6.1 A forester is interested in estimating the total volume of trees in a timber sale. He records the volume for each tree in a simple random sample. In addition, he measures the basal area for each tree marked for sale. He then uses a ratio estimator of total volume.

The forester decides to take a simple random sample of n = 12 from the N = 250 trees marked for sale. Let x denote basal area and y the cubic-foot volume for a tree. The total basal area for all 250 trees, τ_x, is 75 square feet. Use the data in the accompanying table to estimate τ_y, the total cubic-foot volume for those trees marked for sale, and place a bound on the error of estimation.

Trees sampled	Square-foot basal area, x	Volume, y
1	0.3	6
2	0.5	9
3	0.4	7
4	0.9	19
5	0.7	15
6	0.2	5
7	0.6	12
8	0.5	9
9	0.8	20
10	0.4	9
11	0.8	18
12	0.6	13

$$\sum x_i = 0.3 + 0.5 + 0.4 + 0.9 + 0.7 + 0.2 + 0.6 + 0.5 + 0.8 + 0.4 + 0.8 + 0.6 = 6.7$$

$$\sum y_i = 6 + 9 + 7 + 19 + 15 + 5 + 12 + 9 + 20 + 9 + 18 + 13 = 142$$

$$r = \frac{\sum y_i}{\sum x_i} = \frac{142}{6.7} = \mathbf{21.194}, \hat{\tau}_y = r\tau_x = 21.194(75) = \mathbf{1589.55}$$

$$s_r^2 = \frac{1}{n-1}\sum (y_i - rx_i)^2 = 1.7503$$

$$\hat{V}(\hat{\tau}_y) = N\left(\frac{N-n}{n}\right)s_r^2; B = 2\sqrt{\hat{V}(\hat{\tau}_y)} = 2\sqrt{250\left(\frac{250-12}{12}\right)(1.7503)} = \mathbf{186.32}$$

Ans: $\hat{\boldsymbol{\tau}}_{\boldsymbol{y}} = \mathbf{1589.55}; \boldsymbol{B} = \mathbf{186.32}$

6.2 Use the data in Exercise 6.1 to compute an estimate of τ_y, using $N\bar{y}$. Place a bound on the error of estimation. Compare your results with those obtained in Exercise 6.1 Why is the estimate $N\bar{y}$, which does not use any basal-area data, much larger than the ratio estimate? (Look at μ_x and $\bar{x}$. Speculate about the reason for this discrepancy?)

$$\bar{y} = \frac{142}{12} = 11.83; \hat{\tau}_y = N\bar{y} = 250(11.83) = \mathbf{2958.33}$$

$$B = 2\sqrt{\hat{V}(N\bar{y})} = 2\sqrt{250^2\left(1-\frac{12}{250}\right)\left(\frac{26.87879}{12}\right)} = \mathbf{730.13}$$

Conclusion: the estimate $N\bar{y}$ is much larger because s^2 compares y_i observations to an average; whereas s_r^2 compares y_i observations to rx_i.

6.4 A corporation is interested in estimating the total earnings from sales of color television sets at the end of a three-month period. The total earnings figures are available for all districts within the corporation for the corresponding three-month period of the previous year. A simple random sample of 13 district offices is selected from the 123 offices within the corporation. Using a ratio estimator, estimate τ_y and place a bound on the error of estimation. Use the data in the accompanying table and take $\tau_x = 128{,}200$.

