Ceci n'est pas de l'archéologie?
A Bayesian approach to understand
consumption dynamics in Ancient Rome

María Coto Sarmiento

Slide 1 *(portada)*

Hi all,

My name is María. I am doing my PhD in Archaeology in Barcelona Supercomputing Center. Actually, I am the archaeological part of the group who apply quantitative method and complex system theories to study archaeological problems.

I would like to thank you to the committee for organizing these student conferences in a lovely place to talk about Bayesian Method applied to Archaeology.

Let's start to explain what we are doing. Basically, I am a small part of an ERC advanced grant project called Epnet. The acronym stands for Economical and Political Networks, and the project is aimed to explore the dynamics of the Roman Empire economy using a complex systems approach. We use methods such as complex network analysis and computer simulations to understand this economy. This large-scale trade network was active for more than 400 years, so you can imagine the current interest on understanding a economy that survived lots of different crisis during several centuries.

Slide 2 *(outline)*

Having said that, one of the challenges we face is how to compare different hypotheses on the rules governing this market. For this reason, today, I am gonna talk about how to use Bayesian inference to perform model selection with archaeological data.

We are going to divide this talk into six parts. In each parts, we will focus on implementing this approach to understand the economic production in the olive oil context of the Roman Empire.

Slide 3 *(Historical context)*

To put in context, Roman empire is considered one of first integrated economic networks of the Mediterranean trade. Roman empire had an important and big infrastructure to develop this commerce creating a complex network system.

This suggests there were an important organization of commercial network. In fact, this commercial networks were held for hundreds of years even enduring economic instabilities. This complex dynamic can be evidenced of two ways: written sources and archaeological records.

However, usually, these written sources are not partial and its information is not clear, being the most part too much specific without a general information. So, we have to complete this information with archaeological records, in particular, with the pottery studies.

Pottery tell us about what Roman people ate and how the commerce was developed, definitely about the consumption. For example, we can find pottery produced by producers from *Baetica*, current Andalusia in other countries such as England, where were supplied to hundreds of roman legions from Hadrian's wall.

Slide 4 *(Research questions)*

So, we started this project using archaeological data to understand the structure of roman market.

The discussion about this roman market structure is one of the most common topics that roman historians and archeologists argue. Some works suggest that an integrated economy as large as this one would need a free market structure to be stable, while other authors think that the free market only can exist with capitalism and that Rome was a very basic exchange market. So, which hypothesis is correct?

In our case, we propose two questions to develop:

Under what extent the roman economy of 2000 years ago was similar to our own economy?

Did they have a free market or a completely different system?

Slide 5 *(Main challenges)*

However, usually in archaeology, we have problems with the data to understand the past. In particular, we are faced to a noisy information in our study, so we need to consider some issues:

1. One of the main problem that we find are the fragmented sources. So, sometimes it is very difficult to built a hypothesis when your data are not exact and you cannot check or compare these data with written sources (in the case what we are working)

2. Another problem is the high level of uncertainty in archaeology: usually you don't have the complete picture, but only a small percentage of the total. We have to work with incomplete material but also we have to work with incomplete data. I mean, sometimes, there are a poor quality of archaeological data.

3. For that, advanced statistical techniques are necessary because we have to know which hypothesis is more correct to explain better the subject of study. Definitely, we have to select which model provides the best fit to the data with the smallest number of assumptions.

Slide 6 *(Case Study: Monte Testaccio)*

Our principal case of study is Monte Testaccio to explain the roman market structure of one of the most important product of the roman society: olive oil. This product was used in different aspects of the Roman daily life such as consumption, lighting and hygiene. Therefore, Olive oil was produced and shipped in containers called amphorae from different places of the Roman Empire.

As summary,

1. Monte Testaccio is an archaeological dump of amphorae in Roma, Italy. This artificial mound was generated by the accumulation of broken amphorae sherds.

2. It is composed of fragments of broken amphorae (name is *testae*) from Baetica mostly (current place from Andalusia), Africa and Gaul.

3. Amphorae of olive oil were thrown because these could not be re-use due to the rests of oil and, therefore, they were deposited in Monte Testaccio.

4. More or less, around of 53 million of amphorae were deposited during 300 years

This example of study case can help us to understand the roman market economy through archaeological records such as amphorae

Slide 7 (Study case: archaeological data)

The majority of amphorae found in Monte Testaccio belong to Dressel 20. Dressel 20 is an amphora was used to transport olive oil.

- The most of them were marked off with several identifications about their provenance: they are stamps and tituli picti. In this case, we only will use stamps.

- A stamp is identified as tria nomina or three names to define a person or figlina (figlina is a pottery factory). It is not clear the meaning of these stamps: on the one hand, it could be a agent identified as olive oil producer; on the other, it could be a agent identified as pottery factory.

- Tituli picti are inscriptions in ink on the amphorae. The most important information is about the consular year when these amphorae were filled.

Slide 8 (Approach)

So, explained this, the main object of our study is analyse these stamps to know this market. How?

