Reprodeuce of the Paper A General Age-Specific Mortality Model with An Example Indexed by Child or Both Child and Adult Mortality*

Luyuan Hu

2020/12/10

Abstract

Most African countries and nearly one third of all countries need to use mortality models to infer a complete age table of mortality rates for population estimates, projections/projections and other demographic and epidemiological tasks. Models that link child mortality to mortality at other ages are important because almost all countries have measures of child mortality. In this article, our goal is to absorb the essence of "A General Age-Specific Mortality Model with An Example Indexed by Child or Both Child and Adult Mortality" (Clark 2016) by replicating it. The article of Clark (2016) defines a generic parameterized mortality component model (SVD-COMP) using singular value decomposition and calibrates the relationship between child or child/adult mortality and mortality rates at other ages in the mortality table observed in the human mortality database. It also verifies the model by cross-validation and compares the predictive performance of the model with that of the log-quaternion model, which indicates the efficiency of the proposed method.

keywords: Mortality; SVD; HMD; SVD-Comp

1 Introduction

A complete age-disaggregated mortality table is an essential input for the various formal demographic and epidemiological approaches. A key example is the biennial World Population Outlook prepared by the United Nations Population Division (United Nations, Department of Economic and Social Affairs, Population Division 2015). These indicators are generally considered to be reference demographic indicators and are widely used by other national and international agencies as inputs to estimation and modelling efforts.The data contain estimates of mortality, fertility and population size by sex and age from 1950 to the present, as well as projections for the same number of countries in the world by 2100.Therefore, each WPP update must include a complete timeline of deaths at a particular age between 1950 and 2100.

The paper focuses on mortality, finding that Some countries in the developing world, particularly in Africa, where there is no information on child or adult mortality. there are 50 countries in the world (with a total population of nearly 1 billion) with no information on adult mortality, most of them in Africa – 33 countries with a total population of 666 million. Li (2015) shows detailed household surveys on fertility and maternal and child health, and there is essentially at least some up-to-date information on child mortality around the world.

In the absence of a full mortality record in many poor countries, we consider using the mortality models to make up for this loss.

*Code and data are available at: https://github.com/LuyuanHu/Stat304-PS5; https://arxiv.org/abs/1612.01408; https://github.com/sinafala/svd-comp.

The standard approach for producing complete age tables for countries and areas with inadequate data is to take advantage of the fact that they do have information on child mortality.Typically, the model life table is used to infer a full death schedule from ${}_5q_0$.The United Nations Population Division has also (made extensive use of the traditional Coale and Demeny (1966) model life table} and updated {Li (2015)} with incomplete mortality data for many countries and regions of the world), and the Institute for Health Indicators and Assessment (Health) has done the same using the variant Modifed logarithm (Mod-Logit) model Murray et al. (2003).

Life table of commonly used model system – zone model Coale and Demeny (1966) population life table and stability and developing country life table United Nations, Department of Economic and Social Affairs, Population Division (1982), modify the Logit life table system (Mod - Logit) (Murray et al. (2003)). All of these are empirical models in that they summarize observed mortality rates and use that summary to produce mortality prediction tables that are consistent with observed mortality rates.They are both regional and continuous.Regional models identify and replicate commonly observed mortality patterns associated with geographic regions and allow mortality rates to vary within specific patterns within each region.In contrast, the pattern of mortality produced by the continuous model varied steadily.Both methods are essentially two-parameter models.The regional model first identifies a discrete region and then effectively adjusts the level of mortality in a particular region by using continuously changing life expectancy within each region.The continuum model has two parameters that change continuously, such as life expectancy, child mortality or adult mortality.

Murray et al. (2003) enumerate three characteristics required of mortality models: 1) simplicity and ease of use, 2) comprehensive representation of the true variability in sex-age-specific mortality observed in real populations, and 3)validity that is well quantified by comparing age schedules of mortality predicted by the model to corresponding observed life tables. This paper added 1) generality with respect to the underlying model structure, 2) flexibility in terms of input parameters, and 3) an ability to handle a wide range of age groups, including very fine-grained, without having to fundamentally alter the structure of the model.

Clark (2016) defines and describes a new SVD component-based mortality modeling framework that satisfies all of those requirements. The SVD-component framework provides a general, flexible way to model any demographic age schedule as a function of covariates or predictors that are related to age-specific variation in the age schedule. Here the SVD-component framework is demonstrated by creating a mortality model that predicts single-year of age mortality schedules using either ${}_5q_0$ or both (${}_5q_0$, ${}_{45}q_{15}$) as predictors, similar to both the Mod-Logit and Log-Quad models. The resulting model can be used to produce single year of age mortality schedules from ${}_5q_0$ alone that are consistent with observed mortality schedules, and this could be useful for those like the UN Population Division who must manipulate full age schedules of mortality but only have observed values for ${}_5q_0$. The resulting SVD-component model performs better than the current state of the art two-parameter model (Log-Quad), provides predictions by single-year of age, and is easily extensible to include additional predictors beyond child and adult mortality.

2 Data

2.1 Human Mortality Database Life Tables - HMD

The Human Mortality Database (HMD) University of California, Berkeley and Max Planck Institute for Demographic Research (n.d.) contains rigorously cleaned, checked and validated information on deaths and exposure from a number of developed countries where death registration and census data are virtually complete.

The data are aggregated and presented in a wide variety of formats. The objective of this analysis is to capture and characterize as much variability in age-specific mortality as possible, and consequently I chose to use the 1×1 HMD life tables for each sex. Those provide all columns of a standard life table for single calendar years by single year of age from $0 \rightarrow 110+$. Each country provides data for different historical periods, and some countries are subdivided into more specific subpopulations. In the latter situation a 'national population' life table is typically provided that aggregates across the subgroups. Both the national and subgroup populations are included in this analysis to maximize the variability in age-specific mortality schedules in the overall dataset. A few of the 1×1 life tables from the HMD contain problems: The excluded tables have been updated to reflect the updates in HMD through August 21, 2018.

