MH4511 Sampling & Survey

Tutorial 7 Solution AY2025/26 Semester 1

Problem 7.1 (Solution)

This is a 2-stage sampling question. We know $N = 580, n = 12, M_i = 24, m_i = 3$.

To estimate the total:

$$\hat{t}_{unb} = \sum_{i=1}^{12}\sum_{j=1}^{3} w_{ij}y_{ij} = \frac{N}{n}\sum_{i=1}^{n}\hat{t}_i = \sum_{i=1}^{12}\sum_{j=1}^{3}\frac{N}{n}\frac{M_i}{m_i}y_{ij} = \frac{580}{12}\times\frac{24}{3}\times\sum_{i=1}^{12}\sum_{j=1}^{3}y_{ij} = \frac{580}{12}\times\frac{24}{3}\times 131$$

$$= 50653.33$$

$$Var(\hat{t}_{unb}) = N^2\left(1-\frac{n}{N}\right)\frac{s_t^2}{n} + \frac{N}{n}\sum_{i=1}^{n}\left(1-\frac{m_i}{M_i}\right)M_i^2\frac{s_i^2}{m_i}$$

$$= 580^2\left(1-\frac{12}{580}\right)\frac{2611.88}{12} + \frac{580}{12}\sum_{i=1}^{n}\left(1-\frac{3}{24}\right)24^2\frac{s_i^2}{3}$$

$$= 71{,}704{,}779 + \frac{580}{12}\left(1-\frac{3}{24}\right)24^2\frac{54.33}{3}$$

$$= 71{,}704{,}779 + 441{,}187 \approx 8{,}494^2$$

A 95% CI for the total number of worm fragments is approximated by (we use normal approximation):

$$\hat{t}_{unb} \pm z_{\alpha/2}\times SE(\hat{t}_{unb}) = 50653.33 \pm 1.96\times 8494 \approx (34005, 67302)$$

Case	1	2	3	4	5	6	7	8	9	10	11	12	Total	s_t^2
Can 1	1	4	0	3	4	0	5	3	7	3	4	0		
Can 2	5	2	1	6	9	7	5	0	3	1	7	0		
Can 3	7	4	2	6	8	3	1	2	5	4	9	0		
total	13	10	3	15	21	10	11	5	15	8	20	0	131	
$\hat{t}_i$	104	80	24	120	168	80	88	40	120	64	160	0	1048	**2611.88**
Std dev	3.06	1.15	1.00	1.73	2.65	3.51	2.31	1.53	2.00	1.53	2.52	0.00		
s_i^2	9.33	1.33	1.00	3.00	7.00	12.33	5.33	2.33	4.00	2.33	6.33	0.00	54.33	

Problem 7.2 (Solution)

a) Each of the department in the company is a psu.

b) Each of the secretary in the company is a ssu.

c) This is a 2-stage sampling question. We know $N = 20, n = 5$.

$$\hat{t}_{unb} = \frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}y_{ij} = \frac{N}{n}\sum_{i=1}^{n}\hat{t}_i = \frac{N}{n}\sum_{i=1}^{n}M_i\bar{y}_i = \frac{20}{5}\times 14402 = 57608$$

$$SE(\hat{t}_{unb}) = \sqrt{N^2\left(1-\frac{n}{N}\right)\frac{s_t^2}{n} + \frac{N}{n}\sum_{i=1}^{n}\left(1-\frac{m_i}{M_i}\right)M_i^2\frac{s_i^2}{m_i}}$$

$$= \sqrt{20^2\times\left(1-\frac{5}{20}\right)\times\frac{1600000}{5} + \frac{20}{5}\times\ 10616.286} \approx \sqrt{96{,}000{,}000 + 42{,}465.14} \approx 9800$$

d) Using ratio estimator, with $x =$ number of secretaries, $y =$ number of phone calls. Note that the total number of secretaries is unknown.

$$\hat{\bar{y}}_r = \frac{\sum_{i=1}^{n}\hat{t}_i}{\sum_{i=1}^{n}M_i} = \frac{14402}{147} = 97.97$$

$$SE(\hat{\bar{y}}_r) = \sqrt{\frac{1}{\bar{M}^2}\left[\left(1-\frac{n}{N}\right)\frac{s_r^2}{n} + \frac{1}{nN}\sum_{i=1}^{n}\left(1-\frac{m_i}{M_i}\right)M_i^2\frac{s_i^2}{m_i}\right]}$$

$$= \sqrt{\frac{1}{\left(\frac{147}{5}\right)^2}\left[\left(1-\frac{5}{20}\right)\times\frac{182000}{5} + \frac{1}{5\times 20}\times\ 10616.286\right]}$$

$$\approx \sqrt{\frac{1}{(29.4)^2}[27300 + 106.163]} \approx 5.631$$

Department (i)	Number of Secretaries (M_i)	Number of Secretaries Sampled (m_i)	Sample Mean $\bar{y}_i$	Sample Variance s_i^2	$\hat{t}_i = M_i \times \bar{y}_i$	$\left(1-\frac{m_i}{M_i}\right) \times M_i^2\frac{s_i^2}{m_i}$
1	45	9	102	20	4590	3600.000
2	36	7	90	16	3240	2386.286
3	20	4	76	22	1520	1760.000
4	18	4	94	26	1692	1638.000
5	28	6	120	12	3360	1232.000
Total	147	30			14402	10616.286

Problem 7.3 (Solution)

a) Primary sampling units (psu): the 40 blocks of households in the community.
Secondary sampling units (ssu): the 4000 households in the community.

b) We know $N = 40, n = 8$. (See table below.)

Using ratio estimator, $\hat{\bar{y}}_r = \frac{\sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} M_i} = \frac{\sum_{i=1}^{n} M_i \cdot \bar{y}_i}{\sum_{i=1}^{n} M_i} = \frac{1204}{860} = 1.40$

$$s_e^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\bar{y}_i \cdot M_i - \hat{\bar{y}}_r \cdot M_i)^2 = \frac{1}{n-1}\sum_{i=1}^{n}(t_i - \hat{\bar{y}}_r \cdot M_i)^2 = \frac{1}{7} \times 6218 = 888.286$$

$$SE(\hat{\bar{y}}_r) = \sqrt{\left(1-\frac{n}{N}\right)\frac{1}{n \cdot \bar{M}^2} \cdot s_e^2} = \sqrt{\left(1-\frac{8}{40}\right)\frac{1}{8 \cdot \bar{M}^2} \cdot 888.286} = 0.0942,$$

$$\text{where } \bar{M} = \frac{4000}{40} = 100$$

A 95% CI for the average number of newspapers purchased is given by:

$$\hat{\bar{y}}_r \pm t_{\frac{\alpha}{2},n-1} \times SE(\hat{\bar{y}}_r) = 1.40 \pm t_{0.025,7} \times 0.0942 = (1.177, 1.623)$$

$$\text{where } t_{0.025,7} = 2.365$$

Block	Number of households (M_i)	Average number of newspapers purchased ($\bar{y}_i$)	$t_i = M_i \cdot \bar{y}_i$	$\hat{\bar{y}}_r \cdot M_i$	$(t_i - \hat{\bar{y}}_r \cdot M_i)^2$
1	80	1.4	112	112	0
2	85	1.6	136	119	289
3	100	1.5	150	140	100
4	125	1.8	225	175	2500
5	120	1.4	168	168	0
6	125	1.2	150	175	625
7	130	1.0	130	182	2704
8	95	1.4	133	133	0
total	860		1204		6218