Office	Three-month data from previous year, x_i	Three-month data from current year, y_i
1	550	610
2	720	780
3	1500	1600
4	1020	1030
5	620	600
6	980	1050
7	928	977
8	1200	1440
9	1350	1570
10	1750	2210
11	670	980
12	729	865
13	1530	1710

$$\sum x_i = 550 + 720 + 1500 + 1020 + 620 + 980 + 928 + 1200 + 1350 + 1750 + 670 + 729 + 1530 = 13547$$

$$\sum y_i = 610 + 780 + 1600 + 1030 + 600 + 1050 + 977 + 1440 + 1570 + 2210 + 980 + 865 + 1710 = 15422$$

$$r = \frac{\sum y_i}{\sum x_i} = \frac{15422}{13547} = 1.138, \hat{\tau}_y = r\tau_x = 1.138(128200) = 145943.78$$

$$s_r^2 = \frac{1}{n-1}\sum (y_i - rx_i)^2 = 12996.6$$

```
> x=c(550,720,1500,1020,620,980,928,1200,1350,1750,670,729,1530)
> y=c(610,780,1600,1030,600,1050,977,1440,1570,2210,980,865,1710)
> s=c((y-1.138*x)^2)
> sum(s)/12
[1] 12996.6
```

$$\hat{V}(\hat{\tau}_y) = \tau_x^2\left(\frac{N-n}{nN}\right)\left(\frac{s_r^2}{\mu_x^2}\right) = (N\mu_x)^2\left(\frac{N-n}{nN}\right)\left(\frac{s_r^2}{\mu_x^2}\right) = N\left(\frac{N-n}{n}\right)s_r^2$$

$$B = 2\sqrt{\hat{V}(\hat{\tau}_y)} = 2\sqrt{123\left(\frac{123-13}{13}\right)(12996.6)} = 7353.67$$

Ans: $\boldsymbol{\hat{\tau}_y = 145943.78, B = 7353.67}$

6.5 Use the data in Exercise 6.4 to estimate the mean earnings for offices within the corporation. Place a bound on the error of estimation.

$$\hat{\mu}_y = r\mu_x = r\frac{\tau_x}{N} = 1.138\left(\frac{128200}{123}\right) = \mathbf{1186.53}$$

$$\hat{V}(\hat{\mu}_y) = \mu_x^2\hat{V}(r) = \mu_x^2\left(\frac{N-n}{nN}\right)\frac{1}{\mu_x^2}s_r^2 = \frac{N-n}{nN}s_r^2$$

$$B = 2\sqrt{\hat{V}(\hat{\mu}_y)} = 2\sqrt{\frac{123-13}{13(123)}12996.6} = \mathbf{59.80}$$

Ans: $\hat{\boldsymbol{\mu}}_y = \mathbf{1186.53}, \boldsymbol{B} = \mathbf{59.80}$

6.9 A forest resource manager is interested in estimating the number of dead fir trees in a 300-acre area of heavy infestation. Using an aerial photo, she divides the area into 200 1.5-acre plots. Let *x* denote the photo count of dead firs and *y* the actual ground count for a simple random sample of *n* = 10 plots. The total number of dead fir trees obtained from the photo count is $\tau_x = 4200$. Use the sample data in the accompanying table to estimate τ_y, the total number of dead firs in the 300-acre area. Place a bound on the error of estimation.
(use a ratio estimator)

Plot sampled	Photo count, x_i	Ground count, y_i
1	12	18
2	30	42
3	24	24
4	24	36
5	18	24
6	30	36
7	12	14
8	6	10
9	36	48
10	42	54

$$r = \frac{\sum y_i}{\sum x_i} = \frac{306}{234} = \mathbf{1.31}, \hat{\tau}_y = r\tau_x = \frac{306}{234}(4200) = \mathbf{5492.31}$$

$$s_r^2 = \frac{1}{n-1}\sum(y_i - rx_i)^2 = 12.07627$$

```
> x=c(12,30,24,24,18,30,12,6,36,42)
> y=c(18,42,24,36,24,36,14,10,48,54)
> sum(x)
[1] 234
> sum(y)
[1] 306
```

```
> s=c((y-x*306/234)^2)
> sum(s)/9
[1] 12.07627
```

$$\hat{V}(\hat{\tau}_y) = \tau_x^2\left(\frac{N-n}{nN}\right)\left(\frac{s_r^2}{\mu_x^2}\right) = (N\mu_x)^2\left(\frac{N-n}{nN}\right)\left(\frac{s_r^2}{\mu_x^2}\right) = N\left(\frac{N-n}{n}\right)s_r^2$$

$$B = 2\sqrt{\hat{V}(\hat{\tau}_y)} = 2\sqrt{200\left(\frac{200-10}{10}\right)(12.07627)} = \mathbf{428.44}$$

