

股市指数日间变化与夜间跳跃关系的 非参数计量研究

赵 钊

(上海交通大学 安泰经济与管理学院,上海 200240)

摘 要:运用非参数统计分析方法,研究股市指数夜间跳跃与日间变化之间的关系;基于 R 软件,采用基本统计量、散点图与拟合函数图,对全球范围内 7 个股票交易市场的每日交易数据进行分析;以上证综合指数为例,通过渐进正态分布和 Bootstrap 置信带分析,得到夜间跳跃和日间变化拟合度较高的百分比夜间跳跃区间,依据区间内夜间跳跃和日间变化的关系,提出了基于夜间跳跃的股市可获利交易策略。分析认为:夜间跳跃和日间指数变化存在相关关系;在不同的股票市场,其相关关系不同;技术分析对指导股票市场的投资是有效的。

关键词:非参数统计;股市指数;置信带;可获利投资策略

中图分类号:F830.91

文献标志码:A

文章编号:1671-6248(2012)03-0054-06

在观察股市指数时会发现,股市指数不仅在每天的交易时间内不断地变化(日间变化),而且每个交易日的开盘指数与上一个交易日的收盘指数之间也有一个差值(夜间跳跃)。股市指数的日间变化和夜间跳跃从不同侧面反映了各个产业、市场的各种金融变化讯息,在本质上有着内在的联系。运用数理统计的方法,研究股市日间变化和夜间跳跃之间的关系,从而提出股市可获利交易策略,具有重要的实际应用意义。

对股票市场的分析,主要分为技术分析和基本分析。技术分析是指研究过去金融市场的资讯,并结合统计分析方法来预测股票价格的趋势与决定投资的策略,分析股市日间变化与夜间跳跃之间的关系进而提出投资策略。技术分析作为最早的投资分析方法,可追溯到 19 世纪末的道琼斯理论。然而,许多金融经济学家认为技术分析在预测股票价格方面不具有有效性,因为他们认为股票价格是随机游动的,投资者利用技术分析无法预测股票价格,如 Fama

等认为技术分析对股票市场是有效的,即股票价格已充分反映了所有相关的信息,利用技术分析并不能获得超额报酬^[1-2]。然而, Brock 等以 1897 年第一个交易日至 1986 年最后一个交易日道琼斯工业平均指数为研究对象,利用简单技术分析交易规则,如移动平均线法、阻力线法、支撑线法,结果发现技术分析有助于预测道琼斯工业平均指数^[3]。自此以后,许多学者探讨技术分析交易规则是否有助于预测各国股票价格^[4-6],研究结果表明,技术分析对股票市场是有效的。

笔者以全球范围内 7 个股票交易市场从 2005 年 1 月 1 日之后第一个交易日开始至 2012 年 4 月 25 日的每日交易数据为研究对象,运用非参数统计分析方法,研究股市指数夜间跳跃与日间变化之间的关系,并以上证综合指数的分析为例,依据百分比夜间跳跃区间内,夜间跳跃和日间变化的关系,提出了基于夜间跳跃的股市可获利交易策略。

收稿日期:2012-05-15

作者简介:赵 钊(1989-),男,甘肃兰州人,学生。

一、非参数回归理论

回归模型分为参数回归模型和非参数回归模型。参数回归模型的回归函数形式都是预先设定的,只是其中的未知参数需要估计,这一类回归模型最大优点是回归结果可以外延,模型的参数一般都有明确的经济意义,当然这一切需要具有很明确的前提条件,即回归模型的假设成立。如果回归模型及参数的假定与实际情况相背离时,推断和预测就会有较大的偏差,拟合效果也会很差,甚至没有什么实际意义。因此,参数回归模型缺乏稳健性。与参数回归模型相反,非参数回归模型的回归函数形式是未经设定的,其形式约束较少,解释变量与被解释变量的分布也很少限制。因此,它具有更大的适应性。

