

The 31st International Symposium on Forecasting. International Institute of Forecasters, Praga, 2011.

The Lee Carter Method for Estimating and Forecasting Mortality: An Application for Argentina.

Andreozzi, Lucía, Blaconá, María Teresa y Arnesi, Nora Elba.

Cita:

Andreozzi, Lucía, Blaconá, María Teresa y Arnesi, Nora Elba (2011). *The Lee Carter Method for Estimating and Forecasting Mortality: An Application for Argentina. The 31st International Symposium on Forecasting. International Institute of Forecasters, Praga.*

Dirección estable: <https://www.aacademica.org/lucia.andreozzi/38>

ARK: <https://n2t.net/ark:/13683/preH/0mG>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

THE LEE CARTER METHOD FOR ESTIMATING AND FORECASTING MORTALITY: AN APPLICATION FOR ARGENTINA

LUCÍA ANDREOZZI, MARÍA TERESA BLACONÁ, NORA ARNESI

School of Statistics

Faculty of Economics and Statistics

National University of Rosario, Argentina

ABSTRACT

This paper describes the application of the Lee-Carter model to age-specific death rates by gender in Argentina. These rates are available for the period that goes from 1979 to 2006. The index of the level of mortality for each gender, and the shape and sensitivity coefficients for nine age groups were obtained through the Lee-Carter method. The autoregressive moving average (ARIMA) and the space-state (SSM) models are used to forecast the general index for the time period that goes from 2007 to 2011 in order to project life expectancy at birth using life tables.

KEY WORDS: Death rates, Lee-Carter model, ARIMA models, space-state models, Argentina, life expectancy.

I.- Introduction

Life expectancy considerably increased during the 20th century: The National Institute of Statistics and Censuses (INDEC) (INDEC-CELADE (1995), Estimaciones y proyecciones de población. Total del país 1950-2050 (versión revisada), INDEC, serie Análisis Demográfico N° 5, Buenos Aires)¹ points out that life expectancy at birth in Argentina increased from 48,5 years in 1914 to 73,8 years in 2001. On the other hand, mortality trends in highly-industrialized countries have decreased in recent years.

In the fields of Demography and Actuarial Science, there have been many attempts to find an appropriate model that represents mortality. Traditionally, a

¹ INDEC-CELADE (1995), Estimations and Population Projections. Total of the country 1950-2050 (reviewed version), INDEC, No. 5 Demographic Analysis Series, Buenos Aires.

parametric curve, like the ones suggested by DeMoivre (1725), Gompertz (1825), and Weibull (1939), is used to fit annual death rates.

Over the past twenty years, a great number of new approaches were developed in order to forecast mortality by using stochastic models, such as the ones presented by McNown and Rogers (1989, 1992), Bell and Monsell (1991), and Lee and Carter (1992). The Lee-Carter model became one of the most well-known models and it is applied in different countries around the world to forecast age-specific death rates.

On the other hand, Lee-Carter model and its extensions have been used by actuaries for multiple purposes. Essentially, the model assumes that the dynamic of mortality trends over the time is only ruled by a single parameter called mortality index. The mortality forecast is based on the index extrapolation obtained through the selection of an appropriate time series model.

Box-Jenkins models, also known as autoregressive moving average process (ARIMA) (Box y Jenkins, 1976), are usually used on forecasting. Lee and Carter (1996) used an alternative forecast method: the structural time series models, also known as space-state (SSM) models (Harvey, 1989), which mainly differ from their ARIMA counterparts in that they introduce time-variant parameters in their estimation equations. They also compare both methods.

In Argentina, officially, the INDEC develops mortality forecasts using the "component" method (United Nations, 1956). These forecasts are used to calculate population projections by age and gender from a base population. This method consists in individually projecting the variables that compose the population dynamic: mortality, fertility, and migration.

This article presents an alternative method to the one that the INDEC uses nowadays to project mortality. It is the Lee-Carter model. This paper describes the application of such model in Argentina. The primary objective is to apply and estimate the Lee-Carter model on age-specific death rates from Argentina so as to obtain the index of the level of mortality for both men and women. Once that index is obtained, it is possible to make forecasts using ARIMA (Box & Jenkins, 1970) and SSM (Harvey, 1989) time series models.

Section II presents demographic concepts and the Lee Carter Model. In section III the application for Argentina is described. Section IV presents the conclusions and finally section includes V a brief discussion.

II.- Method

By using age-specific death rates, Lee-Carter model (1992) estimates an index of the level of mortality called k . Such estimation is calculated for men, women, and the total population. ARIMA and SSM time series models are used to forecast each index. Both models are compared as regards their goodness of fit and predictive capacity. Once the index of the level of mortality is forecasted, it is possible to predict death rates and life expectancy.