Ans: $\hat{\boldsymbol{\tau}}_y = \mathbf{5492.31}, \boldsymbol{B} = \mathbf{428.44}$

6.10 Members of a teachers' association are concerned about the salary increases given to high school teachers in a particular school system. A simple random sample of n = 15 teachers is selected from an alphabetical listing of all high school teachers in the system. All 15 teachers are interviewed to determine their salaries for this year and the previous year (see the accompanying table). Use these data to estimate R, the rate of change, for N = 750 high school teachers in the community school system. Place a bound on the error of estimation.
(ignore the question "What pattern …")

Teacher	Past year's salary	Present year's salary
1	30400	31500
2	31700	32600
3	32792	33920
4	34956	36400
5	31355	32020
6	30108	31308
7	32891	34100
8	30216	31320
9	30416	31420
10	30397	31600
11	33152	34560
12	31436	32750
13	34192	35800
14	32006	33300
15	32311	33920

```
> x=c(30400,31700,32792,34956,31355,30108,32891,30216,30416,30397,33152,31436,34192,32006,32
311)
> y=c(31500,32600,33920,36400,32020,31308,34100,31320,31420,31600,34560,32750,35800,33300,33
920)
> sum(x)
[1] 478328
> sum(y)
[1] 496518
```

$$\hat{R} = r = \frac{496518}{478328} = \mathbf{1.038};\ s_r^2 = \frac{1}{n-1}\sum(y_i - rx_i)^2 = 51086.04, \bar{x} = \frac{478328}{15} = 31888.53$$

```
> s=c((y-x*496518/478328)^2)
> sum(s)/14
[1] 51086.04
```

$$B = 2\sqrt{\hat{V}(r)} = 2\sqrt{\left(1-\frac{15}{750}\right)\left(\frac{1}{31888.53^2}\right)\frac{51086.04}{15}} = \mathbf{0.003623}$$

Ans: $\hat{\boldsymbol{R}} = \mathbf{1.038}, \boldsymbol{B} = \mathbf{0.003623}$

6.16 Refer to Exercise 6.9. Estimate τ_y by using a regression estimator and place a bound on the error of estimation. Do you think the regression estimator is better than the ratio estimator for this problem?

$$b = \left(\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})\right) \Big/ \left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right) = 1.259372$$

```
> x=c(12,30,24,24,18,30,12,6,36,42)
> y=c(18,42,24,36,24,36,14,10,48,54)
> SSxy=c((y-mean(y))*(x-mean(x)))
> SSxx=c((x-mean(x))^2)
> b=sum(SSxy)/sum(SSxx)
> b
[1] 1.259372
```

$$\hat{\mu}_{yL} = \bar{y} + b(\mu_x - \bar{x}) = 30.6 + 1.259372\left(\frac{4200}{200} - 23.4\right) = 27.5775$$

$$\hat{V}(\hat{\mu}_{yL}) = \frac{N-n}{nN}\text{MSE} = \frac{200-10}{10(200)}13.2 = 1.254$$

```
> SSe=c((y-(mean(y)-b*mean(x)+b)*x)^2)
> SSe/8
 [1]  14.260072 110.277048 139.125396  57.040290  45.226120 159.330240  26.940923
 [8]   2.354912 180.904482 268.922589
> sum(SSe)/8
[1] 1004.382
```

$$\hat{\tau}_{yL} = N\hat{\mu}_{yL} = N[\bar{y} + b(\mu_x - \bar{x})] = 200(27.5775) = \mathbf{5515.5};$$

$$B = 2\sqrt{N^2\hat{V}(\hat{\mu}_{yL})} = 2\sqrt{200^2(1.254)} = \mathbf{448.6}$$

Ans: Regression estimator is not better than the ratio estimator because of small sample size.

