

AS 做市模型深度因子研究

一. 背景

AS 模型的核心思路是**位移**和**间距**，位移说的是围绕做市的价格与 bid, ask 均值的差值，间距指的是挂单的价格与这个做市的价格的价差。其中间距中引入了深度因子，当这个因子大的时候，价差会缩小，这个时候挂单的位置距离做市价格会靠拢，变得容易成交，当这个因子小的时候，价差会变大，挂单成交概率降低。然而，文中并没有给出这个因子的计算方法。本文的目的在于构建一个简单的模型，来实现这个因子的计算。



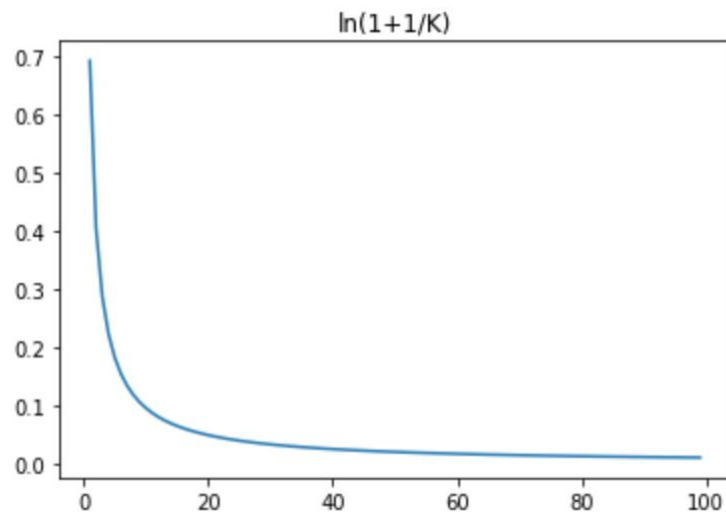
二. 模型构建

2.1 K 的取值范围

首先看公式涉及到深度因子的部分

$$y = \ln \left(1 + \frac{1}{K} \right), \text{ 其中 } K \text{ 为深度因子}$$

绘制函数图像，如下图



