Permutation tests under a rotating sampling plan with clustered data

Jiahua Chen

University of British Columbia

The distribution of lumber strength of any grade may evolve, for example, due to climate change, forest fire, changes in processing methods, and other factors. So, in North America the forest products industry monitors the evolution of their means, percentiles, or other parameters to ensure the wood products meet the industrial standard. For administrative convenience and informativeness, one may adopt a rotating sampling plan by sampling 36 mills in the initial occasion and having six of them replaced in each successive occasion for the next five occasions. The strength data on a specified number, commonly 10 pieces of lumbers from each sampled mill, are obtained. Under such rotating plans the observations on pieces from the same mill are correlated, and the observations on samples from the same mill taken on different occasions are also correlated. Ignoring these correlations may lead to invalid inference procedures. Yet accommodating a cluster structure in parametric models is difficult and entails a high level of misspecification risk. In this paper we explore symmetry in the clustered data collected via a rotating sampling plan to develop a permutation scheme for testing various hypotheses of interest. We also introduce a semiparametric density ratio model to link the distributions of the response variable over time. The combination retains the validity of the inference methods while extracting maximum information from the sampling plan. A simulation study indicates that the proposed permutation tests firmly control the type I error whether or not the data are clustered. The use of the density ratio model improves the power of the tests. We also apply the proposed tests to data from the motivating application. The proposed permutation tests effectively address many real-world issues with trust worth inference conclusions.

Co-authors: Yukun Liu, Carolyn G. Taylor and James Zidek