Outline

1. Functional time series
2. Current state of Australian population forecasting
3. Stochastic population forecasting
Outline

1. Functional time series
2. Current state of Australian population forecasting
3. Stochastic population forecasting
Mortality rates

Australia: male mortality (1921)
Fertility rates

Australia: fertility rates (1921)
Some notation

Let $y_t(x_i)$ be the observed data in period $t$ at age $x_i$, $i = 1, \ldots, p$, $t = 1, \ldots, n$.

$$y_t(x_i) = s_t(x_i) + \sigma_t(x_i) \varepsilon_{t,i}$$

$\varepsilon_{t,i} \overset{iid}{\sim} N(0, 1)$
Some notation

Let $y_t(x_i)$ be the observed data in period $t$ at age $x_i$, $i = 1, \ldots, p$, $t = 1, \ldots, n$.

\[ y_t(x_i) = s_t(x_i) + \sigma_t(x_i) \varepsilon_{t,i} \]

- $\varepsilon_{t,i} \overset{iid}{\sim} N(0, 1)$
- $s_t(x)$ and $\sigma_t(x)$ are smooth functions of $x$. 
Some notation

Let $y_t(x_i)$ be the observed data in period $t$ at age $x_i$, $i = 1, \ldots, p$, $t = 1, \ldots, n$.

$$y_t(x_i) = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i}$$

- $\varepsilon_{t,i} \sim \text{N}(0, 1)$
- $s_t(x)$ and $\sigma_t(x)$ are smooth functions of $x$.
- We need to estimate $s_t(x)$ from the data for $x_1 < x < x_p$. 
Let $y_t(x_i)$ be the observed data in period $t$ at age $x_i$, $i = 1, \ldots, p$, $t = 1, \ldots, n$.

$$y_t(x_i) = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i}$$

- $\varepsilon_{t,i} \sim \text{iid } N(0, 1)$
- $s_t(x)$ and $\sigma_t(x)$ are smooth functions of $x$.
- We need to estimate $s_t(x)$ from the data for $x_1 < x < x_p$.
- We want to forecast whole curve $y_t(x)$ for $t = n + 1, \ldots, n + h$. 
Stochastic population forecasts using FDM

Functional time series model

\[
y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x}
\]
\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where \( \varepsilon_{t,x} \overset{iid}{\sim} N(0, 1) \) and \( e_t(x) \overset{iid}{\sim} N(0, v(x)) \).
Functional time series model

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]
\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \stackrel{iid}{\sim} N(0, 1) \) and \( e_t(x) \stackrel{iid}{\sim} N(0, v(x)) \).

1. Estimate smooth functions \( s_t(x) \) using nonparametric regression.
**Functional time series model**

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]

\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \stackrel{iid}{\sim} N(0,1) \) and \( e_t(x) \stackrel{iid}{\sim} N(0, v(x)) \).

1. Estimate smooth functions \( s_t(x) \) using nonparametric regression.

2. Estimate \( \mu(x) \) as mean \( s_t(x) \) across years.
**Functional time series model**

\[
y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x}
\]

\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where \(\varepsilon_{t,x} \sim \text{iid } N(0,1)\) and \(e_t(x) \sim \text{iid } N(0, v(x))\).

1. Estimate smooth functions \(s_t(x)\) using nonparametric regression.
2. Estimate \(\mu(x)\) as mean \(s_t(x)\) across years.
3. Estimate \(\beta_{t,k}\) and \(\phi_k(x)\) using functional principal components.
Stochastic population forecasts using FDM

**Functional time series model**

\[ y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x} \]

\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \sim \text{iid } N(0, 1) \) and \( e_t(x) \sim \text{iid } N(0, \nu(x)) \).

1. Estimate smooth functions \( s_t(x) \) using nonparametric regression.
2. Estimate \( \mu(x) \) as mean \( s_t(x) \) across years.
3. Estimate \( \beta_{t,k} \) and \( \phi_k(x) \) using functional principal components.
4. Forecast \( \beta_{t,k} \) using time series models.
Functional time series model

\[
y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x}
\]

\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where \(\varepsilon_{t,x} \overset{iid}{\sim} N(0, 1)\) and \(e_t(x) \overset{iid}{\sim} N(0, \nu(x))\).

1. Estimate smooth functions \(s_t(x)\) using nonparametric regression.
2. Estimate \(\mu(x)\) as mean \(s_t(x)\) across years.
3. Estimate \(\beta_{t,k}\) and \(\phi_k(x)\) using functional principal components.
4. Forecast \(\beta_{t,k}\) using time series models.
5. Put it all together to get forecasts of \(y_t(x)\).
Functional time series model

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]

\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \overset{\text{iid}}{\sim} N(0,1) \) and \( e_t(x) \overset{\text{iid}}{\sim} N(0, v(x)) \).

