

《抽樣方法》

試題評析

今年高考抽樣方法考題共有四題，題一為簡單隨機抽樣與比率估計法的綜合考題，只要觀念清楚，並不難作答；題二為傳統統計學信賴區間與分層隨機抽樣的綜合考題，為基本考題，在過去的歷屆考題中已出現多次，亦不難作答；題三為群集隨機抽樣考題，為傳統考法，在過去歷屆考題中也出現過數次，亦易於作答；題四的類似考題也於92年高考中出現過。所以今年高考抽樣方法，普通考生應可拿60分以上，程度好者75分以上應不難。

一、某企業稽查員從該企業全部5000項財產中，以抽出不放回之簡單隨機抽樣，抽出200項財產並核估其實際價值；得到樣本平均數為\$1000，而樣本標準差為\$250。

(一)試估計全部5000項財產實際總價值之95%信賴區間。($Z_{0.025} = 1.96$) (5分)

(二)若希望估計實際總價值之95%信賴區間且估計誤差在\$50,000之內，則需要多大的樣本？(5分)

(三)設根據所抽出的200項財產之帳列價值，得到樣本平均數為\$1,100，樣本標準差為\$240。而且帳列價值與實際價值之樣本相關係數為0.95；同時已知5000項財產帳列總價值為\$5,600,000。試以比率法估計全部5000項財產實際總價值之95%信賴區間。(10分)

(四)若希望以比率法估計實際總價值之95%信賴區間且估計誤差在\$50,000之內，則需要多大的樣本？(10分)

答：

$$(一) \left(\hat{Y} \pm B \right) = \left(N \bar{y} \pm Z s_{\bar{y}} \right) = N \left(\bar{y} \pm Z s_{\bar{y}} \right) = 5000(1000 \pm 1.96 \times 17.3205)$$

$$s_{\bar{y}} = \sqrt{(1-f)} \frac{s}{n} = \sqrt{\left(1 - \frac{200}{5000}\right)} \frac{250}{\sqrt{200}} = 17.3205$$

$$(二) n = \frac{n_0}{1 + \frac{n_0}{N}} \quad n_0 = \left(\frac{NZs}{B} \right)^2 = \left(\frac{5000 \times 1.96 \times 250}{50000} \right)^2 = 2401$$

$$n = \frac{2401}{1 + \frac{2401}{5000}} = \frac{2401}{1 + 0.4802} = \frac{2401}{1.4802} = 1622.07 \approx 1623$$

(三)

$$(1) \text{實際總價值} = \bar{y} = 1000 \quad s_y = 250$$

$$(2) \text{帳列價值} = \bar{X} = 1100 \quad s_x = 240 \quad r_{xy} = 0.95 \quad r = \frac{\bar{y}}{\bar{X}} = \frac{1000}{1100} = 0.909$$

$$(3) s_d^2 = s_y^2 + r^2 s_x^2 - 2r r_{xy} s_y s_x = 250^2 + 0.909^2 \times 240^2 - 2 \times 0.909 \times 0.95 \times 250 \times 240 = 6467.78$$

所以可得 $s_d = 80.4225$

$$(4) \hat{Y}_r = rX = 0.909 \times 5600000 = 5090400$$

$$(5) s_{\hat{Y}_r} = N \sqrt{1-f} \frac{s_d}{\sqrt{n}} = 5000 \sqrt{1 - \frac{200}{5000}} \frac{80.4225}{\sqrt{200}} = 27859.24$$

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

$$(6) \left(\hat{Y}_r \pm B \right) = \left(\hat{Y}_r \pm Zs_{\hat{Y}_r} \right) = (5090400 \pm 1.96 \times 27859.24)$$

$$(四) n = \frac{n_0}{1 + \frac{n_0}{N}} \quad n_0 = \left(\frac{NZs_d}{B} \right)^2 = \left(\frac{5000 \times 1.96 \times 80.4225}{50000} \right)^2 = 248.466$$

$$n = \frac{248.466}{1 + \frac{248.466}{5000}} = \frac{248.466}{1 + 0.0496} = \frac{248.466}{1.0496} = 236.72 \approx 237$$

