Bayesian Demography

By *Jakub Bijak*[1] *and John Bryant*[2]

Keywords: *Applied statistics, Demographic uncertainty, Population processes, Statistical methods*

Abstract: Demography – the scientific study of populations – has had a long relationship with statistical methods. In particular, the last 30 years have witnessed an increasing number of applications of Bayesian statistics. At present, the main areas of interest of Bayesian demography include population forecasting, dealing with inadequate data, and small area estimation, with a few studies on demographic impacts. However, the current gaps in demographic literature, including a lack of theoretical foundations, challenges related to the management of different sources of uncertainty, and the use of new sources of data are also well suited for applications of Bayesian methods. This is where we predict that the next developments will be concentrated, especially if the current challenges, such as those related to computations, can be overcome.

1 Introduction

Demography can be seen as the "study of human populations with respect to their size and structure and of the components of population change" (after: **stat06091**). Its key components include fertility, mortality, migration, as well as auxiliary processes, such as marriage and union dynamics, or health transitions (→ **stat00041**). Formal mathematical description in demography has a long history (→ **stat00095**), dating at least to John Graunt's (→ **stat04443**) *Bills of Mortality*, a 17th-century precursor to contemporary life table analysis (Graunt 1662, see also **stat00010**). Almost as long-lived are the two-way links of demography with probability theory and statistical thinking, such as the 1778 work of Pierre Simon de Laplace (→ **stat01315,** → **stat00207.pub2**) on sex ratios at birth for Paris and London (Courgeau 2012).

Since the 1990s, Bayesian statistical demography has emerged as a distinct methodological thread within the discipline of demography. The current article outlines the history of this dynamically-developing area of population studies, presents some key applications, and makes some forecasts about future developments.

[1] University of Southampton, Southampton, UK; [2] Bayesian Demography Limited, Christchurch, New Zealand

DOI: 10.1002/9781118445112/stat08273

2 From *Political Arithmetick* to Statistical Demography

Since its inception, demography has had strong links with statistical and other mathematical methods, developed for studying human and non-human populations (→ **stat00139**, **stat07462**). The relationships with probability theory (→ **stat03979**) were particularly dynamic. Early methods for analysing population questions (then referred to as 'Political Arithmetick') were probabilistic, only to be superseded by deterministic approaches in the 19th and first half of the 20th century, with statistical methods making a comeback in the late 20th century, aided by the proliferation of survey data (see Courgeau 2012 for a discussion). Contemporary statistical demography can be seen as a branch of applied social statistics (→ **stat00125.pub2**), although one with a long tradition and distinct features: precise description of processes, and strong empirical foundations (Xie 2000). A review of traditional, largely likelihood-based statistical methods (or frequentist, → **stat07913**, see also **stat05859**) used in demography, is provided by Alho and Spencer (2006).

Demographic processes, as any social processes, are characterised by high levels of uncertainty (see **stat02675**), although the fact that many of the people who will be alive in 25 or even 50 years' time are already alive today, and that important demographic variables such as country of birth or age are fixed or evolve deterministically, means that demographic processes are somewhat more predictable than other social and economic phenomena. Population dynamics are described by a cohort-component mechanism of population renewal (→ **stat06091**, for matrix population models in general, see **stat07481**). Typically, uncertainty is lowest for mortality, which has very high biological component (→ **stat06105**), and highest for migration, which is responsive to a vast array of drivers and involves high levels of human agency (→ **stat00098**), with other components falling in-between (see **stat00041**).

The ability of Bayesian methods (→ **stat00207.pub2**) to cope with complex models and integrate different sources of uncertainty, have made them increasingly attractive to statistical demographers. Contemporary statistical demography is becoming increasingly Bayesian, aided by the developments in numerical methods (such as **stat00212**, including **stat00211.pub2**). A recent review of the contemporary literature in Bayesian demography, focusing on forecasts, limited data, and complex models, has been provided in Bijak and Bryant (2016), and a general framework for Bayesian demographic estimation and forecasting is presented in Bryant and Zhang (2018). Selected examples are summarised next.

3 Contemporary Examples

3.1 Population Forecasting: World Population Prospects

Population projections (→ **stat00109**) are demography's most prominent output, and the World Population Prospects, a set of projections for 235 countries and areas published by the United Nations Population Division, are the world's most prominent population projections. Since 2014, the UN Population Division has produced these projections using Bayesian methods (Raftery et al. 2014), with the underlying open source software published in a series of R packages (Ševčíková and Raftery 2016).

