



Working Papers

04/2010

Towards advanced methods for computing life tables

Sixto Muriel de la Riva

Margarita Cantalapiedra Malaguilla

Federico López Carrión

The views expressed in this working paper are those of the authors and do not necessarily reflect the views of the Instituto Nacional de Estadística of Spain

First draft: December 2010

This draft: December 2010

Towards advanced methods for computing life tables

Abstract

INE Spain puts compiled data and literature in a series of tests and comparative analyses in order to select the best-suited methodologies for computing life tables, trying to reach an optimal use of the available data on deaths, better international comparability of mortality indicators and an enhanced approach to measuring mortality over small territorial areas.

Keywords

Mortality tables; risks of dying; regional mortality tables; different methods to build mortality tables

Authors and Affiliations

Sixto Muriel de la Riva, Margarita Cantalapiedra Malaguilla and Federico López Carrión
Directorate of Population Statistics, National Statistics Institute of Spain

Towards advanced methods for computing life tables

Sixto Muriel de la Riva, Margarita Cantalapiedra Malaguilla and Federico López Carrión

Directorate of Population Statistics, National Statistics Institute of Spain

Abstract

Life tables are, of course, a longstanding task of government statistics. Now, a major effort of research and methodological innovation should open the way to optimal utilisation of the available data on deaths, better international comparability of mortality indicators, and a finer approach to the measurement of mortality over small territorial areas.

INE Spain puts compiled data and literature in a series of tests and comparative analyses in order to select the best-suited methodologies for the mentioned objectives. The outcome of this testing stage was a two-part report on our methodological research on life tables.

The first part tries shows three different approaches to constructing a complete annual life table for the Spain resident population, in a growing sophistication scale: the first one assumes that deaths over the reference year at each age are distributed absolutely uniformly; the second, assumes uniform distribution of deaths over the year at each age within a given generation (uniformity in Lexis triangles), following the protocols of the Human Mortality Database¹; the third and genuine method uses observational data from the vital statistics on the date of occurrence of each recorded death to get the best approach to population at risks in specific mortality rates.

Besides, this part shows the considerable distance between uniformity and the temporal distribution (within the observed year) of deaths in low and advance ages. It is followed that the third method provides the most accurate measurement of the biometric functions of a life table.

The second part of the report focus on life tables for sub-national territories (devolved regions and provinces, NUTS-2 and NUTS-3). It analyses three possible ways to construct life tables at disaggregated territorial levels (abbreviated five-yearly age ranges life tables, full life tables with five yearly age ranges output and smoothing survivor function of regional life table according to

¹ Collaborative project sponsored by the University of California at Berkeley, United States and the Max Planck Institute for Demographic Research, Rostock, Germany. / www.mortality.org

national data). It aims to provide an understanding of the short-term (annual) incidence of mortality in each territory while minimising the effect of reporting randomness and of possible inconsistencies between observed deaths and the estimates used as reference population figures. Advantages and disadvantages of every proposed method are deeply analysed and commented.

1 Introduction

Life tables computing constitutes a traditional work in the field of demographic analysis and projections, for the majority of the official statistical offices. In the present paper, several methodological options for the calculus of life tables are shown and discussed, both for the national and for the regional level.

This work has the following objectives:

1. The maximum use of the demographic information generated by the Spain resident population Statistical System, so as to optimise the accuracy of mortality measurements for all the populations considered.
2. To assure the necessary methodological consistency in the definitions and the assumptions underlying the calculus of all the biometric functions of the complete life table, and also in the definition of its derived indicators.
3. As far as possible, to reflect annual behaviour of mortality incidence in the Spain resident population, avoiding the influence of randomness in observed data.
4. To improve the accuracy when measuring mortality incidence at the most advanced ages, given the prevailing aging process of the Spain Population and the important recent advances in life expectancy at those ages.
5. To provide measures of mortality incidence for sub national populations, i.e., regions and provinces, in a regular basis.
6. To support international comparability of the mortality indicators and facilitate their interpretation by users.

Three different options are tested for the construction of complete life tables for the total, male and female Spain resident populations. Another three options are tested for the regional life table. All of them make use of the registered data on deaths occurred among residents in Spain during a one-year interval and the corresponding population stocks exposed to risk of dying. Year 2006 was used for the national life tables and years 2005 and 2006 for the regional and provincials life tables.

2. Methods for computing national life tables

In this first part of the paper three methodological approaches are presented for the computation of life tables at the national level. Differences among the three options are due to the three different hypothesis made about the time lived by the observed population between ages x and $x + 1$.

In the first option, uniform distribution of deaths inside the age interval $(x, x + 1)$ is supposed. For the second option, based on the protocols of the *Human Mortality Database*¹ project, uniform distribution within the two corresponding generations of annual deaths in the age interval $(x, x + 1)$ is assumed. In the third option, observed data on the date of occurrence of each registered death of annual Vital Statistics are used. For all the options birth dates are supposed uniformly distributed along each year.

In general, the functions of the period life table are defined as follows:

The **mortality rate** $m(t, x, s)$, for year t , age x and sex s , is a quotient where the numerator is the number of deaths occurring in the year t among individuals at age x and sex s , and the denominator is the total time, in years, lived by the individuals of age x and sex s , during the year, i.e., the person-years of individuals of age x and sex s during year t .

Now, let a fictitious generation with an incidence of mortality according to the mortality rates that has been defined. So:

The **death probability** $q(x, s)$ is the probability of an individual of sex s of the fictitious generation that reaches the age x lives until next birthday.

The **survivors** $l(x, s)$ is the number of individuals from the fictitious generation that reach the x age for each sex s .

The **deaths** $d(x, s)$ is the number of deceases from the fictitious generation at each age x and for each sex s .

Survivors $l(x, s)$ and **deaths** $d(x, s)$ are obtained recursively:

$$l(0, s) = 100.000$$

$$d(x, s) = l(x, s)q(x, s) \text{ y } l(x + 1, s) = l(x, s) - d(x, s) \text{ for } x = 0, 1, \dots, 99, 100 +$$

The **average number of years lived in the last year of life** for those individuals dying in the age interval $(x, x + 1)$ is denoted by $a(x, s)$.

The **stationary population** $L(x, s)$ is the total time lived for all the individuals of the fictitious generation that are x years old and sex s .

¹ Department of Demography at University of California, Berkeley (USA) and Max Planck Institute for Demographic Research.

The **total years lived** $T(x,s)$ is the total amount of years lived for all the individuals of the fictitious generation aged x or more.

$$T(t, x, s) = T(t, x + 1, s) + L(t, x, s) \text{ for } x = 0, 1, \dots, 99, 100 +$$

Finally, the **life expectancy** is the mean number of years left to live for an individual of age x and sex s .

$$e(t, x, s) = T(t, x, s) / l(t, x, s), \text{ for } x = 0, 1, \dots, 99, 100 +$$

2.1 Option 1.- Assumption of uniform distribution of deaths in the age interval $(x, x+1)$

The main objective of option 1 is to perform a sensitivity analysis, i.e., to show the effect of the alternative options in comparison to uniform distribution.

Under this assumption, the denominator of $m(t, x, s)$ is estimated by the expression $1/2P(t, x, s) + 1/2P(t + 1, x, s)$, assuming uniform distribution of the birthdays of all individuals in the population not dying over the year at a given age.

Hence:

$$m(t, x, s) = \frac{D(t, x, s)}{1/2P(t, x, s) + 1/2P(t + 1, x, s)}$$

Where:

t is the reference year, period or calendar time.

x is the age or completed years, $x = 0, 1, \dots, 99, 100 +$.

s is sex, which may be male, female or both sexes.

$D(t, x, s)$ is the number of deaths between residents in Spain during the year t , at age x and sex s .

$P(t, x, s)$ is the stock of resident population in Spain at 1st January of the year t , with age x and sex s .

