

# 國立臺灣大學資訊系統訓練班

## 第 330 期 Python 初階證券交易分析

### 學習評量

本試卷作答至最後一堂課前一日 以前 (含)，除題目要求撰寫程式外者，皆以書面作答。

#### 作答注意事項

- 總分上限為 100 分
- 第一頁需註明學號與姓名
  - 可互相討論，但需獨立作答。
  - 引用網路資源請註明來源。
- 檔案格式限定 PDF 檔
  - 檔案名稱為學號與版號，如 “PFA330113\_1.pdf”。
  - 限定紙張大小為 A4、字體大小為 12 pt、中文。
  - 禁止手寫掃描或拍照。
  - 程式碼無需繳交，會於第十堂課堂中檢核之。
  - 總頁數不得超過 20 頁。
- 上傳繳交
  - Dropbox Link: <https://www.dropbox.com/request/V5om3MVAxsLYn2sudtRc>
  - 若重複相同名稱上傳，則會系統會自動覆寫上一版本。
  - 若需取得批改結果，則請於第一頁註明電子信箱，老師有空會即時批改並寄送批改結果以供同學修正。

Question:	1	2	3	4	5	6	7	8	9
Points:	8	2	5	10	6	6	6	2	5
Score:									
Question:	10	11	12	13	14	15	16	17	Total
Points:	12	21	20	5	20	18	25	10	181
Score:									

## 基本知識

1. (8 points) 於臺灣證券交易所的官方網站上，其提供了每日收盤行情資料，其中包含了每一檔上市證券在交易日所成交的相關資訊，如 (a) 開盤價(opening price)、(b) 最高價(highest price)、(c) 最低價(lowest price)、(d) 收盤價(closing price)。請分別對以日為單位交易的投資人(已經持倉者與未持倉者)明確說明上述四種價格的用途差異。

Total for Question 1: 8

2. (2 points) 請說明設定停損機制是否可以保證不降低報酬。若回答為「是」，請論證之；若回答為「否」，請舉例之。

Total for Question 2: 2

3. (5 points) 若一個策略 A 之報酬期望值與報酬標準差分別為  $\mu_A$  與  $\sigma_A$ ，且不存在任何一其他與策略 A 相同報酬標準差的策略 B 有高於 A 的報酬期望值，則稱之為 A 為一個「有效率的」策略。即對於所有的策略 B 有  $\sigma_A = \sigma_B$ ，則必然  $\mu_A \geq \mu_B$ 。此時，A 便稱之為有效率。試論證，若一個策略 C 若能藉由停損來提升報酬期望值或降低報酬標準差(其他狀況不變)，則 C 必然不是一個有效率的策略。

Total for Question 3: 5

4. (10 points) 考慮市場僅有兩投資人 A 與 B 能於下一時刻進行買進、只有一證券 C 與持倉者將持有至該證券價格回復至合理價時平倉。同時，該證券 C 現值為 100 元，合理價格為 105 元，且市場該下一時刻僅能有一張的成交量，及以出價較低者之價格完成搓合。若兩投資人以相同價格出價進行交易，則認為其委託單會成交的機率均等。現分別考慮兩種情形 (a) 有預期合理價格，兩人皆認為合理價格為 105 元；(b) 僅預期趨勢，兩人皆認為合理價格高於 100 元，且價格將逐步攀升直至價格過熱後回復合理價格。請說明兩人於理性下所應執行的策略與獲利狀況。即，需具體說明兩人該如何出價，與當證券 C 回復至合理價格時兩人之獲利盈虧。

## Python 基礎

5. (6 points) 請說明 Python 程式語言中是否有物件導向設計中最重要三種功能 (a) 類別(class)、(b) 抽象類別(abstract class)、(c) 介面(interface)。若存在，請說明如何定義(使用)之；若無存在，請說明如何模擬(仿造)之。

Total for Question 5: 6

6. (6 points) 物件導向中的類別大致可以分為兩類 (a) 靜態類別(static class)、(b) 非靜態類別(non-static class)。請各舉一例，同時以 Python 程式語言說明該如何時做，並說明該兩種類型的使用時機差異。

