第五章

5.2

	Stratum			
	I	II	III	
N_i	132	92	27	
σ_i^2	36	25	9	
$N_i\sigma_i$	792	460	81	$\sum N_i\sigma_i = 1333$

From Equation (5.9), $n_i = n\dfrac{N_i o_i}{\sum N_i o_i}$

Then

$$n_1 = 30(792/1333) = 17.82 \approx 18$$
$$n_2 = 30(460/1333) = 10.35 \approx 10$$
$$n_3 = 30(81/1333) = 1.82 \approx 2$$

5.3

	Stratum			
	I	II	III	
N_i	132	92	27	$N = 251$
n_i	18	10	2	
$\bar{y}_i$	8.83	6.7	4.5	
s_i^2	81.56	50.46	24.50	

The estiamte of the total number of man-hours lost during the given month is, from Equation (5.3),

$$\hat{\tau} = N\bar{y}_{st} = 132(8.83) + 92(6.7) + 27(4.5) = 1903.9$$

The estimated variance of $\hat{\tau}$ is, from Equation (5.4),

$$V(\hat{\tau}) = \hat{V}(N\bar{y}_{st}) = \sum N_i^2\left(\frac{N_i - n_i}{N_i}\right)\frac{s_i^2}{n_i} = 114519.9$$

with a bound on the error of estimation

$$B = 2\sqrt{\hat{V}(\bar{y}_{st})} = 676.8$$

5.5

	Stratum			
	I	II	III	
N_i	112	68	39	$N = 219$
c_i	9	25	36	
σ_i^2	2.25	3.24	3.24	

$$n = \frac{\left(\sum N_k\sigma_k / \sqrt{c_k}\right)\left(\sum N_i\sigma_i\sqrt{c_i}\right)}{N^2D + \sum N_i\sigma_i^2}$$

$$B = 2\sqrt{V(\bar{y}_{st})} \text{ or } V(\bar{y}_{st}) = \frac{B^2}{4} = D = 0.1$$

$$\sum\left(\frac{N_k\sigma_k}{\sqrt{c_k}}\right) = \frac{112(1.5)}{3} + \frac{68(1.8)}{5} + \frac{39(1.8)}{6} = 92.18$$

$$\sum N_i\sigma_i\sqrt{c_i} = 112(1.5)(3) + 68(1.8)(5) + 39(1.8)(6) = 1537.2$$

$$\sum N_i\sigma_i^2 = 112(2.25) + 68(3.24) + 39(3.24) = 598.68$$

$$n = \frac{92.18(1537.2)}{219^2(.1) + 598.68} = 26.3 \approx 27$$

To allocate the $n = 27$ to the three strata, use

$$n_i = n\frac{N_i\sigma_i / \sqrt{c_i}}{\sum N_k\sigma_k / \sqrt{c_k}}$$

Then

$$n_1 = 27\frac{112(1.5)/3}{92.18} = 16.40 \quad n_2 = 27\frac{68(1.8)/5}{92.18} = 7.17 \quad n_3 = 27\frac{39(1.8)/6}{92.18} = 3.43$$

Rounding off yields: $n_1 = 16,\ n_2 = 7,\ n_3 = 3$

This adds to total sample size of 26, not 27. Add 1 to one of the sample sizes to achieve n = 27. Add to stratum 3 because 3.43 is closer to the next higher integer than any other sample sizes.

5.8

$$n = \frac{\left(\sum N_i\sigma_i\right)^2}{N^2D + \sum N_i\sigma_i^2} \qquad B = 1, \quad D = \frac{B^2}{4} = .25$$

Use $\frac{\text{range}}{4}$ to estimate σ_i.

