

---

## 一种夏普风格权重的置信区间的估计方法

本文主要内容是：

1. 风格分析建模：夏普风格权重指的是多因子线性回归模型的自变量的系数，对模型适当增加约束条件来解决多重共线性问题。
2. 夏普权重的置信区间：引入衡量统计模型拟合优良性的一个指标。
3. 实证结果：用了一只基金 6 年的数据来进行验证，并选取了 4 个市场指数来说明，夏普风格分析中市场指数的选择是十分重要的

首先介绍一个概念，基金投资风格——指将基金资产在不同的股票之间进行配置的投资策略或者计划。最早是由威廉·夏普在 1988 年提出并使用的，它是一种有限制的线性模型，即使用市场指数的加权组合来试图尽可能地复制投资组合历史收益率的基本模式；而这个加权系数，就叫做“夏普风格权重”。它的大小在一定程度上表示不同的因子、因素对资产组合收益率的解释程度。通过风格分析可以匹配不同的投资者的风险偏好特征，降低投资者的投资选择成本，便于准确评价基金的效益等作用。

本文将对一篇 1997 年发表的文章做一个简要的介绍，切入点是夏普风格权重的置信区间的估计。这里的夏普风格权重指的是文章用到的线性回归模型的自变量的系数。文章发表前，还没有一个明确的测量权重的置信区间的方法，但是作者估计出这些权重的置信区间，并且用蒙特卡洛模拟的方法来验证估计的有效性，这就是此文最大的贡献。

### 一. 风格分析建模

风格分析简单来说就是一种估计资产类别的因子模型，它的基本形式是：

$$R = b_1F_1 + b_2F_2 + b_3F_3 + \dots + b_nF_n + e \quad (1)$$

这其中， $R$  表示的是投资组合收益的时间序列； $F$  代表的是每一个因子收益的时间序列，那么，显而易见的是，每一个  $b$  代表的就是投资组合对每一个因子的敏感程度。假如没有其他限制的话，我们就定义了一个多元线性模型，它的各个自变量是相互独立的。在风格分析中，我们把证券市场指数定义为各个因子用来表示各种资产类别或者投资的方式。

对于传统的多元回归分析来说，由于它的自变量的系数是没有限制的，所以对投资者来说并不能带来直观的感受，另外，传统的多元线性模型还会带来多重共线性的问题-----如果两个自变量具有高度的相关性，比如两个指数代表着投资于同一资产类别的不同方法，那么这两个指数的相关性就会很高，会对模型的结果带来较大的误差。

基于以上的问题，本文对回归系数的值做了两个限制。第一个就是所有的系数取值都限定在 0 到 1 之间；第二个限制就是所有的系数之和为 1。

如果加上这样的限定，每个因子的系数就可以被认为是因子对投资组合的贡献权重大小了。(那具体是如何操作的呢?)；通过加上这两个限制条件，还可以在在一定程度上缓解多重共线性的问题，因为它减少了由于本应该独立的自变量之间的高度相关性所导致的系数的“爆炸”。一旦夏普指数被定义好了之后，我们就可以用目标投资组合的收益的历史数据与夏普指数的历史数据建立多元线性模型， $R^2$  统计量被用来衡量模型的拟合优度。

对于风格分析问题来说，重要的是要保证市场指数之间不能有线性的相关关系，即不能存在多重共线性的问题。刚才提到的对模型适当增加约束条件是一种解决多重共线性的方法，另外，还可以用的方法有：删除一个或者几个具有共线性的变量、增加数据，主成分回归等。但是这几种方法在风格分析中都不是太实用，故采用的是增加模型限制条件的方法。目前一些投资者其实没有发现这个问题，还是选着相关性很高的市场指数来进行分析。他们为了更详尽的描述投资组合的收益情况，不断增加指数的数量和相似性，这样造成的结果就是导致模型的共线性问题变得越来越严重。

## 二. 夏普权重的置信区间

前面提到的夏普权重的置信区间的估计没有一个明确的方法，但是，从直观来说，特定的市场指数的夏普权重的置信区间随着风格分析的标准误差的增加而增加；随着市场指数的不断增多而减小；还随着不同指数之间的“独立性”的增强而减小。所以，为了不断缩小置信区间的范围，我们就需要在市场指数的个数和各个指数之间的独立性两者进行权衡，这时我们就可以引入 AIC 准则，它是衡量统计模型拟合优良性的一个指标，建立在熵的概念上，它的表达式为：

$$AIC = 2K - 2\ln(L) \quad (2)$$

$K$  是模型的参数个数， $L$  是似然函数。我们通常选择 AIC 的值最小的模型。因为随着  $K$  的数量的增加，模型的复杂度也会随之提高，这时似然函数  $L$  也会增加，就会导致 AIC 的值减小。但当  $K$  过大时，似然函数的增速减慢，就会使得 AIC 的值变大。我们选择最小的 AIC 值，既提高了模型的拟合优度，而且加入了惩罚项，使参数尽可能的少，有效防止了过拟合的现象。

