MORIMOTO. T

作業3. 參考答案 by 森元

7.12. ⊛ 假設母體為 $\{S_1, S_2, \cdots, S_N\} \Rightarrow \underbrace{\mu = \sum_{j=1}^{N} \frac{S_j}{N}, \ \sigma^2 = \sum_{j=1}^{N} \frac{(S_j-\mu)^2}{N}}$

(a) $$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \{(X_i - \mu) - (\bar{X} - \mu)\}^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \{(X_i-\mu)^2 - 2(\bar{X}-\mu)(X_i-\mu) + (\bar{X}-\mu)^2\}$$

$$= \frac{1}{n-1} \left\{ \sum_{i=1}^{n} (X_i-\mu)^2 - n(\bar{X}-\mu)^2 \right\}$$

$$E[S^2] = \frac{1}{n-1} \left\{ \sum_{i=1}^{n} \underbrace{E[(X_i-\mu)^2]}_{\to \sigma^2} - n \underbrace{E[(\bar{X}-\mu)^2]}_{\to \frac{\sigma^2}{n}} \right\}$$

由於「Sampling with replacement」,

故 $X_1 \sim X_n$ 可視為 iid 的隨機變數.

- $$E[(X_1-\mu)^2] = \sum_{j=1}^{N} (S_j-\mu)^2 \cdot \underbrace{Pr(X_1 = S_j)}_{= \frac{1}{N}}$$

$$= \sigma^2$$

$$\therefore E[(X_i-\mu)^2] = \sigma^2$$

- $$E[(\bar{X}-\mu)^2] = E\left[\frac{\{(X_1-\mu) + \cdots + (X_n-\mu)\}^2}{n^2}\right]$$

$$= \frac{1}{n^2} \left\{ \sum_{i=1}^{n} \underbrace{E[(X_i-\mu)^2]}_{= \sigma^2} + \underbrace{\sum\sum_{(i \neq j)} cov[X_i, X_j]}_{= 0} \right\}$$

$(\because X_i, X_j$ 為獨立$)$

$\uparrow \because$ iid

$$= \frac{\sigma^2}{n}$$

$$\Rightarrow E[S^2] = \frac{n\sigma^2}{n-1} - \frac{\sigma^2}{n-1} = \frac{n-1}{n-1} \sigma^2 = \sigma^2$$

$\therefore$ S^2 為 σ^2 之不偏估計量

CHA SHIN

注 $\underbrace{E[S^2]}_{=\sigma^2} - E[S]^2 = V[S]$ 若 $E[S]=\sigma \Rightarrow V[S]=0 \Rightarrow S$ 為常數 (a.s)

$(\Pr(S=a)=1)$

7.12

(b) 不一定.

例 $N=3 \cdots \{1,2,3\} \cdots \sigma^2=\frac{2}{3}, \quad \sigma=\frac{\sqrt{2}}{\sqrt{3}}=\frac{\sqrt{6}}{3}$

$n=2 \cdots$

$(X_1, X_2) =$

	$\bar{X}$	S^2	S
(1,1)	1	0	0
(1,2)	1.5	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$
(1,3)	2	2	$\sqrt{2}$
(2,1)	1.5	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$
(2,2)	2	0	0
(2,3)	2.5	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$
(3,1)	2	2	$\sqrt{2}$
(3,2)	2.5	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$
(3,3)	3	0	0

$$\therefore E[S^2] = \frac{1}{9}\left(0+\frac{1}{2}+2+\frac{1}{2}+0+\frac{1}{2}+2+\frac{1}{2}+0\right) = \frac{2}{3} = \sigma^2$$

$$E[S] = \frac{1}{9}\left(0+\frac{1}{\sqrt{2}}+\sqrt{2}+\frac{1}{\sqrt{2}}+0+\frac{1}{\sqrt{2}}+\sqrt{2}+\frac{1}{\sqrt{2}}\right)$$

$$= \frac{4\sqrt{2}}{9} \neq \sigma$$

(c) $\sigma_{\bar{x}}^2 = V[\bar{X}] = \frac{\sigma^2}{n} \quad (\because (a))$

$E\left[\frac{1}{n}S^2\right] = \frac{1}{n}E[S^2] = \frac{1}{n}\sigma^2$

$\therefore$ 證明完成

CHA SHIN

7.12

(d) $T = N\bar{X}$ (maybe)

$V[T] = V[N\bar{X}] = N^2 V[\bar{X}] = \frac{N^2\sigma^2}{n}$

$\therefore \sigma_T^2 = \frac{N^2\sigma^2}{n}$

$E[n^{-1}N^2S^2] = \frac{N^2}{n}E[S^2] = \frac{N^2}{n}\sigma^2$

$\therefore \sigma_T^2 = E[n^{-1}N^2S^2]$ ∴證明完成

(e) 這題你可以將 $S_1 \sim S_N$ 當成 0 或 1.

