

香港統計學會考試



統計學普通證書，2005

試卷 I

時間：3 小時

考生可嘗試**所有**題目  
每一道題或其小題的分數標明在括弧中  
試卷 100 分為滿分  
至少取得 50 分才能獲得通過

試場可提供畫圖紙和標準表格

考生可按照香港統計學會列載於考試指引的規例(文件 Ex1)，使用計算器。

1

普通證書試卷 I 2005

---

本試卷共有 4 頁，採用單面印刷

本頁為第一頁

**題目問題 1** 從第二頁開始

試卷共有八個**題目**

©RSS 2005



問題 1 到 3 與下列的情況有關

某國分爲 50 個行政區。在 2006 年，教育部計劃進行一項對 2003 年期間離校中學生的職業和職業計劃的抽樣調查，希望調查大約 10% 的有關總體。

下列關於該國內每間中學的資料可供使用：規模(在 2006 年 1 月 31 日的學生數目)；地區；按照學生在 2003 年離校日期排列的姓名和聯絡詳情。

1. 這題的答案應與建議的調查有關

- (i) 給出建議兩個不同的二階段抽樣方法去選擇一個合適的樣本。每個方法應採用不同組合的簡單隨機抽樣、分層抽樣、系統抽樣和分組抽樣，這四個抽樣方法的每一個只可採用一次。你可假設適合你用途的抽樣框存在。 (6)
- (ii) 就你每一個建議的抽樣方法，簡要評論它對這調查的一些優點和缺點。 (6)

2.2. 起草一份簡短的自我完成問卷，以從樣本成員獲取得下列的統計資料。

- 離開中學時的月份、年份和年齡
- 2004 和 2005 年每年的職業和從事的地區。(如果被訪者在當年有多過一種職業，你應指示他或她給出提供有關主要職業的資料和工作的日期。請注意到作為學生也被視爲一種職業；被訪者應被要求指出所修讀的主要科目。)
- 長期的職業計劃 (12)

3.3. 從被訪者獲得如問題 2 詳述的資料，以及被訪者的聯繫詳情，將以電子方式儲存。指出若有那個反饋的資訊要編碼回應，哪些回應須給予編碼及：

設計一個適合這用途的數據庫。你的答案應列出域，並給予適當的域名，以及域類型和合適的寬度。

或 設計一個適合這用途的試算表。你的答案應列出變數，並給予適當的變數名，以及它的類型和合適的儲存格寬度。 (12)

問題 4 到 8 與下列建議的統計調查設計有關。調查的主要目的是要測定一個大部份人口在農村經歷貧窮的小國裏，水果和蔬菜價格隨着時間和在不同地區的變化。

這國家分爲 7 個地區，4 個是多山的和 3 個是沿岸的。一多山地區和一海岸已被選爲具有代表性。調查將包括一些在這兩個地區內農村地方舉辦的市場，這些市場被認爲屬於兩分層的其中之一：大市場和小市場。市場樣本會從每個地區內的每一層中抽取，每個市場的本地代表會被訪問生產和供應的背景資料。

在一段四個月的時間裏，每個樣本市場會每月被探訪一次。在每次探訪，會從透過訪問一些商販通過訪問來收集在當天售賣的一些選定的水果和蔬菜的價格。在市場出現的商販和售賣的農產品很可能每天都改變。

4. (i) 說出使用像以上描述的非隨機方法去選擇兩個地區的兩個優點和兩個缺點。 (6)
  - (ii) 在建議的調查中，討論它的困難去
    - (a) 選擇商販做訪問，
    - (b) 決定應獲得那個種水果和蔬菜的價格。在每種情形下，建議怎樣對付這些困難。 (10)
- 5.5. (i) 關於每個市場關於它售賣的農產品的生產和供應、營業次數和擁有攤擋的商販數 目的背景資料是需要的，應向當地代表取得這些資料。對選擇合適的當地代表和從他們收集需要的資料，你會給予什麼忠告？ (6)
- (ii) 簡要評論在今年進行的調查中，獲取的精確細節的如去年產品價格的困難，例如要取得去年的產品價格。 (4)
6. 解釋訪問員偏差的含義。在調查收集價格的過程中，以例子說明它是如何發生的和建議如何減少這些偏差。 (10)

7. 去年在某國的另一地區已完成一個多方面有關被廣泛食用的主要蔬菜的研究。下表的成本和價格以該國的貨幣計算。該表列出在先前的研究中求得的每千克蔬菜價格的標準差， $s_i$ 。它亦列出在建議的調查中，大市場和小市場在海岸地區的數目  $N_i$  和從每層抽選一個市場的預期成本  $c_i$  (以當地貨幣計算)。

| 層   | 市場數目 ( $N_i$ ) | 蔬菜價格的標準差 ( $s_i$ ) | 抽選一個市場的成本 ( $c_i$ ) |
|-----|----------------|--------------------|---------------------|
| 大市場 | 27             | 0.05               | 15                  |
| 小市場 | 53             | 0.07               | 12                  |

- (i) 爲了達到整體樣本大約有 8 個市場，使用均勻抽樣比去計算每層樣本中所需的市場數目。若以這些樣本規模去抽選市場，全部成本是多少？ (4)
- (ii) 使用最佳分配方法抽選樣本規模  $n_i$  去估計一個均值，可令估計的方差爲最小，爲此要按照比例  $N_i s_i / \sqrt{c_i}$  抽取  $n_i$ 。若要在這地區抽選市場的全部成本不多過 105 個當地貨幣，用這方法計算在每一層需要包括在樣本內的市場數目。 (8)
- (iii) 以上(i)和(ii)計算樣本容量規模的兩個方法，你認爲哪個較佳，原因爲何？ (4)
8. 統計調查收集的資料是用來獲得水果和蔬菜的價格指數。通過考以你熟悉的一個官方價格指數作依據，或其他方法，建議如何進行編製。你的答案應指出爲指數作出詳細的定義時，需要作什麼決定，是否需要收集其他資料以計算該指數，以及指數應如何隨着時間而更新。 (12)