**Introduction to Demography**

February 1, 2021

To study changes in populations with stage structure requires demographic theory. Demographic theory links survival probability and reproductive rate at each age to the growth and composition of a population.

Today we'll introduce some of the concepts and parameters important in studying the age structure and growth of a population.

We'll focus on the demography of Canada using primarily census data published by Statistics Canada.

Total Population Size: Consider a population measured at particular points in time, for instance every five years during a census. One can track the population size from census to census to get a historical perspective on population growth:
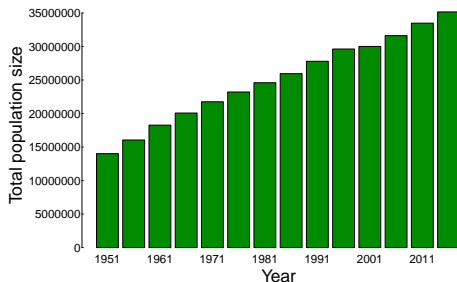
This shows that Canada has grown in the recent past. For instance, between the 1990 and 1995 censuses, the population grew by 1.3% per year (although about half of this growth was due to migration).

In studying a natural population, there are many complications that we are going to ignore. For instance, we are going to ignore emigration (out of the population) and immigration (into the population). Furthermore, males and females generally differ in their demographic parameters but we're going to focus on females.

We'll use five year census points and base our analysis on the 1991 census, so that we may compare our projections to future census years.

```
1991 Canadian Census
********************
Age          Number
Class        Females
********************
0-4          945400
5-9          909700
10-14        901500
15-19        895900
20-24        976800
25-29       1168800
30-34       1221800
35-39       1133000
40-44       1031500
45-49        807500
50-54        656600
55-59        610100
60-64        599400
65-69        571800
70-74        461700
75-79        362600
80-84        237900
85+          197100
********************
Total      13689100
```
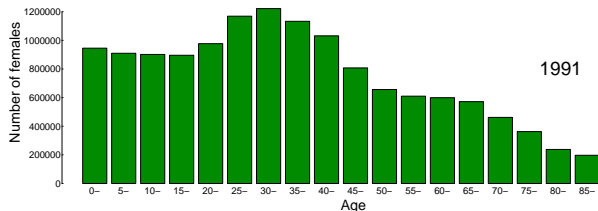
Age distribution: A population can be described by how many individuals there currently are in each age class.

This census data groups individuals into five year age classes.

By convention, the number of females at time t in age class x is $n_{x,t}$.

Shown graphically:

```
1991 Canadian Census
********************
Age              Number
Class            Females
********************
0-4              945400
5-9              909700
10-14            901500
15-19            895900
20-24            976800
25-29           1168800
30-34           1221800
35-39           1133000
40-44           1031500
45-49            807500
50-54            656600
55-59            610100
60-64            599400
65-69            571800
70-74            461700
75-79            362600
80-84            237900
85+              197100
********************
Total          13689100
```
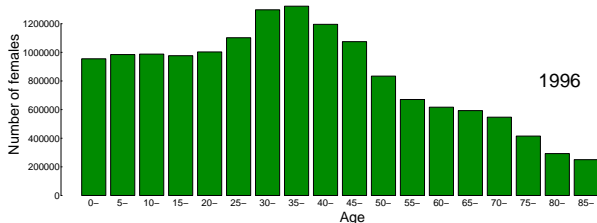
Age distribution: A population can be described by how many individuals there currently are in each age class.

This census data groups individuals into five year age classes.

By convention, the number of females at time t in age class x is $n_{x,t}$.

