



## 量化金融分析师（AQF®）全国统一考试

### 模拟题

适用场次：2024 年 3 月

使用本模拟题，您应该遵守：

1. 本模拟题仅提供给参加 2024 年 3 月份 AQF 全国统一考试的考生，考生仅可以出于准备个人考试的目的查阅和打印本模拟题；
2. 严禁出于任何目的的复制、网络发布和传播、抄袭本模考题内容，如有违反，可能导致违纪或违法行为；

© 版权所有，侵权必究。

量化金融标准委员会

Standard Committee of Quantitative Finance

## 量化金融分析师（AQF®）全国统一考试模拟题

说明：本场考试中的代码都应采用 Python 3.X 版本作答。

1. 单选题（每题 2 分，本部分共 40 分）：只有一个正确答案，选对得 2 分，选错或不选得 0 分。

1.1. Python3 的标准数据类型中，列表和元组有什么不同（ ）

- A. 列表是可变的，元组是不可变的。
- B. 列表支持数字，元组不支持数字。
- C. 列表可以绘图，元组不可以绘图。
- D. 列表可以嵌套，元组不可以嵌套。

参考答案：A

1.2. Python 中，运行以下代码，可以获得（ ）

```
stock_info = {'code': '000001', 'name': '平安银行', 'industry': '银行'}  
stock_info['name']
```

- A. 'name'
- B. '平安银行'
- C. '000001'
- D. '银行'

参考答案：B

1.3. 以 55 元购买一股 X 公司股票，一年之后，收到了 2 元红利，以 58 元卖出股票，你的持有期回报是多少？

- A. 3.45%
- B. 8.62%
- C. 9.09%
- D. 5.45%

参考答案: C

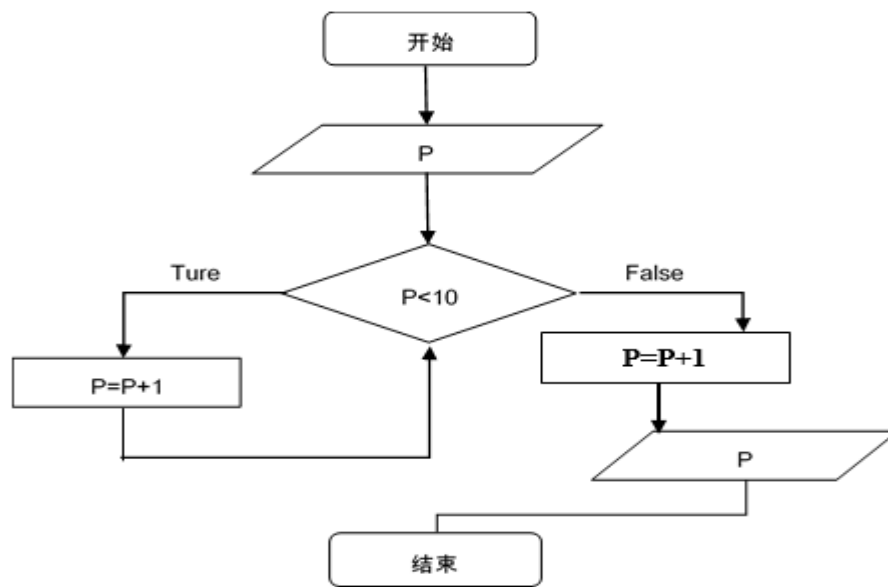
1.4. 某交易员现有两个未完成的限价单。在正常交易时间内, 第一个限价单成交的概率为 0.40, 第二个限价单成交的概率为 0.50, 两个订单都成交的概率为 0.10。则两个订单中至少有一个成交的概率是 ( )

- A. 0.2
- B. 0.3
- C. 0.8
- D. 0.9

Q:  
 $P(A1) = 0.4$   $P(A2)=0.5$   $P(A1*A2)=0.1$   
 $P(A1 + A2) = 1-0.9*0.6*0.5=0.73$

参考答案: C

1.5. 某控制结构如下, 如果输入 P=7, 则会输出 ( )



