

案例作业：场外衍生品估值复核 AI Agent (Derivatives Valuation Review Agent)

限时 **72 小时**（预计有效投入 15–20 小时，无需全职）。本案例难度高于常规作业，且刻意信息密集——时间不够用是设计的一部分，我们同时考察你的**取舍能力**。

案例中的公司、指数、数字均为虚构，仅用于测试目的；市场背景描述仅作情境参考。

一、背景与角色设定

你是会计师事务所金融服务审计团队成员，负责**稳盈量化 1 号私募证券投资基金**（管理人：华远资本管理有限公司）2025 年度财务报表审计中的**衍生金融工具估值复核**。

该基金年末持有两类场外衍生品：

- 一笔雪球型收益凭证**（SNOW-001，名义本金 5,000 万元，交易对手：华澜证券），以对手方对账单价格入账；
- 28 笔场外期权及远期组合**（挂钩 5 个标的），以管理人内部 Excel 模型估值入账。

近年市场上雪球类产品曾在标的指数快速下行阶段出现集中敲入、投资者大额亏损的情形，此类结构的估值与风险披露是监管与审计关注的重点。基金管理人的估值流程存在典型痛点：条款分散在合同 PDF 里、估值依赖对手方报价或表格模型、缺少独立验证。

你的任务：**设计并实现一个“估值复核 Agent”**，先在 SNOW-001 上跑通端到端深度复核并出具复核备忘录，再对 28 笔组合完成一次**批量独立测试**。工具的用户是审计师；定位是**复核 (review / independent price verification)**，不是代客估值。

审计重要性水平（直接使用）：基金年末净资产 3.2 亿元；整体重要性 320 万元，实际执行重要性 240 万元，明显微小差异 16 万元。

二、案例资料

2.1 雪球 SNOW-001：合同条款

见数据包 **雪球条款书_SNOW-001**（协议摘要）。请以合同为准提取产品要素——这本身是工具的一部分。

■ 2.2 雪球 SNOW-001：管理人估值台账与复核记录（摘录）

项目	台账记录
产品代码 / 名义本金	SNOW-001 / 50,000,000 元
年末公允价值（入账值）	49,600,000 元（名义的 99.20%）
估值来源	交易对手华澜证券 2025 年 12 月末对账单
公允价值层级	第二层级

管理人风险管理部对对账单价格的复核记录：

复核模型参数	取值	备注
标的现价	5,330.00 点	2025-12-31 收盘（期初价格的 82.0%）
波动率	16.0%	标的指数近三年历史波动率均值
敲出障碍	6,695.00 点	—
敲入障碍	4,550.00 点	取自估值系统产品主数据
无风险利率 / 股息率	2.0% / 2.5%	—
复核结论	模型价 101.7% ≥ 对账单价 99.2%，对账单估值审慎，通过	蒙特卡洛，5 万路径

存续期市场信息：截至 2025-12-31 已经过 8 个敲出观察日均未敲出（期间最高观察价 6,480 点）；存续期内指数最低 5,010 点（2025 年 11 月），未发生敲入。场内同类小盘指数期权平价隐含波动率：1 年期 27.5%、2 年期 26.0%（见 `market_data.csv`）。

■ 2.3 场外期权/远期组合

见数据包 `portfolio_台账.csv`（28 笔，含管理层估值）与 `market_data.csv`（估值日行情、波动率、利率曲线）。计量约定见 `README_数据包.md`。台账与数据可能存在数据质量问题，同属复核范围；组合中估值异常的笔数未知。

三、交付要求（72 小时内完成，先后顺序与时间分配自行安排）

■ Step 1: 设计文档（不超过 3 页，PPT 或 Word）

1. 复核框架：针对本案例两类持仓，你的 Agent 分别检查什么（列出检查点清单，注明依据：准则/无套利关系/估值常识）；
2. Agent 架构：输入→处理→输出流程图；LLM 与确定性计算的分工边界，幻觉与可复现性的控制机制，输出如何做到可追溯（每条发现附证据链）；

3. **测试策略**：你打算如何证明这个 Agent 本身是可靠的（这与"找出管理层的错误"是两个问题）。

■ Step 2: Agent Demo + 主案例深度复核

- 实现可运行的 Agent（技术形式不限），在 **SNOW-001** 上端到端跑通：条款提取 → 与台账勾稽 → 独立重估 → 差异分析；
- 产出《**SNOW-001 估值复核备忘录**》：管理人估值是否可接受？逐条列出发现（含证据与量化影响）、你的独立重估结果与关键假设、公允价值层级意见、结论与建议调整金额（对照重要性水平）。文风为审计底稿，需可被质询。

■ Step 3: 批量独立测试 + 测试报告

- 用你的 Agent 对 28 笔组合做**批量复核**：逐笔独立重估、按你设定的容差阈值标记异常、按严重程度分级，输出**异常清单**（每条注明问题、量化影响、判断依据）；
- 产出《**测试报告**》：阈值如何设定及理由；批量运行结果汇总（异常笔数、金额影响、按原因分类）；**误报/漏报自评**——你如何知道 Agent 没有漏掉问题、没有虚报问题（例如：用已知答案的自造用例做校验、无套利关系交叉验证、对"干净"交易的通过率分析）；
- 汇总两类持仓的**总体错报金额**，对照重要性水平给出审计结论建议。

■ 提交与答辩

- 72 小时截止时提交：Step 1—3 全部文档 + 代码（仓库或压缩包）+ 一页运行说明 + **AI 协作说明**（≤1 页）：你如何使用 AI 工具，附 1 个你认为最关键的 prompt；并**必须包含一个真实案例：AI 在本作业中给出的某个错误——你如何发现它、如何修正、以及做了什么防止它再犯**（一个错误都没发现，本身会被视为验证不足的信号）；
- 提交后另约 **30 分钟答辩**：10 分钟演示，20 分钟问答（问答比重高于常规，请预留精力）。

四、允许的简化（须在文档中声明）

- 雪球重估可采用 GBM 蒙特卡洛（不要求随机波动率/局部波动率模型）；敲入按每日收盘观察、敲出按月度观察日模拟即可；
- 不要求构建波动率曲面（平价波动率+期限插值即可），微笑/偏度影响可定性讨论；
- 不要求计算 CVA/DVA 的精确金额，但对是否需要调整应有定性判断；
- 美式期权可用二叉树或带下限校验的合理近似；
- 若某项来不及实现，**明确写出**"未覆盖清单"——诚实的未覆盖清单不扣分，被答辩问出来的漏项要扣分。

五、说明

- 允许并鼓励使用任何 AI 工具；但答辩中你须能解释每一个专业判断与你代码的核心逻辑——**AI 可以替你写代码，不能替你负责**；
- 评分权重：专业判断与发现质量 > 独立重估的方法论 > 工程完成度；批量测试部分重点考察你**验证自己工具**的意识；

- 测试报告中的任何声明（如"该类错误可被检出"）都可能在答辩中被现场验证，请确保与代码行为一致。