

《抽樣方法》

試題評析

今年考題，實屬較為傳統的機率抽樣，考試內容皆已在授課內容，考題偏易，僅計算需謹慎並保握時間，程度好的考生可拿90分以上，一般考生可達80分以上。

一、某企業想要了解員工對福利政策的滿意度，採用分層抽樣，進行員工抽查。此企業依薪資分成三層(I、II、III)，各層員工總人數分別為3000、1500、500人。各層依簡單隨機抽樣取300、150、50人，其中一題，

Q：員工對公司的福利政策滿意度，1. 非常滿意 2. 滿意 3. 不滿意 4. 非常不滿意 5. 不知道
調查結果整理如下表所示：

註： $Z_{0.025}=1.96$

		層別			
回答	I	II	III	總和	
1	50	40	30	120	
2	160	75	15	250	
3	60	10	2	72	
4	10	5	1	16	
5	20	20	2	42	
總和	300	150	50	500	

(一)估計員工對福利政策滿意或非常滿意的比例 p ，並求算此比例 p 的95%信賴區間。(10分)

(二)估計II、III層員工滿意度(含滿意及非常滿意)比例差，在95%的信賴水準下，此差異是否顯著？(10分)

試題評析

本題為分層抽樣，在歷屆考題中亦曾出現過，要拿到分數並不困難。

考點命中

高點《抽樣方法講義》第一回，王俊彰編撰，頁47，練習7。

答：

$$(一) \hat{p}_{st} = \sum_{h=1}^L W_h \hat{p}_h = \frac{3000}{5000} \cdot \hat{p}_I + \frac{1500}{5000} \cdot \hat{p}_{II} + \frac{500}{5000} \cdot \hat{p}_{III} = \frac{3000}{5000} \cdot \frac{210}{300} + \frac{1500}{5000} \cdot \frac{115}{150} + \frac{500}{5000} \cdot \frac{45}{50} = 0.74$$

$$\widehat{\text{Var}}(\hat{p}_{st}) = \sum_{h=1}^L W_h^2 (1 - f_h) \frac{\hat{p}_h \hat{q}_h}{n_h - 1} = 0.6^2 \left(1 - \frac{300}{3000}\right) \frac{0.7 \times 0.3}{300 - 1} +$$

$$0.3^2 \left(1 - \frac{150}{1500}\right) \frac{\frac{115}{150} \times \frac{35}{150}}{150 - 1} + 0.1^2 \left(1 - \frac{50}{500}\right) \frac{0.9 \times 0.1}{50 - 1} = 0.000341$$

$$P_{st} \text{的} 95\% \text{信賴區間} \in \left(\hat{p}_{st} \mp Z_{\alpha/2} \sqrt{\sum_{h=1}^L W_h^2 (1 - f_h) \frac{\hat{p}_h \hat{q}_h}{n_h - 1}} \right)$$

$$= (0.74 \mp 0.036212) = (0.703788, 0.776212)$$

(二) $\hat{p}_2 - \hat{p}_3 = -0.13333$

$$\begin{cases} H_0: p_2 = p_3 \\ H_1: p_2 \neq p_3 \end{cases}$$

$$\text{檢定統計量 } Z = \frac{\hat{p}_2 - \hat{p}_3}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_2} + \frac{1}{n_3}\right)}}$$

檢定規則: Reject H_0 if $|Z| > Z_{0.025} = 1.96$

$$|Z| = \left| \frac{\frac{115}{150} - \frac{45}{50}}{\sqrt{\frac{160}{200}\left(1 - \frac{160}{200}\right)\left(\frac{1}{150} + \frac{1}{50}\right)}} \right| = 2.04124 > Z_{0.025} = 1.96$$

\therefore Reject H_0 表示有足夠證據顯示 $p_2 \neq p_3$ ，
即表示第 II、III 層員工滿意度有顯著差異

二、敘述群集抽樣 (cluster sampling) 與分層抽樣的優缺點。(15分)

