

109 年特種考試地方政府公務人員考試試題

等 別：三等考試

類 科：統計

科 目：抽樣方法

甲、申論題部分

- 一、光明社區管理委員會主任委員希望能調查住戶對禁止住戶在社區中庭花園抽菸以免影響其他住戶的意見，因為該社區住戶高達 700 戶，分別分布在 A、B、C、D 四棟大樓，因此主任委員在各棟大樓各用簡單隨機抽樣取出不放回的方式選擇了 50 戶住戶加以調查其意向。各棟大樓戶數及贊成社區中庭花園禁菸之樣本比例如下：

大樓編號	戶數	贊成比例
A	100	0.8
B	300	0.4
C	200	0.5
D	100	0.7

請回答下列問題：

- (一)請問光明社區住戶贊成禁止住戶在社區中庭花園抽菸比例之不偏估計推估值。(5 分)
- (二)請問以上推估值在 95%信心水準下的最大估計誤差。(10 分)
- (三)在試行此一規定三個月後，主任委員想再做一次同樣的調查，請問在抽樣設計及總抽樣數均維持不變的情況下，可以如何來改進推估的精確程度？(10 分)

【解題關鍵】

《考題難易》★

《破題關鍵》分層隨機抽樣比例值估計與樣本數求算屬於課內常考基本題，多個年度皆有類似考題，今年 109 年身障抽樣有類似命題，可參考抽樣方法 P3-91 有相同範例，只要小心計算應可得分。本題的難度在於問到如何來改進推估的精確程度，有多種方式可以改進，但此題應僅是考慮配置問題。

【擬答】

- (一)採用分層隨機抽樣，估計母體比例值

$$p_{st} = \sum_{h=1}^4 W_h p_h = \frac{100}{700} \times 0.8 + \frac{300}{700} \times 0.4 + \frac{200}{700} \times 0.5 + \frac{100}{700} \times 0.7$$

$$= \frac{37}{70} = 0.5286$$

$$(二) s_{p_{st}} = \frac{1}{N} \sqrt{\sum_{h=1}^L N_h (N_h - n_h) \frac{\frac{n_h}{N_h} p_h q_h}{n_h}}$$

$$= \frac{1}{700} \sqrt{100(100 - 50) \frac{0.8 \times 0.2}{49} + \dots + 100(100 - 50) \frac{0.7 \times 0.3}{49}} = \frac{1}{7000} \sqrt{558.1633} = 0.0338$$

住戶贊成禁止住戶在社區中庭花園抽菸比例之 95%估計誤差為

$$Z_{0.025} \cdot s_{p_{st}} = 1.96 \times 0.0338 = 0.0662$$

- (三)原設計採分層抽樣之均等配置，在總樣本數不變下，可採用紐門配置，可改進推估的精確

度。即各分層取 $n_h = \frac{N_h S_h}{\sum_{h=1}^L N_h S_h} \times n$

$$\begin{aligned} \sum_{h=1}^L N_h S_h &= \sum_{h=1}^L N_h \cdot \sqrt{\frac{n_h p_h q_h}{n_h - 1}} \approx \sum_{h=1}^L N_h \sqrt{p_h q_h} \\ &= 100\sqrt{0.8 \times 0.2} + 300\sqrt{0.4 \times 0.6} + \dots + 100\sqrt{0.7 \times 0.3} \\ &= 332.7951 \\ n_1 &= \frac{100 \cdot \sqrt{0.8 \times 0.2}}{332.7951} \times 200 = 24.04, n_1 \text{ 取 } 24 \\ n_2 &= \frac{300 \cdot \sqrt{0.4 \times 0.6}}{332.7951} \times 200 = 88.32, n_2 \text{ 取 } 88 \\ n_3 &= \frac{200 \cdot \sqrt{0.5 \times 0.5}}{332.7951} \times 200 = 60.10, n_3 \text{ 取 } 60 \\ n_4 &= \frac{100 \cdot \sqrt{0.7 \times 0.3}}{332.7951} \times 200 = 27.54, n_4 \text{ 取 } 28 \end{aligned}$$