In addition, for this option the following key expressions for some of the biometric functions are adopted:

The estimated **probability or risk of dying** $q(x, s)$, will be:

$$q(x, s) = \frac{m(x, s)}{1 + (1/2) \cdot m(x, s)} \quad x = 0, 1, \dots, 99 \text{ and } q(100+, s) = 1.$$

The **stationary population** at age x and sex s , for $x = 0, 1, \dots, 99$, on the assumption of uniform distribution of deaths over the year, is estimated as:

$$L(x,s) = l(x+1,s) + 1/2d(x,s) \text{ for } x = 0,1,\dots,99 .$$

Finally, the aggregate of years lived by individuals of the fictitious generation reaching the age of 100, from that age to their deaths, is approximated by:

$$L(100+,s) = \frac{l(100+,s)}{m(100+,s)}$$

2.2 Option 2.- Assumption of uniform distribution of deaths by age and generation

This alternative follows the protocols of *Human Mortality Database*¹ for the construction of life tables. In this case, the denominator of $m(t,x,s)$ is estimated assuming uniform distribution of deaths among individuals of the same age and generation (Lexis triangle) by the expression:

$$1/2P(t,x,s) + 1/2P(t+1,x,s) + 1/6(D(t,x,g,s) - D(t,x,g-1,s))$$

Where:

g is the generation of individuals completing x years in year t , i.e, $t - x = g$

$D(t,x,g,s)$ is the number of individuals dying in year t , at age x , from generation g and sex s .

Hence:

$$m(t,x,s) = \frac{D(t,x,g,s) + D(t,x,g-1,s)}{1/2 \cdot P(t,x,s) + 1/2 \cdot P(t+1,x,s) + 1/6 \cdot (D(t,x,g,s) - D(t,x,g-1,s))} , \text{ for } x = 0,1,\dots,99,100+ .$$

For this option the following key expressions for some of the functions are adopted:

The estimated **probability** or **risk of dying** will be obtained by:

$$q(x,s) = \frac{m(t,x,s)}{1 + (1 - a(x,s)) \cdot m(t,x,s)} \text{ for } x = 0,1,\dots,99 \text{ and } q(100+,s) = 1 .$$

being $a(x,s)$ approximated by:

$$a(x,s) = \frac{1/3D(t,x,g,s) + 2/3D(t,x,g-1,s)}{D(t,x,g,s) + D(t,x,g-1,s)}$$

For the open-ended group the average time lived in the reference year t is estimated for ages of 100 and above and sex s , $a(100+,s)$, as:

¹ Department of Demography at University of California, Berkeley (USA) and Max Planck Institute for Demographic Research.

$$a(100+,s) = \frac{1}{m(100+,s)}$$

2.3 Option 3.- Real distribution of date of death events is taken into account

In this case, the denominator of $m(t,x,s)$ comes from the expression:

$$1/2 \cdot P(t,x,s) + 1/2 \cdot P(t+1,x,s) + \sum_i b(t,x,s,i)$$

Where:

$b(t,x,s,i)$ is defined as the difference (in years) between the date of death and the date of birthday (in year t) of each individual i , of sex s , dying in year t at completed age x .

Hence:

$$m(t,x,s) = \frac{D(t,x,s)}{1/2 \cdot P(t,x,s) + 1/2 \cdot P(t+1,x,s) + \sum_i b(t,x,s,i)}, \text{ for } x = 0,1,\dots,99,100+.$$

For this option the following key expressions for some of the biometric functions are adopted:

The **probability** or **risk** of dying, $q(x,s)$ will be:

$$q(x,s) = \frac{m(t,x,s)}{1 + (1 - a(x,s)) \cdot m(t,x,s)}$$

Being the **average number of years lived in the last year of life** for those individuals dying in the age interval $(x,x+1)$, $a(x,s)$, approximated by the average of the time (years) lived in t for the individual i dying in the age interval $(x,x+1)$ during the observed year t , $a(t,x,s,i)$:

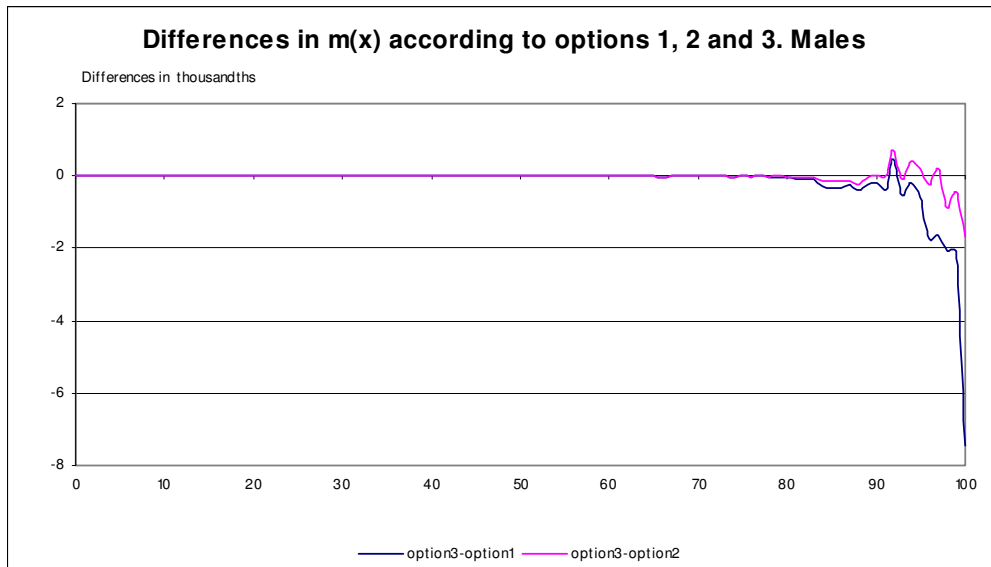
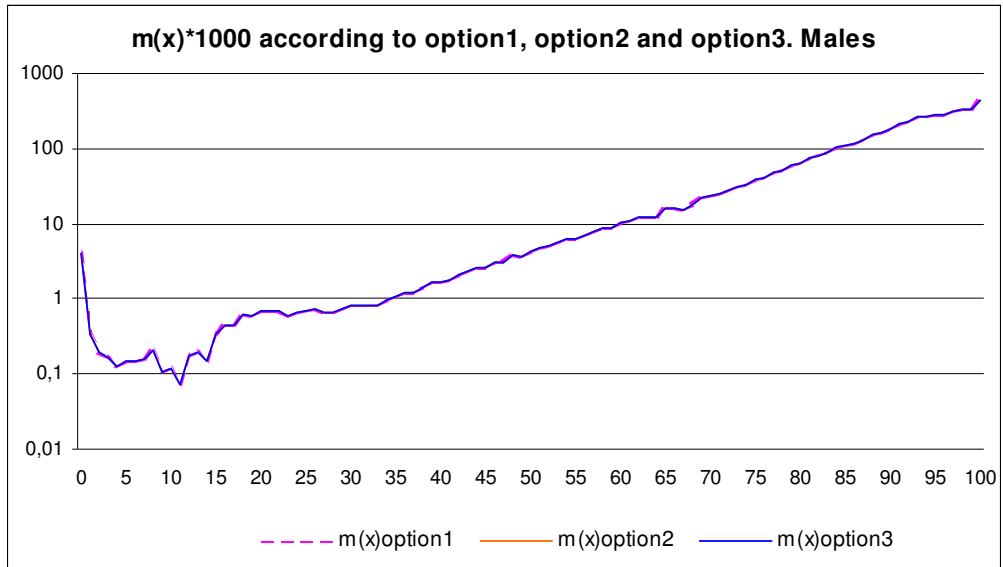
$$a(t,x,s) = \frac{\sum_{i=1}^{D(t,x,s)} a(t,x,s,i)}{D(t,x,s)}, \text{ } x = 0,1,\dots,99.$$