- A. 9
- B. 10
- C. 11
- D. 以上都不对

参考答案: C

1.6. 在投资过程中, 基本面分析和技术分析都是重要的分析方法。某量化研究员正在研究多因子策略, 并整理了一系列候选因子如下, 请问哪个属于技术面因子类型 ( )

- A. PB 因子
- B. EBIT 因子
- C. VR 因子
- D. ROA 因子

参考答案: C

1.7. 假如 x 是一个数组，以下哪一代码表示查看数组的维度？

- A. `print(x.shape)`
- B. `print(x.size)`
- C. `print(x.reshape())`
- D. `print(x.T)`

参考答案: A

1.8. 在 CAPM 模型中，股票的超额收益率是指（ ）

- A. 股票的总收益率减去无风险收益率
- B. 股票的总收益率减去市场收益率
- C. 股票的期望收益率减去无风险收益率
- D. 股票的期望收益率减去市场收益率

参考答案: A

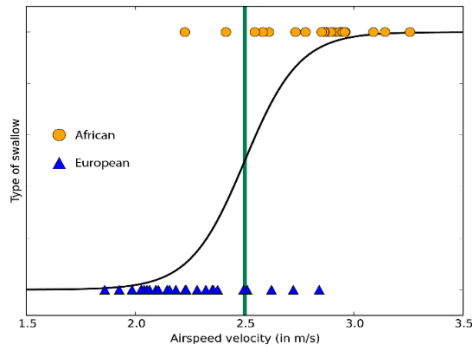
1.9. 现有变量 data，数据类型为 `numpy.ndarray`，具体数据如下：`data = np.array([0.2, 0.2, 1.3, 0.8, 0.6, 1.1])` 以下无法将该一维数组转换为列表数据的是（ ）

- A. `data.list()`
- B. `data.tolist()`
- C. `[i for i in data]`
- D. `list(data)`

参考答案: A

1.10. 机器学习技术目前已经越来越多地应用到量化投资领域，下图为某机器学习模型的

原理，则该图展示的机器学习模型最有可能是（ ）



- A. 逻辑回归
- B. 决策树
- C. 随机森林
- D. 线性回归

参考答案：A

1.11. 7月1日，一位投资者买入两份小麦期货，每份期货规模为100吨小麦。初始保证金为每份合约3,000美元，维持保证金为每份合约1,500美元。7月1日，期货价格为173美元/吨，7月2日，价格上涨为179.75美元，则该投资者在7月2日的保证金账户余额为（ ）

- A. 1,350 美元
- B. 2,300 美元
- C. 4,650 美元
- D. 7,350 美元

$$3000 \times 2 + (179.5 - 173) \times 2 \times 100 = 7350 \text{ \$}$$

参考答案：D

1.12. 线性回归模型通常利用判定系数（R-Squared）来判断模型拟合度。此时，如果在模型中增加一个特征，则下面哪个说法是正确的？

- A. 如果 R-Squared 增加，则这个特征有意义
- B. 如果 R-Squared 减小，则这个特征无意义
- C. 仅看 R-Squared 单一指标，无法确定这个特征是否有意义
- D. 以上说法都不对

参考答案：C

1.13. 以下数据类型中，更新频率最慢的是（ ）

- A. 公司 ESG 评级数据
- B. 十档行情数据
- C. 日成交量数据
- D. 股票分时图数据

参考答案：A

1.14. 李明，AQF，某 FOF 基金经理，希望通过指标了解某策略运行过程中策略达到净值高点后可能出现的最大损失程度，以此评估购入某量化基金产品后可能出现的最大损失，则他应该了解量化基金产品的以下哪项指标（ ）

- A. 波动率
- B. 下行偏差
- C. 最大回撤率
- D. 基金管理规模

参考答案：C

1.15. 下列不属于无监督学习算法的是（ ）

- A. K 均值聚类
- B. 主成分分析
- C. 决策树
- D. 流形学习

参考答案：C

1.16. 李明，某量化研究员，在使用 Jupyter Notebook 编写策略代码的过程中，他可以使用以下哪个魔法命令的功能来查看一个 cell 运行所花费的平均时间？