II.1 Demographic Definitions

This section contains a summary of some demographic concepts used in this article. Firstly, there is a basic concept that is needed to explain the Lee-Carter model:

Age-Specific Death Rates

An age-specific death rate is the ratio of the number of deaths within a specified age group in a given geographic area during a certain period of time to the corresponding population at risk of the same age group, in the same geographic area during a specified time period of study. If the number of deaths is obtained through Vital Statistics, the denominator is estimated on the basis of census data. (Ortega A., 1987), "Tablas de mortalidad", CELADE, Serie E N° 1004, San José, Costa Rica.)

Life Expectancy at Birth

Age-specific life expectancy is an estimation of the average number of the remaining years that a person would be expected to live if current mortality conditions were constant. It is calculated taking into account the age-specific death rates. (Ortega A., 1987)

II.2 Lee Carter Model

The Lee and Carter model, also named LC hereinafter, is a demographic and statistical model that is used to project mortality rates. (Lee, R. D., Carter, L. 1992)

II.2.1 The Model

The model's basic premise is that there is a linear relationship among the logarithm of age-specific death rates $m_{x,t}$ and two explanatory factors: the initial age interval x , and time t . Information is distributed in age intervals, so the interval that begins with the x age will be called "x age interval. The equation that describes this fact is as follows:

$$m_{x,t} = \exp(a_x + b_x k_t + e_{x,t}), \quad t=1, \dots, n \quad x=1, \dots, \omega \quad (1)$$

or applying the logarithm:

$$f_{x,t} = \ln(m_{x,t}) = a_x + b_x k_t + e_{x,t}, \quad t=1, \dots, n \quad x=1, \dots, \omega \quad (2)$$

where

$m_{x,t}$: is the age-specific death rate for the x interval and the year t .

k_t : Is the mortality index in the year t .

a_x : Is the average age-specific mortality.

b_x : Is a deviation in mortality due to changes in the k_t index.

$e_{x,t}$: Is the random error.

ω : Is the beginning of the last age interval.

(Wang, J.Z., 2007)

II.2.2 Fitting the model

The expression (1) shows that there is an equation for each time and age interval. As a result, a system of simultaneous equations needs to be solved so as to estimate values of a_x , b_x , and k_t , which are the solutions to the system. A group of death rates with different r age groups that were analyzed in different moments creates a system of equations containing $2r+n$ unknown factors that correspond to the sum of the r values of a_x , r of b_x , and n of k_t and rxn equations. This system can be written within a matrix such as:

$$\mathbf{M} = \mathbf{A} + \mathbf{b} \cdot \mathbf{k}, \quad (3)$$

where \mathbf{M} is an rxn matrix, and each $M_{i,j}$ element corresponds to the natural logarithm for the age-specific death rate i in the year j .

\mathbf{A} is an rxn matrix. The elements that belong to the same categories are equal, for example: $a_{1j}=a_{2j}=\dots=a_{rj}$ for the same year j , whereas \mathbf{b} is an $rx1$ vector and \mathbf{k} is a $1 \times n$ vector.

At first, there is not a unique solution for this system. Thus, it is necessary to add the following two constraints so as to obtain a unique solution: $\sum_{x=1}^{\omega} b_x = 1$ and $\sum_{t=1}^n k_t = 0$. On using those constraints, a_x coefficients are a simple arithmetic average over time for logarithms of specific rates. Thus, b_x and k_t parameters are uniquely determined. Therefore, a_x coefficients are calculated with the following expression:

$$a_x = \frac{\sum_{t=1}^n \ln(m_{x,t})}{n}. \quad (4)$$

Once the matrix \mathbf{A} values are determined, the system (3) can be rewritten as:

$$\mathbf{M}^* = \mathbf{M} - \mathbf{A} = \mathbf{b} \cdot \mathbf{k} \quad (5)$$

This system provides a unique solution when these constraints are included. Now, \mathbf{b} and \mathbf{k} parameters are to be determined applying the SVD method (Singular Value Decomposition). This method is used to obtain the exact fitting of least squares (M. L. Good, 1969).

Through SVD, \mathbf{M}^* can be written as the product of two matrices. In \mathbf{M}^* , the (i, j) element is the sum of the product of row i of \mathbf{B} and row j of \mathbf{K} , i.e.:

$$m_{i,j} = \sum_{l=1}^r B_{i,l} K_{j,l}^T. \quad (6)$$

Therefore, the decomposition creates r terms that exactly fit the $m_{i,j}$ element of the \mathbf{M}^* matrix. Lee and Carter (1992) suggest using \mathbf{M}^* as the product of \mathbf{b} and \mathbf{k} vectors. On applying SVD, they only took into account the first-order approximation, i.e., \mathbf{M}^* is can be expressed as follows:

$$\mathbf{M}^* \approx \mathbf{B}_1 \mathbf{K}_1^T \quad (7)$$

Thus, $\mathbf{b}_1 = \mathbf{B}$ and $\mathbf{k}_1 = \mathbf{K}$ are determined, which means that there is a first estimation of the parameters of the model. (González Pérez and Guerrero Guzmán, 2007)