Total for Question 6: 6

## 真實帳戶價值與市場除權息影響

7. (6 points) 請分別舉例說明當存在 (a) 除息(cash dividend)、(b) 除權(ex-right)、(c) 現金增資(rights offering) 時，投資人卻於數據處理時未考慮該現象所可能產生的負面影響。

Total for Question 7: 6

8. (2 points) “發行量加權股價指數”(大盤)之還原價“發行量加權股價報酬指數”(大盤還原)乃是採用向後還原的方式還原，而一般所見證券的還原價多採用向前還原的方式，兩者有所不同。請說明此時對於大盤使用向後還原而非向前還原的理由。

Total for Question 8: 2

## 數學理論

9. 我們考慮線性迴歸模型敘述如下：

$$Y_t = X_t\beta + \epsilon_t, t = 1, \dots, T \quad (1)$$

上式中

- $Y_t$  為  $(1 \times 1)$  反應變數在時間  $t$  下之結果；
  - $\beta$  為  $(p \times 1)$  非隨機參數向量；
  - $\epsilon_t$  為  $(1 \times 1)$  隨機誤差項，期望值  $E[\epsilon_t] = 0$  且變異數  $\text{Var}[\epsilon_t] = \sigma_\epsilon^2$ ， $t = 1, \dots, T$ 。同時，對於所有的  $t$  與  $s$ ， $t \neq s$ ， $\text{Cov}[\epsilon_t, \epsilon_s] = 0$ 。
  - $X_t$  為  $(1 \times p)$  預測變數，滿足對於所有  $t$  與  $s$ ， $\text{Cov}[X_t, \epsilon_s] = 0$ 。
- (a) (2 points) 請將方程式寫成向量表達式 ( $y = X\beta + \epsilon$ )。
- (b) (3 points) 請證明最小平方方法的估計為  $\hat{\beta} \equiv (X'X)^{-1}X'y$ 。

Total for Question 9: 5

## 核心概念

10. 現考慮有二證券，證券 A 與證券 B，及一由該二證券所構成之 ETF，其關係式如下：

$$\text{Price}_A(t) = a_A + b_A \text{Factor}(t) + \epsilon_A(t), \quad (2)$$

$$\text{Price}_B(t) = a_B + b_B \text{Factor}(t) + \epsilon_B(t), \quad (3)$$

$$\text{Price}_{\text{ETF}}(t) = c_A \text{Price}_A(t) + c_B \text{Price}_B(t), \quad (4)$$

其中， $a_A$ 、 $a_B$ 、 $b_A$  與  $b_B$  為未知常數係數， $c_A$  與  $c_B$  為已知常數係數， $\text{Factor}(t)$  為一已知影響因子， $\epsilon_A(t)$  與  $\epsilon_B(t)$  為誤差 (errors)。

- (a) (1 point) 請說明如何藉由此三個有價證券建構一不受因子  $\text{Factor}(t)$  的投資組合，且每一證券之權重不得為 0。
- (b) (2 points) 在不存在套利的情況下 (市場具有效率性)，答案 10. (a) 所建構之投資組合是否有交易的價值 (潛在獲利空間)。
- (c) (6 points) 現考慮方程式 (2) 與方程式 (3) 中的係數皆改為已知。即係數  $a_A$ 、 $a_B$ 、 $b_A$  與  $b_B$  皆為已知。請以下兩種組成方式建構不包含因子  $\text{Factor}(t)$  的投資組合，(i) 證券 A 與證券 B、(ii) 證券 A 與 ETF。
- (d) (3 points) 請比較答案 10. (c) 所建構之兩種投資組合的差異。

Total for Question 10: 12

11. 考慮一投資人，欲建構一不受因子影響的投資組合。請在下列不同條件下，完成該投資組合之建構 (需明確給出交易的權重)。

- (a) (3 points) 考慮該投資人已明確選定二證券，證券 A 與證券 B，已知該二證券有線性共用一已知因子，及其共用之關係。即如下關係：

$$\text{Price}_A(t) = a_A + b_A \text{Factor}(t) + \epsilon_A(t), \quad (5)$$

$$\text{Price}_B(t) = a_B + b_B \text{Factor}(t) + \epsilon_B(t), \quad (6)$$

其中， $a_A$ 、 $a_B$ 、 $b_A$  與  $b_B$  為已知常數係數， $\text{Factor}(t)$  為一已知影響因子， $\epsilon_A(t)$  與  $\epsilon_B(t)$  為誤差。

