

“Repeated Measurements on Distinct Scales with Censoring”

Tanzy Love

Visiting Assistant Professor in Department of Statistics
Carnegie Mellon University

Abstract

Repeated Measurements on Distinct Scales with Censoring

We analyze data collected in a somatic embryogenesis experiment carried out on *Zea mays* at Iowa State University. The main objective of the study was to identify the set of genes in maize that actively participate in embryo development. Embryo tissue was sampled and analyzed at various time periods and under different mediums and light conditions. As is the case in many microarray experiments, the operator scanned each slide multiple times in order to find the slide-specific ‘optimal’ laser and sensor settings. The multiple readings of each slide are repeated measurements on different scales with differing censoring; they cannot be considered to be replicate measurements in the traditional sense. Yet it has been shown that the choice of reading can have an impact on genetic inference. We propose a hierarchical modeling approach to estimating gene expression that combines all available readings on each spot and accounts for censoring in the observed values.

We assess the statistical properties of the proposed expression estimates using a simulation experiment. As expected, combining all available scans using an approach with good statistical properties results in expression estimates with noticeably lower bias and root mean squared error relative to other approaches that have been proposed in the literature. Inferences drawn from the somatic embryogenesis experiment which motivated this work changed drastically when data were analyzed using the standard approaches or using the methodology we propose.