ex (1…代表「男性」; 0…代表「女性」)

$\mu = p = \frac{1}{N}\sum_{j=1}^{N} S_j$ … N個人裡面男性的比例

$\hat{p} = \frac{X_1+X_2+\cdots+X_n}{n}$ … p的估計量

↓ 抽出n個人,然後計算男性比例 $(n<N)$

由於 Sampling With Replacement, 故 $X_1 \sim X_n$ 為 iid

且 $Pr(X_i=1) = p$; $Pr(X_i=0) = 1-p$

$\therefore X_1 \sim X_n \sim b(p)$ $(= Bin(1,p))$

$\therefore X_1+X_2+\cdots+X_n \sim Bin(n,p)$ … 期望值 np, 變異數 $np(1-p)$

$E[\hat{p}] = E\left[\frac{X_1+\cdots X_n}{n}\right] = \frac{1}{n}\cdot np = p$

$V[\hat{p}] = E[\hat{p}^2] - \underbrace{E[\hat{p}]^2}_{p^2} = V\left[\frac{X_1+X_2+\cdots X_n}{n}\right] = \frac{1}{n^2}\cdot np(1-p) = \frac{1}{n}p(1-p)$

7.12 (e)

$$E[\hat{p}^2] = \frac{p(1-p)}{n} + p^2$$

$$\therefore E[\hat{p}(1-\hat{p})] = E[\hat{p}-\hat{p}^2] = p - \frac{1}{n}p(1-p) - p^2$$

$$= (p-p^2)(1-\frac{1}{n}) = \frac{n-1}{n}p(1-p)$$

$$\therefore E[\frac{1}{n-1}\hat{p}(1-\hat{p})] = \frac{1}{n}p(1-p)$$

$$\underset{\sigma_{\hat{p}}^2}{\overset{}{V[\hat{p}]}} = V[\frac{X_1+\cdots+X_n}{n}] = \frac{1}{n^2}V[\underbrace{X_1+\cdots+X_n}_{\sim Bin(n,p)}] = \frac{1}{n^2}np(1-p) = \frac{1}{n}p(1-p)$$

$$\therefore \sigma_{\hat{p}}^2 = E[\frac{1}{n-1}\hat{p}(1-\hat{p})]$$

$\therefore \frac{1}{n-1}\hat{p}(1-\hat{p})$ 為 $\sigma_{\hat{p}}^2$ 之不偏估計量.

7.14.

$$I_j = \begin{cases} 1 \cdots & j\text{th 醫院之出院人數少於 } 1000. \\ 0 \cdots & \text{otherwise} \end{cases}$$

$j = 1, 2 \cdots n$ （$n = 25 \cdots$ 樣本數）

$I_1, I_2, \cdots I_{25} \overset{iid}{\sim} b(p)$ （bernoulli 分佈） $(= Bin(1, p)$ （期望 p 變異數 $p(1-p)$））

$I_1 + I_2 + \cdots + I_{25} \sim Bin(n, p)$ （期望值 np. 變異數 $np(1-p)$）

根據中央極限定理，$\sqrt{25}\left(\underbrace{\frac{I_1 + I_2 + \cdots I_{25}}{25}}_{\hat{p}} - p\right) \underset{\text{(近似)}}{\sim} N(0, p(1-p))$

$\therefore 5(\hat{p} - p) \sim N(0, p(1-p))$

$\Rightarrow \hat{p} - p \sim N(0, \frac{1}{25} p(1-p))$

$\Rightarrow \hat{p} \sim N(p, \frac{1}{25} p(1-p))$ $(p = 0.654)$

總數的估計量 $N\hat{p} \sim N(Np, \frac{N^2}{25} p(1-p))$

(of 出院人數少於1000之醫院) $N = 393$

$\therefore N(393 \cdot 0.654, \frac{393^2}{25} \cdot 0.654 (1 - 0.654))$

※ P.S. 若考慮 finite population correction. （有限母體修正）

$$\hat{p} \sim N(p, \frac{1}{25} p(1-p) \cdot (\frac{N-n}{N-1}))$$

但「利用中央極限定理」代表 $I_1 \sim I_{25}$ 為 iid 的隨機變數. 所以可以不用考慮它.

