Homework 7

Nathan Bick

1. Exercises 10.C:1

1. From the $N = 3030$ housing units, the following sample of $n = 20$ housing units is selected.

$\{U_{942}, U_{575}, U_{2503}, U_{2847}, U_{1427}, U_{578}, U_{2145}, U_{1141}, U_{1319}, U_{1079}, U_{370}, U_{545}, U_{2944}, U_{1441}, U_{721}, U_{1216}, U_{403}, U_{17}, U_{411}, U_{772}\}$.

(a) Use the sample data for $y_i = HINCP_i$ to provide a 95% CI for $\bar{Y}_U(HINCP)$.

Firstly, we can take from Note 10.2 that $z_{\alpha/2} = 1.96$. We will proceed with the assumption that we can use the normality approximation for the sampling distribution.

We will use the formula $\left(\bar{y} - z_{\alpha/2}\sqrt{\hat{V}(\bar{y})}, \bar{y} + z_{\alpha/2}\sqrt{\hat{V}(\bar{y})}\right)$.

We get $\bar{y} = 134,581.67$. Note that this includes two units with NA HINCP. For this $\bar{y}$ calculation, I removed those units from n and y_i.

The Standard Error $SE = 39,784.93$.

So the CI is $(56,603, 212,560)$, using the above-calculated statistics.

(b) Use the same sample data to construct a 95% C.I. for t_{HOUSP}.

we follow the same procedure as before. Note that t_{HOUSP} is the total for HOUSP, which we can estimate by $N\bar{y}_{HOUSP}$; $N = 3030$.

we find $\bar{y} = \cancel{887.69}\ 848$, so $\hat{t} = 2,569,440$.

we also know $\sqrt{\hat{V}(\hat{t})} = N\sqrt{\hat{V}(\bar{y})}$, so we have $N\sqrt{\hat{V}(\hat{y})} = 3030 \cdot 198 = 601,440$.

so using these calculated statistics, we get the C.I. $(1,790,617;\ 3,748,262)$.

2. Problems 10.C: 1

1. Verify that the confidence coefficient associated with the procedure $[\hat{t}-5s, \hat{t}+5s]$ in Note 10.5 is $60/70 \approx 85.7\%$. What is the confidence coefficient if the procedure is changed to $[\hat{t}-3s, \hat{t}+3s]$?

I manually copied the data in the pdf into a spreadsheet, at which point I calculated the values of $y-ns$, $y+ns$ for $n \in 3,5$.

That procedure can be found in linked google sheet.

This does confirm the value given for $n=5$, and for $n=3$ the coefficient drops to $\approx 70\%$.

Exercises 11.B:1

1. (i) Repeat Example 11.1 with prob 0.90.

Example 11.1: what size n SRS needed from N=5940 persons in Columbia Community to have $\bar{y}$ for AGEP within 3 years of $\bar{y}_U$ with probability 0.95.

Note that $e=3$, $1-\alpha=0.90$, $z_{\alpha/2}=1.645$

$$\text{so } n = \frac{N\sigma^2}{(N-1)\frac{e^2}{z_{\alpha/2}^2}+\sigma^2} = \frac{5940(391)}{(5939)\frac{(3)^2}{(1.645)^2}+391}$$

$$= 115.299 \approx 116$$

This is a lower number than with 0.95.

(ii) Repeat Example 11.2 with probability 0.90

Example 11.2: what size n SRS need from $N = 5940$ persons in Columbia Community to have estimate $\hat{p}$ for $y = AGEP < 18$ yrs within 0.03 of $p_{AGEP \leq 18}$ with probability 0.95.

Note that $e = 0.03$, $1 - \alpha = 0.90$, $z_{\alpha/2} = 1.645$.

so:

$$n = \frac{N p_{AGEP<18} (1 - p_{AGEP<18})}{(N-1)\frac{e^2}{z_{\alpha/2}^2} + p_{AGEP<18}(1 - p_{AGEP<18})}$$

$$= \frac{5940(0.25)(0.75)}{5939 \frac{(0.03)^2}{(1.645)^2} + (0.25)(0.75)}$$

$$= 514.967 \approx 515.$$

This is a lower number than the n required with 0.95.