Presence and Absence of Women in Early Modern Handwritten News: Random Walks in the Medici Archive

Toth, Gabor Mihaly
gabor.toth@maximilianeum.de
University of Luxembourg

In February 1600 a peculiar news reached the Medici court: a Persian woman ambassador arrived at the Ottoman court in Constantinople. The report came from Wien in a handwritten news sheet (ASF MdP, 3087). How did contemporary news mongers receive this news? How often did they read about women? Did the presence of women increase or decrease over time? Since scholarship on early modern women history has not surveyed handwritten news sheets, there are no answers to these questions (Wiesner-Hanks 2019; Benson 2010; Rublack 2001).

This paper will address them by presenting a completed computational study of a digitized corpus of transcribed handwritten news sheets in Italian from the Medici Archive (today in the Florentine State Archive). The corpus (published in Dooley et al. 2021), henceforth MIA-Euronews Corpus, consists of 1200 news sheets covering a period of 200 years and compiled in the major urban center of early modern Europe; each news sheet is a consecutive list of news items organized into distinct paragraphs. The paper will also present advanced statistical methods hitherto rarely applied in digital humanities: random sampling with replacement, aka bootstrapping (Bradley 1982), combined with Bayesian statistics (Lo 1986).

First, the paper will offer an introduction into handwritten news. 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). 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). Most 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 just *random* and *incomplete* samples from the once existing but today lost corpus of handwritten news. If surviving new sheets are random and incomplete, how can we use them to explore the presence of women in the lost whole of once existing new sheets?

Second, the paper will address the problem of randomness and demonstrate how bootstrapping can harness it. I will draw on a thought experiment:

Consider an AI-enabled robot that is literally a random walker. It keeps wandering in the Medici Archive; it keeps opening folders containing news sheets. It finds that sometimes women are mentioned, sometimes there is a complete silence on them. It thus concludes that women's presence is random in the news sheets.

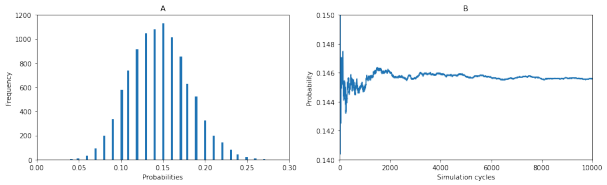

Figure 1. Probability of women occurring in news items randomly sampled from the MIA-Euronews Corpus. Figure A shows the probability distribution as a histogram. Figure B shows how the mean probability is reaching a stable value throughout the repeating random sampling.

However, I will point out that randomness can be investigated given that it is finite. I will demonstrate this by redesigning the thought experiment as a proper bootstrapping process (Aa et al., 1994):

Imagine that our AI-enabled robot randomly selects 100 news sheets and 100 news items at once; then it reports us two pieces of information. First, it tells us in how many news sheets it found the occurrence of at least one woman. Second, it tells us how many news items mentioned at least one woman. Next, we ask our random walker robot to repeat the two tasks over thousands of times.

I implemented this thought experiment algorithmically on the MIA-Euronews Corpus; in my talk I will present the implementation with a real time simulation (Figure 1 A and B) and by focusing on two aspects. First, with the bootstrapping we can map all possible outcomes (women mentioned 10,12, and so on times in hundred news sheets randomly selected) as well as their respective probabilities. Second, the random walk of the robot will sooner or later converge and reach a stable probability. I will point out that harnessing the randomness of a dataset involves studying that dataset from all possible angles (i.e. reshuffling and resampling as the bootstrapping does) and summarizing all possibilities with a probability distribution.

Third, I will discuss the problem of how to get from the evidence collected in the Medici Archive to the lost whole of news sheets circulating in the early modern period. I will rely on an illustrative example:

Consider an unfair coin; we do not know the prior probability of head and tail. We can toss the coin over thousands of times and gather evidence.

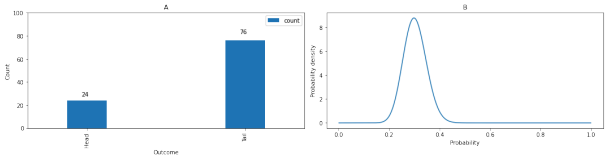

Figure 2. An illustrative example: assessing the bias of an unfair (probability of Head: 0.3) coin by flipping it for a hundred times.

The two outcomes (Head and Tail) are modeled with a binomial random variable. Figure A shows the outcome of the tossing, i.e. the evidence. Figure B shows the most likely prior distribution that could give rise to the evidence gathered throughout the flipping; it renders the outcome of the tossing with the beta distribution, which maps the discrete values (number of Head and Tail) onto a continuous probability space between 0 and 1. The closer the mean value is to 0, the less likely that the coin is biased towards Head and more likely that it is biased towards Tail. The beta distribution on Figure B actually gives a correct guess of the prior probability of Head (i.e. 0.3).