CHA SHIN

$(X_1 \sim X_n : \text{iid})$

7.19

$E[X_j] = \mu \quad V[X_j] = \sigma^2 \quad (j = 1 \sim n)$

根據中央極限定理, $(n \to \infty)$

$$Z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

$$S_{\bar{X}} = \sqrt{V[\bar{X}]} = \frac{\sigma}{\sqrt{n}}$$

② $\because Pr\left(-\infty < \frac{\sqrt{n}}{\sigma}(\bar{X}-\mu) \leq c_2\right) = \Phi(c_2) = 0.95$ $\quad c_2 = 1.645$

$$\| $$

$$Pr\left(-\infty < \bar{X} - \mu \leq \underbrace{\frac{\sigma}{\sqrt{n}}}_{S_{\bar{X}}} c\right)$$

$$\|$$

$$Pr\left(\bar{X} - \frac{\sigma}{\sqrt{n}} c \leq \mu < \infty\right) = \Phi(c)$$

$$\|$$

$$Pr(\bar{X} - c\, S_{\bar{X}} \leq \mu < \infty)$$

$\therefore c_2 = 1.645$ $\quad$ $k = 1.645$

(k)

① $Pr\left(c_1 \leq \frac{\sqrt{n}}{\sigma}(\bar{X}-\mu) < \infty\right) = 0.9$

$$\|$$

$$Pr\left(\frac{\sigma c_1}{\sqrt{n}} \leq \bar{X} - \mu < \infty\right)$$

$$\|$$

$$Pr\left(-\infty < \mu \leq \bar{X} - \frac{\sigma}{\sqrt{n}} c_1\right) = 0.90$$

$\Phi(c_1) = 0.1 \qquad c_1 = \Phi^{-1}(0.9) = -1.28$

$\therefore Pr\left(-\infty < \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} 1.28\right) = 0.9$

$\therefore k = 1.28$

CHA SHIN

7.26

(a) 母體 $= \{x_1, x_2, \dots, x_N\}$

$U_i=1$ 表示被抽出來

$$\{X_1, X_2, \dots X_n\} = \{U_i x_i \mid U_i = 1\}$$

$$\therefore X_1+X_2+\dots X_n = \sum_{\substack{i=1\\U_i=1}}^{N} x_i = \sum_{\substack{i=1\\U_i=1}}^{N} U_i x_i = \sum_{i=1}^{n} U_i x_i$$

$(\because U_i = 0 \text{ or } 1)$

$$\therefore \bar{X} = \frac{1}{n}\sum_{i=1}^{n} U_i x_i$$

(b) $$Pr(U_i=1) = \frac{{}_{N-1}C_{n-1}}{{}_{N}C_{n}} \quad \left(\frac{\text{從 }\{x_1, \dots x_{i-1}, x_{i+1} \dots x_N\}\text{ 選 }n-1\text{ 個}}{\text{從 }\{x_1, \dots x_N\}\text{ 選 }n\text{ 個的組合}}\right)$$

$$= \frac{(N-n)!\,n!}{N!} \cdot \frac{(N-1)!}{(n-1)!\,(N-n)!} = \frac{n}{N}$$

$$E[U_i] = \sum_{U_i=0,1} U_i \cdot Pr(U_i) = 0 \cdot Pr(U_i=0) + 1 \cdot Pr(U_i=1)$$

$$= Pr(U_i=1) = \frac{n}{N}$$

U_i 為 0 或 1 $\quad Pr(U_i=1) = \frac{n}{N}$

$\therefore U_i \sim \text{bernoulli}\left(\frac{n}{N}\right)$

(c) $$E[U_i^2] = 1^2 \cdot Pr(U_i=1) = \frac{n}{N}$$

$$E[U_i]^2 = \left(\frac{n}{N}\right)^2$$

$$\therefore V[U_i] = \frac{n}{N} - \left(\frac{n}{N}\right)^2 = \frac{n}{N}\left(1-\frac{n}{N}\right)$$

(注) $\{x_1 \cdots x_N\} \setminus \{x_i, x_j\}$

$\Rightarrow$ 除了 x_i, x_j 以外的東西

(d) $$E[U_iU_j] = \sum_{\substack{U_i=0,1 \\ U_j=0,1}} \sum u_i \cdot u_j \cdot Pr(U_i=u_i, U_j=u_j)$$

$$= Pr(U_i=1, U_j=1) \quad \left(\because U_i=0 \text{ or } U_j=0 \Rightarrow u_i u_j Pr(U_i=u_i, U_j=u_j)=0\right)$$

$$= \frac{{}_{N-2}C_{n-2}}{{}_NC_n} = \left(\frac{N!}{(N-n)!\,n!}\right)^{-1} \cdot \frac{(N-2)!}{(N-n)!\,(n-2)!}$$

$$= \frac{n(n-1)}{N(N-1)}$$

$\frac{\text{從}\{x_1 \sim x_N\} \setminus \{x_i, x_j\} \text{選} n-2 \text{個}}{\text{從}\{x_1 \sim x_N\} \text{選} n \text{個}}$

