

# 股票特征与未来收益率的检验：Fama-MacBeth截面回归

## Fama-MacBeth截面回归流程介绍

在前面的统计检验设计中，有一个缺点在于检验某变量对于未来收益率影响时，不能很好地控制其他潜在变量的影响。为了克服以上缺陷，Fama-MacBeth截面回归分析是一个很好的统计工具。

沿用前面的设定，假设解释变量用Y（例如未来一个月的收益率）表示，核心解释变量用X（例如企业流通市值）表示，需要控制的协变量为X1，X2...（例如企业价值和企业换手率）。Fama-MacBeth截面回归分析分为两步。第一，对每一个时间t(t=1,2,...,T)，基于横截面样本进行回归分析，记录下每一期横截面回归的回归系数、样本量、调整后R2。第二，Fama-MacBeth最后的回归系数、样本量、调整后R2为T期取平均，系统的统计量为对T期的以上回归系数的时间序列进行统计检验，检验各项回归系数的时间序列是否显著异于零。需要注意的是，这里的统计检验量也需要使用前面介绍的Newey-West调整。

## Fama-MacBeth截面回归举例说明

依然沿用前面的案例，需要研究中国股票市场2000年1月到2021年6月（共计258个月）样本中，控制企业价值和换手率的条件下企业总市值变量是否能够用于预测未来一个月的收益率。总体的回归设定如方程所示：

$$R_{i,t+1} = \hat{\alpha} + \hat{\beta}^{ME}(X\_ME_{i,t}) + \hat{\beta}^{EP}(X\_EP_{i,t}) + \hat{\beta}^{AT}(X\_AT_{i,t}) + \hat{\epsilon}_{i,t}$$

其中，左边是被解释变量，代表第i值股票第t+1个月的月度收益率。X\_ME是核心解释变量，代表第i只股票第t个月月底的A股总市值；X\_EP是控制变量，代表第i只股票第t个月月底的市盈率倒数；X\_AT是控制变量，代表第i只股票第t个月月底异常换手率；epsilon是随机扰动项。在Fama-MacBeth截面回归中，最需要关注的系数是betaME大小和显著性水平，这个系数代表企业市值的变动会给企业未来一个月收益率带来多大影响。

需要注意的是，进行以上回归时，为了剔除异常值对回归系数的影响，一般会对解释变量进行缩尾处理，而对于被解释变量一般不进行缩微处理。此外，由于股票市场横截面的市值变量分布会呈现出长尾分布，这种非标准正态分布也可能对估计系数造成较大影响，一般会对变量取对数处理。

此外，在对方程进行估计时，一般会采用不同的方程设定，给出多个设定下模型估计的结果。例如：

1. 单独把企业总市值回归在未来一个月的收益率上，不添加任何控制变量
2. 把企业总市值回归在未来一个月的收益率上，只添加市盈率倒数作为控制变量
3. 把企业总市值回归在未来一个月的收益率上，只添加异常换手率作为控制变量
4. 按照模型设定，把企业总市值回归在未来一个月的收益率上，同时添加市盈率倒数和异常换手率作为控制变量。

## 使用Fama-MacBeth截面回归分析的具体流程

1. 对t=1到t=T，每个月横截面样本用方程设定进行估计，共计估计258次，记录下每一期横截面回归的回归系数，样本量，调整后R2。
2. Fama-MacBeth最后的回归系数、样本量、调整后R2为T期取平均，系统的统计量为对T期258期的以上回归系数的时间序列进行统计检验，检验各项回归系数的时间序列是否显著异于0

## Fama-MacBeth截面回归总结

然而与前面的方法相比，Fama-MacBeth截面回归分析需要假设变量对于未来收益率的影响是线性的，也就是它只能检验两者是否存在线性关系，这也是该方法的一个局限性。除了上述手动实现的方法之外，还可以调用python的Linermodels库，直接获得Fama-MacBeth的估计结果。

```
linearmodels.panel.model.FamaMacBeth(y,x).fit(cov_type='kernel',bandwidth=4)
```

y: 被解释变量

X: 解释变量和控制变量，如果数据中没有常数项，需要手动添加。

cov\_type: 使用参数'kernel'来获得Newey-West调整后的标准误统计量。

bandwidth:指定Newey-West调整滞后的期数，可以根据Newey和West提供的公式 $L=[4x(T/100)^{(2/9)}]$ 计算得到。