STAT 522

HW4

Satoshi Ido

FEB 14 2024

1. Which of the following SRS designs will give the most precision for estimating a population mean? Show your results. (8 points)

Assume each population has the same value of the population variance S^2.

a. An SRS of size 400 from a population of size 4,000
b. An SRS of size 30 from a population of size 300
c. An SRS of size 3,000 from a population of size 300,000,000

Use the variance of a sample mean estimate:

$$V(\bar{y}) = \frac{S^2}{n}\left(1 - \frac{n}{N}\right)$$

a. $V(\bar{y}_a) = \frac{S^2}{400}\left(1 - \frac{400}{4000}\right) = \frac{9S^2}{4000} = 0.00225S^2$

b. $V(\bar{y}_b) = \frac{S^2}{30}\left(1 - \frac{30}{300}\right) = \frac{9S^2}{300} = 0.03S^2$

c. $V(\bar{y}_c) = \frac{S^2}{3000}\left(1 - \frac{3,000}{300,000,000}\right) = \frac{99999S^2}{300000000} = 0.00033S^2$

∴ Design c will give the most precision. //

2. The 2010 U.S. census listed the following populations for these Arizona cities:

City	Population
Casa Grande	48,571
Gila Bend	1,922
Jerome	444
Phoenix	1,445,632
Tempe	161,719

Suppose that you are interested in estimating the percentage of persons who have been immunized against measles in each city and can take an SRS of persons.

What should your sample size be in each of the 5 cities if you want the estimate form each city to have a margin of error of 4 percentage points? (10 points)

$$|\bar{y}_{str} - \bar{y}_u| \leq 0.04 = e$$

For which cities does the finite population correction make a difference? (2 points)

Hints:

Assume the maximum variance for a proportion, with p = 0.5.

Use z = 1.96 for a 95% CI.

Use the following equations:

$$n_0 = \left(\frac{z_{\alpha/2}S}{e}\right)^2$$

$$n_{SRS} = \frac{n_0}{1+\frac{n_0}{N}} \quad \leftarrow \text{SRS with non-negligible fpc}$$

1. Sample size

Since they are large populations, we assume the population variance (S^2) $\cong p(1-p)$. And we assume $p = 0.5$ to achieve its maximal value. We set $e = 0.04$ to keep M.O.E of 4.

$$\therefore n_0 = \left(\frac{z_{\alpha/2}S}{e}\right)^2 = \frac{(1.96)^2 \cdot (0.5)(0.5)}{(0.04)^2} = 600.25$$

The sample size ignoring the fpc is large compared with the population size, so we make the fpc adjustments for each stratum

Casa Grande : $n = \frac{600.25}{1 + \frac{600.25}{48571}} \cong 593$

Gila Bend : $n = \frac{600.25}{1 + \frac{600.25}{1921}} \cong 458$

Jerome : $n = \frac{600.25}{1 + \frac{600.25}{449}} \cong 256$

Phoenix : $n = \frac{600.25}{1 + \frac{600.25}{1445632}} \cong 600$

Tempe : $n = \frac{600.25}{1 + \frac{600.25}{161719}} \cong 598$ //

FPC

Casa Granda = $\left(1 - \frac{593}{600}\right) \cong 0.02$

Gila Bend = $\left(1 - \frac{458}{600}\right) \cong 0.23\overline{6}$

Jerome $\cong 0.57\overline{33}$

Phoenix $\cong 0$

Tempe $\cong 0.00\overline{33}$

FPCs above show us that Jerome city has the biggest correction difference ($0.57\overline{33}$) //