史上最強
線上申論批閱系統

志光 × 保成 × 學儒

3Q 線上 練題批閱

多位鑽石級名師
線上服務

囊括地方特考、高普考兩大國考，重要考科超過20科一網打盡國考必考題，
暢快提升申論答題力

3Q 線上申論批改，讓你體驗後SAY 3Q

Question 出好題
頂尖名師專業出題，題目緊扣命題趨勢，即便宅在家，也能練好題、答好題、破好題。

Quality 高品質
精選國考師資團隊，品質絕對有保證，系統化出題，題題精選，線上批改，科科重點。

Quietly 更安靜
就算不出門，在家也能練習申論，輕鬆舒適又自在，熟悉又安靜的環境，線上批改超彈性。

■完整課程資訊詳洽全國志光・保成・學儒門市■

二、某鎮農會欲估算該鎮豬農養豬飼料之花費，以簡單隨機抽樣取出不放回的方式於該鎮的 10 戶豬農選擇了 5 戶，詢問其過去一季平均養豬隻頭數及飼料花費。調查所得資料如下：

豬農編號	平均養豬隻頭數	飼料花費(仟元)
1	200	850
2	400	1600
5	100	450
8	50	300
10	250	800

根據此一調查結果，請問你會如何估算該鎮豬農平均每隻豬隻每季飼料花費？請解釋你的想法（可由估計及／或抽樣設計的觀點），並提出推估結果在 95%信心水準下的最大估計誤差。（15 分）

【解題關鍵】

《考題難易》★★

《破題關鍵》比率群集的概念題，是近年來試卷上的常客，也常會是決勝關鍵。今年度 109 年高考與去年 108 年地特考了簡單群集與比率群集的比較，相較之下，本題便屬於較為基礎的試題，但考生需加強描述抽樣方法的能力，請參考抽樣方法 P.12-12 頁關於抽樣設計的描述。

【擬答】

(一)此調查是將母體分為 10 個群集，以簡單隨機抽樣抽出 5 個群集為抽樣單位，觀察其樣本的情況，為群集抽樣。本例欲估計母體平均數，可採比率群集抽樣法即是群集抽樣法中執行比率估計的方法，比率估計法最主要的想法是採用輔助變數來增加估計的準確度，所以在這兩種方法下，皆是將各個集群的母體大小 M_i 視為輔助變數，用以提升推估集群母體參數準確度的估計方法。

$$(\text{二}) \bar{y}_t = \frac{Y_1 + Y_2 + \dots + Y_5}{5} = \frac{850 + 1600 + \dots + 800}{5} = 800$$

$$\bar{M} = \frac{M_1 + M_2 + \dots + M_5}{5} = \frac{200 + 400 + \dots + 250}{5} = 200$$

$$\bar{y}_{rc} = \frac{\bar{y}_t}{\bar{M}} = \frac{800}{200} = 4 \text{ (仟元)} \xrightarrow{\text{估計}} \bar{Y}$$

$$s_c^2 = \frac{\sum_{i=1}^n (Y_i - \bar{y}_{rc} M_i)^2}{n-1} = \frac{\sum_{i=1}^n Y_i^2 - 2\bar{y}_{rc} \sum_{i=1}^n Y_i M_i + \bar{y}_{rc}^2 \sum_{i=1}^n M_i^2}{n-1}$$

$$= \frac{4215000 - 2 \times 4 \times 1070000 + 4^2 \times 275000}{5-1} = 13750$$

$$s_{\bar{y}_{rc}} = \frac{1}{\bar{M}} \sqrt{(1-f) \frac{s_c^2}{n}} = \frac{1}{200} \sqrt{(1-\frac{5}{10}) \frac{13750}{5}} = 0.1854$$