可以看出当 K 的值取值大的时候，y 会趋近于 0。而 K 的值在 (0, 1) 之间的时

候，y 的值变化大，在该区间，K 的灵敏度非常高。这里 K 的区间暂定为

[0.5, 1]，具体范围需要根据测试结果来定。

2.2 K 的算法

2.2.1 准备工作

这里首先定义

a_bid, 买盘前 20 档的挂单的币的个数

a_ask, 卖盘前 20 档的挂单的币的个数

v_bid, 买盘前 20 档的挂单的币的金额

v_ask, 卖盘前 20 档的挂单的币的金额

a_24, 24 小时成交的币的个数

v_24, 24 小时成交的币的金额

$\alpha = a/v$, 显然这里一共有四种情况，这里首先考虑

(1) 公式

$$\alpha_{bid} = a_{bid}/a_{24}$$

$$\alpha_ask = a_ask/a_24$$

(2) 币种

avax, etc, bnb, sand, egld, gala, fil, mana, near, doge

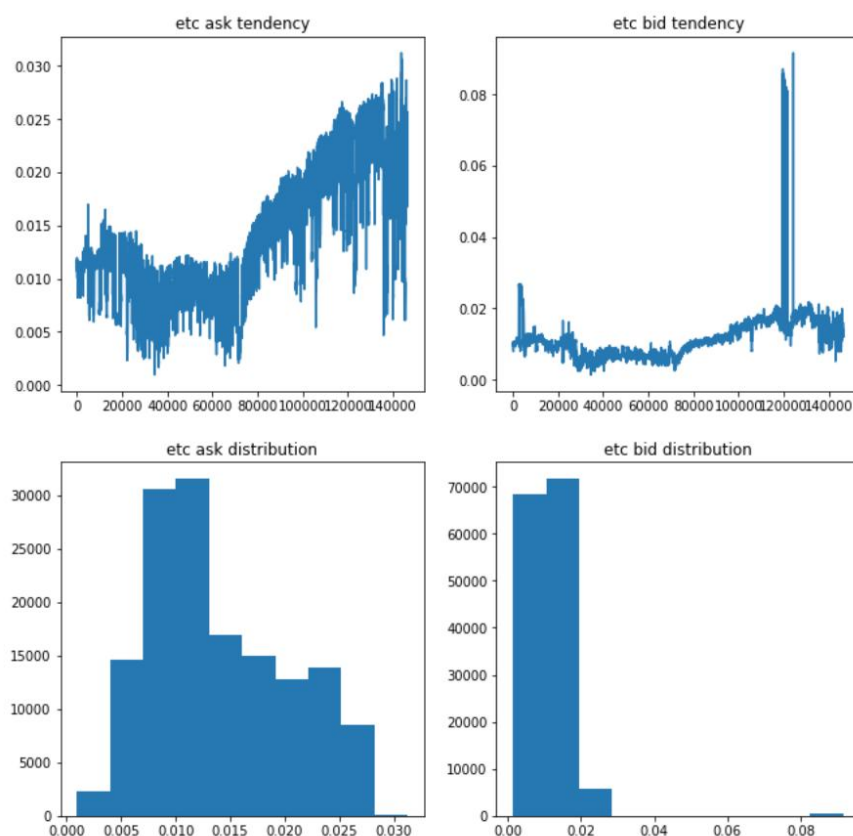
币种为随机选取的 10 种主流币种。

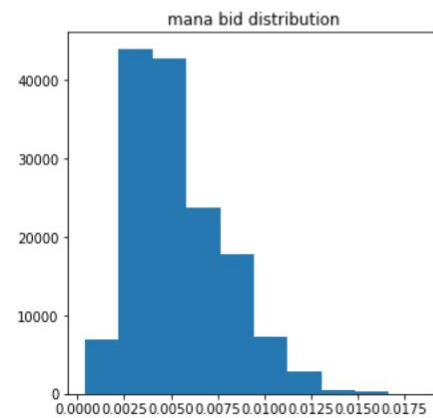
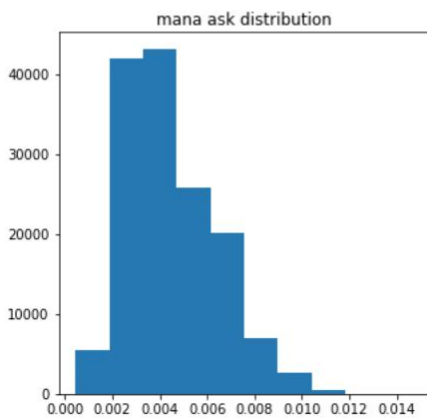
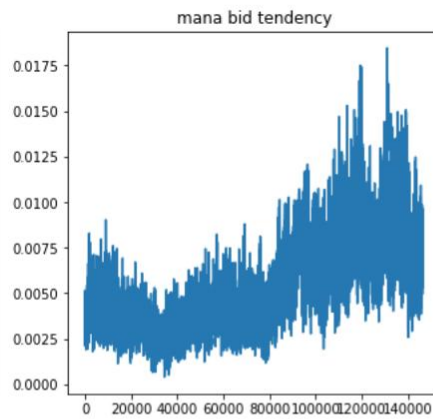
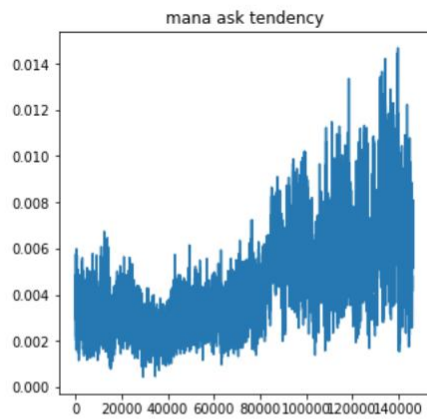
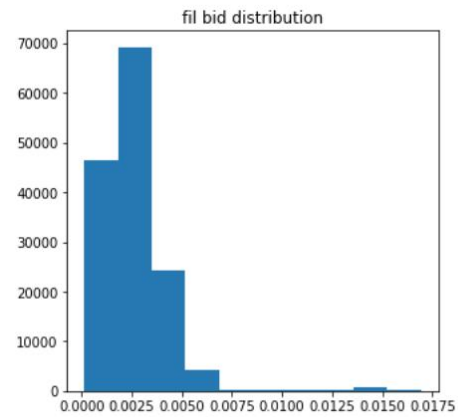
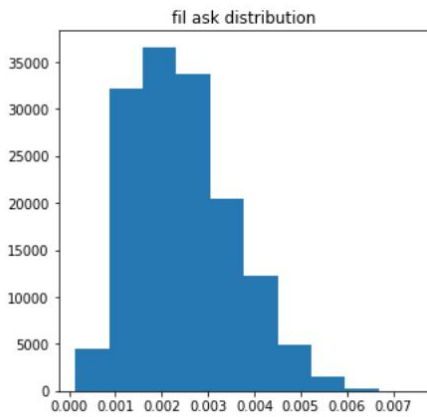
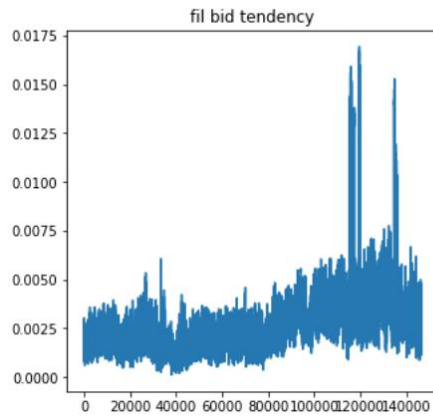
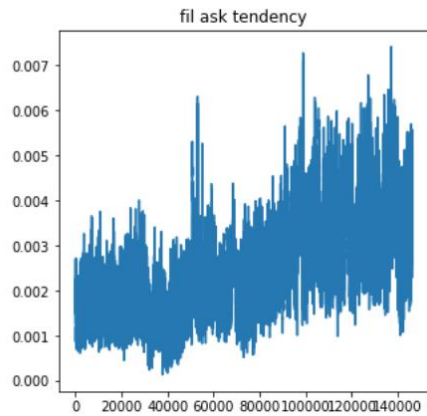
2.2.2 数据获取

