From London to Florence: Harnessing Information Flow in Early Modern Europe with Bayesian Time-To-Event Statistics

Dr. Gabor Mihaly Toth, University of Luxembourg

Early modern Europe witnessed an information revolution featured - in part - by the continual circulation of handwritten news sheets (Salmi-Niklander & Droste, 2019; Davies & Fletcher, 2014). Handwritten news sheets were produced by semi-professional newsagents, postmasters, and diplomats in large urban centers; they collected incoming news, compiled news sheets, and forwarded them to European courts and to other urban centers. From the 1600s handwritten news sheets were complemented with printed news, though they continued to play an important role (Infelise, 2002). Even though handwritten newsletters that once circulated had a significant impact on the history of the continent, their study has remained challenging (Baron & Dooley, 2005). Many of the once existing news sheets did not survive. Those that survived are dispersed in different collections such as for instance the Fugger Archive and the Medici Archive (Barker, 2016; Bauer, 2011). These collections are *incomplete samples* from the once existing but today lost corpus of handwritten news sheets. In light of loss and incompleteness, how can we harness the information flow between large urban centers of early modern Europe?

The goal of this paper, based on an accomplished study, is to present how time-to-event statistics combined with Bayesian inference helped me cope with incompleteness and give insight into information flow (Schober & Vetter, 2018). The focus of the talk will be the methodology; the historical interpretation of the results is the subject of another forthcoming presentation.

In the first part of the presentation, I will outline the general context of early modern handwritten news sheets. I will then present the problem of loss and incompleteness through a collection of approximately 1000 new sheets authored by a certain Amerigo Salvetti (Villani, 2004). This collection (today in the Florentine State Archive; it has been digitized and transcribed; it is available in the MIA database of the Medici Archive Project, https://mia.medici.org/) was produced in London during the mid 17th century. Salvetti was a Florentine diplomat in residence in England; he collected news arriving at London from all over the world; he translated and forwarded them to Florence, to the grand ducal court of the Medici, in handwritten news letters and dispatches. The Salvetti collection of news sheets offers a unique insight into the early modern handwritten news culture. For instance, we can address the question of what featured the information flow between London and Florence. What were the most important geo locations from where news arrived at London and then Florence? How often did news arrive from different sources? Nevertheless, answering these questions remains challenging due to the incompleteness of the Salvetti collection. Even if the Salvetti collection is exceptionally large, it does not contain all newsletters that Salvetti compiled; furthermore, most probably, Salvetti did not forward each news he learned about.

As a result, the straightforward counting of the number of news items arriving from a given location in the Salvetti collection is not sufficient. In short, through the example of the Salvetti collection, I will introduce the audience into the problems that incompleteness brings about and demonstrate the need for more complex quantitative methods.

In the second part, I will discuss the quantitative approaches I applied to resolve the problem set. First, I will explain how I applied the Poisson Process (a basic model in time-to-event statistics) to quantify the frequency of news arriving from a given location such as for instance Copenhagen (throughout the talk I will use Copenhagen as a tangible example). The Poisson process is used to quantify the occurrences of events that happen at a certain rate, though at random (Streit, 2010). The Poisson process also models the time between events. For instance, a volcano erupts periodically but the exact intervals between eruptions are changing and feature a certain degree of randomness. With the Poisson process we can harness how this randomness works and find the most likely amount of time between two eruptions or until the next eruption (Connor & Hill, 1995). By the same token, the Poisson process helped me model the randomness of news arriving from a given location. For instance, in the Salveti collection, there are regular news items from Copenhagen but the amount of time between the arrivals is always changing; sometimes there is a month between two pieces of news, sometimes there are three months, and so on. With the help of the Poisson process, I could summarize the set of all possible intervals; I could then calculate the most likely amount of time one had to wait until the next piece of news arrived from Copenhagen. The Poisson Process also enabled me to quantify the rate with which news arrived from different locations. In short, throughout my talk I will argue that the calculation of the most likely amount of time until the arrival of the next piece of news from a given location is a versatile way to quantify information flow from a given location to another one.

