DURATION MODELS

Lecture Notes 2015-2016

Mortality Tables

Frederic PLANCHET / Olivier DURAND

May 2016

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} I_{]t;\infty[} (T_x)$$

Table of content

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

1. Analysis of mortality

One is interested in the random variable *T* representing the lifetime of an individual; one supposes the individuals of the population initially identical, so that one will be able to have samples resulting from the distribution of *T*.

1.1. Notations

It is convenient to consider the variable T_x representing the residual lifetime of an individual conditional to the fact that he is alive at age *x*, i.e. $T_x =_d \left[T - x \mid T > x \right]$. One can then define the probability of survival between *x* and $x + t$ by:

$$_t p_x = P(T_x > t) = P(T > x + t \mid T > x),$$

and the quotient of mortality between *x* and $x + t$:

$$_t q_x = 1 -_t p_x = P(T_x \le t) = P(T \le x + t \mid T > x).$$

When $t = 1$ it is omitted in the notations, and one writes more simply $q_x =_1 q_x$ and $p_x =_1 p_x$.
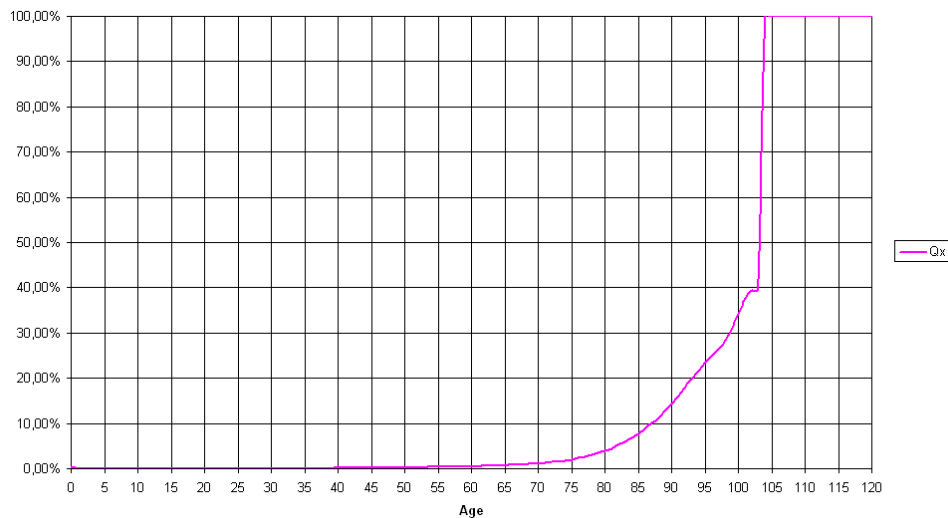
Fig. 1 :   *Instantaneous rate of mortality as a function of age*

These quotients are expressed simply using the survival function of *T*:

$$_t p_x = \frac{S(x+t)}{S(x)}.$$

It is usual to note[1] $l_x = S(x)$; the number of deaths between *x* and $x + t$ is noted $_t d_x = l_x - l_{x+t}$; within the framework of the statistical analysis of the mortality of a cohort one measures the time lived by the individuals of the cohort between *x* and $x + t$, defined by:

---

[1] A normalising constant may be used.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

$$_tL_x = \int_0^t l_{x+u} du\,.$$

Fig. 2 : *Representation of the survival function*

Starting from this indicator one can define the residual lifetime, which is a characteristic indicator of the mortality table:

$$E_x = \int_0^{+\infty} l_{x+u} du = \sum_{i=x}^{\infty} L_x\,.$$

The quotient of mortality $_tq_x$ is calculated by dividing a number of deaths over the period by headcount at the beginning of period; one also calculates the mortality rate, obtained by dividing the number of deaths by average headcount over the period, that is to say:
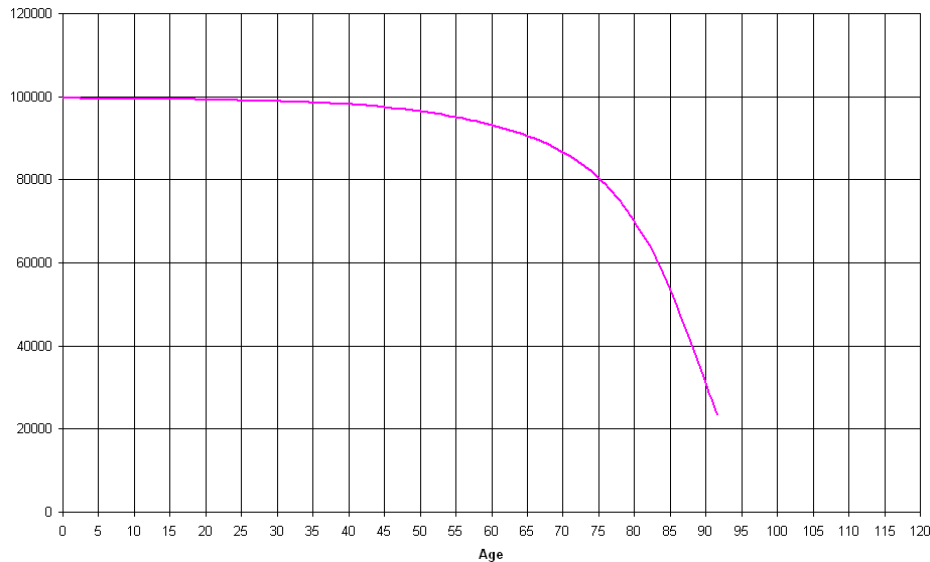
$$_tm_x = \frac{_td_x}{_tL_x}\,.$$

The quotients of mortality are probabilities (numbers without dimension) whereas the rates of death are expressed in reverse time unit and count deaths per person under risk and unit of time. This difference leads to the following relationships with the hazard function, called in this context "instantaneous rate of mortality":

$$\mu_{x+t} = \lim_{h\to 0} h^{-1} P\left(t < T_x \le t+h \,\middle|\, T_x > t\right) = \frac{1}{_tp_x} \frac{\partial}{\partial t}\,_tq_x,$$

Because $h^{-1} P\left(t < T_x \le t+h \,\middle|\, T_x > t\right) = \frac{_{t+h}q_x - _tq_x}{h \,_tp_x}$; thus when h is small, $_hq_x \approx h\mu_x$ and $_hp_x \approx 1 - h\mu_x$. The link between the instantaneous rate of mortality and mortality rate is direct: $\mu_x = \lim_{h\to 0}\,_hm_x$, which justifies *ex-post* the terminology.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

The relation between function of conditional survival and hazard function is written with the following notations:

$$_t p_x = \exp\left(-\int_0^t \mu_{x+s}\, ds\right).$$

1.2. The Lexis diagram

When studying mortality, it is rare to have access to exact information concerning ages and dates of death; these data are generally available in rounded figures: ages in integers and dates in years. In order to correctly determine the raw death rates in this context, one uses a particular formalism: the Lexis[2] diagram.

1.2.1. Presentation

The analysis of the mortality of a given group is based on three measurements of time: the age of individuals, their generation (date of birth) and the observation date; of course these 3 informations are linked and the knowledge of only 2 determines the third.

Each one of these dimensions has however its importance in the determination of the level of mortality:

- ✓ Age: this variable obviously influences the risk of death;

- ✓ Observation date: the risk of death can vary according to circumstances like an epidemic, an extraordinary event (heat wave of summer 2003 for example), *etc.*

- ✓ Generation: phenomena such as the improvement of sanitary conditions and progress of medicine result in modifying the risk of mortality at a given age in the course of time; moreover, one can imagine that the past of a given generation can modify the level of its future mortality: typically, an epidemic intervening on a date *t* and concerning people of age *x* on this date can contribute to decrease the rates of death at the ages higher than *x* for this generation, by resulting in the premature death of the least resistant individuals.

It is then convenient to represent the life of an individual in a rectangular axis system called "Lexis diagram", in the following way:

---

[2] From the name of the german statistician and demograph Wilhelm LEXIS (1837-1914).

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbb{1}_{]t;\infty[}(T_x)$$
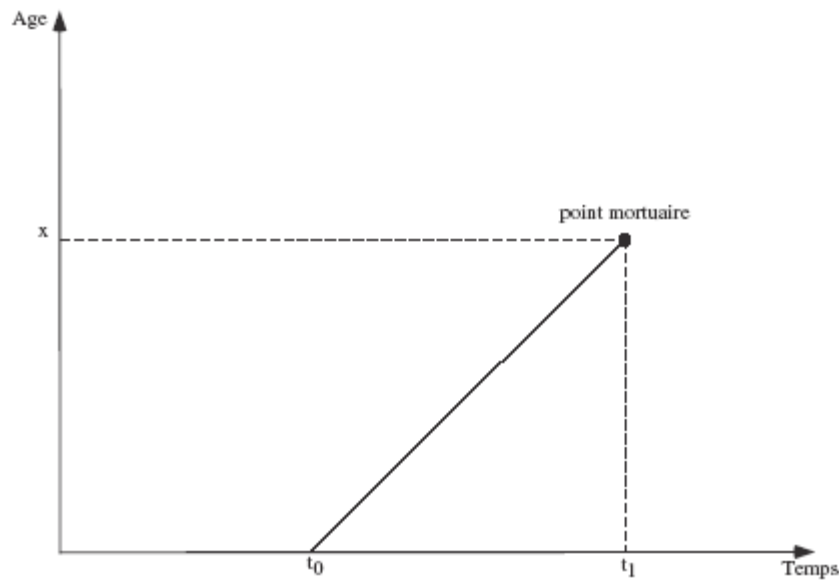
Fig. 3 :   *Lexis diagram*

The life of an individual is thus represented by a line parallel to the first bisectrix, which cuts the x-axis at year of birth and stops at the "funeral point" which is the day of death. By tracing a horizontal band between x and $x+1$ one isolates the individuals died at age x, and by tracing a vertical band between g and $g+1$, one isolates the deaths from the individuals of the generation g. In this formalism, x and g are integers, and x measures the age in completed years.

1.2.2. Lexis diagram and measurement of mortality

The funeral points which are in the square below are associated with the deaths at age x during the year t:
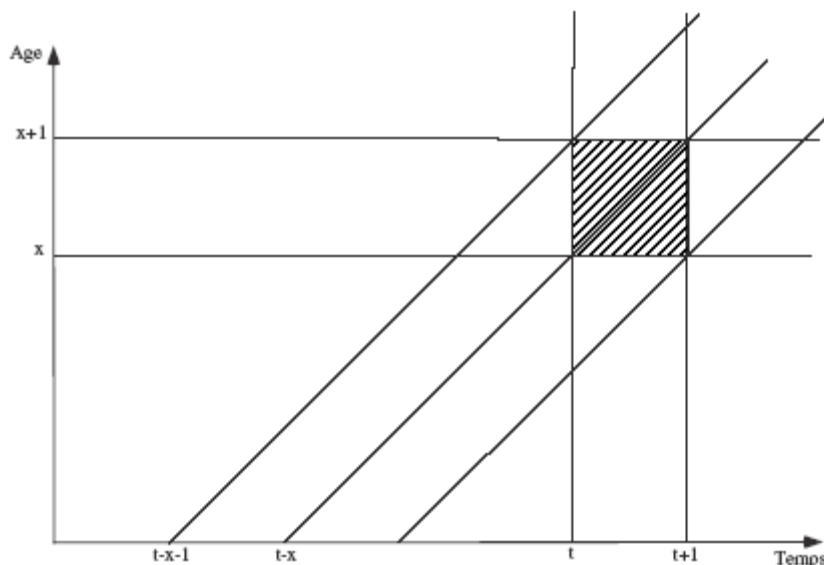
Fig. 4 :   *Identification of the deaths at age x on year t*

