

《抽樣方法》

試題評析

第一題：為簡單隨機抽樣的考題，為基本題型。

第二題：為群集隨機抽樣的考題，亦為基本題，但計算第(3)小題時要特別小心，為群集隨機抽樣與比例機率抽樣法之比較。

第三題：為系統隨機抽樣的考題，自94年度考過之後未再出現，今年再考出來，亦為基本題，但計第(3)小題時會運用到統計學的觀念。

第四題：為分層隨機抽樣的考題，結合統計學的假設檢定來命題，亦為基本題。

綜觀今年抽樣方法考題，題型考法看似平易近人，但計算時在觀念上要特別小心，且此份考題亦把統計學觀念結合在一起，是份很用心出題的考卷。程度好的考生應可以拿到80分左右，一般考生應可拿到65分左右。

一、欲了解平安社區住戶的所得情形，現自平安社區2000個住戶中，以簡單隨機抽樣法抽出100戶，調查每戶每月的所得。調查結果顯示每戶每月的平均所得為12萬元，標準差為3萬元；同時，有15戶每月的所得超過20萬元。

(一)試估計平安社區每戶每月的平均所得 μ 。(5分)

(二)試估計平安社區每月所得超過20萬元的總戶數。(5分)

(三)若希望估計每戶每月平均所得 μ 的信賴度為95%且誤差界限為0.5萬；同時，估計每戶每月所得超過20萬元的比例 p 的信賴度為95%且誤差界限為0.05時，須調查多少住戶？(15分)

註： $z_{0.025} = 1.96$

答：

$$N = 2000 \quad n = 100 \quad \bar{y} = 12 \quad s = 3$$

$$(一) E(\bar{y}) = \bar{Y} = 12。$$

$$(二) p = \frac{15}{100} = 0.15, \quad p \text{ 為每月所得超過20萬元的比例, } \hat{A} = Np = 2000 \times 0.15 = 300。$$

(三)

1.

$$n_0 = \left(\frac{Zs}{B} \right)^2 = \left(\frac{1.96 \times 3}{0.5} \right)^2 = 138.2976 \cong 139$$

$$n = \frac{139}{1 + \frac{139}{2000}} = \frac{139}{1 + 0.0695} = 129.967 \cong 130$$

2.

$$n_0 = \frac{Z^2 \frac{n}{n-1} pq}{B^2} = \frac{1.96^2 \times \frac{100}{100-1} \times 0.15 \times 0.85}{0.05^2} = 197.9 \cong 198$$

$$n = \frac{198}{1 + \frac{198}{2000}} = \frac{198}{1 + 0.099} = 180.163 \cong 181$$

3.由1.與2.中可知，要符合兩者抽樣模式至少需抽出樣本 $n=181$ 。

【版權所有，重製必究！】

高上高普特考 www.get.com.tw/goldenstn¹ 台北市開封街一段2號8樓 02-23318268

【板橋】(02)29557868 【淡水】(02)26259498 【三峽】(02)26735568 【林口】(03)3275388 【羅東】(03)9540923
 【中壢】(03)4256899 【台中】(04)22298699 【逢甲】(04)27075516 【東海】(04)26527979 【中技】(04)22033988
 【彰化】(04)7289398 【台南】(06)2235868 【高雄】(07)2358996

二、工廠人事部門想了解全工廠80條生產線上的員工每年請病假的情形。現自此80條生產線中抽出8條生產線，調查每條生產線的員工人數及每年的病假天數。調查資料如下：

生產線	員工人數	病假天數
1	40	58
2	39	72
3	12	26
4	52	98
5	37	74
6	33	57
7	41	76
8	14	48

(一) 試估計每位員工每年的平均病假天數。(5分)

(二) 試估計全工廠生產線上員工每年的總病假天數，並求算此總病假天數的95%信賴區間。(10分)

(三) 若已知全工廠生產線上員工共有3000位，利用此資料估計全工廠生產線上員工的每年總病假天數。試評估(二)，(三)這兩種估計方法那一種較適合此個案。(10分)

註： $z_{0.025} = 1.96$

答：

(一) 將樣本資料視為簡單隨機樣本

$$\bar{y} = \frac{Y_1 + Y_2 + \dots + Y_8}{M_1 + M_2 + \dots + M_8} = \frac{58 + 72 + \dots + 48}{40 + 39 + \dots + 14} = \frac{509}{268} = 1.89925$$

