European Doctoral School of Demography (EDSD)
E120 Basic Mathematics for Demographers
Assignment Team 1

Liliana Patricia Calderon Bernal
Ainhoa-Elena Leger
Lucas Pitombeira
Marilyn-Anne Tremblay
Paola Vazquez-Castillo

14 October 2020

Exercise 1

A simplified model of population growth can be expressed as:

$$P(t) = P(t_o)e^{r(t-t_o)} \tag{1}$$

where $P(t)$ is the population size at time t, $P(t_0)$ is the population at initial time t_0 and r is the constant growth rate. If interested, see Keyfitz and Caswell (2005, Chap. 1) for a detailed discussion.

Point a

Suppose the growth rate is $r = 0.01$. In how many years the population will double its size with respect to $P(t_0)$?

[Sol.]

From (1),

$$\frac{P(t)}{P(t_0)} = e^{r(t-t_0)} \Rightarrow ln(\frac{P(t)}{P(t_0)}) = r(t-t_0) \Rightarrow \frac{ln(\frac{P(t)}{P(t_0)})}{r} = t - t_0 \tag{a}$$

where, $t - t_0$ is the time elapsed. On the other hand, having the double of the initial population means:

$$\frac{P(t)}{P(t_0)} = \frac{2P(t_0)}{P(t_0)} = 2 \tag{b}$$

Therefore, substituting $(**)$ in $(*)$:

$$t - t_0 = \frac{ln(2)}{r} \tag{c}$$

Note that $eq.(c)$ is the doubling time for any r. Substituting the data of this problem.

$$\Rightarrow t - t_0 = \frac{ln(2)}{0.01} = 69.31$$

Point b

Calculate the growth rate if the population doubles its size in 60 years; if the population triples its size in 100 years; and if the population reduces to 50% of its size in 50 years.

[Sol.]

From (1),

$$r = \frac{ln(\frac{P(t)}{P(t_0)})}{t - t_o} \tag{d}$$

Therefore, substituting in (d):

b1. $r = \frac{ln(2)}{60} \approx 0.0115$

b2. $r = \frac{ln(3)}{100} \approx 0.0110$

b3. $r = \frac{ln(\frac{1}{2})}{50} \approx -0.0139$

Point c

Select three countries of your choice from three different continents. Using data from the World Population Prospects (United Nations, 2017) and applying (1), estimate the growth rate of each of their populations between 1950 and 2015 (ignoring intermediate years).

[Sol.]

We randomly select the countries for the study, based on the countries where the members of the team are native and one from the Asian Tigers. The resulting countries are Italy from Southern Europe, Canada from Northern America and Hong Kong from Eastern Asia.

```
## [1] "Canada"
```

```
## [1] "Hong Kong"
```

```
# c rates 1950-2015
# HongKong
rM1950<- (log(WD[1,4]/WD[1,2]))/(2015-1950)
rM1950
```

```
## [1] 0.01987801
```

```
# Italy
rI1950<- (log(WD[2,4]/WD[2,2]))/(2015-1950)
rI1950
```

```
## [1] 0.004036452
```

```
#Canada
rC1950<- (log(WD[3,4]/WD[3,2]))/(2015-1950)
rC1950
```

```
## [1] 0.01483737
```

Point d

Repeat the same procedure, but in this case compute the population growth between 1990 and 2015.

[Sol.]

```r
#d rates 1990-2015
# HongKong
rM1990<- (log(WD[1,4]/WD[1,3]))/(2015-1990)
rM1990
```

```
## [1] 0.009071091
```

```r
# Italy
rI1990<- (log(WD[2,4]/WD[2,3]))/(2015-1990)
rI1990
```

```
## [1] 0.002401708
```

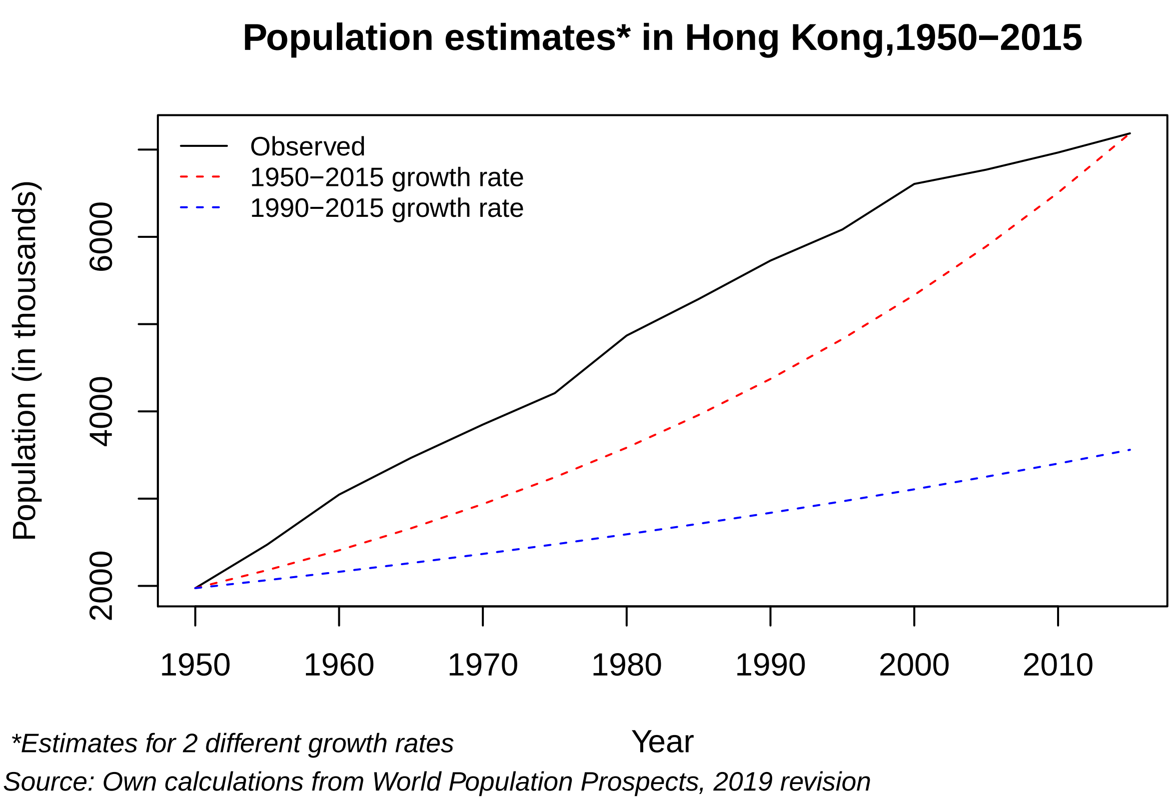
```r
# Canada
rC1990<- (log(WD[3,4]/WD[3,3]))/(2015-1990)
rC1990
```