該鎮豬農平均每隻豬隻每季飼料花費之 95% 估計誤差為

$$Z_{0.025} \cdot s_{\bar{y}_{rc}} = 1.96 \times 0.1854 = 0.3634$$

志光 × 保成 × 學儒

快速上榜

下一個上榜就是你!! 考取見證

一年考取 鄭○賢 109 普考經建行政 狀元 109 高考經建行政 探花 參加奪榜/特訓班的優勢是在考前幾個月逼自己進入備戰狀態，密集且快速的把前面上課的內容完整複習，並且每天固定在表定時間內寫完題目，加上眾人聚在一起凝聚出考前衝刺的氛圍，讓自己更能專注、不懈怠。	5個月考取 王○哲 109 高普考財稅行政 雙料金榜 學習要先從定義開始，了解定義然後熟悉解題流程，有問題就問老師。選擇題最少要寫近3年，考試時細心看題目別掉入陷阱，申論是上榜關鍵，平時多練題庫上場才不會慌，過程論述要清楚，分數自然會高。
7個月考取 翁○樺 109 高普考會計 雙料金榜 我報名的是年度班，且選擇面授課程。後期選擇參加奪榜班，利用補習班的資源多練習題目跟加強申論題寫作。前期我會強迫自己準時補習班報到，後期就是白天去奪榜班念書跟刷題，晚上繼續到補習班趕進度。	7個月考取 顏○凡 109 高普考財稅行政 雙料金榜 當我開始補習時，距離考試已經只剩下7個月，我大部分科目都是照著老師的教導方式準備，只要穩穩跟著進度，不管是學習或之後複習，都會輕鬆很多。規劃好自己的讀書計畫，順順地準備下來，你一定可以得到成果。

三、簡單隨機抽樣可分為取出不放回及取出放回兩大類，樣本觀察值之平均在兩種設計中均為母體平均 μ 之不偏估計，若樣本數為 n ， s_0 為樣本集合，該不偏估計為

$$\bar{y} = \frac{1}{n} \sum_{i \in s_0} y_i$$

請回答下列問題：

- (一)如果均以 $\bar{y}$ 作為 μ 之估計量，雖然均為樣本中觀察值之平均，但在相同的樣本數下，取出不放回的設計會比取出放回相對較有效 (more efficient)，請闡述其原因，並討論其相對有效性與樣本數及母體數之關係。(10 分)
- (二)在取出放回的設計中有另外一個母體平均 μ 之不偏估計，該估計為樣本中不同單元觀察值之平均，亦即若有重複出現之單元，其觀察值只會被計入一次，該不偏估計量為

$$\bar{y}_v = \frac{1}{v} \sum_{i \in s} y_i$$

其中 s 為 s_0 中不同單元所構成之集合， v 為 s 中之單元數目，且 $v \leq s$ 。

抽樣理論中之一基本結果為在簡單隨機抽樣取出放回設計中， $\bar{y}_v$ 亦相對較有效，請說明其理由。(5 分)

【解題關鍵】

《考題難易》★★★

《破題關鍵》簡單隨機抽樣相對效率的比較本屬於課內基本題，但描述過程與討論其相對有效性與樣本數及母體數之關係反倒需要轉一下才知道題目在問什麼。而取後放回的特別抽樣設計 $\bar{y}_v$ 雖然是課外內容，但僅有 5 分的配分，應不構成影響的關鍵。相關的想法可參考抽樣課本 P.2-19 頁之 101 年關務抽樣試題。

【擬答】

(一)採用取後不放回之變異數為 $Var(\bar{y}) = \left(\frac{N-n}{N-1}\right) \frac{\sigma^2}{n}$

採用取後不放回之變異數為 $Var(\bar{y}_w) = \frac{\sigma^2}{n}$

相對有效性 $RE = \frac{Var(\bar{y})}{Var(\bar{y}_w)} = \left(\frac{N-n}{N-1}\right) < 1$

所以，取出不放回的設計會比取出放回相對較有效。

當樣本數 n 越大，取後不放回相對取後放回的效率越佳。

而當母體趨近無限大時，則兩種方法有效性相等。

(二)由 $Var(\bar{y}_w) = E(Var(\bar{y}_w | v)) + Var(E(\bar{y}_w | v))$
 $= E(Var(\bar{y}_w | v)) + Var(\bar{y}_v)$
 $\geq Var(\bar{y}_v)$

