Matematisk statistik
Stockholms universitet
Kandidatuppsats **2023:15**
http://www.math.su.se/matstat

The problem of trend detection in log–linear regression models when switching from single specimen to pooled sample design

Sebastian Nordlund*

June 2023

Abstract

Concentrations of contaminants in animal populations are usually assumed to be log–normally distributed. Contaminant concentrations can either be measured in tissue taken from individuals or by measuring the concentration in a mixture of tissue taken from several different individuals. When individual concentrations are measured, they can be transformed into log–concentrations and their sample mean is often used as a yearly input in linear trend models. When mixtures of tissue taken from different individuals is introduced as the reported concentration in an ongoing observational study where earlier concentrations were reported individually, a bias will propagate and give rise to biased estimations for the coefficient parameters of log–linear trend models (for as long as both types of input are used). This thesis derives a second–order Taylor series approximation for the distribution of the logarithm of such log–normal averages, and uses this to approximate the size of the biased trend estimator as a function of two sources of variance in log–concentration. When a trend estimator is biased, the probability of Type I error will not correspond to the nominal significance levels chosen for analysis. In order to approximate the true probability of Type I error, a parametric bootstrap simulation study is conducted, and a logistic regression model is then fitted to the outcomes of the simulations. A case study on mercury concentration in Baltic herring is also conducted, as a means to illustrate the problem in a practical setting and exemplify what the logistic regression model would predict regarding increased probability of Type I error and how the biased trend estimates differ from trend estimates of bias–corrected data.

---

[*] Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige. E-post: davidsebastiannordlund@gmail.com. Supervisor: Martin Sköld.