6.25 A certain manufacturing firm produces a product that is packaged under two brand names, for marketing purposes. These two brands serve as strata for estimating potential sales volume for the next quarter. A simple random sample of customers for each brand is contacted and asked to provide a potential sales figure y (in number of units) for the coming quarter. Last year's true sales figure, for the same quarter, is available for each of the sampled customers and is denoted by x. The data are given in the accompanying table. The sample for brand I was taken from a list of 120 customers for whom the total sales in the same quarter of last year was 24,500 units. The brand II sample came from 180 customers with a total quarterly sales last year of 21,200 units. Find a ratio estimate of the total potential sales for next quarter. Estimate the variance of your estimator.

Brand I		Brand II	
x_i	y_i	x_i	y_i
204	210	137	150
143	160	189	200
82	75	119	125
256	280	63	60
275	300	103	110
198	190	107	100
		159	180
		63	75
		87	90

```
> x1=c(204,143,82,256,275,198)
> y1=c(210,160,75,280,300,190)
> x2=c(137,189,119,63,103,107,159,63,87)
> y2=c(150,200,125,60,110,100,180,75,90)
> sum(y1)
[1] 1215
> y1SQ=c(y1^2)
> sum(y1SQ)
[1] 279825
> sum(x1)
[1] 1158
> x1SQ=c(x1^2)
> sum(x1SQ)
[1] 249154
```

```
> SSx1y1=sum(c(x1*y1))
> SSx1y1
[1] 263670
> sum(y2)
[1] 1090
> y2SQ=c(y2^2)
> sum(y2SQ)
[1] 149950
> sum(x2)
[1] 1027
> x2SQ=c(x2^2)
> sum(x2SQ)
[1] 131497
> SSx2y2=sum(c(x2*y2))
> SSx2y2
[1] 140210
```

$$\hat{\tau}_{yRC} = N\hat{\mu}_{yRC} = N\frac{\bar{y}_{st}}{\bar{x}_{st}}\mu_x = \frac{\bar{y}_{st}}{\bar{x}_{st}}\tau_x$$

$$\bar{y}_{st} = \frac{1}{N}(N_1\bar{y}_1 + N_2\bar{y}_2) = \frac{1}{300}\left[120\left(\frac{1215}{6}\right) + 180\left(\frac{1090}{9}\right)\right] = \mathbf{153.67}$$

$$\bar{x}_{st} = \frac{1}{N}(N_1\bar{x}_1 + N_2\bar{x}_2) = \frac{1}{300}\left[120\left(\frac{1158}{6}\right) + 180\left(\frac{1027}{9}\right)\right] = \mathbf{145.67}$$

$$r_c = \frac{\bar{y}_{st}}{\bar{x}_{st}} = \frac{153.67}{145.67} = 1.054919; \hat{\tau}_{yRC} = \frac{153.67}{145.67}(24500 + 21200) = \mathbf{48029.84}$$

```
> 153.67/145.67
[1] 1.054919
> SS1=c(((y1-mean(y1)-(x1-mean(x1))*153.67/145.67)^2))
> sum(SS1)
[1] 788.8139
> SS2=c(((y2-mean(y2)-(x2-mean(x2))*153.67/145.67)^2))
> sum(SS2)
[1] 461.7523
```

$$\hat{V}\left(\hat{\tau}_{yRC}\right) = N^2\hat{V}\left(\hat{\mu}_{yRC}\right)$$

$$= 300^2\left[\left(\frac{120}{300}\right)^2\left(\frac{120-6}{120(6)}\right)\left(\frac{1}{5}\right)(788.81) + \left(\frac{180}{300}\right)^2\left(\frac{180-9}{180(9)}\right)\left(\frac{1}{8}\right)(461.75)\right]$$

$$= \mathbf{557095.07}$$

Ans: $\boldsymbol{r_c = 1.054919; \hat{V}\left(\hat{\tau}_{yRC}\right) = 557095.07}$

