， $n (t)$ 的计算过程，和 JK 价格跳[31]的大小。（6.2）给出的规格，价格过程的总的变化是由二次变异 QV [71]测量：

M

$$[X]_t = \text{plim}_{M \rightarrow \infty} \sum_{i=1}^M (X_{t_i} - X_{t_{i-1}})^2,$$

一个分区的时间 $0=t_0 < t_1; \quad t_1 < t_2; \quad \cdots <$

$$[X]_t = [X]_t^c + [X]_t^d = \int_0^t \sigma_u^2 du + \sum_{0 \leq u \leq t} J_u^2.$$

(6.4) 中，第一项也被称为集成的波动 (IV) 在第二任期将跳 QV。鉴于 QV 的定义 (6.3)，并假设一个无摩擦的市场结构，这是很自然的估计价格从 0 到 T 的 QV 平方和的回报，称为已实现方差衡量 RV [6; 14]:

$$RV_t = \sum_{i=1}^{\lfloor t/\Delta \rfloor} (X_{t,i\Delta} - X_{t,(i-1)\Delta})^2,$$

Δ 的长度是高频段和 T / Δ $\lfloor T / \Delta \rfloor$ 是整数部分。我们按照惯例在文献中，并设置 $T = 1$ 。在 (6.5), $X_{t,i\Delta} - X_{t,(i-1)\Delta} = r_{t,i\Delta}$ 因此，长度 Δ 第 i 个高频返回 $M = 1 / \Delta \lfloor \cdot \rfloor$; 可以看出，作为 Δ 缩小到 0 或等价的 $M \rightarrow \infty$, *RV 收敛于价格过程中起作 QV*

11a 的过程是一个局部鞅如果有增加的时间顺序，停止过程是一个鞅 [66] 停止。
12unlike 常规跨度渐近收敛性定理，在这里
填补渐近。

Barndorff-Nielsen 和 Shephard [15; 16]提出了估算和 QV 连续部分，采用共称为双动力变化（BP）：

$$BP_t = \mu_1^{-2} \left(\frac{M}{M-1} \right)$$

在 $E(Z|Z \sim N(0, 1))$ 和 $a > 0$ ，就是 a 近零生。计量可以构建无跳跃的零假设，给出一个了 RV 和 BP 的二维提供个了 RV 和 BP 的二维连续检验分布；在 [51]，表明，BP 是在跳跃的存在一致连续检验分布；在 [51]，的估计，作为影响收益 RT 单的跳经验建议以下统计：跳的概率， $rt_{i,t}$ 和 $rt_{i,t-1}$ 渐跃发统

$$Z_{BNS,t} = RJ_t / \sqrt{0.609 \frac{1}{M} \max \left(1, \frac{TP_t}{BP_t^2} \right)} \sim N(0, 1), \quad (6.7)$$

一天不跳的零假设被拒绝如果 $z_{BNS,T}$ 大于临界值从标准正态分布在某些意义层面。在 Barndorff-Nielsen 和 Shephard [16]的仿真研究，虽然试验有轻度过分拒绝，检测其功率较低跳不过，文献 [16 实证结果；51]发现更多的跳跃的价格比建议的参数模型在 [40]。BP 和多变化的进一步的理论结果是 [19; 78].¹⁴

6.2.2 ABD 和 LM 检验

确定高频价格跳跃，ABD（2007）认为高频返回 RT, J ，表明没有跳空下的

¹³around 20%在 5%级的时候有一个单跳每天占 5%的总价格变化。
¹⁴jump 测试使用的措施实现构建包括，除其他外，阈值为基础的估计 [62, 63]，在 [2]的功率变化率、阈值附 BP 在 [34]，基于分位数的 [31]和 [67 估计，自举技术]。

高频返回 Δ -规模 $1 / 2RT$, J 已经消失的均值和方差接近综合方差四:

$$\lim_{\Delta \rightarrow 0} E \left[\Delta^{-1/2} r_{t,j} \right] = 0,$$

$T \rightarrow$

Δ 假设现货的波动在一个交易日是恒定的, ABD 试验检测价格上涨如果:

$$\left| r_{t,j} \right| / \sqrt{BP_t/M} > \Phi_{1-\beta/2}, \tag{6.9}$$

在 $\Phi_{1-\beta/2}$ 是临界值从标准正态分布 significance level β . Equivalently, under the null of no jumps the high-frequency return $r_{t,j}$ standardized by its spot volatility BP_t/M is a standard normal variable. 的测试的基本原理是在 [17] 的评论, “因此一致...非高斯增量和跳跃是同义词。”

在与测试, 返回 RT, J 点的波动, 通过 T, J 表示, 通过 $\sqrt{BP_t/M}$ 估计, 测试不能直到最后完成交易时, 所有日内收益的实现。在 [58], 相反, 跳跃的检测是在一个实时的基础上完成。

$$\sigma_{t,j} = \sqrt{\frac{1}{k-2} \sum_{i=j-k+2}^{j-1} |r_{t,i}| |r_{t,i-1}|}. \tag{6.10}$$

他们的模拟结果, Lee 和 Mykland [58] 建议 $K = \sqrt{MT}$, 其中 T 是样本中的天数。一年有 250 个交易日, 采样频率为一天 24 分钟的交易时间 (汇率数据) 会给 $K = 268$ (圆整)。