以下小題考慮投資人可使用歷史股價數據與歷史因子數據，即  $\text{Price}_A(1), \dots, \text{Price}_A(T)$ 、 $\text{Price}_B(1), \dots, \text{Price}_B(T)$  與  $\text{Factor}(1), \dots, \text{Factor}(T)$  數據。

- (b) (6 points) 考慮該投資人已明確選定二證券，證券 A 與證券 B，已知該二證券有線性共用一已知因子，然不知其共用之關係。即關係仍如方程式 (5)–(6)，然其中， $a_A$ 、 $a_B$ 、 $b_A$  與  $b_B$  變更為未知常數係數。
- (c) (6 points) 考慮該投資人已明確選定二證券，證券 A 與證券 B，已知該二證券有線性共用一未知因子。即關係仍如方程式 (5)–(6)，然其中， $a_A$ 、 $a_B$ 、 $b_A$  與  $b_B$  變更為未知常數係數， $\text{Factor}(t)$  變更為一未知影響因子。

以下小題若使用者使用檢定作為判斷工具，則假設檢定型二錯誤 (Type II Error) 不會發生。

- (d) (3 points) 考慮該投資人僅知市場之證券 A、證券 B 與證券 C 中，有某二證券線性共用一已知因子，然其二證券並不知道究竟為何。即下方方程式中有一條不成立：

$$\text{Price}_A(t) = a_A + b_A \text{Factor}(t) + \epsilon_A(t), \quad (7)$$

$$\text{Price}_B(t) = a_B + b_B \text{Factor}(t) + \epsilon_B(t), \quad (8)$$

$$\text{Price}_C(t) = a_C + b_C \text{Factor}(t) + \epsilon_C(t), \quad (9)$$

其中， $a_i$  ( $i = A, B, C$ ) 與  $b_i$  ( $i = A, B, C$ ) 為未知常數係數， $\text{Factor}(t)$  為一已知影響因子， $\epsilon_A(t)$  與  $\epsilon_B(t)$  為誤差。

- (e) (3 points) 考慮該投資人僅知市場之證券 A、證券 B 與證券 C 中，有某二證券線性共用一未知因子，然其二證券並不知道究竟為何。即關係仍如方程式 (7-9)，存在一條方程式錯誤，然  $\text{Factor}(t)$  變更為一未知影響因子。

Total for Question 11: 21

12. 請根據產業價值鏈資訊平台 (<https://ic.tpex.org.tw/>) 及答案 11. (b)–11. (e) 撰寫程式，並各自實作一例。

- (a) (5 points) 請合理挑選兩可能共用同一因子的證券，同時假設為真實共用一可能因子 (需選定可取得數據之因子)。根據答案 11. (b) 完成程式並展示一例。
- (b) (5 points) 基於答案 12. (a) 所設定之前提與環境，在不使用因子數據的情況下，根據答案 11. (c) 完成程式並展示之。
- (c) (5 points) 請混入一個明顯不與答案 12. (a) 所挑選之證券共用因子的證券，並根據答案 11. (d) 完成程式並展示之。
- (d) (5 points) 基於答案 12. (a) 所挑選之證券與答案 12. (c) 之證券，根據答案 11. (e) 完成程式並展示之。

Total for Question 12: 20

13. 假設有兩根據課程方式所設計之配對交易投資組合，其為組合 A 與組合 B，令兩組合價值變異數分別為  $\sigma_A^2$  與  $\sigma_B^2$ 、平均價值分別為  $\mu_A$  與  $\mu_B$ ，且考量具有  $\sigma_A^2/\mu_A > \sigma_B^2/\mu_B$  關係。請說明進行本課程的隨機搜尋時會優先選擇何者，而又何者進行交易較能有可觀的獲利。