本文中，估计权重的标准差根据下面这个公式计算：

$$\frac{\sigma_a}{\sigma_{\beta_i} \times \sqrt{n-k-1}} \quad (3)$$

其中， $\sigma_a$  是风格分析的标准误，这里需要说明的是，标准误与标准差是不同的概念。 $\sigma_{\beta_i}$  是无法解释的夏普指数波动， $n$  表示的数据的数量， $k$  表示非 0 权重的指数的个数。

为了验证上述方法的有效性，文章运用了蒙特卡洛方法来模拟系数，构建一个已知的投资组合和市场指数(相当于权重已经知道)，设定一个误差初值，

由蒙特卡洛方法模拟出一系列误差项加入到组合的收益时间序列中。重复 10000 次，得到的权重近似作真实的权重。关于蒙特卡洛的模拟方法，可参考其他解释更加详细的文章。

### 三. 实证结果

文章中的一个实际的例子是用了 **Reich&Tang** 基金 6 年的数据来进行验证，并选取了 4 个市场指数：90 天国债、雷曼兄弟综合指数、罗素 3000 指数、MSCI EAFE 指数 ETF，当然其中有些已经不存在。

**Table 1. Unexplained Sharpe Style Index Volatility**  
(percent per month)

| Index                 | Unexplained Sharpe Style Index Volatility |
|-----------------------|---|
| 90-day Treasury bills | 1.01%                                     |
| Lehman Aggregate      | 0.85                                      |
| Russell 3000          | 2.60                                      |
| MSCI EAFE             | 4.59                                      |

作者将这四个指数分别组合，求出不同组合的拟合优度。

**Table 2. Style Weights and Standard Deviation of Style Weights**  
(standard deviations in parentheses)

|                   | Lehman Aggregate | 90-Day Treasury Bills | MSCI EAFE | Domestic Equity Indexes |        |        |        |        |        |
|-------------------|------------------|-----------------------|-----------|-------------------------|--------|--------|--------|--------|--------|
|                   |                  |                       |           | 1                       | 2      | 3      | 4      | 5      | 6      |
| Index Group 1     | 0.0%             | 10.0%                 | 2.2%      | 87.8%                   |        |        |        |        |        |
| ( $R^2 = 0.868$ ) | (20.1)           | (16.9)                | (3.7)     | (6.5)                   |        |        |        |        |        |
| Index Group 2     |                  | 10.0                  | 2.2       | 87.8                    |        |        |        |        |        |
| ( $R^2 = 0.868$ ) |                  | (5.2)                 | (3.6)     | (5.7)                   |        |        |        |        |        |
| Index Group 3     |                  | 11.6                  | 2.9       | 70.0                    | 15.5%  |        |        |        |        |
| ( $R^2 = 0.872$ ) |                  | (5.3)                 | (3.6)     | (10.3)                  | (7.3)  |        |        |        |        |
| Index Group 4     |                  | 9.9                   | 2.7       | 41.5                    | 30.5   | 10.0%  | 5.4%   |        |        |
| ( $R^2 = 0.873$ ) |                  | (6.0)                 | (3.7)     | (14.4)                  | (12.6) | (14.7) | (12.3) |        |        |
| Index Group 5     |                  | 10.1                  | 2.4       | 29.2                    | 16.3   | 24.2   | 17.8   | 0.0%   | 0.0%   |
| ( $R^2 = 0.872$ ) |                  | (6.0)                 | (3.8)     | (16.2)                  | (13.3) | (20.7) | (15.6) | (14.2) | (14.4) |

Note: Domestic Equity Index Number Assignments

- Index Group 1, 2 1 = Russell 3000
- Index Group 3 1 = Russell 1000  
2 = Russell 2000
- Index Group 4 1 = Russell 1000 Value  
2 = Russell 1000 Growth  
3 = Russell 2000 Value  
4 = Russell 2000 Growth
- Index Group 5 1 = S&P/BARRA Value  
2 = S&P/BARRA Growth  
3 = Sharpe/BARRA Medium Value  
4 = Sharpe/BARRA Medium Growth  
5 = Sharpe/BARRA Small Value  
6 = Sharpe/BARRA Small Growth