The individuals concerned belong to generations t-x and t-x-1. In the same way one can deduct the number of deaths at age x among the generation g:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

ressources-actuarielles.net
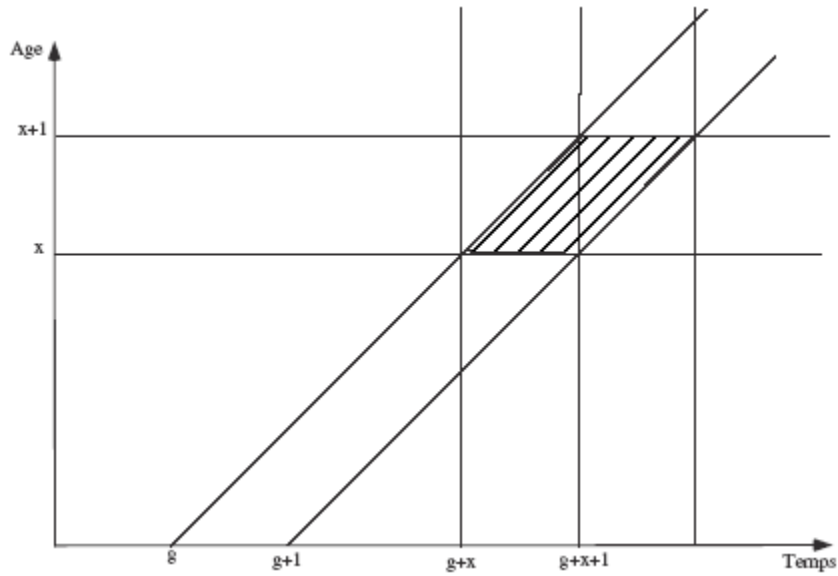
Fig. 5 :   *Identification of the deaths at age x in generation g*

These deaths occurred during years *g+x* and *g+x+1*. One also obtains the number of deaths at age *x* among generation *g* during year *t*:
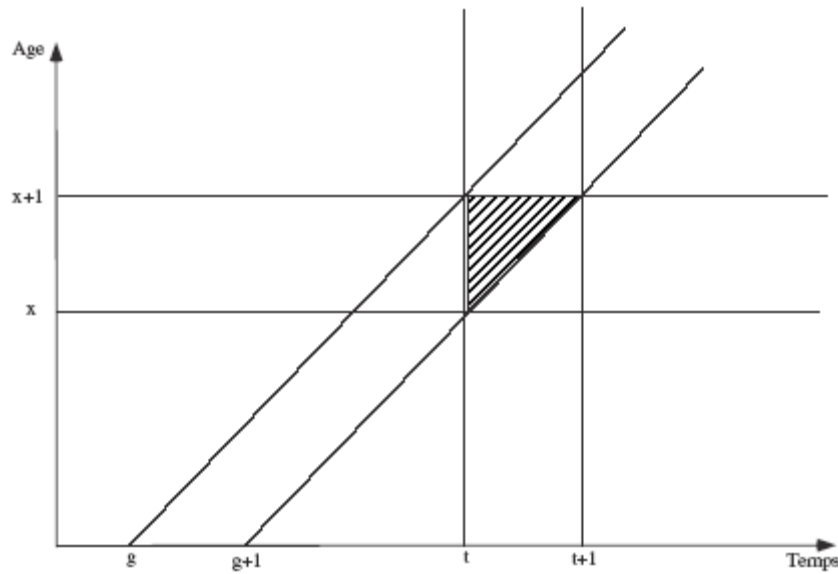
Fig. 6 :   *Identification of the deaths at age x on year t in generation g*

Lastly, one can represent in the same manner the number of deaths during year *t* among the individuals of generation *g*:

7

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} I_{]t;\infty[}(T_x)$$
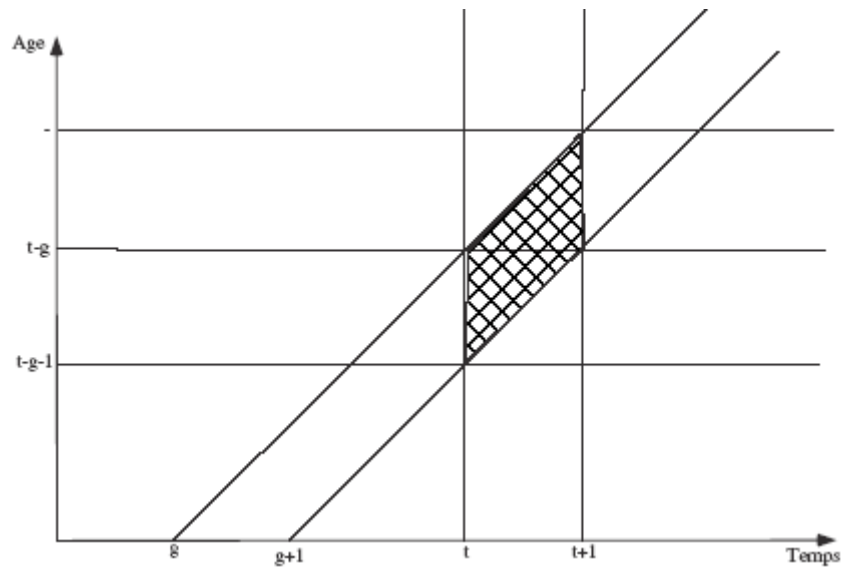
Fig. 7 :   *Identification of the deaths in generation g on year t*

1.3. Longitudinal mortality and transverse mortality

The "natural" measurement of mortality consists in counting the deaths which have occurred during a given period (one year for example), then calculating the rates of death per age by dividing this count by headcount under risk. Which amounts to considering a vertical band of the Lexis diagram.

It can be seen that, should mortality evolve in the course of time, this approach introduces a bias in the measurement of mortality – more precisely, during times of decreasing mortality trend, it results in underestimating the lifetimes (or over-estimating rates of death). Indeed, in this approach one considers individuals of separate generations to calculate the rates of death – the obtained table thus does not represent the mortality of any real generation.
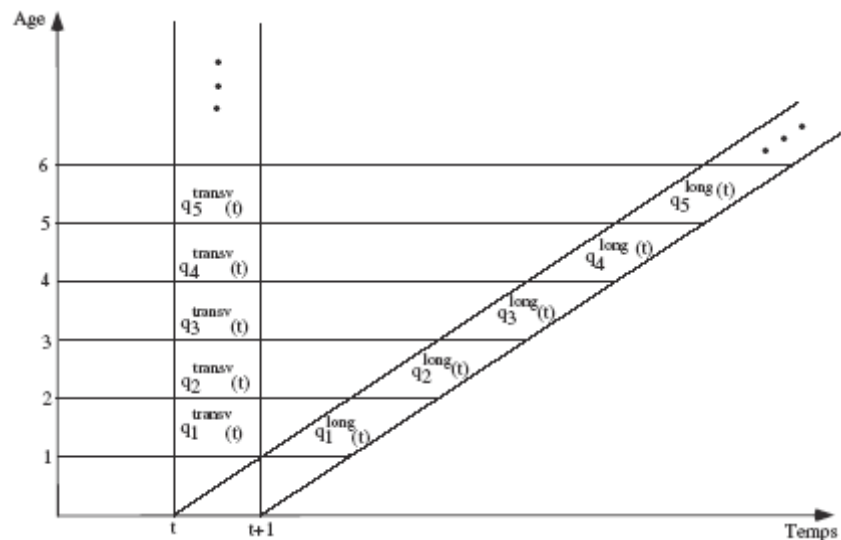
Fig. 8 :   *Longitudinal mortality and transverse mortality*

The actual/real mortality of a generation is obtained by considering the rates along a digonal band, as shown above. One will thereafter need to calculate the quotient of mortality at age x

8

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbb{1}_{]t;\infty[}(T_x)$$

for year t; as one saw in 1.2.2 above, this quotient utilises two generations, $t - x$ and $t - x - 1$. One thus determines the "partial quotients of mortality", by noting $D_{xt}(g)$ the number of deaths at age x for generation g, which occurred on year t:

$$\hat{q}_{xt}^1 = \frac{D_{xt}(t-x)}{l_{x,t+1} + D_{xt}(t-x)}.$$

($l_{x,t+1}$ denotes the number of people of age x as of the date 01.01.t+1); this quotient thus approaches the probability for individuals of generation $t - x$, of dying at age x on year t. One estimates in the same way the probability for individuals of generation $t - x - 1$ of dying at age x on year t:

$$\hat{q}_{xt}^2 = \frac{D_{xt}(t-x-1)}{l_{xt}}.$$

The sought quotient then results from the aggregation of these 2 partial quotients: to survive between one's x^{th} and its $(x+1)^{th}$ birthday, it is necessary to survive from one's x^{th} birthday to the end of the calendar year, then from the end of the calendar year to one's $(x+1)^{th}$ birthday, that is to say:

$$1 - \hat{q}_{xt} = \left(1 - \hat{q}_{xt}^1\right)\left(1 - \hat{q}_{xt}^2\right).$$

When one wants to determine mortality rate at age x for year t, one calculates classically, with obvious notations:

$$\hat{m}_x = \frac{D_{xt}}{(l_{x,t} + l_{x,t+1})/2}.$$

1.4. Distribution of deaths within a year

Available data are often gathered data in which the unit of time is the year. It is then advisable to give oneself a rule of distribution for the deaths within a year. This point was approached previously; three assumptions are classically proposed:

✓ constant instantaneous rates of death between 2 non-integer ages (exponential assumption): ${}_t q_x = 1 - (1 - q_x)^t$;

✓ linear distribution of deaths during a year: ${}_t q_x = 1 - t \times q_x$;

✓ Balducci assumption, which postulates that: ${}_t q_x = \dfrac{t \times q_x}{1 - (1-t)q_x}$.

The Balducci assumption can be dismissed right from the start because it leads to instantaneous rates of mortality decreasing between 2 integer ages; indeed, one finds in this model that:

$$\mu_{x+t} = -\frac{\partial}{\partial t} \ln\left({}_t p_x\right) = \frac{q_x}{p_x + tq_x},$$

which results from $_t p_x = 1 -_t q_x = 1 - \dfrac{tq_x}{p_x + tq_x} = \dfrac{p_x}{p_x + tq_x}$. The choice between the 2 remaining assumptions is not neutral on the appreciation which one will have of the level of mortality. Indeed, if T_x^c and T_x^l are the residual lifetimes respectively in the constant instantaneous rates model and in the linear distribution of deaths model, one has, with obvious notations: $S_x^l(t) \ge S_x^c(t)$, which implies in particular that $e_x^l(t) \ge e_x^c(t)$; the assumption of constant instantaneous rates thus leads to lower lifetimes: thus a prudent assumption in the case of guarantees in the event of death, and less prudent in the case of annuities. However, the difference between the 2 approaches is minor.

To prove the inequality $S_x^l(t) \ge S_x^c(t)$, one sets $t = k + r$, with $k = [t]$ and $0 \le r < 1$ one notes that:

$$S_x^l(t) = P(T_x^l > k + r) =_k p_x (1 - rq_{x+k})$$

and

$$S_x^c(t) = P(T_x^c > k + r) =_k p_x p_{x+k}^r.$$

The inequality to be shown is thus equivalent to $1 - r(1 - p_{x+k}) \ge p_{x+k}^r$, and this last inequality is the direct consequence[3] of $(1 + x)^r \le 1 + rx$ for all $0 \le r < 1$. In the models presented hereafter, the assumption of constant instantaneous rates of mortality between 2 integer ages is made, so that one has $_t q_x = 1 - (1 - q_x)^t$.

1.5. Synthetic indicators of the level of mortality

The characteristics of a mortality table are usually summarised by some indicators: life expectancy and entropy are two important indicators, presented hereafter.

1.5.1. Residual life expectancy

The residual life expectancy is by definition $e_e = E(T_x) = \dfrac{E_x}{l_x}$; one thus has $e_x = \dfrac{1}{l_x} \int_0^{+\infty} l_{x+u} du$;

one deduces in particular from this expression that:

$$\frac{de_x}{dx} = -1 + \mu_x e_x.$$

Indeed, $\dfrac{d}{dx} e_x = \dfrac{-l_x^2 - \left(\dfrac{d}{dx} l_x\right) \int_x^{+\infty} l_u du}{l_x^2}$ and because $\mu_x = -\dfrac{d}{dx} \ln l_x$, the above equality can be written. The discrete version of this formula is simply $e_x = \dfrac{1}{L_x} \sum_{h>0} L_{x+h}$.

---

[3] There even is strict inequality if $r > 0$.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}\left(T_x\right)$$

This expression means that when mortality rates are small, the residual life expectancy decreases by approximately a year every year; on the other hand, when mortality rates become large, one can have a residual life expectancy which increases.