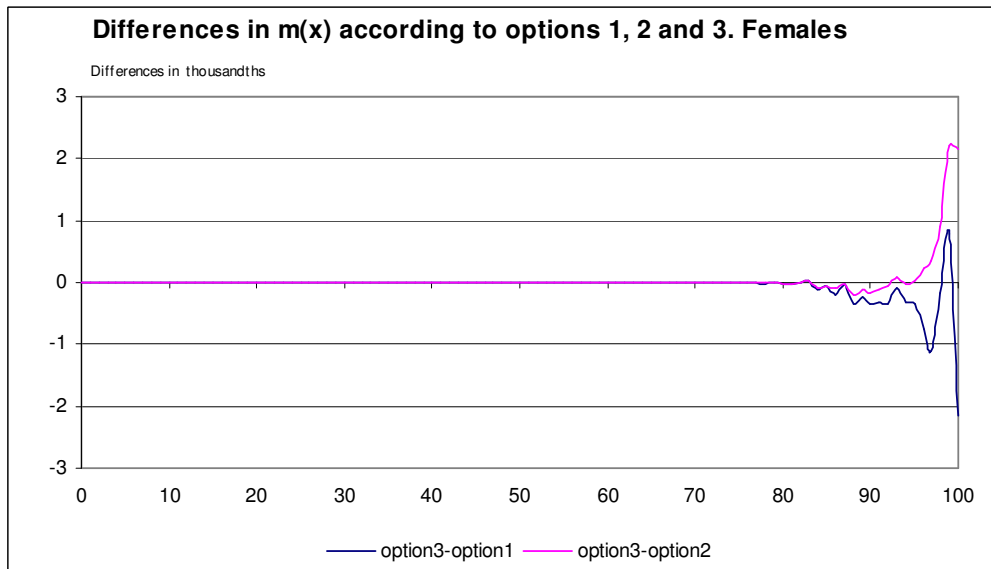
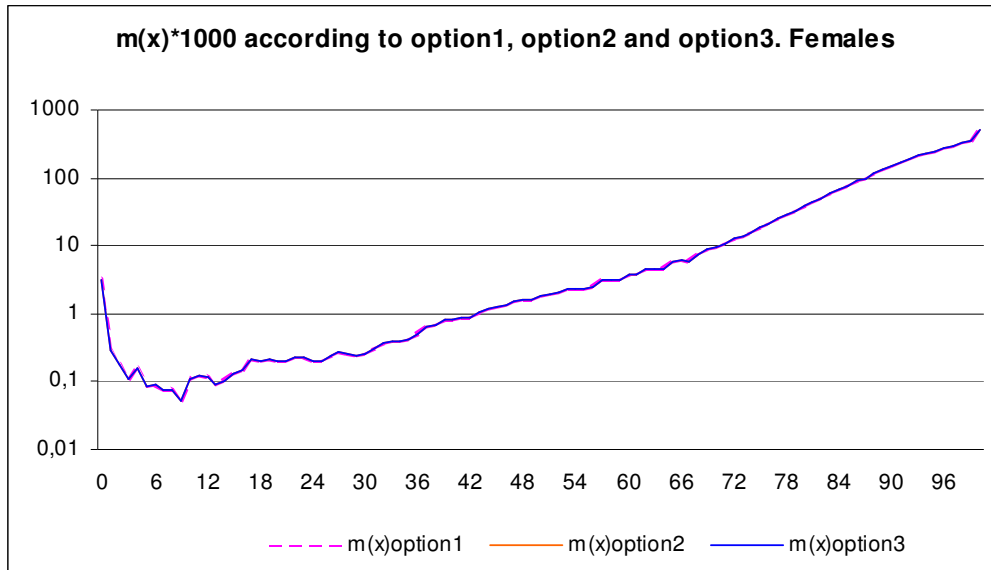
Besides, for the open-ended group the average time lived in the reference year t is estimated for ages of 100 and above and sex s , $a(100+,s)$, as:

$$a(100+,s) = \frac{1}{m(100+,s)}$$

2.4 Discussion

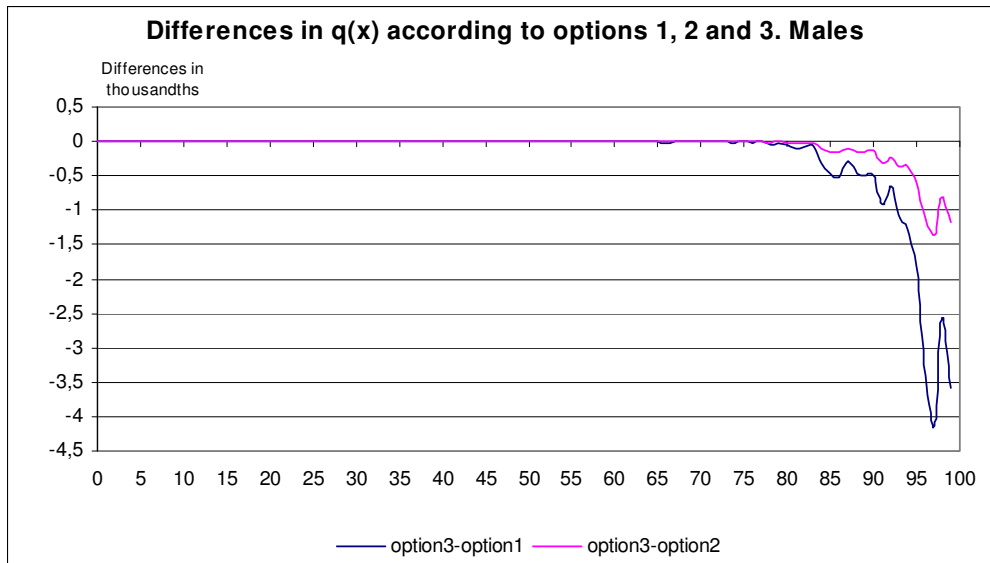
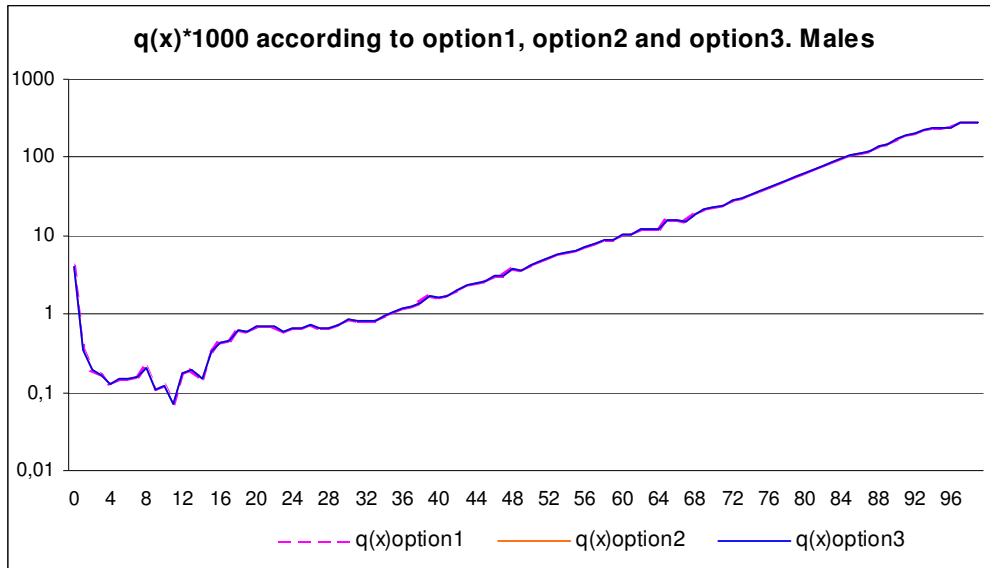
The differences among the specific mortality rates computed by each of the approaches tested were negligible except for the most advanced ages (above 80 years old) where slight differences emerged, both, for males and for females.

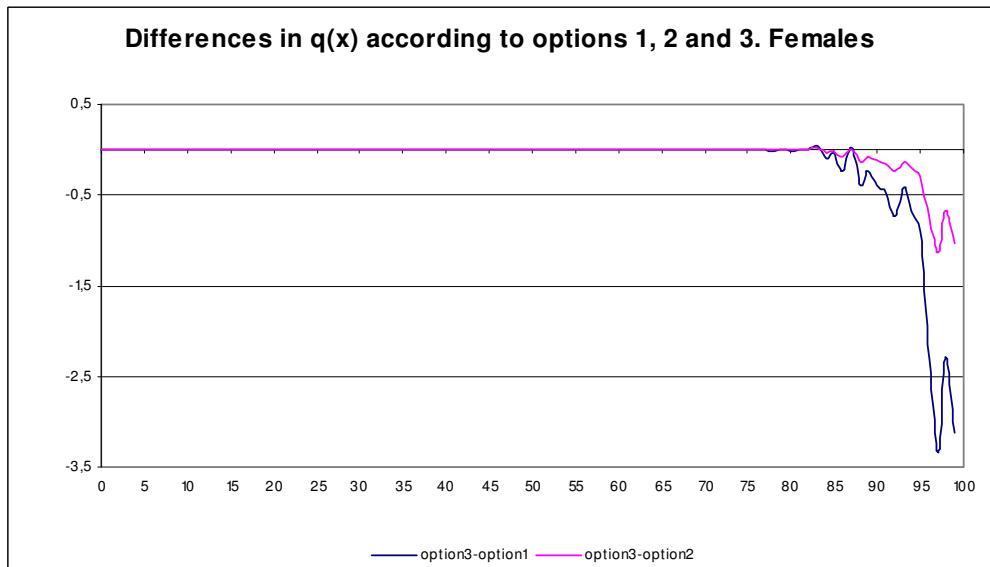
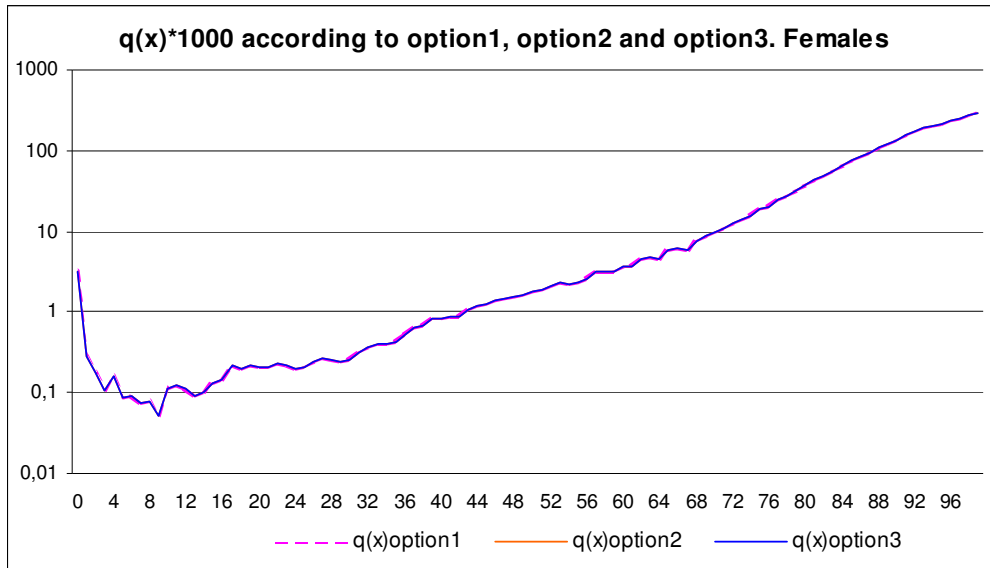