1. **Estimate smooth functions** \( s_t(x) \) **using nonparametric regression**.
2. Estimate \( \mu(x) \) as mean \( s_t(x) \) across years.
3. Estimate \( \beta_{t,k} \) and \( \phi_k(x) \) using functional principal components.
4. Forecast \( \beta_{t,k} \) using time series models.
5. Put it all together to get forecasts of \( y_t(x) \).
1. Creating functional time series

Australia: male death rates (1921–2003)
1. Creating functional time series

Australia: male death rates (1921–2003)
1. Creating functional time series

Monotonic regression splines

- Fit penalized regression spline with a large number of knots.
1. Creating functional time series

Monotonic regression splines

- Fit penalized regression spline with a large number of knots.
- For mortality data, constrain curve to be monotonic for $x > b$. 

Choosing $b = 50$ seems to work quite well in practice for mortality data. Fit is weighted with weights $w_t(x_i) = \sigma - \frac{2}{t(x_i)}$ (based on Poisson deaths). This can be done using a modification of the `gam` function in the `mgcv` package in R.
1. Creating functional time series

Monotonic regression splines

- Fit penalized regression spline with a large number of knots.
- For mortality data, constrain curve to be monotonic for $x > b$.
- Choosing $b = 50$ seems to work quite well in practice for mortality data.
1. Creating functional time series

Monotonic regression splines

- Fit penalized regression spline with a large number of knots.
- For mortality data, constrain curve to be monotonic for $x > b$.
- Choosing $b = 50$ seems to work quite well in practice for mortality data.
- Fit is weighted with weights $w_t(x_i) = \sigma_t^{-2}(x_i)$ (based on Poisson deaths).
1. Creating functional time series

Monotonic regression splines

- Fit penalized regression spline with a large number of knots.
- For mortality data, constrain curve to be monotonic for $x > b$.
- Choosing $b = 50$ seems to work quite well in practice for mortality data.
- Fit is weighted with weights $w_t(x_i) = \sigma_t^{-2}(x_i)$ (based on Poisson deaths).
- This can be done using a modification of the `gam` function in the `mgcv` package in R.
Functional time series model

\[ y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x} \]
\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \sim \text{iid} \ N(0, 1) \) and \( e_t(x) \sim \text{iid} \ N(0, v(x)) \).

1. Estimate smooth functions \( s_t(x) \) using nonparametric regression.
2. **Estimate** \( \mu(x) \) **as mean** \( s_t(x) \) **across years.**
3. Estimate \( \beta_{t,k} \) and \( \phi_k(x) \) using functional principal components.
4. Forecast \( \beta_{t,k} \) using time series models.
5. Put it all together to get forecasts of \( y_t(x) \).
2. Estimate $\mu(x)$

Australia: male death rates (1921–2003)
2. Estimate $\mu(x)$

**Australia: male death rates (1921–2003)**
Stochastic population forecasts using FDM

Functional time series

**Functional time series model**

\[
y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x}
\]

\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where \(\varepsilon_{t,x} \sim \text{iid } N(0, 1)\) and \(e_t(x) \sim \text{iid } N(0, \nu(x))\).

1. Estimate smooth functions \(s_t(x)\) using nonparametric regression.
2. Estimate \(\mu(x)\) as mean \(s_t(x)\) across years.
3. **Estimate** \(\beta_{t,k}\) and \(\phi_k(x)\) **using** functional principal components.
4. Forecast \(\beta_{t,k}\) using time series models.
5. Put it all together to get forecasts of \(y_t(x)\).
3. Functional PC

The optimal basis functions

\[ S_t(x) = \mu(x) + \sum_{i=0}^{K} \beta_{t,i}\phi_i(x) + e_t(x). \]

where \( e_t(x) = \sum_{i=K+1}^{n-1} \beta_{t,i}\phi_i(x) \).
3. Functional PC

The optimal basis functions

\[ s_t(x) = \mu(x) + \sum_{i=0}^{K} \beta_{t,i} \phi_i(x) + e_t(x). \]

where \( e_t(x) = \sum_{i=K+1}^{n-1} \beta_{t,i} \phi_i(x). \)

For a given \( K \), the basis functions \( \phi_i(x) \) which minimize

\[ \text{MISE} = \frac{1}{n} \sum_{t=1}^{n} \int v_t^2(x) \, dx \]

are the principal components.
3. Functional PC

(Ramsay and Silverman, 1997, 2002).

- In FDA, each principal component is specified by a weight function $\phi_i(x)$.
3. Functional PC

(Ramsay and Silverman, 1997, 2002).

- In FDA, each principal component is specified by a weight function $\phi_i(x)$.
- The PC scores for each year are given by:

$$z_{i,t} = \int \phi_i(x) \hat{S}_t(x) dx$$
3. Functional PC

(Ramsay and Silverman, 1997, 2002).

- In FDA, each principal component is specified by a weight function $\phi_i(x)$.
- The PC scores for each year are given by
  \[ z_{i,t} = \int \phi_i(x) \hat{S}_t(x) dx \]
- The aim is to:
3. Functional PC

(Ramsay and Silverman, 1997, 2002).

