Demographic Methods

Monica Alexander

Week 4: Stable populations, population projection

Overview

- Stable populations
- The renewal equation and Euler-Lotka
- Population projection
- Leslie matrices
- Population momentum

Stable populations

Preliminaries

- So far we've looked at various indicators for mortality and fertility
- The most disaggegated form of these are age-specific rates, especially $l(x)$ and $f(x)$
- Today, mostly about how these interact and affect long term growth rates
- NRR recap: what is it, what's the formula
  - Discrete: $NRR = \sum_{15}^{49} {}_nF_x \cdot {}_nL_x \cdot f_{fab}$
  - Continuous: $NRR = \int f(a)l(a)da = \int \phi(a)da$ where $\phi(x)$ is the 'net maternity function'

Preliminaries

- Mortality and fertility rates are changing all the time
- But changes may be small (especially over the short term)
- Every population has inherent characteristics based on the current mortality and fertility rates
- Useful to compare populations based on these inherent characteristics, project forward, etc

Definitions

Stable population: the proportions in each age group do not change over time

- constant mortality rates
- constant growth rate in number of births (and thus every age group)

Special case:

Stationary population: the populations in each age group do not change over time

- constant mortality rates
- constant number of births

Illustration

Assume a very simple and frightening life table:

age x	l_x
0	1
1	0.6
2	0.4
3	0.2
4	0.05
5	0

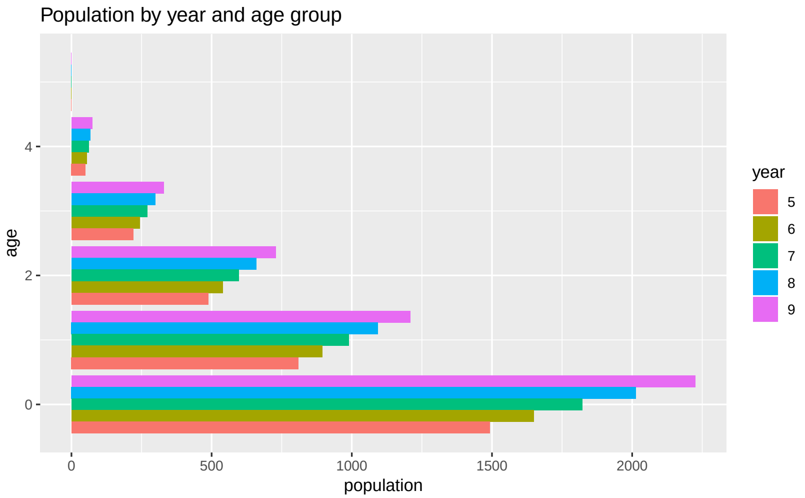
Assume births are growing at a constant annualized growth rate r i.e.

$$B(t) = B(0)e^{rt}$$

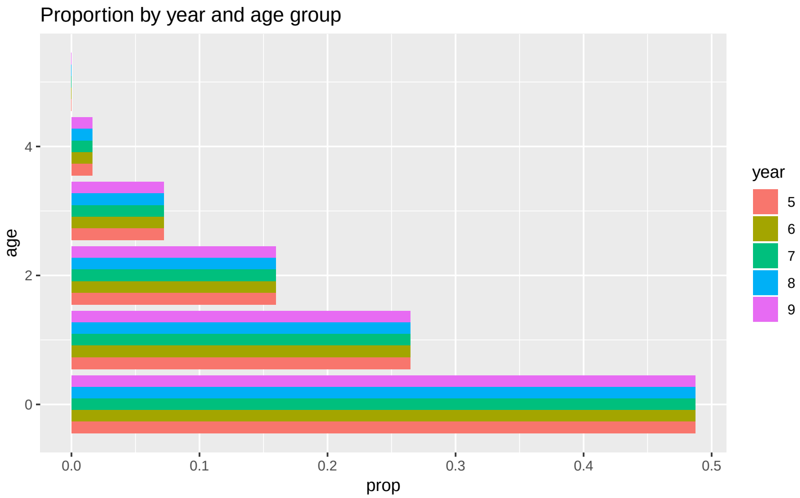
Illustration

age x	l_x	$P(0)$	$P(1)$	$P(2)$	$P(3)$	$P(4)$	$P(5)$
0	1	1000	$1000e^{r}$	$1000e^{2r}$	$1000e^{3r}$	$1000e^{4r}$	$1000e^{5r}$
1	0.6		600	$600e^{r}$	$600e^{2r}$	$600e^{3r}$	$600e^{4r}$
2	0.4			400	$400e^{r}$	$400e^{2r}$	$400e^{3r}$
3	0.2				200	$200e^{r}$	$200e^{2r}$
4	0.05					50	$50e^{r}$
5	0						0

Illustration



Illustration



Renewal equation

Let's go back to thinking about (female) births in relation to the size of the cohort of mothers.