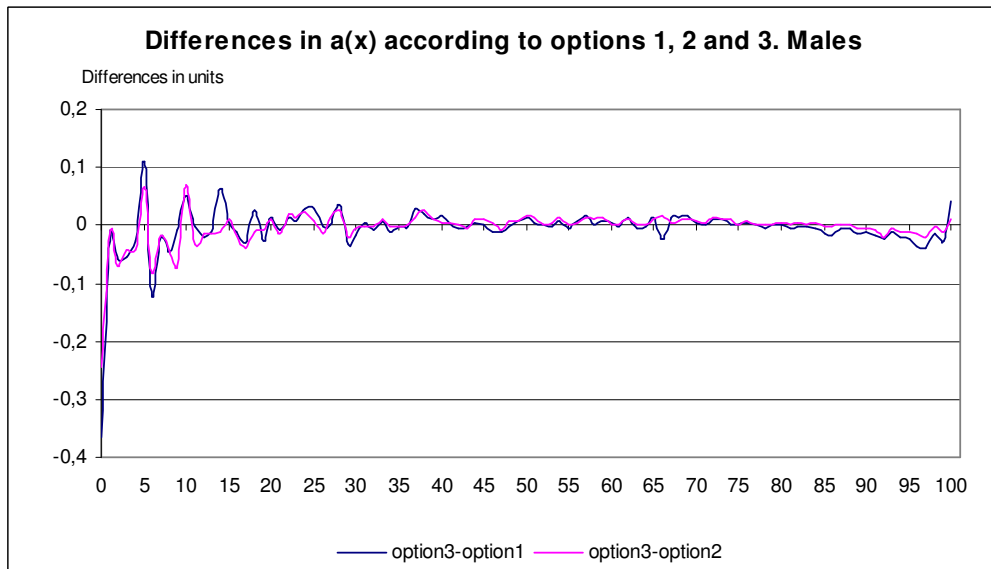
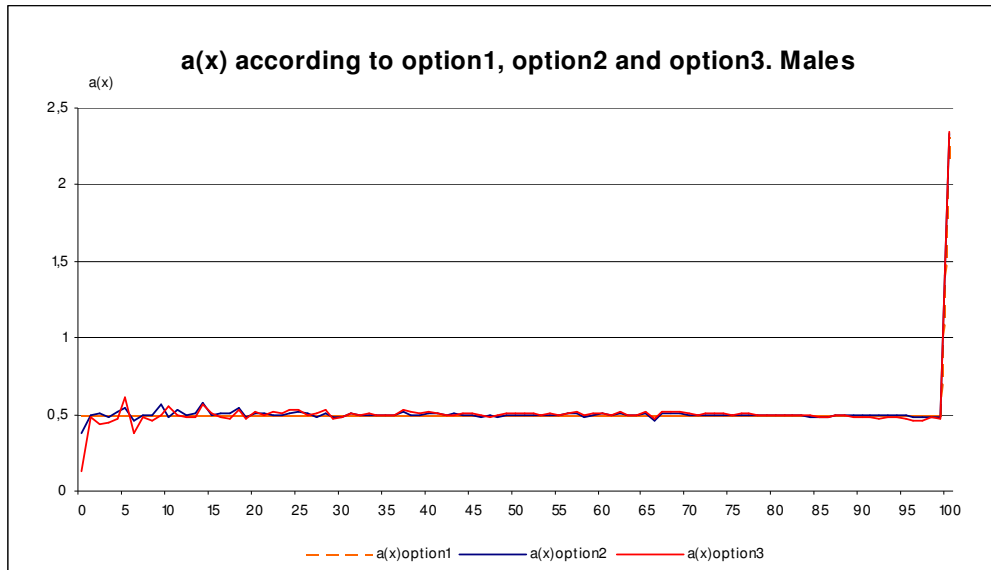
Mentioned differences at advanced ages reflect that deaths deviate from uniform distribution between the ages x and $x+1$: in fact, at advanced ages, deaths are concentrated among the individuals of the youngest generation of the two which complete deaths at the age interval $(x, x+1)$ during the year t .

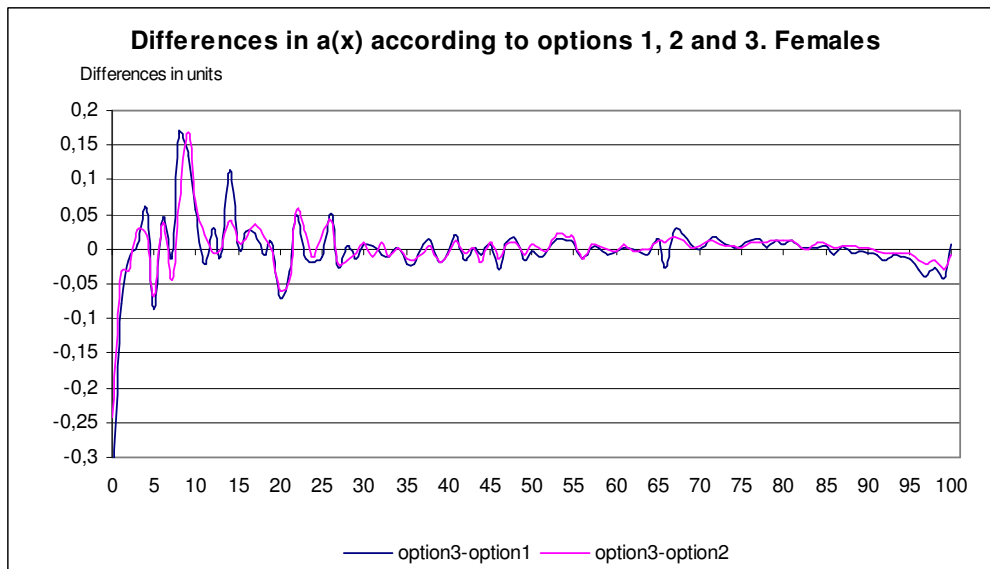
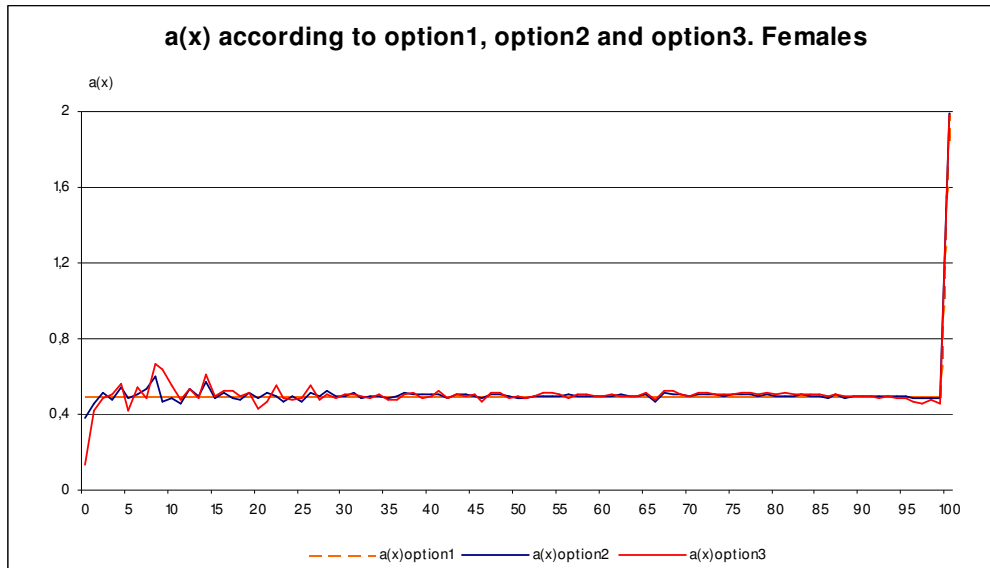
The differences shown among the specific mortality rates are likewise visible for the probabilities of death.





