The University of Chicago
Department of Statistics
Master's Seminar

LI MA
Department of Statistics
The University of Chicago

Correcting the Bias Arising from Cluster Sampling Using Horvitz-Thompson Estimation in a Problem of Estimating the Number of Fish Eggs in a Southern Alaskan Region

Monday, July 31, 2006 at 10:00 AM
110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

The amount of spawning by a fish population is of great interest to environmentalists and the fishing industry in that it reflects both the size and the age composition of that population, and thus provides clues for designing sustainable fishing strategies. For this reason governmental and industrial efforts have been made to measure fish spawning in several regions for various populations. Because fish spawn in huge amounts over extremely vast areas, it is unfeasible to directly count all the eggs laid by a particular population. Hence, attempts to count fish eggs often proceed in two steps—first estimating the average density of eggs in the interested spawning region and then measuring the size of that region; the product of the two gives an estimate of the total amount of spawning. A simple estimator of the average egg density is the average of the egg densities at a number of randomly selected points in the interested region. How to effectively sample these points and how to measure the exact area of the spawning region become the key issues. An economic way to randomly select points in a sea area is through a cluster sampling procedure—first drawing points along the coastline under a (linear) uniform distribution and then selecting points along the transect extending from each of these chosen coastal points into the interested region (in the sea). Due to the geometric properties of the coastline, however, simply averaging the egg densities at the points selected under such a sampling scheme will typically produce a biased estimate of the true average egg density. We introduce a method that corrects this bias using the idea of Horvitz-Thompson estimation and that at the same time bypasses the difficulty of measuring the exact area of the spawning region. An unbiased estimator of the average egg density is introduced that uses weights to account for the coastline geometry and the peculiarity of the cluster sampling procedure. Also addressed is a way to estimate the variance of this new estimator as well as the computational techniques necessary to implement our method. When applied to the data collected in 2004 for herring spawning in a Southern Alaskan region called Sitka Sound, our method corrects significant overestimation in the average egg density and the total number of eggs while leading to no increased variability in the estimation.

Information about building access for persons with disabilities may be obtained in advance by calling the department office at (773) 702-8333.