

《抽樣方法》

試題評析

今年高考抽樣方法的考題是近五年來最簡單的一次，皆以基本觀念為主，甚至出現很久未命題的名詞解釋，淺顯且易作答；題二、三、四與五，皆為基本觀念的計算，只要同學沉穩計算與作答，此份考題要拿滿分並不困難。所以普通考生應可拿80分以上，程度好者可拿90分以上，滿分亦不困難。

一、何謂分層隨機抽樣法 (Stratified random sampling)？(6分) 何謂集體抽樣法 (Cluster sampling)？(6分) 試述其意義並比較其差異。(8分)

答：

(一)分層隨機抽樣法：依若干因素與條件，把抽樣母體分成若干層次，再從這若干層次中抽取若干樣本，此從若干層次中所抽出的樣本為一簡單隨機抽樣法，但是整體綜觀之抽樣，我們稱之為分層隨機抽樣法。

(二)群集隨機抽樣法：依若干因素與條件，把抽樣母體分成若干群集，把每一個群集視為一抽樣單位，以簡單隨機抽樣的方式，從這若干的抽樣單位中抽取出數個抽樣單位，被抽中的抽樣單位其本身所隱含的樣本，即所謂抽取出的樣本，我們稱之為群集隨機抽樣法。

(三)分層隨機抽樣法與群集隨機抽樣法之比較

抽樣方法	抽樣方法說明	優點	缺點
分層隨機抽樣法	每個母體群都考慮到，且每個元素抽出的機率皆相等。	分層隨機抽樣法是相對於其它方法較有效率的。	分層隨機抽樣法是相對於其它方法較耗時的。
群集隨機抽樣法	依母體中的自然標準或是人為標準區分出異質性的cluster，每一cluster抽中機率相等，以每一cluster中的樣本為樣本。	有集中群現象特徵時，其區分與蒐集成本較低。	群集隨機抽樣法相對於其它方法是較無效率，也較不可靠的，因集群效果可能會使異質化反成同質化。

二、將一大為2,000之母體劃分成三層，其各層之大小 N_h 、標準差 S_h 及每單位之調查費用 C_h 如下表所示。今擬採分層隨機抽樣法從此一母體中抽出大小為400之樣本，試分別採：

(一)比例配置法 (Proportional allocation) (8分)

(二)紐門配置法 (Neyman's allocation) (6分)

(三)談明配置法 (Deming's allocation) (6分)

層別	N_h	S_h	C_h
1	1,000	1	3
2	600	2	2
3	400	3	1

推求各層之抽出數。

【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

答：

(一)比例配置法

層別	N_h	$n_h = n \times \frac{N_h}{N}$
1	1000	$400 \times \frac{1000}{2000} = 200$
2	600	$400 \times \frac{600}{2000} = 120$
3	400	$400 \times \frac{400}{2000} = 80$
	$\Sigma N_h = 2000$	

(二)紐門配置法

層別	N_h	S_h	$N_h S_h$	$n_h = n \times \frac{N_h S_h}{\Sigma N_h S_h}$
1	1000	1	1000	$400 \times \frac{1000}{3400} = 118$
2	600	2	1200	$400 \times \frac{1200}{3400} = 141$
3	400	3	1200	$400 \times \frac{1200}{3400} = 141$
	$\Sigma N_h S_h = 3400$			

(三)談明配置

層別	$N_h S_h$	$\sqrt{C_h}$	$\frac{N_h S_h}{\sqrt{C_h}}$	$n_h = n \times \frac{\frac{N_h S_h}{\sqrt{C_h}}}{\Sigma \frac{N_h S_h}{\sqrt{C_h}}}$
1	1000	$\sqrt{3}$	577	$400 \times \frac{577}{2626} = 88$
2	1200	$\sqrt{2}$	849	$400 \times \frac{849}{2626} = 129$
3	1200	$\sqrt{1}$	1200	$400 \times \frac{1200}{2626} = 183$
	$\Sigma \frac{N_h S_h}{\sqrt{C_h}} = 2626$			

【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、(一)何謂簡單隨機抽樣法 (Simple random sampling) ? (3分)

(二)某航空公司的研究員，為瞭解平均每天有多少人網路訂位後未能確實搭乘由甲地飛往乙地的班機，於是以簡單隨機抽樣法抽取50天之乘客網路訂位記錄，得資料如下：

y	0	1	2	3	4	5
f	20	20	5	2	1	2

其中y表示每天已網路訂位未搭乘的人數，而f表示天數；若 μ 表示每天欲搭乘此航空公司由甲地飛往乙地班機網路訂位後未能確實搭乘的平均人數。

1. 試求平均人數 μ 的估計值。(3分)

2. 試求樣本平均數 \bar{y} 之標準誤的估計值。(4分)

3. 試求平均人數 μ 的95%信賴區間？($Z_{0.025}=1.96$) (4分)

4. 若母數Y表示未來400天欲搭乘此航空公司由甲地飛往乙地班機網路訂位後，未能確實搭乘的總人數。

(1) 試求總人數Y的估計值。(3分)

(2) 試求總人數Y的95%信賴區間？(3分)

答：

(一)簡單隨機抽樣法：不考慮任何的外在因素，而母體當中每一個元素或是組成份子被抽中的機率是相等的，我們稱此為簡單隨機抽樣法。

(二)

1.

y	0	1	2	3	4	5	
f	20	20	5	2	1	2	50
yf	0	20	10	6	4	10	50

ANS：

平均人數 μ 為1人。

2.

y	0	1	2	3	4	5	
f	20	20	5	2	1	2	50
yf	0	20	10	6	4	10	50
$y^2 f$	0	20	20	18	16	50	124