We know that the total number of births today is equal to the sum of each age-specific fertility rate multiplied by the number of women at each age:

$$B(t) = \int_0^\infty N(a,t)f(a)da$$

But the number of people at a particular age can be expressed as births a years ago and the probability of survival to age a:

$$N(a,t) = B(t-a)l(a)$$

So then we have

$$B(t) = \int_0^\infty B(t-a)l(a)f(a)da$$

Renewal equation

But in a stable population, we know

$$B(t) = Be^{rt}$$

$$B(t-a) = Be^{r(t-a)}$$

So we get

$$Be^{rt} = \int_0^\infty Be^{r(t-a)} l(a) f(a) da$$

Euler-Lotka equation

Cancelling out terms we get

$$1 = \int_0^{\infty} e^{-ra} l(a) f(a) da$$

Euler (1760) - Lotka (1911) equation.

Based on the age-structure and births of female population, you can estimate the extent to which a population is growing.

Think about the right-hand side as a function of r, $Y(r)$.

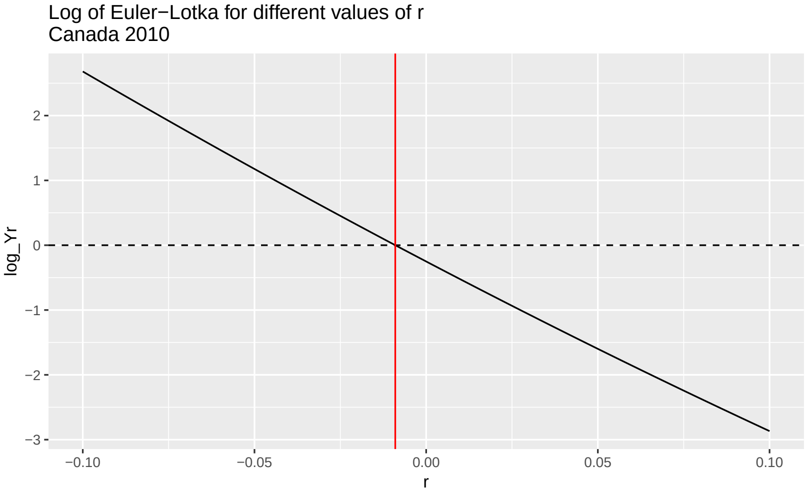
- strictly decreasing function of r
- there will always be a unique r such that $Y(r) = 1$. This value is called:
  - the intrinsic growth rate
  - Lotka's r
  - the Malthusian parameter (related to unrestricted exponential growth)

Notes

$$1 = \int_0^\infty e^{-ra} l(a) f(a) da$$

- what is $l(a)f(a)$?
- what does this tell us about r?

Solving for r



Other stuff about $Y(r)$

$$Y(r) = \int_0^\infty e^{-ra} l(a) f(a) da$$

If $Y(0) = 0$ then $NRR = 1$

Look at the slope of the log of $Y(r)$:

$$\frac{d}{dr} \log Y(r) = \frac{\int_0^\infty -a e^{-ra} l(a) f(a) da}{\int_0^\infty e^{-ra} l(a) f(a) da}$$

At $r = 0$, this is minus the mean age at child-bearing (μ).