因此 $Var(\bar{y}_w) \geq Var(\bar{y}_v)$ ， $\bar{y}_v$ 相對較有效。等號成立於 $E(Var(\bar{y}_w | v)) = 0$ ，也就是說不管選到哪一組 s 來計算平均數， $\bar{y}_w | v$ 的結果都一樣， $\bar{y}_w | v$ 之變異數等於 0，所以只有在 $\bar{y}_v = \bar{y}_w$ 時， $Var(\bar{y}_v) = Var(\bar{y}_w)$ 才會成立。

註 1：根據 Raj 與 Khamis (1958) 的作法，將 V_1 與 V_2 分別視為給定「不同單元」樣本之變異數與整體所有可能樣本之變異數。在取出放回的設計中，樣本平均數之變異數為 $Var(\bar{y}_v) = E_1 V_2(\bar{y}_v | v) + V_1 E_2(\bar{y}_v | v)$

$$\begin{aligned}
&= E_1 V_2 \left(\frac{1}{v} \sum_{i=1}^v y_i \mid v \right) + V_1 E_2 \left(\frac{1}{v} \sum_{i=1}^v y_i \mid v \right) \\
&= E_1 \left(\frac{1}{v} - \frac{1}{n} \right) S^2 + V_1 \left(\frac{1}{n} \sum_{i=1}^n y_i \right) \quad (\text{可參考註 2 之說明}) \\
&= \left(E_1 \left(\frac{1}{v} \right) - \frac{1}{n} \right) S^2 + \left(\frac{1}{n} - \frac{1}{N} \right) S^2 \\
&= \left(E_1 \left(\frac{1}{v} \right) - \frac{1}{N} \right) S^2
\end{aligned}$$

因為樣本中不同單元觀察值只會被計入一次，這些不同單元的集合，即可視為取後不放回的集合，此時樣本平均數之變異數為 $Var(\bar{y}_v) = \left(E\left(\frac{1}{v}\right) - \frac{1}{N} \right) S^2$

樣本變異數估計值為 $var(\bar{y}_v) = \left(\frac{1}{v} - \frac{1}{N} \right) \frac{\sigma^2}{v}$

根據 Pathak (1961)， $E\left(\frac{1}{v}\right) = \frac{1}{N^n} \sum_{j=1}^N j^{n-1}$

所以 $Var(\bar{y}_v) = \left(\frac{1}{N^n} \sum_{j=1}^N j^{n-1} - \frac{1}{N} \right) S^2 = \frac{S^2}{N^n} \sum_{j=1}^{N-1} j^{n-1}$

相對有效性 $RE = \frac{Var(\bar{y}_v)}{Var(\bar{y}_w)} = \frac{\frac{S^2}{N^n} \sum_{j=1}^{N-1} j^{n-1}}{\frac{S^2}{N-1} \frac{1}{n}} = \frac{n \sum_{j=1}^{N-1} j^{n-1}}{(N-1) N^{n-1}}$

所以當樣本數 n 越大時， $\bar{y}_w$ 越有效。

註 2： $E(\bar{y}_v) = E_1 E_2 (\bar{y}_v \mid v) = E_1 E_2 \left(\frac{1}{v} \sum_{i=1}^v y_i \mid v \right)$

$$= E_1 \left(\frac{1}{v} E_2 \left(\sum_{i=1}^v y_i \right) \right) = E_1 \left(\frac{1}{v} \frac{v}{n} \sum_{i=1}^n y_i \right) = E_1 \left(\frac{1}{n} \sum_{i=1}^n y_i \right) = \bar{Y}$$

$$= E_1 \left(\frac{1}{v} E_2 \left(\sum_{i=1}^v y_i \right) \right) = E_1 \left(\frac{1}{v} \frac{v}{n} \sum_{i=1}^n y_i \right)$$

$$= E_1 \left(\frac{1}{n} \sum_{i=1}^n y_i \right) = \bar{Y}$$

其中根據 Raj 與 Khamis (1958) 的作法，將 E_1 與 E_2 分別視為給定「不同單元」樣本之期望值與整體所有可能樣本之期望值。

志光 × 保成 × 學儒 法緒・憲法・公民・行政法・行政學・政治學・財政學
地方自治・公共管理・會計(含中會)・經濟學

測驗易點通

全國首創

O!ops 你又踩雷了嗎? 埋頭苦練, 不如讓老師點通你的學習之路

答題測驗就像玩踩地雷, 總是在賭一把運氣? 已經錯過的題目, 總是一錯再錯?