- (a) (0 points) 當未知是否滿足單因子的線性迴歸前提時，會考慮優先選擇何組合進行交易 (A 或 B)。
- (b) (0 points) 然，當已知滿足單因子的線性迴歸前提時，會考慮優先選擇何組合進行交易 (A 或 B)。
- (c) (5 points) 說明題目 13. (a) 與題目 13. (b) 的關聯 (相同或不相同之原因)。

Total for Question 13: 5

## 延伸思考

14. 現有三可交易證券，證券 A、證券 B 與證券 C，並考慮有線性共用二因子，且不知共用之係數。即如下關係：

$$\text{Price}_A(t) = a_A + b_{1A} \text{Factor}_1(t) + b_{2A} \text{Factor}_2(t) + \epsilon_A(t), \quad (10)$$

$$\text{Price}_B(t) = a_B + b_{1B} \text{Factor}_1(t) + b_{2B} \text{Factor}_2(t) + \epsilon_B(t), \quad (11)$$

$$\text{Price}_C(t) = a_C + b_{1C} \text{Factor}_1(t) + b_{2C} \text{Factor}_2(t) + \epsilon_C(t), \quad (12)$$

其中， $a_i$  ( $i = A, B, C$ ) 與  $b_{ij}$  ( $i = 1, 2, j = A, B, C$ ) 為未知常數係數， $\text{Factor}_1(t)$  與  $\text{Factor}_2(t)$  為一未知影響因子， $\epsilon_A(t)$  與  $\epsilon_B(t)$  為誤差 (errors)。請描述進行配對交易時的權重與現金流。

- (a) 3] 請寫下三證券價格序列受二因子的線性迴歸式 (如  $y(t) = X(t)\beta + \epsilon(t)$  形式)。

以下小題考慮投資人可使用歷史股價數據，即  $\text{Price}_A(t = 1, \dots, T)$ 、 $\text{Price}_B(t = 1, \dots, T)$  與  $\text{Price}_C(t = 1, \dots, T)$  數據。

- (b) (5 points) 請解釋證券 A 之價值是否能改寫為不包含任何因子的表達式。

- (c) (6 points) 請建構不受任何因子影響的投資組合。
- (d) (6 points) 請計算投資人於時間  $t = t_0$  進行買進，於時間  $t = t_1$  進行平倉，所產生的實質現金流 (需考慮放空所需要之資金) 與投資組合價值，並衡量其報酬率 (實質獲利除以本金)。
- (e) (3 points) 請說明最終能獲利的前提。

Total for Question 14: 20

15. 定義  $R^2 = 1 - (\hat{\sigma}_\epsilon^2 / \hat{\sigma}_A^2)$ ，其中  $\hat{\sigma}_A^2$  與  $\hat{\sigma}_\epsilon^2$  分別為方程式 (10) 中  $\text{Price}_A(t)$  與誤差  $\epsilon_A(t)$  的變異數  $\sigma_A^2$  與  $\sigma_\epsilon^2$  之樣本估計值。請說明  $R^2$  值大小與因子共用程度的關係，並完成相關程式。
- (a) (5 points) 解釋  $R^2$  值大小與因子共用程度的關係。
  - (b) (3 points) 當未知所欲投資之投資組合是否有線性的共用因子時，將  $R^2$  作為衡量是否有該關係時，該如何使用。
  - (c) (10 points) 於現有程式架構下新增一個以此  $R^2$  進行隨機搜尋及建構配對交易的功能。

Total for Question 15: 18

16. 請挑選一個布林通道 (Bollinger Bands, BBands) 與隨機指標 (Stochastic Oscillator, KD) 以外的技術指標，並根據該指標之公式說明其使用前提，實作於下有程式架構。