6.26 For Exercises 6.1 and 6.2, a regression estimator could be employed. Compute the relative efficiency of

a. ratio estimation to simple random sampling.

$$\hat{V}(\hat{\tau}) = N^2\left(\frac{N-n}{N}\right)\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\sum_{i=1}^{n}(y_i^2 - n\bar{y}^2)$$

$$= 250^2\left(\frac{250-12}{250}\right)\left(\frac{1}{12}\right)\left(\frac{1}{11}\right)\left[1976 - 12\left(\frac{142}{12}\right)^2\right] = \mathbf{133273.99}$$

$$\hat{V}(\hat{\tau}_y) = N^2\left(\frac{N-n}{nN}\right)\left(\frac{1}{n-1}\right)\left[\sum y_i^2 - 2r\sum x_i y_i + r^2\sum x_i^2\right]$$

$$= 250^2\left(\frac{250-12}{12(250)}\right)\left(\frac{1}{11}\right)\left[1976 - 2\left(\frac{142}{6.7}\right)(91.2) + \left(\frac{142}{12}\right)^2(4.25)\right]$$

$$= \mathbf{8678.56}$$

$$\text{RE}\left(\frac{\hat{\tau}_y}{\hat{\tau}}\right) = \frac{\hat{V}(\hat{\tau})}{\hat{V}(\hat{\tau}_y)} = \frac{133273.99}{8678.56} = \mathbf{15.36}$$

Ans: RE $\left(\frac{\hat{\tau}_y}{\hat{\tau}}\right) = \mathbf{15.35669}$

b. regression estimation to simple random sampling.

```
> y=c(6,9,7,19,15,5,12,9,20,9,18,13)
> var(y)
[1] 26.87879
> sd(y)
[1] 5.184476
> x=c(.3,.5,.4,.9,.7,.2,.6,.5,.8,.4,.8,.6)
> SSxy=c((y-mean(y))*(x-mean(x)))
> SSxx=c((x-mean(x))^2)
> b=sum(SSxy)/sum(SSxx)
> b
[1] 23.40426
```

$$b = \frac{\sum x_i y_i - n\overline{xy}}{\sum x_i^2 - n\bar{x}^2} = \frac{91.2 - 12\left(\frac{6.7}{12}\right)\left(\frac{142}{12}\right)}{4.25 - 12\left(\frac{6.7}{12}\right)^2} = \mathbf{23.40}$$

$$\hat{V}(\hat{\tau}_{yL}) = N^2\left(\frac{N-n}{nN}\right)\left(\frac{1}{n-2}\right)\left[\left(\sum y_i^2 - n\bar{y}^2\right) - b^2\sum x_i^2 - n\bar{x}^2\right]$$

$$= 250^2\left(\frac{250-12}{12(250)}\right)\left(\frac{1}{10}\right)\left[\left(1976 - \left(\frac{142}{12}\right)^2\right) - 23.40^2\left(4.25 - 12\left(\frac{6.7}{12}\right)^2\right)\right]$$

$$= \mathbf{8313.12}$$

```
> SSe=c((y-(mean(y)-b*mean(x)+b)*x)^2)
> sum(SSe)/10
[1] 2.110611
```

$$\text{RE}\left(\frac{\hat{\tau}_{yL}}{\hat{\tau}}\right) = \frac{\hat{V}(\hat{\tau})}{\hat{V}(\hat{\tau}_{yL})} = \frac{133273.99}{8313.12} = 16.03$$

Ans: RE $\left(\frac{\hat{\tau}_{yL}}{\hat{\tau}}\right) = \mathbf{16.03}$

c. regression estimation to ratio estimation.