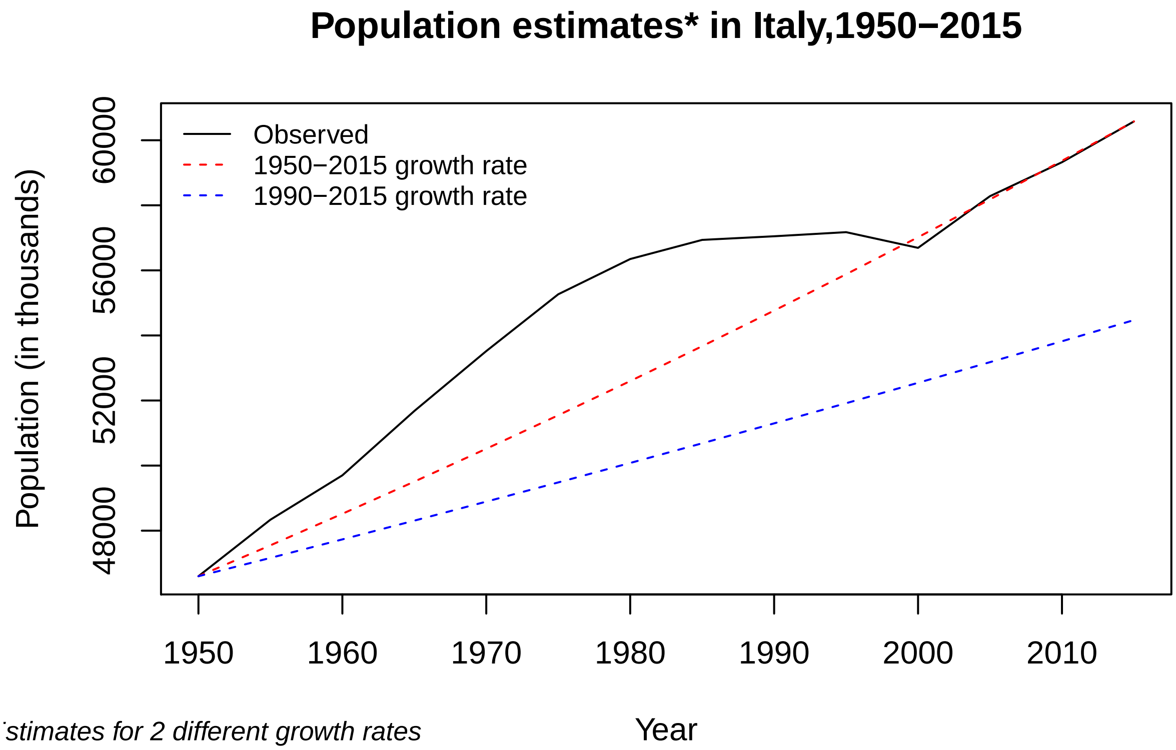
```
## [1] 0.01074288
```

Point e

For each of the three countries, do a single plot with the observed population counts from 1950 until 2015, adding two lines with each of the estimated populations using the results from the two previous questions (one for 1950–2015, and another for 1990–2015).

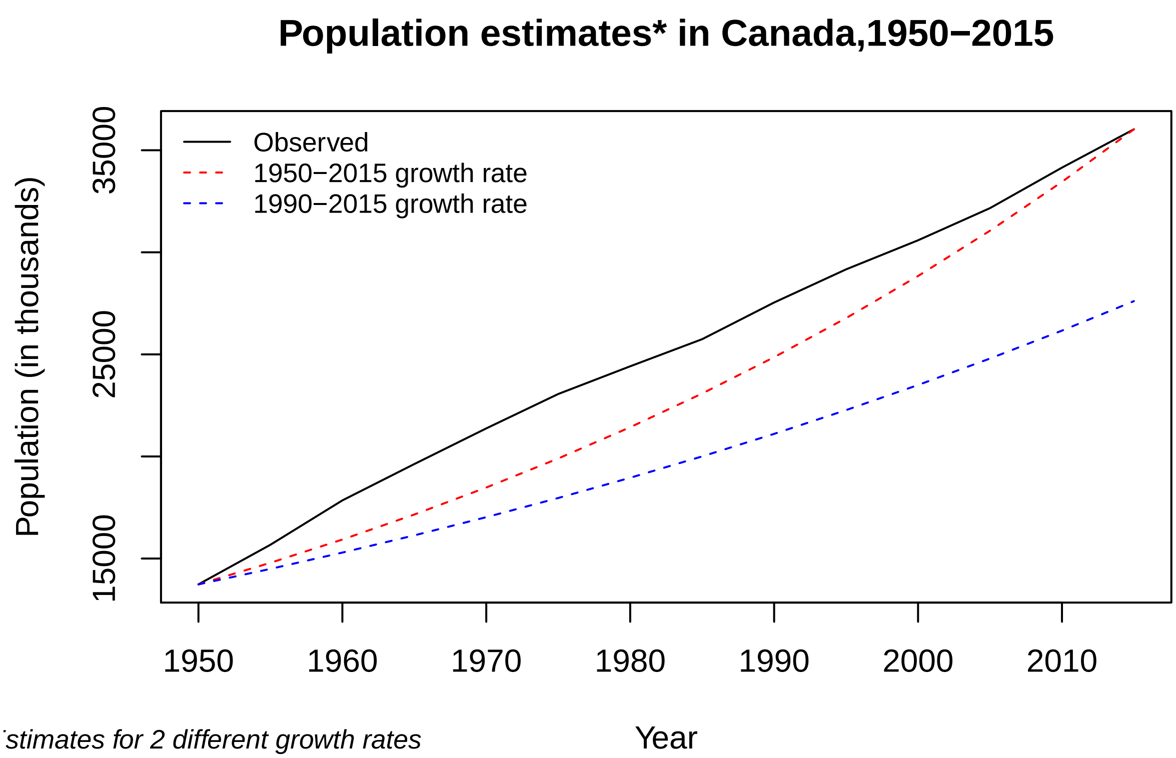
[Sol.]





