

# The Bayesian Approach to missing rich in surveys

Syllabus  
Michel LUBRANO

## Course's objective

Surveys document the evolution of the income distribution. They however suffer from under-sampling and underreporting for high incomes. One of the methods proposed in the literature to solve this question was to introduce a Pareto tail to model high incomes. The crucial question is to know at which threshold to start the Pareto tail. The aim of this lecture is to illustrate how a Bayesian approach can solve many of the problems raised by this question. The first part of the course will be devoted to an introduction to Bayesian inference in econometrics and to the numerical methods useful for integrating posterior distributions, in particular Monte Carlo Markov Chains. Among the Pareto family, the Pareto II combines a relative simplicity and nice properties compared to the simple Pareto I. This will occupy the second part of the course. We finally have to integrate this Pareto tail into a complete model that will allow us to make inference on the threshold where the Pareto tail begins. The method has then to be applied to a set of countries covered by the EU-SILC data set. As a conclusion, a minor part of the literature has proposed to use power law to model the right tail of the income distribution in order to cope for missing poor.

## Outline

The course lasts 12h divided in six sessions of two hours each.

Lecture 1: General introduction and a first look at Bayesian statistics

Lecture 2: The Pareto II distribution and its main properties

Lecture 3: Models for combining a Pareto tail

Lecture 4: Application of MCMC method for inference on the complete model

Lecture 5: Empirical application using SILC data

Lecture 6: The missing poor and power law imputation

## Course materials

Lectures' notes and other materials will be available at  
<https://perso.amse-aixmarseille.fr/lubrano/>

## Grading

Students will be graded on the basis of a final exam consisting in an oral examination with simple questions covering all the lectures.

## References

Concerning Bayesian inference, you can access my lecture notes given in China that will be made available on my web page. You can also have a look at the book:

Bauwens L, M. Lubrano and J.F. Richard (1999) Bayesian Inference in Dynamic Econometric Models. Oxford University Press.

The following articles will be usefull as a complementary material:

Arnold, B. C. (2008). Pareto and generalized Pareto distributions. In Chotikapanich, D., editor, Modeling Income Distribuions and Lorenz Curves, volume 5 of Economic Studies in Equality, Social Exclusion and Well-Being, chapter 7, pages 119–145. Springer, New-York.

Atkinson, A. B. (2017). Pareto and the upper tail of the income distribution in the UK: 1799 to the present. *Economica*, 84(334):129–156.

Bartels, C. and Metzing, M. (2019). An integrated approach for a top-corrected income distribution. *Journal of Economic Inequality*, 17:125–143.

Cabras, S. and Castellanos, M. E. (2011). A Bayesian approach for estimating extreme quantiles under a semiparametric mixture model. *ASTIN Bulletin: The Journal of the IAA*, 41(1):87–106.

Charpentier, A. and Flachaire, E. (2022). Pareto models for top incomes and wealth. *The Journal of Economic Inequality*, 20(1):1–25.

Jenkins, S. P. (2017). Pareto models, top incomes and recent trends in UK income inequality. *Economica*, 84(334): 261–289.

Scarrott, C. and MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Statistical Journal*, 10:33–60.

And for the other side of the income distribution

Masseran, N., Yee, L.H., Safari, M.A.M., Ibrahim, K. (2019). Power law behavior and tail modeling on low income distribution. *Mathematics and Statistics*, 7 (3), 70–77,  
<https://doi.org/10.13189/ms.2019.070303>