- A. `%matplotlib inline`
- B. `%time?`
- C. `%whos`
- D. `%timeit`

参考答案: D

1.17. 数据处理中, 对于数据缺失, 通常不会进行的是 ( )

- A. 判断
- B. 删除
- C. 填充
- D. 保留

参考答案: D

1.18. 以下关于基本面分析的论述错误的是 ( )

- A. 对于具体的个股而言, 影响其价位高低的主要因素在于企业本身的内在素质, 包括财务状况、经营情况、管理水平、技术能力、市场大小、行业特点、发展潜力等一系列因素。
- B. 对于企业的偿债能力分析, 利息保障倍数指标能够有效分析企业的资金风险, 便于提前预警。
- C. 对于企业的营运能力分析, 一般包括应收帐款周转率和应付帐款周转率两种。
- D. 市销率是最近两年在国际资本市场新兴起来的市场比率, 主要用于创业板的企业或高科技企业。在 NASDAQ 市场上上市的公司不要求有盈利业绩, 因此无法用市盈率对股票投资的价值或风险进行判断, 而用该指标进行评判。

参考答案: C

1.19. 李明, AQF, 某量化基金经理, 在进行量化策略回测过程中, 需要计算每日收益率。

已知每日收盘价数据存储在变量 `close` 中, 数据类型为 `pandas.Series`。则下列代码中可以正确计算每日收益率的是 ( )

- A. `close/close.shift(0) - 1`
- B. `close/close.shift(-1) - 1`
- C. `close/close.shift(1) - 1`
- D. `close.pct_change(-1)`

参考答案: C

XXXX	CLOSE	shift(1)	shift(-1)	pct
08.12	1	NAN	2	NAN
08.13	2	1	1.5	1
08.14	1.5	2	NAN	0.025

1.20. 李明, AQF, 某量化基金经理, 在进行蒙特卡洛模拟时想要产生 500 行 5 列服从

标准正态分布的随机数，以下选项不正确的是（ ）

- A. `numpy.random.randn(500, 5)`
- B. `numpy.random.random((500, 5))`
- C. `numpy.random.normal(size=(500, 5))`
- D. `numpy.random.standard_normal((500, 5))`

参考答案：B

2. 多选题（每题 2 分，本部分共 20 分）：有 2-5 个正确答案，全部选对得 2 分，少选得 1 分，选错或不选得 0 分。

2.1. Python3 中有六个标准的数据类型，以下不属于标准数据类型的是（ ）

- A. While
- B. For
- C. Number
- D. Tuple
- E. Dictionary

参考答案：AB

2.2. Python3 的标准数据类型中，以下不属于可变数据类型的是（ ）

- A. Number
- B. Tuple
- C. Set
- D. String
- E. List

参考答案：ABD

2.3. 以下选项中，符合 Python 语言变量命名规则的是（ ）

- A. `_2023`
- B. `keyword88_`
- C. `PRINT_2023`



D. 2023?

E. 2023\_PRINT

**参考答案: ABC**

解析: 标识符可以由字母、数字、下画线( )组成, 其中数字不能放在首位。标识符不能是 Python 关键字, 但可以包含关键字。标识符不能包含空格。

2.4. 下列关于债券交易策略的说法中正确的是 ( )

- A. 卖出凸性策略是购买凸性小的债券, 卖出凸性大的债券
- B. 久期管理策略在预期利率下降时, 增加组合久期获得收益
- C. 持有到期策略属于被动管理策略
- D. 骑乘收益率曲线的投资者会购买比要求的期限稍短的债券
- E. 消极的债券组合管理策略将市场价格假定为公平的均衡交易价格

**参考答案: ABE**

2.5. 李明, 某量化研究员, 在量化交易策略的回测过程中, 他发现回测收益往往会出现高于实盘收益, 可能导致这种情况的原因有 ( )

- A. 回测印花税率高于实际印花税率
- B. 使用了未来函数
- C. 策略模型过拟合或过度优化
- D. 在交易过程中低估了交易成本
- E. 存在幸存者偏差

**参考答案: BCDE**

2.6. 数据清洗是量化策略开发过程中不可缺少的一个环节, 其结果质量直接关系到策略回测的准确度。因此, 在数据分析之前, 研究员往往会花费大量的时间来进行数据清洗工作。

以下数据清洗的做法, 正确的有 ( )

- A. 数据录入过程、数据整合过程都可能会产生重复数据, 为保持数据完整不应随意删除
- B. 对于量级相差大的数据进行 z-score 标准化处理
- C. numpy 中可以使用 fillna 方法替换缺失值数据

D. 在适当情况下, 可以使用某个变量的样本均值、中位数或众数代替无效值和缺失值

????