Shown graphically:

```
1991 Canadian Census
********************
Age         Number
Class       Females
********************
0-4          945400
5-9          909700
10-14        901500
15-19        895900
20-24        976800
25-29       1168800
30-34       1221800
35-39       1133000
40-44       1031500
45-49        807500
50-54        656600
55-59        610100
60-64        599400
65-69        571800
70-74        461700
75-79        362600
80-84        237900
85+          197100
********************
Total      13689100
```
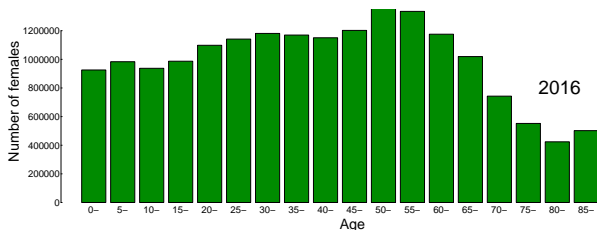
Age distribution: A population can be described by how many individuals there currently are in each age class.

This census data groups individuals into five year age classes.

By convention, the number of females at time t in age class x is $n_{x,t}$.

Shown graphically:

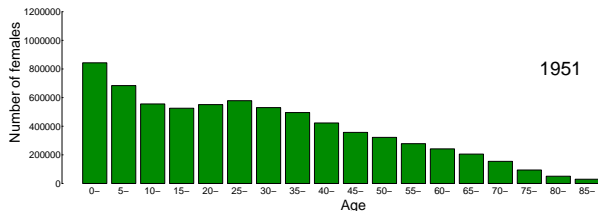| Age Class	Number Females
0-4	945400
5-9	909700
10-14	901500
15-19	895900
20-24	976800
25-29	1168800
30-34	1221800
35-39	1133000
40-44	1031500
45-49	807500
50-54	656600
55-59	610100
60-64	599400
65-69	571800
70-74	461700
75-79	362600
80-84	237900
85+	197100
Total	13689100

1991 Canadian Census

Age distribution: A population can be described by how many individuals there currently are in each age class.
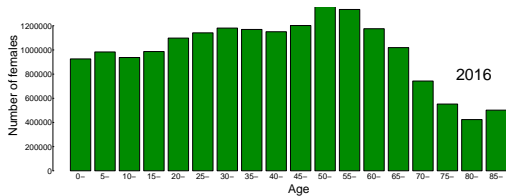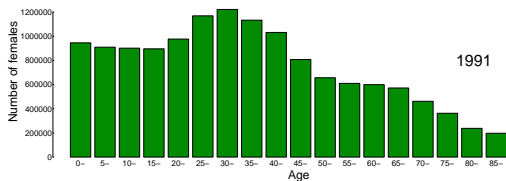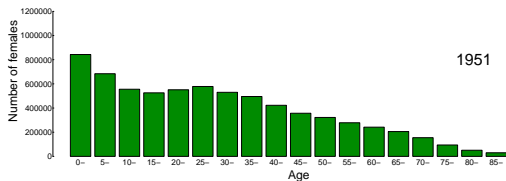
This census data groups individuals into five year age classes.

By convention, the number of females at time t in age class x is $n_{x,t}$.

Shown graphically:

Shown altogether:

```
1991 Canadian Census
********************
Age             Female
Class           Deaths
********************
0-4               1344
5-9                155
10-14              152
15-19              345
20-24              361
25-29              556
30-34              681
35-39              939
40-44             1337
45-49             1659
50-54             2142
55-59             3275
60-64             5032
65-69             7763
70-74             9757
75-79            12949
80-84            14261
85+              27415
********************
Total            90123
```
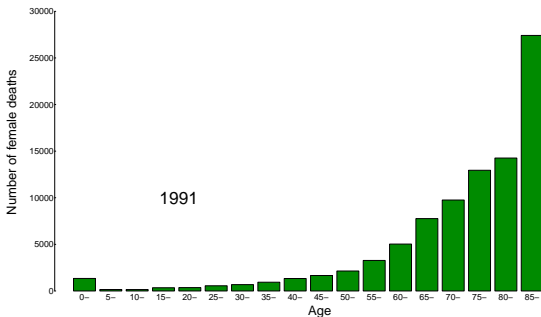
Mortality: Another critical piece of data that we can get from the census is how many individuals die per year, by age class.

