**Chapter 22**

**Sample Survey**

**Learning Objectives**

1. Learn what a sample survey is and how it differs from an experiment as a method of collecting data.

2. Know about the methods of data collection for a survey.

3. Know the difference between sampling and nonsampling error.

4. Learn about four sample designs: (1) simple random sampling, (2) stratified simple random sampling, (3) cluster sampling, and (4) systematic sampling.

5. Lean how to estimate a population mean, a population total, and a population proportion using the above sample designs.

6. Understand the relationship between sample size and precision.

7. Learn how to choose the appropriate sample size using stratified and simple random sampling.

8. Learn how to allocate the total sample to the various strata using stratified simple random sampling.

**Solutions:**

1. a.  = 215 is an estimate of the population mean.

b. 

c. 215  2(2.7386) or 209.5228 to 220.4772

2. a. Estimate of population total = *N* = 400(75) = 30,000

b. Estimate of Standard Error = 



c. 30,000  2(320) or 29,360 to 30,640

3. a. = .30 is an estimate of the population proportion

b. 

c. .30  2(.0437) or .2126 to .3874

4. *B* = 15



A sample size of 73 will provide an approximate 95% confidence interval of width 30.

5. a. = 149,670 and *s* = 73,420



approximate 95% confidence interval

149,670  2(10,040.83)

or

$129,588.34 to $169,751.66

b. = *N*= 771(149,670) = 115,395,570

= *N*= 771(10,040.83) = 7,741,479.93

approximate 95% confidence interval

115,395,570  2(7,741,479.93)

or

$99,912,610.14 to $130,878,529.86

c. = 18/50 = 0.36 and 

approximate 95% confidence interval

0.36  2(0.0663)

or

0.2274 to 0.4926

This is a rather large interval; sample sizes must be rather large to obtain tight confidence intervals on a population proportion.

6. *B* = 5000/2 = 2500 Use the value of *s* for the previous year in the formula to determine the necessary sample size.



A sample size of 337 will provide an approximate 95% confidence interval of width no larger than $5000.

7. a. Stratum 1: = 138

Stratum 2: = 103

Stratum 3: = 210

b. Stratum 1

= 138



138  2(6.3640)

or

125.272 to 150.728

Stratum 2

= 103



103  2(4.2817)

or

94.4366 to 111.5634

Stratum 3

= 210



210  2(8.6603)

or

192.6794 to 227.3206

c. 

= 50.1818 + 46.8182 + 38.1818

= 135.1818





approximate 95% confidence interval

135.1818  2(3.4092)

or

128.3634 to 142.0002

8. a. Stratum 1:  = 200(138) = 27,600

Stratum 2:  = 250(103) = 25,750

Stratum 3:  = 100(210) = 21,000

b.  = 27,600 + 25,750 + 21,000 = 74,350

Note: the sum of the estimate for each stratum total equals 

c. = 550(3.4092) = 1875.06 (see 7c)

approximate 95% confidence interval

74,350  2(1875.06)

or

70,599.88 to 78,100.12

9. a. Stratum 1

= .50



.50  2(.1088)

or

.2824 to .7176

Stratum 2

= .78



.78  2(.0722)

or

.6356 to .9244

Stratum 3

= .21



.21  2(.0720)

or

.066 to .354

b. 

c. 



d. approximate 95% confidence interval

.5745  2(.0530)

or

.4685 to .6805

10. a. 

Rounding up we choose a total sample of 93.







b. With *B* = 10, the first term in the denominator in the formula for *n* changes.



Rounding up, we see that a sample size of 306 is needed to provide this level of precision.







Due to rounding, the total of the allocations to each strata only add to 305. Note that even though the sample size is larger, the proportion allocated to each stratum has not changed.

c. 

Rounding up, we see that a sample size of 275 will provide the desired level of precision.

The allocations to the strata are in the same proportion as for parts a and b.







Again, due to rounding, the stratum allocations do not add to the total sample size. Another item could be sampled from, say, stratum 3 if desired.

11. a.  = 29.5333  = 64.775

 = 45.2125  = 53.0300

b. Indianapolis



29.533  10.9086(.9177)

or

19.5222 to 39.5438

Louisville



64.775  17.7248(.9068)

or

48.7022 to 80.8478

St. Louis



45.2125  (13.7238) (.9487)

or

32.1927 to 58.2323

Memphis



53.0300  18.7719(.9258)

or

35.6510 to 70.4090

c. 

d. 









approximate 95% confidence interval

.4269  2(.0857)

or

.2555 to .5983

12. a. St. Louis total =  = 80 (45.2125) = 3617

In dollars: $3,617,000

b. Indianapolis total =  = 38 (29.5333) = 1122.2654

In dollars: $1,122,265

c. 













approximate 95% confidence interval



48.7821  2(3.8583)

or

41.0655 to 56.4987

In dollars: $41,066 to $56,499

d. approximate 95% confidence interval



233(48.7821)  2(233)(3.8583)

11,366.229  1797.9678

or

9,568.2612 to 13,164.197

In dollars: $9,568,261 to $13,164,197

13. 