	Stratum		
	I	II	
N_i	43	53	$N = 96$
Range	20-5=15	14-3=11	
σ_i	$\frac{15}{4}=3.75$	$\frac{11}{4}=2.75$	

$$n = \frac{\left[43(3.75) + 53(2.75)\right]^2}{(96)^2(.25) + \left[43(3.75)^2 + 53(2.75)^2\right]} = \frac{94249}{3309.5} = 28.48 \approx 29$$

5.10

	Stratum				
	I	II	III	IV	Tot
N_i	86	72	52	30	240
n_i	14	12	9	5	40
$\bar{y}_i$	63.36	183.00	340.557	472.40	
s_i^2	1071.79	9054.18	16794.28	72376.30	

$$\hat{\tau} = N\bar{y}_{st} = \sum N_i \bar{y}_i = 50505.60$$

$$\hat{V}(\hat{\tau}) = \sum N_i^2 \left(\frac{N_i - n_i}{N_i} \right) \frac{s_i^2}{n_i} = 18762430.98$$

$$B = 2\sqrt{\hat{V}(\hat{\tau})} = 8663.12$$

5.11 $$n = \frac{\left(\sum N_i \sigma_i\right)^2}{N^2 D + \sum N_i \sigma_i^2}$$

$$N^2 D = \frac{B^2}{4} = \frac{(5000)^2}{4} = (2500)^2$$

Use s_i^2 to estimate σ_i^2.

$\sum N_i s_i = 86\sqrt{1071.79} + 72\sqrt{16974.28} + 30\sqrt{72376.30} = 24476.20$

$\sum N_i s_i^2 = 86(1071.79) + 72(9054.18) + 52(16794.28) + 30(72376.30) =$
3788666.11

$$n = \frac{24476.20^2}{(2500)^2 + 3788666.11} = 59.68 \approx 60$$

5.13

	Stratum				
	I	II	III	IV	Tot
N_i	97	43	145	68	353
p_i	.9	.9	.5	.5	
c_i	4	4	8	8	
$N_i\sqrt{p_i q_i / c_i}$	14.55	6.45	25.63	12.02	58.65
a_i	.248	.110	.437	.205	
$N_i^2 p_i q_i / a_i$	3413.63	1513.26	12027.53	5640.50	22594.92
$N_i p_i q_i$	8.73	3.87	36.25	17.00	65.85
n_i	39	17	69	33	158

where

$$a_i = \frac{N_i\sqrt{p_i q_i / c_i}}{\sum N_k \sqrt{p_k q_k / c_k}}$$

$$D = \frac{B^2}{4} = \frac{.05^2}{4} = (.025)^2$$

$$n = \frac{\sum N_i^2 p_i q_i / a_i}{N^2 D + \sum N_i p_i q_i} = \frac{22594.92}{(353)^2(.025)^2 + 65.85} = 157.20 \approx 158$$

$$n_i = n a_i$$

5.14

	Stratum				
	I	II	III	IV	Tot
N_i	97	43	145	68	353
n_i	39	17	69	33	
$\hat{p}_i$	.87	.93	.60	.53	
$N_i^2\left(\frac{N_i - n_i}{N_i}\right)\frac{\hat{p}_i\hat{q}_i}{n_i - 1}$	16.74	4.55	38.89	18.53	78.81

$$\hat{p}_{st} = \frac{1}{N}\sum N_i\hat{p}_i = \frac{1}{353}[97(.87) + 43(.93) + 145(.60) + 68(.53)] = .701$$

$$B = 2\sqrt{\frac{1}{N^2}\sum N_i^2\left(\frac{N_i - n_i}{N_i}\right)\frac{\hat{p}_i\hat{q}_i}{n_i}} = 2\sqrt{\frac{78.71}{(353)^2}} = .0503$$

5.15 Total cost $= \sum n_i c_i = 400$

where c_i is the cost of obtaining one observation from stratum i

But, $c_i = c_2 = \frac{c_3}{2} = \frac{c_4}{2}$

Writing the equation in terms of c_1 only gives

$$\begin{aligned}
n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 &= 400 \\
n_1c_1 + n_2c_1 + 2n_3c_1 + 2n_4c_1 &= 400 \\
c_1(n_1 + n_2 + 2n_3 + 2n_4) &= 400 \\
n_1 + n_2 + 2n_3 + 2n_4 &= 400/c_1 = 100 \\
na_1 + na_2 + 2na_3 + 2na_4 &= 100 \\
n(a_1 + a_2 + 2a_3 + 2a_4) &= 100
\end{aligned}$$