由于股市指数受各方面因素影响,对选取的样本数据不能确定其符合某一种既定分布,不能用传统的参数统计方法进行数据的分析,所以,非参数计量方法更适合对股市指数数据进行处理与分析。

(一) 非参数回归函数与真实函数的一致性

一致性是估计量最基本的属性。设有变量 X , 其中包含数据 $x_i (i = 1, 2, 3, \dots, n)$, X 的真实函数为 $f(x)$ 。取 X 的一部分数据 $x_1, x_2, x_3, \dots, x_n$ 进行非参数回归拟合,得到的拟合函数为 $\hat{f}(x)$ 。以下说明当 $n \rightarrow +\infty$ 时, $\hat{f}(x) \rightarrow f(x)$, 即非参数回归函数与真实函数具有一致性需要满足的条件。

根据文献[7]有:

假设 1: 观测值 $x_1, x_2, x_3, \dots, x_n$ 满足独立同分布;

假设 2: 当 $n \rightarrow +\infty$ 时, $h = h_n \rightarrow 0$, 其中 h 为带宽, h 为 n 的函数, 当 $n \rightarrow +\infty$ 时, $h \rightarrow 0$;

假设 3: 当 $n \rightarrow +\infty$ 时, $nh_n \rightarrow +\infty$;

假设 4: K 为 Borel 可测有界实值函数, Ψ 为 K 的变量, 且同时满足:

$$(1) \int K(\Psi) d\Psi = 1,$$

$$(2) \int |K(\Psi)| d\Psi < +\infty,$$

$$(3) \text{当 } |\Psi| \rightarrow +\infty \text{ 时, } |K(\Psi)| \rightarrow 0,$$

$$(4) \text{最小上界 } |K(\Psi)| < +\infty,$$

$$(5) \int K^2(\Psi) d\Psi < +\infty.$$

假设 5: 在任意一点 x_0 , $f(x)$ 连续, 且 $\int |f(x)| \cdot$

$dx < +\infty$ 。

在现实操作中,经济数据不可能是独立同分布的,假设 1 就不成立了。而当所用数据是时间序列数据时,不要求数据独立同分布,假设 1 也成立。

定理 1: 当假设 1 ~ 5 满足时, $\hat{f}(x)$ 与 $f(x)$ 具有弱式一致性, 即当 $n \rightarrow +\infty$ 时, 概率 $P[|\hat{f}(x) - f(x)| < \varepsilon] \rightarrow 1, \varepsilon \rightarrow 0$ 。

定理 1 说明了在数据量足够大时,非参数拟合得到的拟合函数趋向于真实函数。

(二) 置信带的构建

1. 渐进正态分布方法构建置信带

现在假设有数据 $(y_i, x_i) (i = 1, 2, 3, \dots, n)$, 变量 y_i 和 x_i 的非参数回归模型为

$$y_i = m(x_i) + u_i \quad (1)$$

式中: $m(x_i)$ 为回归函数; u_i 为回归函数与真实函数之间的差值, 即残差, 且 u_i 满足独立同分布。

根据文献[7]有:

假设 6: 当 $n \rightarrow +\infty$ 时, $h = h_n \rightarrow 0, nh_n \rightarrow +\infty$;

假设 7: x_i 独立同分布且与残差独立;

假设 8: 同假设 4;

假设 9: 期望 $E|u_i|^{2+\sigma} < +\infty$, 且存在 $\sigma > 0$, 使得 $\int |K(\Psi)|^{2+\sigma} d\Psi < +\infty$, 其中 σ 为标准差。

当假设 6 ~ 9 满足时, $\sqrt{nh}\hat{m}(x) - m(x)$ 服从渐进正态分布 $N(0, f(x)^{-1}\sigma^2 \int K(\Psi)^2 d\Psi)$, 其中 $\hat{m}(x)$ 为 $m(x)$ 的拟合函数。由此, 可以得到回归函数 $y = m(x)$ 的 95% 的置信带为 $\hat{m}(x) \pm 1.96 \sqrt{nh[f(x)^{-1}\sigma^2 \cdot \int K(\Psi)^2 d\Psi]^{1/2}}$ 。

