Abstracts for International Workshop: Survival Analysis for Medical and Health Data: 2025/8/19(Tue): 13:30～17:10

13:30 - 14:25   Session I (Parametric models and inference)

Lachos Davila Victor Hugo;   Nanami Taketomi

14:35 - 16:00 Session II: (Dependence modeling)

Li-Hsien Sun, Sy Han Chiou, Takeshi Emura

16:10-17:00 Session III: (Competing risks)

Ralf Andreas Wilke, IL-Do Ha

Speaker 1: Lachos Davila, Victor Hugo (University of Connecticut, USA)

Title: Mixture cure models with flexible survival and link functions: Parameter estimation via the EM algorithm

Abstract: The mixture cure model is a special type of survival model, and it assumes that the studied population is a mixture of susceptible individuals who may experience the event of interest and cure individuals who will never experience the event. The Weibull model is typically specified as a latency component for the event time and symmetric logistic regression as an incidence component for the probability of being uncured. Challenges arise when incorporating alternative survival distributions and link functions. To address this, we propose a flexible mixture cure rate model, where parameter estimation is performed using the Expectation-Maximization algorithm. This approach is easily implementable in statistical software that supports weighted survival regression models along with the efficient routine glm(), which allows the definition of custom link functions. The proposed method delivers notable gains in computational efficiency, with standard errors as a byproduct. In simulations under varied scenarios and in an analysis of real liver cancer data, we demonstrate the efficiency and flexibility of our proposed model. An EMGCR R package has been developed and is available on the GitHub repository.

Speaker 2: Nanami Taketomi (Nagasaki University, Japan)

Title: Shrinkage estimation for the rate parameter under the exponential distribution with censored survival data

Abstract: The exponential distribution is widely applied in survival and reliability analyses. To estimate the rate parameter of the exponential distribution with censored survival data, maximum likelihood estimation is usually employed. However, in case the number of events or subjects is small for the data, the maximum likelihood estimator can have large bias and variance. In this talk, we present a shrinkage estimator for the rate parameter of the exponential distribution using a penalized log-likelihood by adding a penalty for a large parameter value. We also derive theoretical properties of the penalized maximum likelihood estimator. Simulation results show that the proposed method provides the smaller mean squared error than the maximum likelihood estimation. Finally, we apply the proposed method to a real data.

**Speaker 3**: Li-Hsien Sun (National Central University, Taiwan)

Title: Change point analysis under a copula-based Markov model for Weibull time series

Abstract: Change point detection is a crucial topic which identifies the structure change based on shifts in underlying mechanisms. We propose the copula-based Markov model with the Clayton copula and the Joe copula, where the marginal distribution is the Weibull distribution. The Weibull distribution is widely used in real data analysis such as survival analysis, reliability analysis, and financial engineering. The copula function is applied to deal with nonlinear dependence for sequential data. We derive the likelihood function and obtain

the maximum likelihood estimation (MLE) using the Newton-Raphson method for the change points and parameters. The performance of the proposed model is illustrated through simulation studies. In empirical studies, we analyze the VIX data from the period of 2020 COVID-19 and identify the structure change by the proposed model.

Reference: Sun, L. H., et al. (2025). Change point estimation for Gaussian time series data with copula-based Markov chain models. *Computational Statistics*, 40(3), 1541-1581.

Speaker 4: Sy Han Chiou (Southern Methodist Univ, USA)
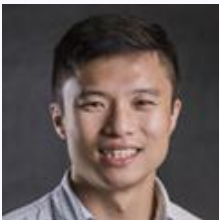
Title: Analyzing bivariate survival data using pseudo-observations

Abstract: Copula models have become increasingly popular in various fields as they provide effective tools for modeling correlated responses. In modeling multivariate survival data, copula models offer flexibility by enabling users to specify both the marginal survival functions and the association structure between them. A semiparametric transformation model is used to define the marginal survival functions, and a conditional Archimedean copula is considered to address the associations among different types of survival times. To expedite computation, pseudo-observations for both the marginal survival and association components are introduced, and inference is implemented using GEE techniques. Additionally, variable selection and goodness-of-fit tests are explored to aid in the selection of appropriate copula models. The effectiveness of the proposed methods is demonstrated through simulations.

Speaker 5: Takeshi Emura (Hiroshima University, Japan)

Title: Inferring median survival under dependent censoring

Abstract: The key difficulty in survival analysis is the proper handling of censoring. So far, existing inference methods for median survival have been developed under the independent censoring assumption, which is too strong for many applications. As a solution, one could develop new methods for dependent censoring regimes. This would be accomplished by combining existing nonparametric methods with copula-graphic estimators developed for survival copula models. We present this methodology with simulations and data examples.

Reference: Emura, T., Ditzhaus, M., Dobler, D., & Murotani, K. (2024). Factorial survival analysis for treatment effects under dependent censoring. *Statistical Methods in Medical Research*, 33(1), 61-79.

Speaker 6: Ralf Andreas Wilke (Copenhagen Business School, Denmark)

Title: On the implications of proportional hazards assumptions for competing risks modelling

Abstract: The assumption of hazard rates being proportional in covariates is widely made in empirical research and extensive research has been done to develop tests of its validity. This paper does not contribute on this end. Instead, it gives new insights on the implications of proportional hazards (PH) modelling in competing risks models. It is shown that the use of a PH model for the cause-specific hazards or subdistribution hazards can strongly restrict the class of copulas and marginal hazards for being compatible with a

competing risks model. The empirical researcher should be aware that working with these models can be so restrictive that only degenerate or independent risks models are compatible. Numerical results confirm that estimates of cause-specific hazards models are not informative about causal patterns in the data generating process.

Speaker 7: IL-Do Ha (Pukyong National University, South Korea)

Title: Robust estimation on competing-risks frailty models for clustered competing-risks data

Abstract: Competing risks data arise when occurrence of an event hinders observation of other types of events, and they are encountered in various research areas including biomedical research. In particular, in the analysis of clustered competing risks data, not only the event times within the same cluster, but also the competing events, are possibly correlated. Such correlation can be modeled via a cluster-specific frailty following, for example, a log-normal distribution. The maximum hierarchical (h-)likelihood method, which gives the (approximate) maximum likelihood estimators (MLEs), has been used for parameter estimation in the competing-risk frailty model with log-normal frailty. However, how robust the method is when the frailty distribution is unknow remains an open question. In this talk, we theoretically establish that the MLEs of the regression coefficients under cause-specific competing-risk frailty model with the log-normal frailty obtained from the maximum h-likelihood estimators (MHLEs) can be robust against mis-specification of the frailty distribution. The proposed method is demonstrated with simulation study and real data sets. In addition, we discuss further extensions.

Keywords: competing risks frailty models; h-likelihood; clustered competing-risks data