Stat 473 - HW 3

Jay Maxwell

2/17/2022

1. Suppose we want to estimate the total number of small farms in 1992.

a) Using your output from (3) in the In-Lab R coding assignment, report the estimated total and its standard error.

The estimated total number small farms in 1992 from our lab assignment was 171707 farms and the SE of that estimate is 28036.57 farms

b) What is the FPC?

The FPC is $(1 - \frac{125}{3041}) = 0.0411$**.**

c) What is the sampling weight associated with each selected county?

The sampling weight is $\frac{3041}{125} = 24.23$

d) Provide an interpretation of this weight.

The sampling weight is the measure of the number of units in the population represented by each sampled unit. So each farm in our sample represent 24.23 farms of our population.

e) Using hand calculations and the result of the sample variance from (3d) in the In-Lab R coding assignment, show how to obtain the standard error you reported in (a) above.

$$\begin{aligned} SE(\hat{t}) &= N\sqrt{\frac{\hat{s}^2}{n}(1 - \frac{n}{N})} \\ &= 3041\sqrt{\frac{11080.412}{125}(1 - \frac{125}{3041})} \\ &= 28036.57 \end{aligned}$$

f) Now, using the results of the total estimator in (3) in the In-Lab again, calculate by hand an estimate for the average number of small farms in 1992 per county as well as its standard error.

$$\begin{aligned} \hat{t} &= N(\bar{y}) \\ 171707 &= 3041(\bar{y}) \\ 56.46 &= \bar{y} \end{aligned}$$

$$\begin{aligned} SE(\bar{y}) &= \frac{SE(\hat{t})}{N} \\ &= \frac{28036.57}{3041} \\ &= 9.22 \end{aligned}$$

2. Now suppose we want to estimate the proportion of counties that have at least 500 farms.

a) Using the results of (1) in the In-Lab R coding assignment, report the proportion and its standard error.

Based on the results from my lab, I found the sample proportion of counties that have at least 500 farms to be 0.52 and the standard error to be approximately 0.0447.

b) Using the reported proportion in (a), calculate by hand the estimated standard error for the proportion and check this result with the standard error reported in (a).

$$\begin{aligned} SE(\hat{p}) &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}(1-\frac{n}{N})} \\ &= \sqrt{\frac{0.52(0.48)}{124}(1-\frac{125}{3041})} \\ &= 0.04393362 \end{aligned}$$

The resulting standard error here of 0.04393362 is very very close, but a little less, than the standard error reported above. Depending on how we choose to round, the difference might not be signficant. This indicates to me that the FPC used in the calculation does not have a significant impact on the findings.

c) Using your result in (a), calculate by hand the estimate for the population total (i.e. the total number of counties that have at least 500 farms), and its standard error. And check them with results obtained by R. Show the intermediate steps.

$$\begin{aligned} \hat{t} &= N * \hat{P} \\ &= 3041 * 0.52 \\ &= 1581.32 \text{ farms} \end{aligned}$$

$$\begin{aligned} SE(\hat{t}) &= N * \hat{p} \\ &= 3041 * 0.04393 \\ &= 133.59 \text{ farms} \end{aligned}$$

```
full_data <- read_csv("agpop-1.csv", show_col_types = FALSE)
names(full_data) <- c("county", "state", "acres92", "acres87", "acres82", "farms92",
    "farms87", "farms82", "largef92", "largef87", "largef82", "smallf92", "smallf87",
    "smallf82", "region")
full_data %>%
    select(county, state, acres92, farms92, largef92, smallf92, region) %>%
    mutate(atLeast500f = farms92 >= 500) -> full_data
cat("number of counties with more than 500 farms:", sum(full_data$atLeast500f))
```

```
## number of counties with more than 500 farms: 1616
```