$$\mathrm{RE}\left(\frac{\hat{\tau}_{yL}}{\hat{\tau}_y}\right) = \frac{\hat{V}\left(\hat{\tau}_y\right)}{\hat{V}\left(\hat{\tau}_{yL}\right)} = \frac{8678.56}{8313.12} = 1.04$$

Ans: RE $\left(\frac{\hat{\mu}_{yL}}{\hat{\mu}_y}\right) = \mathbf{0.829284}$

Additional problem I: Let $\bar{x}$ and $\bar{y}$ be the sample means, s_x^2, s_y^2, and s_{xy} be the sample variances and covariance, $\hat{b} = s_{xy}/s_x^2$ and $\hat{\rho} = s_{xy}/(s_x s_y)$. Show that

$$\frac{1}{n-1}\sum_{i\in\text{sample}}\{y_i-\bar{y}-\hat{b}(x_i-\bar{x})\}^2 = s_y^2+\hat{b}^2 s_x^2 - 2\hat{b}x_{sy} = s_y^2-\hat{b}^2 s_x^2 = (1-\hat{\rho}^2)s_y^2$$

PF:

$\hat{a} = \bar{y}-\hat{b}\bar{x}$,

$$\sum_{i\in\text{sample}}\{y_i-\bar{y}-\hat{b}(x_i-\bar{x})\}^2 = \sum_{i\in\text{sample}}\{y_i-(\bar{y}-\hat{b}\bar{x})-\hat{b}x_i\}^2 = \sum_{i\in\text{sample}}\{y_i-\hat{a}-\hat{b}x_i\}^2$$

$$= (n-2)s_{e,\text{reg}}^2$$

$$\frac{1}{n-1}\sum_{i\in\text{sample}}\{y_i-\bar{y}-\hat{b}(x_i-\bar{x})\}^2 = \frac{n-2}{n-1}s_{e,\text{reg}}^2 = \left(\frac{n-2}{n-1}\right)\left(\frac{n-1}{n-2}\right)(s_y^2+\hat{b}^2 s_x^2-2\hat{b}x_{sy})$$

$$= s_y^2-\hat{b}^2 s_x^2 = (1-\hat{\rho}^2)s_y^2$$

Sampling from Real Populations

6.1 The data set TEMPS in Appendix C (and on the data disk) shows normal temperature (*T*) and amount of precipitation (*P*) for weather stations around the United States. Using the January and March precipitation data as the population of interest, select a sample of *n* stations to answer the following. Choose an appropriate sample size and find a margin of error for each part.
(use the sample size 20)
(a) Estimate the ratio of the average March precipitation to the average January precipitation.

```
> temps <- read.delim("~/Documents/Rutgers/Spring 2020/Stat 476/Homework/temps.txt")
>   View(temps)
> precip1=temps$J.P
> precip3=temps$M.P
> sam=sample(88,20)
> x1=precips1[sam]
Error: object 'precips1' not found
> x1=precip1[sam]
> y1=precip3[sam]
> r=sum(y1)/sum(x1)
> r
[1] 1.213942
```