- In FDA, each principal component is specified by a weight function $\phi_i(x)$.
- The PC scores for each year are given by
  \[ z_{i,t} = \int \phi_i(x)\hat{S}_t(x)dx \]
- The aim is to:
  1. Find the weight function $\phi_1(x)$ that maximizes the variance of $z_{1,t}$ subject to the constraint
     \[ \int \phi_1^2(x)dx = 1. \]
3. Functional PC

(Ramsay and Silverman, 1997, 2002).

- In FDA, each principal component is specified by a weight function $\phi_i(x)$.
- The PC scores for each year are given by
  \[ z_{i,t} = \int \phi_i(x) \hat{S}_t(x) dx \]
  \[ z_{i,t} = \int \phi_i(x) \hat{S}_t(x) dx \]
- The aim is to:
  1. Find the weight function $\phi_1(x)$ that maximizes the variance of $z_{1,t}$ subject to the constraint $\int \phi_1^2(x) dx = 1$.
  2. Find the weight function $\phi_2(x)$ that maximizes the variance of $z_{2,t}$ such that $\int \phi_2^2(x) dx = 1$ and $\int \phi_1(x) \phi_2(x) dx = 0$. 
3. Functional PC

(Ramsay and Silverman, 1997, 2002).

- In FDA, each principal component is specified by a weight function $\phi_i(x)$.
- The PC scores for each year are given by
  $$z_{i,t} = \int \phi_i(x) \hat{s}_t(x) dx$$
- The aim is to:
  1. Find the weight function $\phi_1(x)$ that maximizes the variance of $z_{1,t}$ subject to the constraint $\int \phi_i^2(x) dx = 1$.
  2. Find the weight function $\phi_2(x)$ that maximizes the variance of $z_{2,t}$ such that $\int \phi_i^2(x) dx = 1$ and $\int \phi_1(x) \phi_2(x) dx = 0$.
  3. Find the weight function $\phi_3(x)$ that...
3. Functional PC

Main effects

Interaction

Beta 1

Beta 2

Age

Time

0 20 40 60 80 100

0 20 40 60 80 100

−8 −6 −4 −2

0.0 0.1 0.2 0.3 0.4

−0.6 −0.4 −0.2 0.0 0.2 0.4

−2 −1 0 1 2

1950 1970 1990

−0.6 −0.4 −0.2 0.0 0.2 0.4

1950 1970 1990
3. Functional PC

Recap

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]
\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \overset{iid}{\sim} N(0,1) \) and \( e_t(x) \overset{iid}{\sim} N(0, \nu(x)) \).

- Pure time term excluded as it would make \( \{\beta_{t,k}\} \) correlated.
3. Functional PC

Recap

\[ y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x} \]

\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \) \( \text{iid} \sim N(0, 1) \) and \( e_t(x) \) \( \text{iid} \sim N(0, \nu(x)) \).

- Pure time term excluded as it would make \( \{\beta_{t,k}\} \) correlated.
- We can check if any structure is left in the residuals \( \varepsilon_{t,x} \) (smoothing problem) and \( e_t(x) \) (modelling problem).
3. Functional PC

Smoothing residuals

Year

Age
# 3. Functional PC

## Modelling residuals

<table>
<thead>
<tr>
<th>Year</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>0</td>
</tr>
<tr>
<td>1960</td>
<td>20</td>
</tr>
<tr>
<td>1970</td>
<td>40</td>
</tr>
<tr>
<td>1980</td>
<td>60</td>
</tr>
<tr>
<td>1990</td>
<td>80</td>
</tr>
<tr>
<td>2000</td>
<td>100</td>
</tr>
</tbody>
</table>
**Functional time series model**

\[
y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x}
\]

\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where \(\varepsilon_{t,x} \overset{iid}{\sim} N(0, 1)\) and \(e_t(x) \overset{iid}{\sim} N(0, v(x))\).

1. Estimate smooth functions \(s_t(x)\) using nonparametric regression.
2. Estimate \(\mu(x)\) as mean \(s_t(x)\) across years.
3. Estimate \(\beta_{t,k}\) and \(\phi_k(x)\) using functional principal components.
4. **Forecast** \(\beta_{t,k}\) using time series models.
5. Put it all together to get forecasts of \(y_t(x)\).
4. Forecasting the coefficients

Main effects

Interaction

Basis function 1

Basis function 2

Coefficient 1

Coefficient 2

Mean

0 20 40 60 80 100

−8 −6 −4 −2

0 20 40 60 80 100

−0.6 −0.2 0.0 0.2 0.4

0 20 40 60 80 100

−0.6 −0.2 0.0 0.2 0.4

1960 1980 2000 2020

0 20 40 60 80 100

−0.6 −0.2 0.0 0.2 0.4

1960 1980 2000 2020
4. Forecasting the coefficients

I use exponential smoothing state space models based on Ord, Koehler & Snyder (JASA, 1997) and Hyndman, Koehler, Snyder & Grose (IJF, 2002).
4. Forecasting the coefficients

- I use exponential smoothing state space models based on Ord, Koehler & Snyder (JASA, 1997) and Hyndman, Koehler, Snyder & Grose (IJF, 2002).