II.2.3 Corrections on the first fitting

Usually, the values obtained through the first estimation of the model do not provide an adequate fit to the observed data. Because of this, there may be deviations on the projections, as Lee and Carter (1992), and later on, Bell (1997), point out. Thus, a second step is needed in order to estimate the parameters. This

step uses the a_x and b_x values from the first step so as to obtain a new estimation of k so there should be an observed total number of deaths for the year in question. The objective is to find out that k_t values comply with the following:

$$D_t = \sum_{x=0}^{\omega} N_{x,t} \exp(a_x + b_x k_t + e_{x,t}), \quad (8)$$

where D_t indicates the total number of deaths in the year t , $N_{x,t}$ is the population that belongs to the x age interval in the year t , and ω is the beginning age of the last observed group in mortality tables. (Lee, R. D., Carter, L. 1992)

The R software (Development Core Team (2008)) is used for the model estimation.

II.3 Forecasting age-specific death rates

Once the time series of the k_t index is obtained as expressed in section (II.2.3) it is possible to forecast such index by using ARIMA and SSM models. After forecasting the k_t index, it is possible to obtain the death rates for the projected years. The forecast values of k_{n+h} are replaced in the formula:

$$\hat{m}_{x,n+h} = \hat{m}_{x,n} \exp\left\{\hat{b}_x (\hat{k}_{n+h} - \hat{k}_n)\right\}, \quad h=1,2,\dots \quad x=1,2,\dots,\omega \quad (9)$$

where n is the last year from which data are available; h is the forecast horizon, and x represents the age group.

Death rates forecasts with reference to the last death rate available are obtained through the (9) formula.

Lee and Carter (1992) suggest using an approximate prediction interval to forecast death rates. It is described in (II.4.2). Estimations of the b_x , parameters and standard errors of the k_t projections are used to calculate the interval.

$$PI: \left\{m_{x,t} \exp(2 b_x se_{k_t})\right\} ; \left\{m_{x,t} \exp(-2 b_x se_{k_t})\right\} \quad (10)$$

II.4 Life Expectancy at Birth

Through the forecasted death rates, the classic method of Chiang C. L. (1984) is applied so as to obtain life expectancy at birth.

Life expectancy at x , e_x , is defined as follows:

$$e_x = \frac{T_x}{l_x}; \quad (11)$$

where T_x is the cumulative number of years lived by the cohort population in the age interval and all subsequent age intervals. The l_x value represents the number of living people at the beginning of the x age interval from a population of l_0 newborn babies. This is usually defined as $l_0=100,000$.

III.- Empirical analysis

The following analysis describes the application of the Lee-Carter method (1992) to model age-and-gender-specific death rates in Argentina. A general mortality index is created for each gender. The indexes are forecasted using ARIMA and SSM models. STAMP software (Stamp Version 6.21, Struct Time S.A Copyright 2000-2003 Koopman S.J) is used for the estimation of SSM models.

After forecasts are obtained, it is possible to project age-specific death rates. The corresponding life tables for each gender are carried out using the projected death rates and applying the classic methodology that provides life expectancy at birth. Furthermore, projections of life expectancy at birth from official bodies are included so as to compare the results.

III.1.1 Data

Available data are composed of the population and death values both by age and gender in Argentina from 1979 to 2006. Data from 1979 to 2001 are provided by the WHO's Statistical Information System. The data for the period 2000-2006 is provided by the Dirección de Estadística e Información de Salud del Ministerio de Salud de la Nación (National Center for Health Statistics, Ministry of Health, Argentina). As expected, the information provided by both organizations for the period 2000-2001 is the same, which makes it possible to concatenate the series.

Age groups are determined by the following intervals: the first one goes from 0 to 4 years of age, and then there are ten-year intervals up to the last one of 75 or over. Through available data it is possible to create a maximum of 9 categories.

III.1.2 Fitting the model

The Lee-Carter model for the general mortality index k_t is:

$$m_{x,t} = \exp(a_x + b_x k_t + e_{x,t}), \quad t=1, \dots, 28 \quad x=0, 5, 15, \dots, 75 \quad (12)$$

The a_x and b_x parameters as well as the first estimation of the general mortality index (“raw”) are obtained. Then, the latter is re-estimated using the method suggested by Lee-Carter and taking into account the condition of the equation in (12). This process is carried out for the total series, and for both men and women.

Table 1 shows a_x and b_x estimations for men, women, and the total population. Higher values of b_x appear in the 0-4 age interval, which means that, in such interval, mortality varies substantially when the general mortality index k_t changes. As regards older ages, the value of this parameter is lower, which means that mortality slightly varies in that period of time.

