Lösningsförslag

Uppgift 1

Kortfattad lösning till tentamen i Surveymetodik HSTB21, HSTB52, 2006 03 27

1) $n_0 = \frac{1.96^2 0.5 \cdot 0.5}{0.05^2} \approx 384.16$

$n = \frac{384.16}{1 + \frac{384.16}{2500}} = 332.99 \Rightarrow n = 333$

Uppgift 2

1. Svar: För att lösa denna uppgift använder vi oss av (2.19) i [1, Lohr,s. 38] för att beräkna variansen. Detta ger:

$$\hat{p} = 0.19$$

$$\hat{V}(\hat{p}) = \left(1 - \frac{n}{N}\right)\frac{\hat{p}(1-\hat{p})}{n-1} = \left(1 - \frac{200}{13417}\right)\frac{0.19 \cdot 0.81}{199} \approx 0.028^2$$

$$\hat{p} \pm z_{\alpha/2} SE(\hat{p}) = 0.19 \pm 1.96 \cdot 0.028 \rightarrow [0.135, 0.245]$$

2.Svar: För att lösa denna uppgift använder vi att totalen är proportionen multiplicerat med populationstotalen:

$$\hat{t} = N \cdot \hat{p} = 13417 \cdot 0.19 = 2549.23$$

På liknande sätt beräknar vi variansen och konfidensintervallen:

$$V(\hat{t}) = V(\hat{p} \cdot N) = N^2 V(\hat{p}) = 13417^2 \cdot 0.028^2 = 375.676^2$$

$$\hat{t} \pm z_{\alpha/2} SE(\hat{t}) = 2549.23 \pm 1.96 \cdot 375.676 \rightarrow [1812.905, 3285.555]$$

3.Svar: För att lösa denna uppgift använder vi oss av (2.24) och (2.25) i [1, Lohr,s. 47]. Vi är intresserade av att få ett konfidensintervall av storleken $\hat{p}\pm$ 0.05. Detta innebär att $e = 0.05$ i detta fall. Vi behöver också anta standardavvikelse för populationen. Då $p = 0.5$ uppnås det maximala värdet för $S^2 = 0.25$ $(p(1-p))$, för att vara säkra sätter vi därför S^2 till 0.25. Detta ger:

$$n_0 = \left(\frac{z_{\alpha/2}S}{e}\right)^2 = \frac{1.96^2 \cdot 0.25}{0.05^2} = 384.16$$

som sedan används för att beräkna det nya n:

$$n = \frac{n_0}{1+\frac{n_0}{N}} = \frac{384.16}{1+\frac{384.16}{13417}} = 373.467 \to 374$$

Uppgift 3

$$S^2 = \frac{\sum(y_i-\bar{y}_U)^2}{N-1} = \left[\bar{y}_U = \frac{\sum y_i}{N} = P\text{ , då } Y \text{ antar } 1 \text{ och } 0\right] = \frac{\sum(y_i-P)^2}{N-1} = \frac{\sum(y_i^2-2Py_i+P^2)}{N-1}$$
$$= \frac{\sum y_i^2 - 2P\sum y_i + NP^2}{N-1} = \left[\sum y_i^2 = \sum y_i = \frac{N\sum y_i}{N} = NP\right] = \frac{NP-2NP^2+NP^2}{N-1}$$
$$= \frac{NP-NP^2}{N-1} = \frac{N}{N-1}P(1-P) \quad v.s.v$$

Uppgift 4

· Kortfattad lösning till tentamen i HST B21, Survey metodik med uppsats. 2005 06 14

1) a) x = betyg

$\mu = \frac{1}{6}\sum x_i = 3.5$ (1p)

$S^2 = \frac{1}{5}\left(\sum x_i^2 - 6\mu^2\right) = \frac{1}{5}\left(81 - 6 \cdot 3.5^2\right) \approx \frac{7.5}{5} = 1.5$

b) OSU om 3 nr 1, 2, 6 t ex ger värdena 5, 2, 2

$\hat{\mu} = \frac{1}{3}(5+2+2) = 3$ (1.5p)

$s^2 = \frac{1}{2}\left(\underbrace{\sum x_i^2}_{33} - 3 \cdot 3^2\right) = \frac{6}{2} = 3$

c) $\binom{N}{n} = \binom{6}{3} = \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3} = 20$ (0.5p)

d)

$\bar{x}$	$p(\bar{x})$	
2.667	3/20	= 0.15
3.000	1/20	= 0.05
3.333	6/20	= 0.30
3.667	6/20	= 0.30
4.000	1/20	0.05
4.333	3/20	0.15
	20/20	

(1p)

e) $\sum \bar{x}\, p(\bar{x}) = 3.5 \Rightarrow$ vvr

medel av alla urvalsvar $= 1.5$ (1p)

eller $\sum s^2 p(s^2) = 1.5$

Tot (5p)

1

Uppgift 5

5) a) CI för SUM 95%

(758, 1236) hål

b) $w_{hj} = \frac{N_h}{n_h}$ Stratum 1 har 17 vikter som alla är $\frac{68}{17} = 4$

Stratum 2 har 12 vikter som alla är $\frac{84}{12} = 7$

Stratum 3 har 11 vikter som alla är $\frac{48}{11} = 4,36$

Uppgift 6

Se (2.30) på sidan 52 i Lohr 2nd edition.