Ans: $\boldsymbol{\hat{r} = 1.213942}$

(b) Estimate (with an error bound on) the average March precipitation for all stations, using the sample data on January and March precipitation and the additional knowledge that the average January precipitation for all stations is 2.526136 inches. Compare your answers using ratio regression and difference estimation.

```
> mu1=2.526136
> mean(x1)
[1] 2.08
> mean(y1)
[1] 2.525
> n=20
> N=88
> xbar=mean(x1)
> ybar=mean(y1)
> sx2=var(x1)
> sy2=var(y1)
> sxy=cov(x1,y1)
> yhat=r*mu1
> yhat
[1] 3.066583
> Vhat=(1-n/N)/n*(sy2+r^2*sx2-2*r*sxy)
> Vhat
[1] 0.02954809
> B=2*sqrt(Vhat)
> B
[1] 0.3437911
```

```
> Vhat_ratio=Vhat
> Vhat_ratio
[1] 0.02954809
> bhat=sxy/sx2
> bhat
[1] 0.9657616
> yhat=ybar-bhat*(xbar-mu1)
> yhat
[1] 2.955861
> se2=(sy2+bhat^2*sx2-2*bhat*sxy)*(n-1)/(n-2)
> sqrt(se2)
[1] 0.8186209
> Vhat_reg=(1-n/N)/n*se2
> Vhat_reg
[1] 0.02589178
> Vhat_reg/Vhat_ratio
[1] 0.8762592
```

```
> Vhat_ratio
[1] 0.02954809
> yhat=ybar-(xbar-mu1)
> yhat
[1] 2.971136
> Vhat_dif=(1-n/N)/n*(sy2+sx2-2*sxy)
> Vhat_dif
[1] 0.02462458
> Vhat_dif/Vhat_ratio
[1] 0.8333731
```

Ans: $\boldsymbol{\hat{\mu}_y = 3.066583; B = 0.3437911;}$ **RE** $\boldsymbol{\left(\frac{\hat{\mu}_y}{\hat{\mu}_{yL}}\right) = 0.8762592;}$ **RE** $\boldsymbol{\left(\frac{\hat{\mu}_y}{\hat{\mu}_{yD}}\right) = 0.8333731}$

6.5 The data set SCHOOLS in Appendix C and on the data disk contains information for the 2001–2002 school year on various aspects of education for all 50 states. Select a simple random sample of eight states to answer the following.

(a) Estimate the mean per-pupil expenditure for the United States, with a margin of error.

```
> schools <- read.delim("~/Documents/Rutgers/Spring 2020/Stat 476/Homework/schools.txt")
>   View(schools)
> sam=sample(50,8)
> mu1=mean(schools$Students)
> mu1
[1] 952249.6
> x1=schools$Students[sam]
> y1=schools$Expend[sam]
> r=sum(y1)/sum(x1)
> r
[1] 0.004718303
> n=8
> N=50
> sr2=sum(c(y1-r*x1)^2)/(n-1)
> vhat=(1-n/N)*(sr2/n)/(mean(x1)*mean(x1))
> B=2*sqrt(vhat)
> B
[1] 0.00400049
```

Ans: $\hat{\mu}_{\mathrm{Expend}/Student} = \mathbf{0.004718303}, \boldsymbol{B} = \mathbf{0.00400049}$

(b) Estimate the mean per-capita expenditure for the United States, with a margin of error.

```
> mu2=mean(schools$Pop)
> x2=schools$Pop[sam]
> y2=schools$Expend[sam]
> rhat=sum(y2)/sum(x2)
> rhat
[1] 0.0007769838
> sr2=sum(c(y2-rhat*x2)^2)/(n-1)
> vhat=(1-n/N)*(sr2/n)/(mean(x2)*mean(x2))
> B=2*sqrt(vhat)
> B
[1] 0.0006166054
```

Ans: $\hat{\mu}_{\mathrm{Expend}/Pop} = \mathbf{0.0007769838}, \boldsymbol{B} = \mathbf{0.0006166054}$

(c) Estimate the student–teacher ratio (average number of students per teacher) for the United States, with a margin of error.

```
> mu3=mean(schools$Teachers)
> x3=schools$Teachers[sam]
> y3=schools$Students[sam]
> rhat=sum(y3)/sum(x3)
> rhat
[1] 16.90333
```

```
> sr2=sum(c(y3-rhat*x3)^2)/(n-1)
> vhat=(1-n/N)*(sr2/n)/(mean(x3)*mean(x3))
> B=2*sqrt(vhat)
> B
[1] 3.423083
```

Ans: $\hat{\mu}_{\mathrm{Student}/Teacher} = \mathbf{16.90333}, \boldsymbol{B} = \mathbf{3.423083}$