

# Development of the field of Point Process Learning

Ottmar Cronie

The modern-day capacity to sample data has resulted in data structures which are extensive in both size and complexity. One aspect of the increased complexity is that the data-generating mechanisms under study may exhibit statistical dependence. A further aspect is that, a priori, one typically does not know what the sample size will be, whereby the sample size should be treated as a random variable. Hence, the classical random (iid) sample setting, which was originally motivated by controlled trials and mathematical tractability, may not be suitable for such data structures. Note that ignoring inherent dependencies may lead to (severe) overfitting (Cronie and Van Lieshout, 2018). In contrast, *point processes* (Baddeley et al., 2015; Daley and Vere-Jones, 2008) may be described as *generalised random samples* where we allow for the sample points to be dependent and for the sample size to be random. Consequently, they have become ubiquitous in analyses of spatially and/or temporally referenced event data, so-called *point patterns*, where examples include collections of disease cases, fire incidents, earthquakes and accidents (Baddeley et al., 2015; Daley and Vere-Jones, 2008).

The lack of a statistical learning theory for point processes has recently been addressed by Cronie et al. (2021) for the first time, leading to the first general statistical learning theory for point processes, which is referred to as *Point Process Learning*. It is based on the combination of two new building blocks, namely point process cross-validation and point process prediction.

A range of different topics related to the development of the field of point process learning can be explored as master's thesis projects. Some of these will, in particular, be related to i) the spatial-temporal modelling/analysis of Covid-19 in Sweden and ii) traffic accident analysis. The projects will be supervised by Ottmar Cronie, Mathematical sciences, who may be contacted for further details: ottmar@chalmers.se.

## References

- Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: methodology and applications with R*. CRC press.
- Cronie, O., Moradi, M., and Biscio, C. A. (2021). Statistical learning and cross-validation for point processes. *arXiv preprint arXiv:2103.01356*.
- Cronie, O. and Van Lieshout, M. N. M. (2018). A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika*, 105(2):455–462.
- Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. Springer-Verlag New York, second edition.