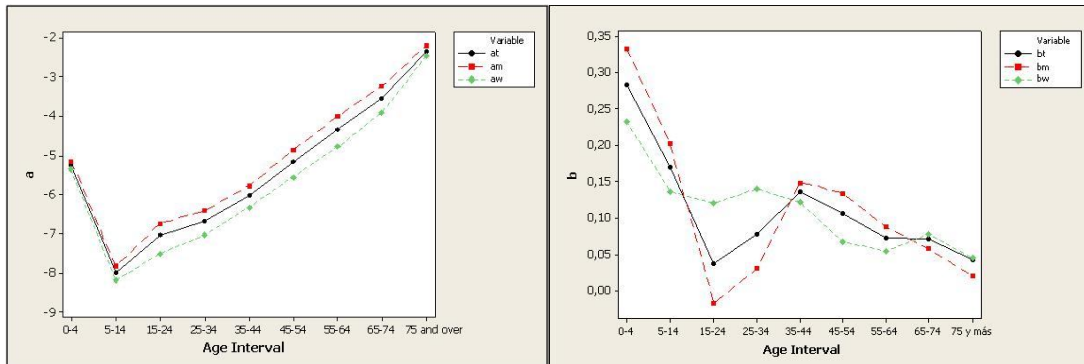
What is notorious is that the estimated value b_{15} for men, which belongs to the 15-24 age group, is negative. See Figure 1 as follows:

Table 1: Estimations of a_x and b_x , parameters for men, women, and the total population in Argentina.

Age Group	Total		Men		Women	
	a_x	b_x	a_x	b_x	a_x	b_x
0-4	-5.24167	0.28414	-5.14386	0.33404	-5.35429	0.23262
5-14	-7.99261	0.16979	-7.83313	0.20265	-8.18974	0.13717
15-24	-7.03527	0.03720	-6.72832	-0.01741	-7.50422	0.12124
25-34	-6.66634	0.07869	-6.40873	0.03053	-7.02750	0.14107
35-44	-6.01734	0.13688	-5.78227	0.14905	-6.32116	0.12217
45-54	-5.14825	0.10667	-4.85319	0.13503	-5.55004	0.06805
55-64	-4.33273	0.07287	-4.00105	0.08816	-4.77527	0.05459
65-74	-3.55048	0.07152	-3.22951	0.05801	-3.90812	0.07838
75 and over	-2.34548	0.04223	-2.19382	0.01994	-2.45140	0.04471

Figure 1 shows the estimations of the “shape” parameter (a_x) just as Lee and Carter define it. These estimations show the way in which mortality behaves through age intervals. In Argentina, it has a similar behavior to that of most of the countries: it is high at the beginning of life, then it declines until it reaches a minimum in the 5-14 interval; then it rapidly increases up to the 15-24 interval, then it shows a relatively slight increase up to the 35-45 interval and from that moment on it rises until it reaches the first age levels (Ortega A., 1987).

Figure 1. Estimation of a and b parameters for men, women and the total population in Argentina



Source: Compiled by the author based on data provided by the WHO and the Ministry of Health. Argentina.

In a following stage, first estimations and re-estimations of the general index of mortality are calculated for Argentina. The names of the indexes are: k_m for men, k_w for women and k_t for the total population.

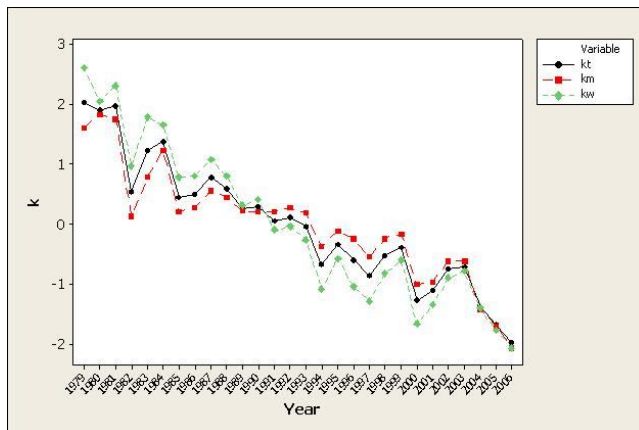
The series of these general indexes clearly tend to decrease over time (Figure 2). Moreover, there is a possible outlier value in 1982, which will be taken into account for modeling. For the first half of the period under study the same figure shows a notable increase on k_w with regard to k_m and k_t . The series behavior presents a similar shape for both genders from 1990 onwards.

ARIMA and SSM models are used for forecasting the k_t , k_m , and k_w indexes. In order to select among the different ARIMA models, the Akaike and Schwartz criteria for model selection are used. Said models include a dummy variable that represents the outlier value of 1982.

