

## M2 Research internship in applied mathematics

**Identifiability of dependence structures between event times and censoring in survival analysis.**

### Context

Modern survival analysis [1] routinely deals with complex follow-up data, where event times (e.g. relapse, death, recovery, disability) might not be observed for all subjects. Instead, follow-up may be interrupted by administrative end of study, loss to follow-up, or occurrence of competing events. In such situations, the standard assumption of independent censoring is often unrealistic: the mechanism that censors an individual can itself be related to their underlying risk. Ignoring this dependence may bias classical estimators (Kaplan–Meier, Cox models, log-rank tests, etc.) and lead to misleading conclusions for prognosis or risk evaluation.

On the other hand, since Tsiatis [2] in 1975, the dependence structure hidden behind a censoring mechanism is known not to be nonparametrically identifiable. Recent developments [3], [4], [5] shows that it might be however parametrically identifiable, under conditions depending on the parametric models. However, outside a few simple parametric copula families[6], [7], little is known about the extent of these identifiability conditions.

This internship will contribute to the development of new statistical methods to model dependence structures underlying censored observations using copulas and related tools, in order to better understand when reliable inference is still possible under censoring. The goal is to combine a focused literature review with first numerical experiments and reproducible code, laying the groundwork for more advanced methodological work on the topic.

### Objectives

The main objective of the internship is to explore, from both a theoretical and practical perspective, how survival analysis can accommodate dependence between event times and censoring times. More concretely, the work will focus on:

- Conducting a targeted literature review on dependent censoring, copula-based survival models and identifiability issues in censored data.
- Develop mathematical tools and theoretical concepts to classify the parametric families into identifiable or not, include specific cases of non-smooth dependence structures.
- Implementing small-scale simulation studies (in e.g., Julia, R or Python) to illustrate the impact of dependent censoring on classical survival estimators and to test the practical identifiability of the developed models.
- Depending on progress and interest, carrying out an illustrative case study on a publicly available dataset (e.g., large-scale registry or cohort data) to showcase the practical relevance of the methods investigated during the internship.

The internship will take place at SESSTIM in a research environment combining applied mathematics and public health. Communication of the obtained results will be an integral part of the work. Depending on mutual interest and available funding, the internship may naturally evolve into a PhD project on advanced dependence modelling under censoring.

### Candidate's profile

The following are strictly required:

- Master 2 level in applied mathematics, statistics and/or computer sciences, with at least an introductory knowledge of probability and statistical inference.
- Capacity to 1) write in LaTeX, 2) collaborate with Git, and 3) code in at least one scientific language (Julia, Python, R)
- Very good written English.

---

Aix-Marseille Université – Faculté des sciences médicales et paramédicales – 27, boulevard Jean Moulin 13385 Marseille Cedex 5 – France Tél : (+33) 04 91 32 46 00 / Courriel : [sesstim-u1252@inserm.fr](mailto:sesstim-u1252@inserm.fr) / Site : <https://sesstim.univ-amu.fr/>

Prior experience in survival analysis (hazard, censoring, Kaplan–Meier, regression models) and/or copula modelling is a plus, but not strictly mandatory if the candidate is strongly motivated to learn these topics. We do not require prior knowledge of the Julia programming language (which may be used for the numerical part), only solid programming skills in at least one scientific language (R, Python, Julia, ...) and the willingness to adapt.

## Additional information

- **Length:** 4 to 6 months, starting from February to April 2026.
- **Location:** At SESTIM, on the *Faculté des Sciences Médicales et Paramédicales* in Marseille, France.
- **Wages:** Regulatory internship stipend at the Université Aix-Marseille (4€35/hour for 35h/week, about 600€/month)
- **To apply:** Please send your application with a resume and a few motivational lines to [oskar.laverny@univ-amu.fr](mailto:oskar.laverny@univ-amu.fr) with [Internship2026] in the subject. You may add links to previous scientific (related or not) projects you did.

## References

- [1] T. R. Fleming et D. P. Harrington, *Counting processes and survival analysis*, vol. 625. John Wiley & Sons, 2013.
- [2] A. Tsiatis, « A nonidentifiability aspect of the problem of competing risks. », *Proc. Natl. Acad. Sci.*, vol. 72, n° 1, p. 20-22, janv. 1975, doi: 10.1073/pnas.72.1.20.
- [3] C. Czado et I. V. Keilegom, « Dependent censoring based on copulas », 14 avril 2021, *arXiv*: arXiv:2104.06872. doi: 10.48550/arXiv.2104.06872.
- [4] N. W. Deressa et I. V. Keilegom, « Copula Based Cox Proportional Hazards Models for Dependent Censoring », *J. Am. Stat. Assoc.*, vol. 119, n° 546, p. 1044-1054, avr. 2024, doi: 10.1080/01621459.2022.2161387.
- [5] G. Crommen, N. W. Deressa, M. D'Haen, J. Ding, I. Willems, et I. Van Keilegom, « Recent advances in copula-based methods for dependent censoring », *SORT-Stat. Oper. Res. Trans.*, p. 3-42, 2025.
- [6] R. B. Nelsen, *An introduction to copulas*, 2nd ed. in Springer series in statistics. New York: Springer, 2006.
- [7] H. Joe, *Dependence modeling with copulas*. CRC press, 2014.