- These provide a stochastic framework for exponential smoothing forecasts.
4. Forecasting the coefficients

- I use exponential smoothing state space models based on Ord, Koehler & Snyder (JASA, 1997) and Hyndman, Koehler, Snyder & Grose (IJF, 2002).

- These provide a stochastic framework for exponential smoothing forecasts.

- The models shown are equivalent to...
4. Forecasting the coefficients

- I use exponential smoothing state space models based on Ord, Koehler & Snyder (JASA, 1997) and Hyndman, Koehler, Snyder & Grose (IJF, 2002).
- These provide a stochastic framework for exponential smoothing forecasts.
- The models shown are equivalent to
damped Holt’s method
4. Forecasting the coefficients

- I use exponential smoothing state space models based on Ord, Koehler & Snyder (JASA, 1997) and Hyndman, Koehler, Snyder & Grose (IJF, 2002).
- These provide a stochastic framework for exponential smoothing forecasts.
- The models shown are equivalent to
  - damped Holt’s method
  - ARIMA(0,1,2)
4. Forecasting the coefficients

- I use exponential smoothing state space models based on Ord, Koehler & Snyder (JASA, 1997) and Hyndman, Koehler, Snyder & Grose (IJF, 2002).
- These provide a stochastic framework for exponential smoothing forecasts.
- The models shown are equivalent to:
  - damped Holt’s method
  - ARIMA(0,1,2)
- Univariate models are ok because the series are uncorrelated.
Functional time series model

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]

\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

where \( \varepsilon_{t,x} \overset{iid}{\sim} N(0,1) \) and \( e_t(x) \overset{iid}{\sim} N(0, v(x)) \).

1. Estimate smooth functions \( s_t(x) \) using nonparametric regression.
2. Estimate \( \mu(x) \) as mean \( s_t(x) \) across years.
3. Estimate \( \beta_{t,k} \) and \( \phi_k(x) \) using functional principal components.
4. Forecast \( \beta_{t,k} \) using time series models.
5. **Put it all together to get forecasts of** \( y_t(x) \).
5. Forecasts of \( y_t(x) \)

\[
y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x}
\]

\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where \( \varepsilon_{t,x} \overset{iid}{\sim} N(0, 1) \) and \( e_t(x) \overset{iid}{\sim} N(0, \nu(x)) \).
5. Forecasts of $y_t(x)$

$$y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x}$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)$$

where $\varepsilon_{t,x} \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$ and $e_t(x) \overset{\text{iid}}{\sim} \mathcal{N}(0, v(x))$.

Let $\mathcal{I} = \{y_t(x_i); t = 1, \ldots, n; i = 1, \ldots, p\}$.

- $\mathbb{E}[y_{n+h}(x) \mid \mathcal{I}, \Phi] = \hat{\mu}(x) + \sum_{k=1}^{K} \hat{\beta}_{n+h|n,k} \hat{\phi}_k(x)$. 
5. Forecasts of $y_t(x)$

\[
y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x}
\]

\[
s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

where $\varepsilon_{t,x} \overset{\text{iid}}{\sim} N(0, 1)$ and $e_t(x) \overset{\text{iid}}{\sim} N(0, \nu(x))$.

Let $I = \{ y_t(x_i); t = 1, \ldots, n; i = 1, \ldots, p \}$.

- $E[y_{n+h}(x) | I, \Phi] = \hat{\mu}(x) + \sum_{k=1}^{K} \hat{\beta}_{n+h|n,k} \phi_k(x)$.

- $\text{Var}[y_{n+h}(x) | I, \Phi] = \sigma_{n+h}^2(x) + \hat{\sigma}_{\mu}^2(x) + \sum_{k=1}^{K} \nu_{n+h|n,k} \hat{\phi}_k^2(x) + \nu(x)$

where $\nu_{n+h|n,k} = \text{Var}(\beta_{n+h,k} | \beta_{1,k}, \ldots, \beta_{n,k})$. 
5. Forecasts of $y_t(x)$

Australia: male death rate forecasts (2004 and 2023)
5. Forecasts of $y_t(x)$

Australia: male death rate forecasts (2004 and 2023)
5. Forecasts of $y_t(x)$

Australia: male death rate forecasts (2004 and 2023)

Log death rate
Age
5. Forecasts of $y_t(x)$

Australia: male death rate forecasts (2004 and 2023)
Some references

Some references


Some references


Some references

Outline

1. Functional time series
2. Current state of Australian population forecasting
3. Stochastic population forecasting
The Australian Bureau of Statistics provide population “projections”.