The probabilistic UN method derives parametric decline or increase curves for fertility and life expectancy using Bayesian hierarchical models (→**stat00232**, **stat08140**). For most countries, there are only a handful of data points available for estimation. If each country's curve was estimated using only data for that country, the estimates would be unstable. The Bayesian hierarchical model avoids this instability by

DOI: 10.1002/9781118445112/stat08273

sharing information (‘borrowing strength’) across countries, so that each country’s curve is a compromise between data for that particular country, and the typical profile across all countries.

The new methods provide much more comprehensive, internally consistent, and interpretable measures of uncertainty, in the form of quantiles and prediction intervals, than the fertility variants previously used by the UN (Fosdick and Raftery 2014). At each step in the development process, held-back data were used to test the accuracy of the point estimates and uncertainty measures, including the calibration of errors (predictive intervals). This approach has now become standard in demographic forecasting of whole populations, as well as individual components of change, where elements of Bayesian time series analysis are also used (→ **stat00219.pub2**, **stat06763**; see Bijak and Bryant 2016).

3.2 Incomplete and Noisy Data

Much of the data that demographers work with is incomplete and subject to measurement error (→ **stat04144**). Problems occur in custom-designed data sources such as censuses and surveys, and also in administrative databases (→ **stat05227.pub2**) and big data (→ **stat07979**). When asked about births, for instance, mothers may only include children who are still alive. Population registers that in principle contain only residents of the country, may in practice include people who have emigrated but failed to notify the authorities. Demographers have long emphasised the importance of diagnosing, and correcting for, flaws in their input data, and have developed many techniques for doing so (Moultrie et al. 2013). Bayesian demographers are reformulating many of these methods in statistical terms.

Bayesian statistics is well-suited to dealing with incomplete and noisy data. From a Bayesian perspective, missing data is just another type of unknown quantity, to be modelled and estimated in much the same way as any other. Informative prior distributions (→ **stat00243.pub2**) can be used to capture substantive knowledge about errors in datasets. When a statistical model is already hierarchical, as it generally is in Bayesian statistics, adding one more layer representing measurement processes is natural and convenient. Markov chain Monte Carlo methods (→ **stat07189**) provide a powerful set of tools for fitting such models.

Recently, Bayesian hierarchical models have been used, for instance, to estimate maternal mortality rates with uncertainty measures (→ **stat06103**) for 183 countries in a way that carefully accounts for likely errors in the input data (Alkema et al. 2016). In countries with little or no good data, the final estimates lean on results from the theoretical model; in countries with good data, they are close to country-specific observations, incorporating a substantive knowledge about reporting systems (→ **stat00019**).

Similarly, Schmertmann and Hauer (2019) build on a decades-old technique where the total fertility rate (→ **stat03412**) is estimated from the ratio between children and women in a population. By setting up a probabilistic model that brings in information on age patterns for fertility and mortality, and that allows for random variation in numbers of births, they are able to produce accurate estimates of total fertility rates, with realistic uncertainty measures. These methods apply to other mammalian species as well as humans.

Subject matter experts who know about data collection can also offer insights into likely errors in data. The weaker the data are, the more valuable these insights become. Bayesian methods allow demographers to include them in their models in a transparent and rigorous way, as informative priors. Techniques exist for eliciting these priors so that they reflect the beliefs of the experts (→ **stat00231.pub2**, **stat03871**). The IMEM (Integrated Model of European Migration) project, which produced a migration flow matrix for 31 European countries based on inconsistent and incomplete data, exemplifies this approach (Raymer et al 2013). Similarly, Azose and Raftery (2019) use migrant stocks data to reconstruct global flow matrices.

DOI: 10.1002/9781118445112/stat08273

3.3 Disaggregated Estimates

Although most of the academic literature on demographic estimates and projections has focused on the national level, most potential applications for demographic estimates and projections require numbers at the local level. When planning a school, a highway, or a supermarket, for instance, local-level detail is essential. Moreover, users typically require the same level of age-sex detail as they get at the national level, with other variables such as ethnicity or education, also being in high demand.