(e) $$Cov[U_i, U_j] = \underbrace{E[U_iU_j]}_{\frac{n(n-1)}{N(N-1)}} - E[U_i]E[U_j]$$

$$\frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2$$

$$= \frac{n(n-1)N - n^2(N-1)}{N^2(N-1)} = \frac{n^2N - nN - n^2N + n^2}{N^2(N-1)}$$

$$= \frac{-n(N-n)}{N^2(N-1)}$$

(f) $$V[\bar{X}] = V\left[\frac{1}{n}\sum_{i=1}^N U_i x_i\right] \quad ((注) \ x_i \cdots \text{常數})$$

$$= \frac{1}{n^2} V\left[\sum_{i=1}^N x_i U_i\right] = \frac{1}{n^2} E\left[\left\{x_1\left(U_1 - \frac{n}{N}\right) + \cdots + x_N\left(U_N - \frac{n}{N}\right)\right\}^2\right]$$

$$= \frac{1}{n^2}\left\{ \sum_{j=1}^N \underbrace{E\left[x_j^2\left(U_j - \frac{n}{N}\right)^2\right]}_{} + \sum_{\substack{i \\ (i \neq j)}}\sum_{j} \underbrace{E\left[x_i x_j\left(U_i - \frac{n}{N}\right)\left(U_j - \frac{n}{N}\right)\right]}_{} \right.$$

$$= x_j^2 V[U_j] \qquad\qquad = x_i x_j \underbrace{Cov[U_i, U_j]}_{\frac{-n(N-n)}{N^2(N-1)}}$$

$$= x_j^2 \cdot \frac{n}{N}\left(1 - \frac{n}{N}\right)$$

7.26

(f)
$$= \frac{1}{n^2}\left\{ \frac{n}{N}\left(1-\frac{n}{N}\right)\sum_{j=1}^{N} x_j^2 + \frac{-n(N-n)}{N^2(N-1)}\sum_{(i\neq j)}\sum x_i x_j \right\}$$

$$= \frac{1}{n^2}\cdot\frac{n(N-n)}{N^2}\left\{ \underbrace{\sum_{j=1}^{N} x_j^2}_{N(\mu^2+\sigma^2)} - \frac{1}{N-1}\underbrace{\sum_{(i\neq j)}\sum x_i x_j}_{(x_1+\cdots x_N)^2-(x_1^2+\cdots x_N^2)} \right\}$$

$$= N^2\mu^2 - N\mu^2 - N\sigma^2$$
$$= N(N-1)\mu^2 - N\sigma^2$$

$$= \frac{1}{n^2}\frac{n(N-n)}{N^2}\left\{ N(\mu^2+\sigma^2) - N\mu^2 + \frac{N}{N-1}\sigma^2 \right\}$$

$$= \frac{1}{n^2}\cdot\frac{n(N-n)}{N^2}\underbrace{\left\{ N\sigma^2 + \frac{N}{N-1}\sigma^2 \right\}}_{\frac{N^2\sigma^2}{N-1}}$$

$$= \frac{1}{n^2}\cdot\frac{n(N-n)}{N^2}\cdot\frac{N^2\sigma^2}{N-1} = \frac{1}{n}\cdot\frac{N-n}{N-1}\cdot\sigma^2$$

注
$$\sum_{i\neq j}\sum x_i x_j = x_1(x_2+x_3+\cdots+x_N) + x_2(x_1+x_3+\cdots+x_N) + \cdots + x_N(x_1+\cdots+x_{N-1})$$

（全部写出来）

$$= x_1(x_1+\cdots x_N) - x_1^2 + x_2(x_1+\cdots x_N) - x_2^2 + \cdots + x_N(x_1+\cdots x_N) - x_N^2$$

$$= (x_1+\cdots x_N)(x_1+\cdots x_N) - (x_1^2+\cdots x_N^2)$$

CHA SHIN

(採訪者可以控制) $\begin{cases} \cdot\ p \cdots \text{問到敏感問題的機率（例如：「您是否曾經考試作弊過」）} \\ \cdot\ 1-p \cdots \text{問到一般問題的機率（例如：「您的血型是A型嗎」）} \end{cases}$