試題評析	本題為說明比較題，要拿到分數需撰寫詳細一點。
考點命中	高點《抽樣方法講義》第一回，王俊彰編撰，頁6。

答

	分層抽樣	群集抽樣
定義	依若干因素，將母體分成不重複的若干層次，再從每一層中選取一組簡單隨機樣本，此抽樣程序稱為分層隨機抽樣。	將母體大小為M的抽樣母體，分成N個相似的群集，並以此群集為抽樣單位，然後隨機抽出n個群集，此抽樣程序稱為群集隨機抽樣。
原則	組內變異小，組間變異大	群內變異大，群間變異小
優點	1.能將大樣本，按重要變項予以畫分 2.樣本太大無法列出時，需使用該技術。 3.可與其它技術結合使用。 4.可靠性高 5.利於比較	若各階層有加權錯誤，可能導致偏差。
缺點	1.容易蒐集被研究者的資料 2.經濟省事 3.簡單易行	若組群的數目小易致偏差。

三、某一經濟學者想要了解臺北市某區的家戶電力使用量的費用狀況，在沒有全區戶數名單及經濟成本考量下，採用群集抽樣(cluster sampling)。總共有200鄰，簡單隨機抽取20鄰，所得資料如下表：

(單位：千元)

抽樣鄰號	戶數	抽樣鄰戶 總電費	抽樣鄰號	戶數	抽樣鄰戶 總電費
1	150	524.4	11	81	282.9

2	69	239.1	12	50	207.3
3	125	361.6	13	117	417.5
4	72	230.4	14	64	227.6
5	93	303.7	15	73	240.9
6	79	274.0	16	47	156.5
7	90	315.5	17	82	293.6
8	84	298.6	18	176	630.2
9	80	266.3	19	122	432.5
10	101	347.8	20	60	214.8

(一)估計此區一戶的平均月支電費及估計標準誤 (the error of estimation)。(10分)

(二)估計此區居民的總電費及其估計標準誤。(10分)

(三)假設鄰近行政區的居民結構與此區相似，欲估計鄰近行政區總電費，在估計標準誤為5000千元情況下，請估算須抽取多少群集？(10分)

試題評析 本題為M未知的群集抽樣考題，計算需花一些時間，並且需小心謹慎，要拿到分數並不困難。

考點命中 高點《抽樣方法講義》第二回，王俊彰編撰，頁6，例題1。

答：

$$(一) \hat{\mu}_{icl} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n m_i} = \frac{6265.2}{1815} = 3.451901$$

$$\widehat{\text{Var}}(\hat{\mu}_{icl}) = \frac{1}{M^2} (1-f) \frac{s_r^2}{n} = \frac{1}{m^2} (1-f) \frac{s_r^2}{n} = \frac{1}{\left(\frac{1815}{20}\right)^2} \left(1 - \frac{20}{200}\right) \frac{451.3409}{20}$$

$$= 0.002466$$

$$iis_{\hat{\mu}_{icl}} = 0.049661$$

$$\text{where } s_r^2 = \frac{\sum_{i=1}^n (y_i - \bar{y}_{cl} m_i)^2}{n-1} = 451.3409$$

$$(二) \hat{\mu}_{ocl} = N_i \bar{y}_t = N_i \bar{N} \frac{\sum_{i=1}^n y_i}{n} = 200_i \bar{N} \frac{1815}{20} = 18150$$

$$\widehat{\text{Var}}(\hat{\mu}_{ocl}) = N^2 (1-f) \frac{s_t^2}{n} = 200^2 \left(1 - \frac{20}{200}\right) \frac{13128.7278}{20} = 23631710.021$$

$$iis_{\hat{\mu}_{ocl}} = 4861.2457$$

$$\text{where } s_t^2 = \frac{\sum_{i=1}^n (y_i - \bar{y}_t)^2}{n-1} = 13128.7278$$

$$(三) \sqrt{N^2 (1-f) \frac{s_t^2}{n}} = 5000_i \sqrt{200^2 \left(1 - \frac{n}{200}\right) \frac{13128.7278}{n}} = 5000$$

$$i \div 200^2 \left(1 - \frac{n}{200}\right) \frac{13128.7278}{n} = 25000000$$

$$i \div \frac{525149112}{n} = 2625745.5579 + 25000000$$

$$j \div n = 19.009; j \div \text{取} n = 20$$

四、某手機公司欲估計某款手機三個月（一季）的總營收，已知去年同季的各分店營收（ x_i ），公司根據246家分店，採簡單隨機抽樣抽取20家分店，資料如下表：（單位：千元）