As the main objective of the model is forecasting, the $PSMAPE^2$ ("Post Sample Mean Absolute Percentual Error") is fundamental for the selection of the model. This measure is obtained by modeling 25 observations and comparing the last three observations with three out-of-sample predictions obtained through the model.

² $PSMAPE = \left(\frac{1}{H} \sum_{h=1}^H |e_{t+h}| \right) 100\%$ where $e_{t+h} = \frac{\text{real value} - \text{predictive value } h \text{ step ahead}}{\text{real value}}$, and H is the number of out-of-sample predictions.

Figure 2. Estimation of the general index of mortality k for men, women, and the total population in Argentina.



Source: Compiled by the author based on data provided by the WHO and the Ministry of Health. Argentina.

An ARIMA(0.1.2) model is selected based on goodness of fit criteria and predictive capacity. It is a first order integrated model, which means that there is a stochastic trend. Furthermore, the inclusion of a significant constant in each model shows that there is a deterministic trend in the three series.

Once the model is selected, it is possible to generate forecasts for the 5 years starting in 2006, Figure 3 shows the estimations of the corresponding 95% prediction intervals.

The basic structural model expressed in terms of trend and irregular components for the general index of mortality k_t can be written as:

$$y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad t=1, \dots, 28.$$

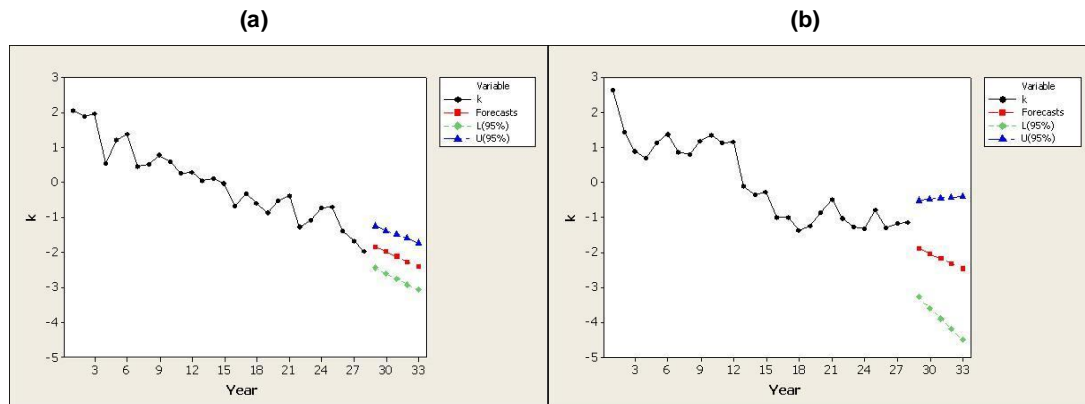
$$\text{Level:} \quad \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_{1t} \quad (13)$$

$$\text{Slope:} \quad \beta_t = \beta_{t-1} + \eta_{2t}$$

where μ_t y β_t represent the trend and ε_t represents the irregular component with distribution $N(0, \sigma_\varepsilon^2)$ the errors η_{1t} and η_{2t} are $\eta_{it} \sim IN(0, \sigma_i^2)$. $i=1,2$ independent of each other and of σ_ε^2 .

Estimated slope hyperparameters are null for the total population and for men, which means that in both cases there is a deterministic behavior for this component. On the other hand, as regards women such parameter presents a non-zero value that reflects a stochastic behavior. On all three cases the level presents a stochastic behavior as well as the irregular component.

Figure 3. Projections and 95% forecast intervals for the estimation of the kt index using the ARIMA(0,1,2) (a) and SSM (b) models for the total population in Argentina.



Source: Compiled by the author based on data provided by the WHO and the Ministry of Health. Argentina.

Once the SSM model is estimated, projections of the general mortality index are obtained for men, women, and the total population. Forecast intervals are wider than those generated by the ARIMA models and show a significant increase over time.

From the forecasts generated by the ARIMA and SSM models, death rate forecasts are calculated. Thus forecast values of k_{2006+h} are replaced in the formula:

$$\hat{m}_{x,2006+h} = \hat{m}_{x,2006} \exp\left\{\hat{p}_x \left(\hat{k}_{2006+h} - \hat{k}_{2006}\right)\right\} \quad h=1,2,\dots,5 \quad x=0,5,\dots,75$$

They are calculated with reference to the year 2006 from which information is available. The forecast horizon is represented by h , and x indicates the beginning of the age interval for which the death rate is calculated.

Tables 2 and 3 show the age-specific death rates that were calculated using the ARIMA and SSM forecasts of the kt index. Such tables include the corresponding 95% prediction intervals. There is a clear coincidence of the estimated values in both models. However, ARIMA models provide more accurate forecasts.

Table 2. Age-specific death rates and 95% confidence intervals for men, women, and the total population in Argentina, estimated using ARIMA forecasts for 2007-2011.

