

instead of being able to calculate the psu total t_i from all students in a sampled class, we estimate the psu total by $\hat{t}_i = M_i \bar{y}_i$. We also calculate the within-psu variances s_i^2 .

TABLE 5.7

Calculations using formulas for math scores in Example 5.7.

School	M_i	\bar{y}_i	s_i^2	\hat{t}_i	$M_i^2 \left(1 - \frac{m_i}{M_i}\right) \frac{s_i^2}{m_i}$	$M_i^2 (\bar{y}_i - \hat{y}_r)^2$
9	163	34.75	74.51	5664.25	86841	70336
17	180	40.80	111.01	7344.00	159855	1909563
18	114	37.85	124.87	4314.90	66906	290396
22	367	27.95	109.31	10257.65	696046	3604196
35	109	46.10	50.31	5024.90	24401	2000806
43	219	32.20	162.80	7051.80	354749	40855
46	318	30.60	86.57	9730.80	410178	643681
55	259	36.35	141.61	9414.65	438284	698572
62	311	35.40	97.83	11009.40	442693	501495
75	263	24.60	69.52	6469.80	222134	5024481
Sum	2303			76282.15	2902087	14784382

We use the ratio estimator to estimate the mean math score. From (5.30),

$$\hat{y}_r = \frac{\sum_{i \in \mathcal{S}} \hat{t}_i}{\sum_{i \in \mathcal{S}} M_i} = \frac{76282.15}{2303} = 33.12.$$

Or we can, equivalently, use (5.31) to calculate \hat{y}_r using the sampling weights (as is done by most software packages). The weight for student j in school i is:

$$w_{ij} = \frac{N}{n} \frac{M_i}{m_i} = \frac{75}{10} \frac{M_i}{20}.$$

Using (5.31), the estimated mean again is

$$\hat{y}_r = \frac{\sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}_i} w_{ij} y_{ij}}{\sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}_i} w_{ij}} = \frac{572116.1}{17272.5} = 33.12.$$

The weights do not allow us to calculate the standard error, however. We need the clustering information to do that. From Table 5.7,

$$s_r^2 = \frac{1}{n-1} \sum_{i \in \mathcal{S}} M_i^2 (\bar{y}_i - \hat{y}_r)^2 = \frac{14784382}{9} = 1642709$$

and $\bar{M} = 2303/10 = 230.3$. The output in Table 5.6 uses the with-replacement variance

$$\hat{V}_{\text{WR}}(\hat{y}_r) = \frac{s_r^2}{n\bar{M}^2} = \frac{1642709}{(10)(230.3)^2} = 3.097$$

with $\text{SE}_{\text{WR}}(\hat{y}_r) = \sqrt{3.097} = 1.76$.