通过这个例子，作者想要说明的是：夏普风格分析中市场指数的选择是十分重要的，并且，对投资组合的先验信息有一定的了解也是必要的，因为这样可以使我们更好的推断后验信息。

这种方法的应用是非常广泛的，比如，在基金分类中，可以根据夏普权重的相似性将某些基金分为一类，可以认为是一种投资风格的基金；另外，从投资这的角度来说，由于权重是随着时间不断变化的，投资者可以把这些夏

普权重的变化当作是基金经理改变其投资策略的信号,有助于投资者根据自身的风险偏好调整其的投资选择。当然,这个方法还是有其一定的局限性的。就是没有给出一个计算夏普权重置信区间的明确方法,只是在处理的时候期待它服从一个正态分布,从而可以使用传统的 T 统计量来进行检验。

#### 四. 私募云通风格分析

在分析股票型基金时,由于持仓信息无法实时获取,客户希望仅通过基于收益的方法就能够对可获知净值变动的股票型基金进行较为及时的分析,即求其对每个因子的平均风险暴露。这种做法的假设前提是:投资组合的风格是相对固定的,相对于组合中对于个股权重的变化,对于每个因子(特别是风格因子)的风险暴露能够保持一定的连续性。因此,私募云通以 6 个月作为最小回归单位,新成立 6 个月内的基金不予进行回归。

为了提高外部评价的准确性,私募云通采用多个因子模型,以便客户对私募证券投资基金(偏股型或股票型)进行业绩归因,具有多个模型比对参考的优点。目前所采用的 3 个因子分别是市场因子、规模因子和价值因子。私募云通新版可视化终端 FOF EASY【2.0】即将上线,您可以看到:

表 1.三因子回归结果(样例)

| 因子名称             | 因子暴露    | 标准误    | t 值     | 置信上限   | 置信下限    |
|------------------|---------|--------|---------|--------|---------|
| 规模因子(*)          | -0.0847 | 0.1797 | -0.4712 | 0.2871 | -0.4565 |
| 价值因子             | -0.3735 | 0.2100 | -1.7787 | 0.0609 | -0.8079 |
| 市场因子(***)        | 0.3258  | 0.0783 | 4.1638  | 0.4877 | 0.1639  |
| F 值              | 8.4227  |        |         |        |         |
| R <sup>2</sup> 值 | 52.34%  |        |         |        |         |