Shown graphically:

Generally, the number of deaths in the population is less useful than the proportion of individuals in each age class that die between each census period.

The proportion of individuals in each age class that die between each census period is obtained by multiplying the number that died per year in each age class by the length of the age class and dividing by the total number of individuals in each age class:
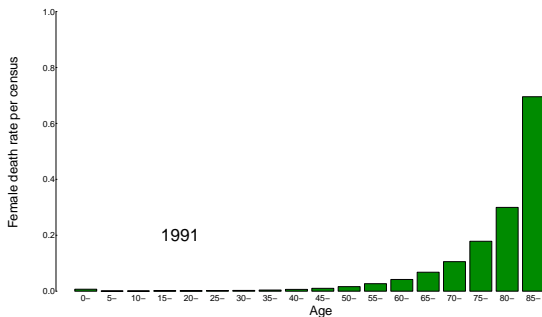
For example, in the third age class, in which there were 901500 females, there were 152 deaths (in the census year). Over five years, this means a female's probability of death would be $5 * 152/901500 = 0.00084$.

The probability of dying from age class x to age class $x + 1$ is denoted by d_x.

Another important parameter is the probability of surviving from age class x to age class $x + 1$, p_x, which is simply $1 - d_x$.

Sometimes, survival data are given from birth to the current age class. The probability of surviving from birth to the beginning of age class x is generally written as l_x.

Fortunately, one can determine l_x from p_x and vice versa, using the fact that:

$$l_x = p_{x-1} l_{x-1}$$

$$l_x = p_1 p_2 \cdots p_{x-1}$$

```
1991 Canadian Census
********************
Age
Class          Births
********************
0-4                 0
5-9                 0
10-14             265
15-19           24180
20-24           80723
25-29          150024
30-34          107560
35-39           33107
40-44            4124
45-49             138
50-54               0
55-59               0
60-64               0
65-69               0
70-74               0
75-79               0
80-84               0
85+                 0
********************
Total          400121
```

Fertility: Finally, we must determine the number of babies that individuals in each age class have.

These figures are for all births. We are following females, however, and only 195916 of these births (48.96%) were girls. *[A male-biased sex ratio at birth is a general phenomenon among humans.]* We'll use this proportion to esimate the number of female offspring born in 1991 to mothers in each age class.

Shown graphically:
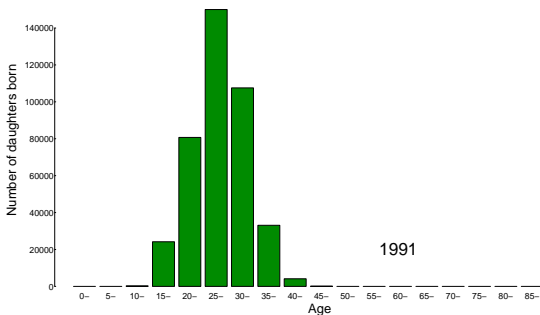
Again, the total number of daughters born is less useful than the number of daughters born per female in each age class during the census period.

This is obtained by multiplying the number of daughters born by the length of the age class and dividing by the total number of females in each age class:
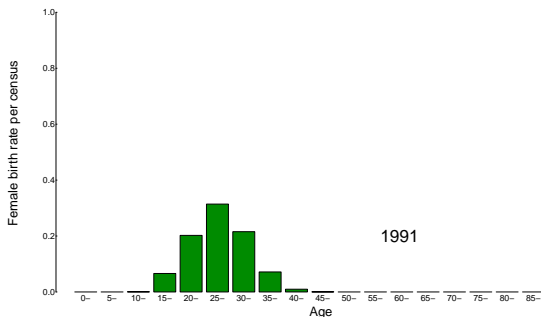
For example, in the fifth age class, in which there were 976800 females, there were 80723 births (in the census year), of which $0.4896 * 80723$ would be expected to be female. Over five years, this means a female would be expected to have $5 * 0.4896 * 80723/976800 = 0.2023$ female offspring.