分店	上一年同季營收 (x_i)	本年本季營收 (y_i)	分店	上一年同季營收 (x_i)	本年本季營收 (y_i)
1	1500	1600	11	1350	1570
2	928	977	12	720	780
3	620	600	13	670	980
4	550	610	14	957	1020
5	1750	2210	15	610	710
6	1530	1710	16	1100	1450
7	1200	1440	17	1450	1560
8	729	865	18	800	920
9	1020	1030	19	780	850
10	980	1050	20	1300	1450

(一)利用比例估計式及去年同季總營收（ τ_x ）=256400千元，估算今年本季總營收（ τ_y ）和其估計標準誤。（10分）

(二)利用比例估計式，估算分店的平均本季營收及估計標準誤。（10分）

(三)利用迴歸估計式，估算分店的平均本季營收及估計標準誤。（15分）

試題評析	本題為比率與迴歸抽樣的綜合考題，在歷屆考題中亦曾出現，計算需花一些時間，並且小心謹慎，要拿到分數並不困難。
考點命中	高點《抽樣方法講義》第三回，王俊彰編撰，頁18，練習5。

答：

$$(一) \hat{R} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = 1.13814$$

$$\hat{\sigma}_{y_r} = r \hat{\sigma}_x = 1.13814 \sqrt{256400} = 291808.84$$

$$\widehat{\text{Var}}(\hat{\sigma}_{y_r}) = N^2(1-f) \frac{s_r^2}{n} = 246^2 \left(1 - \frac{20}{246}\right) \frac{11091.1692}{20} = 30831232.1422$$

$$s_r = \hat{\sigma}_{y_r} = 5552.5879$$

$$\text{where } s_r^2 = \frac{\sum_{i=1}^n (y_i - rx_i)^2}{n-1} = \frac{S_{yy} - 2rS_{xy} + r^2S_{xx}}{n-1} = 11091.1692$$

$$S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = 3489857.8$$

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = 2399577.2$$

$$S_{xy} = \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}) = 2806090.6$$

$$(二) \hat{t}_{y_r} = N\hat{\mu}_{y_r} \rightarrow \hat{\mu}_{y_r} = \frac{\hat{t}_{y_r}}{N} = 1186.2148$$

$$\widehat{\text{Var}}(\hat{t}_{y_r}) = N^2\widehat{\text{Var}}(\hat{\mu}_{y_r}) \rightarrow \widehat{\text{Var}}(\hat{\mu}_{y_r}) = \frac{\widehat{\text{Var}}(\hat{t}_{y_r})}{N^2} = \frac{30831232.1422}{246^2} = 509.4724$$

$$i \ddot{s}_{\hat{\mu}_{y_r}} = 22.5715$$

$$(三) \hat{\mu}_{y_{lr}} = \bar{y} + b(\mu_x - \bar{x}) = 1169.1 + 1.1694 \times \left(\frac{256400}{246} - 1027.2 \right) = 1186.7304$$

$$b = \frac{S_{xy}}{S_{xx}} = 1.1694$$

$$\widehat{\text{var}}(\hat{\mu}_{y_{lr}}) = (1-f) \times \frac{\text{MSE}}{n} = (1-f) \times \frac{S_{yy} - b^2 S_{xx}}{n-2} = \left(1 - \frac{20}{246} \right) \times \frac{11580.26}{20}$$

$$= 531.9388$$

$$i \ddot{s}_{\hat{\mu}_{y_r}} = 23.0638$$

【版權所有，重製必究！】