“The projections are not intended as predictions or forecasts, but are illustrations of growth and change in the population that would occur if assumptions made about future demographic trends were to prevail over the projection period.

While the assumptions are formulated on the basis of an assessment of past demographic trends, both in Australia and overseas, there is no certainty that any of the assumptions will be realised. In addition, no assessment has been made of changes in non-demographic conditions.”

ABS 3222.0 - Population Projections, Australia, 2004 to 2101
ABS population projections

The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates

No objectivity.
No dynamic changes in rates allowed
No variation allowed across ages.
No probabilistic basis.
Not prediction intervals.

Most users use the “Medium” projection, but it is unrelated to the mean, median or mode of the future distribution.
The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates
- No objectivity.
The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates
- No objectivity.
- No dynamic changes in rates allowed
The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates
- No objectivity.
- No dynamic changes in rates allowed
- No variation allowed across ages.
The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates
- No objectivity.
- No dynamic changes in rates allowed
- No variation allowed across ages.
- No probabilistic basis.
The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates
- No objectivity.
- No dynamic changes in rates allowed
- No variation allowed across ages.
- No probabilistic basis.
- Not prediction intervals.
The ABS provides three projection scenarios labelled “High”, “Medium” and “Low”.

- Based on assumed mortality, fertility and migration rates
- No objectivity.
- No dynamic changes in rates allowed
- No variation allowed across ages.
- No probabilistic basis.
- Not prediction intervals.

Most users use the “Medium” projection, but it is unrelated to the mean, median or mode of the future distribution.
ABS population projections

What do these projections mean?
What do these projections mean?
Outline

1. Functional time series
2. Current state of Australian population forecasting
3. Stochastic population forecasting
Forecasts represent median of future distribution.
Stochastic population forecasts

- Forecasts represent median of future distribution.
- Percentiles allow information about uncertainty.
Stochastic population forecasts

- Forecasts represent median of future distribution.
- Percentiles allow information about uncertainty
- Prediction intervals with specified probability coverage for population size and all derived variables (total fertility rate, life expectancy, old-age dependencies, etc.)
Stochastic population forecasts using FDM

- Forecasts represent median of future distribution.
- Percentiles allow information about uncertainty
- Prediction intervals with specified probability coverage for population size and all derived variables (total fertility rate, life expectancy, old-age dependencies, etc.)
- The probability of future events can be estimated.
Stochastic population forecasts

- Forecasts represent median of future distribution.
- Percentiles allow information about uncertainty.
- Prediction intervals with specified probability coverage for population size and all derived variables (total fertility rate, life expectancy, old-age dependencies, etc.)
- The probability of future events can be estimated.
- Economic planning is better based on prediction intervals than point forecasts.
Stochastic population forecasts

- Forecasts represent median of future distribution.
- Percentiles allow information about uncertainty.
- Prediction intervals with specified probability coverage for population size and all derived variables (total fertility rate, life expectancy, old-age dependencies, etc.)
- The probability of future events can be estimated.
- Economic planning is better based on prediction intervals than point forecasts.
- Stochastic models allow true policy analysis.
Demographic growth-balance equation

\[ P_{t+1}(x + 1) = P_t(x) - D_t(x, x + 1) + G_t(x, x + 1) \]
\[ P_{t+1}(0) = B_t - D_t(B, 0) + G_t(B, 0) \]

\( x = 0, 1, 2, \ldots \)
Demographic growth-balance equation

\[ P_{t+1}(x+1) = P_t(x) - D_t(x, x+1) + G_t(x, x+1) \]

\[ P_{t+1}(0) = B_t - D_t(B, 0) + G_t(B, 0) \]

\[ x = 0, 1, 2, \ldots \]

\[ P_t(x) = \text{population of age } x \text{ at 1 January, year } t \]

\[ B_t = \text{births in calendar year } t \]

\[ D_t(x, x+1) = \text{deaths in calendar year } t \text{ of persons aged } x \text{ at the beginning of year } t \]

\[ D_t(B, 0) = \text{infant deaths in calendar year } t \]

\[ G_t(x, x+1) = \text{net migrants in calendar year } t \text{ of persons aged } x \text{ at the beginning of year } t \]

\[ G_t(B, 0) = \text{net migrants of infants born in calendar year } t \]
Key ideas

- Build a stochastic functional model for each of mortality, fertility and net migration.
Key ideas

- Build a stochastic functional model for each of mortality, fertility and net migration.
- Treat all observed data as functional (i.e., smooth curves of age) rather than discrete values.
Key ideas

- Build a \textit{stochastic functional model} for each of mortality, fertility and net migration.
- Treat all observed data as \textit{functional} (i.e., smooth curves of age) rather than discrete values.
- Use the models to simulate future \textit{sample paths} of all components giving the entire age distribution at every year into the future.
Key ideas