2. Bootstrap 方法构建置信带

除了渐进正态分布方法, 另外还有一种 Bootstrap 方法可以用来构建置信带, 且受到了更多的关注。将数据 $x_1, x_2, x_3, \dots, x_n$ 重新排序 j 次, 得到 j 组重新排序的数据 $x_1^j, x_2^j, x_3^j, \dots, x_n^j$, 用第 j 组数据做回归, 因为数据只是重排, 并没有改变, 所以拟合函数 $\hat{f}^j(x)$ 的期望 $E[\hat{f}^j(x)] = \hat{f}(x)$, 再用上面的方法求得的置信带是对应于期望 $E[\hat{f}(x)]$ (而不是 $f(x)$) 的置信带。这时直接用真实值和回归值之间的差值, 即可求得对应于 $f(x)$ 的置信带, 这就是 Bootstrap 构建置信带的方法。根据文献[7]的分析, Bootstrap 的范围误差比渐进正态分布的要小。

二、数据的处理与分析

(一) 基本统计量分析

笔者对全球 7 个股票市场, 即上证综合指数、深

圳成分股指数、道琼斯指数、纳斯达克指数、德国指数、日经指数和香港恒生指数,从 2005 年 1 月 1 日之后第一个交易日开始至 2012 年 4 月 25 日的每日交易数据进行分析(数据来源为雅虎金融数据库)。数据包含 7 个项目,分别为交易日期、开盘价、最高价、最低价、收盘价、成交量和调整后的收盘价。

每日开盘价 O_i 与上一个交易日的收盘价 C_{i-1}

和自变量夜间跳跃 P_i 的关系为

$$P_i = O_i - C_{i-1} \quad i = 1, 2, 3, \dots, n \tag{2}$$

每日调整后收盘价 C_i 与当日的开盘价 O_i 和因变量指数变化 E_i 的关系为

$$E_i = C_i - O_i \quad i = 1, 2, 3, \dots, n \tag{3}$$

夜间跳跃和日间变化的基本统计量见表 1、2。

由表1、2可知:夜间跳跃和日间变化的平均值聚集

表 1 夜间跳跃基本统计量

股指指数	最小值	中值	平均值	最大值	1 分位数	3 分位数	全距	标准差	峰度	偏度
上证综合指数	-247.51	-0.350	-1.814 0	261.54	-7.040	4.660 0	509.05	23.860 100	27.864 330	0.589 771
深圳成分股指数	-829.20	-1.735	-9.240 0	547.70	-32.068	24.235 0	1 376.90	92.420 800	18.483 250	-1.271 740
道琼斯指数	-86.32	-0.500	-1.064 8	90.41	-2.520	0.597 5	176.73	7.640 305	52.200 450	2.186 709
纳斯达克指数	-118.82	0.930	0.369 9	104.80	-7.120	7.980 0	223.62	16.654 190	9.618 939	-0.375 320
德国指数	-363.19	2.850	1.232 0	245.46	-14.560	19.270 0	608.65	38.182 120	14.510 690	-0.921 170
日经指数	-276.31	7.650	2.849 0	236.32	-56.575	66.843 0	512.63	89.177 100	2.611 701	-0.219 100
香港恒生指数	-5.70	17.500	12.801 0	31.66	2.535	21.040 0	37.36	9.461 988	1.413 637	-0.164 830