The number of female offspring per female in age class x is denoted by m_x.

Notice that the total period fertility rate (Σm_x) is only 0.88. That is, even if all females survive from birth until the end of their reproductive period, the number of females born to a cohort of women will not fully replace that cohort.

With these parameters of a population in hand, we can turn next to analysing how a population will change over time.

In summary, we have a whole bunch of data about the Canadian population, from which we can infer:

$n_{x,t}$: number of females at time t in age class x

d_x: probability of dying while in class x (not surviving to class $x+1$)

p_x: probability of surviving from class x to class $x+1$ ($p_x = 1 - d_x$)

l_x: probability of surviving from birth to the beginning of age class x

m_x: number of female offspring per female in age class x

We will assume that individuals censused within an age class are actually at the beginning of that age class.

In particular, the number of individuals within the $0-4$ age class will be the total number of individuals born since the previous census. The number of individuals in the $5-9$ age class will then be those that were in the $0-4$ age class and that survived as $0-$, $1-$, $2-$, $3-$, and $4-$year olds.

Different texts make different assumptions here (e.g., that censused individuals are at the end of their age class).

Step 1: The first step is to determine the transition matrix for the population from census point to census point.

First, the females in age class x reproduce (at rate m_x). The females then move up to age class $x + 1$ if they survive to the next census (with probability p_x).

Therefore, given the population composition at census t, at census $t + 1$:

1. the number of newborns is $\Sigma_t(n_{x,t} \cdot m_x)$

2. the number of individuals in other age classes is $n_{x-1,t} \cdot p_{x-1}$

3. the number of individuals in the last age class (ω) is $n_{\omega-1,t}p_{\omega-1} + n_{\omega,t}p_\omega$, since some individuals live past 90.

Population changes can more easily be described using matrix notation:

$$\vec{n}_{t+1} = L\vec{n}_t$$

where

$$L = \begin{bmatrix} m_1 & m_2 & m_3 & \cdots & m_{\omega-1} & m_\omega \\ p_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & p_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{\omega-1} & p_\omega \end{bmatrix} \qquad \vec{n}_t = \begin{bmatrix} n_{1,t} \\ n_{2,t} \\ n_{3,t} \\ n_{4,t} \\ \vdots \\ n_{\omega,t} \end{bmatrix}$$

L is known as a **Leslie Matrix** for population projection.

We can then proceed through our steps to find the general solution.

Step 2: Determine the ω eigenvalues of the Leslie matrix.

Step 3: Make a diagonal matrix, D, with one eigenvalue in each of the diagonal positions.

Step 4: Determine the eigenvectors associated with each eigenvalue.

Step 5: Make a transformation matrix, A, whose columns are the eigenvectors (placed in the same order as the eigenvalues in matrix D).

Step 6: Write the general solution of the linear equations as:

$$\vec{n}_t = AD^t A^{-1} \vec{n}_0$$

For large matrices, this can be a real mess! So we will use a simpler approximation.

If we place the **leading eigenvalue** (the eigenvalue with largest magnitude) in the first position, D^t will become more and more similar over time to the following matrix:

$$\widetilde{D}^t = \begin{bmatrix} \lambda_1^t & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

This is because, if λ_1 is larger than all the other eigenvalues, λ_1^t will be *much larger* than λ_i^t (for i other than 1), and the influence of these other eigenvalues will become negligible after enough time has passed.

Whether or not this will give a sufficiently accurate approximation depends on how much time has passed and on the difference in magnitude between the leading eigenvalue and the other eigenvalues.

An approximate projection of the population is then:

$$\vec{n}_t \approx A\widetilde{D}^t A^{-1} \vec{n}_0$$

$$\approx \lambda_1^t \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{n1} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \end{bmatrix} \vec{n}_0$$

where $\vec{u}$ is the first column of A and $\vec{v}$ is the first row of A^{-1}.