From a practical point of view, that means that the graph of the e_x is almost aligned on a line of slope - 1 until around 75 years, then curving up as can be seen on the graph below:
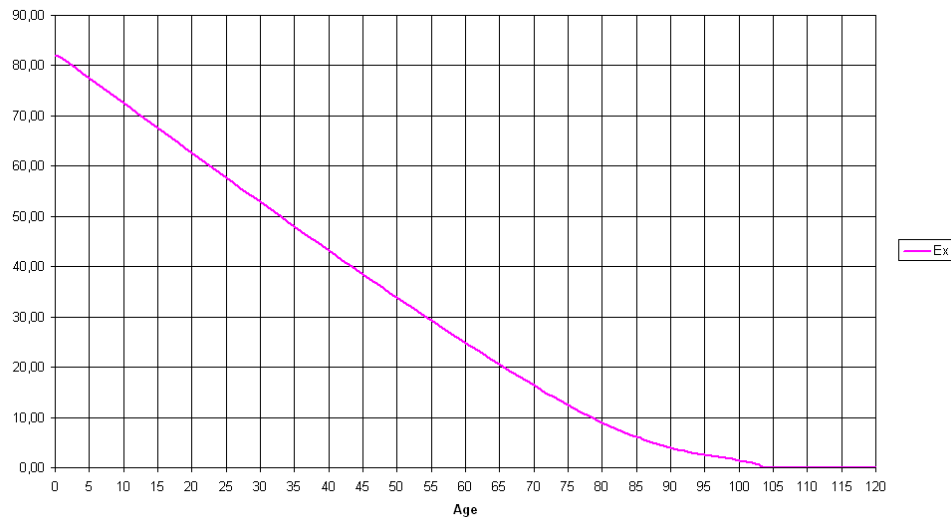
Fig. 9 :   *Residual life expectancy as a function of age*

From 75 years old onwards, a polynomial adjustment of order 2 works in general properly (which provides a simple parameterisation of a period mortality table). It can be noted that the residual life expectancy can be interpreted as the price of a continuous life annuity with zero discount rate.

1.5.2. Entropy

The decrease of mortality rates at young ages, yet with an ultimate lifespan seemingly remaining constant, has for consequence a phenomenon of "orthogonalisation" of the mortality tables, the number of people dying at high ages[4] becoming larger and larger:

---

[4] This phenomenon is accompanied by a decrease of the variance of lifetime with time.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$
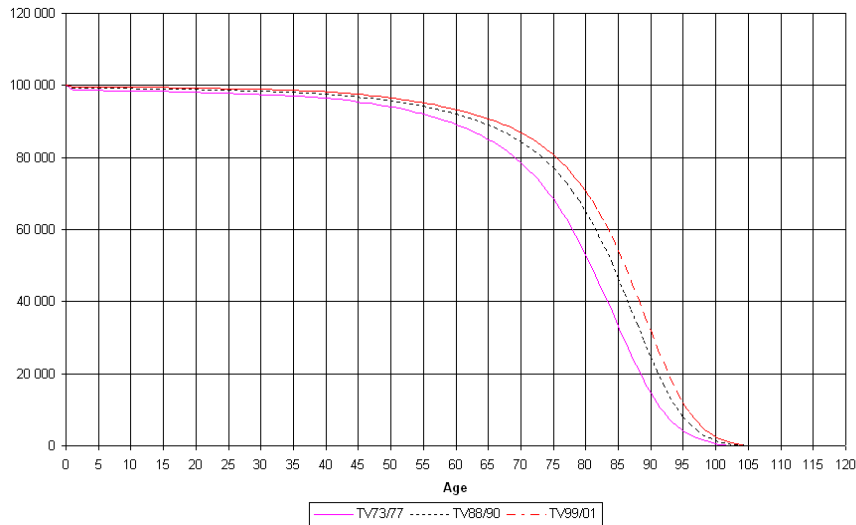
Fig. 10 :   *Illustration of the phenomenon of orthogonalisation of mortality tables*

The entropy proposes to measure this phenomenon; one defines it by:

$$H = -\frac{\displaystyle\int_0^{+\infty} l_x \ln(l_x)\, dx}{\displaystyle\int_0^{+\infty} l_x\, dx}.$$

As one has $\mu_x = -\dfrac{d}{dx}\ln(l_x)$, one can rewrite this quantity in the form:

$$H = \frac{\displaystyle\int_0^{+\infty} l_x \mu_x e_x\, dx}{l_0 e_0}.$$

The discrete version of this formula is $H = -\dfrac{\displaystyle\sum_{h>0} L_{x+h} \ln(L_{x+h})}{\displaystyle\sum_{h>0} L_{x+h}}$.   Entropy thus divides the

average number of "lost years" because of death by the number of possible years "in store" at date 0.

One can notice that $H = 0$ if and only if all of the deaths occur at the same age and that $H = 1$ corresponds to the extreme opposite situation in which the instantaneous rate of mortality is constant: this measure is therefore well adapted to the assessment of the orthogonalisation phenomenon. Entropy went from approximately 50 % at the end of the 19th century to 15 % today.

2. Some indicators

The objective of this section is to provide some useful orders of magnitude on the level of mortality. One also illustrates the manner of quantifying the impact on the mortality of specific characteristics of the population, by taking the example of the smoker/non-smoker criterion.

12

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \, \mathbb{1}_{]t;\infty[} \left(T_x \right)$$

2.1. General data

The life expectancies at birth and at 60 years old, as well as the rate of death at this age, are indicated in the table below:

	Women			Men		
	TV73/77	TV88/90	TV99/01	TD73/77	TD88/90	TD99/01
Birth	76.5	80.2	82.2	68.6	72.0	74.7
Age 60	20.9	23.5	25	16.1	18.3	19.9
q60	0.77%	0.57%	0.48%	1.90%	1.57%	1.18%

This table clearly reveals disparities between men and women:

	Women/men		
	TV73/77	TV88/90	TV99/01
Birth	112 %	111 %	110 %
Age 60	130 %	128 %	126 %
q60	41 %	37 %	41 %

One can also see a decreasing mortality trend:

	Women		Men	
	TV88/90/TV73/77	TV99/01/TV88/90	TD88/90/TD73/77	TD99/01/TD88/90
Birth	105 %	102 %	105 %	104 %
Age 60	112 %	106 %	114 %	109 %
q60	74 %	85 %	82 %	75 %

The mortality gap between men and women translates in the fact that, in developed countries[5], a sample of people older than 60 show 70 men for 100 women, and a sample of people older than 80 show 44 men for 100 women.

2.2. Impact of smoking

The impact of smoking on mortality is illustrated on the basis of:

✓ The study of male smoker/nonsmoker Canadian tables;

✓ Epidemiological studies conducted by the Ministry for Health in France.

2.2.1. Study of Canadian tables

Nonsmoking characteristic of policyholders has a favorable impact on their mortality. In Canada, studies led to the establishment, for men older than 30, of mortality tables differentiated by this characteristic. Their study enables us to quantify the impact of smoking on mortality.

The following graph shows the evolution of annual mortality rates with age and depending on whether the man smokes or does not smoke.

---

[5] As per a 1998 study from the *US bureau of the census*.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

ressources-actuarielles.net
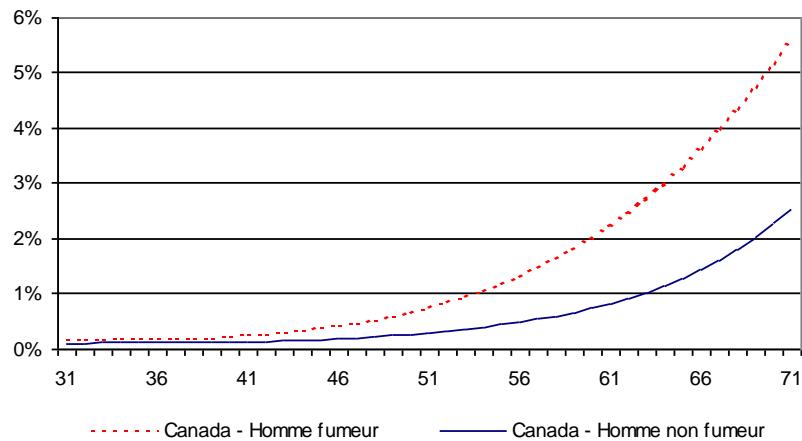
Fig. 11 :   *Comparison of male smoker vs nonsmoker rates of death*

As one could expect it, mortality rates of the smokers are systematically higher than those of the nonsmokers. The reduction of mortality of nonsmokers versus smokers shows a maximum at age 59 (63.35 %).
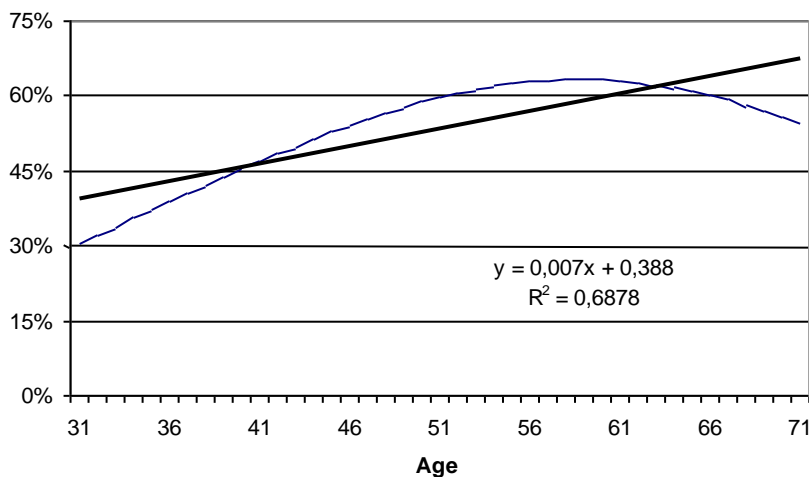
Fig. 12 :   *Relative reduction of nonsmokers rates of death compared to that of smokers (men)*

The average reduction between ages 31 and 71 amounts to 53.4 %. This analysis makes it possible to measure the impact of smoking on mortality. Compared to the total population, the reduction for nonsmokers is about 20 % on average between ages 31 and 70.

14

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$
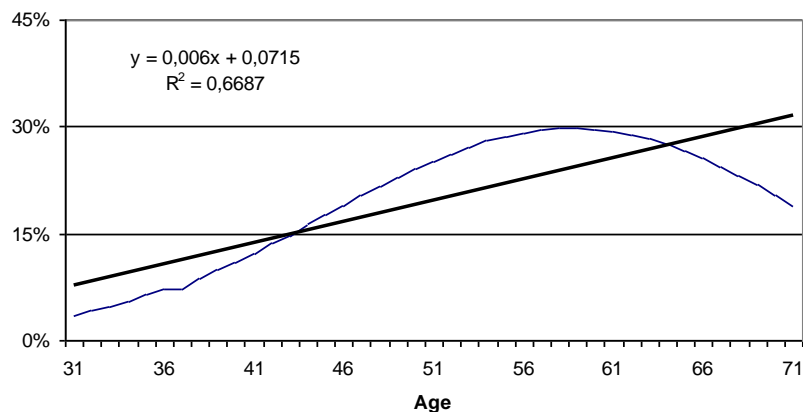
Fig. 13 :   *Relative reduction of nonsmokers rates of death compared to the general population (men)*

A maximum is reached at age 58 with a reduction of almost 30%.

2.2.2. Study of French Ministry for Health

The study "smoking and mortality: epidemiologic aspects" provides interesting indicators making it possible to quantify the under-mortality of nonsmokers. The key elements can be summarised as follows:

- ✓ Between ages 39 and 65, 1 death out of 3 for men is related to smoking and 1 death out of 16 among women.

- ✓ Between ages 35 and 49, 40 % of men and 29 % of women are regular smokers. These percentages decrease respectively to 28 % and 14 % between ages 50 and 64.