stimates for 2 different growth rates

: Own calculations from World Population Prospects, 2019 revision



stimates for 2 different growth rates

: Own calculations from World Population Prospects, 2019 revision

Point f

Briefly discuss the differences among the three countries, and which of the estimated growth rates are better according to the graphs.

[Sol.]

From the graph we can note that in the three countries, the blue line, which is considering the growth rates form 1990 to 2015, produce lower population estimates compared to the one considering the 1950-2015 rate. This makes sense considering that these three countries experimented the second phase of the demographic transition somewhen between 1950-1990, meaning that the fertility rates lowered within the period producing lower population estimates. As there was a disrupting event, it is better considering the rates calculated with data form the beginning and the end of the estimation period.

Comparing the three countries, they differ the most on their population sizes, but also in the speed of their growth (measured with the rate of growth), which is also a phenomena that can be appreciated in the graph by comparing the slopes of the lines.

Exercise 2

One of the most well-known age at marriage distributions for historical populations in the demographic literature is the Coale-McNeil model (Coale and McNeil, 1972), which is based on a three-parameter double-exponential function. Fitting the curve to Swedish female data from the eighteenth century, Coale and McNeil obtain a "standard distribution" for the ages at first marriage, defined as

$$g_s(z) = 0.1946 \cdot \exp\left[-0.174(z-6.06) - e^{-0.2881(z-6.06)}\right] \tag{2}$$

where $z \geq 0$ stands for the age at marriage (age 0 is simply the minimum age at marriage). Trussel (1976) discusses (2) ans suggests a new distribution that includes a scaling parameter $k > 0$, given by

$$g(z) = g_s\left(\frac{z}{k}\right) \cdot \frac{1}{k}, \quad z \geq 0 \tag{3}$$

This scaling parameter makes the distribution of ages at first marriage more fexible. When $k = 1$, we obtain the standard distribution described by Coale and McNeil (1972).

Point a

Estimate the modal age at first marriage of the Coale-McNeil model with respect to the scaling parameter k by computing $g'(z)$, using Trussell's (1976) approach.

[Sol.]

We can express the composition of functions (3) substituting (2) as

$$g_s(z) = \frac{1}{k} \cdot 0.1946 \cdot \exp\Big[-0.174\Big(\frac{z}{k} - 6.06\Big) - e^{-0.2881\left(\frac{z}{k}-6.06\right)}\Big].$$

The modal age at first marriage is the age in which more marriages are occurring. It is the value in which the function $g_s(z)$ has its maximum. Therefore, we need to find the value in which the first derivative of the function equals 0 and verify that it is a local maximum. The first derivative, applying the chain rule, is obtained as

$$g_s'(z) = \frac{1}{k} \cdot 0.1946 \cdot e^{-0.174\left(\frac{z}{k}-6.06\right) - e^{-0.2881\left(\frac{z}{k}-6.06\right)}} \cdot \Big[-\frac{0.174}{k} - e^{-0.2281(\frac{z}{k}-6.06)}\Big(-\frac{0.2881}{k}\Big)\Big]. \tag{a}$$

We now look for the value M such that $g_s'(M) = 0$. The first part of the expression (a) (outside the squared brackets) is always different from 0 because of $k > 0$ and the exponential function being strictly positive. Then our search is limited to

$$\frac{-0.174}{k} - e^{-0.2281(\frac{z}{k}-6.06)}(-\frac{0.2881}{k}) = 0$$

Rearranging the expression

$$e^{-0.2881(\frac{z}{k}-6.06)} = \frac{0.174}{0.2881}$$

and taking the logarithm

$$-0.2881(\frac{z}{k} - 6.06) = \ln(\frac{0.174}{0.2881})$$

we obtain the modal age at first marriage M with respect to the scaling parameter k

$$M = k\Big[6.06 - (\frac{1}{0.2881})\ln(\frac{0.174}{0.2881})\Big]. \tag{b}$$

We need to check that it is a maximum, i.e. that the second derivative is negative at the modal age at first marriage M. If we call

$$A(z) = -0.174\Big(\frac{z}{k} - 6.06\Big) - e^{-0.2881\left(\frac{z}{k}-6.06\right)}$$

then the expression (a) can be rewritten as

$$g_s'(z) = \frac{1}{k} \cdot 0.1946 \cdot e^{A(z)} \cdot A'(z).$$

The second derivative which we need to search for the maximum is thus

$$g_s''(z) = \frac{0.1946}{k}\Big[e^{A(z)} \cdot A'(z)^2 + e^{A(z)} \cdot A''(z)\Big] = \frac{0.1946}{k}e^{A(z)}\Big[A'(z)^2 + A''(z)\Big]. \tag{c}$$