$\vec{u}$ is the **right eigenvector** associated with λ_1 and $\vec{v}$ is the **left eigenvector** associated with λ_1.

$\begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \end{bmatrix} \vec{n}_0$ is the "adjusted initial population size" (more below).

Once the general solution is found, iteration is simple: just multiply by λ_1 every generation. λ_1, therefore, represents the **long-term growth rate** of the population.

Note: The first column of A is also called the **right eigenvector**, which must solve $L\vec{u} = \lambda_1 \vec{u}$. Similarly, the first row of A^{-1} is called the **left eigenvector**, which must solve $\vec{v}L = \lambda_1 \vec{v}$. Typically when the term "eigenvector" is used, without specifying right or left, it is a reference to the right eigenvector.

Important: To use this approximate solution, the length of the left eigenvector must be adjusted so that $\vec{v} \cdot \vec{u} = 1$.

The left eigenvector contains information about the **reproductive value** of each age class. When multiplied by the initial population vector, it gives the initial population size adjusted for the fact that some age classes (e.g., teenagers) have greater reproductive value than others (e.g., post-menopausal).

The right eigenvector is referred to as the **stable age distribution** (more on this later).

Let's first analyse the Canadian data using a simplified version of the census data.

We'll group ages from $0-29$, $30-59$, and $60+$ to get only three classes.

The general Leslie matrix for this case would be:

$$L = \begin{bmatrix} m_1 & m_2 & m_3 \\ p_1 & 0 & 0 \\ 0 & p_2 & p_3 \end{bmatrix}$$

For the Canadian population, I get:
$\vec{m} = [0.5834, 0.2973, 0]$
$\vec{p} = [0.9851, 0.9348, 0.1399]$
$\vec{n}_{1991} = [5798100, 5460500, 2430500]$

One way of finding out the future age distribution is to numerically power up the matrix and calculate $\vec{n}_t = L^t \vec{n}_0$ to get:

$\vec{n}_{2021} = [5005209, 5711850, 5444516]$ (in one census)
$\vec{n}_{2051} = [4617419, 4930754, 6101159]$ (in two censuses)
$\vec{n}_{2081} = [4159039, 4548732, 5462859]$ (in three censuses)

This indicates that the proportion of females in the oldest age class ($60+$) should have risen rapidly from 18% in 1991 to 34% in 2021 to 39% in 2051, where it should remain roughly constant.

It also indicates that the population size of females should have grown by 18% between 1991 and 2021, but decrease by 3% between 2021 and 2051, decrease by 9% between 2051 and 2081, and by roughly 9% every 30 years thereafter.

In 1991, the population was very young and growing more rapidly because of this. As the Canadian population aged, its growth rate should have declined.
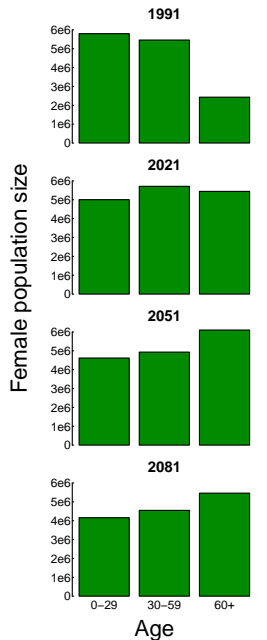
Let's now use the simpler method described above, where we assume that the leading eigenvalue dominates the system.

For this matrix, the eigenvalues are [0.9064, -0.3231, 0.1399].

After an initial period of time, a linear system will approach a stable distribution, where the ratio of individuals in each class is constant and given by the eigenvector associated with the leading eigenvalue.

In this case, the eigenvector associated with the leading eigenvalue is: $[0.2930, 0.3185, 0.3885]$ (It has been normalized to sum to one.)

