

常用描述性统计概念

1. 随机变量

如果一个随机变量 x 最多只能取可数的不同值，称为离散型随机变量。

如果一个随机变量 x 取值无法一一列出，可以取遍某个区间的任意数值，称为连续性随机变量。

2. 方差和标准差

方差和标准差是风险评估指标。

方差是标准差的平方数。

先求出平均值 R ，然后计算每个数 R_i 和平均数 R 的差的平方 $(R - R_i)^2$ ，将结果相加，最后得出标准差。

3. 分位数

3.1. 上分位数

将给定数从小到大排列，从最大值开始往最小值数，最大值为第一个数。

那么找到第三个数，就是上 3% 分位数。找到第 5 和第 6 个数，将两数求平均值，则为上 5.5% 分位数。

3.2. 下分位数

将给定数从小到大排列，从最小值开始往最大值数，最小值为第一个数。

那么找到第三个数，就是下 3% 分位数。找到第 5 和第 6 个数，将两数求平均值，则为下 5.5% 分位数。

4. 中位数

将给定数从小到大排列，从最小值开始往最大值数，最小值为第一个数。

取中间数即为中位数，如果给定的一组数是偶数个，则取中间两个数的平均数即为中位数。

中位数评价结果更为合理和切合实际。 中位数能够免疫极端值的影响，较好的反映投资策略的真实水平。

5. 相关系数

相关系数取值范围 $[-1, 1]$ 。

- -1 代表完全负相关
- 1 代表完全正相关
- 0 代表完全不相关
- $(-1, 0)$ 代表负相关
- $(0, 1)$ 代表正相关

在实际应用中，如果两只股票都上涨表示正相关，上涨幅度相同，则表示完全正相关。