Nevertheless, the calculation of these probabilities of death is also influenced by differences in estimated average number of years lived in the last year of life for those individuals dying in the age interval $(x, x+1)$, $a(x, s)$. In particular, the differences are noteworthy in the approximation by each method of $a(x, s)$. Not even option 2 produces a good approximation of observed reality. The most accurate approximation is provided by option 3.

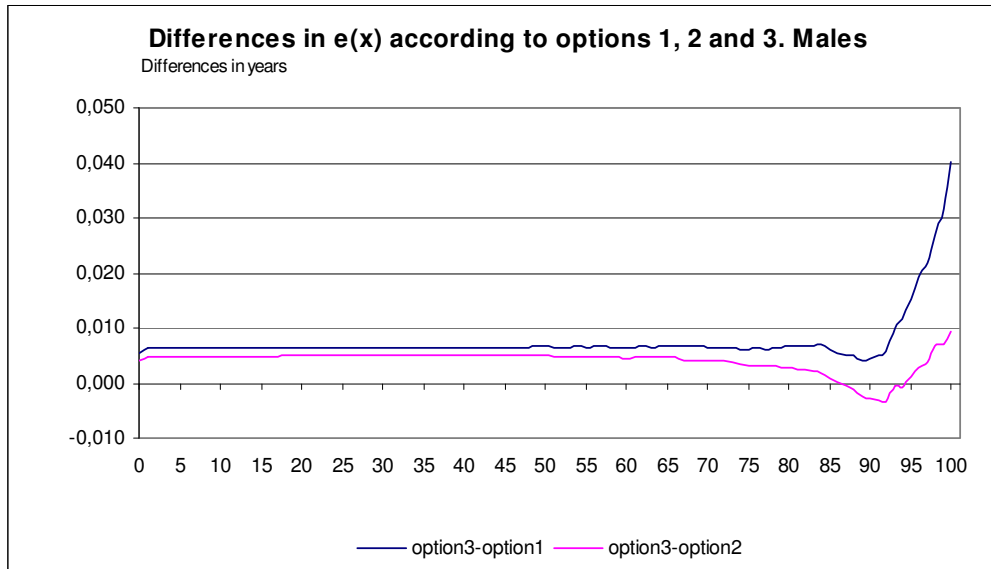




Finally, the differences emerging in each of the mortality table functions resulting from each method have only a negligible effect on the resulting values of life expectancy at birth, as shown in the following charts.

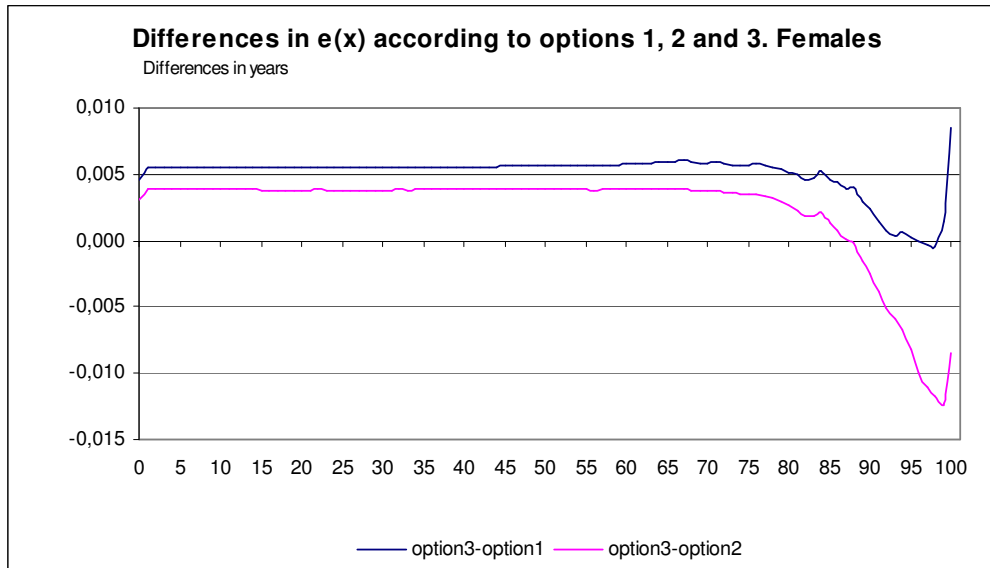
Life expectancy according to options 1, 2, 3. Males

Age	e(x)			Differences	
	option1	option2	option3	option3-option1	option3-option2
0	77,589	77,591	77,595	0,005	0,004
5	72,971	72,973	72,978	0,006	0,005
15	63,072	63,074	63,079	0,006	0,005
25	53,407	53,408	53,413	0,006	0,005
35	43,778	43,779	43,784	0,006	0,005
45	34,427	34,428	34,433	0,006	0,005
55	25,668	25,670	25,674	0,007	0,005
65	17,709	17,711	17,715	0,007	0,005
75	10,825	10,828	10,831	0,006	0,003
85	5,830	5,835	5,836	0,006	0,001
95	3,131	3,145	3,146	0,015	0,001



Life expectancy according to options 1, 2, 3. Females

Age	e(x)			Differences	
	option1	option2	option3	option3-option1	option3-option2
0	84,089	84,091	84,094	0,005	0,003
5	79,418	79,420	79,424	0,006	0,004
15	69,487	69,489	69,492	0,005	0,004
25	59,614	59,615	59,619	0,006	0,004
35	49,773	49,774	49,778	0,006	0,004
45	40,125	40,127	40,131	0,006	0,004
55	30,743	30,745	30,748	0,006	0,004
65	21,656	21,658	21,662	0,006	0,004
75	13,280	13,283	13,286	0,006	0,003
85	6,757	6,761	6,762	0,005	0,001
95	3,192	3,201	3,192	0,000	-0,008



Option 3, shows a slight improvement in life expectancy at each age, as a result of the slight decrease in mortality rates in the upper ages. The life expectancy gains are more significant at those ages, where present improvements of such indicator are concentrated.

3. Methods for computation of life tables at lower territorial levels

The second part of this paper addresses the questions for an accurate measurement of mortality at lower territorial levels, trying to minimize the limitations originated by the small size of the observed population.

Three different options are tested to construct annual mortality tables for the regions and the provinces of Spain. For the first approach, abbreviated life tables by five-year age intervals are computed. For the second option in a first step, complete life tables for all sub national populations considered are calculated; in a second step, aggregated results by five-year age intervals are obtained. The third method consists of adjusting the survivors series of the complete mortality tables using the *Brass logits model*¹ to derive the rest of the life table series.

3.1 Option 1.- Abbreviated life table

Under this option, the denominator of the specific mortality rate, $m(t,x,n,s)$, is computed just assuming uniform distribution of birthdays of all individuals of the population not dying over the year at the age x by means of the following expression:

$$1/2 \cdot P(t,x,n,s) + 1/2 \cdot P(t+1,x,n,s) + \sum_i b(t,x,n,s,i)$$

Where:

$P(t,x,n,s)$ is the stock of resident population in the considered region, at January the 1st of year t , in the age interval $[x,x+n)$ and sex s .

$b(t,x,n,s,i)$ is defined as the difference (in years) between the dates of death and birth in the year t , of each individual i , of sex s , dying in year the t with age in the interval $[x,x+n)$.

Hence:

$$m(t,x,n,s) = \frac{D(t,x,n,s)}{1/2 \cdot P(t,x,n,s) + 1/2 \cdot P(t+1,x,n,s) + \sum_i b(t,x,n,s,i)} \quad \text{for } x = 0,1,5 \dots w + .$$

Where:

$D(t,x,n,s)$ is the number of deaths of individuals with age in the interval $[x,x+n)$, of sex s , among the resident population in the considered region during the year t .

