Program

Speaker 1: Mengjiao Peng (East China Normal University, China)

Title: *Joint mean variance screening for ultrahigh dimensional categorical predictors with multiple survival outcomes*

Abstract: Medical research often involves collecting ultrahigh-dimensional gene features, where the number of gene features is much larger than the sample size, creating challenges in survival prediction and disease mechanism research in precision medicine. The focus of this article is to extract important categorical covariate features from vast amount of genomic data for further survival outcome prediction. The complexity of the problem is heightened when studying two types of survival endpoints, one being possibly censored by the other, known as semi-competing risks outcomes. We propose the Joint Mean Variance Index (JMV) as a measure of the correlation between outcomes and a categorical covariate, aiming in identifying important gene features in ultrahigh-dimensional multiple survival outcomes data. The proposed screening procedure is model-free, robust to model misspecification, heavy tails, and outliers. Additionally, we show that the screening procedure exhibits sure independence and ranking consistency. Extensive numerical simulations are conducted to confirm the theoretical properties.

Speaker 2: Zili Zhang (Huaqiao University, China)

Other authors: Christiana Charalambous, Peter Foster

Title: *A Gaussian copula joint model for longitudinal and time-to-event data with random effects*

Abstract: Longitudinal and survival sub-models are two building blocks for joint modelling of longitudinal and time-to-event data. Extensive research indicates separate analysis of these two processes could result in biased outputs due to their associations. Conditional independence between measurements of biomarkers and event time process given latent classes or random effects is a conventional approach for characterising the association between the two sub-models while taking the heterogeneity among the population into account. However, this assumption is difficult to validate because of the unobservable latent variables. Thus, a Gaussian copula joint model with random effects is proposed to accommodate the scenarios where the conditional independence assumption is questionable. The conventional joint model assuming conditional independence is a special case of the proposed model when the association parameters in the Gaussian copula shrink to zero. Simulation studies and real data application are carried out to evaluate the performance of the proposed model with different correlation structures. In addition, personalised dynamic predictions

of survival probabilities are obtained based on the proposed model and comparisons are made to the predictions obtained under the conventional joint model.

Speaker 3: Takeshi Emura (the ISM, Japan)

Other authors: Marc Ditzhaus, Dennis Dobler, Kenta Murotani
Title: *Factorial survival analysis for treatment effects under dependent censoring*
Abstract: Factorial analyses offer a powerful nonparametric means to detect main or interaction effects among multiple treatments. For survival outcomes, e.g. from clinical trials, such techniques can be adopted for comparing reasonable quantifications of treatment effects. The key difficulty to solve in survival analysis concerns the proper handling of censoring. So far, all existing factorial analyses for survival data were developed under the independent censoring assumption, which is too strong for many applications. As a solution, the central aim of this article is to develop new methods in factorial survival analyses under quite general dependent censoring regimes. This will be accomplished by combining existing results for factorial survival analyses with techniques developed for survival copula models. As a result, we will present an appealing F-test that exhibits sound performance in our simulation study. The new methods are illustrated in real data analysis. We implement the proposed method in an R function surv.factorial(.) in the R package *compound.Cox*.

Speaker 4: Dongdong Li (Harvard Pilgrim Health Care Institute, USA)

Title: *Proportional hazards regression models for interval-censored outcome with interval-censored covariates*
Abstract: Identifying predictors for features of viral rebound trajectories after antiretroviral therapy (ART) interruption is central to HIV cure research. Motivated by the need to assess whether the time to viral suppression after ART initiation is predictive of the time to viral rebound after ART interruption, we investigate modeling approaches that relate an interval-censored outcome (e.g., time to viral rebound) and an interval-censored covariate (e.g., time to viral suppression). We develop estimation and inference procedures for fitting a proportional hazards regression model when both the outcome and a covariate can be interval-censored, without making parametric distributional assumptions about baseline hazard functions, through the use of an Expectation-Maximization algorithm. As some participants experienced multiple episodes of ART initiation and interruption, we further extend the proposed method to accommodate the clustering effect of multiple observations from the same participant. We evaluate the finite-sample performance of the proposed method for both independent and clustered data settings through simulation studies. To illustrate, we assess the

effect of time to viral suppression after ART initiation on time to viral rebound after ART interruption using data from the Zurich Primary HIV Infection cohort.