7.29

$q = Pr(\text{YES} \mid \text{敏感的問題})$

↑我們感興趣的未知參數

我們已知血型分佈，

$\underset{\text{known}}{t} \overset{def}{=} Pr(\text{YES} \mid \text{一般的問題})$

雖然採訪者並不知道被採訪者回答的問題是什麼，

但由於 p 與 t 已知，故可估計 q。

$$I_j = \begin{cases} 1 \cdots \text{第}j\text{個被採訪者回答「YES」} \\ 0 \quad\quad\quad\quad \text{〃} \quad\quad\quad \text{「NO」} \end{cases} \quad (j=1\sim n)$$

$$Pr(I_j=1) = Pr(\text{YES} \mid \text{敏感問題})\cdot Pr(\text{敏感問題}) + Pr(\text{YES} \mid \text{一般問題})\cdot Pr(\text{一般問題})$$

$$\underbrace{Pr(\text{YES} \mid \text{敏感問題})}_{q}\cdot \underbrace{Pr(\text{敏感問題})}_{p} + \underbrace{Pr(\text{YES} \mid \text{一般問題})}_{t}\cdot \underbrace{Pr(\text{一般問題})}_{1-p}$$

$$= qp + t(1-p)$$

$$\therefore I_j \overset{iid}{\sim} \text{bernoulli}(qp+t(1-p)) \quad (j=1\sim n)$$

根據弱大數法則，$\dfrac{I_1+I_2+\cdots+I_n}{n} \xrightarrow{P} qp+t(1-p)$

$$\therefore \frac{\bar{I}-t(1-p)}{p} \xrightarrow{P} q. \qquad \therefore \underset{(\text{估計量})}{\hat{q}} \overset{def}{=} \frac{\frac{I_1+\cdots+I_n}{n}-t(1-p)}{p} \quad ⌟$$

※ (Method of Moments Estimator)

$$X \overset{def}{=} I_1+I_2+\cdots+I_n \qquad X \sim Bin(n,\ qp+t(1-p)) \quad (\text{二項分布})$$

$$\cdot\ E[X] = n\{qp+t(1-p)\} \qquad \therefore E[\hat{q}] = \frac{qp+t(1-p)-t(1-p)}{p} = q \quad ⌟$$

$$\cdot\ V[\hat{q}] = V\left[\frac{X-nt(1-p)}{np}\right] = V\left[\frac{X}{np}\right] = \frac{1}{(np)^2}V[X]$$

$$= \frac{1}{(np)^2}\, n\{qp+t(1-p)\}\{1-qp-t(1-p)\}$$

$$= \frac{1}{np^2}\{qp+t(1-p)\}\cdot\{1-qp-t(1-p)\} \quad ⌟$$

CHA SHIN

$N=2000.$

7.35. μ…母體的平均，σ^2…母體的變異數

(a) 令 $X_1 \sim X_{25}$ 代表抽出來的樣本。

$(X_1=104, X_2=109, \cdots, X_{25}=97)$

$E[X_j] = \mu \quad (j=1\sim25)$

$$\underbrace{E\left[\frac{X_1+\cdots+X_{25}}{25}\right]}_{\bar{X}} = \mu$$

樣本數越多，預測越精準，以 $\bar{X}$ 作為 μ 的

估計量。 $\therefore \bar{X} = \frac{1}{25}(104+109+\cdots+97) \quad (=98.04)$

(b) $\frac{1}{n-1}\sum_{j=1}^{n}(X_j-\bar{X})^2$ 為 σ^2 之不偏估計量。 (133.7067)

$(n=25)$

另外 $V[\bar{X}] = \frac{1}{n}\cdot\frac{N-n}{N-1}\sigma^2. \quad (n=25, N=2000)$

P.S. 由於 N 很大，故 $\frac{N-n}{N-1} \approx 1$

$\therefore V[\bar{X}] \approx \frac{\sigma^2}{n}$ (跟 $X_1 \sim X_n$: i.i.d 的狀況差不多。)

無限母體

(c) 雖然我們現在考慮的是有限母體，

但由於 N 很大，所以假設 $X_1 \sim X_n$ 為 i.i.d 的樣本。

另外，我們也把 (b) 的 $\frac{1}{n-1}\sum_{j=1}^{n}(X_j-\bar{X})^2$ 當作母體的變異數，

7.35

(c) 由於 $X_1 \sim X_n$ $(n=25)$ 為 iid 的隨機變數（假設）

且 $\sigma^2 < \infty$，故 $\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) \xrightarrow{d} N(0,1)$

（∵ 中央極限定理）

$$Pr\left(-1.96 \le \frac{\sqrt{n}}{\sigma}(\bar{X}-\mu) \le 1.96\right) = 0.95$$

$$= Pr\left(\frac{-\sigma}{\sqrt{n}} \cdot 1.96 \le \bar{X}-\mu \le \frac{\sigma}{\sqrt{n}} \cdot 1.96\right)$$

$$= Pr\left(\bar{X} - \frac{\sigma}{\sqrt{n}} \cdot 1.96 \le \mu \le \bar{X} + \frac{\sigma}{\sqrt{n}} \cdot 1.96\right) = 0.95$$

