Likelihood-based approach for handling interval-censored covariates in generalized linear models

Andrea Toloba[1], Guadalupe Gómez Melis[1], Klaus Langohr[1]

[1] Universitat Politècnica de Catalunya·BarcelonaTech
Department of Statistics and Operations Research

The development of methods to address censored covariates has gained significant attention in recent years. Although the problem itself is not new, its presence in real-world data has often been overlooked. Interval-censored covariate data, in particular, is frequently replaced by a single imputed value, which is known to introduce bias and underestimate variance. While recent methods have emerged for handling discrete time-to-event covariates, these approaches are often limited to survival analysis contexts, leaving other applications unaddressed.

In this talk, we shift our focus to analytical chemistry, specifically to data related to the quantification of compounds in mixtures. Compounds are often defined by multiple analytes, each measured via liquid chromatography and subject to analyte-specific detection and quantification limits. This chemical technique results in interval-censored data for the overall quantity of a compound. Our motivating example originates from metabolomics, exploring the association between circulating carotenoids—molecules present in the bloodstream—and cardiometabolic health. Advancing research in this area requires fitting generalized linear models for cardiometabolic biomarkers while incorporating interval-censored circulating carotenoid levels as a covariate.

Building on this example, we present an extension of the GEL algorithm[1], which was originally developed for time-to-event interval-censored covariates in linear models. The GEL algorithm is an EM-type method that alternates between estimating the distribution of the censored covariate and maximizing the model's likelihood function. However, like other recent approaches in the literature, it relies heavily on the assumption that the censored covariate has a discrete support, which limits its applicability. Our extension overcomes this limitation by handling interval-censored covariates nonparametrically and regardless of the distribution's support, broadening its usability to a wide range of applications.

[1] Gómez, Espinal and Lagakos (2003) Inference for a linear regression model with an interval-censored covariate. Stat in Med 22(3):409-25

Keywords: regression modeling; censored covariates; interval censoring