

# Two-phase outcome-dependent sampling with bivariate sequential survival data

Terry Lin

MSc candidate in Statistics

Department of Mathematics and Statistics, Memorial University

June 26, 2017

01:00 - 02:00 pm @ HH-3017

**Abstract:** In some observational studies, the covariates of interest might be expensive to measure although the outcome variable could easily be obtained. In this situation, a cost-efficient two-phase outcome-dependent sampling design could be employed to measure the expensive covariate for more informative subjects. In the first phase, all members of a random sample from a population or a cohort are measured for the outcome variable and inexpensive covariates. In the second phase, a subset of the cohort is selected based on the outcome variable, and the expensive covariate is measured only for the selected individuals. Case-cohort design is a commonly used outcome-dependent sampling design in time-to-event analyses. In generalized case-cohort design, in which the selection probability depends only on the event indicator, a random sub-sample of individuals who experienced the event are selected, along with a random sub-sample of those with censored event times. It was previously shown that when the selection probability in phase two depends on observed event time and censoring time in addition to the event indicator, the efficiency of the design might increase. Efficient design has a lower variance of the coefficient estimate of the expensive covariate in the regression model. In this study, we consider bivariate sequential time-to-event data, which consists of gap times between two events observed in sequence, as the outcome variables. The objective of this study is to explore efficient two-phase sampling designs for a predetermined phase two sample size. We consider sampling designs depending on the event indicators and gap times. A likelihood-based method is used to estimate the associations between the expensive covariate and the two gap times. We show that when the selection probability in phase two depends on the two observed gap times and censoring times in addition to their event indicators, the efficiency of the design might improve.