(二) 以群集隨機抽樣法估計母體總和值

$$1. \bar{y}_t = \frac{Y_1 + \dots + Y_8}{8} = \frac{58 + 72 + \dots + 48}{8} = \frac{509}{8} = 63.625$$

$$2. \hat{Y}_{cl} = N \cdot \bar{y}_t = 80 \times 63.625 = 5090$$

$$3. s_e^2 = \frac{\sum_{i=1}^8 (Y_i - \bar{y}_t)^2}{n-1} = \frac{1}{8-1} \left[\sum_{i=1}^8 Y_i^2 - \frac{\left(\sum_{i=1}^8 Y_i \right)^2}{n} \right] = \frac{1}{7} \left[35633 - \frac{259081}{8} \right]$$

$$= 463.9821$$

$$4. s_{\hat{Y}_{cl}} = N \sqrt{1-f} \frac{s_e}{\sqrt{n}} = 80 \sqrt{1-0.1} \frac{21.5402}{\sqrt{8}} = 80 \times 0.9486 \times 7.6156 = 577.9326$$

$$5. B = Z \cdot s_{\hat{Y}_{cl}} = 1.96 \times 577.9326 = 1132.7478$$

$$6. [5090 \pm 1132.7478]$$

【版權所有，重製必究！】

高上高普特考 www.get.com.tw/goldensun2 台北市開封街一段2號8樓 02-23318268

【板橋】(02)29557868 【淡水】(02)26259498 【三峽】(02)26735568 【林口】(03)3275388 【羅東】(03)9540923
 【中壢】(03)4256899 【台中】(04)22298699 【逢甲】(04)27075516 【東海】(04)26527979 【中技】(04)22033988
 【彰化】(04)7289398 【台南】(06)2235868 【高雄】(07)2358996

(三)以比例機率抽樣法估計母體總和值

$$1. \hat{Y}_{ppz} = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{Z_i} = \frac{M}{n} \sum_{i=1}^n \frac{Y_i}{M_i} = \frac{3000}{8} \left(\frac{58}{40} + \frac{72}{39} + \dots + \frac{48}{14} \right) = 6133.8519$$

$$2. s_{\hat{Y}_{ppz}} = \sqrt{\frac{M^2 \sum_{i=1}^n (\bar{y}_i - y_i)^2}{n(n-1)}} = \sqrt{\frac{3000^2 \times (35.9301 - 33.4422)}{8(8-1)}} = 632.3298$$

$$3. B = Z \cdot s_{\hat{Y}_{ppz}} = 1.96 \times 632.3298 = 1239.3664$$

$$4. [6133.9519 \pm 1239.3664]$$

5. 由(二)與(三)中可知，此題每一群集母體的大小差異很大，所以適合採用比例機率抽樣法估計母體總和值。

三、調查A地區的離婚狀態，以系統抽樣法抽得1969年至2009年間A地區的離婚對數的年份資料如下表：

年	1969	1974	1979	1984	1989	1994	1999	2004	2009
離婚對數	85	77	93	179	408	736	889	890	875

(一)估計這段時間A地區的總離婚對數，並求算此總離婚對數的95%信賴區間。(10分)

(二)利用這段時間的平均離婚對數來預測2010年A地區的離婚對數適當嗎？請解釋之。(5分)

(三)利用連續的差數 (successive differences) 估計A地區總離婚對數的變異數，且評估利用連續的差數所產生的變異數估計式與利用簡單隨機抽樣的估計所產生的變異數估計式那種較適合此個案。(10分)

註： $z_{0.025} = 1.96$

答：

(一)每間格五年之離婚對數為y

1. 平均離婚對數

$$\sum_{i=1}^9 y_i = 4232 \quad \sum_{i=1}^9 y_i^2 = 3110050$$

$$\bar{y}_{sy} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{9} \cdot 4232 = 470.2222$$

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i \right)^2}{n} \right] = \frac{1}{8} \left[3110050 - \frac{17909824}{9} \right] = \frac{1}{8} [1120069.556]$$

$$= 140008.6944$$

$$s = 374.1773$$

$$\text{var}(\bar{y}_{sy}) = (1-f) \frac{s^2}{n} = \left(1 - \frac{9}{41} \right) \frac{140008.6944}{9} = 12141.6754 \quad s_{\bar{y}_{sy}} = 110.1893$$

$$[470.2222 \pm 1.96 \times 110.1893]$$

2. 總離婚對數

$$\hat{Y}_{sy} = N \cdot \bar{y}_{sy} = 41 \times 470.2222 = 19279.1102$$

【版權所有，重製必究！】

高上高普特考 www.get.com.tw/goldensun³ 台北市開封街一段2號8樓 02-23318268

【板橋】(02)29557868 【淡水】(02)26259498 【三峽】(02)26735568 【林口】(03)3275388 【羅東】(03)9540923
 【中壢】(03)4256899 【台中】(04)22298699 【逢甲】(04)27075516 【東海】(04)26527979 【中技】(04)22033988
 【彰化】(04)7289398 【台南】(06)2235868 【高雄】(07)2358996

$$s_{\hat{Y}_{sy}} = N \cdot s_{y_{sy}} = 41 \times 110.1893 = 4517.7613$$

$$[19279.1102 \pm 1.96 \times 4517.7613]$$

- (二) 1. 由表中的離婚對數可知從1969~1999為一個明顯上升的趨勢，雖1999~2009為一平穩狀態。所以從抽樣中所得到的估計變異數會高估於真正的變異數。
2. 依此圖表要預測2010年的離婚對數是超出目前資料所能提供的，所以離婚率的樣本平均值並不是一個用來預測2010離婚率的好估計值。