¹ Brass, William. Methods for estimating fertility and mortality from limited and defective data.(1975)

Then, the following key expressions for some of the biometric functions are adopted:

The **probability or risk of dying**, $q(x,n,s)$ is derived by:

$$q(x,n,s) = \frac{n \cdot m(t,x,n,s)}{1 + n \cdot (1 - a(x,n,s)) \cdot m(t,x,n,s)} \text{ for } x = 0,1,2,..,99 .$$

$$q(100+,s) = 1$$

Being he **average number of years lived in the last year of life** for those individuals dying in the age interval $(x,x+n)$, $a(x,n,s)$, approximated by the average of the time (years) lived in t for the individual i dying in the age interval $(x,x+n)$ during the observed year t , $a(x,n,s,i)$:

$$a(t,x,n,s) = \frac{\sum_{i=1}^{N(t,x,s)} a(t,x,n,s,i)}{D(t,x,n,s)} , x = 0,1,5,...,w - 5 .$$

Besides, for the open-ended group the average time lived in the reference year t is estimated for ages of w and above and sex s , $a(w+,s)$, as:

$$a(w+,s) = \frac{1}{m(w+,s)}$$

3.2 Option 2.- Aggregated results from complete life tables

The starting-point for this alternative is a complete life table, constructed, for each sub-national population, according to the third option concerning total population, as described in the first part of the present paper. From the survivors function of this complete life table, a life table with aggregated results by groups of ages is defined, as described in the following points.

The **survivors function**, $l(x,s)$, for the life table with aggregated results takes the values of the same function of the complete life table for ages $x = 0,1,5,10,15,...,w - 5, w$.

Then, the **deaths series**, $d(x,n,s)$, can be calculated as:

$$d(x,n,s) = l(x,s) - l(x+n,s) \text{ for } x = 0,1,5,10,15,...,w - 5, w + \text{ and } n = 1,4 \text{ or } 5 \text{ (as applicable).}$$

From the deaths and the survivors functions, the **probability of dying** within the age interval $[x,x+n)$, $q(x,n,s)$, is estimated by the expression:

$$q(x,n,s) = \frac{d(x,n,s)}{l(x,s)} \text{ for } x = 0,1,5,10,15,...,w - 5 .$$

$$q(w+,s) = 1$$

And, the **stationary population** at ages $[x, x+n)$ and sex s , $L(x, n, s)$ is developed from the same function of the starting complete table, $L(x, s)$, as follows:

$$L(x, n, s) = \sum_{x \leq y \leq x+n} L(y, s), \text{ for } x = 0, 1, 5, 10, 15, \dots, w-5, w+.$$

Besides, **average number of years lived in the last year of life** by individuals of sex s dying within the age interval $[x, x+n)$ during year t , $a(x, n, s)$, is obtained from:

$$a(x, n, s) = 1 - \frac{n \cdot l(x, s) - L(x, s)}{n \cdot d(x, s)} \text{ for } x = 0, 1, 5, 10, 15, \dots, w-5.$$

$$a(w+, s) = \frac{1}{m(w+, s)}$$

Eventually, the **mortality rates** series, $m(x, n, s)$, is deduced from the following formula:

$$m(x, n, s) = \frac{d(x, n, s)}{L(x, n, s)} \text{ for } x = 0, 1, 5, 10, 15, \dots, w-5, w+.$$

3.3 Option 3.- Complete life tables adjusted by the Brass logits method

The *Brass method* establishes a functional relation between the survivors function, $l(x, s, h)$, of the complete life table of each territorial unit h and the survivor function of the complete life table for Spain, $l(x, s, \text{Spain})$:

$$\text{logit}(x, s, h) = \alpha(s, h) + \beta(s, h) \cdot \text{logit}(x, s, \text{Spain}) + \varepsilon(x, s, h)$$

Where:

$\alpha(s, h)$ and $\beta(s, h)$ are the parameters of the regression model and $\varepsilon(x, s, h)$ is the random error term in the model, and:

$$\text{logit}(x, s, h) = \frac{1}{2} \cdot \ln\left(\frac{100.000 - l(x, s, h)}{l(x, s, h)}\right)$$

$$\text{logit}(x, s, \text{Spain}) = \frac{1}{2} \cdot \ln\left(\frac{100.000 - l(x, s, \text{Spain})}{l(x, s, \text{Spain})}\right)$$

The parameters of the linear model between logistic transformation of the survivor functions for the reference territorial unit h are estimated by *Least Squared Method*. Then estimated survivors function, $\hat{l}(x, s, h)$, results from the estimated model.

After that, the following expressions for some of the key biometrics functions of the complete life table are adopted:

$$\hat{d}(x, s, h) = \hat{l}(x, s, h) - \hat{l}(x + 1, s, h)$$

$$\hat{q}(x, s, h) = \frac{\hat{d}(x, s, h)}{\hat{l}(x, s, h)}, \text{ for } x = 0, 1, \dots, 99 .$$

$$\hat{l}(0, s, h) = 100.000$$

$$\hat{d}(100+, s, h) = \hat{l}(100+, s, h)$$

$$\hat{q}(100+, s, h) = 1$$

In addition to that, the robustness of infant mortality in all territorial levels considered allows to impute the **probability of dying at age 0** with the observed value.

Besides, the **average number of years lived in the last year of life** for those individuals dying in the age interval $[x, x + 1)$, $a(x, s, h)$, is taken 1/2, except for the age 0 where the observed value over studied population during the year t is taken.

Finally, the **stationary population**, $L(x, s, h)$, is calculated as follows:

$$\hat{L}(x, s, h) = \hat{l}(x + 1, s, h) + \hat{a}(x, s, h) \cdot \hat{d}(x, s, h), \quad x = 0, 1, \dots, 99 .$$

And for the open-ended age group is approximated by:

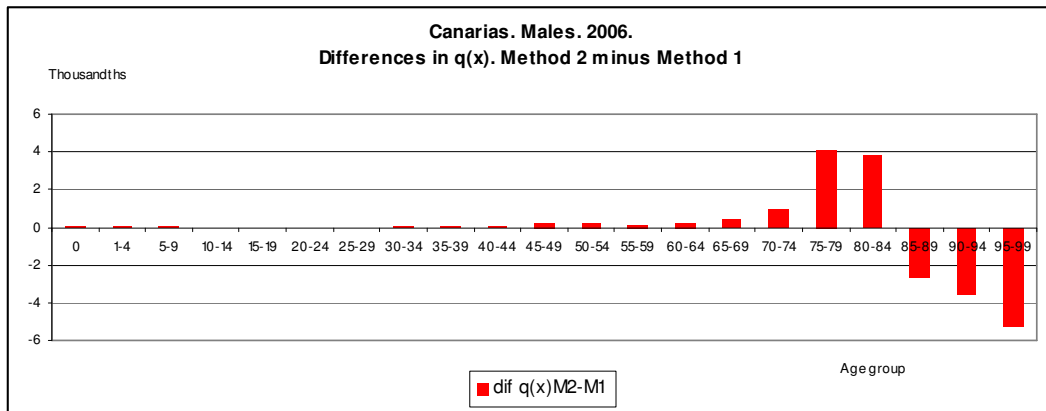
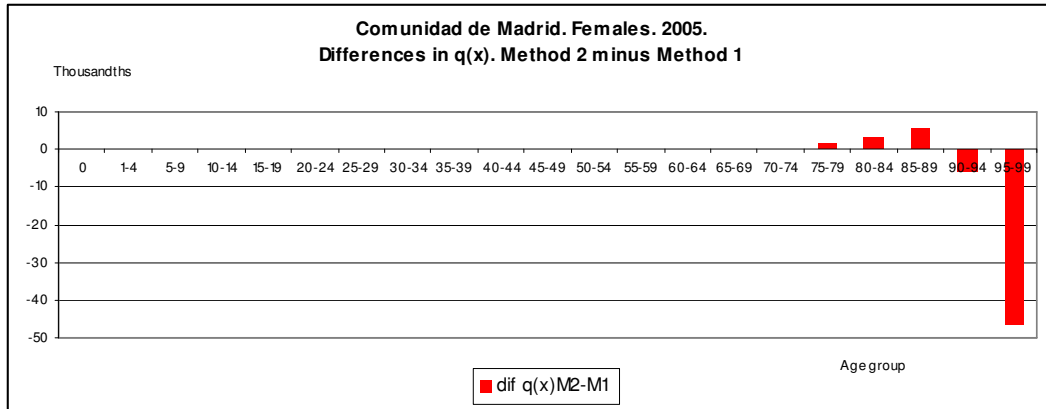
$$\hat{L}(100+, s, i) = \frac{\sum_{95 \leq x \leq 99} \hat{l}(x, s, i)}{\sum_{95 \leq x \leq 99} l(x, s, \text{Spain})}$$