Speaker 5: IL DO HA (Dept. of Statistics, Pukyong National University, South Korea)

Other authors: Hangbin Lee (Seoul National University), Youngjo Lee (Seoul National University)
Title: *Deep Neural Network for Semi-parametric Frailty Models via H-likelihood*
Abstract: Recently, deep neural network (DNN) has provided a major breakthrough to enhance prediction in various areas. For prediction of clustered time-to-event data, we propose a new DNN-based frailty model (DNN-FM). An advantage of the proposed model is that the joint maximization of the new h-likelihood provides maximum likelihood estimators for fixed parameters and best unbiased predictors for frailties. Thus, the proposed DNN-FM is trained by using a negative profiled h-likelihood, constructed by profiling out the non-parametric baseline hazard, as a loss function. Numerical studies show that the proposed method improves the prediction performance of the existing methods (e.g. DNN-Cox model) in terms of Brier score and C-index. A real data analysis shows that the inclusion of subject-specific frailties to the DNN-Cox model helps to improve the risk prediction of the DNN-Cox model. We also present an online learning framework of the proposed DNN model for large data set.
Keywords: Deep neural network, Frailty models, H-likelihood, Random effect.

Speaker 6: Xinyuan Song (The Chinese University of Hong Kong, HK)

Title: *A tree-based Bayesian accelerated failure time cure model for estimating heterogeneous treatment effect*
Abstract: Estimating heterogeneous treatment effects has drawn increasing attention in medical studies, considering that patients with divergent features can undergo a different disease progression even with identical treatment. Such heterogeneity can co-occur with a cured fraction for biomedical studies with a time-to-event outcome and further complicates the quantification of treatment effects. This study considers a joint framework of Bayesian causal forest and accelerated failure time cure model to capture the cured proportion and treatment effect heterogeneity through three separate Bayesian additive regression trees. Under the potential outcomes framework, conditional and sample average treatment effects within the uncured subgroup are derived on the scale of log survival time subject to right-censoring, and treatment effects on the scale of survival probability are derived for each individual. Bayesian backfitting Markov chain Monte Carlo algorithm with the Gibbs sampler is conducted to estimate the causal effects. Simulation studies show the satisfactory performance of the proposed method. The proposed model is then applied to a breast cancer dataset extracted from the SEER database to demonstrate its usage in detecting heterogeneous treatment effects and cured

subgroups.

Speaker 7: Liming Xiang (Nanyang Technological University, Singapore)

Title: Multiple imputation for competing risks analysis with covariates subject to detection limits
Other authors: Yilin Wu, Rui Huang, Huixia Judy Wang.
Abstract: Competing risks occur commonly in survival analysis when subjects may experience multiple types of events and the occurrence of the primary event of interest can be precluded by a competing event. Challenges arise for the analysis of competing risks data with covariates subject to censoring due to detection limits. We propose a semiparametric multiple imputation method for inference under the subdistribution hazard model. Our proposal automatically utilizes the information from the exactly observed covariate values and the outcome data to efficiently impute the censored covariates based on rejection sampling. We establish consistency and asymptotic normality of the resulting estimator and demonstrate its promising finite sample performance via extensive simulations. Finally, we illustrate the application of the method with the data from a study of community acquired pneumonia.

Speaker 8: Feng-Chang Lin (University of North Carolina at Chapel Hill, USA)

Title: *A more efficient estimator for competing risks data with missing cause of failure*
Abstract: When a competing risk observation misses the cause of failure, one common approach, other than a complete-case analysis, is to utilize the cause-specific hazards to impute the missing indicator with a probability of being the type. However, in addition to the baseline markers used as covariates for the cause-specific hazards, the same markers are likely collected when the event happens. Unfortunately, such information is usually observed but seldom used. In this presentation, we will demonstrate how we can utilize the transition information of markers to provide a more accurate probability estimation of the missing indicator and, therefore, provide a more efficient estimator for the parameter estimation. To demonstrate our method's feasibility, we will apply the newly developed approach to two P. vivax malaria recurrent infection data and Pseudomonas aeruginosa (Pa) infections in young cystic fibrosis (CF) patients.