



Bandwidth selection in kernel estimation of relative risk

Ya-Mei Chang*

Tamkang University, New Taipei City, Taiwan – yamei628@gmail.com

The motivation of this paper is arising from analyzing tree data in Pinjar study area, Western Australia. We are interested in estimating the tree death risk which is similar to estimating disease risk in epidemiology. Many literatures have used kernel estimations to estimate disease risk and discussed different criteria to select bandwidth parameters in kernel estimations. Some literatures have recommended constraining the bandwidth parameters to be equal to reduce bias. The equal-bandwidth constraint is reasonable when the underlying densities of case and control data are nearly equal. However, our tree data is on the contrary. The distribution of living trees is dense and uniform, but the distribution of dead trees is sparse and non-uniform. Whether to constrain equal bandwidths becomes a concern in estimating the tree death risk. In this work, we compare some existing cross-validation criteria of bandwidth selection and investigate the influence of the equal-bandwidth constraint through simulations. The sizes of data sets in the simulations are similar to our tree data.

Keywords: bandwidth selection; kernel estimation; relative risk; cross-validation; point pattern process; large data set.