By making the assumption that these proportions are homogeneous on the indicated age segments, it is possible to estimate nonsmokers' under-mortality compared to smokers. Let us note:

- ▪ q_{tabac} the rate of over-mortality related to smoking,

- ▪ q the mortality rate of nonsmokers,

- ▪ π_F the proportion of smokers,

- ▪ π_{NF} the proportion of nonsmokers,

- ▪ δ the proportion of deaths due to smoking.

$$\frac{\pi_F \times q_{tabac}}{\pi_F \times (q + q_{tabac}) + \pi_{NF} \times q} = \delta$$

Thus:

$$\frac{q_{tabac}}{q} = \frac{\delta}{\pi_F \times (1-\delta)}$$

The rate of under-mortality of nonsmokers compared to smokers is thus written:

$$1 - \frac{q}{q + q_{tabac}} = 1 - \frac{\pi_F \times (1-\delta)}{\delta + \pi_F \times (1-\delta)}$$

The rates calculated based on the elements of the study of the Ministry for Health are summarised in the following table:

	Men	Women
Age 35-49	55.56 %	20.41 %
Age 50-64	64.10 %	32.26 %

Concerning men, the rates of under-mortality of nonsmokers compared to smokers (56 % and 64 %) are comparable with those resulting from the Canadian tables (53 %). These same rates are lower for women; however, female smoking is more recent and its impact less well assessed than that of men.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}\left(T_x\right)$$

By supposing that the proportions of smokers referred to above are homogeneous on all age segments, rates of under-mortality of nonsmokers compared to general population are given by:

	Men	Women
Age 35-49	33.33 %	6.25 %
Age 50-64	31.71 %	6.25 %

The male rates are slightly higher than what is observed with the Canadian tables. The female rates are clearly lower than the male rates.

3. The construction of custom mortality tables

We are in the parametric context; the construction of a custom table systematically comprises two stages: first of all the estimate of raw rates – by age, or age and generation in the case of prospective tables – then the adjustment of these raw rates to a parametric model.

In practice one can distinguish two situations: first of all, the "reference" situation in which one disposes of data in sufficient quantity to build a reliable table. But in certain cases it may be that the available data are not sufficient to determine in a precise enough way the structure of the table, and one will then seek to simply position the mortality of the studied group with respect to a benchmark mortality, which will provide the general structure.

3.1. Period tables

3.1.1. Full construction

The standard approach of construction of a mortality table within a parametric framework was described previously, it is therefore not detailed again here. It will only be reminded that it is based on the choice of a parametric form for the hazard function, with Makeham model as a reference model – the parameters estimate being carried out by the method of maximum likelihood.

In some specific situations, one will be able however to turn to other models, such as Poisson regressions; the typical example of application of such models is the analysis of the mortality of a rare event, such as the consequences of the exposure to asbestos. The justification of the interest of the Poisson distribution for "rare" events comes from the observation that the binomial distribution $\mathrm{B}\left(n, \dfrac{\lambda}{n}\right)$ converges in law towards $\mathrm{P}(\lambda)$ when $n \to \infty$, which is easily verified with the Laplace transform.

As the number of deaths is very weak compared to risk exposure, one can use a Poisson distribution as model for the number of deaths per age and period.

The following table gives the number of deaths caused by mesothelioma[6] noted by age group, for five periods, as well as the population under risk for this period:

---

[6] A cancer resulting from exposure to asbestos.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

	1970-74		1975-79		1980-84		1985-89		1990-95	
25-29	0	10041742	1	10978690	1	10602254	1	10680272	1	10791607
30-34	2	7720583	1	10038396	0	11005461	1	10651073	1	10837520
35-39	2	8074903	3	7589268	5	9904593	5	10922900	6	10657919
40-44	5	8510762	7	7879250	7	7457766	10	9761988	11	10853140
45-49	9	8211522	9	8220829	14	7662805	14	7265550	17	9464014
50-54	10	7173352	18	7821153	22	7866442	26	7354438	24	7022582
55-59	16	4824443	20	6743790	32	7372021	41	7446988	41	7071006
60-64	28	6069611	26	4404567	38	6213936	58	6813378	68	6988969
65-69	33	5371770	42	5298248	41	3889820	63	5575185	84	6148376
70-74	34	4157113	49	4371284	56	4387290	56	3277849	72	4829840
75-79	24	2432745	37	3018047	53	3254297	73	3391145	64	2511709
80-84	10	1229739	22	1467570	35	1878692	54	2112437	63	2362417
85-89	7	527277	11	560756	16	691452	23	927740	31	1123450

If one wants to explain the deaths according to age and period, one can choose two series of parameters, (a_i) and (c_j) each describing the effect of a given age segment and a given cohort. To satisfy the constraints of positivity (deaths are a positive number), one can propose a multiplicative model $a_i \, c_j$. One will choose for example to model the number of expected deaths $\lambda_{ij} = E(d_{ij})$ with a model of the form $\ln\left(\dfrac{\lambda_{ij}}{N_{ij}}\right) = a_i + c_j$ or, in an equivalent way $\lambda_{ij} = N_{ij} \exp(a_i + c_j)$. In a cell, there is finally a likelihood related to the Poisson distribution: $\left(N_{ij} \exp(a_i + c_j)\right)^{d_{ij}} \exp\left(-N_{ij} \exp(a_i + c_j)\right) / d_{ij}!$, and global likelihood is obtained by multiplying likelihoods of each cell.

The application of this type of models to the construction of prospective tables is presented in 3.2.2 below.

3.1.2. Use of a benchmark

The use of a benchmark consists in seeking a "positioning" of the custom table compared to a given benchmark table – which can be for example an INSEE table.

Many models are possible, but the most common approach consists in applying a reduction (or increase) rate to the rates of the benchmark table, which consists in seeking a coefficient α such that $q_x^{ex} = \alpha \times q_x^{ref}$. While remembering that the quotient of mortality is the discrete version of the hazard rate μ_x (with the relation $\mu_x = -\ln(1 - q_x)$ if one makes the assumption of constant hazard function between two integer ages), one notices that this model is thus a proportional random model in which one supposes the basic hazard function to be known. More precisely, if it is supposed that $\mu_x^{ex} = \alpha \times \mu_x^{ref}$ on the one hand and that $\mu_x = -\ln(1 - q_x)$ on the other hand, one obtains the following relation between the quotients of mortality:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \, \mathbb{I}_{]t;\infty[} (T_x)$$

$$1 - q_x^{ex} = \left(1 - q_x^{ref}\right)^{\alpha},$$

relation which, to the first order when the rates are small, is equivalent to $q_x^{ex} \approx \alpha \times q_x^{ref}$. It was seen that in this context an estimator of the type "ordinary least squares" could be proposed for $\theta = -\ln(\alpha)$. The estimator of α is thus obtained:

$$\hat{\alpha} = \exp\left(\gamma - \frac{1}{n} \sum_{i=1}^{n} \ln H_{ref}(x_i) \right) = e^{\gamma} \left[\prod_{i=1}^{n} H_{ref}(x_i) \right]^{1/n}$$

With γ the constant of Euler[7] and $H_{ref}(x) = -\ln S_{ref}(x) = -\ln L_x^{ref}$ the cumulated hazard function.

One can also consider as a selection criterion for the parameter α the difference between the number of deaths observed and the theoretical number of deaths associated with the "corrected" table. By noting L_x^{obs} the headcount under risk at age x in the population considered, the number of deaths predicted by the "corrected" table at age x is $\alpha \times q_x^{ref} \times L_x^{obs}$. If one constrains the total number of deaths predicted to equal the number observed, one obtains the following estimate of α:

$$\hat{\alpha} = \frac{\sum_x D_x^{obs}}{\sum_x q_x^{ref} \times L_x^{obs}}.$$

An alternative approach consists in reasoning age by age and considering statistics of the Chi-square type defined by:

$$\chi(\alpha) = \sum_{i=1}^{n} L_x^{obs} \frac{\left(q_x^{obs} - \alpha \times q_x^{ref}\right)^2}{\alpha \times q_x^{ref}}$$

and to seek the value of α which minimises this distance.

3.2. Prospective tables

The objective of prospective tables is to take into account the future evolution of mortality; the usual methods first seek to adjust past tendencies, then to extrapolate them to the future. The prospective approach consisting in integrating into the future the effect of future medical progress is not examined here.

The models used adjust raw rates calculated by methods such as that presented in 1.3 above to a parametric model, making it possible on the one hand to smooth sampling fluctuations and on the other hand to project the rates evolution in the future, by extrapolation.

We have raw rates indexed by age x and calendar year t, which typically take the following form:

---

[7] The value of which is approximately 0,577215665.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$
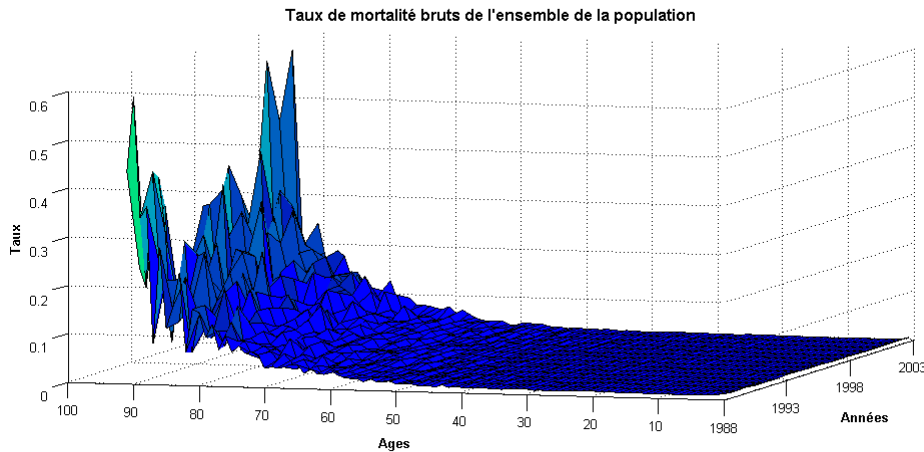
ressources-actuarielles.net

Fig. 14 :   *Raw death rates per year*

The passage of raw quotients of mortality to instantaneous rates of mortality, which is the variable modelled in some approaches, is carried out via an assumption on the distribution of deaths within a year (see 1.4 above); if one makes the assumption of constant instantaneous rates in each square of the Lexis diagram, one obtains the following estimator:

$$\mu_{xt}^* = -\mathbf{ln}\left(1 - \hat{q}_{xt}\right)$$

3.2.1. The Lee-Carter model

It is a method of extrapolation of past tendencies initially used on American data, which quickly became a standard (see the original article Lee and Carter [1992]). The modeling used for the instantaneous rate of mortality is the following one:

$$\mathbf{ln}\,\mu_{xt} = \alpha_x + \beta_x k_t,$$

or, in an equivalent way

$$\mathbf{ln}\,\mu_{xt}^* = \alpha_x + \beta_x k_t + \varepsilon_{xt}$$

The idea of the model is thus to adjust to the series (indexed twice by x and t) of the logarithms of instantaneous rates of death a parametric structure (deterministic) to which is added a random phenomenon; the criterion of optimisation will consist in maximising the variance explained by the model, which is equibalent to minimising the variance of errors.

The parameter α_x is interpreted as the average value of $\mathbf{ln}\left(\mu_{xt}\right)$ with time. It is verified that $\dfrac{d\,\mathbf{ln}\left(\mu_{xt}\right)}{dt} = \beta_x \dfrac{dk_t}{dt}$ and it is deduced that the coefficient β_x translates the sensitivity of instantaneous mortality to age x compared to the general evolution k_t, in the sens that $\dfrac{d\,\mathbf{ln}\left(\mu_{xt}\right)}{dk_t} = \beta_x$. In particular, the Lee-Carter model assumes that this sensitivity is constant with time. This constraint of the model can appear relatively strong:

- For any age x the quotients of relative variations of mortality rates at different dates do not depend on age x. If the relative variation of mortality rate at age 50 in

19

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} I_{]t;\infty[}\left(T_x\right)$$

2000 were 80 % of what it was in 1990 this coefficient of 80 % is retained for all ages;

- For the same date *t* the quotients of relative variations of mortality rates at different ages do not depend on the date *t*. If in 2000 the relative variation of mortality rate at age 20 is 50 % of the relative variation at age 50 this coefficient of 50 % will apply to any future or past date.

Lastly, one can notice that the form of the model implies the homoscedasticity of mortality rates, which is obviously false in practice. This disadvantage will be examined more in details in 3.2.2 below.