- Build a *stochastic functional model* for each of mortality, fertility and net migration.
- Treat all observed data as *functional* (i.e., smooth curves of age) rather than discrete values.
- Use the models to *simulate future sample paths* of all components giving the entire age distribution at every year into the future.
- Compute future births, deaths, net migrants and populations from simulated rates.
Key ideas

- Build a *stochastic functional model* for each of mortality, fertility and net migration.
- Treat all observed data as *functional* (i.e., smooth curves of age) rather than discrete values.
- Use the models to simulate future sample paths of all components giving the entire age distribution at every year into the future.
- Compute future births, deaths, net migrants and populations from simulated rates.
- Combine the results to get *age-specific stochastic population forecasts*. 
The available data

In most countries, the following data are available:

\[ P_t(x) = \text{population of age } x \text{ at } 1 \text{ January, year } t \]

\[ E_t(x) = \text{population of age } x \text{ at } 30 \text{ June, year } t \]

\[ B_t(x) = \text{births in calendar year } t \text{ to females of age } x \]

\[ D_t(x) = \text{deaths in calendar year } t \text{ of persons of age } x \]
The available data

In most countries, the following data are available:

\begin{align*}
P_t(x) &= \text{population of age } x \text{ at 1 January, year } t \\
E_t(x) &= \text{population of age } x \text{ at 30 June, year } t \\
B_t(x) &= \text{births} \text{ in calendar year } t \text{ to females of age } x \\
D_t(x) &= \text{deaths} \text{ in calendar year } t \text{ of persons of age } x
\end{align*}

From these, we can estimate:

\begin{itemize}
  \item \( m_t(x) = D_t(x)/E_t(x) \) = central death rate in calendar year \( t \);
\end{itemize}
The available data

In most countries, the following data are available:

- \( P_t(x) \) = population of age \( x \) at 1 January, year \( t \)
- \( E_t(x) \) = population of age \( x \) at 30 June, year \( t \)
- \( B_t(x) \) = births in calendar year \( t \) to females of age \( x \)
- \( D_t(x) \) = deaths in calendar year \( t \) of persons of age \( x \)

From these, we can estimate:

- \( m_t(x) = D_t(x)/E_t(x) \) = central death rate in calendar year \( t \);
- \( f_t(x) = B_t(x)/E_{tF}(x) \) = fertility rate for females of age \( x \) in calendar year \( t \).
Australia’s start-of-year population

Australia: female population (1921–2004)
Mortality rates

Australia: male death rates (1921–2003)
Net migration

We need to *estimate* migration data based on difference in population numbers after adjusting for births and deaths.
Net migration

We need to estimate migration data based on difference in population numbers after adjusting for births and deaths.

Demographic growth-balance equation

\[ G_t(x, x+1) = P_{t+1}(x+1) - P_t(x) + D_t(x, x+1) \]
\[ G_t(B, 0) = P_{t+1}(0) - B_t + D_t(B, 0) \]

\( x = 0, 1, 2, \ldots \)
Net migration

We need to estimate migration data based on difference in population numbers after adjusting for births and deaths.

Demographic growth-balance equation

\[ G_t(x, x + 1) = P_{t+1}(x + 1) - P_t(x) + D_t(x, x + 1) \]

\[ G_t(B, 0) = P_{t+1}(0) - B_t + D_t(B, 0) \]

\( x = 0, 1, 2, \ldots \)

Note: “net migration” numbers also include errors associated with all estimates. i.e., a “residual”.
Net migration

Australia: male net migration (1922–2003)
Net migration

Component models

- Data: age/sex-specific mortality rates, fertility rates and net migration.
Component models

- **Data**: age/sex-specific mortality rates, fertility rates and net migration.
- **Models**: Five functional time series models for mortality (M/F), fertility and net migration (M/F) assuming independence between components.
Stochastic population forecasts using FDM

Component models

- Data: age/sex-specific mortality rates, fertility rates and net migration.
- Models: Five functional time series models for mortality (M/F), fertility and net migration (M/F) assuming independence between components.
- For each component:

\[
\begin{align*}
y_t(x) &= s_t(x) + \sigma_t(x)\varepsilon_{t,x} \\
s_t(x) &= \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \varphi_k(x) + e_t(x)
\end{align*}
\]
Let $g_\lambda(u) = \begin{cases} \log(u) & \lambda = 0; \\ \frac{x^\lambda - 1}{\lambda} & \lambda > 0. \end{cases}$

- **Mortality rates:**
  
  \[ y_t(x_i) = g_0(m_t(x_i)) \text{ where } m_t(x_i) = \text{empirical mortality rate at age } x_i. \]
Let $g_\lambda(u) = \begin{cases} 
\log(u) & \lambda = 0; \\
\frac{x^\lambda - 1}{\lambda} & \lambda > 0.
\end{cases}

- **Mortality rates:**
  \[ y_t(x_i) = g_0(m_t(x_i)) \text{ where } m_t(x_i) = \text{empirical mortality rate at age } x_i. \]

- **Fertility rates:**
  \[ y_t(x_i) = g_{0.45}(p_t(x_i)) \text{ where } p_t(x_i) = \text{empirical fertility rate at age } x_i. \]
Functional time series

Let \( g_\lambda(u) = \begin{cases} 
\log(u) & \lambda = 0; \\
\frac{u^{\lambda-1}}{\lambda} & \lambda > 0.
\end{cases} \)

- **Mortality rates:**
  \( y_t(x_i) = g_0(m_t(x_i)) \) where \( m_t(x_i) = \) empirical mortality rate at age \( x_i \).

