A Practical Guide to Analyzing Complex Survival Data with Kaplan-Meier

Luís Meira-Machado[1]

[1]lmachado@math.uminho.pt, Department of Mathematics and Centre of Mathematics, University of Minho, Portugal

The estimation of survival has been a widely researched topic in statistical and medical literature. Among the commonly used estimators, the Kaplan-Meier product-limit estimator is preferred due to its nonparametric nature, which does not rely on any assumptions about the lifetime probability distribution. The estimator is computed using a product of elementary probabilities, which calculates conditional survival probabilities. The redistribution to the right algorithm is another method that is employed to estimate the Kaplan-Meier estimator of survival by redistributing the mass of a censored subject equally among those who fail or are censored at later times. In this work, we present additional alternative representations of the Kaplan-Meier estimator and discuss their applications and advantages of its usage. One of these representations defines the estimator as a sum of weights, which is useful in estimating various quantities in the context of multi-state models. The estimator can also be represented as a weighted average of identically distributed terms, where the weights are obtained by using the inverse probability of censoring. We also demonstrate how these formulations can be applied to estimate different quantities, particularly in the context of multi-state models. Additionally, we will also focus on methods that aim to reduce the variability of the estimators and allow for the estimation of some of these quantities conditionally on covariates. To illustrate these methods, we include two real data examples from medicine.

Keywords: Kaplan-Meier, Multi-state models, Survival Analysis