Introduction

For estimating mortality and fertility patterns several parametric and non-parametric techniques have been proposed. The parametric ones are non linear models that represent the mortality pattern as a function of age and a number of parameters. (e.g. Heligman-Pollard, 1980; Hartmann, 1987; Forfar, et al. 1988; Kostaki, 1992; Hannerz, 1999). The non-parametric approaches does not involve functional forms or parameters of such forms. In general non-parametric methods apply to very wide families of distributions rather than only to families specified by a particular form.

As regards to mortality estimation two broad strategies are utilized for representing the mortality pattern of a population. The first one is based on the use of model life tables that can appropriately represent the mortality experience. The second one refers to the use of some graduation technique, applied to the empirical age-specific death frequencies for extracting the underlying probabilities of dying under the assumption that the later follows a smooth pattern. The mortality smoothing problem is of specific interest since the age-specific mortality rates exhibit a complex though typical pattern with significantly varying levels and curvature throughout the ages. Therefore many standard graduation techniques can produce inaccurate results, especially at the ends of the age interval. In the literature several models have given accurate results such as the Heligman-Pollard eight and nine parameter formula (Heligman and Pollard, 1980; Kostaki, 1992) as well as kernel estimates (Hardle, 1990; 1991; Bowman and Azzalini, 1997; Peristera and Kostaki, 2007; Peristera, 2008).

As regards to fertility estimation there is a vast literature about models that estimate the corresponding pattern. Various parametric and non parametric graduation techniques have been proposed for estimating age-specific female fertility patterns e.g. the Coale-Trussell function (Coale and Trussell, 1974; 1978), the Beta, the Gamma and the Hadwiger distributions (Hoem et al., 1981; Hadwiger, 1940; Gilje, 1969; Yntema, 1969), cubic splines (Hoem and Rennermalm, 1978; Gilks, 1986) the parametric models proposed by Peristera and Kostaki (P-K models) in 2007. Recently, new models have been proposed for estimating the distorted fertility pattern consisting of two humps, one at earlier ages and a second at later ages, indicating heterogeneity of these populations in the fertility behavior (Chandola et al., 1999; Peristera and Kostaki, 2007).

In this work emphasis is given in the use of wavelet estimates for estimating mortality and fertility patterns from various populations. The theoretical properties as well as their application for graduation purposes of demographic rates will be examined. Various types of wavelet estimates are evaluated for several populations. In order to test the reliability of the produced results we compare them with other methods either parametric or non-parametric.

Methodological Issues

Wavelet estimates is a general mathematical tool with applications in different scientific areas. Lately, wavelets are used for statistical analyses purposes, such as non-parametric regression analysis, density estimation, time series analysis, etc. (Abramovich et al., 2000; Antoniadis, 1997). However, they have not been evaluated so far in the case of demographic data. Various types of wavelets estimates have been proposed in the literature. The choice depends on the type of the problem as well as in the structure of the data. Furthermore the choice of the appropriate degree of smoothness is required for using these estimates.

In non-parametric regression, the goal is to recover an unknown function, say g, based on sampled data that are contaminated with noise. Denoising techiques provide a very effective and simpleway of finding structure in data sets without the imposition of a parametric regression model (as in linear or polynomial regression for example). Only very general assumptions about g are made such as that it belongs to a certain class of functions. Thus, simple denoising algorithms that use the wavelet transform consist of three steps:

1) Calculate the wavelet transform of the noisy signal.

2) Modify the noisy wavelet coefficients according to some rule.

3) Compute the inverse transform using the modified coefficients. Traditionally, for the second step of the above approach there are two kinds of denoising methods; namely, linear and nonlinear techniques

The nonparametric regression literature was arguably dominated by (nonlinear) wavelet shrinkage and wavelet thresholding estimators. These estimators are a new subset of an old class of nonparametric regression estimators, namely orthogonal series methods. Moreover, these estimators are easily implemented through fast algorithms so they are very appealing in practical situations (Donoho and Johnstone, 1994; Donoho et al., 1995). Several new wavelet based curve smoothing procedures have been recently proposed, and one of the purposes of this review is to present few of them under the general concept of penalized least squares regression.

The term wavelets is used to refer to a set of orthonormal basis functions generated by dilation and translation of a compactly supported scaling function (or father wavelet), φ, and a mother wavelet, ψ, associated with an r-regular multiresolution analysis of L2(R). A variety of different wavelet families now exist that combine compact support with various degrees of smoothness and numbers of vanishing moments (Daubechies, 1992).

Consider the standard univariate nonparametric regression setting $y_i = g(t_i) + \sigma \varepsilon_i$, $i = 1, . ., n$, where ε_i are independent $N(0, 1)$ random variables and the noise level σ may be known or unknown. We suppose, without loss of generality, that t_i are within the unit interval [0, 1]. The goal is to recover the underlying function g from the noisy data, $y = (y_1, . . . , y_n)^T$, without assuming any particular parametric structure for g.

One of the basic approaches to handle this regression problem is to consider the unknown function g expanded as a generalised Fourier series and then to estimate the generalized Fourier coefficients from the data. The original (nonparametric) problem is thus transformed to a parametric one, although the potential number of parameters is infinite. An appropriate choice of basis for the expansion is therefore a key point in relation to the efficiency of such an approach. A ‘good’ basis should be parsimonious in the sense that a large set of possible response functions can be approximated well by only few terms of the generalized Fourier expansion employed. Wavelet series allow a parsimonious expansion for a wide variety of functions, including inhomogeneous cases. It is therefore natural to consider applying the generalized Fourier series approach using a wavelet series.