（σ^2… 在此假設 $\frac{1}{n-1}\sum_{j=1}^{n}(X_j-\bar{X})^2$ 跟 σ^2 一致）

$$= Pr(93.50732 \le \mu \le 102.5727) \quad (=0.95)$$

$$\therefore [93.50732, 102.5727]$$

$$Pr(2000 \cdot 93.50732 \le 2000\mu \le 2000 \cdot 102.5727) = (0.95)$$

$$\therefore [187014.6, 205145.4]$$

CHA SHIN

7.49. 參考課本 7.4 Estimation of a Ration

(a) 我們有興趣的數值 $= r = \frac{\mu_Y}{\mu_X}$

根據(弱)大數法則 $\frac{X_1+\cdots X_n}{n} \xrightarrow{P} \mu_X$；$\frac{Y_1+\cdots+Y_n}{n} \xrightarrow{P} \mu_Y$ 且

$\Rightarrow \frac{\left(\frac{Y_1+\cdots Y_n}{n}\right)}{\left(\frac{X_1+\cdots X_n}{n}\right)} \xrightarrow{P} \frac{\mu_Y}{\mu_X}$。

故此，以 $\frac{\left(\frac{Y_1+\cdots Y_n}{n}\right)}{\left(\frac{X_1+\cdots+X_n}{n}\right)}$ 作為 $r = \frac{\mu_Y}{\mu_X}$ 之估計量 $(\hat{r})$

$$(R=)\ \frac{\bar{Y}}{\bar{X}} = \frac{150/50}{(3000/50)} = \frac{1}{20} = 0.05$$

(b) 直接利用課本的方法：

$$S_R^2 = \frac{1}{n}\cdot\left(1-\frac{n-1}{N-1}\right)\cdot\frac{1}{\bar{X}^2}\left(R^2 S_X^2 + S_Y^2 - 2RS_{XY}\right) \quad (\text{近似})$$

(由於N很大，故無視有限母體修正) →

註
$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{49}(45000)$$
$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2 = \frac{1}{49}(200)$$
$$S_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) = \frac{1}{49}(2000)$$

$$\therefore S_R^2 = \frac{1}{50}\left(1-\frac{49}{999}\right)\cdot\frac{1}{60^2}\left(0.05^2\cdot\frac{45000}{49} + \frac{200}{49} - 2\cdot 0.05\cdot\frac{2000}{49}\right)$$
$$\doteqdot 1.2129476\cdot 10^{-5}$$

$\therefore S_R = 0.003482739726\ldots$

$\Phi^{-1}(0.95)=1.645$

為了得 90% 信賴區間，(勉強) 把 S_R^2 當作 R 的變異數.

(且假設 N 服從常態分佈)

$\Rightarrow R \sim N(r, S_R^2)$

$Pr\left(-1.645 \leq \frac{R-r}{S_R} \leq 1.645\right) = 0.90$

$Pr\left(R-1.645 \cdot S_R \leq r \leq R+1.645 \cdot S_R\right) = 0.90$

$\therefore$ 求 $R \pm 1.645 \cdot S_R$

$= [0.04427, 0.05573]$

(c).

$\hat{\mu_y} = 3 \quad \left(\frac{150}{50}\right) = \bar{Y}$

$\hat{\tau} = N \cdot \hat{\mu_y} = 3000$ (總數 $\hat{\tau}$ 之估計量)

$\overset{\shortparallel}{N\bar{Y}}$

N很大，所以在此無視有限母體修正.

$V[N\bar{Y}] = N^2 \cdot \frac{1}{n} \cdot V[Y_j] = N^2 \cdot \frac{1}{n} \cdot \sigma_Y^2 \approx 1000^2 \cdot \frac{1}{50} \cdot \frac{200}{49}$

以 S_Y^2 ($\overset{*}{(a)}$) 作為 σ_Y^2 之估計量.

$\shortparallel \frac{200}{49}$

$N\bar{Y} \sim N\left(\underset{總數}{\tau}, 1000^2 \cdot \frac{1}{50} \cdot \frac{200}{49}\right)$

95% 信賴區間

$\therefore \underbrace{N\bar{Y}}_{\substack{\shortparallel \\ 3000}} \pm \underbrace{\Phi^{-1}(0.975)}_{1.96} \cdot \sqrt{1000^2 \cdot \frac{1}{50} \cdot \frac{200}{49}}$

$\therefore [2440, 3560]$

CHA SHIN

7.49

(d) 「植被覆蓋面積」 × r = 「禽鳥總數」

(「 〃 (總面積) × R = 「 〃 」之估計量)

$(\rho=0.667)$

禽鳥總數跟植被覆蓋面積有高度相關性，

故 ratio estimate 較有效，((d) 的方法較佳)

7.53

$$\begin{pmatrix} 19.60199 & 8.401796 \\ 22.93532 & 5.621891 \\ 19.45274 & 7.363184 \\ 16.61092 & \end{pmatrix}$$

(a) • proportional allocation ...