	2007	2008	2009	2010	2011
0-4	0.00328 0.00276-0.0039	0.00316 0.00264-0.00377	0.003035 0.00252-0.00365	0.002920 0.00242-0.00353	0.002809 0.00231-0.00341
5-14	0.00029 0.00027-0.00032	0.00029 0.00026-0.00032	0.000281 0.00025-0.00031	0.000274 0.00024-0.00031	0.000268 0.00024-0.0003
15-24	0.00082 0.0008-0.00083	0.00081 0.00079-0.00083	0.00081 0.00079-0.00083	0.0008 0.00078-0.00082	0.0008 0.00078-0.00082
25-34	0.00104 0.00099-0.0011	0.00103 0.00098-0.0011	0.00102 0.00097-0.00107	0.00101 0.00096-0.00107	0.00100 0.00095-0.00106
35-44	0.00195 0.00179-0.00211	0.00191 0.00175-0.00205	0.00187 0.00171-0.00205	0.00184 0.00168-0.00202	0.00181 0.00164-0.00198

45-54	0.00477 0.00447-0.00518	0.00470 0.0044-0.00512	0.00463 0.00432-0.00506	0.00457 0.00425-0.005	0.0045 0.00418-0.00495
55-64	0.011453 0.01096-0.01197	0.011340 0.01083-0.01187	0.011228 0.01071-0.01177	0.011117 0.01059-0.01167	0.011007 0.01047-0.01157
65-74	0.025009 0.02395-0.02612	0.024766 0.02368-0.0259	0.024526 0.02341-0.02569	0.024289 0.02316-0.02548	0.024053 0.02289-0.02526
74 and over	0.088449 0.08621-0.09074	0.087942 0.08564-0.09031	0.087438 0.08508-0.08987	0.086936 0.08396-0.08942	0.086438 0.08396-0.08898

Table 3. Age-specific death rates and 95% confidence intervals for men, women, and the total population in Argentina, estimated using SSM forecasts for 2007-2011.

	2007	2008	2009	2010	2011
0-4	0.00324 0.00217-0.00482	0.00311 0.00198-0.00488	0.00299 0.00182-0.00492	0.00287 0.00166-0.00496	0.00276 0.00153-0.00499
5-14	0.00029 0.00023-0.00037	0.00029 0.00022-0.00037	0.00028 0.00021-0.00037	0.00027 0.0002-0.00038	0.00027 0.00019-0.00038
16-24	0.00082 0.00077-0.00086	0.00081 0.00076-0.00086	0.00081 0.00076-0.00086	0.0008 0.00075-0.00086	0.0008 0.00074-0.00086
25-34	0.00104 0.00093-0.00116	0.00103 0.00091-0.00116	0.00102 0.00089-0.00117	0.00101 0.00086-0.00117	0.001 0.00084-0.00117
35-44	0.00193 0.0016-0.00234	0.0019 0.00152-0.00236	0.00186 0.00146-0.00237	0.00183 0.0014-0.00237	0.00179 0.00135-0.00238
45-54	0.00475 0.00409-0.00575	0.00468 0.00395-0.0058	0.00461 0.00382-0.00586	0.00454 0.0037-0.0059	0.00447 0.00358-0.00595
55-64	0.01142 0.0103-0.01264	0.0113 0.01068-0.01268	0.01119 0.00984-0.01271	0.01107 0.00962-0.01274	0.01096 0.00941-0.01277
65-74	0.024928 0.02255-0.02756	0.024680 0.02204-0.02764	0.024434 0.02155-0.0277	0.024191 0.02108-0.02776	0.023951 0.02063-0.0278
74 and over	0.08828 0.0832-0.09366	0.08776 0.08209-0.09383	0.08724 0.081-0.09396	0.08673 0.07997-0.09407	0.08622 0.07895-0.09415

Table 4. Estimation of life expectancy at birth for men, women, and the total population in Argentina based on death rate forecasts obtained through ARIMA and SSM models.

Year	Total		Men		Women	
	ARIMA	SSM	ARIMA	SSM	ARIMA	SSM
2007	74.77	74.77	71.03	71.01	78.05	78.05
2008	75.03	75.03	71.25	71.14	78.33	78.33
2009	75.29	75.29	71.46	71.27	78.61	78.61
2010	75.55	75.55	71.66	71.40	78.90	78.90
2011	75.82	75.82	71.87	71.52	79.18	79.18

Estimation of life expectancy at birth for the total population presents a slightly difference between both forecast methods. (See Table 4.)

Table 5 shows the official estimations (INDEC) of life expectancy at birth projected for Argentina. In general, life expectancy forecasts obtained through the Lee Carter model are higher than the estimations provided by the INDEC for 2005-2010.

Table 5. Life expectancy at birth in Argentina projected by the INDEC.

