

※ 考生請注意：本試題不可使用計算機。請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

一、選擇題 (每題 4 分，共 12 題，共計 48 分) 請選出最佳答案

1. 現有一顆六面的公正骰子，初投 2 次，都出現數字 5，請問投第 3 次出現數字 5 的機率為何？

- (A) 比出現其他數字的機率來得大
- (B) 比出現其他數字的機率來得小
- (C)  $1/6$
- (D)  $1/216$ 。

2. 已知國內某一大學來自南部的學生所占全校學生的比率是 60%，現從此大學隨機抽出 8 位學生，請問在抽出的樣本中，出現至少 7 位來自南部的學生之機率為何？

- (A) 0.1064
- (B) 0.0896
- (C) 0.0168
- (D) 0.9104。

3. 對一個連續的均勻分配 (uniform distribution) 來說，機率密度函數  $f(x)$  的高度

- (A) 不能大於 1
- (B) 對於  $x$  的每個值是相同的
- (C) 隨著  $x$  的增加而增加
- (D) 隨著  $x$  的增加而減小。

4. 如果從兩個群體各隨機抽取獨立的大樣本，則兩個樣本均值之差距的抽樣分佈

- (A) 可以由帕松分佈 (Poisson distribution) 近似
- (B) 變異數為 1
- (C) 可以由常態分佈近似
- (D) 平均數為 1。

5. 某大學生認為，統計學期末考試的平均分至少是 78 分，他計劃抽樣，以測試他的想法。請問正確的假設檢定是？

- (A)  $H_0: \mu < 78$        $H_a: \mu \geq 78$
- (B)  $H_0: \mu \leq 78$        $H_a: \mu > 78$
- (C)  $H_0: \mu \geq 78$        $H_a: \mu < 78$
- (D)  $H_0: \mu > 78$        $H_a: \mu \leq 78$ 。

6. 下表列出國內某大公司研究所學歷與大學學歷員工的薪水差異，

	研究所學歷	大學學歷
抽樣數	64	36
薪水樣本平均 (NTD1,000)	44	41
母體變異數	128	72

根據上表，兩母體平均數間差距之95%信賴區間等於？

- (A) 1.04 to 4.96  
 (B) -4.84 to 11.84  
 (C) -1.96 to 1.96  
 (D) -0.92 to 6.92。

7. 決定係數 (coefficient of determination,  $R^2$ ) 較大顯示觀察值更接近於

- (A) 解釋變數的平均值  
 (B) 被解釋變數的平均值  
 (C) 最小平方法的配適線  
 (D) 原點。

8. 給定下列資料，其中誤差平方和 (SSE)=0.4130。

被解釋變數 (y)	解釋變數 (x)
15	4
17	6
23	2
17	4

則此迴歸的斜率為

- (A) -2.4  
 (B) -1.2  
 (C) 0.707  
 (D) -1.5。

9. 當被解釋變數為二元 (0, 1) 之變量，下列何種估計模型最不適用？

- (A) 最小平方法(OLS)  
 (B) Probit 模型  
 (C) Logit 模型  
 (D) Tobit 模型。

10. 下列哪一個原因最不可能讓迴歸的解釋變數具有內生性 (endogeneity) ?

- (A) 遺漏變數
- (B) 異質變異數
- (C) 衡量誤差
- (D) 解釋變數與被解釋變數互為因果。

11. 請問最佳線性不偏估計式 (BLUE) 當中的最佳 (Best) 意指甚麼？

- (A) 用最小平方法得到的誤差平方總和 (SSE) 最小
- (B) 用最小平方法得到的誤差平方總和最大
- (C) 在線性不偏的估計式中變異數最小
- (D) 在線性不偏的估計式中的係數最大。

12. 隨機選取 15 人並都給予品牌 A 和品牌 B 的果汁。2 個人喜歡品牌 A，13 個人喜歡品牌 B，我們想知道客戶是否鍾愛某一品牌更勝於另一者。當信賴水準=95%，虛無假設應該

- (A) 被拒絕
- (B) 不被拒絕
- (C) 修改
- (D) 以上都不正確。

## 二、問答題 (6 小題、共 52 分)

假設您取得義大利北部的米蘭與中部佛羅倫斯兩地之房價交易資料，於 2016 年 3 月與 2017 年 1 月都有交易的房子共 360 筆，經整理，這些房子於 2016 年 3 月的平均價格如下：

	米蘭	佛羅倫斯
樣本數	160	200
平均價格	30 萬歐元	32 萬歐元

若您把這些資料拿去進行以下的迴歸分析：

$$y_i = \alpha + x_i\beta + e_i,$$

其中， $y_i$  為房價， $x_i$  為虛擬變數 (若房屋  $i$  的地點位於佛羅倫斯其值為 1，若位於米蘭其值為 0)，且假設資料都符合簡單迴歸的古典假設，請回答下列問題：

- (A) [12 分] 請證明斜率  $\beta$  的估計式恰為兩樣本平均價格的差距。
- (B) [6 分]  $\beta$  的估計值為？
- (C) [12 分] 若您將這些房子在兩地於 2017 年 1 月的平均價格也整理出來，發現平均價格如下

	米蘭	佛羅倫斯
樣本數	160	200
平均價格	29 萬歐元	29 萬歐元

且假設您經過仔細考察，也發現兩地房價均下跌可能都肇因於義大利的總體景氣不佳，但佛羅倫斯房價之所以下降這麼多，乃肇因於 2016 年 8 月位於義大利中部的大地震。若想研究地震是否造成佛羅倫斯的房價下跌，請問您要如何以迴歸分析進行驗證？(請仔細寫下迴歸式，並說明變數定義、虛無假設、以及拒絕虛無假設的條件)

(D) [6 分] 承上題，因地震造成房價下跌的平均值為多少？

(E) [8 分] 若您在使用 2016 年 3 月的資料分析時，懷疑兩地房價的變異數不同。請問在進行迴歸分析後，該如何進一步檢定？(請仔細寫下檢定的方式，虛無假設，以及拒絕虛無假設的條件)

(F) [8 分] 承上題，若兩地房價的變異數不同，則使用最小平方法會可能產生什麼問題？這些問題，又該如何修正？