E. 对于数据中的离群值，单独进行标准化处理

**参考答案：BC**

2.7. 在投资过程中，基本面分析和技术分析都是重要的分析方法。某量化交易员正在研究多因子策略，并整理了一系列候选因子如下，请问哪些属于技术面因子类型（）

- A. PB 因子
- B. CCI 因子
- C. WR 因子
- D. MACD 因子
- E. ROA 因子

**参考答案：BCD**

2.8. 关于基本面多因子模型，正确的是（）

- A. 基本面多因子模型的基本假设是具有不同“属性”的股票，在市场上应该有不同的收益率；
- B. 基本面多因子模型主要解释变量是可观察到的股票（上市公司）自身的基本属性，比如市盈率、市值大小等；
- C. 基本面多因子模型的主要分析方法是进行横截面分析，以确定股票收益率对因子的敏感性（因子载荷）。
- D. 基本面多因子模型的主要分析方法是进行时间序列分析，以确定股票收益率对因子的敏感性（因子载荷）。
- E. 在实际操作中，基本面多因子模型效果要明显好于其他两类模型，是现在的多因子模型研究的主流；

**参考答案：BCE**

2.9. 李明，某量化基金经理，把某股票一周的收益率存储在变量 `arr` 中，`arr = np.array([-0.06, -0.13, 0.19, -0.11, 0.02])`，他现在想要判断这一周的股票收益率是否都小于 0，如果都小于 0 返回 `True`，否则返回 `False`。可以实现该功能的代码为（）

- A. `np.all(arr<0)`
- B. `np.any(arr<0)`

- C. `np.where(arr<0)`
- D. `np.all(np.where(arr<0,True,False))`
- E. `np.sign(arr<0)`

参考答案: AD

2.10. 李明, AQF, 某量化研究员通过历史模拟的方法计算 VaR, 得出 50ETF 一日 99%和 95%的 VaR 值分别是-0.0518 和-0.0265, 分别存储在变量 VaR\_1 和 VaR\_5 中, 以下打印此结果的代码正确的是 ( )

- A. `f'50ETF 一日 95%的 VaR 为{VaR_5*100:.2f}%, 99%的 VaR 为{VaR_1*100:.2f}%'`
- B. `'50ETF 一日 95%的 VaR 为%.2f%%, 99%的 VaR 为%.2f%%' % (VaR_5*100, VaR_1*100)`
- C. `'50ETF 一日 95% 的 VaR 为 {:.2f}% , 99% 的 VaR{:.2f}%'.format(VaR_5*100, VaR_1*100)`
- D. `'50ETF 一日 95%% 的 VaR 为 %.2f%% , 99%% 的 VaR 为 %.2f%%' % (VaR_5*100, VaR_1*100)`
- E. `f'50ETF 一日 95%的 VaR 为{VaR_5*100:.2f}%%, 99%的 VaR 为{VaR_1*100:.2f}%%'`

参考答案: ACD

3. 解答题 (每题 4 分, 本部分总计 40 分): 按步骤、得分点给分。

3.1. 假设有一张面值为 1,000 元的公司债券, 以 8%的息票率每半年付息一次, 付息时间分别为 1 月 1 日和 7 月 1 日。如果当前时间点为 2023 年 4 月 1 日, 而债券将于 2033 年 7 月 1 日到期, 市场要求回报率为 5%, 计算出这只债券的净价是?