一點就通!

常考題型 **知識強化**

同樣的出題範圍一考再考, 卻還是選不出答案, 測驗題不能硬背, 唯有讓老師帶你一觀出題知識的原貌, 弄清題目在考什麼, 才是唯一正解。

易錯題型 **觀念釐清**

彙整全國最大公職王線上網站測驗中, 考生最高頻率答錯的試題, 針對試題透徹分析出最易混淆的考點, 加強授課、觀念釐清。

■完整課程資訊詳洽全國志光・保成・學儒門市■

四、欣民里里長希望能將里內獲得整建該里休閒公園之經費用於該公園內規劃一狗公園, 作為里民飼養家犬之活動空間, 藉以減少里民於公園遛狗時可能造成之環境污染及人犬衝突, 因此里長希望能先進行可行性評估, 其中包括對里民的問卷抽樣調查, 問卷內容包括:

1. 家中是否養狗。
2. 若有養狗, 請問飼養隻數。
3. 是否贊成狗公園之規劃。

在該里的 500 戶中, 里長以簡單隨機抽樣取出不放回的方式選擇了 50 戶作為樣本, 調查資料顯示其中有 10 戶養狗。10 戶有養狗的樣本戶對另外兩題之答案彙整如下:

家戶編號	68	124	136	158	248	314	379	401	452	479
家犬隻數	1	1	1	3	1	1	4	1	1	1
是否贊成	是	是	是	否	是	否	是	否	是	否

(一)該里家戶養狗比例之不偏估計推估值為何? 及此一比例之 95%信賴區間? (5 分)

(二)請問該里養狗之家戶平均養狗隻數之不偏估計, 以及其 95%信賴區間? (10 分)

(三)請問該里養狗總隻數之不偏估計, 以及其 95%信賴區間? (10 分)

(四)調查資料中沒有養狗的家戶贊成設置狗公園的戶數是 16 戶, 但是有人質疑此一抽樣設計或許沒有考慮家戶年齡結構, 因為有高齡長者的家戶可能希望規劃友善樂齡運動空間, 經重新審視資料, 樣本中有 20 戶為家中有 65 歲以上高齡里民, 而其中有 2 戶贊成, 而戶籍資料顯示欣民里中有 350 戶家中有 65 歲以上長者, 請問根據此一資料, 在考量家戶年齡結構之要求下, 贊成設置狗公園之比例不偏估計為何? 你是否認為此一估計量較為恰當? 請說明理由。 (10 分)

(五)若有里民質疑樣本數過少以至於養狗戶數之推估精確度不足, 若欲重新以相同抽樣設計執行此一調查, 但要將推估養狗戶數比例控制在 95%信心水準下不超過正負 5%, 請問樣本戶數至少需要多少? (10 分) (10 分)

【解題關鍵】

《考題難易》★★

《破題關鍵》簡單隨機抽樣比例值、平均值、總和值估計雖屬於課內基本題，但平均數與總和值的估計需注意題意上的差別，108 年身障、104 年高考有類似命題，可參考抽樣方法 P2-39 與 P2-49 有相同範例。本題的難度在於第四題的分層設計，其實觀念不難，只要如擬答把資料整理成習慣的表格，便可迎刃而解。

【擬答】

(一)該里家戶養狗比例之不偏估計推估值為

$$p = \frac{10}{50} = 0.2 \xrightarrow{\text{估計}} P$$

$$s_p = \sqrt{(1-f) \frac{s^2}{n}} = \sqrt{(1-f) \frac{\frac{n}{n-1} pq}{n}} \\ = \sqrt{(1-\frac{50}{500}) \frac{0.2 \times 0.8}{50-1}} = 0.0542$$

