Estimation and projections software for general use

Ludi Simpson
University of Manchester
ludi.simpson@manchester.ac.uk
Paper to IUSSP conference 2009, Marrakech.
Submitted to session 1901.

Projection of population is a sophisticated sub-discipline of demography providing mathematics to estimate and combine components of population change in a dynamic picture of changing population structure. Mathematical treatment has been developed for variously multi-regional, probabilistic, and socially disaggregated systems, and has stimulated many computer programs to support and implement population projections. Many projection models are implemented for specific applied projects; their computer implementations rarely outlive those projects and are rarely documented systematically. Some computer programs for projections have been made available for use beyond their authors, in an effort to make more efficient the implementation of demographic advances for a range of users.

There are tensions between on the one hand the technically demanding mathematics for consistent and accurate implementation of demographic components, in a model that recognises complex multi-dimensional reality and the uncertainty of the future, and on the other hand the data available for a population projection and the capability of non-specialists to use the software.

The paper reviews the solutions to these tensions, which diversely cater for different research environments and different purposes. The paper then details one implementation in the Excel platform which has been taken up by local and national government (www.ccsr.ac.uk/popgroup). This software was developed in the UK, initially to respond to the needs of local government, to replicate sub-national government forecasts, to test the impact on population of plans for future house-building, and to understand the local dynamics of ethnically diverse populations. It was the interest in both large areas for which many data were available, and in small areas and ethnic groups for which few and non-standard data were available, which led to a flexible approach to data entry and to assumptions where data were not available.

Excel platform

Users' own data will be held in diverse formats, but the widespread use of Microsoft Office products means that users will have their own knowledge to convert from their particular data format into Excel. Similarly, transformation of the detailed projection results into particular reporting software does not require new skills, if Excel is familiar to the user. Thus, while the software has extensive data entry and data analysis routines prepared in VBA, the use of the Excel platform means that these routines are helpful tools but are not frustrating limitations to the experienced user.

Mathematical approach to integrate rates and events

The mathematical approach in the forecast demands minimally an initial population and a set of default demographic rates. The projection framework allows a number of 'groups' which are often geographical areas but may also be social or ethnic groups. For any annual period of the forecast demographic rates are optional and may take the form of age-specific rates, differentials from the default rates, or overall indicators such as the Total Fertility Rate. Counts of events may also be provided for any period as totals or disaggregated by age and sex. An existing population estimate may also be supplied, again as a total or disaggregated by age and sex, as may a constraint for expected housing or employment.

The software uses the rates to compute an initial set of components and population with full single year of age detail, for the next period. As a second stage, this initial population (and the rates and components) are adjusted to be consistent with any counts of events supplied for the period. Finally, if the user has provided any constraints of existing population or housing or employment estimates, these are used to further adjust the demographic rates.

The integration of rates with counts of events and existing estimates has strong implications. It allows the system to provide the user with estimated population as well as forecasts. It also estimates demographic rates for periods in the past when only counts of events had been provided.

Three example scenarios

To illustrate the versatility of this approach, three projection scenarios will be followed. In the first scenario, detailed data for each component of change are known, from a national statistics agency's previous work. These form the basis for replicating other forecasts in detail, in order to test the impact of alternative assumptions. This is often the scenario in developed countries where national and sub-national demographic projections are the subject of discussion, feedback and improvement.

In the second scenario, counts of births, deaths are available since a census, without demographic rates. This scenario is common for any areas which have not been subject to demographic study, perhaps because the areas are relatively small or because the areas are not reported in regular official statistics. With suitable model schedules of fertility and mortality and migration provided by the user, the software uses the birth and death counts to estimate fertility and mortality indicators for the past periods. These, when compared to the model schedules, provide differentials for a projection.

The second scenario is extended when population estimates are available for one or more occasions after the base population, for example from a second census or based on administrative records. These are sufficient for the software to estimate different levels and rates of migration for each area, which can then also be used in a projection to the future. This scenario is a particular example of using the projection software in a 'training phase' to learn from past events, prior to the projection. A past demographic time series is an additional output.

The third scenario incorporates targets of future housing. In practice it may be an extension of either of the two scenarios above. When housing developments are

regulated by legislation, the planning authorities provide permission for building conditional on provision for the impact of the new population which new housing may attract, for example requiring new roads or new school buildings. There are various levers with which to make a population projection consistent with targets of future house-building. For example, if the target housing levels are greater than implied by an uncontrolled population projection, the target may be plausibly met by higher vacancy rates (more empty housing), or by lower average household size (or equivalently, higher headship rates), or by greater in-migration, or by less out-migration, or by a combination of these. At present, the software allows the user to manually alter any of these levers, and automates the adjustment of migration flows (in any combination chosen by the user) to meet the target.