表 2 日间变化基本统计量

股指指数	最小值	中值	平均值	最大值	1 分位数	3 分位数	全距	标准差	峰度	偏度
上证综合指数	-311.42	2.160	2.438 00	257.15	-14.980	23.740	568.57	50.206 390	8.752 556	-0.791 25
深圳成分股指数	-1 037.22	9.005	12.790 00	902.69	-63.623	106.103	1 939.91	204.243 200	6.383 386	-0.416 26
道琼斯指数	-774.17	7.485	1.564 00	925.19	-49.750	58.877	1 699.36	132.066 900	9.483 210	-0.168 29
纳斯达克指数	-163.43	1.460	0.106 70	109.65	-13.560	15.000	273.08	26.058 770	5.565 908	-0.387 34
德国指数	-502.49	3.190	0.059 23	508.54	-31.470	36.820	1 011.03	73.700 780	8.584 182	-0.171 05
日经指数	-942.40	-3.440	-3.941 00	1 039.63	-64.540	61.740	1 982.03	136.460 600	10.040 580	-0.466 70
香港恒生指数	-33.20	-17.340	-12.790 00	5.20	-21.020	-2.500	38.40	9.495 601	1.429 391	0.167 56

在 0 附近;根据夜间跳跃和日间变化的标准差分析,深圳成分股指数在这 7 支指数中是最不稳定的指数,相对而言,上证综合指数在这 7 支指数中是比较稳定的;香港恒生指数夜间跳跃和日间变化的峰度都小于 3,说明其是变化比较平缓的指数。

(二)非参数回归分析

对 7 个股票市场的夜间跳跃和日间变化数据,运用 Local-Linear 非参数回归分析,得到 7 个股票的数据量、带宽、剩余标准差和判定系数 R^2 (R^2 表示夜间跳跃和日间变化之间的相关性),见表 3。

由表 3 可知,带宽从小到大排序为上证综合指数、纳斯达克指数、深圳成分股指数、道琼斯指数、日经指数、香港恒生指数、德国指数,带宽越大,回归函数越平滑,但也越不精确。剩余标准差表示拟合值和真实值之间的差值,其值越小,拟合精度越好。由表 3 可知:纳斯达克指数拟合精度最好,香港恒生指数最差,上证综合指数的相关性最高, $R^2 = 0.351\ 027$ 。

表 3 各股市非参数回归数据

股指指数	数据量	带宽	剩余标准差	R^2
上证综合指数	1 833	0.207 911	1 646.605	0.351 027
深圳成分股指数	1 732	1.535 218	31 051.120	0.257 769
道琼斯指数	1 498	3.460 744	16 006.040	0.082 921
纳斯达克指数	1 841	0.493 226	573.237	0.158 172
德国指数	1 869	20.927 420	5 266.971	0.030 107
日经指数	1 792	6.071 607	16 229.620	0.128 500
香港恒生指数	1 834	16.762 700	46 159.200	0.100 264

1. 散点图和拟合函数图

运用 R 软件^[8],绘制各股票市场夜间跳跃和日间变化的散点图和拟合函数图^[9-10],见图 1、2。

7 个股市的散点图基本分为 3 个类型。上证综合指数、道琼斯指数的散点图(图 1(a))呈“十”字型,表明在夜间跳跃的特定区间内,夜间跳跃和日间变化的关系较为明显;而日经指数的散点图(图 1(b))呈“一”字型,这是由于日本的许多经济政策跟随美国而变动,美国市场上的信息会对日本市场造