Period	Life Expectancy at Birth		
	Total	Men	Women
2005-2010	75.24	71.56	79.06
2010-2015	76.13	72.45	79.95

Once the official data are available for years 2007, 2008, and 2009, both models (ARIMA and SSM) can be evaluated as regards their predictive capacity. The PSMAPE is used as a diagnostic measure. Tables 6 and 7 show the obtained PSMAPE values for the ARIMA and SSM models. Space-state forecasts are slightly outperformed by the ARIMA forecast, except for the 5-14 age group. Note, however, that the forecasts for 15-24 and 25-34 intervals present an average PSMAPE larger than the 5%, this could be due to the fact that the death rates for this age intervals present high variability.

Table 6. PSMAPE for one, two, and three-step ahead forecasts and the average PSMAPE (ARIMA forecasts).

ARIMA	2007	2008	2009	Average
0-4	0,8	1,8	1,7	1,4
5-14	5,0	3,8	0,6	3,1
15-24	3,5	9,0	10,0	7,5
25-34	3,2	6,6	9,7	6,5
35-44	0,1	0,8	3,7	1,5
45-54	0,7	0,0	0,6	0,4
55-64	2,6	1,3	0,7	1,5
65-74	4,5	0,6	1,2	2,1
74 & over	7,2	0,3	1,1	2,9

Table 7. PSMAPE for one, two, and three-step ahead forecasts and the average PSMAPE (SSM forecasts).

Space-State	2007	2008	2009	Average
0-4	0,5	3,2	3,2	2,3
5-14	4,3	3,1	0,5	2,6
15-24	3,6	9,1	10,1	7,6
25-34	3,6	6,9	10,1	6,8
35-44	0,6	0,1	4,4	1,7
45-54	1,2	0,5	1,1	0,9
55-64	3,0	0,9	0,3	1,4
65-74	4,8	0,3	0,8	2,0
74 & over	7,4	0,1	0,9	2,8

IV.- Conclusion

This paper describes the application of the Lee-Carter model to age-specific death rates by gender in Argentina. These rates are available for the period that goes from 1979 to 2006. The index of the level of mortality for each gender, and the shape and sensitivity coefficients for nine age groups were obtained through the Lee-Carter method. The general index of mortality for 2007-2011 is forecasted using ARIMA and SSM models. It is important to highlight that the period under study in this article represents the maximum period from which data are available. The behavior of the estimated parameters is similar to the one observed in most of the countries in which this analysis was carried out. The general mortality index (k_t) is a time series that presents low variability and it clearly shows a decreasing trend. Forecasts models such as ARIMA(0,1,2) with constant and SSM that were used to project the k index present an adequate fit. However, there is a difference in quality if the corresponding prediction intervals are compared. SSM models present wider intervals than the ARIMA models. In a second stage, age-specific death rates are predicted using forecasts of the k_t index obtained from both models. According to the forecast method used, such rates show differences from the third or fourth decimal place onwards.

Even though the estimations of death rates and life expectancy are similar for both forecast models, prediction intervals for the k_t index are slightly more accurate when using the ARIMA model.

The Lee Carter method in combination with Arima and Space-State models successfully predict future death rates. However, long term forecast are necessary in the field of Demography and obtaining them depends on the available data.

IV.- Discussion

When estimating the parameters. it is important to take into consideration that the period under study includes the year 1982, the year in which the Falklands Conflict took place. This could be the event that generates the outlier value of b_{15} for men. Another possible explanation is based on the fact that studies show that in Argentina there was an increase of deaths from violent causes among men from 15

to 24 years of age. An investigation carried out by the Epidemiological Research Center of the National Academy of Medicine. Buenos Aires (Serfaty, 2003) points out that from 1991 to 2000 there was an increase in death rates due to violent causes and shows that such rates were duplicated for young men. Moreover, there is a striking difference between men and women within the intervals 15-24 and 25-34. i.e. the values are much lower for men than for women.

In Argentina, United States (Lee and Carter, 1992), and México (González Pérez and Guerrero Guzmán, 2007) the life expectancy forecasts obtained through the LC method are higher than those generated by the methods used by the statistical official bodies. The Lee-Carter method appears as a method with probabilistic support. Furthermore, it is of simple application and it generates multiple measures and results that describe current and future mortality. The application of this method in Argentina gives successful results, such as in many other countries. It is important to take into account that the period under study is considerably shorter than those of Sweden (Wang. 2007), United States (Lee and Carter, 1992), and Chile (Lee and Rofman. 1994). These three studies use time periods of more than 100 years. The period under study affects the number of forecasts that can be generated. Since the LC model is based exclusively on historical data of mortality and population, it is essential to have reliable information for considerable periods of time. That shows how important it is for a given country, region or city to obtain data in an efficient way and to keep the records updated.

REFERENCES

Alho. J.M. (1990). "*Stochastic methods in population forecasting*". International Journal of Forecasting 6. 521-530.