Rounding up we see that a sample size of 28 is necessary to obtain the desired precision.







b. 







This is the same as proportional allocation . Note that for each stratum



14. a. 

 = 300(15) = 4500



b.  = [ 95 – 15 (7) ]2 + [ 325 – 15 (18) ]2 + [ 190 – 15 (15) ]2 + [ 140 – 15 (10)]2

= (–10)2 + (55)2 + (–35)2 + (–10)2

= 4450



= 300(1.4708) = 441.24

 = [ 1 – .3 (7) ]2 + [ 6 – .3 (18) ]2 + [ 6 – .3 (15) ]2 + [2 – .3 (10) ]2

= (–1.1)2 + (.6)2 + (1.5)2 + (–1)2

= 4.82



c. approximate 95% confidence

Interval for Population Mean:

15  2(1.4708)

or

12.0584 to 17.9416

d. approximate 95% confidence

Interval for Population Total:

4500  2(441.24)

or

3617.52 to 5382.48

e. approximate 95% confidence

Interval for Population Proportion:

.30  2(.0484)

or

.2032 to .3968

15. a. 

 = 600(80) = 48,000



b.  = [ 3500 – 80 (35) ]2 + [ 965 – 80 (15) ]2 + [ 960 – 80 (12) ] 2

+ [ 2070 – 80 (23) ] 2 + [ 1100 – 80 (20) ] 2 + [ 1805 – 80 (25) ] 2

= (700) 2 + (–235) 2 + (0) 2 + (230) 2 + (–500) 2 + (–195) 2

= 886,150



approximate 95% confidence

Interval for Population Mean:

80  2(7.6861)

or

64.6278 to 95.3722

c. = 600(7.6861) = 4611.66

approximate 95% confidence

Interval for Population Total:

48,000  2(4611.66)

or

38,776.68 to 57,223.32

d.  = [ 3 – .1 (35) ] 2 + [ 0 – .1 (15) ] 2 + [ 1 – .1 (12) ] 2 + [4 – .1 (23) ] 2

+ [ 3 – .1 (20) ] 2 + [ 2 – .1 (25) ] 2

= (–.5) 2 + (–1.5) 2 + (–.2) 2 + (1.7) 2 + (1) 2 + (–.5) 2

= 6.68



approximate 95% confidence

Interval for Population Proportion:

.10  2(.0211)

or

.0578 to .1422

16. a. 

Estimate of mean age of mechanical engineers: 40 years

b. 

Estimate of proportion attending local university: .70

c.  = [ 520 – 40 (12) ] 2 + · · · + [ 462 – 40 (13) ] 2

= (40) 2 + (–7) 2 + (–10) 2 + (–11) 2 + (30) 2 + (9) 2 + (22) 2 + (8) 2 + (–23) 2

+ (–58) 2

= 7292



approximate 95% confidence

Interval for Mean age:

40  2(2.0683)

or

35.8634 to 44.1366

d.  = [ 8 – .7 (12) ] 2 + · · · + [ 12 – .7 (13) ] 2

= (–.4) 2 + (–.7) 2 + (–.4) 2 + (.3) 2 + (–1.2) 2 + (–.1) 2 + (–1.4) 2 + (.3) 2

+ (.7) 2 + (2.9) 2

= 13.3



approximate 95% confidence

Interval for Proportion Attending Local University:

.70  2(.0883)

or

.5234 to .8766

17. a. 

Estimate of mean age: 36.9737 years

b. Proportion of College Graduates: 128 / 304 = .4211

Proportion of Males: 112 / 304 = .3684

c.  = [ 17 (37) – (36.9737) (17) ] 2 + · · · + [ 57 (44) – (36.9737) (44) ] 2

= (.4471) 2 + (–174.0795) 2 + (–25.3162) 2 + (–460.2642) 2 + (173.1309) 2

+ (180.3156) 2 + (–94.7376) 2 + (400.4991) 2

= 474,650.68



approximate 95% confidence

Interval for Mean Age of Agents:

36.9737  2(2.2394)

or

32.4949 to 41.4525

d.  = [ 3 – .4211 (17) ] 2 + · · · + [ 25 – .4211 (57) ] 2

= (–4.1587) 2 + (–.7385) 2 + (–2.9486) 2 + (10.2074) 2 + (–.1073) 2 + (–3.0532) 2