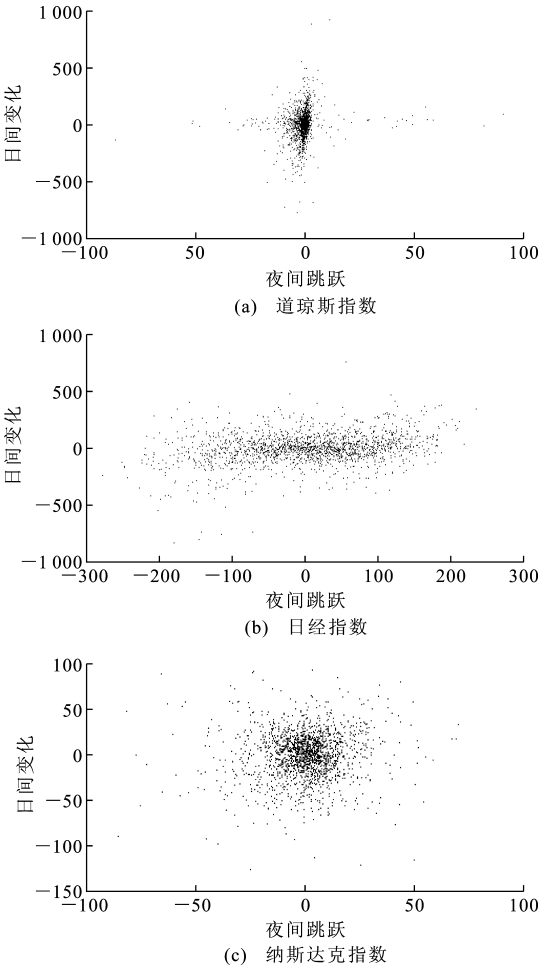


图 1 夜间跳跃和日间变化的散点图

成影响,在夜间跳跃中已经反映了一些信息,所以在日间指数变化中所反映的信息量就减少了,反映在图上就是扁平状;其余 4 个股市的散点图呈“球”型,表明夜间跳跃和日间变化的关系不显著,也可看作是市场比较有效。纵观 7 个股市夜间跳跃和日间变化的拟合函数图,可以看到在夜间跳跃为 0 的邻域,上证综合指数和道琼斯指数日间变化随夜间跳跃的变化起伏明显,说明夜间跳跃对日间变化影响显著;日经指数日间变化随夜间跳跃变化起伏不大,说明夜间跳跃对日间变化有影响,但影响不显著;其余 4 个股市,日间变化基本不随夜间跳跃的变化而变化,说明夜间跳跃对日间变化没有影响。夜间跳跃对日间变化的影响区域,不仅反映了拟合函数的带宽,也反映了各个股票市场上 2 个变量拟合程度较高的区域。

2. 置信带

由上述分析可知,上证综合指数夜间跳跃和日间变化关系较为明显,相关性最高($R^2 = 0.351\ 027$),

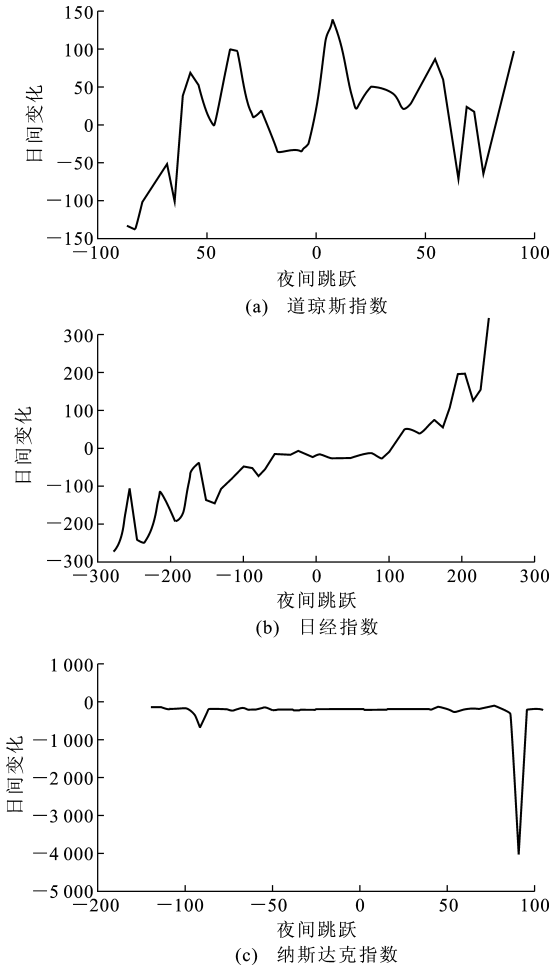


图 2 夜间跳跃和日间变化的拟合函数图

因此,选择上证综合指数为研究对象,通过置信带(区间)分析,确定夜间跳跃和日间变化拟合度较高的区域。为了便于比较,需要将夜间跳跃和日间变化的值百分比化,即分别用夜间跳跃和日间变化除以当日收盘价再乘以 100 得到新的变量:百分比夜间跳跃和百分比日间变化,然后用新变量做非参数回归分析。分别用渐进正态分布和 Bootstrap 方法求置信带,并比较在 2 种置信带方法下拟合度较高的夜间跳跃区间^[11-13]。