with

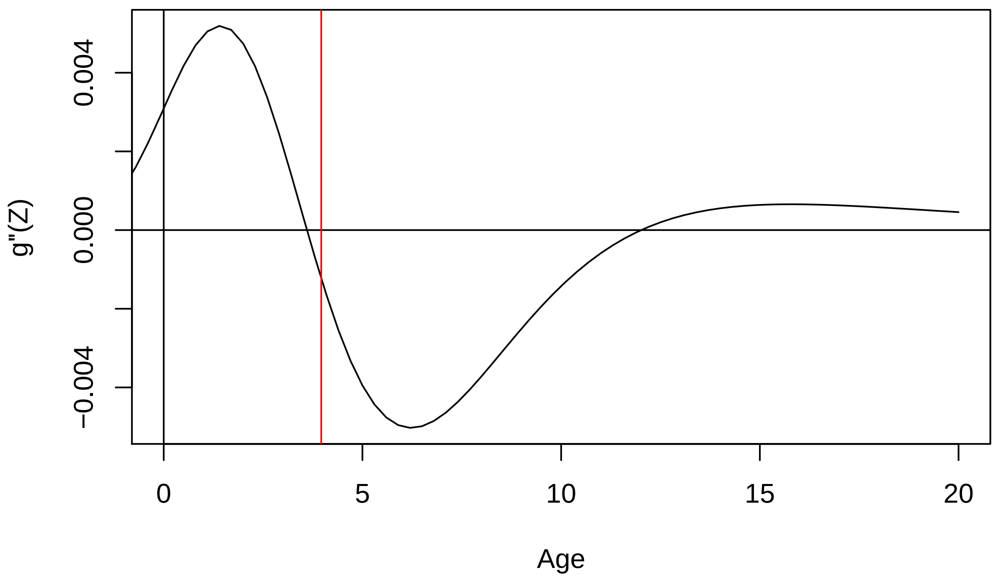
$$A'(z) = -\frac{0.174}{k} - e^{-0.2281(\frac{z}{k}-6.06)}\Big(-\frac{0.2881}{k}\Big), \quad A''(z) = -e^{-0.2281(\frac{z}{k}-6.06)}\Big(\frac{0.2881}{k}\Big)^2.$$

Let's study the sign of expression (c) at the modal age at first marriage (b). Remind that k is a positive constant, the exponential is always positive, and $A'(M) = 0$. Therefore, we need to focus only on $A''(M)$. After substitution of M we get

$$A''(M) = -\frac{0.174 * 0.2881}{k} < 0 \quad \text{for } k > 0.$$

We can also plot the $g_s''(z)$ to check graphically its sign at the modal age at marriage with R. For instance, for $k = 1$ it is negative in $M = 3.96$, which confirms that M is a maximum.

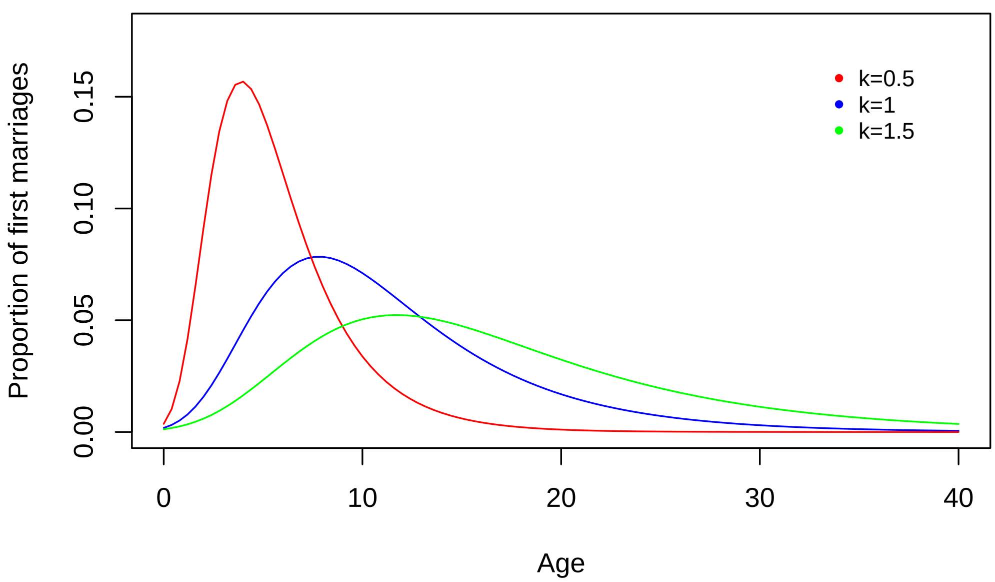
```
## [1] 3.963655
```



Point b

In a single plot, sketch $g(z)$ for 3 different values of k: $k = 1$, one value of k larger than 1, and another smaller. Plot for values $0 \leq z \leq 40$.

[Sol.]



Point c

Briefly discuss which is the effect of parameter k in the shape of this distribution, and the location of the modal age at first marriage.

[Sol.]

Let's compute the modal age at death for the three values of k specified above.

```
k <- c(0.5, 1, 1.5)
M <- k * (6.06 - (1/0.2881) * log(0.174/0.2881))
cbind(k, M)
```

```
##        k         M
## [1,] 0.5  3.905134
## [2,] 1.0  7.810268
## [3,] 1.5 11.715403
```

The parameter k shapes the distribution after the minimum age at marriage. It ranges from 0.2 (very rapid increase rate of marriage after the minimum age) to 1.7 (very slow). A value of k equal to 0.5 indicates a rate of marriage that is double that in the standard population ($k = 1$). When k is smaller than 1 the risk of first marriage is concentrated at the very young ages, when k is larger than 1 the risk is distributed over all ages.