3.4 Discussion

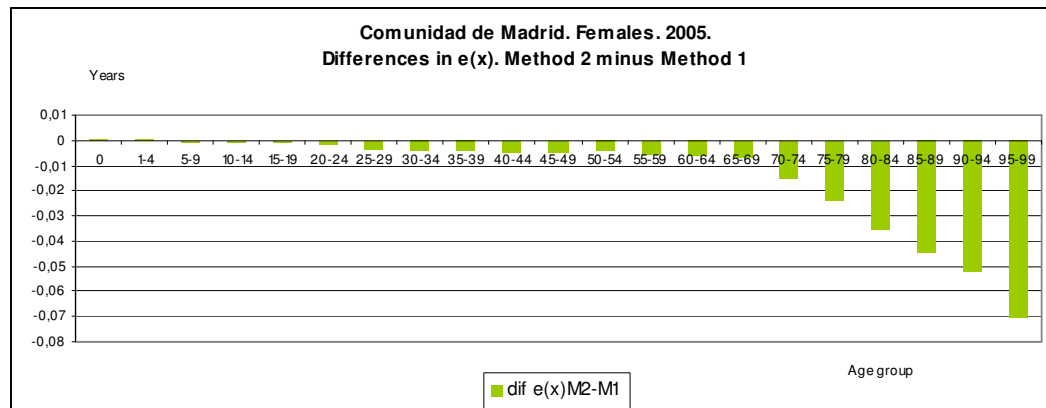
The option 1 suffers from the limitation of not isolating the effect of the age structure within each age interval. On the contrary, option 2 eliminates the effect of the age structure of the population, even within every age interval.

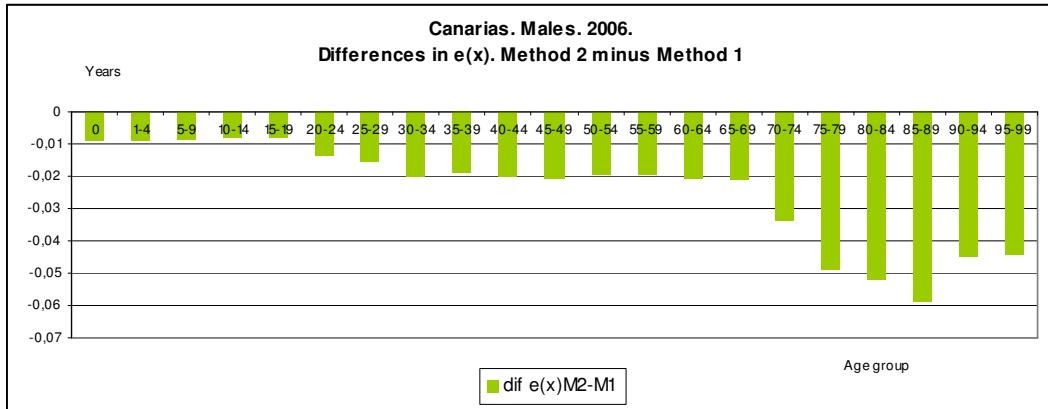
In addition, the deviations in the risks of death between the two methods must therefore arise from this limitation of option 1 with respect to option 2, which leads to a less accurate measurement of the incidence of mortality in the reference population independently of the age structure of the population. The deviations increase with distance from uniformity of the distribution of deaths in each age interval. The effect is imperceptible in young and adult age groups, but becomes visible in the upper age intervals.

There follow some examples exhibiting the differences between the results discussed:

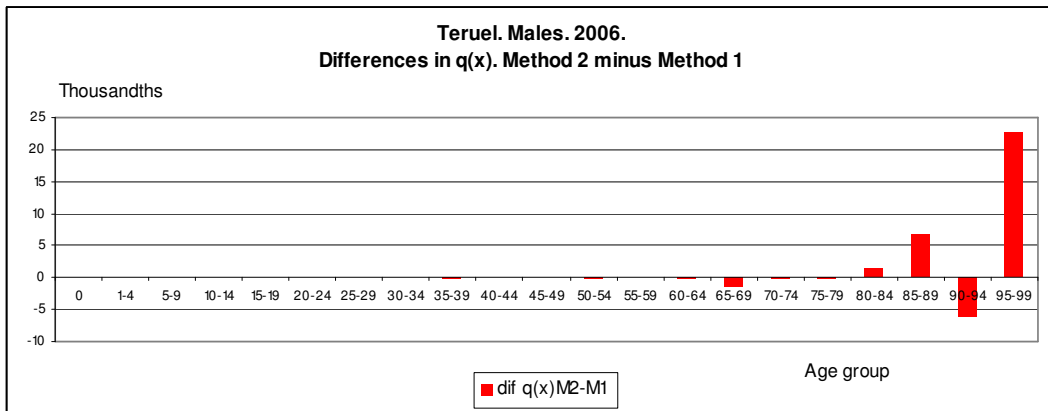


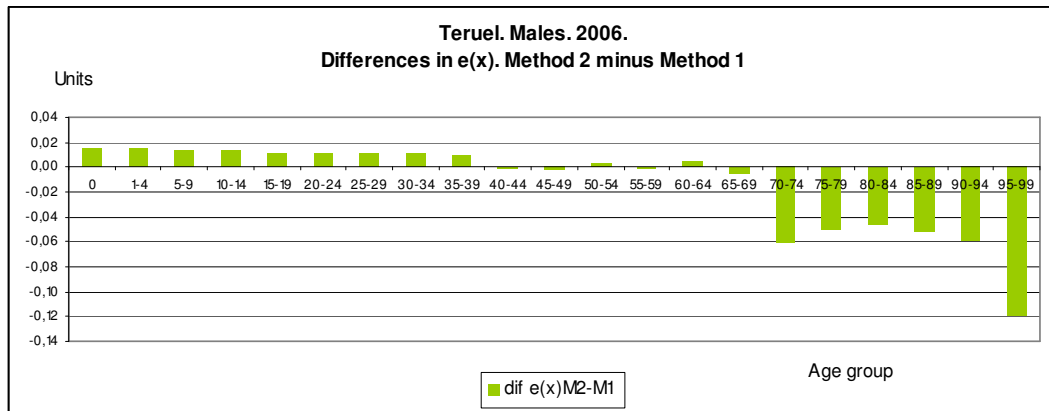
These deviations in the risks of death obtained by each option in the upper age intervals entail that option 1 quantifies life expectancies at all ages that are slightly above the values produced by option 2:





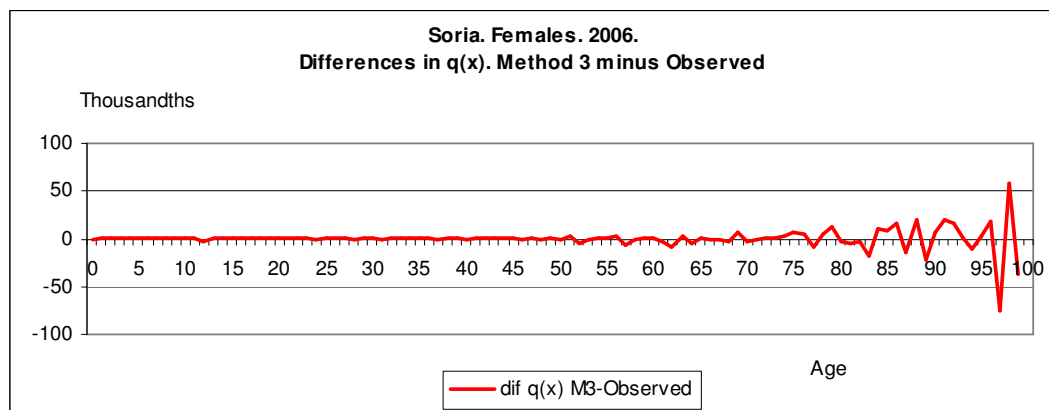
However, the randomness of data on deaths by age recorded in a given year in the smaller regions and provinces, and the greater relative error that the population estimates used for such territorial units may have, at advanced ages especially, do of course distort the visibility of this effect. This is the case of territories such as Teruel where the deviations between risks of death and life expectancies resulting from the two methods are harder to interpret:



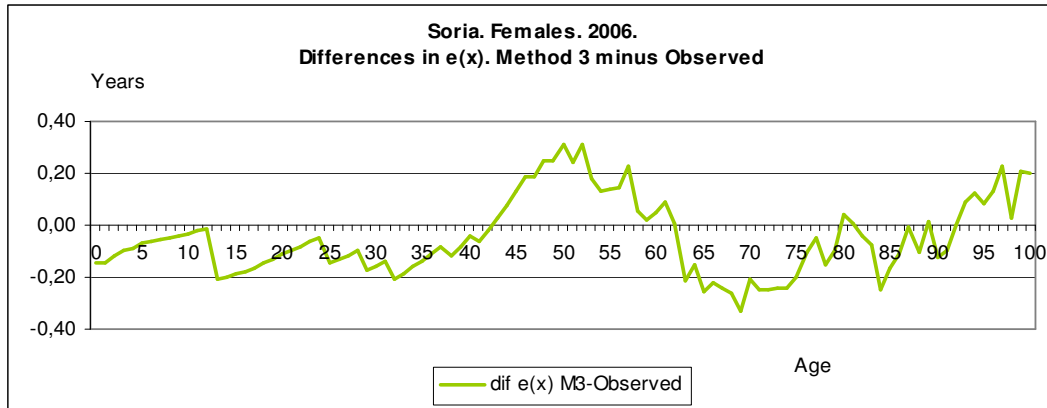


But both options have the drawback of providing less detailed results than a complete table by simple ages. However, they mitigate the randomness of results that any complete life table for a small population displays and, in general, they avoid the problems of a very low or null number of deaths observed at some simple ages in such populations.

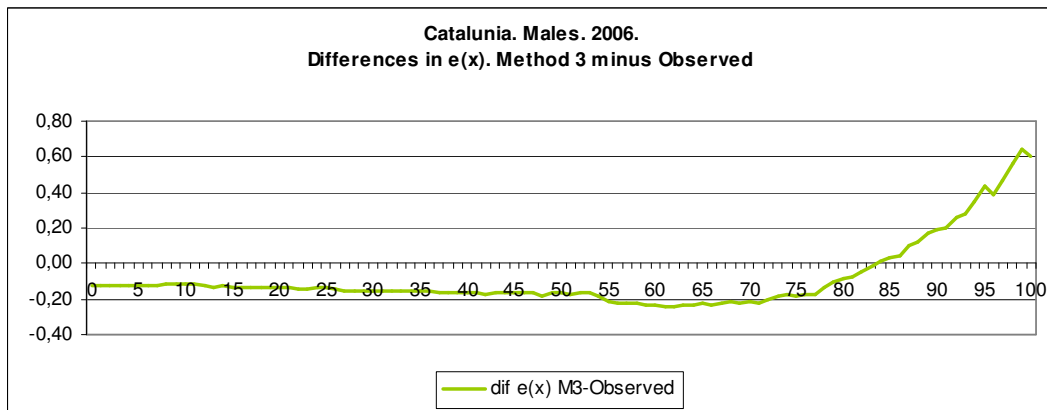
On the other hand, option 3 allows for constructing a complete life table. The method emerges from an adjustment of observed data designed to eliminate intrinsic randomness and to avoid inconsistencies imported by the population estimates used, while achieving the aim of measuring the incidence of mortality for the reference period. If we compare the results of this methodological option to the results of a complete life table constructed directly, at ages under 90, no significant differences are perceptible in the risks of death, as a general rule. Besides, deviations increase with age and decrease with population size. Above age 90, the differences are larger, especially in the smaller populations. By way of example, we present the differences between the risks of death resulting from option 3 and those provided by a complete life table calculated directly for small populations, such as the female population of Soria in 2006:



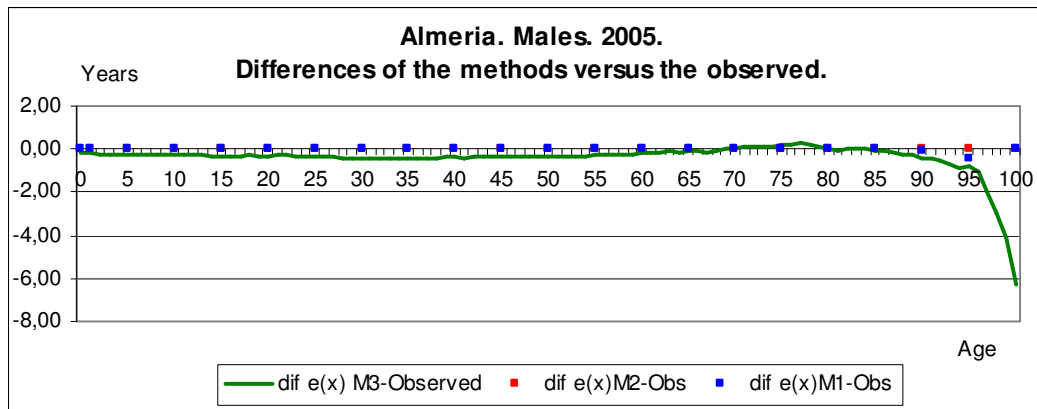
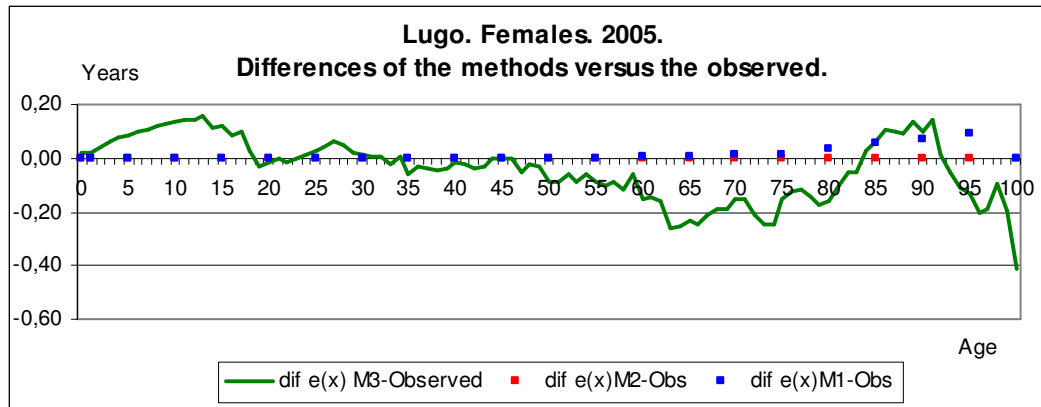
Respect to life expectancies at birth resulting from option 3 and from a complete life table constructed directly, the differences are negligible in most cases. In fact, they exceed six months only in most extreme cases. Furthermore, the differences are also seen to decrease with reference population size.



In addition to all that, the main problem of option 3 is the distortion imported into life expectancies at the most advanced ages with respect to observed life expectancies upon direct construction. This is observed, for example, in the male population of Cataluña in 2006:



The distortion created by the Brass method in the upper age intervals is absent from options 1 and 2, given the results obtained even in the smaller populations:



This methodological approach, on the other hand, is the only one that provides mortality curves and other functions by simple ages, while avoiding the undesirable effects of the intrinsic randomness of observed annual deaths in small populations and potential inconsistencies between population estimates used as a reference and the results obtained.

4 Conclusions

According to the discussions related to national life tables, the three options are designed to capture the annual behaviour. They also keep consistency in the definitions and calculation assumptions stated. Nevertheless, option 3 is the one that maximizes the use of the available data from Spanish statistical system, getting maximum accuracy in measuring the incidence of mortality and the living time over the whole range of ages, specially in advanced ages.

However, this third option for the national life tables is not suitable for regional life tables due to the random factors that punish small size populations. Besides, first and third option proposed for regional life tables are rejected since they are not consistent with national life tables and they present some other mentioned handicaps. On the other hand, the second option keeps all the good properties followed by the chosen option for national tables.

References

J.R.Wilmoth, K.Andreev, D.Jdanov, and D.A.Glei. Methods Protocol for the Human Mortality Database.

<http://www.mortality.org/Public/Docs/MethodsProtocol.pdf>

Julio Vinuesa and Dolores Puga. Técnicas y ejercicios de Demografía. (2007)

Brass, William. Methods for estimating fertility and mortality from limited and defective data.(1975)

INE. Tablas de mortalidad. Metodología. (2009)

http://www.ine.es/daco/daco42/mortalidad/metodo_9107.pdf