該里家戶養狗比例之 95%信賴區間為

$$p \pm Z_{0.025} \cdot s$$

$$\Rightarrow (0.2 \pm 1.96 \times 0.0542) \Rightarrow (0.0937, 0.3063)$$

(二)該里養狗之家戶平均養狗隻數之不偏估計為

$$\bar{y} = \frac{1+1+\dots+1}{10} = 1.5 \xrightarrow{\text{估計}} Y$$

$$s_y = \sqrt{(1-f) \frac{s^2}{n}} = \sqrt{(1-\frac{10}{500}) \frac{1.1667}{10}} = 0.3381$$

$$\text{其中 } s^2 = \frac{\sum_{i=1}^{10} (y_i - \bar{y})}{10-1} = 1.1667$$

平均養狗隻數之 95%信賴區間為

$$\Rightarrow 1.5 \pm 1.96 \times 0.3381$$

$$\Rightarrow (0.8373, 2.1627)$$

(三)該里養狗總隻數之不偏估計

$$\hat{Y} = N\bar{y} = 500 \times \frac{1+1+\dots+1}{50} = 500 \times 0.3 = 150 \xrightarrow{\text{估計}} Y$$

$$s_{\hat{Y}} = N \sqrt{(1-f) \frac{s^2}{n}} = 500 \sqrt{(1-\frac{50}{500}) \frac{0.5816}{50}} = 51.16$$

$$\text{其中 } s^2 = \frac{\sum_{i=1}^{50} (y_i - \bar{y})}{50-1} = 0.5816$$

養狗總隻數之 95%信賴區間為

$$\Rightarrow 150 \pm 1.96 \times 51.16$$

$$\Rightarrow (49.7264, 250.2736)$$

(四)有養狗 10 戶家戶中，有 6 戶贊成設置狗公園，沒有養狗 40 戶家戶中，有 16 戶贊成設置狗公園，代表全部樣本中，有 22 戶贊成。

現將母體分為兩層，家中有 65 歲以上高齡里民為第 1 層，家中沒有 65 歲以上高齡里民為第 2 層，可將資料整理如下：

層別	戶籍資料數 N_h	樣本數 n_h	贊成數 a_h
家中有 65 歲以上里民	350	20	2
家中沒有 65 歲以上里民	150	30	20
總和	500	50	22

分層隨機抽樣贊成設置狗公園之比例不偏估計為 $p_{st} = \sum_{h=1}^2 W_h p_h = \frac{350}{500} \times \frac{2}{20} + \frac{150}{500} \times \frac{20}{30} = 0.27$

當層跟層差異大時，採用分層隨機抽樣相較簡單隨機抽樣之精確度較高。本題的分層為家中是否有 65 歲以上里民，很明顯兩個分層的差異很大，所以採用分層抽樣較為恰當。

(五)採用簡單隨機抽樣比例值估計下，樣本數為

$$n_0 = \frac{Z^2 s^2}{B^2} = \frac{Z^2 \frac{n}{n-1} pq}{B^2} \approx \frac{Z^2 pq}{B^2} = \frac{1.96^2 \times 0.2 \times 0.8}{0.05^2} = 245.86 \approx 246$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{246}{1 + \frac{246}{500}} = 164.88$$

n 取 165

註：從題意上看起來不需要考慮分層隨機抽樣的總樣本數。

要您上榜

志光 × 保成 × 學儒 一次繳費輔導至考取

高普考取班 8 大保障

學費省很大 全年課程不間斷，一次繳清學費輔導至考取。	課程最完整 完整課程循環，基礎班→正規班→專題課→總複習…等，全部擁有。	上榜賺獎金 報名考取班第一年考取同職等考試，頒發高額獎學金。	學習最便利 輔導期間可依自己時間選擇面授或視訊學習，提高學習效率。
師資最多元 重點科目安排多元師資，雙循環教學，可旁聽加強弱科，強化上榜實力。	加選最超值 輔導期間要加選其科目增加考試機會，加選另享專案優惠。	榜單最實在 年年榜單見證，錄取人數最多，錄取率最高，奪榜實力全國第一。	公約有保障 考取班簽訂公約，保障您的權利與義務至考取為止。

■ 完整課程資訊詳洽全國志光・保成・學儒門市 ■