LECTURE 10: RATIO ESTIMATION

7.1 Ratio Estimation under SRS

Suppose we draw an SRS of size n from a population of size N, and we would like to estimate the ratio of two population totals X and Y,

$$R = X/Y = \bar{X}/\bar{Y}\,.$$

R is called population ratio.

For example, in a community of senior citizens, we might want to estimate the ratio of total pharmaceutical to total medical expenses per month (or per year).

An obvious estimator of R (called ratio estimator) is

$$r = x'/y' = \bar{x}/\bar{y}\,.$$

The ratio estimator r is usually a biased estimator of R since $\mathbb{E}(r) = \mathbb{E}(\bar{x}/\bar{y}) \neq \mathbb{E}(\bar{x})/\mathbb{E}(\bar{Y}) = \bar{X}/\bar{Y} = R$ for all $\bar{x}$ and $\bar{y}$. But the bias may be considered negligibly small.

It can be shown that

$$SE(r) \approx \left(R/\sqrt{n}\right)\left(V_X^2 + V_Y^2 - 2\rho_{XY}V_XV_Y\right)^{1/2}\sqrt{\frac{N-n}{N-1}}$$

where the coefficients of variation are

$$V_X = \sigma_X/\bar{X},\ V_Y = \sigma_Y/\bar{Y},$$

and the correlation coefficient

$$\rho_{XY} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{N\,\sigma_X\sigma_Y}\,.$$

Definition A $100(1-\alpha)\%$ confidence interval for a population ratio R is

$$r \,\pm\, z_{1-\alpha/2}\,\widehat{SE}(r)$$

where

$$\widehat{SE}(r) = \left(r/\sqrt{n}\right)\left(\hat{V}_X^2 + \hat{V}_Y^2 - 2\hat{\rho}_{XY}\hat{V}_X\,\hat{V}_Y\right)^{1/2}\sqrt{\frac{N-n}{N-1}}\,,$$

$$\hat{V}_X^2 = \left(\frac{N-1}{N}\right)\left(\frac{s_x^2}{\bar{x}^2}\right), \quad \hat{V}_Y^2 = \left(\frac{N-1}{N}\right)\left(\frac{s_y^2}{\bar{y}^2}\right),$$

and

$$\hat{\rho}_{XY} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x\, s_y}.$$

EXAMPLE (page 194) We are given that $r = 0.3191$, $\hat{\rho}_{XY} = 0.99$, $s_x = 54826.6$, $s_y = 141033.6$, $\bar{x} = 128571.4$, $\bar{y} = 402857.1$, $\hat{V}_X^2 = 0.1591$, $\hat{V}_Y^2 = 0.1072$. We obtain $\widehat{SE}(r) = 0.0040$, hence a 95% CI for R is

$$0.3191 \pm (1.96)(0.0040) = [0.31126, 0.32694].$$