依照 N_i 的比例來決定每個 stratum 的樣本數。

$$\begin{cases} n_1+n_2+\cdots+n_7=100 \\ n_1:n_2:\cdots n_7 = N_1:N_2:N_3:\cdots:N_7 \end{cases}$$

$= 20, 23, 19, 17, 8, 6, 7$

• optimal allocation $(N=N_1+N_2+\cdots+N_7)$

$$\bar{X}_s = \frac{N_1}{N}\overline{X_{1\cdot}} + \frac{N_2}{N}\cdot\overline{X_{2\cdot}} + \cdots \frac{N_7}{N}\overline{X_{7\cdot}}$$

$$\left(\overline{X_{i\cdot}} = \frac{1}{n_i}(X_{i1}+\cdots+X_{i,n_i})\right)$$

求 $\begin{pmatrix} n_1 \\ \vdots \\ n_7 \end{pmatrix}$ 使得 $V[\bar{X}]$ 為最小。 $(n_1+\cdots+n_7=100)$

$$V[\bar{X}_s] = \left(\frac{N_1}{N}\right)^2\cdot\frac{\sigma_1^2}{n_1} + \cdots + \left(\frac{N_7}{N}\right)^2\cdot\frac{\sigma_7^2}{n_7}$$

(注 雖然有限母體，但為了簡單，假設 $X_{i1} \sim X_{i,n_i}$ 為 iid 的隨機變數)

$$f\begin{pmatrix} n_1 \\ \vdots \\ n_7 \end{pmatrix} = V[\bar{X}_s]$$

但有限制條件，故利用 Lagrange 未定乘數法。

$(n_1+\cdots n_7=100)$

CHA SHIN

正確的數字　9.56006　12.10429
17.92416　8.58891
17.25997　15.22968
19.33294

$$L(n_1 \sim n_7;\lambda) = \left(\frac{N_1}{N}\right)^2 \cdot \frac{\sigma_1^2}{n_1} + \cdots \left(\frac{N_7}{N}\right)^2 \cdot \frac{\sigma_7^2}{n_7} - \lambda(n_1 + \cdots + n_7 - 100)$$

$$\begin{cases} \dfrac{\partial L}{\partial n_1} = 0 \Rightarrow \dfrac{-\sigma_1^2}{n_1^2} \cdot \left(\dfrac{N_1}{N}\right)^2 - \lambda = 0 \\ \vdots \\ \dfrac{\partial L}{\partial n_7} = 0 \Rightarrow \dfrac{-\sigma_1^2}{n_7^2} \left(\dfrac{N_7}{N}\right)^2 - \lambda = 0 \end{cases}$$

$$\therefore -\lambda = \left(\frac{\sigma_1}{n_1}\right)^2 \cdot \left(\frac{N_1}{N}\right)^2 = \left(\frac{\sigma_2}{n_2}\right)^2 \cdot \left(\frac{N_2}{N}\right)^2 = \cdots = \left(\frac{\sigma_7}{n_7}\right)^2 \cdot \left(\frac{N_7}{N}\right)^2$$

(抱歉 (b) 寫在後面)

$$\Rightarrow \underbrace{n_1 : n_2 : \cdots : n_7 = \frac{N_1}{N}\sigma_1 : \frac{N_2}{N}\sigma_2 : \cdots : \frac{N_7}{N}\sigma_7}$$

故依照 $\frac{N_i \sigma_i}{N}$ 的比例來算 $n_1 \sim n_7$
$(i = 1 \sim 7)$ $\Rightarrow (10, 18, 17, 19, 12, 9, 15)$

(c)

$$\mu = 所有的平均 = \frac{\mu_1 N_1 + \cdots + \mu_7 N_7}{N_1 + N_2 + \cdots + N_7} = \left(\frac{1}{N_1 + \cdots N_7} \sum_{i=1}^{7} \sum_{j=1}^{N_i} X_{ij}\right)$$

$$N \stackrel{def}{=} N_1 + N_2 + \cdots + N_7$$

$\underbrace{(26.31085)}$

σ^2（全部加起來的情況的變異數＝母體的變異數）

$$= \frac{1}{N} \sum_{i=1}^{7} \sum_{j=1}^{N_i} = \frac{1}{N} \sum_{i=1}^{7} \sum_{j=1}^{M_i} (X_{ij} - \mu)^2$$

我們不知道 X_{ij} 的值

$$= \frac{1}{N} \sum_{i=1}^{7} \sum_{j=1}^{N_i} \{ \underbrace{(X_{ij} - \mu_i) + (\mu_i - \mu)}_{= X_{ij} - \mu} \}^2$$