参考答案:

首先计算 2023 年 1 月 1 日的价值:

$$N = 21, I/Y = 2.5, PMT = 40, FV = 1000 \rightarrow PV = -1,242.77$$

随后计算 2023 年 4 月 1 日的价值, 即全价:

$$1,242.77 \times (1 + 2.5\%)^{0.25 \times 2} = 1,258.21$$

最后, 计算净价:

$$1,258.21 - 40 \times \frac{90}{180} = 1,238.21$$

3.2. 假设去年某公司支付每股股利为 1.1 元，预计在未来日子里该公司股利按每年 5% 的速率增长。假定要求收益率为 8%，目前该公司的股票每股价格为 35.5 元，你建议当前持有该股票的投资者出售股票吗？（要求写出分析步骤）

**参考答案：**

利用 DDM 永续增长模型，该公司股票估值为  $1.1 \times (1 + 0.05) / (0.08 - 0.05) = 38.5$  元，当前股票价格为 35.5 元，因此股票被低估，建议持有该股票的投资者继续持有该股票。

3.3. 李明，AQF，某量化研究员，在策略研究的过程中，生成了一系列日期数据 `dates=pd.Series(['20230301', '20230302', '20230303', ...])`，但是日期的数据类型为字符串，不方便之后时间序列数据的处理，请你帮他把日期格式改为 `datetime` 格式。

**参考答案：**

```
dates = pd.to_datetime(dates, format='%Y%m%d')
```

3.4. 李明，AQF，某量化研究员，获得了某股票 2022 年的价格数据存储于变量 `k_df` 中，数据类型为 `pandas.DataFrame`，部分数据如下：

date	open	high	low	close	volume
2022/1/17	16.24306	16.53573	16.14551	16.45768	5679.0
2022/1/18	16.52597	16.88693	16.50646	16.76011	6789.0
2022/1/19	16.59426	16.9162	16.49671	16.6528	3567.0
2022/1/20	16.71133	16.85766	16.53573	16.73084	7897.0
2022/1/21	16.58451	16.63329	16.22355	16.25282	3210.0
2022/1/24	16.39915	16.51622	16.12599	16.37964	4982.0
2022/1/25	16.37964	16.39915	16.11624	16.28208	8901.0

在研究策略的过程中，他希望计算出成交量 20 日的滚动之和，以便观察成交的活跃程度，并将计算的结果存储在变量 `k_df` 新的一列“vol”中，请写出相应代码。

**参考答案：**

```
k_df['vol'] = k_df['volume'].rolling(20).sum()
```

3.5. 李明，AQF，某量化基金经理，在进行策略研究时需要了解各行业情况，他把股票信息存储于变量 `stock_data` 中，数据类型为 `pandas.DataFrame`，存储数据为股票当天的市盈率数据（`pe_ratio`）、负债权益比数据（`debt_to_equity_ratio`）、行业分类

(industry)。部分数据如下：

	pe_ratio	debt_to_equity_ratio	industry
Stock1	31.72	0.82	Energy
Stock2	10.48	5.12	Energy
Stock3	21.69	2.56	Auto
Stock4	11.81	3.43	Auto

请使用 `groupby`，找出 `pe` 均值最大的行业，将该行业的名称保存在变量 `pe_max_industry` 中。（已完成的第三方库导入：`import numpy as np` 和 `import pandas as pd`）

**参考答案：**

```
pe_data = stock_data.groupby('industry')['pe_ratio'].mean()
pe_max_industry = pe_data.idxmax()
```

3.6. 李明，AQF，某量化研究员，想要选取因子符合特定条件的股票。因子数据如下：

```
factor_a = {'600000':0.02, '600004':0.04, '600005':0.08, '600006':0.01,
            '600007':0.1, '600008':0.05, '600009':0.09, '600010':0.03}
factor_b = {'600000':13, '600004':2, '600005':6, '600006':12, '600007':7,
            '600008':8, '600009':11, '600010':16}
```

研究员给出的条件为：因子 `a` 值不低于 `0.04` 且因子 `b` 值大于 `10` 的股票代码储存在一个名为 `result` 的列表中，当该列表中的股票达到 `2` 只时正常停止选取。请给出该过程的代码。