**Solution:** 本題以 BBands 做為範例。

- (a) (0 points) 挑選一個技術指標 (BBands 與 KD 以外)，並寫下定義，及說明為何挑選該指標。

**Solution:** 令一投資組合 A 價格  $\text{Price}_A(t)$ ，定義簡單移動平均

$$\text{MA}(t; N) \equiv \frac{1}{N} \sum_{t_0=t-N+1}^t \text{Price}_A(t)$$

與其於時間  $t$  時，近  $N$  筆之樣本變異數序列為  $\sigma_N^2(t)$ 。根據維基百科中 BBands 條目 ([https://en.wikipedia.org/wiki/Bollinger\\_Bands](https://en.wikipedia.org/wiki/Bollinger_Bands))，在給定了簡單移動平均線樣本數  $N$  (觀察範圍) 與標準差數量  $K$  (軌道寬度) 下，BBands 可定義三條軌道 (線) 如下：

$$\begin{aligned} \text{middleBB}(t) &\equiv \text{MA}(t; N), \\ \text{lowerBB}(t) &\equiv \text{middleBB}(t) - K\sigma_N^2(t), \\ \text{upperBB}(t) &\equiv \text{middleBB}(t) + K\sigma_N^2(t). \end{aligned}$$

- (b) (10 points) 請藉由該技術指標的定義推論該指標該進行交易的訊號為何，同時說明適合使用該指標的證券所應需具備的性質。

**Solution:**

**闡述指標意圖** 首先，由  $\text{middleBB}(t)$ 、 $\text{lowerBB}(t)$  與  $\text{upperBB}(t)$  之定義，可知該指標考慮了應用於其之證券其於時間  $t$  之股價分布有一隨機分佈  $D(t)$ ，而三條軌道則提供一個參考範圍。再者，由定義上關於兩軌道  $\text{lowerBB}(t)$ 、 $\text{upperBB}(t)$  與  $\text{middleBB}(t)$  上下對稱，可知該  $D(t)$  應為對稱分佈，或其通常股價範圍應為該二軌道之間，而於兩軌道區間外則為罕見事件。

**保留原意圖下之延伸改良** 這也意味著，想  $\text{lowerBB}(t)$  與  $\text{upperBB}(t)$  一起向上 (或向下) 調整是可行的 (兩軌道調整幅度並不一定相等)。且如果使用者認為分佈並不對稱，也可以對單一軌道或多個軌道進行修改。

**產生交易訊號之資訊** 若  $\text{middleBB}(t)$ 、 $\text{lowerBB}(t)$  與  $\text{upperBB}(t)$  需具備提供交易訊號之用途，則因三軌道必然無法交錯，導致需藉由軌道外之相關資訊以提供訊號。而由於三軌道為股價之函數，且  $\text{LU}(t) \equiv [\text{lowerBB}(t), \text{upperBB}(t)]$  區間描述了股價之正常範圍，故可考慮以股價與其之關聯作為提供訊號之依據。又由於將  $\text{LU}(t)$  區間視為正常範圍，這導致了  $\text{middleBB}(t)$  與股價關係無法提供交易訊號。即僅能以股價與  $\text{lowerBB}(t)$  及  $\text{upperBB}(t)$  之關聯產生交易訊號。故此，我們考慮股價與  $\text{lowerBB}(t)$  之交錯與股價與  $\text{upperBB}(t)$  之交錯作為交易訊號。

**交易訊號外之相關訊息 (補充)** 股價是否長期位於  $\text{middleBB}(t)$  (簡單移動平均) 之上 (或下)，仍可提供近期股價之成長趨勢 (持續上漲或持續下跌)。

**產生交易訊號之方法 (開倉)** 由於我們將該  $\text{lowerBB}(t)$  視為正常範圍之下界，故可考慮股價與  $\text{lowerBB}(t)$  之交錯進行買入的開倉時機點。因 BBands 考量了股價之樣本變異數與時間相依，代表該指標已考慮股價與時間之相依性。故我們也可考慮股價之變動具有連續性，代表股價與  $\text{lowerBB}(t)$  之交錯進行買入的開倉時機點須為股價向上穿越  $\text{lowerBB}(t)$ ，否則將有持續下跌之風險。類似地，我們也可以定義股價向下穿越  $\text{upperBB}(t)$  作為放空之時機。

