Statistica Sinica **36** (2026), 1-22
doi:https://doi.org/10.5705/ss.202023.0285

GLOBAL GROUP TESTING AND SCREENING WITH DYNAMIC EFFECTS

Ying Cui and Limin Peng*

Emory University

Abstract: Identifying outcome-related variables is of general research interest in biomedical research. This task can be complicated by the presence of dynamic (or varying) variable effects that often manifest meaningful scientific mechanisms. Appropriately accounting for possible dynamic effects is crucial to avoid depreciating some important variables. In this work, we propose a model-free testing and screening framework by adopting a global view pertaining to the concept of interval quantile independence. The new framework not only permits robust identification of variables dynamically associated with an outcome, but also offers the flexibility to perform group testing that simultaneously evaluates multiple continuous or discrete covariates. We show that the key testing strategy can naturally evolve into unconditional and conditional screening procedures for ultra-high dimensional settings that enjoys the desirable sure screening property. We demonstrate good practical utility of the proposed methods via extensive simulation studies and a real application to a microarray data set.

Key words and phrases: Dynamic effects, hypothesis testing, interval quantile independence, variable screening.

1. Introduction

A general question arising from many biomedical studies is to determine whether some covariates are relevant to a study outcome. For example, in a genetics study, it is often of interest to identify a group of genes that contribute to the variations of a known disease marker or symptom (Subramanian et al., 2005; Efron and Tibshirani, 2007; Newton et al., 2007, for example). Addressing such an interest, however, may be complicated by the presence of dynamic (or varying) covariate effects. The key issue relates to the way how the relevance or importance of a covariate is defined. For instance, in the context of variable screening, the importance of a covariate was ranked by marginal correlation (Fan and Lv, 2008), maximum marginal likelihood estimate of a generalized linear model (Fan and Song, 2010) or a generalized marginal utility function (Fan, Samworth and Wu, 2009), and generalized correlation (Hall and Miller, 2009). These approaches involve an assumed linear or generalized linear relationship between the outcome and covariates or transformation thereof, which implicitly

*Corresponding author. E-mail: lpeng@emory.edu