Bayesian hierarchical models are well-suited to disaggregated estimates and projections. With data classified by dimensions such as age, sex, region, and time, demographers have traditionally calculated demographic rates directly, by dividing the number of events by the population at risk (→ **stat04534**). Within a statistical model, the rates can be also predicted based on age, sex, region, time, and other covariates such as income. An estimate from the Bayesian hierarchical model is a compromise between the direct estimate, which plays the role of the data, and the prediction, which plays the role of the prior. The more data are available for a particular cell within the classification, the further the estimate is pulled towards the direct one; the less data are available, the more the estimate relies on predictions. This is an effective way to smooth, allowing sensitivity to local variation when there is enough data to support it, yet giving reasonable answers elsewhere.

Amongst recent examples, Alexander et al. (2017), present a Bayesian hierarchical model for estimating subnational mortality rates by age, sex, and small area. The model improves on the direct estimates by pulling the results towards overall means. New Zealand's national statistical office, has, since 2015, used Bayesian hierarchical models to construct its life tables (→ **stat06037**; see Statistics New Zealand 2015).

Zhang and Bryant (forthcoming) present an extreme example of modelling with sparse data by constructing estimates and forecasts of origin-destination mortality rates, by sex and single year of age, for eight regions in Iceland. Excluding structural zeros, 66% of cells in the migration data have counts of zero. The model is nevertheless able to extract a strong enough signal to produce sensible migration rates, for past and future years, and held-back data suggest that the model is reasonably well calibrated, indicating that demographers may need to revise their rules of thumb on how much disaggregation is too much.

Besides the key areas of application listed above, Bayesian methods are also successfully used in a vast array of specific applications, from event history analysis (→ **stat06013**, see also → **stat06060**), to actuarial analysis focusing on life insurance (→ **stat06074**), to disease mapping (→ **stat06102.pub2**), and paleodemography (→ **stat03399**) – see Bijak and Bryant (2016) for specific examples.

4 Knowledge Gaps and Current Developments

One important gap of contemporary demography is its lack of theoretical foundations: historically, the discipline has been largely data-driven, with theoretical developments lagging behind (Xie 2000, Burch 2018). Filling this gap can be achieved in different ways, not least, as suggested by Burch (2018), by employing formal analytical and computational models. Such models, including microsimulations and agent-based models (→ **stat07981**) would enable making greater use of demographic and wider social theory, while maintaining empirical rigour. Bayesian methods can help with the design, construction and analysis of such models, by providing tools for uncertainty quantification (→ **stat07205**).

Another gap is linked to description and management of uncertainty of population forecasts across a spectrum of temporal horizons, depending on the particular user or policy needs. The perspectives can range from very short-term, akin to early warnings in finance (→ **stat04323**), to long-range, secular trends, enabling making high-level statements about population developments a few generations ahead (→

DOI: 10.1002/9781118445112/stat08273

stat00109), for example to provide input to climate models or similar endeavours. The challenges of describing uncertainty are particularly acute in the case of migration, with its high volatility in the short term, and long-range uncertainty being difficult to assess (Bijak 2010; although see Azose et al. 2016 for a recent attempt). In such cases, additional knowledge, for example elicited from experts, can be incorporated in the Bayesian models via informative prior distributions (→ **stat05938**).

Another area of demography and population statistics where Bayesian methods have a large and still underutilised potential, especially in official statistical agencies, is related to harnessing the opportunities offered by new data sources (such as **stat07979**), as well as new ways of dealing with existing data, including administrative sources and vital registers (→ **stat05227.pub2**, **stat00019**). In official statistics, this would require a shift of perspective, from a traditional, largely design-based approach to population statistics, to model-based inference, the latter typically associated with the Bayesian paradigm (→ **stat01616**). One especially promising example are methods for combining different data sources (→ **stat03679**) which in the Bayesian paradigm come with inevitable statements about measurement uncertainty (→ **stat04144**). Further development of such tools will help address the methodological challenges of the emerging field of data science (→ **stat08150**).

Computational challenges remain one of the main barriers to wider use of Bayesian methods in demography. Bayesian methods are almost always slower than traditional frequentist methods and they require applied demographers to learn about algorithms and diagnostics. The explosive growth in the number of packages implementing Bayesian methods, and the development of modern and efficient modelling languages, such as Stan (Carpenter et al. 2017) are, however, increasingly leading to faster and more user-friendly software.