Exercise 3

The force of mortality of the Weibull and logistic models are, respectively,

$$\mu(x) = ax^b \quad \text{and} \quad \mu(x) = c + \frac{ae^{bx}}{1 + \alpha e^{bx}} \tag{4}$$

(Thatcher et al., 1998). Derive, step by step, the corresponding survival functions l(x) and distribution of deaths d(x) for each of these two models, taking into account that

$$l(x) = e^{-\int_0^x \mu(s)ds} \quad \text{and} \quad d(x) = \mu(x)l(x)$$

Try to simplify the final expressions as much as possible. Provide some plots of these survival functions and distributions of deaths for reasonable values of parameters of a, b, c, and α.

Weibull model

The survival function is given by

$$l(x) = \exp\Big(-\int_0^x \mu(s)ds\Big) = \exp\Big(-\int_0^x as^b ds\Big) = \exp\left[-a\cdot\frac{s^{b+1}}{b+1}\Bigg|_0^x\right] = \exp\left[-a\cdot\frac{x^{b+1}}{b+1} + a\cdot\frac{0^{b+1}}{b+1}\right] = \exp\left[-a\cdot\frac{x^{b+1}}{b+1}\right]$$

and the death distribution by

$$d(x) = \mu(x)l(x) = ax^b \exp\Big(-\frac{a}{b+1}x^{b+1}\Big).$$

Logistic model

The survival function can be calculated as

$$l(x) = \exp\Big(-\int_0^x c + \frac{ae^{bx}}{1+\alpha e^{bx}} ds\Big) = \exp\Big(-\int_0^x c\ dx - \frac{a}{\alpha b}\int_0^x \frac{\alpha b e^{bx}}{1+\alpha e^{bx}} dx\Big)$$

We can recognize the second integral to be a derivative of the form

$$d(\ln|f(x)|) = \frac{f'(x)}{f(x)}$$

that gives us

$$l(x) = \exp\Big(-cx\big|_0^x - \frac{a}{\alpha b}\ln|1+\alpha e^{bx}|\big|_0^x\Big) = \exp\Big[-cx - \frac{a}{\alpha b}\Big(\ln(1+\alpha e^{bx}) - \ln(1+\alpha)\Big)\Big].$$

Thus,

$$l(x) = \exp\Big[-cx - \frac{a}{\alpha b}\Big(\ln\frac{1+\alpha e^{bx}}{1+\alpha}\Big)\Big] = \exp\Big[-cx - \Big(\ln\frac{1+\alpha e^{bx}}{1+\alpha}\Big)^{\frac{a}{\alpha b}}\Big] = e^{-cx}\cdot e^{-\ln\left(\frac{1+\alpha e^{bx}}{1+\alpha}\right)^{\frac{a}{\alpha b}}} = e^{-cx}\cdot\Big(\frac{1+\alpha e^{bx}}{1+\alpha}\Big)^{-\frac{a}{\alpha b}}$$

The death distribution is

$$d(x) = \mu(x)l(x) = \Big(c + \frac{ae^{bx}}{1+\alpha e^{bx}}\Big)\Big(e^{-cx}\cdot\Big(\frac{1+\alpha e^{bx}}{1+\alpha}\Big)^{-\frac{a}{\alpha b}}\Big)$$

$$d(x) = c\cdot\left(e^{-cx}\cdot\Big(\frac{1+\alpha e^{bx}}{1+\alpha}\Big)^{-\frac{a}{\alpha b}}\right) + \Big(\frac{ae^{bx}}{1+\alpha e^{bx}}\Big)\left(e^{-cx}\cdot\Big(\frac{1+\alpha e^{bx}}{1+\alpha}\Big)^{-\frac{a}{\alpha b}}\right).$$

Plots for the Weibull model

To obtain an indication of the values of the parameters of a and b to consider, we first study the behaviour of the force of mortality. With this parametrization of the Weibull model, b is the shape parameter.

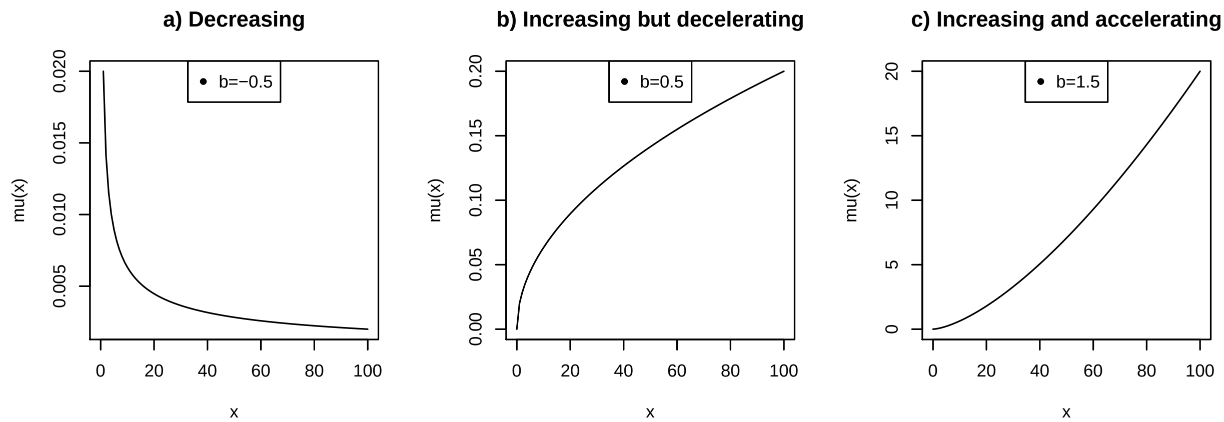
We calculate the first and second derivative.