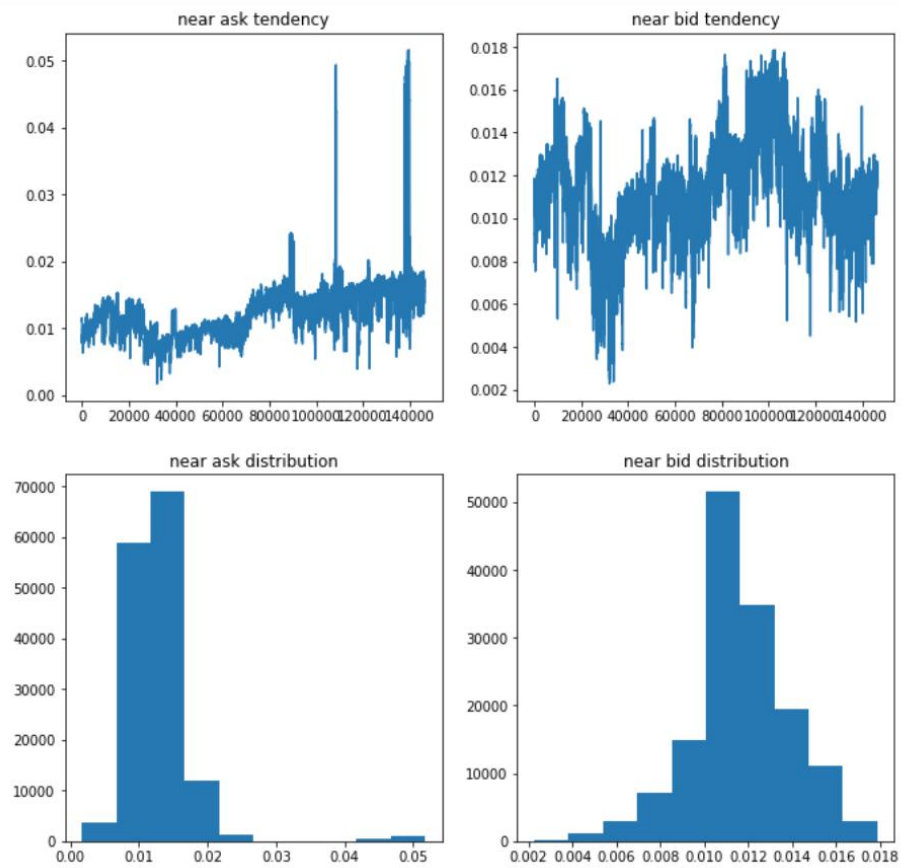
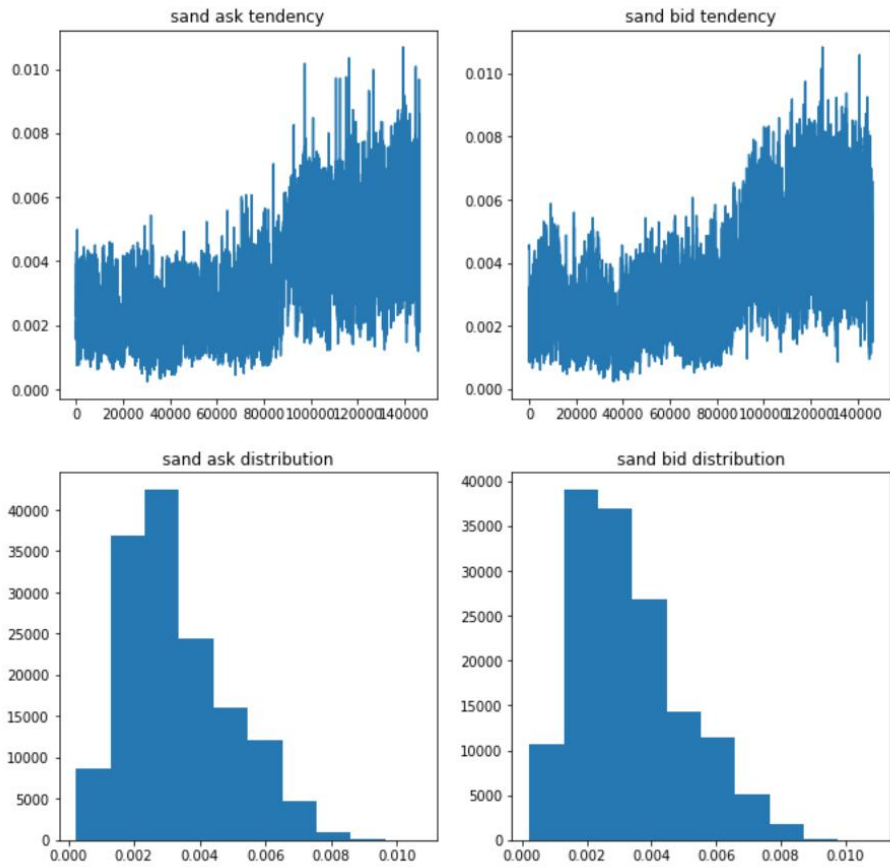
从币安接口获取 146000 条数据（时长大概在 40 个小时），并保存下来。