**参考答案：**

```
result = []
for code in factor_a.keys():
    if len(result) >= 2:
        break
    if factor_a[code] <= 0.04 and factor_b[code] > 10:
        result.append(code)
```

3.7. 李明，AQF，某量化基金经理在研究股票策略时，获取了下列数据，数据类型是 `pandas.DataFrame`，他想要筛选出符合下列条件的股票：市净率大于 `3`、流动比率在 `100` 到 `200` 之间。并以列表的形式储存在变量 `stock_lst` 中，请写出该筛选过程的代码。

0	code	pb	current_ratio
---	------	----	---------------

1	600567	5.46	199.53
2	000987	1.98	23.10
3	600812	3.12	217.77
4	300230	2.23	223.08
5	000789	9.11	79.13
6	300113	2.98	118.11

参考答案:

```
stock_lst = list(data[(data['pb'] > 3) & ((data['current_ratio'] > 100)
& (data['current_ratio'] < 200))]['code'])
```

```
stock_lst = []
for data['pb'] > 3 and (data['current_ratio'] > 100 and data['current_ratio'] < 200):
    stock_lst.append(data['code'])
```

3.8. 李明, AQF, 某量化基金经理, 在量化回测的过程中, 获取了某股票的每日价格数据,

并将其存储在变量 `data` 中, 数据类型为 `pandas.DataFrame`, 部分数据如下:

date	open	close
2020-02-23	0.891	0.876
2020-02-24	0.876	0.876
2020-02-25	0.877	0.880
2020-02-28	0.878	0.872

现需编写一段代码用于计算该股票的收阳线的天数和收阴线的概率, 并打印出来。

参考答案:

```
data['up']=data['close']-data['open']
print('收阳线天数为: ', len(data[data['up'] > 0]))
print('收阴线概率为: ', len(data[data['up'] < 0]) / len(data))
```

3.9. 李明, AQF, 某量化研究员, 为了进行策略研究, 从数据平台上下载了某股票的部分日交易数据, 包括收盘价、交易量和成交均价, 该数据储存在变量 `stock_df` 中, 数据类型为 `pandas.DataFrame`。

date	close	volume	vwap
2022-01-10	35.22	1374600	36.62
2022-01-11	36.79	2108200	35.83
2022-01-12	35.55	1002400	36.39
2022-01-13	36.04	1872100	37.66
2022-01-14	36.38	1779200	36.01
2022-01-17	35.10	1395500	35.73

为了进一步分析，他想先根据交易量和成交均价计算出成交金额，然后计算 60 日成交金额的移动平均，并在该 DataFrame 中新增一列 `amount_ma`，将该计算结果储存在这一列中。请写出以上计算的相应代码。

**参考答案：**

```
stock_df['amount'] = stock_df['volume'] * stock_df['vwap']
stock_df['amount_ma'] = stock_df['amount'].rolling(60).mean()
```

3.10. 李明，AQF，某量化基金经理，正在分析股票价量策略。他获取了某股票的价格数据计算出了信号和收益，并挑选了策略对标的指数，并把这些数据和该指数数据存储于 `data` 变量中，数据类型为 `pandas.DataFrame`，部分数据如下：

	position	return	index return
2022-01-04	0	NaN	NaN
2022-01-05	1	0.01586	0.00422
2022-01-06	0	0.02210	0.00814
2022-01-07	0	-0.01712	-0.00490
2022-01-08	1	0.02826	0.01233

其中，`data` 的行索引为 `DatetimeIndex` 数据类型，“`position`”列为策略的持仓信号，1 表示满仓持有，0 表示无持仓，“`return`”列表示根据收盘价算出的股票日收益率，“`index return`”列为指数的日收益率，李明现希望计算该策略相对该指数的累计超额收益，并存储于变量 `data` 新的一列“`cum_hedge_r`”中，请写出相应代码。

**参考答案：**

```
data['r'] = data['position'] * data['return']
data['hedge_r'] = data['r'].fillna(0) - data['index return'].fillna(0)
data['cum_hedge_r'] = (1 + data['hedge_r']).cumprod()
```