In order to make the model identifiable, it is advisable to add constraints on the parameters; indeed, for all non zero constant *c* the model is invariant through the following transformations:

$$\left(\alpha_x, \beta_x, k_t\right) \rightarrow \left(\alpha_x, \frac{\beta_x}{c}, c \times k_t\right)$$

$$\left(\alpha_x, \beta_x, k_t\right) \rightarrow \left(\alpha_x - c\beta_x, \beta_x, k_t + c\right)$$

It is thus advisable to impose two constraints on the parameters. The following constraints are retained:

$$\sum_{x=x_m}^{x_M} \beta_x = 1 \text{ and } \sum_{t=t_m}^{t_M} k_t = 0 .$$

One then obtains the parameters by a criterion of least squares (nonlinear):

$$\left(\hat{\alpha}_x, \hat{\beta}_x, k_t\right) = \mathbf{arg\,min} \sum_{x,t} \left(\mathbf{ln}\,\mu_{xt}^* - \alpha_x - \beta_x k_t\right)^2$$

Next, we solve this program of optimisation, under the constraints of identifiability. The number of parameters to be estimated is high, it is equal to $2 \times \left(x_M - x_m + 1\right) + t_M - t_m - 1$.

3.2.1.1. Parameters estimation

With respect to $\left(\alpha_x\right)$, since:

$$\frac{\partial}{\partial \alpha_x} \sum_{y,t} \left(\mathbf{ln}\,\mu_{yt}^* - \alpha_y - \beta_y k_t\right)^2 = 2\left(t_M - t_m + 1\right)\alpha_x - 2\sum_{t=t_m}^{t_M} \left(\mathbf{ln}\,\mu_{xt}^* - \beta_x k_t\right),$$

one finds by taking into account the constraint $\sum_{t=t_m}^{t_M} k_t = 0$ that:

$$\hat{\alpha}_x = \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \mathbf{ln}\,\mu_{xt}^* .$$

In other words, $\hat{\alpha}_x$ is the temporal average, at age *x*, of the instantaneous rates of death (on the logarithmic scale). One then considers the matrix $Z = \left(z_{xt}\right)$ of the rates centered with respect to the temporal dimension:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[} (T_x)$$

$$z_{xt} = \ln \mu_{xt}^* - \hat{\alpha}_x .$$

Z is a matrix of dimension $(x_M - x_m +1, t_M - t_m +1)$. The form of the model amounts to seeking an approximation of Z as a product of 2 vectors of the form $Z \approx \hat{\beta}\hat{k}'$, so that the decomposition is optimal in the sense of the least squares criterion, *i.e.* explains the greatest possible share of the total variance. The resolution of this problem requires the decomposition in eigenvalues of matrix Z, which one puts in the form:

$$Z = \sum_{i \geq 1} \sqrt{\lambda_i} v_i u_i'$$

with $\lambda_1 \geq \lambda_2 .. \geq 0$ the eigenvalues of $Z'Z$, u_i the standardised[8] eigenvector of $Z'Z$ associated with λ_i, and v_i the eigenvector associated to the same eigenvalue for ZZ'. This is justified by the fact that, since $Z'Zu_i = \lambda_i u_i$, one has $ZZ'(Zu_i) = \lambda_i Zu_i$, and thus the two transposed matrices have the same eigenvalues with identical orders of multiplicity. Moreover, if u_i is an eigenvector of $Z'Z$ then Zu_i is an eigenvector of ZZ' associated with the same eigenvalue. If one poses $v_i = \frac{1}{\sqrt{\lambda_i}} Zu_i$, one sees that $Zu_i u_i' = \sqrt{\lambda_i} v_i u_i'$, which while adding and taking into account the orthonormality of the eigenvectors, leads to the decomposition of Z.

One thus proposes as an approximation $Z \approx \sqrt{\lambda_1} v_1 u_1'$, with as a measure of the quality of this approximation the share of explained inertia $\frac{\lambda_1}{\sum \lambda_i}$. One finally obtains the following estimators of β and k:

$$\hat{\beta} = \frac{1}{\sum v_{1j}} v_1 \text{ et } \hat{k} = \sqrt{\lambda_1} \sum v_{1j} u_1 .$$

The objective is to use the results of this fitting to extrapolate mortality rates for $t > t_M$; the idea is to analyse the series of $(\hat{k}_t)$, which capture information on the temporal evolution of these rates in order to fit an ARIMA type model.

At this stage one has a first estimate of the parameters of the model; however, if one is in a situation in which the headcount under risk is very large, one can think that sampling fluctuations over the total number of death per age must be very small. It is then advised to first proceed to fitting the model parameters so that the number of deaths forecasted by the model each year equals the observed number of deaths. Since the risk exposure is $L_{xt} = \frac{D_{xt}}{\mu_{xt}^*}$ with $D_{xt} = D_{xt}(t-x) + D_{xt}(t-x-1)$, this constraint is expressed by:

$$\sum_{x=x_m}^{x_M} D_{xt} = \sum_{x=x_m}^{x_M} L_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t\right),$$

---

[8] $u_i' u_i = 1$

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

the variable being $\hat{k}_t$. The function $F(k) = \sum_{x=x_m}^{x_M} L_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x k) - \sum_{x=x_m}^{x_M} D_{xt}$ is introduced, so that

the constraint above is expressed by $F(k) = 0$. the form of the function F ensures the unicity of the root if it exists. The research of the root can be done by an algorithm of the Newton-Raphson type, while posing:

$$k_{i+1} = k_i - \frac{F(k_i)}{F'(k_i)}$$

with initial value $k_0 = \hat{k}_t$ and the stop criterion $\frac{k_{i+1} - k_i}{k_i} \le \varepsilon$ while taking for example $\varepsilon = 10^{-7}$.

A new estimator $\hat{\hat{k}}_t$ is obtained. But the series $\left(\hat{\hat{k}}_t \right)$ must be corrected to respect the constraint

of identifiability $\sum_{t=t_m}^{t_M} k_t = 0$, which results in posing:

$$k_t^* = \hat{\hat{k}}_t - \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{\hat{k}}_t.$$

We then have to correct $(\hat{\alpha}_x)$ so that the equality between deaths predicted by the model and observed deaths remains valid, which leads to:

$$\alpha_x^* = \hat{\alpha}_x + \frac{\hat{\beta}_x}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{\hat{k}}_t.$$

The "typical" shape of obtained parameters is as follows, first of all for the parameters which are a function of age:
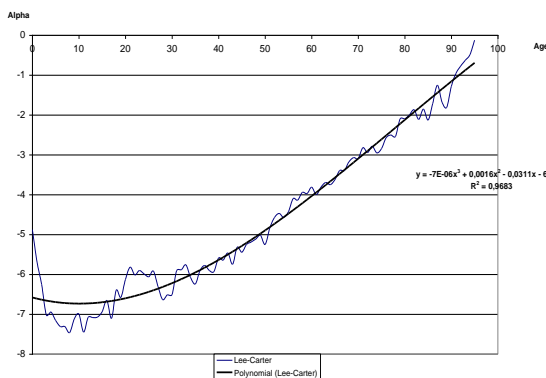
Parameters α_x Parameters β_x

Fig. 15 :   *Alpha coefficients*

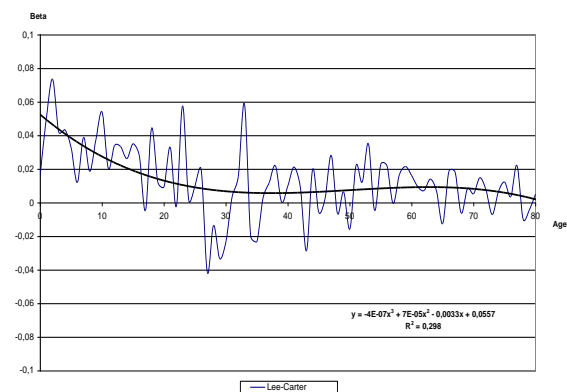Fig. 16 :   *Beta coefficients*

then for the temporal component:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{I}_{]t;\infty[}(T_x)$$
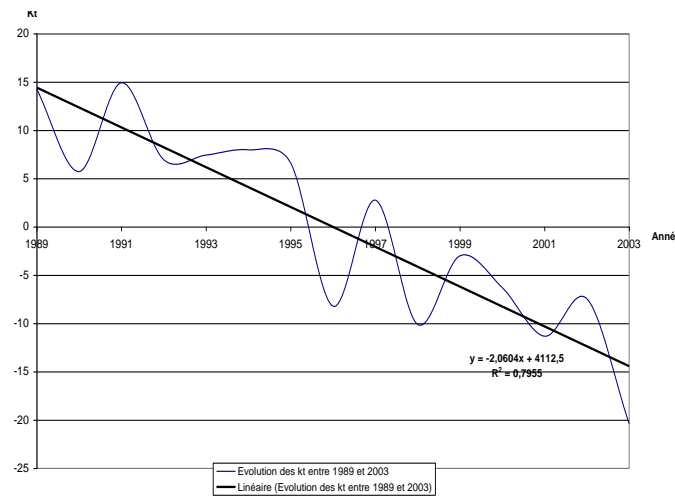
Fig. 17 :   *Evolution of the temporal tendency*

3.2.1.2. Extrapolation of the temporal component

The series (k_t^*) then has to be modelled in order to extrapolate future rates; for that, an ARIMA model is generally used[9], but any other time series model can be used. However, taking into account the pace of the graph above, the simplest modeling that one can imagine is for example a linear regression of these coefficients:

$$k_t^* = at + b + \varepsilon_t$$

with (ε_t) a Gaussian white noise.

3.2.1.3. Closing of the table

The estimate of the parameters of the Lee-Carter model requires that one has a complete rectangular matrix of rates of death (μ_{xt}^*); in practice the estimated raw amounts show great instability at high ages, because of the small available headcount. Furthermore, it can happen that the data are not available any more beyond a particular age. Several methods exist to fill the table before carrying out the fitting, or *ex-post* (referred to as "closing of the mortality table"). One can consult on the subject Denuit and Quashie [2005].

As an illustration, one presents here the method of Coale and Kisker (Coale and Kisker [1990]); the method consists in extrapolating mortality rates at high ages (up to *x=110* years for example) by using the following formula[10]:

$$\hat{\mu}_x = \hat{\mu}_{65} \times e^{g_x(x-65)}$$

g_x indicating the average rate of growth of μ_x between ages 65 and *x*. One thus calculates the coefficients g_x up to a specific age, after which they are extrapolated in order to be able to recompose the rates μ_x. Coale and Kisker indeed empirically noticed that the curves of the g_x generally show a peak in the 80 years old neighbourhood before decreasing linearly. They consequently proposed the following equation:

---

[9] While following the approach of Box and Jenkins.

[10] Index *t* is omitted in order to make notations lighter.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \, \mathbf{1}_{]t;\infty[}(T_x)$$

$$g_x = g_{80} + s \times (x - 80), \ x \geq 80.$$

Finally, one can use the following formula to extrapolate beyond age 80 the instantaneous rates of mortality:

$$\hat{\mu}_x = \hat{\mu}_{x-1} \times e^{g_{80} + s(x-80)}, \ x \geq 80.$$

The following values of parameters are used:

$$s = -\frac{\ln\left(\hat{\mu}_{79} + 31 \times g_{80}\right)}{465} \ \text{and} \ g_{80} = \frac{\ln\left(\dfrac{\hat{\mu}_{80}}{\hat{\mu}_{65}}\right)}{15}.$$

Thus, the mortality rates smoothed from raw data are directly obtained by the Lee-Carter smoothing method for ages lower than 80. For ages equal to or higher than 80, if the sample is not large enough, one resorts to the method of Coale and Kisker: which builds mortality rates at high ages based on smoothed (by Lee-Carter) rates at ages 65 and 80.

The question of the closing of the table is important in the case of the construction of a table for reserving life annuities. One will however note that this will not be so important, should high age beneficiaries be a minority in the life annuities portfolio under study.

Indeed, let us consider the simple example in which one computes the net present value of a life annuity on a single head using the mortality table TF00-02; one compares the calculation carried out with the complete table to that carried out with the table closed in a prudent way by making the rate of death a constant from age 95 onwards. So, carrying out this calculation for a head aged 75 using a discount rate of 2.5 %, should we consider that the rate of death remains stable from age 95 (and that the survivors all die brutally at age 120), the reserve would only have to be increased by 0.7 % (and approximately 2.5 % at age 85).

The difference between two closing methods in terms of reserving is truly significant only at very high ages (see for example Delwarde and DENUIT [2006]).