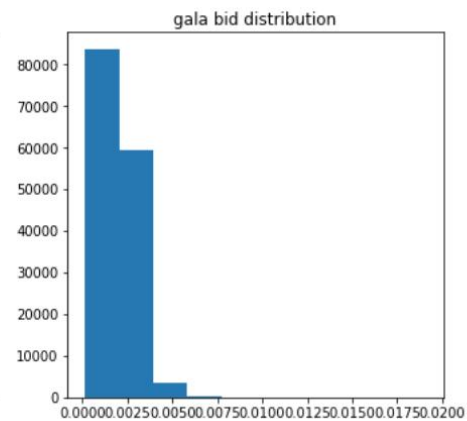
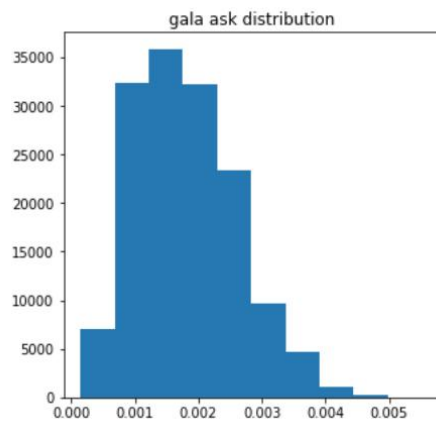
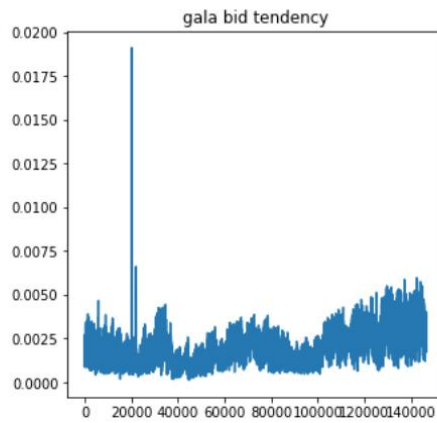
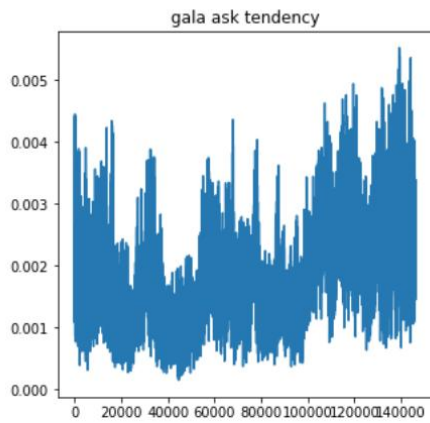
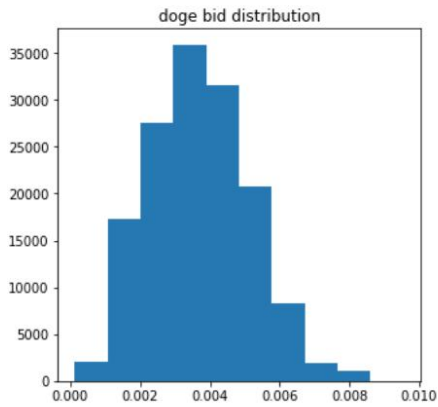
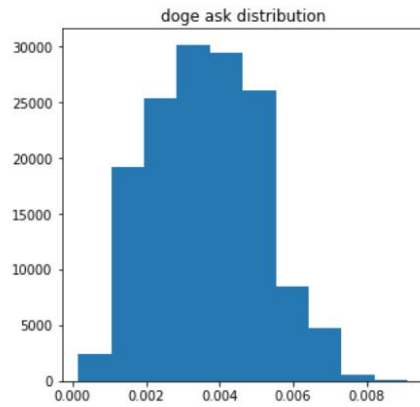
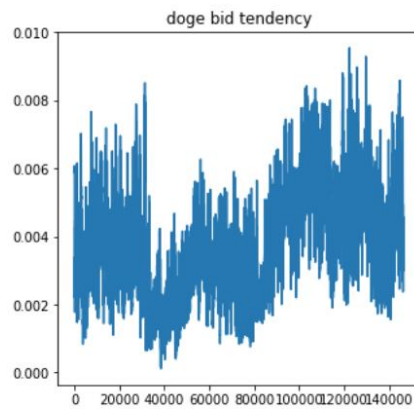
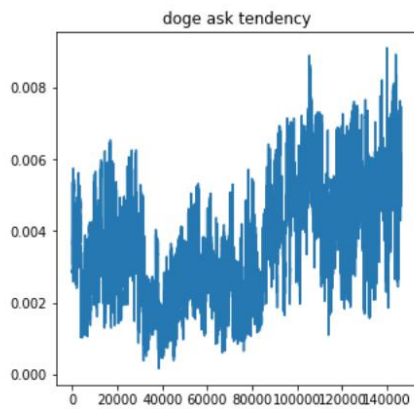
2.2.3 数据展示

