STAT 521: Homework Assignment 1

Due on Feb. 7th, 2008

Problem 1: Identifying sample survey terms:

A survey is conducted to find the average weight of cows in a region. A list of all farms is available for the region, and 50 farms are selected at random. Then the weight of each cow at the 50 selected farms is recorded.

(a) What is the target population?

(b) What is the element?

(c) What is the sampling unit?

(d) What is the frame?

(e) List at least two possible sources of nonsampling errors.

Problem 2: (Result 2.3 in the lecture note)

For a fixed sample size design (i.e. n is a fixed number), prove that the variance of the π estimator can alternatively be written as

$$V(\widehat{t}_\pi) = -\frac{1}{2}\sum_{k\in U}\sum_{l\in U}\Delta_{kl}(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l})^2,$$

and its unbiased estimator is

$$\widehat{V}(\widehat{t}_\pi) = -\frac{1}{2}\sum_{k\in S}\sum_{l\in S}\frac{\Delta_{kl}}{\pi_{kl}}(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l})^2,$$

provided that all $\pi_{kl} > 0$.

Problem 3:

According to Result 2.2 and Result 2.3 in the lecture note, for a fixed sample size design, there exist two unbiased estimators for the variance of the π estimator, i.e.

Result 2.2:

$$\widehat{V}(\widehat{t}_\pi) = \sum_{k\in S}\sum_{l\in S}\frac{\Delta_{kl}}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_k}{\pi_k},$$

Results 2.3:

$$\widehat{V}(\widehat{t}_\pi) = -\frac{1}{2}\sum_{k\in S}\sum_{l\in S}\frac{\Delta_{kl}}{\pi_{kl}}(\frac{y_k}{\pi_k} - \frac{y_k}{\pi_k})^2.$$

Although both of them are unbiased estimators for the same quantity, it is not necessarily true that these two estimators are the same in general. Prove that under SI (simple random design without replacement) of size n out of N, these two estimators are identical and they are given by

$$\widehat{V}(\widehat{t}_\pi) = N^2(1 - \frac{n}{N})\frac{S_n^2}{n},$$

where S_n^2 is the sample variance defined as $S_n^2 = \frac{1}{n-1}\sum_{k\in S}(y_k - \bar{y}_n)^2$ and $\bar{y}_n$ is the sample mean.