US Regional Differences in Healthy and Total Life Expectancy for Persons with and without Diabetes: An Illustration of Two Multistate Life Table Methods

Scott M. Lynch
Duke University

J. Scott Brown
Miami University

[NOTE: this paper is being prepared for publication in reply to a recent journal article. In that article, the key goal is illustrating the methodology; for the PAA submission, the key concern is the substance]

Abstract

Type 2 diabetes has become increasingly prevalent in the US. While some view diabetes as a health outcome itself, diabetes contributes to numerous poor health outcomes, including cardiovascular disease, stroke, and various neuropathies. Although regional disparities in diabetic prevalence are well known, no research has examined whether life expectancy and health differences between diabetics and nondiabetics exist across regions; that is: is diabetes equally consequential to health throughout the country? Furthermore, does the region in which one was raised have an impact over one's current region of residence? In this paper, we use data from a long-term panel (the Health and Retirement Study) and cross-sectional study (the National Health Interview Survey) to examine the interaction between region of residence and diabetes status in predicting total and healthy life expectancy. We employ two recently developed multistate life table methods, and illustrate some extensions of them, in answering these questions.

Introduction

A recent article in this journal (Demographic Research) described a number of software packages/programs that produce multistate life tables. Over the last several decades, a variety of methods and packages have been developed, each receiving some use. Thus, to be sure, the recent paper made an important contribution by providing a single source in which to learn about many methods. However, some methods were discussed in greater depth than others, and the discussion of at least one method was incorrect. Specficially, the method described in Lynch and Brown (2005 and 2010) is considerably different than what described.

In particular, the methods described in Lynch and Brown (2005) and Lynch and Brown (2010) are actually two different methods. Lynch and Brown (2005) describe a method for generating distributions of multistate (and single and multiple decrement) life tables using panel data. Lynch and Brown (2010) describe a method for generating distributions of multistate life tables using cross-sectional data coupled with a collection of period mortality rates: an extension of Sullivan's method.

Both methods consist of two parts: (1) estimation of a hazard model of some form, and (2) construction of multistate life tables from the parameter estimates in step (1). The important feature of both methods is the use of Gibbs sampling to produce a joint sample from the posterior distribution for the hazard model parameters rather than the use of maximum likelihood estimation, which only produces a point estimate for the parameter and a point estimate for the standard errors. Standard errors for one parameterization (i.e., hazard model parameters) are not easily transformable into standard errors for another parameterization (i.e., life table parameters), necessitating the use of the delta method as simply an approximation. With a sample from the joint distribution of the original hazard model parameters obtained via Gibbs sampling, however, the transformation from hazard model standard errors to life table parameter standard errors is straightforward.

Specifically, Gibbs sampling produces a sample of hazard model parameters. Each set of hazard model parameters is combined with a covariate profile to provide an age-specific matrix of transition probabilities (one matrix for each age across the age range). Finally, a radix population is post-multiplied by each age-specific transition probability matrix to produce the various multistate life table quantities that are of interest, including the survivorship matrix (lx), the person-years matrix (Lx) and the state expectancy matrix (ex). Repeating this process for all of the Gibbs samples therefore yields a sample of life tables. Life table quantities of interest can then be sorted and empirical intervals for them can be constructed much as one might do using bootstrapping. Indeed, the key differences between the Gibbs sampling and a bootstrapping approach—other than the paradigmatic/theoretic differences between Bayesian and classical inference—is that bootstrapping requires resampling of the data, which may present problems when state spaces have few cases (i.e., a bootstrap sample could potentially end up with no one in a particular state). Since Gibbs sampling does not involve resampling the data, small cells are not an issue for it.

Another advantage of a Bayesian approach is the ability to incorporate prior information into the analysis, something that our 2005 and 2010 papers did not highlight but which we will highlight in our examples here.

Our "software" for both methods is currently written in R, a freely available software package. Software for our 2005 method was initially written in Unix-based C, but by 2010, we had rewritten our software in R, and our 2010 software was written in R only. Importantly, however, we do not see our key contribution to be the software, but rather the methodology. While we have written the software for a specific data structure (person-spells), a specific age-parameterization for poor health and mortality (probit; i.e., s-shaped), a specific hazard model (i.e., a discrete time bivariate probit), and a specific set of life table assumptions (e.g., linear person years lived within age intervals), the methodology is not as restrictive as the software. In other words, our specific models and assumptions are not a requirement of the general methodology, as we will illustrate in this paper.

In this paper, our goal is to illustrate both the 2005 (panel) and 2010 (cross-sectional) methodologies. Because much of the methodological detail is presented in the earlier papers, we do not repeat that here. Instead, we discuss the methods more generally. We illustrate the

methods using panel data from 7 waves of the HRS (biannual from 1998-2010) and repeated cross-sectional data from the 1997-2013 NHIS in examining total and healthy life expectancy for diabetics and nondiabetics, with a particular focus on understanding US regional differences in how diabetes translates into poor health and death.

Substantive Background

Diabetes has become increasingly prevalent over the last few decades, in conjunction with the rising obesity rate in the US. Although some view diabetes itself as a health outcome, diabetes is a precursor to many distal health outcomes, including cardiovascular disease, cerebrovascular disease, vision impairment, neuropathies, amputations, etc. The management of diabetes is crucial to preventing or limiting health consequences and premature mortality. However, health knowledge is known to vary across the US, and regional disparities in a variety of health outcomes are well established.

Data

Some Methodological Details

Results

Discussion and Conclusions

References

Lynch, Scott M. and J. Scott Brown. (2005). "A New Approach to Estimating Life Tables with Covariates and Constructing Interval Estimates of Life Table Quantities." *Sociological Methodology* 35:177-225.

Lynch, Scott M. and J. Scott Brown. (2010). "Obtaining Multistate Life Table Distributions for Highly-Refined Subpopulations from Cross-Sectional Data: A Bayesian Extension of Sullivan's Method." *Demography* 47(4):1053-1077.