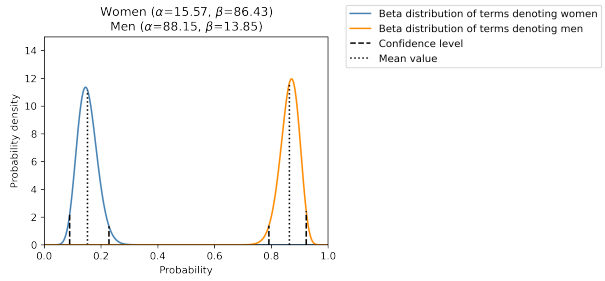

Figure 3. The probability of women's and men's presence or absence in news items randomly selected from the MIA-Euroews Corpus.

Absence and presence are modeled with a binomial random variable; the plot renders the variable with the beta distribution, which maps the discrete values (absence or presence) onto a continuous probability space between 0 and 1. The closer the mean value is to 0, the more likely that the outcome of the random variable is absence (as in case of women); conversely, the closer the mean value is to 1, the more likely that the outcome of the random variable is presence (as in case of men).

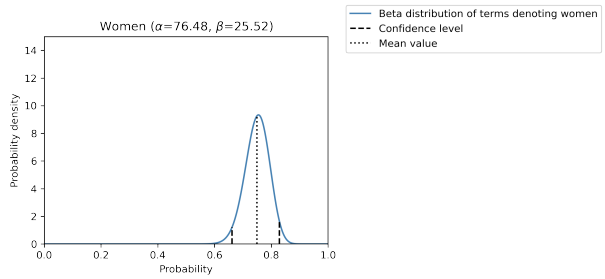

Figure 4. The probability of women's absence or presence in complete news sheets randomly selected from the MIA News Corpus. For further explanation of the plot, see the legend under Figure 3.

I will accomplish a live simulation and show that the prior probability of Head and Tail do condition the evidence we gather by tossing the coin (Figure 2). Iwill transpose this idea to the study of women's presence and absence in the lost corpus of early modern news sheets; I will reframe the problem in terms of Bayesian statistics as follows: women's presence or absence in the lost whole can be viewed as the prior; the result of bootstrapping is the evidence, which is conditioned by the prior. I will then show how Bayesian formula, Bayesian updating, and the beta distribution as conjugate prior, enable us to develop a plausible model representing women's presence or absence in the lost whole of early modern news sheets (Gupta & Nadarajah, 2004). I will also demonstrate how this procedure can be used to compare the data corpus split into two periods: pre and post 1600.

As a conclusion, I will evaluate the insights gathered throughout the experiments described above in light of the current scholarship on early modern women history. I will argue that in handwritten news there was no complete silence about women (Figure 4) but the news world was still dominated by the presence of men. I will also show that women's presence in the news did not change throughout the period. I will discuss the impact of the methodology presented by highlighting how different projects, such as for instance the Euronews Project and the Medici Archive Project, applied it to survey early modern information culture. Finally, I will reflect on the problem of "revolutionary potential" as framed in the Call: can we revolutionize historical research by studying the randomness of sources and archival assets with the help of methods from the AI toolbox ?

Bibliography

Aa, Christopher Z. Mooney / **Robert Duval** (1994): "Bootstrapping: A nonparametric approach to statistical inference", in: *Journal of the American Statistical Association* 89, https://doi.org/10.2307/2290969

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

Baron, Sabrina Alcorn / **Brendan, Dooley** (2005): *The politics of information in early modern Europe*. London: Routledge.

Bauer / **Oswald** (2011): *Zeitungen vor der Zeitung: Die Fuggerzeitungen (1568-1605) und das frühmoderne Nachrichtensystem.* Berlin: Akademie Verlag.

Benson, Pamela. Joseph (2010): *Invention of the Renaissance Woman: The Challenge of Female Independence in the Literature and Thought of Italy and England*. University Park, PA: Penn State Press.

Bradley, Efron (1982): *The jackknife, the bootstrap, and other resampling plans.* Society for Industrial and Applied Mathematics. http://dx.doi.org/10.1137/1.9781611970319.ch5.

Davies, Simon / **Fletcher, Puck** (2014): *News in early modern Europe: Currents and connections*. Leiden: Brill.

Dooley, Brendan / **Kreuze, Wouter** / **Boerio, Davide** / **Allori, L.orenzo** / **Mansutti, Sara** / **Paltrinieri, Carlotta** (2021): *Euronews XML corpus.* 10.5281/zenodo.5524182.

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

Gupta, Arjun. K. / **Nadarajah, Saralees** (2004): *Handbook of beta distribution and its applications*. Boca, Raton: CRC Press.

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

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

Rublack, Ulinka (2001): *The crimes of women in early modern Germany*. Oxford: Oxford University Press.

Wiesner-Hanks, Merry E. (2019) *: Women and gender in early modern Europe.* Cambridge University Press.