注:\*, \*\*, \*\*\*分别代表在 90%, 95%, 99%的置信度下统计显著

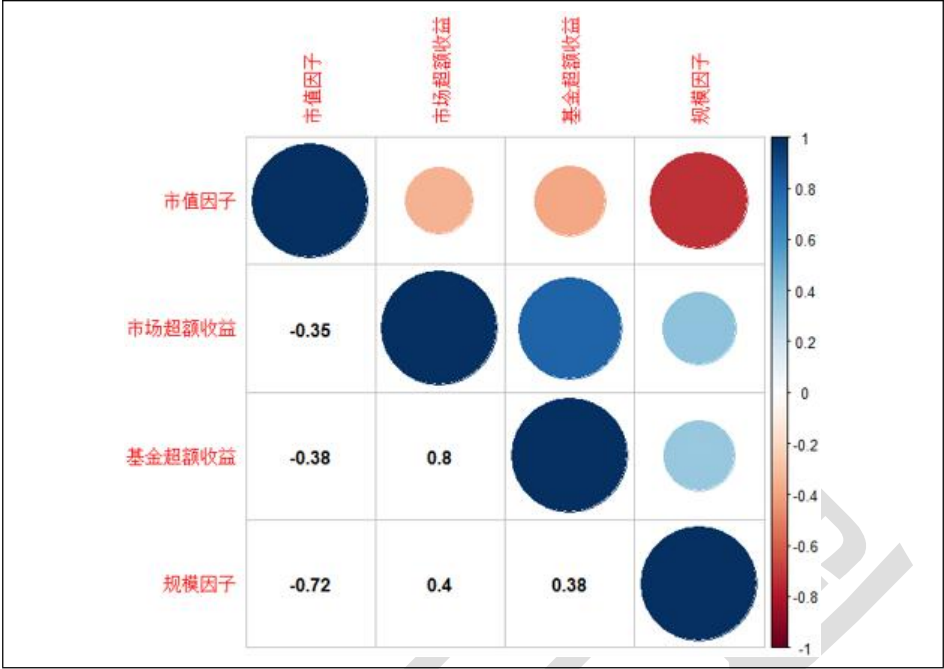


图 1.风格因子相关性强度图（样例）

将收益归因结果从高到低排序，正红负绿。

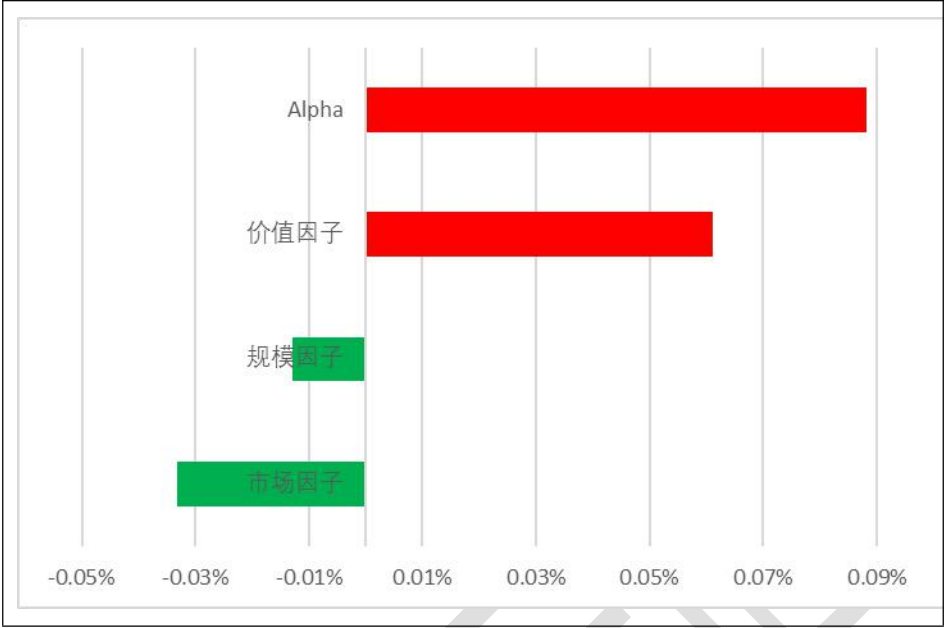


图 2.三因子模型收益归因结果（样例）

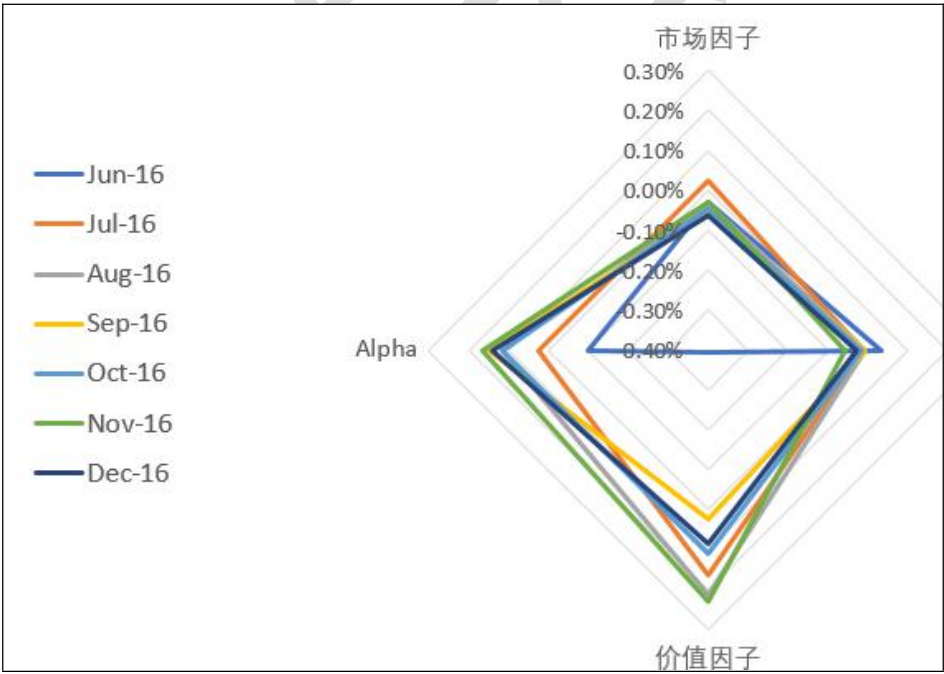


图 3.三因子模型收益归因结果（样例）

将风险归因结果从高到低排序，风险归因结果均为正值。