Other stuff about $Y(r)$

Do a Taylor Series expansion of $\log Y(r)$ around 0, and you get:

$$0 = \log(\mathrm{NRR}) - \mu r + \dots$$

So

$$r \approx \frac{\log(NRR)}{\mu}$$

i.e. we can approximate r if we know the NRR and make some assumption about the mean age at childbearing (/generation length)

Stable age distributions

We know the number of people in each age group in a stable population is

$$N(a,t) = B(t)e^{-ra}l(a)$$

So the proportion of people in age group a is

$$c(a,t) = \frac{B(t)}{\int N(a,t)da}e^{-ra}l(a) = be^{-ra}l(a)$$

Where b is the birth rate (constant over time). All proportions must equal one so

$$1 = \int c(a,t)da = b\int e^{-ra}l(a)da$$

Stable age distributions

So in a stable population, the birth rate is:

$$b = \frac{1}{\int e^{-ra} l(a) da}$$

Expresses the birth rate in terms of the growth rate and survival in a population

Stationary population identity

When $r = 0$, the population is stationary. And

$$b = \frac{1}{\int e^{-ra} l(a) da}$$

implies

$$b = \frac{1}{\int l(a) da}$$

- what is the denominator?
- what is the birth rate in Canada?

Population projection

Cohort component projection framework

Cannan (1895).

Elaborating on the demographic accounting identity to be disaggregated by age (and sex).

- model the age distribution of a population, not just the size
- model the components of population growth, not just overall growth.

The idea is to:

1. project each age group forward, based on underlying mortality rates
2. calculate the number of births based on projected fertility rates
3. adjust for migration, if applicable

Cohort Component Projection in the wild

Estimating and projecting subnational populations of women of reproductive age in Kenya

- important population to know about (fertility, family planning, maternal mortality)
- age patterns and components of change are important / interesting (not just population size)

Issue: data

- decennial censuses give us (reasonably reliable) counts by age and area
- mortality data? (decent at national level)
- migration data? (some available through surveys)

CCP goes Bayes

$$\eta_{r,c,a+1} = \eta_{r,c,a} \cdot (1 - \rho_{r,c,a}) + \phi_{r,c,a}$$

with for region r and birth cohort c:

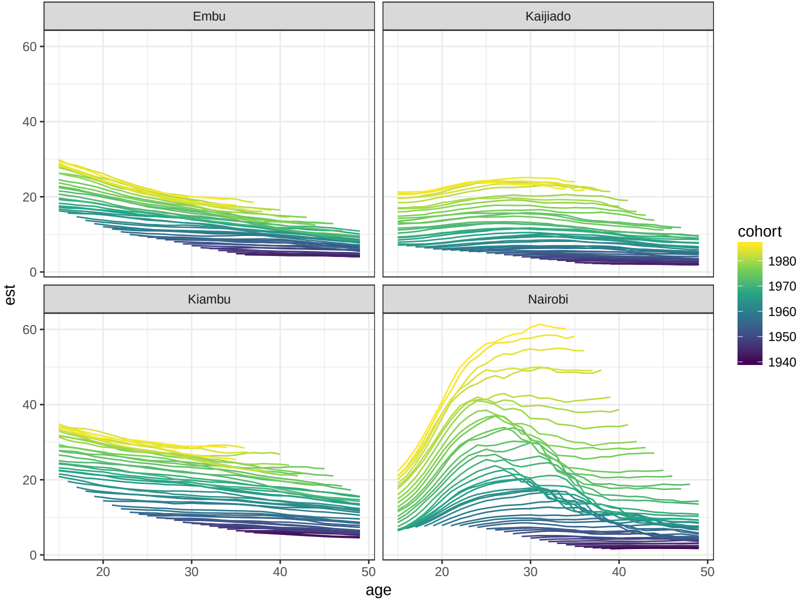
- $\eta_{r,c,a}$ the number of women of age $a = 15, 16, \ldots, 49$,
- $\rho_{r,c,a}$ the expected proportion of women who die between age a and $a + 1$
- $\phi_{r,c,a}$ the expected number of net-migrants into/out of region r at age a

CCP goes Bayes

$$\eta_{r,c,a+1} = \eta_{r,c,a} \cdot (1 - \rho_{r,c,a}) + \phi_{r,c,a}$$

- some observations for η
- model on ρ (SVD based on national schedules)
- model on ϕ (based on patterns observed in surveys)
- incorporate different sorts of uncertainty in different data sources

Example results



Leslie matrices

Leslie matrices

Leslie (1945). Cohort component projection viewed as a Markov process.

- Define a N by N matrix, where each cell refers to an age group and N is the total number of age groups you have.
- Individuals start in the column age and end up in the row age (this is the opposite to usual, sorry folks)
- Multiply a start population age vector by a Leslie matrix to get the population for the next time step.

