中证100(000903.SH)和中证500(000905.SH)周频数据

基准对比：沪深300指数数据

参数估计期：2007-2009年（训练模型）

样本外测试期：2010-2012年（验证策略）

**因子计算流程**

1.动量因子（3月RSI差值）：

1.分别计算中证100/500的14日RSI

2.取最近63个交易日（3个月）的均值

3.计算差值：RSI\_diff=RSI\_500-RSI\_100

2.反转因子（PSY差值）：

6月PSY：窗口：126个交易日

计算：6个月上涨天数/总天数×100

12月PSY：窗口：252个交易日

差值：PSY\_500-PSY\_100

3.成交量因子（1月成交比）：窗口：21个交易日

Σ(中证500成交额)/Σ(中证100成交额)

**建模流程**

Logit模型构建：

**估算概率**p\_i

**邻近样本聚合法**

* 1. 计算当前样本点与历史所有样本点在**所有标准化自变量**上的**欧式距离**。
  2. 选取距离最近的N\_i个样本 (N\_i取10-15，权衡精度与代表性)。
  3. 统计这N\_i个样本中，**小盘股实际跑赢的次数(**n\_i**)**。
  4. 估算当前样本点的概率：p\_i=n\_i/N\_i。

数据标准化（报告中没说明）：z-score标准化

欧式距离

3.信号生成规则：

概率区间操作头寸

P>65%多小盘/空大盘100%

P<35%多大盘/空小盘100%

其他空仓0%

回测验证

1.测试框架：

for每月末:

计算当前因子值→预测概率→生成信号→计算下月收益

2.绩效指标：

核心指标：年化收益、夏普比率、最大回撤

辅助指标：胜率、盈亏比、超额收益

基准对比：沪深300指数

六、关键参数配置

参数值说明

RSI窗口14日标准设置

PSY窗口6/12月中期+长期

成交比窗口21日1个月

最近邻数15平衡偏差方差