Uppgift 7

2) a) $\hat{p}_{str} = \frac{1}{105}\left(35 \cdot \frac{11}{20} + 40 \cdot \frac{10}{20} + 30 \cdot \frac{5}{20}\right) = 0.445$

$SE^2[\hat{p}_{str}] = \frac{1}{105^2}\left(35^2 \cdot \frac{0.55 \cdot 0.45}{19}\left(1 - \frac{20}{3500}\right)\right.$

$\left. + 40^2 \frac{0.5^2}{19}\left(1 - \frac{20}{4000}\right) + 30^2 \frac{0.25 \cdot 0.75}{19}\left(1 - \frac{20}{3000}\right)\right)$

$SE[\hat{p}_{str}] = 0,06433$

95% KI $\quad 44.5\% \pm 12.6\%$

~~b)~~ c) p_1 = andel i Lambohov $\quad p_2$ = andel i innerstaden

$\hat{p}_1 - \hat{p}_2 \pm 1.96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

urvalsfraktionen ≈ 0

$30.0\% \pm 29.7\%$

Intervallet missar 0:an precis. Det är med stor sannolikhet skillnad i attityd beroende på om man bor i Lambohov eller i innerstaden

d) $n_i = \frac{60 \, N_i \sqrt{\hat{p}_i(1-\hat{p}_i)}}{\sum_i N_i \sqrt{\hat{p}_i(1-\hat{p}_i)}}$

ger $n_1 = 21 \quad n_2 = 24 \quad n_3 = 15$

Uppgift 8

a) Som ett första steg beräknar vi punktskattningen (3.2) i [1].

$$\hat{\bar{y}}_{str} = \sum_{h=1}^{H} \frac{N_h}{N} \bar{y}_h$$

där resultatet för respektive strata framgår av tabellen nedan.

	N_h	$\bar{y}_h$	$\frac{N_h}{N} \cdot \bar{y}_h$
1	2902	1502.67	244.08
2	10730	2015.26	1210.33
3	4234	2894.03	685.85

Detta ger att:

$$\begin{aligned} \hat{\bar{y}}_{str} &= 244.08 + 1210.33 + 685.85 \\ &= 2140.25586 \end{aligned}$$

Sedan beräknas variansen med hjälp av:

$$\hat{V}(\hat{\bar{y}}_{str}) = \sum_{h=1}^{H} \left(1 - \frac{n_h}{N_h}\right) \left(\frac{N_h}{N}\right)^2 \frac{s_h^2}{n_h}$$

Notera att vi bara har observerat (fått svar från) n_{hr} i respektive strata. Respektive del framgår av tabellen nedan:

	N_h	n_{rh}	s_h	$\left(\frac{N_h}{N}\right)^2$	$1 - \frac{n_{rh}}{N_h}$	$\frac{s_h{}^2}{n_{rh}}$	$\left(1 - \frac{n_{rh}}{N_h}\right) \left(\frac{N_h}{N}\right)^2 \frac{s_h{}^2}{n_{rh}}$
1	2902	98.00	306.18	0.03	0.97	956.59	24.39
2	10730	378.00	459.84	0.36	0.96	559.40	194.67
3	4234	159.00	734.33	0.06	0.96	3391.45	183.32

Detta ger således att

$$\begin{aligned} \hat{V}(\hat{\bar{y}}_{str}) &= 24.39 + 194.67 + 183.32 \\ &= 402.37258 \end{aligned}$$

Konfidensintervallen kan sedan beräknas på följande sätt

$$\hat{\bar{y}}_{str} \pm z_{\alpha/2} \cdot \sqrt{\hat{V}(\hat{\bar{y}}_{str})} = 2140.25586 \pm 2.576 \cdot 20.05923 \rightarrow [2088.58329, 2191.92843]$$

b)
Designvikterna beräknas på följande sätt.

$$d_h = \frac{1}{\pi_h} = \frac{1}{\frac{n_h}{N_h}} = \frac{N_h}{n_h}$$

Det ger följande resultat i vårt exempel:

	N_h	n_h	$\frac{N_h}{n_h}$
1	2902	162	17.91
2	10730	601	17.85
3	4234	237	17.86

c)

$$n_{6-} = \frac{N_{6-}}{N} * n = \frac{2902}{17866} * 1000 = 162.43$$

$$n_{2-6} = \frac{10730}{17866} * 1000 = 600.58$$

$$n_{<2} = \frac{4234}{17866} * 1000 = 236.99$$

d)

$$\hat{V}(\hat{\bar{y}}_{osu}) = (1 - \frac{n_r}{N}) * \frac{s^2}{n_r} = (1 - \frac{635}{17866}) * \frac{699.18^2}{635} = 742.485$$

$$\text{Designeffekt} = \frac{\hat{V}(\hat{\bar{y}}_{str})}{\hat{V}(\hat{\bar{y}}_{osu})} = \frac{402.373}{742.485} = 0.542$$