

Population size estimation based on erroneous capture-recapture data

Li-Chun Zhang

University of Southampton, Southampton, UK; Statistics Norway, Oslo, Norway

E-mail: L.Zhang@soton.ac.uk

Abstract

Log-linear models have long been used for population size estimation based on capture-recapture data subjected to under enumeration. We extend the scope to allow for additional erroneous enumeration, by introducing a concept of pseudo conditional independence and developing new classes of log-linear and -odds models accordingly. For selection among models that achieve the same fitted log-likelihood, which can occur for incomplete categorical data, we explore the use of latent discrepancy measures. Potential applications include replacing traditional costly Census with alternative register-based approaches, sizing of dynamic clandestine population using relevant but erroneous lists, evaluation of diagnostic efficiency of multiple tests, *etc.*

Key Words: Coverage of register, pseudo conditional independence, log-linear and -odds models, maximum likelihood estimation, latent log-likelihood discrepancy.