1. 本文目的

本文旨在使用中国上海证券交易所的数据，检验Fama-French的三因子模型有效性水平。本文所选取的数据处理方法，第一部分皆尽量复刻Fama(1996)中对于美股市场检验的方法。第二部分，则增加了一些作者本人的修改，以更好的验证三因子模型在中国上证股票中的有效性水平。

1. 数据选取和处理方法

考虑到2000年前中国股市市场还过于不完善，17年开始掀起了一阵“价值股”的行情的“特殊时期数据”。因此，本文先选择检验了2000-2016年上海证券交易所交易的股市月度数据进行处理。数据来自于CSMAR数据库。

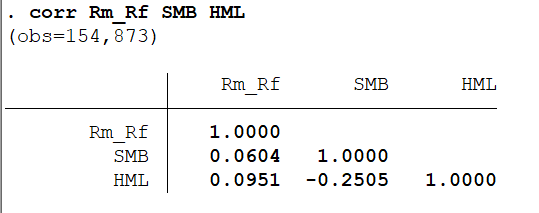
本文中所选取的市场收益率来源于CSMAR所公布的市场月度收益率数据。在Fama(1996)的论文中无风险利率选择美联储1月期国库券的利率。本文中所选取的无风险利率水平，考虑到在早期，中国还没有发行国债，因此选用的是人民银行公布的三个月定期存款利率折算的连续复利月收益率来衡量无风险收益率。

（由于在06年，中国股市基本完成了股权分置改革；08年由于金融危机影响，中国股票经历了股灾，基本到09年危机才基本消除。因此本文除了对2000-2016年整体的数据进行处理分析，还会将时间按照06年前，06-09年，10年往后进行分类，做比较分析。）

关于SMB（流通市值因子）和HML（账面价值比因子）的处理。对于流通市值(ME）分类：本文中，模仿Fama，我将股票按照市值进行排序，根据中位数进行分割。小市值分类为“S”，大市值分类为“B”。对于账面价值比分类(BE/ME)：将股市按照账面价值和流通市值的比，进行排序。高的30%为高账面市值比“H”，低的30%为低账面市值比“L”，其余为中等账面市值比“M”。通过高低组之间的加权回报率（权重为流通市值），计算溢价因子。

表[1.1]为各因子之间相关系数矩阵表。可以看出，因子之间基本没有相关性，基本不存在多重共线性的问题。

表[1.1]：



表[1.1]各因子间相关系数矩阵表

在对于各模型进行检验中，模仿Fama的方案，我将股市按照市值分成5组（分位数），按照账面价值比再分为5组，交叉得到25组数据。

表[1.2]—表[1.5]是对于25组的相关指标分组描述性统计。

表[1.2]:



表[1.2]2000-2016年分组加权回报率的描述性统计表

表[1.3]:



表[1.3]2000-2005年分组加权回报率的描述性统计表

表[1.4]:



表[1.4]2006-2009年分组加权回报率的描述性统计表

表[1.5]:

表[1.5]2010-2016年分组加权回报率的描述性统计表

1. 单因子模型检验

单因子模型即CAPM模型。对于单因子模型进行检验：

*R*(*t*)−*Rf*(*t*)=*a*+*b*(*RM*(*t*)−*Rf*(*t*))+*e*(*t*) (2.1)



将式(1.1)进行异方差稳健的OLS回归。统计量披露采取和Fama（1996）一文中对于三因子模型相同的数据披露表述。

1. 双因子模型检验

对市值因子和账面价值比因子进行处理。

*R*(*t*)−*Rf*(*t*)=*a*+*sSMB*(*t*)+*hHML*(*t*)+*e*(*t*)



四．三因子模型检验

对Fama——French三因子模型进行处理。

*R*(*t*)−*Rf*(*t*)=*a*+*b*(*RM*(*t*)−*Rf*(*t*))+*sSMB*(*t*)+*hHML*(*t*)+*e*(*t*)



五．结论

1. CAPM模型总体来看基本可以在0.01情况下拒绝原假设（|t|>2.58,双边检验）。R-square基本在0.4左右，可以较好解释中国上证股市的风险溢价。但应该还存在未能被纳入解释变量的其他因子。

2. 双因子模型总体来看并无太多解释力度。

3. Fama-French三因子模型基本可以在0.01情况下拒绝原假设（|t|>2.58,双边检验）。R-square基本在0.45左右，可以一定程度上解释中国上证股市的风险溢价。但应该还存在未能被纳入解释变量的其他因子。

4. 比对三因子模型和CAPM模型，在上证股市中，R-square并无在三因子模型中显著提升。这意味着市值因子和账面价值比因子的解释力度不够强。