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} = \frac{\sum_{i=1}^n y_i^2 f - \frac{\left(\sum_{i=1}^n yf\right)^2}{n}}{50-1} = \frac{124 - \frac{(50)^2}{50}}{49} = \frac{74}{49} = 1.5102$$

$$s = \sqrt{s^2} = \sqrt{1.5102} = 1.2289$$

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{1.2289}{\sqrt{50}} = \frac{1.2289}{7.071} = 0.1738$$

ANS：樣本平均數之標準誤的估計值為0.1738。

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

3.

$$B = 1.96 \times 0.1738$$

$$(\bar{y} \pm B) = (1 \pm 1.96 \times 0.1738)$$

$$\text{ANS: } (\bar{y} \pm B) = (1 \pm 1.96 \times 0.1738)。$$

4.

$$(a) \hat{Y} = N\bar{y} = 400 \times 1 = 400$$

ANS: 總人數Y 的估計值為400。

(b)

$$s_{\hat{y}} = Ns_y = 400 \times 0.1738 = 69.52$$

$$B = 1.96 \times 69.52$$

$$(\hat{Y} \pm B) = (400 \pm 1.96 \times 69.52)$$

$$\text{ANS: } (\hat{Y} \pm B) = (400 \pm 1.96 \times 69.52)$$

四、一房地產業務員為瞭解某地區三年來房地產價值的改變率，而於該地區全部之2,000戶房地產中，以簡單隨機抽樣法在該地區中抽取20戶房地產，並得其三年前每戶的地價(x)及目前的地價(y)相關資料(單位：百萬元)，如下表所示：

	樣本大小	平均數	標準差	相關係數
x	20	5	0.5	0.8
y	20	8	0.4	

若已知該地區三年前每戶平均地產價值為 $\bar{X} = 4$ 百萬元，試求：

(一)該地區這三年來地產價值的改變率R的估計值？(5分)

(二)此改變率的95%信賴區間？(5分)

(三)該地區今年每戶平均地產 \bar{Y} 的比率估計值？(5分)

(四)該地區今年每戶平均地產 \bar{Y} 的95%信賴區間？(5分)

答：

(一)

$$r = \frac{8}{5} = 1.6$$

ANS: 三年來地產價值改變率的估計值為1.6。

(二)

$$(1) s_d^2 = s_y^2 + r^2 s_x^2 - 2r r_{xy} s_y s_x = 0.4^2 + 1.6^2 \times 0.5^2 - 2 \times 1.6 \times 0.8 \times 0.5 \times 0.4 = 0.288$$

$$f = \frac{20}{2000} = 0.01 \quad \text{var}(r) = \frac{1-f}{(\bar{X})^2} \cdot \frac{s_d^2}{n} = \frac{1-0.01}{(4)^2} \cdot \frac{0.288}{20} = 0.061875 \cdot 0.0144 = 0.000891$$

$$s_r = 0.02984$$

$$B = 1.96 \times 0.02984$$

$$(r \pm B) = (1.6 \pm 1.96 \times 0.02984)$$

【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

ANS : $(r \pm B) = (1.6 \pm 1.96 \times 0.02984)$ 。

(三)

$$\bar{y}_r = r \cdot \bar{X} = 1.6 \times 4 = 6.4$$

ANS : 該地區平均每戶地產 \bar{Y} 的比率估計值為 6.4 (百萬元)。

(四)

$$s_{\bar{y}_r} = \bar{X} s_r = 4 \times 0.02984 = 0.11936$$

$$B = 1.96 \times 0.11936$$

$$(\bar{y}_r \pm B) = (6.4 \pm 1.96 \times 0.11936)$$

ANS : $(\bar{y}_r \pm B) = (6.4 \pm 1.96 \times 0.11936)$ 。

五、某大都市社會局的業務承辦人員，為瞭解該市住在療養院內之院民的年齡狀況，從調閱之資料知該市共有 $N=100$ 間療養院，照顧了 $M=3,000$ 位院民。由於資料甚多，僅以簡單隨機抽樣法抽出 4 間療養院作為一個集體樣本，全面檢視這 4 間療養院所有院民之年齡，並計算得各院院民之平均年齡如下表所示：

療養院	院民數	平均年齡
1	23	70
2	30	75
3	25	60
4	40	70

(一) 試估計該市療養院中所有院民平均年齡之集體估計值？(10分)

(二) 試求該市療養院中所有院民平均年齡之 95% 信賴區間？(10分)

答：

(一)

$$\bar{y}_{cl} = \frac{N}{M} \bar{y}_t = \frac{100}{3000} \times \frac{1610 + 2250 + 1500 + 2800}{4} = \frac{1}{30} \times 2040 = 68$$

ANS : 療養院中所有院民平均年齡之集體估計值為 68。

(二)

$$s_e^2 = \frac{\sum_{i=1}^4 (Y_i - \bar{y}_t)^2}{n-1} = \frac{1098200}{3} = 366067 \quad s_e = 605.0347 \quad f = \frac{4}{100} = 0.04$$

$$s_{\bar{y}_{cl}} = \frac{N}{M} \sqrt{1-f} \frac{s_e}{\sqrt{n}} = \frac{100}{3000} \sqrt{1-0.04} \frac{605.0347}{\sqrt{4}} = 0.03333 \times 0.9798 \times 302.5174 = 9.8792$$

$$B = 1.96 \times 9.8792$$

$$(\bar{y}_{cl} \pm B) = (68 \pm 1.96 \times 9.8792)$$

ANS : $(\bar{y}_{cl} \pm B) = (68 \pm 1.96 \times 9.8792)$

【版權所有，重製必究！】