The performance of the resulting wavelet estimator depends on the penalty and the regularization parameter λ. To select a good penalty function, Antoniadis and Fan (2001) and Fan and Li (2001) proposed three principles that a good penalty function should satisfy. In that, as regards to the choice of the penalty parameter λ in finite sample situations an optimal choice is important. Given the basic framework of function estimation using wavelet threshold and its relation with the regularization approach with penalties non smooth at 0, there are a variety of methods to choose the regularization parameter λ. Solo (2001) in his discussion of the paper by Antoniadis and Fan (2001) suggests a data based estimator of λ similar in spirit to the SURE selection criterion used by Donoho and Johnstone (1994), and provides an appropriate simple SURE formula for the general penalties studied in this review. Another way to address the optimal choice of the regularization parameter is Generalized Cross Validation (GCV). Cross validation has been widely used as an automatic procedure to choose the smoothing parameter in many statistical settings. The classical cross-validation method is performed by systematically expelling a data point from the construction of an estimate, predicting what the removed value would be and, then, comparing the prediction with the value of the expelled point. One way to proceed is to use the approach adopted by Jansen et al. (1997). It is clear that this is an area where further careful theoretical and practical work is needed.

References

Abramovich F., Bailey T. C., Sapatinas T. (2000). Wavelet analysis and tis statistical applications. The Statistician, 49,1, 1-129
Antoniadis A. (1997). Wavelets in statistics: a review (with discussion). Italian Journal of Statistics, 6, 97-144
Antoniadis, A. & Fan, J. (2001) Regularization of wavelets approximations (with discussion). Journal of American Statistical Association, 96, 939-967
Bowman, A.W. and Azzalini, A. (1997). Applied Smoothing Techniques for data analysis: The kernel approach with S-plus illustrations. Oxford Science Publications, Oxford
Chandola. T., Coleman, D. A., Horns, R. W. (1999). Recent European fertility patterns: fitting curves to 'distorted' distributions, Population Studies, 53, 3, 317-329
Coale, A.J. and Trussell, T. J. (1974). Model fertility schedules: variations in the age structure of childbearing in human populations, Population Index, 40, 2, 185-258
Coale, A.J. and Trussell, T. J. (1978). Technical note: finding the two parameters that specify a model schedule of marital fertility, Population Index, 44, 2, 203-214
Daubechies, I. (1992). Ten Lectures on Wavelets. Philadelphia: SIAM
Donoho, D.L., & Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81, 425-455
Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., & Picard, D. (1995). Wavelet shrinkage: asymptopia? (with discussion). Journal of Royal Statistical Society, B, 57, 301-337
Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of American Statistical Association, 96 1348–1360
Forfar, DO., McCutcheon, J.J., and Wilkie, A.D. (1988). On graduation by mathematical formula, Journal of the Institute of Actuaries, 115, 1-135
Gilje, E. (1969). Fitting curves to age-specific fertility rates: some examples, Statistical Review of the Swedish National Central Bureau of Statistics III, 7, 18-134

Gilks, W. R. (1986). The relationship between birth history and current fertility in developing countries, Population Studies, 40, 3, 437-455
Hadwiger, H. (1940). Eine analytische reprodutions-funktion fur biologische Gesamtheiten, Skandinavisk Aktuarietidskrift, 23, 101-113
Hannerz, H. (1999). Methodology and applications of a new law of mortality. Lund: Department of Statistics. University of Lund, Sweden

Hardle, W. (1990). Applied Non-parametric Regression. Cambridge University Press, Cambridge
Hardle, W. (1991). Smoothing Techniques with Implementation in S. Springer-Verlag, New-York
Hartmann, M. (1987). Past and Recent Attempts to Model Mortality at All Ages, Journal of Official Statistics, 3, 19-36
Heligman, M. and Pollard, J.H. (1980). The age pattern of mortality, Journal of the Institute of Actuaries, 107, 49-80
Hoem, J. M., Madsen, D., Nielsen, J. L., Ohlsen, E., Hansen, H. O., Rennermalm, B. (1981). Experiments in modelling recent Danish fertility curves, Demography, 18, 231-244
Hoem, J. M. and Rennermalm, B. (1978). On the statistical theory of graduation by Splines, Working Paper No. 14, University of Copenhagen, Laboratory of Actuarial Mathematics
Jansen, M., Malfait, M. & Bultheel, A. (1997). Generalised cross-validation for wavelet thresholding. Signal. Processing., 56, 33–44.
Kostaki, A. (1992). Nine-parameter version of the Heligman Pollard Formula, Mathematical Population Studies, 3, 4, 277-288
Peristera, P. (2008). Statistical methodology for estimating patterns of demographic phenomena. PhD Thesis. Athens University of Economics and Business
Peristera, P. and Kostaki, A. (2007). Modeling Fertility in modern populations. Demographic Research, 16, 6, 141-194.
Solo, V. (2001). Invited discussion on "Regularization of wavelets approximations" by A. Antoniadis and J. Fan. Journal of American Statistical Association, 96 963–965

Yntema, L. (1969). On Hadwiger's fertility function, Statistical Review of the Swedish National Central Bureau of Statistics III, 7, 113-117