按上述思路将股票市场的数据进行处理,处理后的数据如表 4 所示。

采用渐进正态分布和 Bootstrap 方法分别计算股票市场置信带,调整百分比夜间跳跃的范围,得到拟合程度较高,即接近日间变化为 0 的区间的图见图 3、4。

由图 3、4 可知:百分比日间变化为 0 时,说明不论百分比夜间跳跃如何变化,都影响不到百分比日间变化,则百分比夜间跳跃与百分比日间变化没有

表 4 数据百分比化处理

次数	夜间跳跃	日间变化	夜间跳跃/收盘价	日间变化/收盘价	百分比夜间跳跃	百分比日间变化
1	-38.91	-152.61	-0.002 77	-0.010 870	-0.277 02	-1.086 510
2	-106.88	-174.66	-0.007 76	-0.012 690	-0.776 50	-1.268 930
3	-45.57	-6.75	-0.003 32	-0.000 490	-0.332 34	-0.049 230
4	-13.88	-123.30	-0.001 02	-0.009 080	-0.102 25	-0.908 300
5	-1.44	-42.03	-0.000 11	-0.003 110	-0.010 64	-0.310 610
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1 831	-61.74	77.37	-0.002 94	0.003 682	-0.293 85	0.368 242
1 832	-39.68	-346.57	-0.001 92	-0.016 800	-0.192 39	-1.680 390
1 833	-18.22	70.99	-0.000 88	0.003 433	-0.088 12	0.343 326

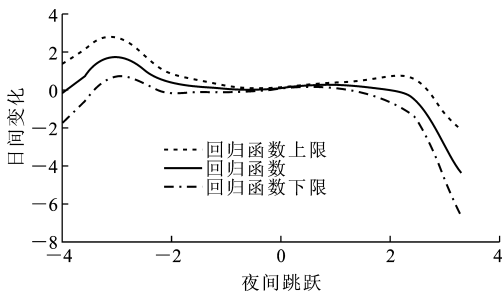


图 3 上证综合指数渐进正态分布的置信带

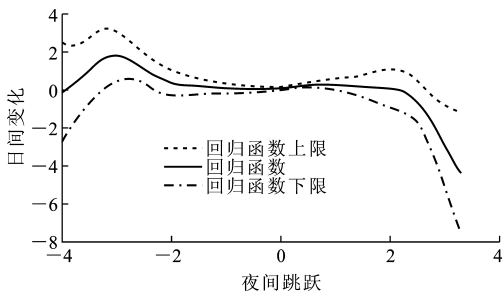


图 4 上证综合指数 Bootstrap 的置信带

显著关系。在置信带图中,作百分比日间变化为 0 的直线,凡置信带包含直线百分比日间变化为 0 的区间,百分比夜间跳跃与百分比日间变化没有显著的关系;反之,置信带没有包含直线百分比日间变化为 0 的区间,百分比夜间跳跃与百分比日间变化有显著关系^[14-15]。

由图 3、4 还可知:在区间 $(-3.5, -2.5)$ 、 $(0.2, 1)$ 和区间 $(-3.2, -2.5)$ 、 $(0, 1)$ 内日间变化与夜间跳跃有显著关系,其交集为 $(-3.2, -2.5)$ 、 $(0.2, 1)$ 。依此,可以观察每日的夜间跳跃数据,当其处于此区间中时,根据此区间内日间变化与夜间跳跃的关系,决定买入/卖出上证综合指数基金或等价的股票组合,以达到获利目的。当夜间跳跃不在此区间时,不存在基于夜间跳跃的可获利交易策略。