二、某工廠有3000位員工，為瞭解員工暴露在污染環境中工作之等級對於肺活量之影響，隨機抽出50位員工並檢查其肺活量，得到以下摘要資料：

暴露等級	人數	平均肺活量	標率差
高暴露 $N_1 = 1850$	30	72	9
中暴露 $N_2 = 700$	12	88	10
低暴露 $N_3 = 450$	8	90	8

(一) 試估計高暴露於污染環境全體員工之平均肺活量之95%信賴區間。(10分)

(二) 試估計全體員工之平均肺活量之95%信賴區間。(10分)

(三) 設已知全體員工在污染環境中，高暴露人數為 $N_1 = 1850$ 人，中暴露人數為 $N_2 = 700$ 人，低暴露人數為 $N_3 = 450$ 人。試以事後分層方式估計全體員工之平均肺活量之95%信賴區間。(10分)

答：

(一)

$$s_{p1} = \sqrt{\left(1 - \frac{30}{1850}\right) \cdot \frac{9^2}{30}} = 1.6298 \quad B = Z \cdot s_{p1} = 1.96 \times 1.6298 = 3.1944$$

$$[\bar{y}_1 \pm B] = [72 \pm 3.1944]$$

(二)

$$\bar{y}_{st} = \sum W_h \bar{y}_h = \frac{1850}{3000} \cdot 72 + \frac{700}{3000} \cdot 88 + \frac{450}{3000} \cdot 90 = 78.43$$

$$\text{var}(\bar{y}_{st}) = \sum W_h^2 (1 - f_h) \cdot \frac{s_h^2}{n_h} = 1.0101 + 0.4459 + 0.1768 = 1.6328$$

$$s_{\bar{y}_{st}} = 2.1743 \quad B = 1.96 \times 2.1743 = 4.2616$$

$$[\bar{y}_{st} \pm B] = [78.43 \pm 4.2616]$$

(三)

$$\text{var}_p(\bar{y}_{st}) = \left(\frac{1}{n} - \frac{1}{N} \right) \sum W_h s_h^2 + \frac{1}{n^2} \sum (1 - W_h) s_h^2 \approx \frac{1}{n} \sum W_h s_h^2 + \frac{1}{n^2} \sum (1 - W_h) s_h^2$$

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

$$\text{var}_p(\bar{y}_{st}) \cong \frac{1}{n} \sum W_h s_h^2 + \frac{1}{n^2} \sum (1 - W_h) s_h^2 = \frac{1}{n} \sum \frac{N_h}{N} s_h^2 + \frac{1}{n^2} \sum \left(1 - \frac{N_h}{N}\right) s_h^2$$

$$\text{var}_p(\bar{y}_{st}) \cong \frac{1}{50} \left[\frac{1850}{3000} \cdot 9^2 + \frac{700}{3000} \cdot 10^2 + \frac{450}{3000} \cdot 8^2 \right] + \frac{1}{50^2} \left[\left(1 - \frac{1850}{3000}\right) \cdot 9^2 + \left(1 - \frac{700}{3000}\right) \cdot 10^2 + \left(1 - \frac{450}{3000}\right) \cdot 8^2 \right] = 1.722686$$

$$s_{y_{pst}}^- = 1.31251 \quad B = 1.96 \times 1.31251 = 2.57251$$

$$[78.4333 \pm B] = [78.4333 \pm 2.57251]$$

三、某大都市有50家市立醫院，為瞭解急診服務品質，以一階段簡單集體 (cluster) 抽樣抽出10家醫院，被抽到的醫院去年所有外傷急診案例，將逐一被檢視是否為未盡責致死。以下資料為樣本中每家醫院去年外傷急診病人數與未盡責致死人數。

醫院代號	外傷急診病人數	外傷急診中未盡責致死人數	醫院代號	外傷急診病人數	外傷急診中未盡責致死人數
1	200	2	6	240	2
2	600	5	7	320	3
3	350	4	8	280	2
4	220	0	9	180	1
5	250	1	10	100	0

(一) 試估計該都市所有市立醫院外傷急診中未盡責致死比率之95%信賴區間。(10分)

(二) 若希望試估計該都市所有市立醫院外傷急診中未盡責致死比率之95%信賴區間，而且估計誤差在0.05%以內，則需要抽多少家醫院？(10分)

