**《construct.py》软件功能说明书**

版本：v1.0 日期：2025-08-18

一、概述

**1. 目标**

构建单文件命令行工具 construct.py，实现六条曲线（SPX、QQQ、Mean-Variance、Minimum-Variance、Equally Weighted、DRL 权重策略）的统一回测、指标计算与图表导出，并与通用金融口径一致。默认频率为月度(M)。

2. 范围

输入包括指数/个股价格、四类权重表；输出包括指标汇总表、资金曲线/回撤曲线（csv+png），以及按月/季/年三种粒度的图表与（可选）收益分布图。

**二、功能性需求（FRD）**

FRD-1 输入与配置

- 指数：SPX.csv、QQQ.csv，价格列统一为 close。

- 个股价格：包含 gvkey、datadate、prccd、ajexdi；复权价 adj\_close\_q = prccd / ajexdi。

- 权重表：mean\_weighted\_rf.xlsx、minimum\_weighted\_rf.xlsx、equally\_weighted\_rf.xlsx、drl\_weight.csv。

- 支持通过 CLI 覆盖输入文件名（可配置）。

- 非交易日停留前值（forward fill）。

FRD-2 频率与年化口径

- 统一通过 freq 控制年化：D=252, W=52, M=12, Q=4, Y=1。

- 默认 freq=M；所有指标（Sharpe、Sortino、IR、年化收益/波动、VaR/CVaR）按同一 freq 年化。

- 输出图表需分别按月/季/年三种粒度导出。

FRD-3 权重对齐与约束

- 不允许卖空（权重<0 报错）。

- 同一 trade\_date 权重和应为 1；若偏离则按比例归一并打印日志（包含原始和归一后的和）。

- 提供 unify\_quarterly\_weights(dict, anchor\_key) 将多策略 trade\_date 对齐到锚定策略。

FRD-4 组合构建与交易成本

- 仅在调仓日发生交易，成本 = |换手| × 价格 × TRANSACTION\_COST。

- TRANSACTION\_COST 作为文件级静态变量，默认 0.1% (=0.001)，在源码顶部注明可修改。

FRD-5 收益序列与重采样

- 从日度持仓与价格计算策略日收益；若需要其它频率，先用几何合成(1+r)连乘-1重采样到目标 freq。

FRD-6 KPI 计算（与通用口径一致）

- 绝对收益：累计收益、年化收益(均值×n)、年化波动(标准差×√n)。

- 风险调整：Sharpe=(年化收益-年化无风险)/年化波动；Sortino 用年化下行动波动（MAR 为年化输入并换算为期率）；Calmar=年化收益/|最大回撤|；IR=年化超额收益/年化跟踪误差（基于策略-基准差序列）。

- 风险控制：最大回撤与回撤持续时间（基于资金曲线与历史峰值）、VaR/CVaR(95%/99%)（历史模拟法，对损失分布）。

- IR 基准默认 SPX；额外生成相对 QQQ 的图表（通过 --ir-benchmark QQQ 触发）。

FRD-7 图表与导出

- 对每个对象（六条曲线）导出：equity.csv/png、drawdown.csv/png。

- 生成月/季/年三套图表；可选导出收益分布直方图与箱线图。

- 汇总指标表：risk\_metrics\_summary.csv。

FRD-8 CLI 参数

--freq {D,W,M,Q,Y} 默认 M

--rf FLOAT 年化无风险，默认 0.02

--mar FLOAT 年化最低可接受收益，默认 0.0

--cost FLOAT 交易成本，默认忽略该参数，使用源码静态变量 TRANSACTION\_COST=0.001

--start/--end 起止日期（YYYY-MM-DD）

--out PATH 输出目录，默认 test\_back

--ir-benchmark STR IR 基准（SPX/QQQ），默认 SPX；若为 QQQ，补充输出相对 QQQ 的图表

--price-file INDEX/SEC 输入文件名可配置（指数/个股）

--weights-\* 各权重表文件名可配置

**三、非功能性需求（NFR）**

- 可复现：同参同数据输出一致。

- 可审计：详细日志（参数、样本区间、freq、权重归一、交易成本、IR 基准、输出文件名）。

- 性能：对大规模价格矩阵支持分块/懒加载（如需）。

- 代码质量：类型注解、单元测试、PEP8。

**四、数据规范**

- 指数：列至少包含 date, close。

- 个股：列至少包含 gvkey, datadate, prccd, ajexdi，复权价 adj\_close\_q；按 gvkey 透视到宽表。

- 权重：列至少包含 gvkey, trade\_date, weight。

**五、算法与公式**

- 重采样：compound 合成 (1+r)连乘-1。

- 年化：n=periods[freq]；ann\_ret=mean\*n；ann\_vol=std\*sqrt(n)；downside 同理。

- Drawdown：equity/cummax - 1；持续时间取最大连续为负段长度。

- VaR/CVaR：对损失 -ret 求分位与尾部均值（95%、99%）。

**六、流程（Pipeline）**

1) 读取并标准化价格（前向填充、对齐日历、截取区间）；

2) 读取权重、做非负与和=1 校验，不满足则归一并告警；

3) 统一调仓日（锚定策略）；

4) 构建组合（调仓日计成本）、得到日收益；

5) 对策略/基准收益重采样至 freq；

6) 计算 KPI；

7) 导出汇总表、三套图表（M/Q/Y）与时序 csv；

8) 记录日志与运行摘要。

**七、主要模块/函数**

- main()：解析 CLI，组织流程

- load\_prices()/load\_index()：读入指数/个股并前向填充

- load\_weights()：读入四类权重并校验/归一

- unify\_quarterly\_weights()：按锚定策略对齐 trade\_date

- build\_portfolio()：生成持仓、价值、成本、日收益

- resample\_returns()：几何重采样到 D/W/M/Q/Y

- compute\_drawdown()：资金曲线→回撤与持续时间

- compute\_metrics()：统一 KPI（Sharpe/Sortino/Calmar/IR/MaxDD/VaR/CVaR）

- save\_timeseries()/save\_histograms()：导出 csv/png

- summarize\_and\_dump()：汇总表导出

**八、目录结构（建议）**

project/

construct.py

data/

test\_back/ # 输出

README.md

requirements.txt

**九、验收标准（抽样）**

- 生成 risk\_metrics\_summary.csv，含 6 条曲线全部 KPI 字段；

- 生成 6×3 套 equity/drawdown 图（M/Q/Y）；

- IR 基准=SPX 时有 IR 列；指定 --ir-benchmark QQQ 时新增相对 QQQ 的 IR 图；

- 权重和≠1 的样本在日志中出现“normalize weights”字样并给出原始/归一值；

- 交易成本静态变量在源码顶端可见。

**十、可配置项**

- 输入文件名、价格列名、freq、rf/mar、IR 基准、输出目录、图表 DPI/风格。

（完）