

類 科：統計

科 目：抽樣方法

考試時間：2 小時

座號：_____

※注意：(一)可以使用電子計算器，須詳列解答過程。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

(三)本科目除專門名詞或數理公式外，應使用本國文字作答。

- 一、(一)假設 $F = \{u_1, u_2, u_3, u_4\}$ 是一個僅僅包含四個元素的小規模有限母體 (finite population)，而四個元素的研究變數 (study variable) 值分別為 $y_1=1$ 、 $y_2=3$ 、 $y_3=3$ 以及 $y_4=9$ 。我們採用不置回的簡單隨機抽樣 (simple random sampling without replacement) 從母體 F 之中抽出樣本大小 (sample size) 為 $n=2$ 的樣本組合，並以 Y_1, Y_2 來表示此樣本組合中的兩個樣本數據 (註：採用大寫英文字母 Y ，表示樣本數據皆為隨機變數)。試求 Y_1 與 Y_2 的聯合機率分佈 (joint probability distribution) 以及 Y_1 與 Y_2 的共變異數 (covariance)。(15 分)
- (二)假設前一小題之中抽出樣本大小為 $n=3$ 的樣本組合，並以 Y_1, Y_2, Y_3 來表示此樣本組合中的三個樣本數據。試求 Y_1 與 Y_3 的共變異數。(5 分)

- 二、(一)假設某一個有限母體可被完整地切割為 L 個互不相交的層別 (strata)，其中第 h 層之層大小 (stratum size) 為 N_h ，第 h 層之第 j 個元素的研究變數值為 y_{hj} ， $h=1, \dots, L$ ， $j=1, 2, \dots, N_h$ 。母體大小 (population size)、母體平均數 (population mean) 以及母體變異數 (population variance) 的數學式分別為

$$\text{母體大小： } N = N_1 + N_2 + \dots + N_L ;$$

$$\text{母體平均數： } \mu = \frac{1}{N} \sum_{h=1}^L \sum_{j=1}^{N_h} y_{hj} ;$$

$$\text{母體變異數： } \sigma^2 = \frac{1}{N} \sum_{h=1}^L \sum_{j=1}^{N_h} (y_{hj} - \mu)^2 。$$

其次，母體中第 h 層之層平均數 (stratum mean) 以及層變異數 (stratum variance) 的數學式分別為

$$\text{第 } h \text{ 層之層平均數： } \mu_h = \frac{1}{N_h} \sum_{j=1}^{N_h} y_{hj} ;$$

$$\text{第 } h \text{ 層之層變異數： } \sigma_h^2 = \frac{1}{N_h} \sum_{j=1}^{N_h} (y_{hj} - \mu_h)^2 。$$

試證明下列等式：

$$\sigma^2 = \sum_{h=1}^L W_h \sigma_h^2 + \sum_{h=1}^L W_h (\mu_h - \mu)^2 ,$$

其中 $W_h = \frac{N_h}{N}$ 為第 h 層的層比重 (stratum weight)， $h=1, \dots, L$ 。(10 分)

- (二)使用分層隨機抽樣 (stratified random sampling) 之後，抽樣者應如何推估 (未知其值的) 母體平均數？需要什麼假設條件？請詳細說明。(4 分)
- (三)在使用分層隨機抽樣之前，抽樣者必須事先決定將要採用何種較為適當的配置，例如比例配置 (proportional allocation)、尼門配置 (Neyman allocation)。請詳細說明在使用事後分層 (post-stratification) 的方法之前，抽樣者是否也必須事先決定採用何種配置？(5 分)
- (四)請詳細說明在什麼情況之下，抽樣者需要借助於雙重抽樣 (double sampling) 之方法，來進行分層隨機抽樣並推估母體的平均數 (或是推估母體的其他參數)。(5 分)

(請接背面)

類 科：統計
科 目：抽樣方法

三、某水果商訂購了一卡車的椰子，並打算採用比值估計（ratio estimation）之方法來推估整輛卡車裝載所有的椰子如果全數被剖開之後的椰子汁總重量。該水果商採用不置回的簡單隨機抽樣之方法從整輛卡車的椰子之中抽出了 10 顆椰子。試說明該水果商應該如何利用被抽出的 10 顆椰子來取得樣本數據以及如何進行比值估計，並寫出點估計量的數學式。（15 分）

四、迴歸分析（regression analysis）課程中的最小平方估計量（least squares estimator），在抽樣方法的理論中有何可用之處？請詳細說明並寫出相關的估計量數學式。（10 分）

五、(一)某一間大學的學術部門打算採用二階段集群抽樣（two-stage cluster sampling）之方法來推估校內全體專任教師在前一個學年度的學術論文平均產量（亦即，推估任何一位專任教師在前一個學年度的學術論文平均發表篇數），並且將校內的每一個科系視為一個集群。試就比值估計（ratio estimation）以及不偏估計（unbiased estimation）兩種情況分別說明如何進行二階段集群抽樣以及如何推估，並寫出點估計量的數學式。（13 分）

(二)假設前一小題之中的學術部門打算改變為採用集群抽樣（cluster sampling）之方法來推估校內全體專任教師在前一個學年度的學術論文平均產量，並且依舊將校內的每一個科系視為一個集群。試就比值估計（ratio estimation）以及不偏估計（unbiased estimation）兩種情況分別說明如何進行集群抽樣以及如何推估，並寫出點估計量的數學式。（13 分）

[註] 集群抽樣又稱之為單階段集群抽樣（single-stage cluster sampling）

六、在什麼情況之下，使用系統抽樣（systematic sampling）之方法來推估母體的平均數，其推估之效果會優於使用簡單隨機抽樣的推估效果？請詳細說明。（5 分）