$$\mu'(x) = abx^{b-1}, \quad \mu''(x) = ab(b-1)x^{b-2}.$$

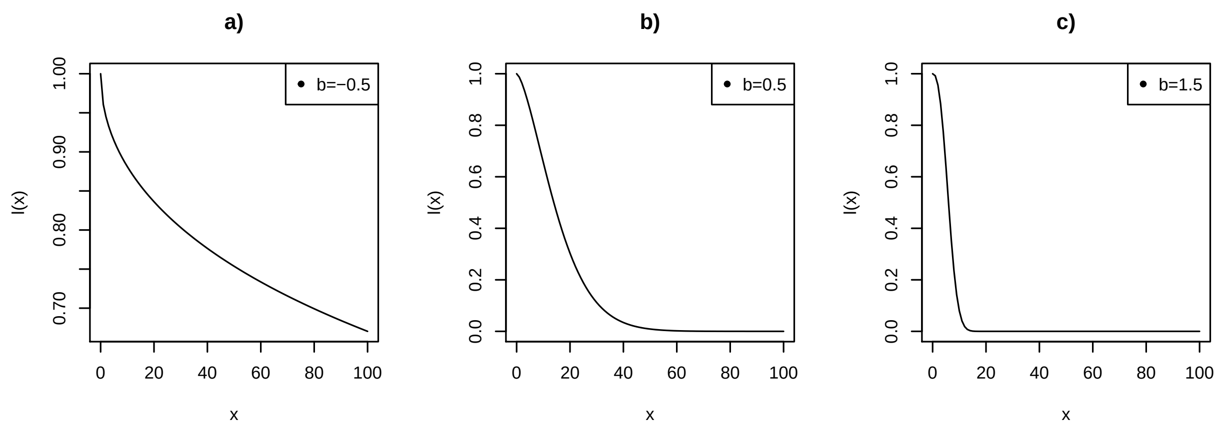
Three cases for the force of mortality can be highlighted (see Table below): a decreasing trend ($b < 0$), an increasing but decelerating trend ($0 < b < 1$) and an increasing and accelerating trend ($b > 1$). See Bebbington, Lai, and Zitikis (2011) for more details on these cases.

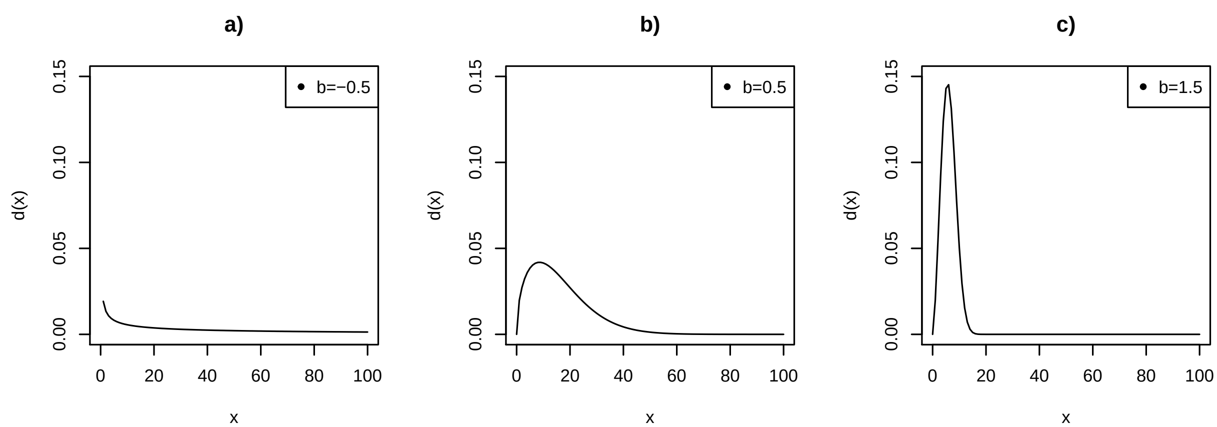
Values of b	I derivative	II derivative
$b < 0$	$\mu'(x) < 0$	$\mu''(x) > 0$
$0 < b < 1$	$\mu'(x) > 0$	$\mu''(x) < 0$
$b > 1$	$\mu'(x) > 0$	$\mu''(x) > 0$

Based on these three cases, we fix the parameter $a = 0.02$ and take b equal to -0.5, 0.5, 1, 1.5. The plot below shows the resulting plots.



We now plot the survival functions and distributions of deaths of the Weibull distribution in the three cases.





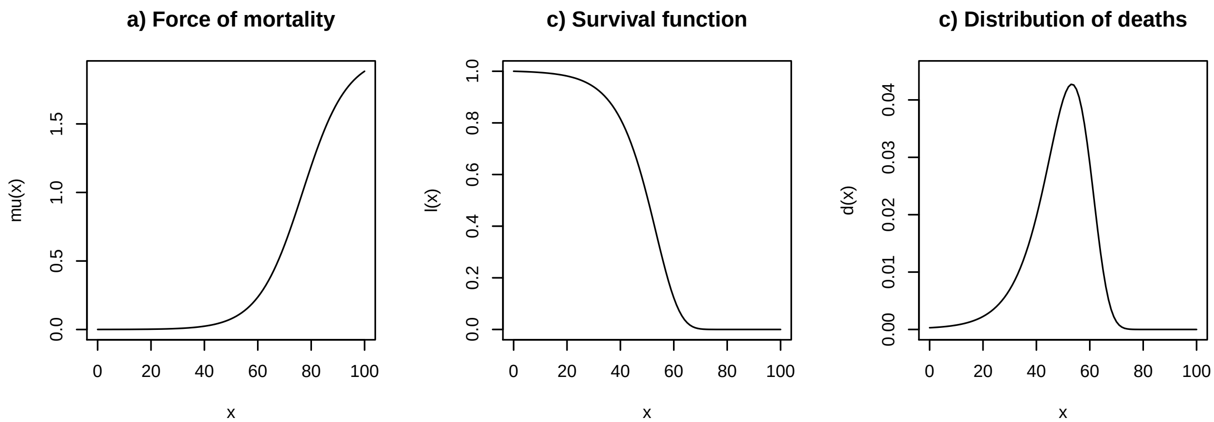
In the case where the force of mortality is decreasing and high in infancy (a), the survival curve is higher than in the other cases and decreases slowly, and the distribution of death if flat except in the first ages. As $b > 0$ increases (b, c) the survival curve becomes more and more sharply decreasing and the distribution of death compressed in the first ages.