- **Fertility rates:**
  \( y_t(x_i) = g_{0.45}(p_t(x_i)) \) where \( p_t(x_i) = \) empirical fertility rate at age \( x_i \).

- **Net migration:**
  \( y_t(x_i) = \) empirical net migration at age \( x_i \).
Mortality: female

Mortality: female

Main effects

Interaction

Beta 1

Beta 2
Mortality: female

Australia: forecast female log death rates (2004, 2023)
Stochastic population forecasts using FDM

Mortality: female

Australia: forecast female log death rates (2004, 2023)
Stochastic population forecasts using FDM

Fertility

Australia fertility rates (1921–2003)

Age

Fertility rate

Fertility rates for Australia from 1921 to 2003, showing variability across different age groups.
Fertility

Main effects

Interaction

Beta 1

Beta 2
Fertility

Australia: forecast fertility rates (2004, 2023)
Migration: male

Stochastic population forecasts using FDM

Migration: male

Main effects

Interaction

Beta 1

Beta 2
Migration: male

Migration: male

Male net migration

Age
Number people

-2000 -1000 0 1000 2000 3000

0 20 40 60 80 100
Migration: female

Migration: female

Main effects

Interaction

Beta 1

Beta 2

Time

Time
Migration: female

Migration: female

Female net migration

Number people

Age
Simulation

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]

\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

- For each of \( m^F_t(x) \), \( m^M_t(x) \), \( f_t(x) \), \( G^F_t(x, x+1) \), and \( G^M_t(x, x+1) \):
Simulation

\[ y_t(x) = s_t(x) + \sigma_t(x)\varepsilon_{t,x} \]
\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

For each of \( m^F_t(x), m^M_t(x), f_t(x), G^F_t(x, x + 1), \) and \( G^M_t(x, x + 1) \):

- Generate random sample paths of \( \beta_{t,k} \) for \( t = n + 1, \ldots, n + h \) conditional on \( \beta_{1,k}, \ldots, \beta_{n,k} \).
Simulation

\[ y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x} \]
\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

- For each of \( m^F_t(x) \), \( m^M_t(x) \), \( f_t(x) \), \( G^F_t(x, x + 1) \), and \( G^M_t(x, x + 1) \):
  - Generate random sample paths of \( \beta_{t,k} \) for \( t = n + 1, \ldots, n + h \) conditional on \( \beta_{1,k}, \ldots, \beta_{n,k} \).
  - Generate random values for \( e_t(x) \) and \( \varepsilon_{t,x} \).
Simulation

\[ y_t(x) = s_t(x) + \sigma_t(x) \varepsilon_{t,x} \]
\[ s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

- For each of \( m^F_t(x) \), \( m^M_t(x) \), \( f_t(x) \), \( G^F_t(x, x+1) \), and \( G^M_t(x, x+1) \):
  - Generate random sample paths of \( \beta_{t,k} \) for \( t = n + 1, \ldots, n + h \) conditional on \( \beta_{1,k}, \ldots, \beta_{n,k} \).
  - Generate random values for \( e_t(x) \) and \( \varepsilon_{t,x} \).

- Use simulated rates to generate \( B_t(x) \), \( D^F_t(x, x+1) \), \( D^M_t(x, x+1) \) for \( t = n + 1, \ldots, n + h \), assuming deaths and births are Poisson.
Simulation

Demographic growth-balance equation used to get population sample paths.

**Demographic growth-balance equation**

\[
P_{t+1}(x + 1) = P_t(x) - D_t(x, x + 1) + G_t(x, x + 1)
\]

\[
P_{t+1}(0) = B_t - D_t(B, 0) + G_t(B, 0)
\]

\[x = 0, 1, 2, \ldots\]
Simulation

Demographic growth-balance equation used to get population sample paths.

Demographic growth-balance equation

\[ P_{t+1}(x + 1) = P_t(x) - D_t(x, x + 1) + G_t(x, x + 1) \]
\[ P_{t+1}(0) = B_t - D_t(B, 0) + G_t(B, 0) \]

\( x = 0, 1, 2, \ldots \)

- 10000 sample paths of population \( P_t(x) \), deaths \( D_t(x) \) and births \( B_t(x) \) generated for \( t = 2004, \ldots, 2023 \) and \( x = 0, 1, 2, \ldots \).
Simulations

Demographic growth-balance equation used to get population sample paths.