三、结 语

运用非参数回归分析方法,对全球 7 个股票市场夜间跳跃和日间指数变化之间关系进行研究。夜间跳跃和日间指数变化存在相关关系,在不同的股票市场,其相关关系不同。通过渐进正态分布和 Bootstrap 置信带分析,可获得夜间跳跃和日间变化拟合度较高的百分比夜间跳跃区间,依据区间内夜间跳跃和日间变化的关系,预测股市指数的变化趋势,从而制定基于夜间跳跃的股市可获利交易策略。由此说明,技术分析对指导股票市场的投资是有效的。从理论上讲,本文研究方法也适用于其他股市和个股的投资。

参考文献:

[1] Fama E F. Efficient capital markets: a review of theory and empirical work[J]. The Journal of Finance, 1970, 25 (2): 383-417.

[2] Fama E F, French K. The cross-section of expected stock returns[J]. The Journal of Finance, 1992, 47 (2): 427-465.

[3] Brock W, Lakonishok J, LeBaron B. Simple technical trading rules and the stochastic properties of stock returns[J]. The Journal of Finance, 1992, 47 (5): 1731-1764.

[4] Bessembinder H, Chan K. The profitability of technical trading rules in the Asian stock markets [J]. Pacific-Basin Finance Journal, 1995, 3 (2): 257-284.

[5] Hudson R, Dempsey M, Keasey K. A note on the weak from efficiency of capital markets: the application of simple technical trading rules to UK stock prices from 1935 to 1994 [J]. Journal of Banking and Finance, 1996, 20 (6): 1121-1132.

- [6] Parisi F, Vasquez A. Simple technical trading rules of stock returns; evidence from 1987 to 1998 in Chile[J]. Emerging Markets Review, 2000, 1(2): 152-164.
- [7] Pagan A, Ullah A. Nonparametric econometrics [M]. Cambridge: Cambridge University Press, 1999.
- [8] Hayfield T, Racine J S. Nonparametric econometrics; the np package[J]. Journal of Statistical Software 2008, 27 (5): 1-32.
- [9] 范依梅. “有效市场假说”综述[J]. 集团经济研究, 2007(9Z): 266.
- [10] 戴丽娜. 非参数技术理论及应用综述[J]. 新乡学院学报: 自然科学版, 2008; 25(3): 14-17.
- [11] 张锐力, 董彦峰. 有效市场假设及其检验方法: 文献综述[J]. 现代营销: 学苑版, 2011(8): 84-85.
- [12] 张晨曦, 姜姣娇. 有效市场假说的理论综述[J]. 商业文化, 2011(4): 372-373.
- [13] 邓 杰, 唐国兴. 中国股票市场技术交易规则有效性的实证研究[J]. 华东经济管理, 2009, 23(5): 135-140.
- [14] 纪华东. 中国股票市场有效性研究综述[J]. 市场周刊: 理论研究, 2011(1): 85-86.
- [15] 张亦春, 周颖刚. 中国股市弱式有效研究综述[J]. 当代财经, 2001(8): 30-32.

Nonparametric research on the relationship between intraday movement and overnight gap of stock indices

ZHAO Zhao

(Antai School of Economics and Management, Shanghai Jiaotong University,
Shanghai 200240, China)

Abstract: This paper adopts the method of nonparametric statistic analysis to conduct a research on the relationship between the overnight gap and intraday movement of stock indices. The daily trading data of 7 stock markets around the world were processed and analyzed through basic statistic analysis, scattering diagram and the function figure based on R software. The analysis result shows that there is a relationship between overnight gap and intraday movement, and the relation is different with the stock market. Taking Shanghai composite index analysis data as an example, it is found that the percentage gap area as the gap and change fits better by asymptotic and Bootstrap confidence band analysis. And a profitable trading strategy based on stocks overnight gap was proposed according to the relationship between gap and change in the percentage gap interval. This also demonstrates that technical analysis is efficient to guide the stocks investment.

Key words: nonparametric statistics; stock market index; confidence band; profitable investment strategy