
Portfolio Construction and Analytics 读书笔记

目录	1
Contents	1
Contents	1
1 资产管理的介绍	1
2 随机变量、概率分布和重要的统计概念	1
3 常见的分布函数介绍	1
4 统计学模型	1
5 模型模拟	1
5.1 蒙特卡罗模拟	1
5.1.1 选择分布函数	1
5.1.2 理解蒙特卡罗模拟的输出	1
5.2 为什么采用蒙特卡罗模拟	2
5.2.1 多个输入变量和混合分布	2
5.2.2 合并相关	2
5.2.3 模型评估	2
5.2.4 模拟多少次?	3
5.2.5 随机数的生成	3

- 1 资产管理的介绍
- 2 随机变量、概率分布和重要的统计概念
- 3 常见的分布函数介绍
- 4 统计学模型
- 5 模型模拟

5.1 蒙特卡罗模拟

假设你现在有1000美元。你计划用这笔资金投资标普500指数。 C_0, C_1 分别表示期初和期末（一年以后）的资金。 $r_{0,1}$ 表示期间资金的回报率。则有：

$$C_1 = (1 + r_{0,1})C_0$$

在 $[t, t + 1]$ 区间内资金的回报率可以表示成：

$$r_t = \frac{P_{t+1} - P_t + D_t}{P_t}$$

P_t :标普500指数在 t 时刻的价格

D_t :在区间 $[t, t + 1]$ 内的标普500指数的股息

5.1.1 选择分布函数

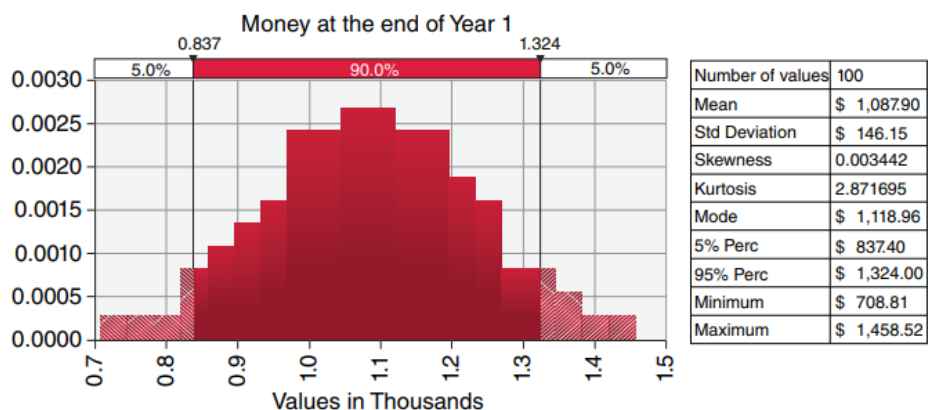
为了采用蒙特卡罗模拟我们的模型，我们首先要确定模拟数据的分布函数。这里有两种方法确定数据的分布函数。第一种是采用历史数据，第二种是随机产生一种给定分布的数据序列。例如，我们可以生成一组均值为 μ ，方差为 σ 的正态分布的数据。

5.1.2 理解蒙特卡罗模拟的输出

作为蒙特卡罗模拟的一个例子，我们模拟了标普指数一年以后的收益情况。我们生成了100个服从正态分布的随机数（均值为8.79%，方差为14.65%）图表中的直方图展示了未来一年的收益率分布：

根据统计学的知识我们可以知道，我们有95%的置信度认为一年后的收益率分布的区间：

$$\left(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}} \right) = (1058.9, 1116.9)$$



5-1 未来一年标普500收益率分布

5.2 为什么采用蒙特卡罗模拟

上面的例子给出了蒙特卡罗模拟的基本方法，我们采用上述方法，给出了一年以后收益率的分布情况。值得注意的是，如果标普500收益率不满足正态分布，那么我们的模拟就不准确了。下面的例子是一个更加复杂的模拟。