Note: the Leslie matrix is mostly structural zeroes (why?)

	0-5	5-10	10-15	15-20	20-25
0-5	X	X	X	X	X
5-10	X	0	0	0	0
10-15	0	X	0	0	0
15-20	0	0	X	0	0
20-25	0	0	0	X	0

Leslie matrix

- subdiagonals are mortality
- first row is fertility

$$\begin{bmatrix} \text{kids} & \text{kids} & \text{kids} & \dots & \text{kids} \\ \text{survivors} & 0 & 0 & \dots & 0 \\ 0 & \text{survivors} & 0 & \dots & 0 \\ \hdotsfor{5} \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Leslie matrix subdiagonals

The ratio of people surviving to this age group compared to the number in the last age group, i.e.

$$\frac{{}_nL_{x+n}}{{}_nL_x}$$

$$\begin{bmatrix} \text{kids} & \text{kids} & \text{kids} & \dots & \text{kids} \\ \frac{{}_5L_{10}}{{}_5L_0} & 0 & 0 & \dots & 0 \\ 0 & \frac{{}_5L_{15}}{{}_5L_{10}} & 0 & \dots & 0 \\ \hdotsfor{5} \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Leslie matrix first row

Something like fertility rates, but need to account for mortality (of both women and babies).

$$ {}_nL_0 \cdot \frac{1}{2}({}_nF_x + {}_nF_{x+n} \cdot \frac{{}_nL_{x+n}}{{}_nL_x}) \cdot f_{fab} $$

Note that fertility usually assumed to be zero for ages less than 15, so the first couple of elements of the first row will be 0.

Canada 2010

$$\begin{bmatrix} 0 & 0 & 0.0137 & \ldots & 0 \\ 0.9995 & 0 & 0 & \ldots & 0 \\ 0 & 0.9996 & 0 & \ldots & 0 \\ \hdotsfor{5} \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

We can reconstruct NRR from a Leslie matrix

Note we can get the NRR from a Leslie matrix:

$$NRR = \sum A_{1,j(x)} \frac{{}_nLx}{{}_nL_0}$$

Project population forward

Call Leslie matrix A. Say we have a vector of population counts by age group, $K(t)$. Then

$$K(t+5) = AK(t)$$

but also, if we assume stable rates:

$$K(t+10) = A \cdot AK(t) = A^2K(t)$$

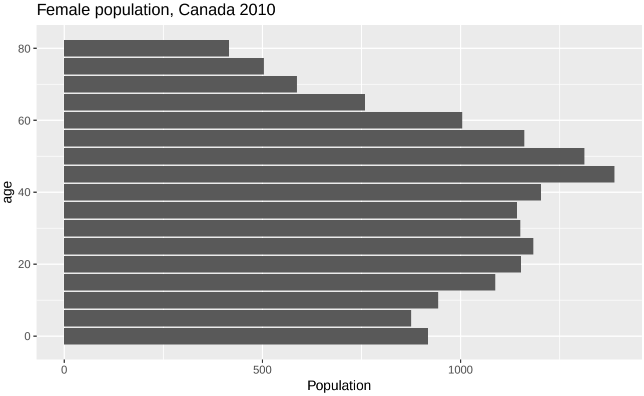
and in general

$$K(t+m \cdot n) = A^mK(t)$$

Project Canada's population in 2010 to 2050

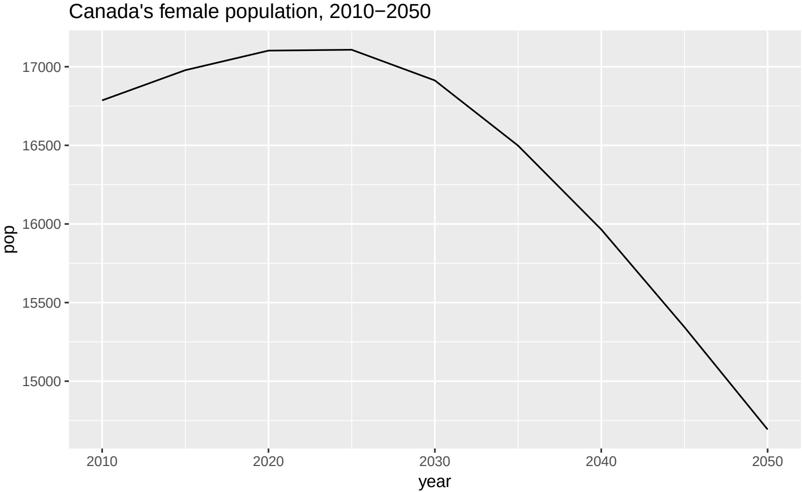
Our starting population vector:

$$K(2010) = [917, 876, 944, 1088, 1152, \ldots 503, 416]$$



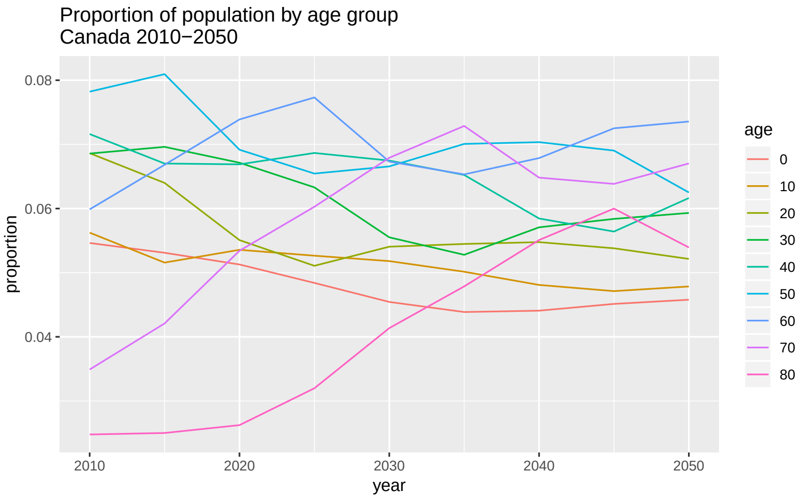
Do you think population size will increase or decrease in 40 years?

Project Canada's population in 2010 to 2050

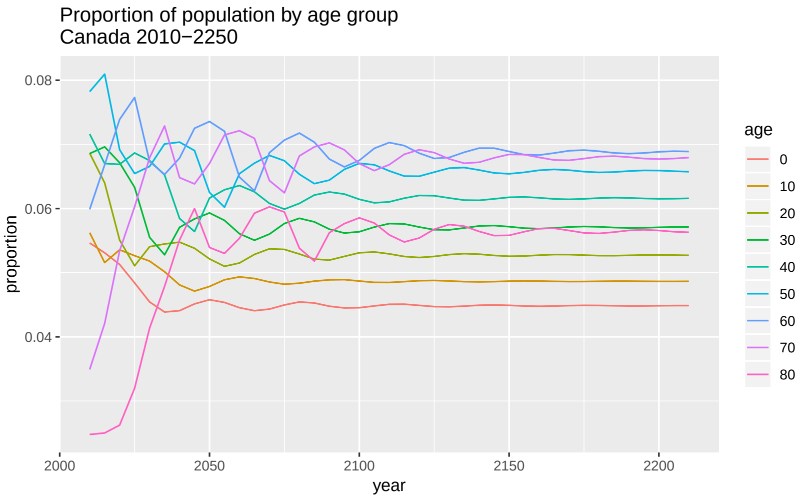


What will the proportions by age look like?

Project Canada's population in 2010 to 2050



Project Canada's population in 2010 to 2250



Some characteristics of A

Think about the special case when

$$AK = \lambda K$$

- what are λ and K?
- λ is an eigenvalue, and K is an eigenvector

Some characteristics of A

If we do an eigen-decomposition of A:

$$A = U \Delta U^{-1}$$

- The leading eigenvalue is e^{nr}
- The first right eigenvector is the stable age distribution

