

1163-62-450

Natalie C. Gasca* (ncgasca@uw.edu) and **Robyn L. McClelland**. *Sparse Partial Least Squares Regression to Model the Relationship Between Correlated Diet Data and Risk of Heart Disease Events*. Preliminary report.

When relating the concurrent impact of foods on heart disease risk factors, nutritional researchers encountered frequent unintuitive results due to the strong correlations between some foods. Supervised dimension reduction methods such as sparse partial least squares (PLS) can reduce diet data in a tailored way to heart disease by selecting a few dietary patterns and foods while conserving predictive ability of the risk factor. To directly relate diet to the risk of incident disease, the Cox proportional hazards model is a standard approach that models the time to first event as a function of observed covariates. We propose incorporating sparse PLS into the Cox proportional hazards model to study the relationship between correlated data and heart disease risk while utilizing variable selection and imposing fewer model assumptions than existing methods. We compare sparse and non-sparse PLS Cox methods in their ability to create interpretable and predictive patterns, via simulations. We also apply the methods to diet data and time-to-coronary-disease in the Multi-Ethnic Study of Atherosclerosis. We will discuss in which settings this proposed method can construct interpretable patterns related to time-to-disease-events by leveraging variable selection and dimension reduction. (Received September 07, 2020)