Tutorial 4 (chapter 3 part 2)

Simple random sampling

Exercise:

If a simple random sample without replacement (WOR) of 10 universities are selected from a population of 50 universities in a particular country. The numbers of statistics professors in the sample university are: 23, 14, 38, 11, 7, 31, 9, 18, 12, 25. Solve by R:

1. Estimate the mean number of statistics professors in this population, $\bar{y}$.
2. Estimate the variance of your estimator, $\{v(\bar{y})\}$.
3. Estimate the total number of statistics professors in this population, y'.
4. Estimate the variance of your estimator, $\{v(y')\}$.
5. How many possible samples can be drawn.
6. Give an approximate 95% confidence interval for the population mean.
7. Give an approximate 95% confidence interval for the population total.

By R:

```
> #Exercise 2:
> #The population data of of 10 universities.
> y=c(23, 14, 38, 11, 7, 31, 9, 18, 12, 25)
> #1.sample mean
> ybar= mean(y)
> ybar
[1] 18.8
>
> #Population size and sample size
> N=50
> n=10
>
> # to obtain sample variance s^2.
> #The denominator N - 1 is used in R.
> ssquare=var(y)
> ssquare
[1] 104.4
>
> #2.variance of ybar
> vybar=((N-n)/ (N*n))*(ssquare)
> vybar
[1] 8.352
>
```

```
> #3.sample total
> ytotal= N*ybar
> ytotal
[1] 940
>
> #4.variance of ytotal
> vytotal=(N^2)* vybar
> vytotal
[1] 20880
>
> #5. all samples
> allsamples=choose(N,n)
> allsamples
[1] 10272278170
>
>
> #6.to obtain CL for sample mean
> CIL= ybar- (1.96*(sqrt(vybar)))
> CIL
[1] 13.13563
> CIU= ybar+ (1.96*(sqrt(vybar)))
> CIU
[1] 24.46437
>
> #7.to obtain CL for sample total
> CIL= ytotal- (1.96*(sqrt(vytotal)))
> CIL
[1] 656.7817
> CIU= ytotal+ (1.96*(sqrt(vytotal)))
> CIU
[1] 1223.218
```

Exercise 3.10:

Assume that the sample of 32 women drawn from the population of 680 deceased women in exercise 3.7 is a preliminary sample. Examine, whether this sample size is sufficient to estimate the average age with a margin of error of 5 years? If not, how many more deceased women need to be selected in the sample? (WR)

Solution:

$$N = 680, \qquad n = 32, \qquad B = 5$$

$$n = \frac{s^2}{D}, \qquad D = \frac{B^2}{4} = 6.25$$

$s^2 = 154.0312$ from exercise 3.7

Therefore ,

$$n = \frac{s^2}{D} = \frac{154.0312}{6.25} = 24.645 \approx 25$$

It means that the size of preliminary sample of 32 is sufficient for estimating the population mean with desired precision.

```
> y=c(49, 66, 56, 47, 71, 58, 67,74,66,47, 59,60,51,67
+     ,61,72,63,69,46,33,57,67,28,69,70,44,54,32,50,68,74,60)
> y
 [1] 49 66 56 47 71 58 67 74 66 47 59 60 51 67 61 72 63 69 46 33 57 67 28 69 70 44
[27] 54 32 50 68 74 60
> ybar= mean(y)
> ybar
[1] 57.96875
> ssquare=var(y)
> ssquare
[1] 154.0312
> ssquare/6.25
[1] 24.645
```

Exercise 3.11:

A car dealer is feeling concerned over the complaints received in the office of the manufacturer regarding the free service provided by him to the newly purchased cars. To assess the seriousness of the problem, the dealer decided to draw a WR random sample of 70 buyers out of the total of 1400 individuals who had purchased cars through him during the last one year. Twenty-one buyers included in the sample graded service provided by him as unsatisfactory. Estimate the percentage of buyers feeling unsatisfied with the service provided, and construct a suitable level confidence interval for it.

Solution:

The giving information are $N = 1400, n = 70, a = 21$

$p = \frac{21}{70} = 0.3$, $q = 1 - 0.3 = 0.7$

Unbiased estimator WR : $v(p) = \frac{pq}{n-1} = \frac{0.3*0.7}{70-1} = 0.0030$

$$se(p) = \sqrt{0.0030} = 0.0552$$

The confidence limit for p is $p \pm 1.96 * se(p)$

$$0.3 \pm 1.96 * 0.0552$$

The C.I is $[0.191808, 0.408192]$

Note: rounding to the fourth decimal place

Exercise 3.12:

An investigator wishes to estimate the proportion of students in a university whose fathers are graduates. To arrive at the estimate, a WOR simple random sample of 67 students was drawn from a total of 1400 students. On contacting the sampled students, it was found that the fathers of 46 students had not graduated. Estimate the proportion of students whose fathers were at least graduates. Also, set the confidence interval for population proportion.

Solution:

The giving information are $N = 1400, n = 67, a = 67 - 46 = 21$
(Because the 46 is the number of fathers were not graduates and the question asks about graduates)
$p = \frac{21}{67} = 0.3134,\ \ q = 1 - 0.3134 = 0.6866$
Unbiased estimator WOR : $v(p) = \frac{(N-n)}{N}\frac{pq}{n-1} = 0.0031$

$$se(p) = \sqrt{0.0031} = 0.0557$$

The confidence interval for p is $p \pm 1.96 * se(p)$

$$0.3134 \pm 1.96 * 0.0557$$

The C.I is $[\ 0.2042, 0.4226]$
Note: rounding to the fourth decimal place

If the question additionally asks about the total number of students whose fathers were at least gradates with 95% C.I

Then the answer is
$A' = N * p = 1400 * 0.3134 = 438.76 \approx 439$

The confidence limit for y' is $N * p \pm \big(1.96 * se(p)\big) * N$
[285.9192, 591.6008] ≈[286, 592]

Exercise 3.13:

Assume the WOR sample of 67 students in exercise 3.12 as the preliminary sample. If the permissible error could be taken as .1, determine how many additional students will have to be selected to estimate the proportion in question with specified precision?

Solution:

$$N = 1400, \quad n = 67, \quad B = 0.1$$

$$n = \frac{Npq}{(N-1)D + pq}, \quad D = \frac{B^2}{4} = 0.0025$$

$p = 0.3134,\ q = 0.6866$ from exercise 3.12

Therefore,

$$n = \frac{Npq}{(N-1)D + pq} = \frac{154.0312}{6.25} = 81.14155 \approx 81$$

It means that already selected sample of size of 67 in (3.12) is not sufficient for estimating the population proportion with desired precision. Therefore, $81 - 67 = 14$ more students need to be selected.

Distribution by R:

```
> qnorm(p=1-0.05/2)
[1] 1.959964
> qt(p=1-0.05/2,df=9)
[1] 2.262157
>
```