

test

Jon Minton

2022-07-26

Aim

The aim of this appendix is to present results for an alternative method of attempting to detect discontinuities in the series of life expectancy changes for different countries. The main approach used was segmented regression using the `segmented` package. The segmented regressions were fit to the data itself y t . This alternative approach instead involves fitting data to the first differences $y(t) - y(t - 1)$ t .

The alternative approach identified some similar breakpoints/change points to the main, segmented approach. However it differed for many populations in identifying adjacent or near-adjacent change points in the data. Further inspection of these pairs of change points indicated that the approach developed inadvertently operated as a form of outlier detector algorithm, removing from the series extreme values. Though extreme values are a genuine phenomenon within mortality change data, they were not the primary purpose of the analysis, and so this approach was not used within the main manuscript.

Description of approach

- We first fit a ‘null’ (no change) model to the data. These are models that represent the assumption that no change occurred in trends in life expectancy over time.
- Then we fit a series on one-break models to the data. These are models that represent the assumption that a single significant change occurred to the series over time. From this series of one-break models, we attempt to find the best one-break model.
- We then fit a series of two-break models to the data. These are models that represent the assumption that there were two significant changes to the series over time. From this series of two-break models, we attempt to find the best two-break models.
- We now have three models to compare: the null model (M_0), the best one-break model (M_1), and the best two-break model (M_2). Although both M_1 can be compared directly¹ with M_0 , and M_2 can be compared directly with M_0 , M_2 cannot necessarily be compared directly with M_1 . Instead, the triplet of models (M_0 , M_1 and M_2) are compared indirectly using BIC² [BIC stands for Bayesian Information Criterion, like the similar AIC (An Information Criterion or Akaike’s Information Criterion), is a penalised model fit score. By penalised this means that the fit of the model to the data (more specifically its log likelihood) is calculated, then a ‘penalty’ is applied to this score based on the complexity of the model. BIC and AIC differ only according to how the penalty is applied, with BIC tending to penalise more complex models more severely than AIC. This means BIC will tend to be more conservative in selecting models, providing some protection against overfitting. Both AIC and BIC can be used to compare both nested and non-nested models based on the same dataset, unlike the F-test, the Lagrange Multiplier Test, and so on. Measures of model fit should *not* be used to compare models fit to different datasets.², which like AIC provides a penalised model fit score. Lower BIC scores

¹By ‘compared directly’, we mean that one model specification can be expressed as a restricted/constrained version of another model, the unrestricted model, with one or more terms in the unrestricted model set to fixed values, usually zero, in the restricted model. Such models can be compared directly using an F-test

²For example, it would be wrong to conclude that a model with an R^2 or adjusted R^2 fit to one dataset of 0.80 is ‘better’ than a model with an R^2 or adjusted R^2 of 0.50 fit to a different dataset, even though R^2 and its variants are often (mis)interpreted in this way

indicate better fit to the data, and so, for each dataset, the model (M_0 , M_1 or M_2) with the lowest BIC will be selected.

- For the best of the three models for each dataset, the breakpoint or breakpoints (if any), and the model predictions, will be presented and visualised.

Breaks in the data are estimated by fitting the following model specifications:

- M_0 : $\frac{de_x}{dt} \sim \alpha$
- M_1 : $\frac{de_x}{dt} \sim \alpha + \beta T$, where

$$\begin{array}{ccc} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{array}$$

$$T = \begin{cases} 1 & \text{if } t \geq \tau \\ 0 & \text{otherwise} \end{cases}$$

- M_2 : $\frac{de_x}{dt} \sim \alpha + \beta_1 T_1 + \beta_2 T_2$, where

$$T_1 = \begin{cases} 1 & \text{if } \tau_1 \leq t < \tau_2 \\ 0 & \text{otherwise} \end{cases}$$