Statistica Sinica **37** (2027), 1-22
doi:https://doi.org/10.5705/ss.202024.0003

RANK-BASED INFERENCE FOR THE ACCELERATED FAILURE TIME MODEL WITH PARTIALLY INTERVAL-CENSORED DATA

Taehwa Choi†, Sangbum Choi† and Dipankar Bandyopadhyay*

Sungshin Women's University, Korea University and Virginia Commonwealth University

Abstract: This paper presents a unified rank-based inferential procedure for fitting the accelerated failure time model to partially interval-censored data. A Gehan-type monotone estimating function is constructed based on the idea of the familiar weighted log-rank test, and an extension to a general class of rank-based estimating functions is suggested. The proposed estimators can be obtained via linear programming and are shown to be consistent and asymptotically normal via standard empirical process theory. Unlike common maximum likelihood-based estimators for partially interval-censored regression models, our approach can directly provide a regression coefficient estimator without involving a complex nonparametric estimation of the underlying residual distribution function. An efficient variance estimation procedure for the regression coefficient estimator is considered. Moreover, we extend the proposed rank-based procedure to the linear regression analysis of multivariate clustered partially interval-censored data. The finite-sample operating characteristics of our approach are examined via simulation studies. Data example from a colorectal cancer study illustrates the practical usefulness of the method.

Key words and phrases: Accelerated failure time, clustered data, empirical processes, Gehan statistics, interval-censoring, left-censoring, survival analysis.

1. Introduction

Interval-censored (IC) responses are frequently encountered in various biomedical settings and health studies, requiring periodic follow-ups or inspections (Bogaerts, Komárek and Lesaffre, 2017), where the time to an event of interest is only known to fall within a particular interval. Instances of exact failure times observed for certain subjects within the same study lead to partially interval-censored data, which can be categorized into two groups: partly interval-censored (PIC) (Kim, 2003) and doubly-censored (DC) (Cai and Cheng, 2004). In addition to a certain number of exact observations (failures/events), DC data are either left-censored or right-censored (e.g., HIV/AIDS studies), whereas PIC data contain additional interval-censoring (e.g., cancer trials).

*Corresponding author. E-mail: dbandyop@vcu.edu
†Contributed equally to this work.