Demographic growth-balance equation

\[
\begin{align*}
  P_{t+1}(x + 1) &= P_t(x) - D_t(x, x + 1) + G_t(x, x + 1) \\
  P_{t+1}(0) &= B_t - D_t(B, 0) + G_t(B, 0)
\end{align*}
\]

\[x = 0, 1, 2, \ldots\]

- 10000 sample paths of population \(P_t(x)\), deaths \(D_t(x)\) and births \(B_t(x)\) generated for \(t = 2004, \ldots, 2023\) and \(x = 0, 1, 2, \ldots\).

- This allows the computation of the empirical forecast distribution of any demographic quantity that is based on births, deaths and population numbers.
**Forecasts of life expectancy at age 0**

<table>
<thead>
<tr>
<th>Year</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>70</td>
</tr>
<tr>
<td>1980</td>
<td>75</td>
</tr>
<tr>
<td>2000</td>
<td>80</td>
</tr>
<tr>
<td>2020</td>
<td>85</td>
</tr>
</tbody>
</table>

**Forecast female life expectancy**

**Forecast male life expectancy**
### Forecasts of TFR

<table>
<thead>
<tr>
<th>Year</th>
<th>TFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920</td>
<td>1500</td>
</tr>
<tr>
<td>1940</td>
<td>2000</td>
</tr>
<tr>
<td>1960</td>
<td>2500</td>
</tr>
<tr>
<td>1980</td>
<td>3000</td>
</tr>
<tr>
<td>2000</td>
<td>3500</td>
</tr>
<tr>
<td>2020</td>
<td>4000</td>
</tr>
</tbody>
</table>

**Forecast Total Fertility Rate**

![Graph showing forecast total fertility rate](image-url)
Forecast population pyramid for 2023, along with 80% prediction intervals. Dashed: actual population pyramid for 2003.
Twenty-year forecasts of total population along with 80% and 95% prediction intervals. Dashed lines show the ABS (2003) projections, series A, B and C.
Stochastic population forecasts using FDM

Population forecasts

Twenty-year forecasts of total population along with 80% and 95% prediction intervals. Dashed lines show the ABS (2003) projections, series A, B and C.
Old-age dependency ratio forecasts

<table>
<thead>
<tr>
<th>Year</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920</td>
<td>0.10</td>
</tr>
<tr>
<td>1940</td>
<td>0.15</td>
</tr>
<tr>
<td>1960</td>
<td>0.20</td>
</tr>
<tr>
<td>1980</td>
<td>0.25</td>
</tr>
<tr>
<td>2000</td>
<td>0.30</td>
</tr>
<tr>
<td>2020</td>
<td></td>
</tr>
</tbody>
</table>
Advantages of stochastic simulation approach

- Functional data analysis provides a way of forecasting age-specific mortality, fertility and net migration.
Advantages of stochastic simulation approach

- Functional data analysis provides a way of forecasting age-specific mortality, fertility and net migration.

- Stochastic age-specific cohort-component simulation provides a way of forecasting many demographic quantities with prediction intervals.
Advantages of stochastic simulation approach

- Functional data analysis provides a way of forecasting age-specific mortality, fertility and net migration.

- Stochastic age-specific cohort-component simulation provides a way of forecasting many demographic quantities with prediction intervals.

- No need to select combinations of assumed rates.
Advantages of stochastic simulation approach

- Functional data analysis provides a way of forecasting age-specific mortality, fertility and net migration.
- Stochastic age-specific cohort-component simulation provides a way of forecasting many demographic quantities with prediction intervals.
- No need to select combinations of assumed rates.
- True prediction intervals with specified coverage for population and all derived variables (TFR, life expectancy, old-age dependencies, etc.)
Extensions

- We intend extending this to Australian states, and to other countries.

Software and papers:
www.robhyndman.info
Extensions

- We intend extending this to Australian states, and to other countries.
- We intend extending the individual models to:
  - allow dependencies between sexes
  - allow cohort effects
  - allow interaction between fertility and migration
Extensions

- We intend extending this to Australian states, and to other countries.
- We intend extending the individual models to:
  - allow dependencies between sexes
We intend extending this to Australian states, and to other countries.

We intend extending the individual models to:

- allow dependencies between sexes
- allow cohort effects
Extensions

- We intend extending this to Australian states, and to other countries.
- We intend extending the individual models to:
  - allow dependencies between sexes
  - allow cohort effects
  - allow interaction between fertility and migration?
We intend extending this to Australian states, and to other countries.

We intend extending the individual models to:
- allow dependencies between sexes
- allow cohort effects
- allow interaction between fertility and migration?
Extensions

- We intend extending this to Australian states, and to other countries.
- We intend extending the individual models to:
  - allow dependencies between sexes
  - allow cohort effects
  - allow interaction between fertility and migration?

Software and papers:

[www.robhyndman.info]