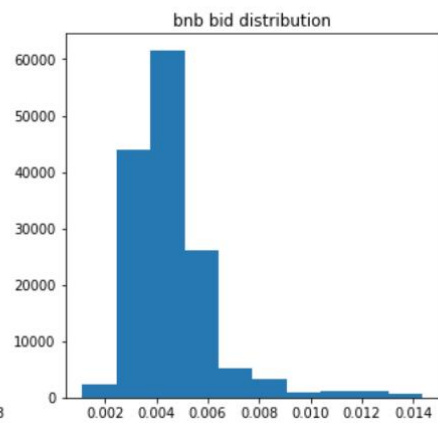
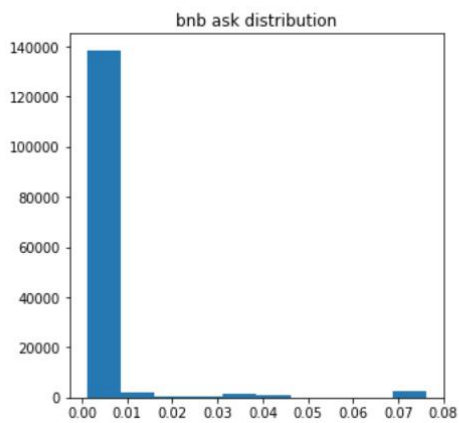
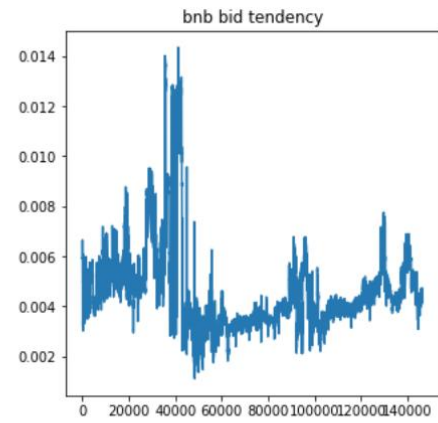
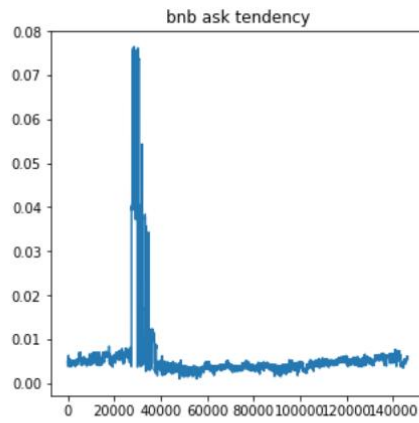
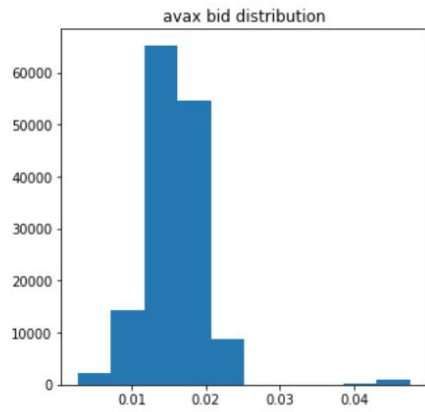
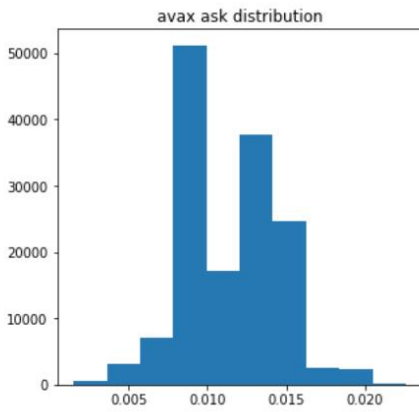
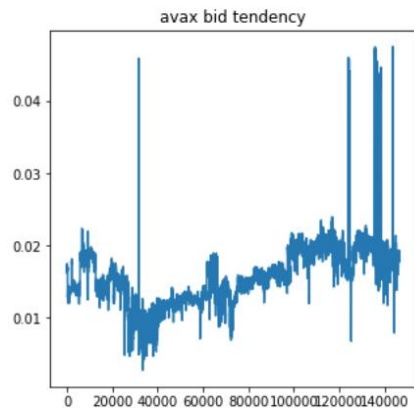
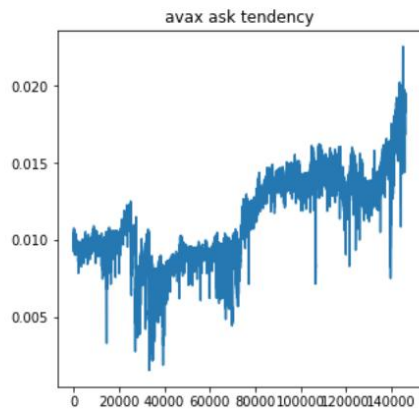
每个币种都有 4 张图，2 张为 α_bid 的走势图中、直方图，2 张为 α_ask 的走势图中、直方图。