- We want to analyse the economic activities from Monte Testaccio through studying of the stamps.

- Specifically, we want to analysee the frequencies of distribution of olive oil in the Mediterranean with the study of the amphorae from Monte Testaccio. We want to measure the frequencies of the distribution of the stamps to know the structure of the market: was this a monopoly or there were thousands of small producers? Can we identify patterns similar to current markets?

- For that, we are going to use stamps as data to measure this frequency because they give us data about the producers, identified with a unique stamp.

Slide 9 *(A method to analyse)*

Now, how we can to create a model to solve our questions with the roman market?

As we said, first, we are going to define a statistical model to each hypothesis and after, we apply Bayesian Model to know which of them explain better the evidence.

- For that, we explore the codes found in the amphorae Dressel 20 with the producer information.

- We are interested to know the frequency distribution of producers. To achieve this goal we treat each code as the identity of a different producer. In this way, if we count how many stamps with the same code appear in Monte Testaccio we are effectively quantifying the size of this producer. For example if we have a stamp, say 'PNN', which has been found 30 times, and another one called 'LCF' found 3 times, it is reasonable to say that PNN identifies an olive oil producer much bigger than LCF.

Slide 10 *(Hypothesis)*

Based on our knowledge with the data, we have defined four hypotheses:

- The demand of olive oil was covered by small-size agents (producers)

- There was a coexistence of small - size producers and major producers

- Concentration of fertile land in a few hands created lots of mid-sized agents, similar to current European agriculture. As we can see with informations about the post-war periods.

- We are going to use Pareto principle. It consists this frequency distribution follows a Power-law where 80 % of the market was dominated by 20% of the producers. This is a typical structure seen in current markets.

Slide 11 (*Quantitative analysis of archaeological data*)

So, our questions can be solved applying Bayesian Method to study the ancient roman market. In this case, we have to explain why we use Bayesian Method in Archeology.

- First at all, Quantitative method provides an excellent scientific basis to the archaeological data. In fact, despite to be a potential tool it was not used until the last 20 years. One of the main reason of its popularity was the advance of the new computational methods such as the ABC and Monte Carlo. We believe that this framework is able to face to the risk of the archaeological data.

- Secondly, to difference of hypothesis testing, Bayesian method allows to know which model is more suitable from the evidence. In fact, it can be used to evaluate the probability of a hypothesis based in observed evidences. Also, we receive updated data constantly due to new archaeological diggs, and this approach is perfect to update our knowledge based on new evidence.

- and Lastly, it is a effective tools in contexts where the level of uncertainty is high as archaeology where exists a poor quality of exact data

- We have translated these hypotheses into 4 different statistical models using statistical methods based on the frequency distribution. Once defined this models, we apply Bayesian inference.

Slide 12 (*Model 1: Poisson pero también intro*)

First, in this picture, we use Poisson distribution to define the frequency of the producers. You can see that the meaning of x-axis is the number of stamps and the y-axis is the density of frequency of them.

Slide 13 (*Model 2: Negative binomial*)

Secondly, we use Negative binomial distribution to add variance to Poisson distribution. We want to add the existence of major producers in this roman market. They coexisted with small-scale producers.

Slide 14 *(Model 3: Lognormal distribution)*

Third, we can know if these small producers survive to the increase and concentration of land in hands of few producers. For that, we use Lognormal distribution. As we said, It was used to calculate the distribution in farm size structure in England over the post war period.

Slide 15 *(Model 4: Pareto distribution)*

Last, we use Pareto distribution because we want to know if Roman Market had a similar structure to current markets.

Slide 16 *(Markov Chain Monte Carlo Simulation)*

Once done, we used for running Markov Chain Monte Carlo similation (MCMC) to generate a posterior distribution. All were made using R statistical package nd R2Jags JAGS.

Through multiple sampling, Monte Carlo searches the best way to get an evidence more suitable, generating a posterior beliefs from our knowledge.

Slide 17 *(DIC)*

Finally, we have chosen Deviance Information Criterion, a useful tool when the posterior distribution are obtained by Monte Carlo. This modelling determines which model is better to know the evidence. In this table, we want to show the DIC value of the four models, being smaller more suitable to quantify the evidence.

Slide 18 *(Results)*

So, our results suggest:

- Results demonstrate that the model 4 (pareto) of the four frequency distributions can fit better to the hypothesis proposed.

- So, the results show a high frequency of different stamps. It means that exists a important presence of the small-size producers that they coexisted with the large producers. In fact, this commerce was not only focused in large producers.

Slide 19 *(discussion)*

As discussion, we can speak about...

- The possible existence of a free market of olive oil production in the Roman Empire.

- Also, we suggest the possibility of a low variability with respect to the current market

- From a methodological perspective the Bayesian method is a useful tool for archaeology, where we always have different hypotheses to explain what we see in the fragmented and noisy archaeological record..

Slide 20 *(final)*

Thank you very much