BRYANT J. and ZHANG J. L., 2018, Bayesian demographic estimation and forecasting, Boca Raton, CRC Press, 292 pages

Daniel Courgeau

Bryant J. and Zhang J. L., 2019, *Bayesian demographic estimation and forecasting*, Boca Raton, CRC Press, 292 pages.

Many books have argued the usefulness, or even the necessity, of Bayesian methods in the social sciences and statistics. The originality of this volume is its presentation of the general application to demography and population projections. Nonetheless, numerous articles in statistics have already presented and discussed this topic, which is increasingly important in the field. Arbel and Costemalle (2016) in particular use a Bayesian approach to reconcile two different sources for estimates of immigration flows in France.(1)

To begin, it may be useful to detail what distinguishes the frequentist or objectivist approach commonly used in demography from the Bayesian or epistemic approach, which remains little used. Few articles published to date in *Population*, in particular, feature Bayesian methods.

The notion of probability originated with Pascal's treatise of 1654,(2) while demography was launched by Graunt a few years later, in 1662.(3) Graunt took an objectivist approach to probabilities (defined in terms of the object of study). It was not until the 18th century, in 1763, that Bayes proposed an epistemic notion of probability (defined in terms of the knowledge that humanity can have of objects).(4) He was followed by Laplace, who used it specifically for demographic problems, such as a comparison of the births of boys and girls in 1778.(5)

How do these two principal concepts differ? The following summarizes the more complete and detailed presentation found in my book *Probability and Social Science* (2012).(6) The objectivist approach assumes that the probability of an event exists independently of the statistician, who tries to estimate it through successive experiments. As the number of trials tends to infinity, the ratio of the cases where the event occurs to the total number of observations will tend towards this probability. But the very hypothesis that this probability exists cannot be clearly demonstrated. Bruno de Finetti, a great defender of Bayesian approaches, said clearly in 1974 that probability does not exist objectively—that is, independently of the human mind.(7)

(1) Arbel J., Costemalle V., 2016, Estimation des flux d'immigration: réconciliation de deux sources par une approche bayésienne, *Économie et Statistique*, 483–485, 121–149.

(2) Pascal B., 1654, *Traité du triangle arithmétique, avec quelques autres traités sur le même sujet*, Paris, Guillaume Desprez.

(3) Graunt J., 1662, *Natural and political observations mentioned in a following index, and made upon the bills of mortality*, London, printed by Tho. Roycroft for John Martin, James Allestry, and Tho. Dicas.

(4) Bayes T. R., 1763, An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society of London*, 53, 370–418.

(5) Laplace P.-S., 1778, Mémoire sur les probabilités, *Mémoires de l'Académie royale des sciences de Paris*, 1781, 227–332.

(6) Courgeau D., 2012, *Probability and social science*, Dordrecht, Heidelberg, London, and New York, Springer.

(7) de Finetti B., 1974, *Theory of probability*, 2 vols., London and New York, John Wiley & Sons.

 DOI: 10.3917/popu.2004.0618

The epistemic approach, in contrast, focuses on the knowledge that we can have of a phenomenon. An experiment, a survey sample, or a more or less exhaustive demographic census provides us with new information on the phenomenon we are studying. The epistemic statistician can then take advantage of this information to improve their a priori opinion of its probability, using Bayes' theorem to calculate its a posteriori probability. Of course, this estimate depends on the chosen a priori probability, but when the choice is made with appropriate care, the result will be considerably improved relative to the results of an approach based on the concept of objective probability.

When it comes to using these concepts to make a decision, the two approaches differ even more. When an objectivist provides a 95% confidence interval for an estimate, they can only say that if they were to draw a large number of samples of the same size, 95% of the obtained confidence intervals would be expected to contain the estimated parameter. Clearly, this complex definition does not fit with what might be expected of it. The Bayesian, in contrast, starting from initial hypotheses, can clearly state that a Bayesian 95% confidence interval, or credibility interval, in fact indicates the probability that the estimated parameter is found within it.

It may thus be wondered why the Bayesian approach, which seems better suited to the social sciences, has taken so long to gain acceptance among researchers in these domains. One important reason is the complexity of the calculations involved, which ordinary computers have only recently become powerful enough to perform. To understand this, one need only read Laplace's excellent 1778 article, which presents the complex calculations and approximations required to solve a simple problem involving comparing the births of girls and boys. A second reason is the desire for an objective demography, which is expected not to appeal to personal judgements. This notion of objective demography is increasingly challenged by analyses involving increasingly complex interactions between events and projections for subpopulations based on small samples.

John Bryant and J. L. Zhang's book takes a pedagogical approach, guiding the reader progressively towards the use of Bayesian methods. They first provide the main elements of these methods using demographic examples: exchangeability, informative and non-informative a priori distributions, hierarchies, calculating the a posteriori distribution, various possible validity tests, etc.

They then analyse different demographic situations in great detail using existing data. Infant mortality in Swedish counties from 1995 to 2015 and projections to 2025; life expectancy in Portugal from 1990 to 2015 and projections to 2035; health expenditures in the Netherlands from 2003 to 2011 and projections to 2020; internal migration in Iceland from 2000 to 2014 and projections to 2020; fertility of Cambodian provinces estimated using the 2008 census and the 2010 survey; change over time in the New Zealand population by region from 2008 to 2016, and estimates of internal migration; the growth of the Chinese population from 1990 to 2015. They thereby demonstrate that estimating past

populations and projecting future populations are the same problem, which is best solved using Bayesian methods.

Unfortunately, the book lacks a more precise comparison of estimates and projections produced with the two approaches. For example, it would have been very useful to see how much Bayesian projections improve on those derived using frequentist methods. To my knowledge, only Bijak (2011) has compared the results of such forecasts not only with observations, but also with frequentist forecasts.[8] Bijak's results clearly demonstrate the superiority of Bayesian methods. The frequency with which empirical observations fall within the confidence intervals for the Bayesian projection is always much higher than the corresponding figure for the frequentist projection.

Despite this oversight, I strongly recommend this book. In it, demographers and statisticians will find population estimates and projections that are considerably better justified and much more accurate than those produced using classical methods. Moreover, in 2014, the UN adopted a Bayesian methodology for its official demographic projections by sex and age. Should that not encourage demographers to use Bayesian estimates more frequently?

Daniel Courgeau

(8) Bijak J., 2011, *Forecasting international migration* (Springer Series on Demographic Methods and Population Analysis, 24), Dordrecht, Heidelberg, London, and New York, Springer.