In demography, this distribution is known as the **stable age distribution**.

Once at the stable age distribution, every age class and hence the whole population will get multiplied by the leading eigenvalue every census.

In this example, the leading eigenvalue is 0.9064, indicating that the population will shrink by $1 - 0.9064 = 0.0936$ (or 9.36%) every census.

We can then use these values to project into the future.

To use the approximate projection, we first find the left eigenvector:
$\vec{v} = [2.5157, 0.8250, 0]$ (chosen such that $\vec{v} \cdot \vec{u} = 1$).

This indicates that $0 - 29$ year olds have about three times the reproductive value of $30 - 59$ year olds, and that the reproductive value of females over 60 is zero.

Adjusting the initial population vector by the reproductive value gives
$\vec{v} \cdot \vec{n}_{1991} = 1.91 \times 10^7$.

This is larger than the female population size in 1991 (1.37×10^7), because the 1991 population is fairly young relative to the stable age distribution.

Our estimates of the future population are then λ^t times the stable age distribution, $\vec{u} = [0.2931, 0.3185, 0.3884]$, times the adjusted initial population size of 1.91×10^7:

$\vec{n}_{2021} = [5070924, 5511469, 6721841]$ (in one census)
$\vec{n}_{2051} = [4596189, 4995490, 6092549]$ (in two censuses)
$\vec{n}_{2081} = [4165898, 4527817, 5522170]$ (in three censuses)

This correctly predicts that the proportion in the oldest age class ($60+$) should be 39%.

This approximation suggests that the population size of females should grow by 26% (rather than 18%) between 1991 and 2021, then decrease by 9% (rather than 3%) between 2021 and 2051, and then decrease by 9% every 30 years thereafter.

Although this method should technically only work well for longer time periods, it still does very well in this case even over the short time period measured here.

After only two censuses, the approximate method performs nearly as well as the recursions using the full Leslie matrix. There is a short, initial phase, however, in which the non-leading eigenvalues still exert an influence on the population dynamics.

The phase is particularly short in this example, since the non-leading eigenvalues are all small (-0.3231 and 0.1399) and so these terms rapidly become negligible in the general solution.

In summary: the leading eigenvalue does estimate long term growth, its eigenvector does estimate the long term ratio between classes, but the short term growth can be influenced by the fact that the population isn't yet at a stable age distribution.

The above analysis with only three age classes provides only a crude demographic picture of the Canadian population.

Let's now return to an analysis using the full data matrix.

$$L = \begin{bmatrix}
0 & 0 & 0.0007197 & 0.06608 & 0.2023 & 0.3142 & 0.2155 & 0.07154 & 0.009788 & 0.0004184 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.9929 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.9991 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.9992 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.9981 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.9982 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.9976 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.9972 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9959 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9935 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9897 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9837 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9732 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.958 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9321 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8943 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8214 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7003 & 0.3045
\end{bmatrix}$$

This is an example where R comes in really handy!

The leading eigenvalue for this matrix is 0.9772.

It's eigenvector is [0.05059, 0.05140, 0.05256, 0.05374, 0.05489, 0.05607, 0.05724, 0.05841, 0.05953, 0.06052, 0.06130, 0.06171, 0.06146, 0.06025, 0.05747, 0.05260, 0.04422, 0.04603]. (I've normalized the eigenvector to sum to one.)

This tells us many things. There will about 5.1% of the population under age 5 at the stable age distribution and 20.8% under age 20. Similarly, there will be 20.0% of the population over age 70 and 32.2% over age 60.

We can compare these figures to their 1991 values of 9.2% of the population over age 70 and 17.8% over age 60, to infer that the population will in the future be composed of nearly twice as many older individuals then there were in 1991, expressed as a fraction of the population.

Furthermore, when the population does reach the stable age distribution, it will shrink by 2.28% during every five year census period.