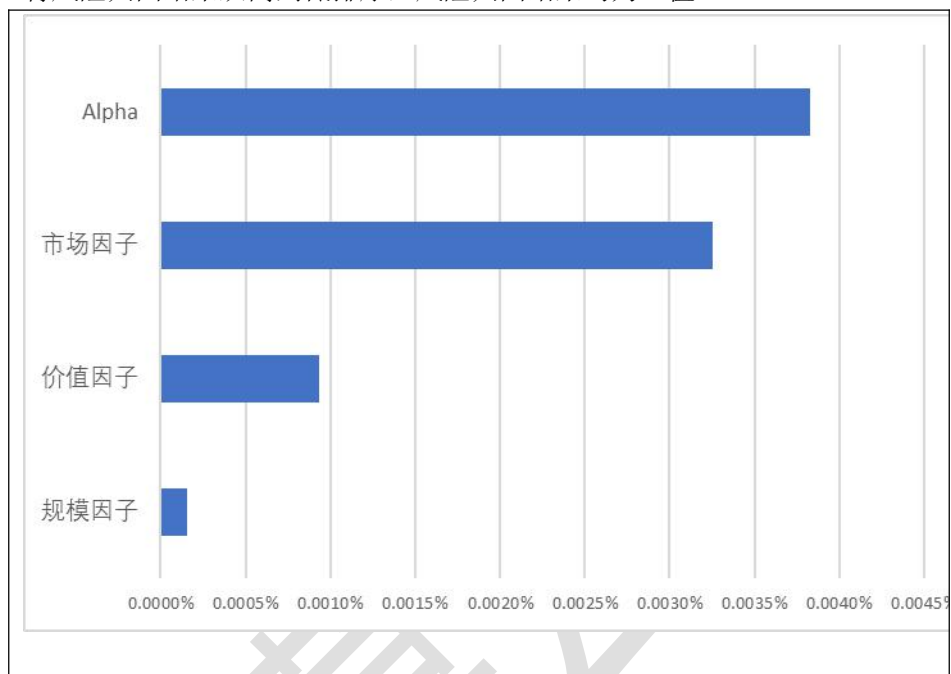
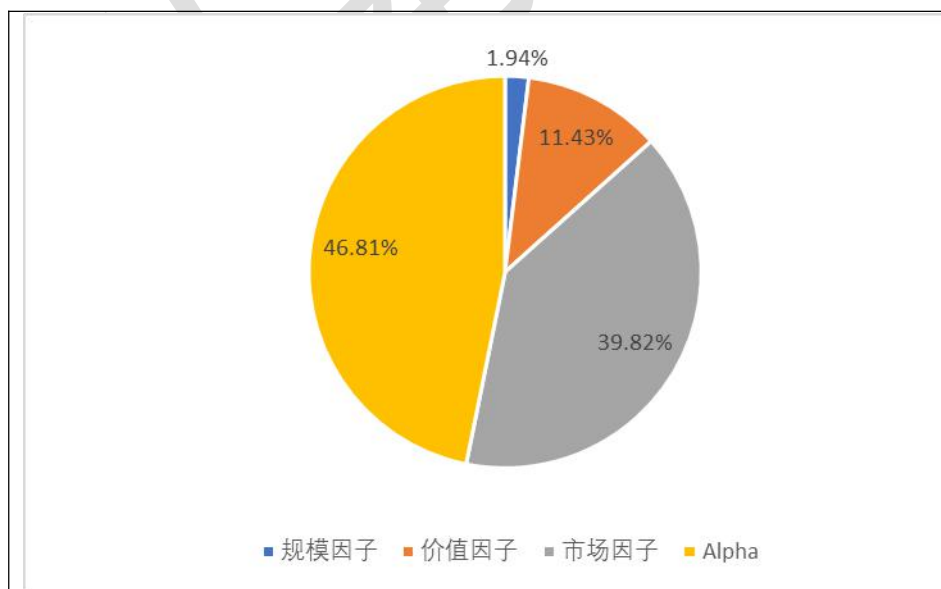


图 4.三因子模型风险归因结果（样例）



---

图 5.三因子模型风险归因结果（样例）  
等等多种收益、风险归因图像可视化展示，欢迎使用下方联系方式。

**参考文献：**

[1].Lobosco、DiBartolomeo: Approximating the Confidence Intervals for Sharpe Style Weights.

私募股权通