3.2.2. The log-Poisson model

The Lee-Carter model rests on the assumption of homoscedasticity of mortality rates, which constitutes a strong and not very realistic assumption: indeed, the variance of the rates of death grows at high ages, noticeably because of lowering survivors headcount. One can illustrate this fact in two manners; first of all, one considers the French census as at 01.01.2005, which is supposed to die according to table TV1999/2001. The variance of the rates of rough deaths that one would observe can be approximated by $\dfrac{q_x(1 - q_x)}{L_x}$, and one notes the following evolution:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$
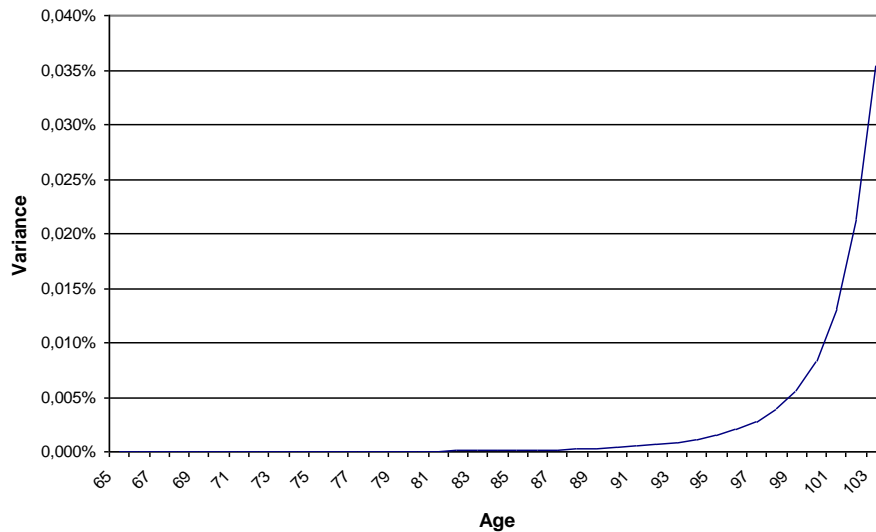
Fig. 18 :   *Variance of the rate of death with age*

One notes a very strong increase after age 85. In a more direct way, when one carries out a fit by the Lee-Carter method, one can analyse the variance of the residues, and confront the observations with the assumption of heteroscedasticity. One obtains graphs with the following shape[11]:
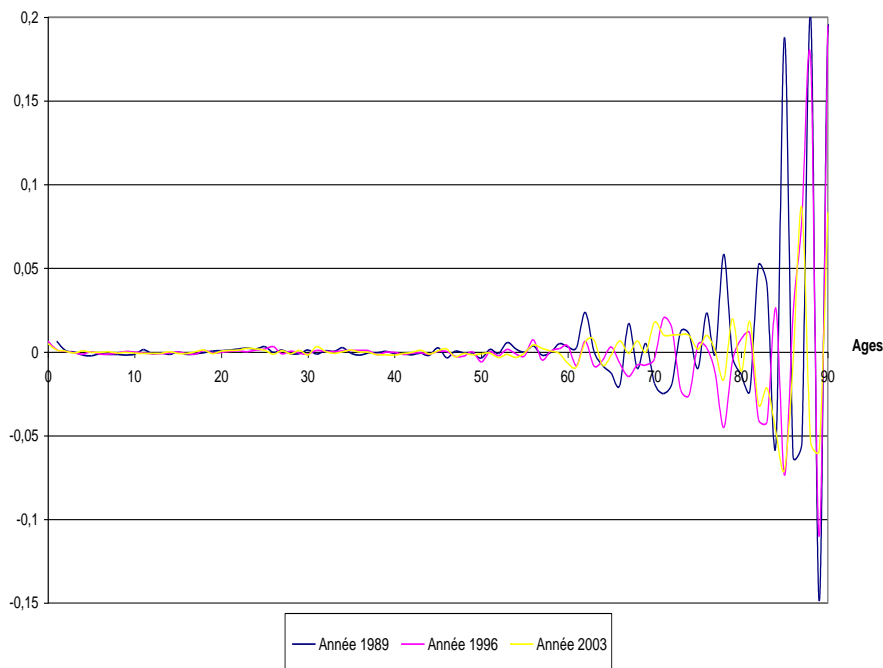
Fig. 19 :   *Residues analysis in the Lee-Carter model*

Moreover, the criterion retained to estimate the parameters in the Lee-Carter method is not of "maximum of likelihood" type.

Furthermore, it can seem natural to directly model the number of deaths rather than the instantaneous rate of mortality. If D_{xt} indicates the number of deaths at age x on year t, and

---

[11] See LELIEUR [2005].

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[} (T_x)$$

L_{xt} the risk exposure, D_{xt} is then a random variable of which the expectancy will be modelled as follows:

$$E(D_{xt}) = L_{xt} \times \mu_{xt}.$$

The log-Poisson model, proposed by Brouhns and *al.* [2002], is an adaptation of the Lee-Carter model which integrates these various elements. It will be noted that the equality above is the direct consequence of the assumption of constant μ_{xt} on each square of the Lexis diagram. Indeed, since the risk exposure is equal to:

$$L_{xt} = \int_0^1 S(x+u,t+u)\,du$$

and $S(x+u,t+u) = S(x,t)\exp\left(-\int_0^u \mu(x+v,t+v)\,dv\right)$, a constant μ_{xt} leads to:

$$S(x+u,t+u) = S(x,t)\exp(-u \times \mu(x,t))$$

then $L_{xt} = S(x,t)\int_0^1 \exp(-u \times \mu(x,t))\,du$ and thus:

$$L_{xt} = \frac{S(x,t)}{\mu(x,t)}\big(1 - \exp(-\mu(x,t))\big) = \frac{S(x,t)}{\mu(x,t)} q(x,t),$$

which establishes the result.

The idea is to model the number of deaths at age *x* on year *t* by a Poisson distribution, as done in 3.1 above, by supposing that D_{xt} follows a Poisson distribution of parameter $L_{xt}\mu_{xt}$ with $\mu_{xt} = \exp(\alpha_x + \beta_x k_t)$. The expression of the rate of instantaneous death is identical to that proposed in the Lee-Carter model, with the same interpretation of the various parameters. In particular, the model will be identifiable only with constraints on the parameters, and one can retain the same ones as those used by Lee and Carter. Lastly, one can note that to pass from the model of Lee-Carter to this Poissonnian model is equivalent to passing from a linear model to a generalised linear model with logarithm as the function of link[12].

Since one has $P(D_{xt} = d) = \frac{(L_{xt}\mu_{xt})^d}{d!}\exp(-L_{xt}\mu_{xt})$ with $\mu_{xt} = \exp(\alpha_x + \beta_x k_t)$, the log-likelihood[13] of the model is written (plus or minus an additive constant):

$$\mathbf{ln}\,L(\alpha,\beta,k) = \sum_{x,t} \big\{ D_{xt}(\alpha_x + \beta_x k_t) - L_{xt}\mathbf{exp}(\alpha_x + \beta_x k_t) \big\}$$

One thus has a simple expression of log-likelihood; the equations of likelihood do not have an analytical solution because of the presence of the nonlinear term $\beta_x k_t$ and must be solved

---

[12] See RENSHAW [1991].

[13] It is a likelihood only if we use real headcounts under risk; if headcounts are normalised starting from an initial headcount of L_0, we end up with pseudo-likelihood.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{[t;\infty[}(T_x)$$

numerically; one can for example use an algorithm of Newton-Raphson and use the process suggested in 3.2.1.1 above with objective function F to be cancelled the vector of scores $\left(\frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial k} \right)'$; which leads to the following relations of recurrence:

$$\hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i - \frac{\sum_t \left(D_{xt} - L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i \right) \right)}{-\sum_t \left(L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i \right) \right)}$$

$$\hat{k}_t^{i+1} = \hat{k}_t^i - \frac{\sum_x \left(D_{xt} - L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^i \right) \right)\hat{\beta}_x^i}{-\sum_x \left(L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^i \right) \right)\left(\hat{\beta}_x^i \right)^2}$$

$$\hat{\beta}_x^{i+1} = \hat{\beta}_x^i - \frac{\sum_t \left(D_{xt} - L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^{i+1} \right) \right)\hat{k}_t^i}{-\sum_t \left(L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^{i+1} \right) \right)\left(\hat{k}_t^{i+1} \right)^2}$$

The initial values are free, one will simply choose values for $\hat{\beta}_x^0 \neq 0$ to avoid divisions by 0. In order to have the constraints of identifiability verified, it is advisable to fit the parameters thus estimated, by posing:

$$k_t^* = \left(\hat{k}_t - \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t \right) \sum_x \hat{\beta}_x$$

$$\beta_x^* = \frac{\hat{\beta}_x}{\sum_x \hat{\beta}_x}$$

$$\alpha_x^* = \hat{\alpha}_x + \frac{\hat{\beta}_x}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t$$

The estimated values of the parameters are rather close to those obtained by the Lee-Carter model, as can be seen on the graphics shown in appendix[14]. The extrapolation of the temporal component is then carried out in the same manner as in the Lee-Carter model.

3.2.2.1. Obtaining confidence intervals[15]

In practice, such custom tables will generally be used to calculate residual life expectancies, to obtain lifetimes of annuity beneficiaries (current and future); more precisely they could be used to calculate net present values of life annuities, of the form:

$$a_{xt} = \sum_{i \geq 0} v^{i+1} \prod_{j=0}^{i} \mathbf{exp}\left(-\mu_{x+i,t+i} \right)$$

---

[14] Graphs taken from BROUHNS et al. [2002].
[15] Only the principe of the method is described here ; for the detailed approach, see HADERER [2003].

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

With $v = \frac{1}{1+r}$ the discount factor. Beyond the point estimation of a_{xt} which rises from modeling μ_{xt}, one wishes to measure the associated precision, and thus to obtain confidence intervals. Two sources of risk are combined in this case, on the one hand the fluctuations of sampling of the Poissonnian model of regression, and on the other hand uncertainty related to the prediction of k_t for $t > t_M$.

The estimate of the parameters of the model (for $t \le t_M$) by the method of the maximum of likelihood makes it possible to conclude that the vector $\left(\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t\right)$ is asymptotically distributed according to a normal distribution. One can then build confidence intervals for a_{xt} by the following simulation method:

- ✓ Generation of a realisation $\left(\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t\right)$ based on the normal distribution;

- ✓ starting from the above realisation, estimation of the projection parameters of the ARIMA model associated with k_t;

- ✓ simulation of a trajectory of k_t for $t > t_M$;

- ✓ determination of a realisation of the variable of interest (for example a_{xt}).

By renewing the operation one obtains an empirical distribution of the variable of interest, and then, in particular, a confidence interval.

When the sample size is very large one can consider that sampling fluctuations become negligible, and remove the first stage of the algorithm.

This kind of application will be developed within the framework of stochastic mortality models.

3.2.3. Log-linear models

In the choice of a model suitable for structuring a set of historical data, "flexibility" of the model as well as its "fidelity" to the data are directly related to the number of introduced parameters. The choice of a very flexible model is generally done to the detriment of predictive qualities (a completely nonparametric model does not authorise any prediction).

Thus the Lee-Carter or Log-Poisson models may appear very parameterised. Furthermore, in the context of insurance portfolio data, the size of which is noticeably smaller than that at country level, the high number of parameters can lead to irregularities which are consequences of sampling fluctuations. This phenomenon is highlighted in Lelieur [2005].

In this context it can be useful to turn to less parameterised alternative models bringing into play analytical expressions related to ages or years (or both).