Plots for the logistic model

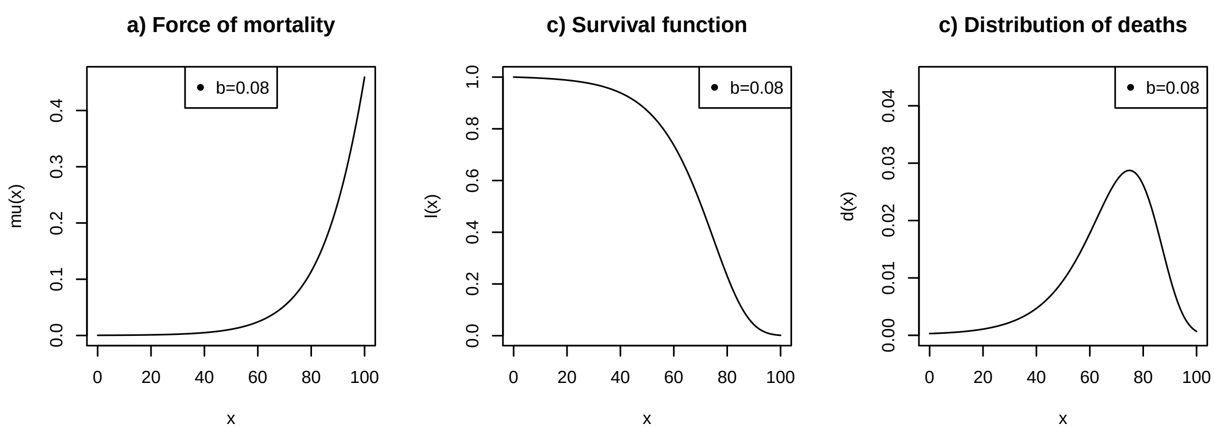
We refer to the article of Thatcher (1999) for the choice of the parameters to plot some logistic models. This article considers the long-term changes in the pattern of mortality and shows the trend in the parameters of the fitted model.

According to Thatcher (1999), the parameter b governs the relative rate at which the force of mortality increases with age. The estimates for b are quite stable between 0.08 and 0.12. The parameter c captures the force of mortality at young adult ages (30-35 years), whereas the force of mortality in the middle range of ages (around 50 years) depends on the parameter α.

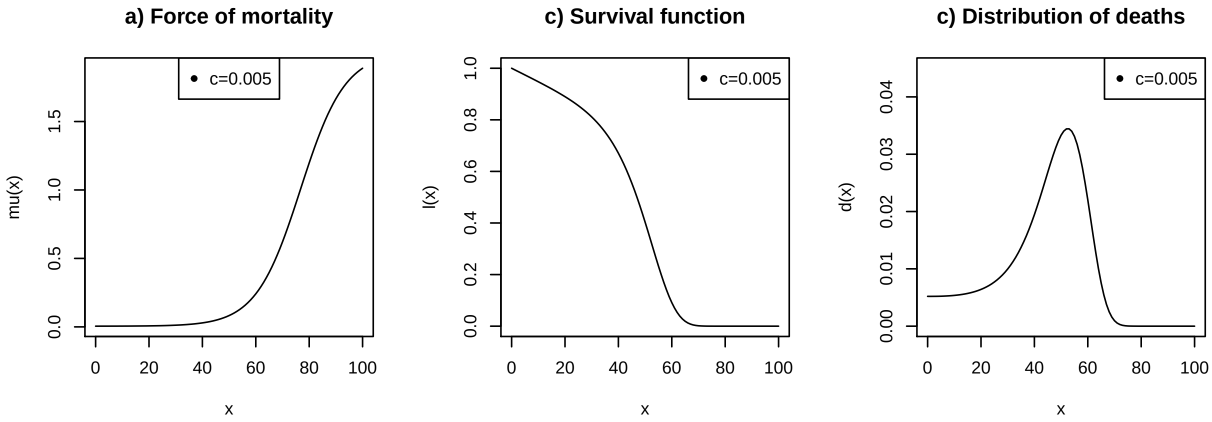
We choose at first $b = 0.12$, $c = 0.0001$, $a = 0.0002$, $\alpha = 0.0001$.



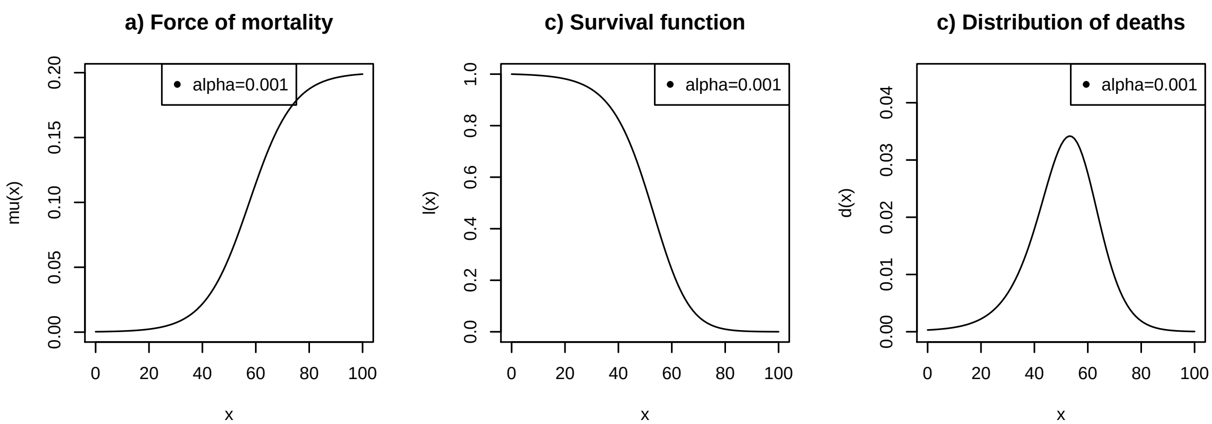
We now change $b = 0.08$ (keeping $c = 0.0001$, $a = 0.0002$, $\alpha = 0.0001$). The force of mortality becomes steeper at older ages, the survival curve is more rectangular and the distribution of deaths is more shifted to older ages.