The left eigenvector of reproductive values is $\vec{v} = [3.2481, 3.1968, 3.1265, 3.0554, 2.7764, 2.0597, 0.9944, 0.2724, 0.0339, 0.0014, 0, 0, 0, 0, 0, 0, 0]$ (normalized so that $\vec{v} \cdot \vec{u} = 1$).

When this eigenvector is multiplied by the vector of female population size from 1991, the adjusted initial population size becomes $\vec{v} \cdot \vec{n}_{1991} = 1.82 \times 10^7$ (again larger than the initial female population size of 1.36×10^7).

Taking this adjusted initial population size, 1.82×10^7, and multiplying by the stable age distribution by $(0.9772)^t$ provides an estimate of the female age distribution in future censuses.

Of course, with R, it is straightforward to iterate the Leslie matrix to predict future population states.
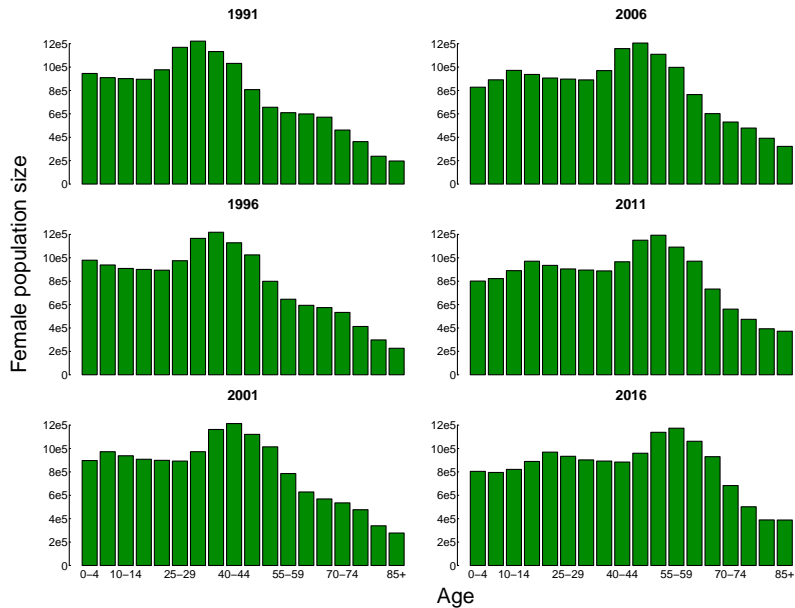
Finally, we can do a check by comparing the expected age distribution in 1996 to the actual census:
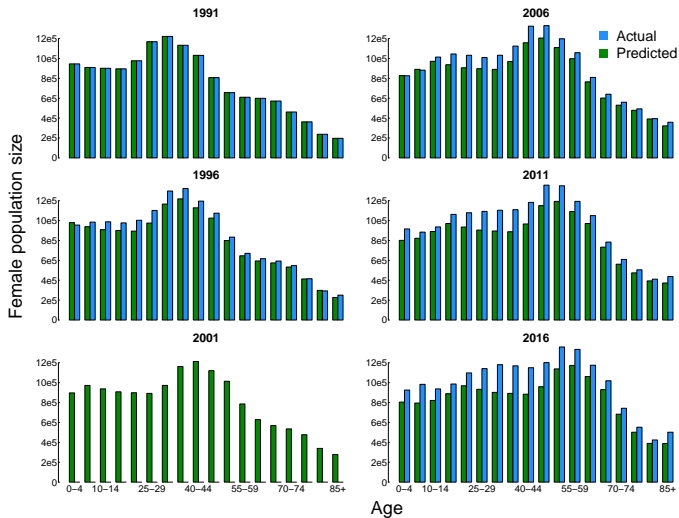
These comparisons show that our projections fall short.

These discrepancies are due to three main factors:

1. Immigration and emigration

2. Too coarse an age distribution

3. A changing life table

Changing age distribution of births in Canada