Besides, the influences of age x and year t on mortality rates $q_x(t)$ are expressed via the introduction of the *logit*:

$$\mathbf{lg}(q_{xt}) = \mathbf{lg}_x(t) = \mathbf{ln}\left(q_{xt} / (1 - q_{xt})\right).$$

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \, \mathbb{I}_{]t;\infty[}\left(T_x\right)$$

ressources-actuarielles.net

Logit for low mortality rates is not very different from the variable $\mathbf{ln}\left(\mu_{xt}\right)$ of the Lee-Carter model but it can be materially different at high ages. It has the advantage of varying in $\left]-\infty,+\infty\right[$, which simplifies the implementation of regression models. The typical shape of a logit is the following one (obtained with TV 1999/2001):
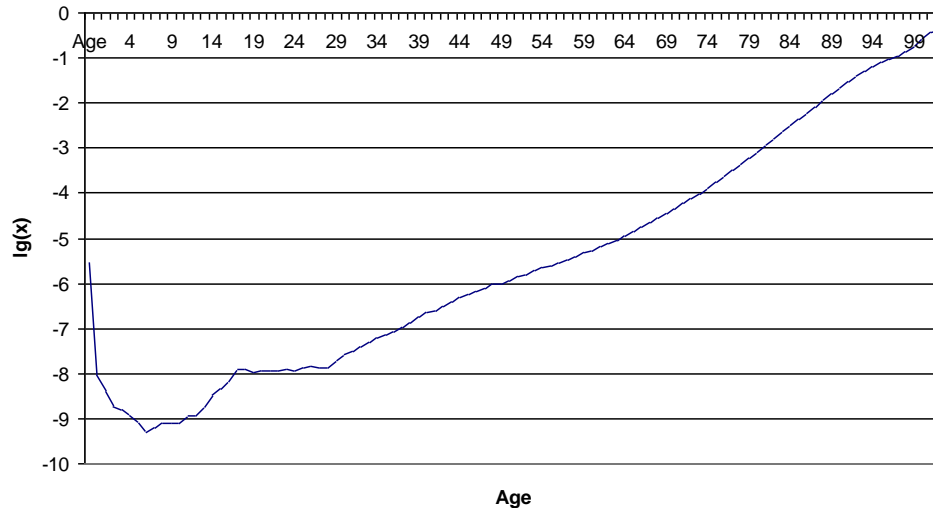

Fig. 20 :   *Logit of rate of death as a function of age (table TV 99-01)*

One is thus led to introduce log-linear models. The basic model of this family imposes a linear trend with respect to time:

$$\mathbf{lg}_x\left(t\right) = \alpha_x + \beta_x \times t + \varepsilon_{xt}.$$

The residues are supposed iid (and thus homoscedastic), which makes it possible to use the standard results of the ordinary linear model (with the calendar year t as explanatory variable, under x fixed). This parameterisation is close to that of the Lee-Carter model in which one would have supposed $k_t = t$ and replaced $\mathbf{ln}\left(\mu_{xt}\right)$ by $\mathbf{lg}_x\left(t\right)$. This model is used in particular for the construction of TPG 1993 tables. One has explicit expressions for the parameters. Indeed, by remembering that in the linear model $y_i = ax_i + b + \varepsilon_i$ one has:

$$\hat{a} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x} \times \bar{y}}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2} = \frac{\mathbf{cov}(x,y)}{\mathbf{var}(x)} \text{ and } \hat{b} = \bar{y} - \hat{a}\bar{x},$$

with $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ one obtains the expressions of the coefficients α_x and β_x.

One empirically notes a very strong correlation between the series $\left(\alpha_x\right)$ and $\left(\beta_x\right)$, which results in proposing an alternative to the model in which these two coefficients are bound; that leads to the following model:

$$\mathbf{lg}_x\left(t\right) = a\beta_x + b + \beta_x \times t + \varepsilon_{xt} = \left(a + t\right)\beta_x + b + \varepsilon_{xt}$$

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

ressources-actuarielles.net

The number of parameters to be estimated significantly decreases to $2 + x_M - x_m + 1$ instead of $2 \times (x_M - x_m + 1)$ in the preceding model and $2 \times (x_M - x_m + 1) + t_M - t_m - 1$ with the Lee-Carter model. However, the problem of least squares becomes nonlinear, which complicates the estimate of the parameters[16]. In practice one must resort to digital methods whereas in the first version of the model one has a direct explicit expression of the parameters.

The linear drift can appear unrealistic on the long run, and one notes for example on American data a deceleration of the tendency. One can then seek models allowing to introduce for very long run forecasts some exogenic information translating a foreseeable deceleration of the drift. This can be carried out with the following models:

$$\mathbf{lg}_x(t) = \alpha_x + \beta_x \times t + \gamma_x \times t^\delta + \varepsilon_{xt}.$$

In these models the estimates also reveal a very strong correlation between the estimates of parameters (α_x), (β_x) and (γ_x), which encourages to propose two new models in posing $\alpha_x = a\gamma_x + b$ and $\beta_x = c\gamma_x + d$ and leads to the specification:

$$\mathbf{lg}_x(t) = b + d\,t + \gamma_x\left(t^\delta + c\,t + a\right) + \varepsilon_{xt}.$$

The resolution of the associated criterion of least squares does not call for any particular observation.

3.2.4. The shifted logistic model

One considers here the model suggested by Bongaart [2004] and defined by:

$$\mu_{xt} = \frac{\alpha(t)\exp(\beta x)}{1 + \alpha(t)\exp(\beta x)} + \gamma(t).$$

Since the reverse of the logistic function $\mathbf{lg}(x) = \mathbf{ln}\left(\dfrac{x}{1-x}\right)$ is $y \to \dfrac{e^y}{1+e^y}$, it can be deduced by writing:

$$\mu_{xt} - \gamma(t) = \frac{\exp(\beta x + \delta(t))}{1 + \exp(\beta x + \delta(t))}$$

with $\delta(t) = \mathbf{ln}(\alpha(t))$ that this model can be also written:

$$\mathbf{lg}(\mu_{xt} - \gamma(t)) = \beta x + \mathbf{ln}(\alpha(t)).$$

This model is in fact a generalisation of the Makeham model $\mu_x = \alpha\exp(\beta x) + \gamma$ (Makeham [1860]) suggested by Thatcher [1999] by posing $\mu_x = \dfrac{\alpha\exp(\beta x)}{1 + \alpha\exp(\beta x)} + \gamma$ which can be adapted to the case of variable rates of deaths in the course of time. This adjustment of the

---

[16] Which has to be globally carried out, and not age by age anymore.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

Makeham model is justified originally by the will to correct the over-estimating of conditional rates of deaths at high ages observed in practice.

The fact that β is independent from time is the consequence of the fact that it is noted empirically that this parameter depends little on time.

The estimate of the parameters can be carried out by a method of nonlinear least squares by minimising $\varepsilon = \sum_{x,t} n_{xt} \frac{(\hat{q}_{xt} - q_{xt})^2}{q_{xt}}$ with n_{xt} the risk exposure for the age and year considered[17].

Once the model is fitted on past values, the extrapolation of future mortality is carried out by an extrapolation, *via* techniques of time series, of coefficients $\alpha(t)$ and $\gamma(t)$. This parameterisation has the advantage to be less constrained in the temporal dimension than Lee-Carter or log-Poisson, extrapolation resting on 2 parameters and not just one.

The calculation of q_x as a function of $\mu_x = \frac{\alpha \exp(\beta x)}{1 + \alpha \exp(\beta x)} + \gamma$ is carried out *via*

$$q_x = 1 - \exp\left(-\int_x^{x+1} \mu(u)\,du\right) = 1 - \exp\left(-\int_x^{x+1}\left(\frac{\alpha \exp(\beta u)}{1 + \alpha \exp(\beta u)} + \gamma\right)du\right).$$

While posing $v_{\alpha,\beta}(u) = 1 + \alpha \exp(\beta u)$ one notices that $\frac{\alpha \exp(\beta u)}{1 + \alpha \exp(\beta u)}\,du = \frac{1}{\beta}\frac{dv}{v}$, which leads after some handling to:

$$q_x = 1 - e^{-\gamma}\left[\frac{v_{\alpha,\beta}(x)}{v_{\alpha,\beta}(x+1)}\right]^{1/\beta}.$$

In other words, the survival function of this model is $S(x) = e^{-\gamma x}\left(\frac{v_{\alpha,\beta}(x)}{v_{\alpha,\beta}(0)}\right)^{-\frac{1}{\beta}}$. The term "shifted model" is justified by the following observation: if one considers only the component of mortality rate associated with ageing,

$$\mu_{xt}^s = \frac{\alpha(t)\exp(\beta x)}{1 + \alpha(t)\exp(\beta x)}$$

then for a fixed year t_0 one can write for $t \geq t_0$:

$$\mu_{xt}^s = \frac{\alpha(t_0)\exp(\beta(x - \Lambda(t)))}{1 + \alpha(t_0)\exp(\beta(x - \Lambda(t)))}$$

with $\Lambda(t) = -\frac{1}{\beta}\ln\left(\frac{\alpha(t)}{\alpha(t_0)}\right)$.

---

[17] In practice this criterion is close to a discretised maximum likelihood (see lecture notes on parametric models).

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbb{1}_{]t;\infty[}(T_x)$$

3.2.5. Use of time series

The Lee-Carter model, after having adjusted on the historical data the parameters, α, β and k, proposes to regard the continuation of k_t as a time series to obtain the prospective values of the rates. One is thus led to pose:

$$k_t^* = at + b + \varepsilon_t$$

This approach can be transposed within the framework of the logistic models, with an aim of reducing the number of parameters. One then seeks to parameterise the function $x \rightarrow \mathbf{lg}_x(t)$ to take into account the influence of the year t in a nonparametric way, then to model the introduced time series. A model of the following form is thus considered:

$$\mathbf{lg}_x(t) = f(x,\theta_t) + \varepsilon_{xt}$$

where the function $f(x,\theta_t)$ is chosen, for the sake of implementation simplicity, linear with respect to the (vectorial) parameter θ_t. In a second phase the series (θ_t) is modelled.

The form retained for f is that of a cubic spline with nodes at the ages $(x_i, i = 1,\ldots,p)$. The form of the function f with p nodes is then the following one:

$$f\left(x,a,b,c,d,e_1,\ldots e_p\right) = a_t + b_t x + c_t x^2 + d_t x^3 + \sum_{i=1}^{p} e_{i,t}\left[\left(x - x_i\right)^+\right]^3$$

In practice, a simplified version of this model in which only the parameter a_t depends on time provides reliable results. By observing that the modeling of a_t through a linear regression similar to that carried out for k_t, one can build an entirely parametric version of the model while proposing:

$$f\left(x,a,b,c,d,e_1,\ldots e_p\right) = a + \lambda t + bx + cx^2 + dx^3 + \sum_{i=1}^{p} e_i\left[\left(x - x_i\right)^+\right]^3$$

3.2.6. Models with benchmarks

If one does not have sufficient data to structure the complete table correctly, one can imagine to use the structure of an existing benchmark table and simply to position the mortality of the group considered compared to this reference.

Two approaches are possible to achieve this goal, they are presented briefly hereafter.

3.2.6.1. Logistic regression

When one wishes to position a table compared to another, it can appear natural to carry out the regression of the logits of raw rates on the logits of benchmark table, which leads to the following model, suggested initially in Brass [1971]:

$$\mathbf{ln}\left(q_{xt} / \left(1 - q_{xt}\right)\right) = a \times \mathbf{ln}\left(q_{xt}^{ref} / \left(1 - q_{xt}^{ref}\right)\right) + b + \varepsilon_{xt},$$

or:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

$$\mathbf{lg}_x(t) = a \times \mathbf{lg}_x^{ref}(t) + b + \varepsilon_{xt}.$$

The implementation of this approach if one retains a criterion of the type "least squares" is very simple, since we have a linear regression within the framework of an ordinary linear model. One thus has an explicit expression of the parameters *a* and *b* (see 3.2.3 above).

It allows an easy extrapolation of the logits of the custom rates in the age segments for which the custom data would be insufficient.

One can adapt the criterion of optimisation used to take into account the context of use of the tables while rather retaining:

$$(\hat{a}, \hat{b}) = \mathbf{arg\,min}\left[e_{60}^{liss\acute{e}}(a,b) - e_{60}^{nonliss\acute{e}} \right],$$

under the following constraint:

$$e_{60}^{liss\acute{e}}(a,b) - e_{60}^{nonliss\acute{e}} > 0.$$

Where $e_{60}^{liss\acute{e}}(a,b)$ indicates the residual life expectancy at age 60, function of the parameters *a* and *b*, calculated based on the regression on the logits and $e_{60}^{nonliss\acute{e}}$ – indicating the residual life expectancy at age 60 calculated based on raw data. One then loses the explicit character of the expression of the parameters. The detail of the approach is presented in Lelieur [2005].

