

111年公務人員高等考試三級考試試題

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。

(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、南興社區疑似爆發某流行病之群聚感染，為規劃後續防疫措施，亟需先行瞭解該社區染病人數。現有兩種檢測方法可用於檢測是否染病，檢測方法 A 可迅速得到結果，但其精確程度較不理想，而檢測方法 B 雖可提供相當精確之結果，但其所需檢驗程序繁複，若要爭取時效，僅能施測 30 位。今先以方法 A 對該社區所有之 500 位居民進行檢測，驗得其陽性比例為 0.6，再以簡單隨機抽樣取出不放回之方式，由全社區之居民中選擇 30 位以方法 B 進行施測，驗得其陽性為 15 人，而這 30 位受測者以方法 A 施測之陽性人數為 13 人，且 30 人中有 26 位以方法 A 及方法 B 檢驗之結果相同。請回答下列問題：

(一)若僅以方法 B 施測所得之結果推估本社區染病人數，請問其估計值及 95%信賴區間為何？(5 分)

(二)是否可以整合兩種方法推估本社區染病人數？若您認為可以，請敘述您採用的推估方式、本社區染病人數推估值及其 95%信賴區間，同時請問在上述兩種推估方式中，您會採取那一種及您的理由。(15 分)

二、在某大學所進行的某滿意度調查中，樣本為以簡單隨機抽樣取出不放回所選擇出的 $n=400$ 位大學部學生，該校大學部學生共 5000 名，400 位學生依其年級分類及各年級滿意樣本比例整理如下表：

	大一	大二	大三	大四
樣本數 (x_i)	200	100	50	50
滿意比例 (p_i)	0.4	0.8	0.8	0.8

則全體學生滿意比例之不偏估計為 $\frac{200 \cdot 0.4 + 100 \cdot 0.8 + 50 \cdot 0.8 + 50 \cdot 0.8}{400} = 0.6$

(一)根據註冊組提供之資料，該校大學部學生之各年級比例如下：

	大一	大二	大三	大四
母體比例 (Q_i)	0.3	0.3	0.2	0.2

請檢定在 $\alpha=0.05$ 之水準下樣本比例與母體比例是否相符(請參考下表之卡方分配 95%百分位數)。(5 分)

自由度	1	5	10	20
$\chi^2_{0.95}$	3.84	11.07	18.31	31.41

(二)根據子題(一)之結論，請問滿意度 0.6 的結果是否應加以修正？請敘明理由，若您認為需要修正，請提供修正後之母體比例。(10 分)

三、金芒鄉是主要的芒果產區，鄉內農戶有專門種植芒果之專業果農，也有混種其他作物而僅部分種植芒果之農戶，專業芒果之農戶名冊可由農林漁牧業普查資料中獲得，但兼種芒果之農戶時有變動，無法經由普查資料中確認，惟為求精準推估金芒鄉本年度芒果產值，亦須將兼種芒果之農戶之芒果產值列入。在本鄉一共 500 戶農戶中，有 100 戶為專業種植芒果的農戶，而此 100 戶中，有 5 戶為大型農戶，95 戶為中小型農戶。本調查以下列方式選擇樣本農戶進行金芒鄉本年度之調查：5 戶大型專業芒果農戶全數加以調查。接下來在全金芒鄉之 11 個村中，以簡單隨機抽樣取出不放回之方式選擇其中 6 個村，每個樣本村中再以簡單隨機抽樣取出不放回之方式選擇 5 戶中小型專業芒果農戶，若該村中小型專業芒果農戶數不足或等於 5 戶，則該村中所有中小型專業芒果農戶全數加以調查。另外以簡單隨機抽樣取出不放回之方式在其餘農戶中選擇其中 50 戶作為樣本戶。

本調查各類農戶所得資料綜整如下：

(1)大型專業芒果農戶：平均值為 125 萬元，標準差為 20 萬元。

(2)中小型專業芒果農戶：6 個樣本村之資料如下：

村別	文新村	文景村	文佑村	金河村	同億村	同欣村
總戶數	16	4	3	12	2	13
平均值(萬元)	30	25	30	25	20	30
標準差(萬元)	5	10	6	5	8	8

(3)其餘農戶：50 戶樣本戶中，有 10 戶種植芒果，而這 10 戶之平均芒果產值為 12 萬元，標準差為 8 萬元。

請回答下列問題（請以萬元為單位作答）：

(一)請提出兩種不同估計方式，推估金芒鄉中小型專業芒果農戶之年度芒果總產值，以及其相對應之變異數估計，並比較您提出之方法於本調查之適用性。(20 分)

(二)若以其他農戶之 50 戶樣本戶中，10 戶有兼種芒果之農戶的平均芒果產值 12 萬元，做為全體兼種芒果農戶之平均芒果產值推估值，請問其變異數估計值為何？(5 分)

(三)請推估金芒鄉兼種芒果農戶之年度芒果總產值，並提出其變異數估計值。(10 分)

四、今有一母體，其中有三個單元，各單元編號 i 及其主要母體變量值 y_i 如下表：

i	1	2	3
y_i	7	5	3

若要從該母體中選擇 2 個單元作為樣本，在不考慮抽樣順序及可重複選擇之情形下，共有 6 種可能之樣本組合，考慮一抽樣設計為每組樣本組合 $s=(i, j)$ ，被選擇到之選擇機率為 $p(s)=(i+j)/24$ ，例如樣本組合 $(2,3)$ 之選擇機率為 $5/24$ ， $(1,1)$ 被選擇之機率為 $2/24$ ……等。

請回答下列問題（建議以分數計算及作答）：

(一)請問在本抽樣設計下，第 i 個單元被選到之包含機率（Inclusion Probability）若為 π_i ，則 π_2 為何？（5 分）

(二)請問若以樣本平均 $\bar{y} = \frac{\sum_{i \in s} y_i}{2}$ 作為母體平均 $\mu = \frac{\sum_{i=1}^3 y_i}{3}$ 之估計量，則 $\bar{y}$ 在本設計下推估 μ 之偏誤（Bias）為何？（5 分）

(三) $\bar{y}$ 在本設計下之均方誤差（Mean square error；MSE）為何？（5 分）

(四)請就本設計提出一不偏估計量，請問您提出之估計量在樣本組合分別為 $(2,3)$ 及 $(3,3)$ 的情況下之估計值為何？並證明或以本母體為例驗證該估計量為不偏。（10 分）

(五)請試述您要如何評估您所提出之不偏估計與 $\bar{y}$ 在本設計下孰優孰劣？（不須計算實際數值）（5 分）