3 Mortality Model

Traditional model life tables (e.g. United Nations, Department of Economic and Social Affairs, Population Division 1955, @ledermann1969nouvelles, @coale1966, @united1982model, @murray2003, @wilmoth2012flexible) take an inductive, empirically-driven approach to identify and parsimoniously express the regularity of mortality with age based on observed relationships in large collections of high quality life tables. An alternative, sometimes deductive approach, can be found in the wide variety of parametric or functional-form mortality models that define age-specific measures of mortality in an analytical form, sometimes with interpretable parameters. Brass (1971) developed an innovative new approach with his two-parameter 'relational' model that has been extended and refined in many ways. More recently the Log-Quad modelcombines empirical and functional-form approaches to mortality models.

Population forecasting has motivated another important family of related mortality models. Forecasting generates many iterations of both age-specific mortality and fertility into the future, and those are usually based on a summary of the corresponding age-specific mortality and fertility in the past. Hence there is an immediate need to represent full age schedules and their dynamics compactly. This led to the widespread use of dimension-reduction or data compression techniques to reduce the dimensionality of the problem so that only a few parameters are necessary to represent age schedules and their dynamics.
Ledermann and Breas (1959) appear to have been the first to use principal components analysis (PCA) to summarize age-specific mortality and generate model life tables, and this approach was refined by many subsequent investigators. Following the early use of PCA to build model life tables, PCA and related methods like the singular value decomposition (SVD) have been widely used and refined by forecasters to create time series models of mortality and fertility. Bell (1997) provides a comprehensive summary of this line of development in various fields, dominated by actuarial science and applications in forecasting.

3.1 Model Scales

This analysis is conducted on life table probabilities of dying for those who survive to the beginning of each one-year age group. Single year probabilities ${}_1q_x$ are taken directly from the HMD life tables, five-year probabilities ${}_5q_x$ are calculated as ${}_5q_x = 1 - \prod_{a=x}^{x+4} (1 -_1 q_a)$, and ${}_{45}q_{15}$ is calculated as ${}_{45}q_{15} = 1 - \prod_{a=15}^{59} (1 -_1 q_a)$. `Child mortality` refers to ${}_5q_0$ and 'adult mortality' refers to ${}_{45}q_{15}$.

The natural scale of the models described below is the full real line, so life table probabilities of dying q are transformed using the *logit* function

$$logit(x) = \ln\left(\frac{x}{1-x}\right)$$

so that their transformed values occupy the full real line. Outputs from the models are transformed back to the probability scale with range [0,1] using the *expit* function, inverse of the *logit*

$$expit(x) = \frac{\mathrm{e}^x}{1+\mathrm{e}^x}.$$

4 Results

5 Discussion

5.1 First discussion point

Mortality models often have inbuilt identification issues challenging the statistician. The statistician can choose to work with well-defined freely varying parameters, derived as maximal invariants in this paper, or with ad hoc identified parameters which at first glance seem more intuitive, but which can introduce a number of unnecessary challenges. In this paper

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

References

Bell, William R. 1997. "Comparing and Assessing Time Series Methods for Forecasting Age-Specific Fertility and Mortality Rates." *Journal of Official Statistics* 13.

Brass, William. 1971. "On the Scale of Mortality." In *Biological Aspects of Demography*, edited by William Brass, 69–110. Taylor; Francis: London, UK.

Clark, Samuel J. 2016. "A General Age-Specific Mortality Model with an Example Indexed by Child or Child/Adult Mortality." *arXiv Preprint arXiv:1612.01408.*

Coale, Ansley J., and Paul Demeny. 1966. *Regional Model Life Tables and Stable Populations.* Princeton University Press.

Ledermann, Sully. 1969. "Nouvelles Tables-Types de Mortalité." In. INED Traveaux et Documents 53. Paris: I.N.E.D. Travaux et Documents; Presses Universitaires de France.

Ledermann, Sully, and Jean Breas. 1959. "Les Dimensions de La Mortalité." *Population* 14 (4): 637–82.

Li, Nan. 2015. "Estimating Life Tables for Developing Countries." 2014/4. http://www.un.org/en/development/desa/population/publications/pdf/technical/TP2014-4.pdf: United Nations Department of Economic; Social Affairs Population Division; United Nations.

Murray, Christopher JL, Brodie D Ferguson, Alan D Lopez, Michel Guillot, Joshua A Salomon, and Omar Ahmad. 2003. "Modified Logit Life Table System: Principles, Empirical Validation, and Application." *Population Studies* 57 (2): 165–82.

United Nations, Department of Economic and Social Affairs, Population Division. 1955. *Age and Sex Patterns of Mortality: Model Life-Tables for Under-Developed Countries.* New York: United Nations Department of International Economic; Social Affairs Population Division.

———. 1982. *Model Life Tables for Developing Countries.* 77. New York: United Nations Department of International Economic; Social Affairs Population Division.

———. 2015. *World Population Prospects: The 2015 Revision.* New York: United Nations.

University of California, Berkeley, and Max Planck Institute for Demographic Research. n.d. *Human Mortality Database.* http://www.mortality.org or http://www.humanmortality.de.

Wilmoth, John, Sarah Zureick, Vladimir Canudas-Romo, Mie Inoue, and Cheryl Sawyer. 2012. "A Flexible Two-Dimensional Mortality Model for Use in Indirect Estimation." *Population Studies* 66 (1): 1–28.