+ (–.2128) 2 + (.9973) 2

= 141.0989



approximate 95% confidence

Interval for Proportion of Agents that are College Graduates:

.4211  2(.0386)

or

.3439 to .4983

e.  = [ 4 – .3684 (17) ] 2 + · · · + [ 26 – .3684 (57) ] 2

= (–2.2628) 2 + (–.8940) 2 + (–2.5784) 2 + (3.6856) 2 + (–3.8412) 2 + (1.5792) 2

+ (–.6832) 2 + (5.0012) 2

= 68.8787



approximate 95% confidence

Interval for Proportion of Agents that are Male:

.3684  2(.0270)

or

.3144 to .4224

18. a.  = 0.19



Approximate 95% Confidence Interval:

0.19  2(0.0206)

or

0.1488 to 0.2312

b. = 0.31



Approximate 95% Confidence Interval:

0.31  2(0.0243)

or

0.2615 to 0.3585

c. = 0.17



Approximate 95% Confidence Interval:

0.17  2(0.0197)

or

0.1306 to 0.2094

d. The largest standard error is when  = .50.

At  = .50, we get



Multiplying by 2, we get a bound of *B* = 2(.0262) = 0.0525

For a sample of 363, then, they know that in the worst case (= 0.50), the bound will be approximately 5%.

e. If the poll was conducted by calling people at home during the day the sample results would only be representative of adults not working outside the home. It is likely that the Louis Harris organization took precautions against this and other possible sources of bias.

19. a. Assume (*N* – *n*) / *N* ≈ 1

= 0.70



b. = 0.63



c. The estimate of the standard error in part (b) is larger because is closer to .50 in part (b).

d. Approximate 95% Confidence interval:

0.55 ± 2(.0245)

or

0.6510 to 0.7490

e. Approximate 95% Confidence interval:

0.31 ± 2(.0258)

or

0.5784 to 0.6816

20. a. 

Approximate 95% Confidence Interval for Mean Annual Salary:

23,200  2(204.9390)

or

$22,790 to $23,610

b.  = 3000 (23,200) = 69,600,000

 = 3000 (204.9390) = 614,817

Approximate 95% Confidence Interval for Population Total Salary:

69,600,000  2(614,817)

or

$68,370,366 to $70,829,634

c. = .73



Approximate 95% Confidence Interval for Proportion that are Generally Satisfied:

.73  2(.0304)

or

.6692 to .7908

d. If management administered the questionnaire and anonymity was not guaranteed we would expect a definite upward bias in the percent reporting they were “generally satisfied” with their job. A procedure for guaranteeing anonymity should reduce the bias.

21. a. = 1/3



Approximate 95% Confidence Interval:

.3333  2(.0840)

or

.1653 to .5013

b.  = 760 (19 / 45) = 320.8889

c. = 19 / 45 = .4222



Approximate 95% Confidence Interval:

.4222  2(.0722)

or

.2778 to .5666

d. 





= 1019.1571 + 3012.7901 + 513.2400 = 4545.1892



Approximate 95% Confidence Interval:

.3717  2(.0482)

or

.2753 to .4681

22. a.  = 380 (9 / 30) + 760 (12 / 45) + 260 (11 / 25) = 431.0667

Estimate approximately 431 deaths due to beating.

b. 



= (380) (380 – 30) (9 / 30) (21 / 30) / 29 + (760) (760 – 45) (12 / 45) (33 / 45) / 44 +

(260) (260 – 25)(11 / 25) (14 / 25) / 24

= 4005.5079



Approximate 95% Confidence Interval:

.3079  2(.0452)

or

.2175 to .3983

c. 



= (380) (380 – 30) (21 / 30) (9 / 30) / 29 + (760) (760 – 45) (34 / 45) (11 / 45) / 44 +

(260) (260 – 25) (15 / 25) (10 / 25) / 24

= 3855.0417



Approximate 95% Confidence Interval:

.7116  2(.0443)

or

.6230 to .8002

d. = 1400 (.7116) = 996.24

Estimate of total number of victims is 996.

23. a. 



Rounding up, we need a sample size of 171 for the desired precision.

b. 









24. a. 

Estimate of mean age is approximately 75 years old.

b. 

 = [12 – .35 (14) ] 2 + [ 2 – .35 (7) ] 2 + [30 – .35 (96) ]2

+ [ 8 – .35 (23) ] 2 + [ 10 – .35 (71) ] 2 + [ 22 – .35 (29) ] 2

= (7.1) 2 + (–.45) 2 + (–3.6) 2 + (–.05) 2 + (–14.85) 2 + (11.85) 2

= 424.52



Approximate 95% Confidence Interval:

.35  2(.0760)

or

.198 to .502

= 4800 (.35) = 1680

Estimate of total number of Disabled Persons is 1680.