5.2.1 多个输入变量和混合分布

假设你为未来养老而计划做一笔投资，投资的本金为1000美元，投资的周期为30年。假设标普500指数服从均值为 μ ，方差为 σ 的正态分布。记期初的资金为 C_0 ，期末的资金为 C_{30} 。容易知道，期末的收益可以写成：

$$r_{0,t} = (1 + r_{0,1})(1 + r_{1,2}) \dots (1 + r_{t-1,t}) - 1$$

期末的收益取决于这30个正态分布。

5.2.2 合并相关

如果我们投资的是两类标的：国债和股票。这二者的收益率存在着负相关。那么我们又该如何模拟呢？不妨假设股票和国债收益率的相关系数为-0.2。国债收益率服从均值为4%，方差为7%的正态分布。我们采用蒙特卡罗进行了500次模拟，下图展示了30年后收益率的分布情况：

5.2.3 模型评估

如何评估模型的好坏呢？我们考虑下面两种不同的策略：

A：50%资金投资股票，50%资金投资债券

B：30%资金投资股票，70%资金投资债券

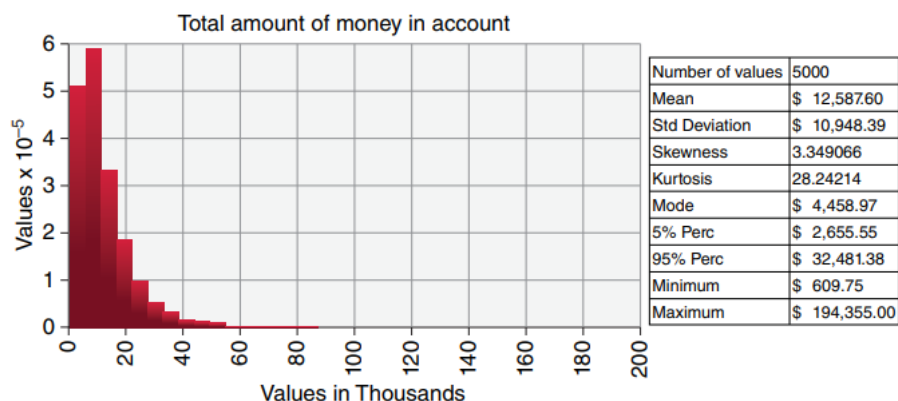
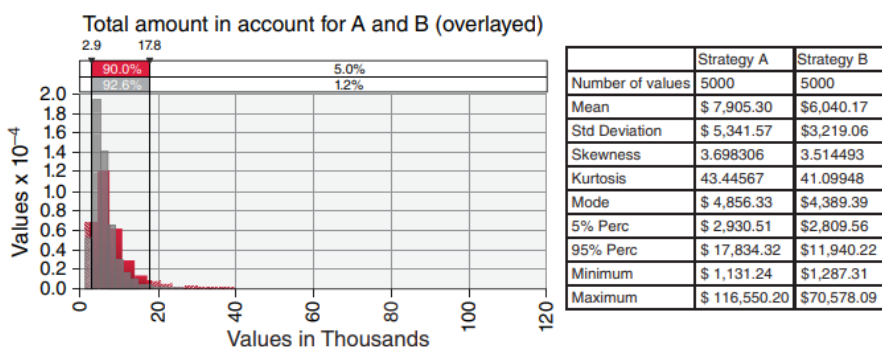


Exhibit 5.3 Output distribution for amount of capital after 30 years.

我们对投资组合A和投资组合B都进行了400次模拟。从模拟的结果我们可以看出，30年后，投资组合A的平均收益大于投资组合B。但是投资组合A的变异系数也大于投资组合B。这说明投资组合A的风险比B大。课本上的图标罗列了二者的均值、方差、峰度、偏度。从中我们可以直观地看到两个投资组合的差异。



5.2.4模拟多少次？

模拟的数据至少要大于等于30才具有统计学意义。在上述的例子中我们模拟了400个样本用于分析。一般来说提高样本的数量有助于我么能提高模拟的精度。

5.2.5随机数的生成

Excel、R、MATLAB 都可以很方便地生成随机数，可参考各软件的文档、百度等。