An increase of c to 0.005 (keeping $b = 0.12$, $a = 0.0002$, $\alpha = 0.0001$) produces an increase in the number of deaths at younger ages. The survival curve is then lower and the distribution of deaths higher until age 35.



Finally, increasing α to 0.001 (keeping $b = 0.12$, $c = 0.0001$, $a = 0.0002$), more deaths seem to occurr at the middle ages and the distribution of deaths is wider around age 50.



Exercise 4

A man aged 30 marries a woman aged 25. Express in symbols the expected number of years they will both be alive. If μ_x is constant and equal to 0.02, evaluate the symbolic expression.

[Sol.]

The expected number of years they will both be alive can be obtained with the definition of the average given a distribution $f_x(t)$.

$$e_x^0 = \int_0^\infty t \cdot f_x(t)dt = \int_0^\infty t \cdot_t p_x \mu_{x+t} = \int_0^\infty t \cdot \left[- \frac{\partial p_x(t)}{\partial t} dt \right]$$

Using integration by parts we obtain

$$\int_0^\infty t \cdot \left[- \frac{\partial p_x(t)}{\partial t} dt \right] = t(-{}_tp_x)\Big|_{t=0}^\infty - \int_0^\infty -{}_tp_x dt$$

Thus, the expected number of years is given by $e_x^0 = \int_0^\infty {}_tp_x$.

The joint probability of the man and the woman to live together is expressed ad the product of the two individual probabilities, as they are independent.

$$_tp_{30:25} =_t p_{30} \cdot_t p_{25}$$

Let's first express the generic probability $_tp_x$ for the man and the woman. From the definition of a probability between times x and $x+t$ we obtain

$$_tp_x = \frac{l_{x+t}}{l_x} = \frac{e^{-\int_0^{x+t}\mu(s)ds}}{e^{-\int_0^{x}\mu(s)ds}} = e^{-\int_0^{x+t}\mu(s)ds+\int_0^{x}\mu(s)ds} = e^{-\int_x^{x+t}\mu(s)ds} = e^{-\int_0^{t}\mu(x+s)ds}$$

As μ is constant then $_tp_x = e^{-\mu t}$. Thus the joint probability becomes

$$_tp_{30:25} = e^{-\mu t}e^{-\mu t} = e^{-0.02t}e^{-0.02t}.$$

Putting all together, the expected number of years that the man and the woman will be alive is

$$e_x^0 = \int_0^{Inf} e^{-0.02t}e^{-0.02t} = 25\text{years}.$$

Alternatively, we can consider the time they will be alive as a joint decrement process, because the risk of dying conjunctly are whether one dies first or the other. So what needs to be calculated is $e^o_{30:25}$.

On one hand we know that $e^o_x = \int_x^\infty \frac{l(x+n)}{l(x)dn}$ and $_np_x = \frac{l(x+n)}{l(x)}$ $\therefore$ $e^o_x = \int_0^\infty {_tp_x}dt$ and by definition, $_tp_x = e^{-\int_x^{x+n}\mu(a)da}$.

On the other hand, if we assume independence $_tp_{x:y} =_t p_x *_t p_y$. In this case, this is a convenient expression considering that the joint-life status survives if and only if both survive. Then:

$$e^o_{30:25} = \int_0^\infty tp_{30:25}dt = \int_0^\infty tp_{30} \cdot_t p_{25}dt = \int_0^\infty e^{-\int_{30}^{30+t}\mu(a)da} \cdot e^{-\int_{25}^{25+t}\mu(a)da}$$

$$= \int_0^\infty e^{0.02a|_{30}^{30+t}} \cdot e^{0.02a|_{25}^{25+t}} = \int_0^\infty e^{0.02(30+t-30)} \cdot e^{0.02(25+t-25)}$$

$$= \int_0^\infty e^{-0.02t} \cdot e^{-0.02t}dt = \int_0^\infty e^{-0.04t}dt = \frac{e^{-0.04t}}{0.04} = 25$$

Therefore,

the expected number of years each one in and independent way will be alive is in average 50 years, however this time is 25 considering the joint expectancy.This is an empiric example that, when the mortality force remains constant the life expectancy is the same at every age, because there are not differences in the "way" people from age to age dies (or among sexes). This is an unrealistic scenario that can only happen if there is an homogenous population without biological aging or mortality risk sex/age/social differences.

References

Bebbington, Mark, Chin-Diew Lai, and Ričardas Zitikis. 2011. "Modelling Deceleration in Senescent Mortality." *Mathematical Population Studies* 18 (1): 18–37.

Preston, Samuel H, Patrick Heuveline, and Michel Guillot. 2001. "Demography, Measuring and Modeling Population Processes." *Population and Development Review* 27 (2): 365–67.

Team, R Core, and others. 2013. "R: A Language and Environment for Statistical Computing." Vienna, Austria.

Thatcher, A Roger. 1999. "The Long-Term Pattern of Adult Mortality and the Highest Attained Age." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162 (1): 5–43.