**產生交易訊號之方法 (平倉)** 由於  $\text{LU}(t)$  之間皆為股價之正常範圍，故無法預期股價會落於區間中的哪一個價格 (任何一個該區間之股價皆可能)。然而，落於正常範圍之上半部  $\text{LM}(t) \equiv [\text{lowerBB}(t), \text{middleBB}(t)]$  與正常範圍之下半部  $\text{MU}(t) \equiv [\text{middleBB}(t), \text{upperBB}(t)]$  為正常範圍下中會發生之約大半可能範圍。故，可約定上述開倉後之平倉時機定義為股價進入此兩部分中之某一部分之中 ( $\text{LM}(t)$  或  $\text{MU}(t)$ ) 時進行。又因若定義做多之平倉區間為  $\text{LM}(t)$  與做空之區間為  $\text{MU}$ ，皆會於開倉時即刻平倉，無法獲利。據此，我們定義做多時之平倉時機為進入上半部，做空之平倉時機為進入下半部，即穿越  $\text{middleBB}(t)$  之時刻。

**所依賴之股價特性** 由於損益之計算依賴於開倉與平倉時之價差，這導致了 (1) 當股價向上穿越  $\text{lowerBB}(t)$  進行開倉 (做多)，而平倉時價格不得低於買入價格；(2) 當股價向下穿越  $\text{upperBB}(t)$  進行開倉 (做空)，而平倉時價格不得高於買入價格。因此，必需對於股價於開倉與平倉之間進行一定限制。由於於時間  $t$  時進行開倉是依據該時之  $\text{lowerBB}(t)$  數值 (或  $\text{upperBB}(t)$  數值)，故可推斷平倉之區間亦為該時之  $\text{MU}(t)$  (或  $\text{LM}(t)$ )，或認為未來之區間下界 (或上界) 不應低於 (或高於) 開倉之數值。這意味著，該指標並未考量股價為具有趨勢性，但可具有一定幅度的周期性，即均值回復之特性。

(c) (5 points) 當應用於具有趨勢型證券性質的證券與均值回復性質的證券分別會有何現象。

**Solution:** 根據答案 16. (c) 中對於該指標所依賴之股價特性敘述，可知當股價趨勢明確時，將可能造成虧損；而當股價具有小於某一定幅度的周期性時，將可能獲利。

**根據股票性值修改使用方式** 若股價具有趨勢性時，可在修改開倉與平倉之定義下繼續使用該指標。於正趨勢時，開倉僅定義為向上穿越  $\text{middleBB}(t)$  時進行做多，而向下穿越  $\text{upperBB}(t)$  時進行平倉；反之亦然。我們僅需事先檢定是否具有趨勢性即可，該技術指標不會因而無法使用。如若無法知道是否具有趨勢性，則可以藉由設定兩種  $K$  值， $K_1 \leq K_2$ ，以製作具有五條軌道之布林通道。即，如下五條軌道：

$$\text{middleBB}(t) \equiv \text{MA}(t; N),$$

$$\text{lowerBB}_i(t) \equiv \text{middleBB}(t) - K_i \sigma_N^2(t), i = 1, 2,$$

$$\text{upperBB}_2(i) \equiv \text{middleBB}(t) + K_i \sigma_N^2(t), i = 1, 2.$$

當股票價格位於  $\text{lowerBB}_1(t)$  與  $\text{upperBB}_1(t)$  之間，視為均值回復之證券進行操作；當股票價格於  $\text{lowerBB}_2(t)$  以下或  $\text{upperBB}_2(t)$  以上，視為趨勢型證券進行操作。

- (d) (10 points) 請於現有程式下撰寫答案 16. (b) 所描述之技術指標，並將其應用於答案 15. (c) 所隨機搜尋並建構之投資組合上。
17. (10 points) 請挑選一個與答案 16. 的技術指標類型之相似指標，再完成一次題目 16.。並闡述其兩指標的應用前提(場合)差異。

**Solution:** 如隨機指標 (stochastic oscillator, KD) 雖與布林通道皆為均值回復類型之指標，但其未成熟隨機值 (raw stochastic value, RSV) 要求了應用指標之標的價格不得有盤整價格收斂之現象。

Total for Question 17: 10