Alho. J.M. (1991). "Effect of aggregation on the estimation of trend in mortality". Mathematical Population Studies 3. 53-67.

Alho. Juha M. (2000). "*A statistical look at Modeen's forecast of the population of Finland in 1934.*" Yearbook of Population Research in Finland. Vol. 36. 2000. 107-20 pp.

Alho. J.M.; Spencer B.D. (1985). "Uncertain Population Forecasting" Journal of the American Statistical Association 80 (1985): 306-314.

Bell W R. (1997). "Comparing and assessing time series methods for forecasting age specific fertility and mortality rates." *Journal of Official Statistics*. 13 (3): 279–303.

Bell. W.R.. Monsell. B.. (1991). "Using principal components in time series modeling and forecasting of age-specific mortality rates." *Proceedings of the American Statistical Association. Social Statistics Section*. pp. 154-159.

Booth. H; Maindonald. J. and Smith. L. (2002). "Age-Time Interactions in Mortality Projection: Applying Lee-Carter to Australia." *Working Papers in Demography*,85.

Booth. H. & Tickle. L. (2003). "The future aged: New projections of Australia's elderly population." *Australasian Journal on Ageing*. 22(4). pp. 196-202.

Carter. L. (1996). "Forecasting U.S. Mortality: A comparison of Box-Jenkins ARIMA and Structural Time Series Models." *The Sociological Quarterly* 37. No 1:127-44.

Chia. N.C. and A.K.C. Tsui (2003). "Life annuities of compulsory savings and income adequacy of the elderly in Singapore." *Journal of Pension Economics and Finance*. March 2003

Chiang. C. L. (1984). "The Life Table and its Applications". Malabar (FL). Robert E. Krieger Publ. Co.

DeMoivre. (1725). "Annuities upon Lives" printed by W.P. London. England.

González Pérez. C.Y. Guerrero Guzmán. V. M. (2007) "Pronósticos estadísticos de mortalidad y su impacto sobre el Sistema de Pensiones de México". Published on-line by the National Commission for the Pension System web site http://www.consar.gob.mx/premio_pensiones/pdf/2007/ganadores/Primer_lugar.pdf

Gompertz, B. (1825). "On the nature of the function expressive of the law of human mortality, and on the mode of determining the value of life contingencies". *Phil. Trans. Roy. Soc.* 115:513-585.

INDEC. 2004 "Estimaciones y proyecciones de población. Total del país. 1950-2015".

INDEC-CELADE (1995). "Estimaciones y proyecciones de población. Total del país 1950-2050 (versión revisada)". INDEC. Serie Análisis Demográfico N° 5. Buenos Aires

Lee. R. D.. Carter. L. (1992). "Modeling and Forecasting the Time Series of U.S. Mortality." *Journal of the American Statistical Association* 87:659-71.

Lee R D. Miller T. (2001). "Evaluating the performance of the Lee-Carter method for forecasting mortality." *Demography*. 38 (4): 537–549.

Lee. R.; Miller. T. (2000). "*Evaluating the performance of Lee Carter mortality forecasts.*" University of California. Berkeley.

Lee. R. D. and Rofman. R. (1994). "*Modeling and Forecasting Mortality in Chile.*" *Notas* 22. No. 59:182-213.

Li N. Lee R D. Tuljapurkar S. (2004). "*Using the Lee-Carter method to forecast mortality for populations with limited data.*" *International Statistical Review*. 72. 1: 19–36.

McNown. R. F.. and Rogers. A. (1989). "*Forecasting Mortality: A Parameterized Time Series Approach.*" *Demography*. 26. 645-660.

McNown. R. F.. and Rogers. A. (1992). "*Forecasting Cause-Specific Mortality Using Time Series Methods.*" *International Journal of Forecasting* 8 (1992) 413-423.

Makeham W. (1860). "*On the Law of mortality and the construction of annuity tables*". *Journal of the Institute of Actuaries and Assurance Magazine* 8.

Ministerio de Salud. Dirección de Estadística e Información de Salud. Programa Nacional de Estadísticas de Salud. (2000). "*Modelos de Formularios e Instructivos del Sistema de Estadísticas Vitales*". Buenos Aires. Argentina.

Ortega A. (1987). "*Tablas de mortalidad*". CELADE. Serie E N° 1004. San José. Costa Rica.

Serfaty. E (2003) "*Mortalidad por causas violentas en adolescentes y jóvenes de 10 a 24 años. Argentina 1991 a 2000*". Centro de Investigaciones Epidemiológicas de la Academia Nacional de Medicina de Buenos Aires.

Wang. J.Z.. (2007) "*Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model*" Mathematical Statistics Stockholm University.

Wilmoth. J. R. (1993). "*Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality change.*" Technical Report. Department of Demography". University of California. Berkeley.