(三)

1. 連續的差數估計變異數

$$sd = \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2}{2n(n-1)} = \frac{\sum_{i=1}^9 \sum_{j=1}^9 (x_i - x_j)^2}{2 \times 9 \times (9-1)} = \frac{(85-77)^2 + \dots + (890-875)^2}{144}$$

$$= 70004.34722$$

2. 簡單隨機抽樣估計變異數

$$\text{var}(\bar{y}) = (1-f) \frac{s^2}{n} \cong \frac{s^2}{n} = \frac{140008.6944}{9} = 15556.5216$$

3.

- (1) 如以變異數值來做判斷，則簡單隨機抽樣估計變異數會較連續的差數估計變異數為優良。
(2) 但是在系統隨機抽樣下，因為未給定N值，所以亦不適合採用簡單隨機抽樣來計算變異數估計式，相對較有效率的方式應以系統隨機抽樣來估計其變異數。

四、某企業員工的薪資分為 I，II 兩層，其中 I 層有 400 名員工，II 層有 600 名員工。為了解員工對薪資的滿意程度，分別從 I，II 層員工中各抽出 100 名員工調查。調查結果依滿意程度及性別加以分類，得各項人數如下表：

	I			II		
	滿意	不滿意	沒意見	滿意	不滿意	沒意見
男	60	15	5	25	3	2
女	10	7	3	40	21	9

- (一) 試估計該企業員工對薪資滿意的比例 p，並求算此比例 p 的 95% 信賴區間。(10分)
(二) 試估計該企業對薪資不滿意的女性員工總人數。(5分)
(三) 試估計 I，II 層員工對薪資不滿意的比例差。試問在 95% 信賴水準下，此比例差是否有顯著？(10分)

註： $z_{0.025} = 1.96$

【版權所有，重製必究！】

高上高普特考 www.get.com.tw/goldensun⁴ 台北市開封街一段 2 號 8 樓 02-23318268

【板橋】(02)29557868 【淡水】(02)26259498 【三峽】(02)26735568 【林口】(03)3275388 【羅東】(03)9540923
【中壢】(03)4256899 【台中】(04)22298699 【逢甲】(04)27075516 【東海】(04)26527979 【中技】(04)22033988
【彰化】(04)7289398 【台南】(06)2235868 【高雄】(07)2358996

答：

(一)

$$1. p_{st} = \sum_{h=1}^2 W_h p_h = \frac{400}{1000} \times 0.7 + \frac{600}{1000} \times 0.65 = 0.67$$

$$2. s_{pst} = \frac{1}{N} \sqrt{\sum_{h=1}^L N_h (N_h - n_h) \frac{p_h q_h}{n_h - 1}} = \frac{1}{1000} \sqrt{933.6558} = 0.0305$$

$$3. [0.67 \pm 1.96 \times 0.0305]$$

(二)

1. 女性不滿意薪資比例

$$p_{st} = \sum_{h=1}^2 W_h p_h f_h = \frac{400}{1000} \times \frac{7}{100} + \frac{600}{1000} \times \frac{21}{100}$$

$$= 0.4 \times 0.07 + 0.6 \times 0.21 = 0.154$$

$$2. \hat{A}_{st} = N p_{st} = 1000 \times 0.154 = 154$$

(三)

$$\text{第一層薪資不滿意員工 } \hat{p}_1 = \frac{22}{100} = 0.22$$

$$\text{第二層薪資不滿意員工 } \hat{p}_2 = \frac{24}{100} = 0.24$$

$$\hat{p}_1 - \hat{p}_2 = 0.22 - 0.24 = -0.02$$

$$v(\hat{p}_1 - \hat{p}_2) = \frac{\hat{p}_1 \cdot \hat{q}_1}{n_1} + \frac{\hat{p}_2 \cdot \hat{q}_2}{n_2} = \frac{0.22 \times 0.78}{100} + \frac{0.24 \times 0.76}{100} = 0.00354$$

$$s_{\hat{p}_1 - \hat{p}_2} = 0.0595 \quad B = 1.96 \times 0.0595 = 0.11662$$

$$[-0.02 \pm 0.1162]$$

因信賴區間包含0，所以無法說明 I，II 層員工對薪資不滿意比例之間有顯著差異。

【版權所有，重製必究！】

高上高普特考 www.get.com.tw/goldensun⁵ 台北市開封街一段2號8樓 02-23318268

【板橋】(02)29557868 【淡水】(02)26259498 【三峽】(02)26735568 【林口】(03)3275388 【羅東】(03)9540923
 【中壢】(03)4256899 【台中】(04)22298699 【逢甲】(04)27075516 【東海】(04)26527979 【中技】(04)22033988
 【彰化】(04)7289398 【台南】(06)2235868 【高雄】(07)2358996