Besides computation, there are also other impediments to a wider uptake of Bayesian methods in demography and official population statistics. They include insufficient statistical training amongst the users and producers of demographic research, and challenges related to communicating uncertainty to the users and other stakeholders (Bijak and Bryant 2016). Some of the possible solutions come from applying the statistical decision theory in this context (→ **stat00215**), and from ensuring that the robustness of the results is thoroughly checked, especially with respect to the choice of priors, but also other elements of inference, such as the likelihood or model choice (→ **stat00210**). These challenges notwithstanding, Bayesian approaches help demography maintain formal statistical rigour, for which it is renowned, while allowing the discipline to address modern, increasingly complex research challenges which are beyond the reach of traditional statistical methods.

Acknowledgments

Jakub Bijak gratefully acknowledges the European Research Council grant ERC-CoG-2016-725232 *Bayesian Agent-Based Population Studies* and the ESRC Centre for Population Change (ES/R009139/1).

Related Articles

Demography; Demography including Marriage, Fertility, Migration and Population Projections; Population Projection; Social Statistics; Stochastic Demography; Bayesian Inference; Markov Chain Monte Carlo Algorithms; Bayesian Forecasting; Hierarchical Models - Theory; Inference, Design-Based vs. Model-Based; Uncertainty; Data Fusion; Probability Theory; Measurement Error and Uncertainty; Prior Distribution; Demographic Stochastic Models; Data Science

DOI: 10.1002/9781118445112/stat08273

References

[1] Alexander M, Zagheni E and Barbieri M (2017). A flexible Bayesian model for estimating subnational mortality. *Demography*, 54(6), 2025–2041.

[2] Alho JM and Spencer BD (2005) *Statistical Demography and Forecasting*. Berlin–Heidelberg: Springer.

[3] Alkema L, Chou D, Hogan D, et al. (2016) Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *Lancet*, 387(10017), 462–474.

[4] Azose JJ and Raftery AE (2019) Estimation of emigration, return migration, and transit migration between all pairs of countries. *PNAS*, 116(1), 116–122.

[5] Azose JJ, Ševčíková H and Raftery AE (2016). Population projections with migration uncertainty. *PNAS*, 113(23), 6460–6465.

[6] Bijak J (2010) *Forecasting International Migration in Europe: A Bayesian View*. Dordrecht: Springer.

[7] Bijak J and Bryant J (2016) Bayesian demography 250 years after Bayes. *Population Studies*, 70(1), 1–19.

[8] Bryant J and Zhang JL (2018) *Bayesian Demographic Estimation and Forecasting*. Chapman and Hall/CRC.

[9] Burch TK (2018) *Model-Based Demography: Essays on Integrating Data, Technique and Theory*. Cham: Springer.

[10] Carpenter B, Gelman A, Hoffman MD,... Riddell A (2017) Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32.

[11] Courgeau D (2012) *Probability and social science. Methodological relationships between the two approaches*. Dordrecht: Springer.

[12] Fosdick BK and Raftery AE (2014) Regional Probabilistic Fertility Forecasting by Modeling Between-Country Correlations. *Demographic Research*, 30: 1011–1034.

[13] Graunt J (1662) *Natural and political observations upon the bills of mortality*. Roycroft: London.

[14] Moultrie TA, Dorrington RE, Hill AG, Hill K, Timæus IM and Zaba B (2013) Tools for demographic estimation. International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/

[15] Raftery AE, Alkema L and Gerland P (2014) Bayesian Population Projections for the United Nations. *Statistical Science*, 29(1), 58–68.

[16] Raymer J, Wiśniowski A, Forster JJ, Smith PW and Bijak J (2013) Integrated modeling of European migration. *JASA*, 108(503), 801–819.

[17] Schmertmann CP and Hauer ME (2019) Bayesian estimation of total fertility from a population's age–sex structure, *Statistical Modelling*, 19(3), 225–247.

[18] Ševčíková H and Raftery AE (2016) bayesPop: Probabilistic Population Projections. *Journal of Statistical Software*, 75(5), 1–29.

[19] Statistics New Zealand (2015). New Zealand period life tables: Methodology for 2012–14. www.stats.govt.nz

[20] Xie Y (2000) Demography: Past, Present and Future. *JASA*, 95 (450), 670–673.

[21] Zhang JL and Bryant J (forthcoming) Bayesian disaggregated forecasts: Internal migration in Iceland. In: S Mazzuco and N Keilman (eds.) *New Perspectives on Population Forecasting*. Cham: Springer.

DOI: 10.1002/9781118445112/stat08273