In practice, I parsed all geolocations in the Salvetti collection, extracted the time stamps of their occurrences, and inferred the temporal intervals between the occurrences. I could then calculate the rate of news arriving from each geolocation, in the Salvetti collection.

In the third part, I will briefly show how I applied Bayesian statistics to address the incompleteness of the Salvetti collection (Lo, 1986).

In the conclusion, I will critically reflect on the application of Bayesian inference. First, I will discuss the limitations of the methodology. Specifically, I will address the problem of how to model change over time given that the current setting of the methodology treats the Salvetti collection as one monolithic whole. Second, I will reflect on the question, to which extent can Bayesian statistics help students of media history move beyond individual collections (such as for instance the Salvetti collection) and gather insights into information flow in early modern Europe?

References

Aboura, K., & Agbinya, J. I. (2013, July). Bayesian analysis of stay times in a system. *2013*

Pan African International Conference on Information Science, Computing and Telecommunications (PACT). http://dx.doi.org/10.1109/scat.2013.7055107

Barker, S. (2016). 'Secret and Uncertain': A history of Avvisi at the court of the medici grand dukes. In N. Moxham (Ed.), *News Networks in Early Modern Europe* (pp. 716–738). Brill. http://dx.doi.org/10.1163/9789004277199_032

Baron, S. A., & Dooley, B. (2005). *The politics of information in early modern Europe*. Routledge.

Bazett, T. (2022). Summary: Bayesian inference. In *Bayesian Inference*. Springer International Publishing. http://dx.doi.org/10.1007/978-3-030-95792-6_7

Bauer, O. (2011). *Zeitungen vor der Zeitung: Die Fuggerzeitungen (1568-1605) und das frühmoderne Nachrichtensystem*. Oldenbourg Verlag.

Berger, J. O. (1980). Bayesian analysis. In *Statistical Decision Theory* (pp. 89–168). Springer New York. http://dx.doi.org/10.1007/978-1-4757-1727-3_4

Connor, C. B., & Hill, B. E. (1995). Three nonhomogeneous Poisson models for the probability of basaltic volcanism: Application to the Yucca Mountain region, Nevada. *Journal of Geophysical Research: Solid Earth*, *100*(B6), 10107–10125. https://doi.org/10.1029/95jb01055

Davies, S., & Fletcher, P. (2014). *News in early modern Europe: Currents and connections*. BRILL.

Droste, H., & Salmi-Niklander, K. (2019). *Handwritten newspapers: An alternative medium during the early modern and modern periods*. BoD - Books on Demand.

Infelise, M. (2002). *Prima dei giornali: Alle origini della pubblica informazione, secoli XVI e XVII*.

Lo, A. Y. (1986). Bayesian statistical inference for sampling a finite population. *The Annals of Statistics*, *14*(3). https://doi.org/10.1214/aos/1176350061

Riddell, A., & Betancourt, M. (2021). Reassembling the english novel, 1789–1919. *Journal of Cultural Analytics*, *6*(1). https://doi.org/10.22148/001c.19102

Schober, P., & Vetter, T. R. (2018). Survival analysis and interpretation of time-to-event data. *Anesthesia & Analgesia*, *127*(3), 792–798. https://doi.org/10.1213/ane.0000000000003653

Streit, R. L. (2010). The poisson point process. In *Poisson Point Processes* (pp. 11–55). Springer US. http://dx.doi.org/10.1007/978-1-4419-6923-1_2

Tiihonen, I. L. I., Tolonen, M., & Lahti, L. (2021, October). Probabilistic Analysis of Early Modern British Book Prices. In *CHR 2021 Proceedings of the Conference on Computational Humanities Research* 2021.

Villani, S. (2004). Per la progettata edizione della corrispondenza dei rappresentanti toscani a Londra: Amerigo Salvetti e Giovanni Salvetti Antelminelli durante il "Commonwealth" e il Protettorato (1649-1660). *Archivio Storico Italiano,* 162 (1), 109 - 125.