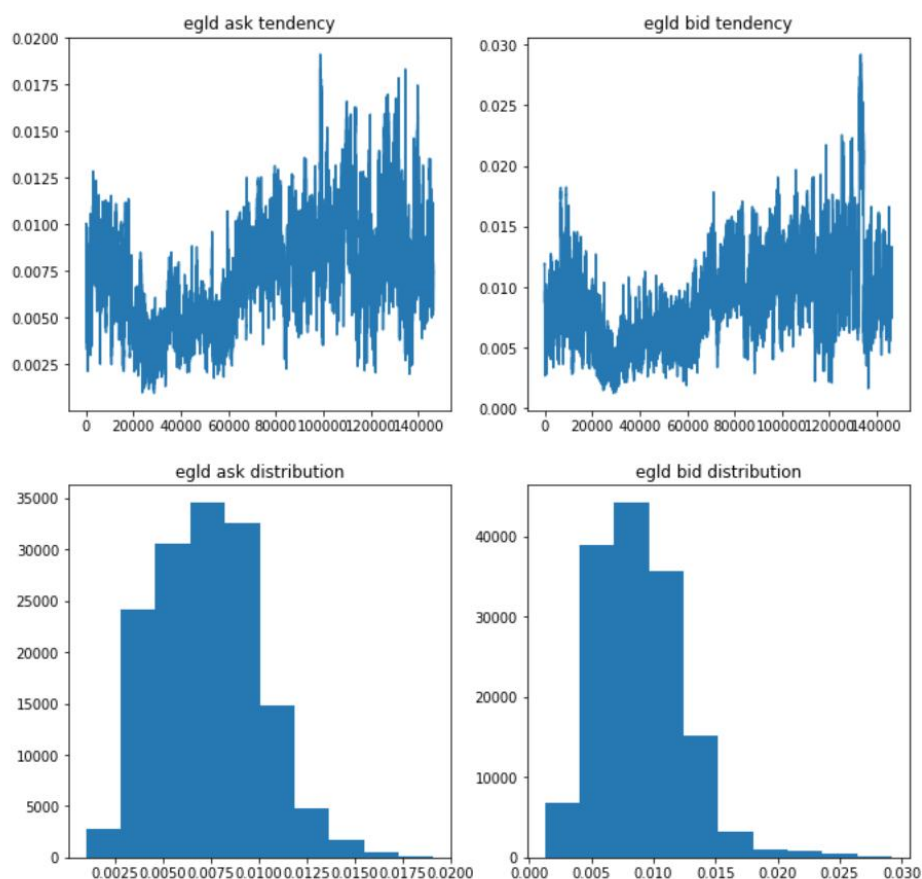












可以看到,

- ① 走势图整体上表现出平稳的特征, 即在某一个值附近震荡, 直方图呈现为有偏的正太分布。
- ② 会有一些极值出现
- ③ 市值高的币种盘口特征越稳定

2.2.3 数据处理

- ① 去掉极值, 计算 1 分位和 99 分位, 分为为 X_{min} , X_{max}
- ② 将 α 的值锁定在 1 分位到 99 分位之间, 利用 `clamp` 函数
- ③ 数据标准化处理

2.2.4 K 的计算公式

$$K = 0.5 * \text{clamp}(a, X_{min}, X_{max}) / (X_{max} - X_{min}) + 0.5,$$

其中 a 为当前盘口前 20 档的挂挡数, `clamp` 函数表示当 $a < X_{min}$ 的时候, 函数返回 X_{min} , 当 $a > X_{max}$ 的时候, 函数返回 X_{max}

三. 后续

3.1 尚未完成的工作

因为没有存放成交额的数据，需要找时间再存放一次。

40 档的数据只有实时的，没有历史数据，也需要时间存放

3.2 因子的大小

里面分位数，数据标准化的区间，都还需要在实际过程中进行优化

3.3 数据模型评价

深度因子模型是建立在 AS 模型上面的，因此当 AS 模型失效的时候，深度因子模型即便再精确也是南辕北辙，所以，在实际测试的时候，测试的结果需要分析是 AS 模型的自身的不足，还是深度因子模型的不精确。