README

Introduction

This repository houses the R code used for simulation studies within the project titled "Two-phase designs with failure time processes subject to non-susceptibility." Inside, a folder labeled *utility* contains all utility functions. Additionally, there's a running script named *sim_script.R* designed for executing simulations.

utility

This folder contains five *.r* files.

design.r and *design_withV.r*: contains functions for phase II designs

em_coxph.r: contains functions for esitmation via EM algorithm

avar_louis.r: contains functions for asymptotic variance computation for EM

opt_str_design_gau.r: contains functions for optimal stratified designs, which are not feasible in practice

sim_script.R

This is a script for one simulation study. A for-loop is used for repeated simulations. In each simulation run, 3 steps are included: 1) data generation; 2) two-phase design implementation; and 3) estimation and inference.

sample_dt.phII.csv

A toy data set showing the data structure after a phase II sub-sampling:

- del: event status;
- time: observed times;
- x: expensive covariate, NA for missing;
- v: inexpensive covariate;
- id: subject id;
- r: a binary indicator suggesting if the subject was selected into the phase II sub-sample.