ABD 和 LM 检验确定检验统计量作为绝对的标准化回归 $|RT, J| / T, J$, 与现货的波动 T, J 估计不同。选择一个临界区, 及 (2007) 考虑一日

显著性水平 从 M 的最大观测分布

intraday statistics, $\max_{j=1..M} (|r_{t,j}| / \sigma_{t,j})$. Under the null hypothesis, each of the M statistics is independent standard normal, and hence the distribution function of $\max_{j=1..M} (|r_{t,j}| / \sigma_{t,j})$ is given by the probability of M 没有成功的独立的二项试验:

$$\Pr \left(\max_{j=1..M} \frac{|r_{t,j}|}{\sigma_{t,j}} \leq u \right) = \Pr (\text{Bin} (M, \beta) = 0) = (1 - \beta)^M, \tag{6.11}$$

在 $\text{Bin} (m, \quad)$ 成功的概率是 二项式变量, 这也是盘中的意义层

面，即 $\alpha = 2(1 - F\Phi(U))$ ， $F\Phi(\cdot)$ 标准正态累积分布函数。所有接受概率 $(1 - \alpha)M$ (6.11)，等于 $1 - \alpha$ 等：

$$\alpha = 1 - (1 - \alpha)^{1/M}. \tag{6.12}$$

Rasmussen[69]和 boudt 等人。[28]指出，调整 α 和 M 之间 (6.12) 是的 \sim IDA 与多重比较检验 K 校正。一个共同的选择是 Bonferroni 过程，使 $\alpha = \alpha / M$ [1]。在他们的模拟研究，及 (2007) 发现他们的测试是中度过大，有 2 分钟 0.153% 实证大小返回给定 0.1% 级。作为一个结果，一个更有效的校正是必须的。15

由于多个测试解决过度拒绝问题，李和米克兰[58]考虑渐近分布的最大测试统计样本大小 $n = Mt$ 。没有跳空下，与 T, J 在 (6.10) 的 LM 检验统计量的标准的正常比例是渐近独立。表示统计锰最大的观察，李和米克兰[58]使用极值理论表明，有两个常数 CN 和 SN，锰由这些常数的限制标准 Gumbel 分布， $N \rightarrow \infty$ ：

$$c_n = \sqrt{2\log(n)/c} - [\log(\pi) + \log(\log(n))] / (2c\sqrt{2\log(n)}),$$
$$s_n = 1 / (c\sqrt{2\log(n)}), \tag{6.14}$$

15 相关研究认为 bajgrowicz 和 Scaillet[9]，把问题控制在多个测试由 Benjamini 和 Hochberg 提出 [23] 的错误发现率。

在 \sim IDA 与 K 校正和耿贝尔临界值模拟表 6.1，1% 级。

Frequency	2 min	5 min	10 min
M	195	78	39
Šidàk correction	0.0105	0.0104	0.0099
Gumbel: c_n in (6.14)	0.0047	0.0039	0.0030
Gumbel: c_n in (6.16)	0.0121	0.0103	0.0084

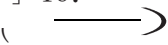
表 6.1 报告的 \sim IDA 与 K 校正和耿贝尔临界值的平均抑制率从 1000 独立同分布 $N(0, 1)$ 样品在 1% 日线级别 变量。每个样本长度为 1 年，即 250 个交易日；每一天的 6.5 个交易日小时采样每 2，5，和 10 分钟，因此 M 是 195，78，和 39。

in which $c = \frac{\sqrt{2\log(n)}}{2\pi}$. The LM test detects a jump in $r_{t,j}$ if:

$$\left(\frac{|r_{t,j}|}{\sigma_{t,j}} - c_n \right) / s_n > -\log(-\log(1 - \alpha)), \tag{6.15}$$

高斯变量，我们有信心在拒绝零跳跃在大， J 。

在这里，我们注意到，在 (6.14)，为中国的方程包含一个打印错误，并在 [3] 中给出正确的表达式：46] 16：



$$c_n = \sqrt{2 \log(n)} - [\log(4\pi) + \log(\log(n))] / (2$$

重要的是，缺失的因素“4”在 CN 解释非常保守，在 [58] 矮小的结果。量化这一因素的影响，我们模拟了 1000 份数 $n \in (0, 1)$ 变量和计算这些样品的平均拒绝率给出了一个 1% 的水平。每个样本长度为 1 年，即 250 个交易日；每一天的 6.5 个交易日采样每 2, 5, 和 10 分钟，因此 M 是 195, 78, 和 39。

16 在 [35, 43, 46, 77] 也可以看。

表现良好。CN 的正确表达了拒绝率接近 5 分钟频率标称 1%级；相反，用 cn Gumbel 分布的临界值 (6.14) 是严重不足的三个频率。

我们总结了这一部分，总结了许多经验的作品，已经实施了这两项测试。表 6.2 列出了一些这些研究；可以看出，大部分已经注意到组织真正的噪声对测试的影响。这个问题是减轻了采样在一个安全的频率或直接占使用噪声鲁棒波动性的措施。例如，Bos 等人。[27]使用简明释义估计 jacod 等人。[52]在有噪音和跳跃的存在下，获得第四的一致估计。另一方面，boudt 等人。[28]表明，ABD 和 LM 检验进行调整，以避免失真的大小是必要的。在拒绝域方面，大多数研究只是依照 ABD 和 LM 检验了。