Using the sampling fractions from Exercise 5.13,

$$n = 100/[.248 + .110 + 2(.437) + 2(.205)] = 60.90 \cong 61$$

$$\begin{aligned}
n_1 &= na_1 = 61(.248) = 15.13 \approx 15 \\
n_2 &= na_2 = 61(.110) = 6.71 \approx 7 \\
n_3 &= na_3 = 61(.437) = 26.66 \approx 27 \\
n_4 &= na_4 = 61(.205) = 12.50 \approx 12
\end{aligned}$$

Total cost = 15(4) + 4(4) + 27(8) + 12(8) = \$400.

5.17

No. of Employees	Frequency	$\sqrt{\text{Frequency}}$	Cumulative $\sqrt{\text{Frequency}}$
0-10	2	1.41	1.41
11-20	4	2.00	3.41
21-30	6	2.45	5.86
31-40	6	2.45	8.31
41-50	5	2.24	10.55
51-60	8	2.83	13.38
61-70	10	3.16	16.54
71-80	14	3.74	20.28
81-90	19	4.36	24.64
91-100	13	3.61	28.25
101-110	3	1.73	29.98
111-120	7	2.65	32.62

L = 4 strata, 32.62 / 4 = 8.155

Stratum boundaries should be as close as possible to: 8.155, 16.312, 24.468
Choose boundaries of 8.31, 16.54, 24.64.

Stratum 1: 0-40 employees
Stratum 2: 41-70 employees
Stratum 3: 71-90 employees
Stratum 4: 91-120 employees

5.19

	Stratum I	Stratum II
$\bar{y}_i$	63.47	64.30
s_i^2	1.07	1.30
n_i	8	7

N_1 and N_2 are unknown. Assume that they are equal. Let N' represent both of these terms.

$$\hat{\mu} = \bar{y}_{st} = \frac{1}{N}\sum N_i\bar{y}_i = \frac{1}{2N'}\left[N'\bar{y}_1 + N'\bar{y}_2\right] = \frac{1}{2}(\bar{y}_1 + \bar{y}_2) = \frac{1}{2}(63.47 + 64.30) = 63.88$$

$$B = 2\sqrt{\frac{1}{N^2}\sum N_i^2\frac{s_i^2}{n_i}} = 2\sqrt{\frac{N'^2}{(2N')^2}\sum\frac{s_i^2}{n_i}} = 2\sqrt{\frac{1}{4}\sum\frac{s_i^2}{n_i}} = 2\sqrt{\frac{1}{4}\left(\frac{1.07}{6} + \frac{1.30}{6}\right)} = .628$$

The shipment appears to be below the standard in average weight.

5.21 **(a)** $\hat{p} = \frac{\sum y_i}{n} = \frac{6+10}{100} = .16$

$$B = 2\sqrt{\frac{\hat{p}\hat{q}}{n-1}} = 2\sqrt{\frac{.16(.84)}{99}} = .074 \quad \text{(ignoring fpc)}$$

(b) $\hat{p}_{st} = \frac{1}{N}\sum N_i \hat{p}_i = \sum \frac{N_i}{N}\hat{p}_i = .6\frac{6}{38} + .4\frac{10}{62} = .16$

$$B = 2\sqrt{\frac{1}{N^2}\sum N_i^2 \frac{\hat{p}_i\hat{q}_i}{n_i - 1}} = 2\sqrt{\sum\left(\frac{N_i}{N}\right)^2 \frac{\hat{p}_i\hat{q}_i}{n_i - 1}}$$

$$= 2\sqrt{(.6)^2\left(\frac{6}{38}\right)\left(\frac{32}{38}\right)\left(\frac{1}{37}\right) + (.4)^2\left(\frac{10}{62}\right)\left(\frac{52}{62}\right)\left(\frac{1}{61}\right)} = .081$$

There seems to be no good reason to poststratify here.

5.27 **(a)**

	Stratum		Total
	I	II	
N_i	20	26	
σ_i	25	47.5	
$N_i\sigma_i$	500	1235	1735
w_i	.29	.71	
$N_i\sigma_i^2$	12500	58662.5	71162.5

where

$\frac{\text{range}}{4}$ is used to estimate σ_i.