One can also retain the alternative $\mathbf{lg}_x(t) = a_x\,\mathbf{lg}_x^{ref}(t) + b_x + \varepsilon_{xt}$ with coefficients depending on age (or year). It is a model of this type which was used to build tables TGH and TGF 05 (*cf* Planchet [2006]).

Lastly, one can observe that when $\mathbf{ln}\left(q_{xt}^{ref} / \left(1 - q_{xt}^{ref}\right)\right) = \alpha_x + \beta_x \times k_t$, i.e. if the underlying mortality structure is described by a model of the Lee-Carter type, then:

$$\mathbf{ln}\left(q_{xt} / (1 - q_{xt})\right) = a \times (\alpha_x + \beta_x \times k_t) + b + \varepsilon_{xt} = (a \times \alpha_x + b) + a \times \beta_x \times k_t + \varepsilon_{xt}$$

and thus the adjusted model is also of Lee-Carter type with the same temporal tendency. Only the coefficient of sensitivity β_x is transformed into $a \times \beta_x$. This is a way to carry out a "level positioning" of custom mortality, the tendency of the reference being re-used directly.

3.2.6.2. Positioning with respect to a benchmark

One can also seek, in a set of exogenic prospective tables available, the benchmark table period $[t, t+h]$ the "closer" to the period $\left[t^{ex}, t^{ex} + h\right]$ resulting from custom data. This leads to using shifted benchmark tables as custom tables.

The concept of "closest" supposes the use of some sort of distance between the two tables. Various approaches are possible on this level: Chi-square on $q_x(t)$, distance deduced from the residual expectancies or their integrals (which represents the net present value of a portfolio of life annuities in which all ages are equi-represented and with zero discount rate) – the advantage of this indicator is the "smoothing" of fluctuations.

These models will not be developed here.

4. Criteria of validation of the model

The criteria of validation of the model provide decision-making supporting information within the framework of the selection of the most relevant model. The relevance is appreciated with respect to the context of use of the proposed tables – often for prospective tables: the liability valuation of life annuities.

This results in particular in paying a special attention to the representation of the residual life expectancies.

4.1. Fidelity to data

The first of the requirements which a model must meet is to be faithful to the data which were used to calibrate it. This fidelity can be examined *a priori* in two manners:

- ✓ Through mortality rates $q_x(t)$, $(x,t) \in [x_0, x_1] \times [a_0, a_1]$;

- ✓ Through the residual life expectancy in the range $[x_0, x_1]$, defined by
  $$e_t(x, x_1) = E\left[\mathbf{min}(X - x, x_1 - x) | X > x\right]:$$

The second criterion is justified by the fact that the use of prospective tables is mainly directed towards life annuity liability calculations. The residual life expectancy represents the the liability or reserve under a zero discount rate. The audit of conditional expectancies is thus impossible to avoid.

The least parameterised models are in theory and in general most faithful. Nevertheless, this statistical logic is not always respected, calibration being done on $\mathbf{lg}_x(t)$ or $\mathbf{ln}(\mu_{xt})$ and not on the elements selected to appreciate the fidelity of the model (mortality rate $q_x(t)$ and residual life expectancy $e_t(x, x_1)$). One can however, and before implementation, note the following:

- ✓ The Lee-Carter model can result in notably underestimating mortality rates at high ages (from 85-90 years). Indeed, the algorithm of reference built on a maximum likelihood approach favours the early ages ("younger") and in addition the relation $\mu_{xt} = -\mathbf{ln}(1 - q_{xt})$ is based on the assumption of constant instantaneous rate of death between two integer ages – a debatable assumption at high ages;

- ✓ One can expect that the least parameterised models fit best the "residual irregularities" of raw data which constitutes an all the more important handicap as data size is small;

- ✓ As far as residual life expectancies are concerned, the irregularities of raw and fitted tables are classically crushed and only potential and systematic drifts of the models on the q_x stand out.

One can note at this stage that to avoid having to use the "formula of passage" $\mu_{xt} = -\mathbf{ln}(1 - q_{xt})$ in the Lee-Carter model, it is possible to model directly $\mathbf{lg}_x(t)$ rather than $\mathbf{ln}(\mu_{xt})$ by writing:

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{[t;\infty[}(T_x)$$

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = \alpha_x + \beta_x k_t + \varepsilon_{xt}.$$

4.2. Comparison of modelled values and observations

The comparison sees the number of exits observed and the number of exits estimated by the model on an interval $[x, x+1[$ or a paving stone $[x, x+1[\times[t, t+1[$. It is a central tool of the validation of a custom distribution and appreciation of its level of prudence.

The calculation of the exits observed does not raise a difficulty and simply implies to define a convention of round-off. The whole part of the moment of exit is generally used. The number of exits estimated by the model is simply the expectation of the random variable "many exits"; one has, while being limited to the interval1 $[x, x+1[$:

$$P(D_i = 1) = \int_e^{x+1} \left(1 - \frac{S_X(c)}{S_X(e)}\right) F_C(dc)$$

By assuming a constant hazard function on this interval, $1 - \frac{S_X(c)}{S_X(e)} = 1 - e^{-\mu_x \times (c-e)}$ to avoid the explicit intervention of the distribution of censoring $F_C(dc)$ in the calculation, one supposes μ_x rather small so that $1 - e^{-\mu_x \times (c-e)} \approx \mu_x \times (c-e)$, then

$$P(D_i = 1) = \mu_x \times \int_e^{x+1} (c-e) F_C(dc).$$

The number of exits expected on the interval $[x, x+1[$ is thus written, by supposing constant hazard function on this interval and by using an approximation of the survival function, valid when the rate of hazard μ_x is small, $E(D) = \sum_{i=1}^{n} P(D_i = 1) = \mu_x \times E(x)$; using the conditional probability of adjusted exit, one thus has:

$$E(D) = -\ln(1 - q(x)) \times E(x).$$

One can prevent making the assumption that μ_x is small and thus only use the assumption of constant hazard function by using:

$$E(D) \approx \sum_{i=1}^{n} \left(1 - e^{-\mu_x \times (T_i - E_i)}\right) = \sum_{i=1}^{n} \left(1 - (1 - q(x))^{(T_i - E_i)}\right).$$

When the conditional probability of exit is small, this expression becomes $E(D) = q(x) \times E(x)$, formula often used by the experts.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \, \mathbf{1}_{]t;\infty[}\left(T_x\right)$$

ressources-actuarielles.net

4.3. The stability of estimates

The choice of the ages and years segments from which one must generate the forecasts is important since the estimates of the parameters depend materially on this choice. Indeed, it is possible that these differences, if they exist, generate different forecasts.

With regard to age estimates, the choice of the age segment should not bear notable incidence. Thus if one retains for example the age segment [50 to 70] and the age segment [60 to 80], the common ages (from 60 to 70) must have close estimates. That is not the case in the Lee-Carter and log-Poisson models:
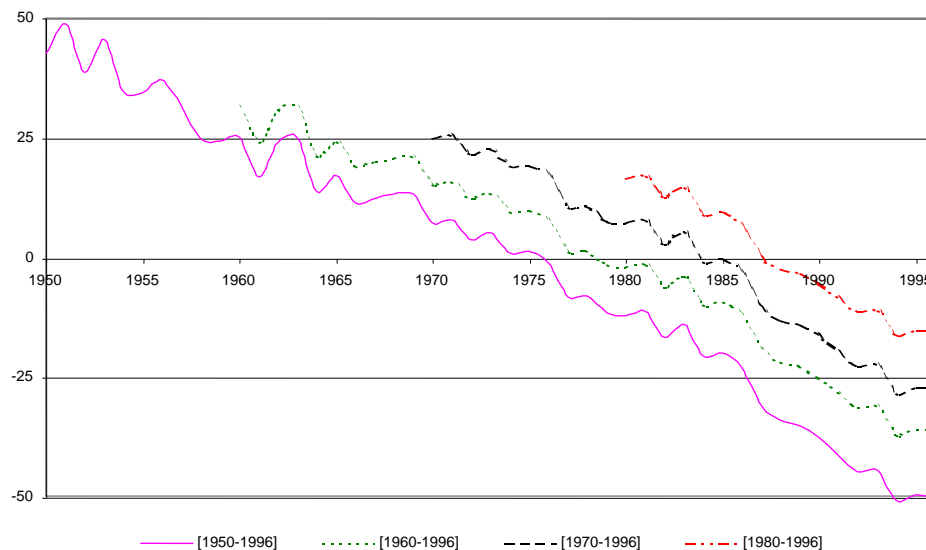

Fig. 21 :   *Estimated trends in Lee-Carter and impact of observation period*

For the models in which estimates are made age by age, their very nature insures equality of estimates when the age segments vary. For other models which take jointly into account all ages within an age segment, it can be expected that parameters estimates will depend (more or less strongly) on the chosen age segment. However, this instability should not disturb the estimation of logits, rates of death and residual life expectancies, provided the models present high levels of fidelity.

Note that for the Lee-Carter model (the less parameterised in age) the identification constraint $\sum \beta_x = 1$ has for mechanical consequence estimations differences – limited however to translations.

4.4. Prospective capacity

It can be noted that, as a general rule, the more parametric a model is, the better its capacity to be used for prospective purposes. This remark leads to favouring parametric approaches.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} I_{]t;\infty[}(T_x)$$

5. References

BRASS W. [1971] « *On the Scale of Mortality* », Biological aspects of demography, W. Brass (ed.), Londres, Taylor and Francis, 69-110.

BONGAARTS J. [2004] « *Long-Range Trends in Adult Mortality : Models and Projection Methods* », Population Council, n°194.

BROUHNS N., DENUIT M. [2002], « *Risque de longévité et rentes viagères : II. Tables de mortalité prospectives pour la population belge* », BAB.

BROUHNS N., DENUIT M. [2002], « *Risque de longévité et rentes viagères : I. Evolution de la mortalité en Belgique de 1880 à nos jours* », BAB.

BROUHNS N., DENUIT M., VERMUNT J.K. [2002], « *A Poisson log-bilinear regression approach to the construction of projected lifetables* », IME.

COALE A., KISKER E.E. [1990]. « *Defects in data on old age mortality in the united states : New procedures for calculating approximatively accurate mortality schedules and life tables at the highest ages* ». Asian and Pacific Population Forum 4,1-31.

CZADO C., DELWARDE A., DENUIT M. [2004], « *BAYESIAN POISSON LOG-BILINEAR MORTALITY PROJECTIONS* »

DEBON A., MARTINEZ-RUIZ F., MONTES F. [2004], « *Dynamic Life Tables: A Geostatistical Approach* », IME Congress

DENUIT M., QUASHIE A. [2005], « *Modèles d'extrapolation de la mortalité aux grands âges* ». Institut des Sciences Actuarielles et Institut de Statistique Université Catholique de Louvain, Louvain-la-Neuve, Belgique.

DENUIT M., MAGIS C., WALHIN F. [2004], « *La mortalité, un phénomène en pleine mutation: quelle solution pour le marché des rentes ?* », Working Paper.

FRASER H.B., WEITZ J. S. [2003], « *Explaining Mortality Rate Plateaus* », Proc. Nat. Acad. Sci.

HADERER M. [2003], « *Réassurance du risque de longévité* », ISFA, Mémoire d'actuariat.

LELIEUR V. [2005] « *Construction de tables de mortalité prospectives : le cas des petites populations* », ISFA, Mémoire d'actuariat.

PLANCHET F. [2006] « *Tables de mortalité d'expérience pour des portefeuilles de rentes – note de présentation* », Publication de l'Institut des Actuaires.

RENSHAW A.E. [1991] « *Actuarial graduation practice and generalized linear and non-linear models* », Journal of the Institute of Actuaries, vol. 118, pp. 295–312.

SITHOLE T., HABERMAN S., VERRALL R.J. [2000] « *An investigation into parametric models for mortality projections, with applications to immediate annuitants and life office pensioners* », Insurance: Mathematics and Economics, vol. 27, pp. 285–312.

THATCHER A.R. [1999] « *The Long-term Pattern of Adult Mortality and the Highest Attained Age.* », Journal of the Royal Statistical Society 162 Part 1: 5-43.

THATCHER A.R., KANNISTO V., VAUPEL J.W. [1998]. *The Force of Mortality at Ages 80 to 120*. Odense: Odense University Press.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \, \mathbf{1}_{]t;\infty[} \left(T_x \right)$$

ressources-actuarielles.net

6. Appendix