Some characteristics of A

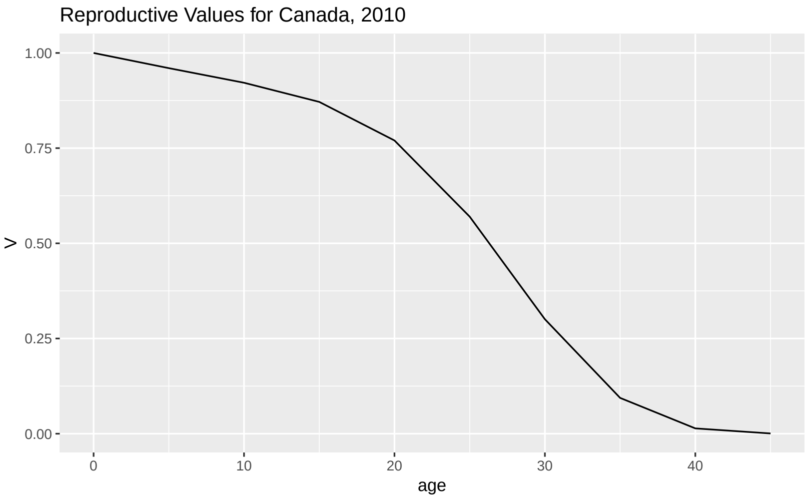
$$A = U\Delta U^{-1}$$

It turns out that V^*, which is U^{-1} multiplied so the first element is 1, are Fisher's reproductive values:

$$v(x) = \int_x = e^{-r(x-y)} \frac{l(y)}{l(x)} f(x) dy$$

which measures the (discounted) number of future female children that will be born to a woman aged x (their contribution to future population growth).

Reproductive values



Reproductive values: historical context

- First introduced by Ronald Fisher in his book 'Fundamental theorem of Natural Selection'
- Fisher interested in understanding reproduction in order to increase birth rates of 'upper class'
- He was a staunch eugenicist and believed the poor should be bred out

FUNDAMENTAL THEOREM OF NATURAL SELECTION 29

case among a people by no means precocious in reproduction, it would be surprising if, in a state of society entailing marriage at or soon after puberty, the age of maximum reproductive value should fall at any later age than twelve. In the Australian data, the value at birth is lower, partly by reason of the effect of an increasing population in setting a lower value upon remote children and partly because of the risk of death before the reproductive age is reached. The value shown is probably correct, apart from changes in the rate since 1911, for such a purpose as assessing how far it is worth while to give assistance to immigrants in respect of infants (though of course, it takes no account of the factor of eugenic quality), for such infants will usually emigrate with their parents; but it is overvalued from the point of view of Natural Selection to a considerable extent, owing to the capacity of the parents to replace a baby lost during lactation. The

Properties of A

We know

$$v(x) = \int_x = e^{-r(x-y)} \frac{l(y)}{l(x)} f(x) dy$$

We also know the stable age structure (normalized so b = 1) is

$$c(x) = l(x) e^{-rx}$$

We get $v(x)$ from the first left eigenvector and $c(x)$ from the first right eigenvector. It turns out that

$$\int_x v(x) c(x) dx = \mu_r$$

i.e. the mean age at childbearing in the stable population.

So the first element of $V^* U$ (where $V*$ is U^{-1} multiplied so that the first element is 1) multiplied by the age interval (in our case, 5) is μ_r. For Canada 2010, this is 30.6 years.

Population momentum

Background

- We've seen that even if we project constant rates forward in time, there are some residual patterns based on past trends
- What would happen if the NRR in Nigeria suddenly dropped to 1?
- What would happen if the NRR in Germany suddenly increased to 1?

Even though $r = 0$, there would still be some residual growth / decline.

This idea is called **population momentum**

Formalization

Stable momentum (Keyfitz)

- Assume there is a sudden change in fertility with no change in mortality, such that NRR $= 1$.
- Assume these new rates persist, so we end up with a stationary distribution.
- We can compare the size of the ultimate population and population before the drop.

Keyfitz scenario

If $B(U)$ is the ultimate number of births and $B(-\epsilon)$ is births before the drop, Keyfitz showed that

$$B(U) \approx \frac{B(-\epsilon)}{\sqrt{NRR}}$$

where NRR is pre-drop. We can convert this to population size K using the stationary population identity:

$$\begin{aligned}
b(U)K(U) &\approx \frac{b(-\epsilon)K(-\epsilon)}{\sqrt{NRR}} \\
\frac{K(U)}{K(-\epsilon)} &\approx \frac{b(-\epsilon)}{b(U)\sqrt{NRR}} \\
&\approx \frac{b(-\epsilon)e_0}{\sqrt{NRR}}
\end{aligned}$$

For Canada in 2010 this ratio was 0.856.

Population momentum, Canada