CHA SHIN

7.53

(c) … $= \frac{1}{N}\sum_{i=1}^{7}\sum_{j=1}^{N_i}\underbrace{\left\{(X_{ij}-\mu_i)^2+2(X_{ij}-\mu_i)(\mu_i-\mu)+(\mu_i-\mu)^2\right\}}$

① $\sum_{i=1}^{7}\sum_{j=1}^{N_i}(X_{ij}-\mu_i)^2 = \sum_{i=1}^{7} N_i\sigma_i^2$ $\left(\because\ \sigma_i^2=\frac{1}{N_i}\sum_{j=1}^{N_i}(X_{ij}-\mu_i\right.$

$+$

② $\sum_{i=1}^{7}2(\mu_i-\mu)\underbrace{\sum_{j=1}^{N_i}(X_{ij}-\mu_i)}_{\overset{\shortparallel}{0}}$ $\left(\because\ \mu_i=\frac{X_{i1}+\dots X_{iN_i}}{N_i}\right)$

$+$

③ $\sum_{i=1}^{7}\sum_{j=1}^{N_i}(\mu_i-\mu)^2 = \sum_{i=1}^{7}N_i(\mu_i-\mu)^2$

$= \frac{1}{N}\left\{\sum_{i=1}^{7}N_i\sigma_i^2+\sum_{i=1}^{7}N_i(\mu_i-\mu)^2\right\}$ → $\underset{\sim\sim\sim}{620.1678}$

$(N=N_1+N_2+\dots+N_7)$

(d) 根據(b), $V[\overline{X}_s]=\frac{1}{10}\left(\left(\frac{N_1}{N}\right)^2\sigma_1^2+\dots\left(\frac{N_7}{N}\right)^2\cdot\sigma_7^2\right)$

若 $Y_1,\dots Y_n$ 來自（全部加起來的）母體

$V[\overline{Y}]=\frac{\sigma^2}{n}$ (∵ 無視有限母體修正)

$=\frac{1}{10}\left(\left(\frac{N_1}{N}\right)^2\sigma_1^2+\dots+\left(\frac{N_7}{N}\right)^2\sigma_7^2\right)$

$\therefore n=\frac{\sigma^2}{\frac{1}{10}\left\{\left(\frac{N_1}{N}\right)^2\sigma_1^2+\dots\left(\frac{N_7}{N}\right)^2\sigma_7^2\right\}}=\underset{\sim\sim\sim}{139.3551}$

(e) $V[\overline{X}_s]=\left(\frac{N_1}{N}\right)^2\frac{\sigma_1^2}{70\left(\frac{N_1}{N}\right)}+\dots\left(\frac{N_7}{N}\right)^2\cdot\frac{\sigma_7^2}{70\left(\frac{N_7}{N}\right)}$

$=\frac{1}{70}\left\{\left(\frac{N_1}{N}\sigma_1^2+\dots+\frac{N_7}{N}\sigma_7^2\right)\right\}$

7.53

(e) $\therefore n = \dfrac{\sigma^2}{\frac{1}{70}\left\{\left(\frac{N_1}{N}\right)\sigma_1^2 + \cdots \left(\frac{N_7}{N}\right)\sigma_7^2\right\}} = 126.4621$

(b) ① proportional

$$Var\left[\frac{N_1}{N}\overline{X_{1\cdot}} + \cdots \frac{N_7}{N}\overline{X_{7\cdot}}\right]$$

$$= \left\{\left(\frac{N_1}{N}\right)^2 \cdot \frac{\sigma_1^2}{100 \cdot \frac{N_1}{N}} + \cdots + \left(\frac{N_7}{N}\right)^2 \cdot \frac{\sigma_7^2}{100 \cdot \frac{N_7}{N}}\right\}$$

$$= \frac{1}{100}\left(\underbrace{\frac{N_1}{N}\sigma_1^2 + \cdots + \frac{N_7}{N}\sigma_7^2}_{343.2788}\right)$$

$$\approx 3.432788.$$

② optimal

$$Var\left[\frac{N_1}{N}\overline{X_{1\cdot}} + \cdots + \frac{N_7}{N}\overline{X_{7\cdot}}\right]$$

$$= \left(\frac{N_1}{N}\right)^2 \cdot \frac{\sigma_1^2}{n_1} + \cdots \left(\frac{N_7}{N}\right)^2 \cdot \frac{\sigma_7^2}{n_7} = 2.8917741$$

③ Simple Random Sampling

根據 (c) $\sigma^2 = 620.1678$

$$\therefore V\left[\frac{Y_1 + \cdots + Y_{100}}{100}\right] = \frac{620.1678}{100} = 6.201678$$

CHA SHIN