Detecting Simultaneous Change-points in Aligned Sequences

David Siegmund
Department of Statistics
Stanford University
Stanford, CA 94305, USA
siegmund@stanford.edu

Abstract

We discuss the problem of detecting local signals that occur at the same location in multiple one dimensional noisy sequences, with particular attention to relatively weak signals that may occur in a small fraction of the sequences. We model these signals as temporary shifts in the mean of independent normally distributed observations; and to detect these shifts, we propose statistics that combine data across sequences. We provide approximations to the p-value of these statistics and show that they have better power properties and provide a more easily interpreted summary of the data than do procedures based on a separate analysis for each sequence. In particular, we examine the case where the signal is a temporary shift in the mean of independent Gaussian observations. The formulation of the model is motivated by the problem of detecting recurrent DNA copy number variants in multiple samples, and our results are illustrated by applications to data involving DNA copy number changes. We also indicate an application to online detection by distributed sensors of a change in the local environment of some of the sensors.

This is joint research with Benjamin Yakir and Nancy Zhang.