STAT 521: Homework 3

Due on February 29, 2024

Problem 1: (30 pt) Assume that a simple random sample of size n is selected from a population of size N and (x_i, y_i) are observed in the sample. In addition, we assume that the population mean of x, denoted by $\bar{X}$, is known.

1. Use a Taylor linearization method to find the variance of the product estimator $\bar{x}\bar{y}/\bar{X}$, where $(\bar{x}, \bar{y})$ is the sample mean of (x_i, y_i).

2. Find the condition that this product estimator has a smaller variance than the sample mean $\bar{y}$.

3. Prove that if the population covariance of x and y is zero, then the product estimator is less efficient than $\bar{y}$.

Problem 2: (20 pt) In a population of 10,000 businesses, we want to estimate the average sales $\bar{Y}$. For that, we sample $n = 100$ businesses using simple random sampling. Furthermore, we have at our disposal the auxiliary information "number of employees", denoted by x, for each business. It is known that $\bar{X} = 50$ in the population. From the sample, we computed the following statistics:

- $\bar{y}_n = 5.2 \times 10^6$ \$ (average sales in the sample)
- $\bar{x}_n = 45$ employees (sample mean)
- $s_y^2 = 25 \times 10^{10}$ (sample variance of y_k)
- $s_x^2 = 15$ (sample variance of x_k)
- $r = 0.8$ (sample correlation coefficient between x and y)

Answer the following questions.

1. Compute a 95% confidence interval for $\bar{Y}$ using the ratio estimator.
2. Compute a 95% confidence interval for $\bar{Y}$ using the regression estimator based on the simple linear regression of y on x (with intercept).

Problem 3: (10 pt)

Under the setup of the Week 6 (Part 1) lecture, prove the last two equalities in page 18. That is, show that

$$
\begin{aligned}
Cov\left(\frac{1}{N_1}\sum_{i=1}^{N} T_i e_i(1), \frac{1}{N_0}\sum_{i=1}^{N}(1-T_i)'_i \mathbf{B}_0 \mid \mathcal{F}_N\right) &= 0 \\
Cov\left(\frac{1}{N_0}\sum_{i=1}^{N}(1-T_i) e_i(0), \frac{1}{N_0}\sum_{i=1}^{N}(1-T_i)'_i \mathbf{B}_0 \mid \mathcal{F}_N\right) &= 0
\end{aligned}
$$

Problem 4: (20 pt)

Under the setup of the Week 6 (Part 2) lecture,

1. Prove Lemma 3.

2. Show that the final weight in (13) satisfies a hard calibration for $\mathbf{x}_1$ in the sense that

$$\sum_{i \in A} \hat{\omega}_i \mathbf{x}_{1i} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_{1i}.$$