答：

$$(一) N = 50 \quad n = 10 \quad \hat{M} = N \cdot \frac{200 + 600 + \dots + 180 + 100}{10} = 50 \times 274 = 13700$$

$$(1) \bar{A}_t = \frac{A_1 + A_2 + \dots + A_{10}}{10} = \frac{2 + 5 + \dots + 1 + 0}{10} = 2$$

$$(2) P_{cl} = \frac{\hat{A}}{\hat{M}} = \frac{N \bar{A}_t}{\hat{M}} = \frac{50}{13700} \cdot 2 = 0.0073$$

$$(3) s_e^2 = \frac{\sum_{i=1}^{10} (A_i - \bar{A}_t)^2}{n-1} = \frac{1}{n-1} \left[\sum_{i=1}^{10} A_i^2 - \frac{\left(\sum_{i=1}^{10} A_i\right)^2}{n} \right] = \frac{1}{10-1} \left[64 - \frac{20^2}{10} \right] = 2.666$$

$$s_e = 1.6327 \quad f = \frac{n}{N} = \frac{10}{50} = 0.2$$

$$(4) s_{pcl} = \frac{N}{M} \sqrt{(1-f)} \frac{s_e}{\sqrt{n}} = \frac{50}{13700} \sqrt{1-0.2} \frac{1.6327}{\sqrt{10}} = 0.00169$$

$$(5) B = Z \times s_{pcl} = 1.96 \times 0.00169 = 0.0033124$$

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

$$(6) (P_{cl} \pm B) = (0.0073 \pm 0.0033124)$$

(二)

$$(1) n = \frac{n_0}{1 + \frac{n_0}{N}} \quad n_0 = \left(\frac{NZs_e}{MB} \right)^2 = \left(\frac{50 \times 1.96 \times 1.6327}{13700 \times 0.05\%} \right)^2 = 545.61$$

$$(2) n = \frac{n_0}{1 + \frac{n_0}{50}} = \frac{545.61}{1 + \frac{545.61}{50}} = \frac{545.61}{11.9122} = 45.8 \cong 46$$

四、假設某次研究調查希望從5所小學之中抽出2所小學（不得重複抽取），而且每所小學被抽中的機率必須與該小學的學生人數成正比。以下資料為這5所小學的學生人數：

學校代號	S1	S2	S3	S4	S5
學生人數	200	200	600	400	600

(一) 詳細說明你的抽樣步驟以達成上述之要求。(10分)

(二) 試求出各小學被抽中之機率。(10分)

答：

(一) 抽樣程序

抽樣的程序過程可以分成如下的步驟：

- (1) 選擇母體
- (2) 選擇適合的調查方法
- (3) 選擇抽樣型式
- (4) 選擇抽樣方法
- (5) 選擇抽樣單位
- (6) 選定誤差範圍
- (7) 選定樣本大小

(二)

(1) 每所學校被抽中的機率與學生人數成正比，所以可表示如下：

學校	S1	S2	S3	S4	S5
學生	200	200	600	400	600
機率	1/10	1/10	3/10	2/10	3/10

(2) 五所學校中取2所，學校不得重複抽取，茲以S1學校被抽中的機率如下說明：

$$(a) \frac{\binom{1}{1} \binom{1}{1} \binom{3}{0} \binom{2}{0} \binom{3}{0}}{\binom{10}{2}}$$

$$(b) \frac{\binom{1}{1} \binom{1}{0} \binom{3}{1} \binom{2}{0} \binom{3}{0}}{\binom{10}{2}}$$

【版權所有，重製必究！】

$$(c) \frac{\binom{1}{1} \binom{1}{0} \binom{3}{0} \binom{2}{1} \binom{3}{0}}{\binom{10}{2}}$$

$$(d) \frac{\binom{1}{1} \binom{1}{0} \binom{3}{0} \binom{2}{0} \binom{3}{1}}{\binom{10}{2}}$$

把(a)+(b)+(c)+(d)=即是S1學校被抽中機率，其餘S2、S3、S4、S5類推之。

高點 · 高上高普特考

【版權所有，重製必究！】

高上高普特考 www.get.com.tw/goldenstn⁵ 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