$\sigma_1 \approx \frac{100-0}{4} = 25$ for small plants (stratum I)

$\sigma_2 \approx \frac{200-10}{4} = 47.5$ for large plants (stratum II)

$$a_i = \frac{N_i\sigma_i}{\sum N_i\sigma_i}$$

$a_1 = \frac{500}{1735} = .29 \qquad a_2 = \frac{1235}{1735} = .71$

(b) $B = 100$

$$N^2 D = \frac{B^2}{4} = \frac{(100)^2}{4} = 2500$$

$$n = \frac{\left(\sum N_i\sigma_i\right)^2}{N^2 D + \sum N_i\sigma_i^2} = \frac{(1735)^2}{2500 + 71162.5} = 40.87 \approx 41$$

$n_1 = na_1 = 41(.29) = 11.9 \approx 12$

$n_2 = na_2 = 41(.71) = 29.1 \approx 29$

Since there is only 26 "large" plants, we allocate $n_1 = 15$, $n_2 = 26$.

5.31

	Stratum		
	I	II	III
w_i	.5	.1	.4
No	417	29	240
	31.4	17.6	21.8
Yes	913	136	860
	68.6	82.4	78.2
Tot	1330	165	1100
	100%	100%	100%

(a) $\hat{p}_{st} = \sum \frac{N_i}{N}\hat{p}_i = \sum a_i\hat{p}_i = (.5)(.686) + (.1)(.824) + (.4)(.782) = .738$

$$\hat{V}(\hat{p}_{st}) = \sum a_i^2 \frac{\hat{p}_i\hat{q}_i}{n_i - 1} \text{ (ignoring the fpc)}$$

$$= (.5)^2 \frac{(.686)(.314)}{1329} + (.1)^2 \frac{(.824)(.176)}{164} + (.4)^2 \frac{(.782)(.218)}{1099} = .74 \times 10^{-4}$$

(b) $\hat{p}_1 - \hat{p}_2 = = .686 - .824 = -.138$

$$B = 2\sqrt{\hat{V}(\hat{p}_1 - \hat{p}_2)} = 2\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 2\sqrt{\frac{(.686)(.314)}{1330} + \frac{(.824)(.176)}{165}} = .0645$$

(c) $\hat{p}_1 - \hat{p}_2 = = .686 - .782 = -.096$

$$B = 2\sqrt{\hat{V}(\hat{p}_1 - \hat{p}_2)} = 2\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 2\sqrt{\frac{(.686)(.314)}{1330} + \frac{(.782)(.218)}{1100}} = .036$$

5.32 **(a)** $\bar{y}_{st} = \sum a_i\bar{y}_i = (.5)(7.63) + (.1)(7.74) + (.4)(6.55) = 7.209$

$$B = 2\sqrt{\hat{V}(\bar{y}_{st})} = 2\sqrt{\sum a_i^2 \frac{s_i^2}{n_i}} = 2\sqrt{(.5)^2 \frac{(.15)^2}{1347} + (.1)^2 \frac{(.35)^2}{163} + (.4)^2 \frac{(.11)^2}{1095}} =$$

.0073

(b) $\bar{y}_1 - \bar{y}_2 = 7.63 - 7.74 = -.11$

$$B = 2\sqrt{\hat{V}(\bar{y}_1 - \bar{y}_2)} = 2\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2\sqrt{\frac{(.15)^2}{1347} + \frac{(.35)^2}{163}} = .055$$

The confidence interval just touches 0, so there is no strong evidence that the residents have a higher mean than the non-resident anesthesiologists.

$\bar{y}_2 - \bar{y}_3 = 7.74 - 6.55 = 1.19$

$$B = 2\sqrt{\hat{V}(\bar{y}_2 - \bar{y}_3)} = 2\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2\sqrt{\frac{(.35)^2}{163} + \frac{(.11)^2}{1095}} = .055$$

The confidence interval overlaps 0, so there is evidence that the residents have a higher mean than the nurses.