Specification of the Lee-Carter approach

1.1 Notations and abbreviations

\( y_{\text{age}, \text{year}} \): number of events of age in year

\( n_{\text{age}, \text{year}} \): population size of age in year

\( \alpha_{\text{age}} \): age-specific effect of age

\( \beta_{\text{age}} \): age-period adjustment term of age

\( \kappa_{\text{year}} \): period effect at year

\( E(\cdot) \): expectation function

\( \text{sign}(\cdot) \): sign function of a value, 1 for positive and -1 for negative.

\( \ell(\cdot) \): log-likelihood function

\( \text{ARIMA}(p, d, q) \): Autoregressive integrated moving average with \( p \) number of time lags of the autoregressive term, \( d \) order of differencing term, and \( q \) number of time lags of the moving-average term.
1.2 Model structure and assumptions

Lee-Carter model [1] was defined as

\[
\log(E(y_{\text{year}, \text{age}})) = \alpha_{\text{age}} + \beta_{\text{age}} \kappa_{\text{year}} + \log(n_{\text{year}, \text{age}})
\]

Two constraints are needed for ensuring identifiability:

\[
\sum_{\text{year}} \kappa_{\text{year}} = 0 \\
\sum_{\text{age}} \beta_{\text{age}} = 1
\]

In implementation, we used a Poisson-regression-based approach [2] for the Lee-Carter model

\[
y_{\text{year}, \text{age}} \sim \text{Poisson}(\mu_{\text{year}, \text{age}})
\]

\[
\mu_{\text{year}, \text{age}} = n_{\text{year}, \text{age}} \exp(\alpha_{\text{age}} + \beta_{\text{age}} \kappa_{\text{year}})
\]

Thus, the log-likelihood function is

\[
\ell(\alpha_{\text{age}}, \beta_{\text{age}}, \kappa_{\text{year}} | y_{\text{year}, \text{age}}, n_{\text{year}, \text{age}}) = -\mu_{\text{year}, \text{age}} + y_{\text{year}, \text{age}} \log(\mu_{\text{year}, \text{age}}) - \log(y_{\text{year}, \text{age}}!)
\]

\[
= y_{\text{year}, \text{age}}(\alpha_{\text{age}} + \beta_{\text{age}} \kappa_{\text{year}}) - n_{\text{year}, \text{age}} \exp(\alpha_{\text{age}} + \beta_{\text{age}} \kappa_{\text{year}}) + \text{constant}
\]

1.2.1 Comparator models

Two reduced models were considered as comparators:

**Age-Period model** assumed \( \beta_{\text{age}} \) are the same for all age groups, so the model reduced to

\[
y_{\text{year}, \text{age}} \sim \text{Poisson}(\mu_{\text{year}, \text{age}})
\]

\[
\mu_{\text{year}, \text{age}} = n_{\text{year}, \text{age}} \exp(\alpha_{\text{age}} + \kappa_{\text{year}})
\]

**Age-Trend model** assumed \( \beta_{\text{age}} \) are the same for all age groups and a linear period effect, so the model reduced to
\[ y_{\text{year,age}} \sim \text{Poisson}(\mu_{\text{year,age}}) \]
\[ \mu_{\text{year,age}} = n_{\text{year,age}} \exp(\alpha_{\text{age}} + \text{year} \times \kappa) \]

### 1.3 Model fitting

The maximum likelihood estimation of this model is to solve

\[
\arg\max_{\alpha_{\text{age}}, \beta_{\text{age}}, \kappa_{\text{year}}} \ell(\alpha_{\text{age}}, \beta_{\text{age}}, \kappa_{\text{year}} | y_{\text{year,age}}, n_{\text{year,age}})
\]

subject to

\[
\sum_{\text{year}} \kappa_{\text{year}} = 0 \\
\sum_{\text{age}} \beta_{\text{age}} = 1
\]

We employed Newton methods as described in Brouhns et al. [2] to this task. Our implementation used `StMoMo::fit` function from package StMoMo [3].

### 1.4 Modelling and forecasting

For period effects, \( \kappa_{\text{year}} \), we employed the Box-Jenkins method, which uses autocorrelation function (ACF), partial autocorrelation function (PACF), and extended ACF if necessary to specified a time-series model of \( \kappa_{\text{year}} \). This specification and modelling were implemented using functions, `TSA::acf`, `TSA::pacf`, and `TSA::eacf` from package `TSA` [4].

### 1.5 Bootstrap

The bootstrap simulation employed the semi-parametric bootstrap by Renshaw and Haberman [5]. Our implementation used `StMoMo::simulation` function from package StMoMo [3]. In general, we generated 10,000 simulations for each presented result.
1.6. Measurements of goodness of fit

Since the likelihood-based LCM we applied is a special case of the ordinary Poisson regression [6], the measurements of goodness of fit of Poisson regression can be directly applied.

**Akaike information criterion (AIC):** by definition,

\[ 2k - 2\hat{\ell}(\hat{\alpha}_{age}, \hat{\beta}_{age}, \hat{\kappa}_{year}|y_{year,age}, n_{year,age}) \]

, where \(\hat{\ell}(.)\) is the log-likelihood function given estimated parameters, \(k\) is the number of parameters, which equals the sum of the numbers of \(\hat{\alpha}_{age}, \hat{\beta}_{age}, \hat{\kappa}_{year}\) minus two constraints.

**Bayesian information criterion (BIC):** by definition,

\[ \log(o)k - 2\hat{\ell}(\hat{\alpha}_{age}, \hat{\beta}_{age}, \hat{\kappa}_{year}|y_{year,age}, n_{year,age}) \]

, where \(k\) is the number of parameters as above and \(o\) is number of observations.

**Deviance residuals:** We used the deviance residuals defined in Colin Cameron and Trivedi [6] to assess the goodness of fit.

\[ \text{sign}(y_{age,year} - E(y_{age,year})) \sqrt{2\left[y_{age,year}\log \frac{y_{age,year}}{E(y_{age,year})} - (y_{age,year} - E(y_{age,year}))\right]} \]
Bibliography


