

# Upper and lower bounds for annuities and life insurance from incomplete mortality data

Filipe Costa de Souza<sup>1</sup>

 <https://orcid.org/0000-0001-9903-5403>

E-mail: filipe.costas@ufpe.br

<sup>1</sup> Universidade Federal de Pernambuco, Centro de Ciências Sociais Aplicadas, Departamento de Ciências Contábeis e Atuariais, Recife, PE, Brazil

Received on 01.05.2018 – Desk acceptance on 02.15.2018 – 2<sup>nd</sup> version approved on 08.04.2018 – Ahead of print on 03.25.2019

Associate Editor: Luís Eduardo Afonso

## ABSTRACT

This study aimed to set upper and lower bounds for the expected present value of whole life annuities and whole life insurance policies from incomplete mortality data, generalizing previous results on life expectancy. Since its inception, in the 17<sup>th</sup> century, actuarial science has been devoted to the study of annuities and insurance plans. Thus, setting intervals that provide an initial idea about the cost of these products using incomplete mortality data represents a theoretical contribution to the area and this may have major applications in markets lacking historical records or those having little reliability of mortality data, as well as in new markets still poorly explored. For both the continuous and discrete cases, upper and lower bounds were constructed for the expected present value of whole life annuities and whole life insurance policies, contracted by a person currently aged  $x$ , based on information about the expected present value of these respective financial products subscribed to by a person of age  $x + n$  and the probability that an individual of age  $x$  survives to at least age  $x + n$ . Through the bounds of a continuous annuity, in an environment where the instantaneous interest rate is equal to zero, the results shown also set bounds for the complete life expectancy, which implies that the contribution of this research generalizes previous results in the literature. It was also found that, for both annuities and insurance plans, the length of constructed intervals increases as the data gap size increases and it decreases as the survival curve becomes more rectangular. Illustratively, bounds for life expectancy at 40 and 60 years of age, for the 10 municipalities showing the highest life expectancy at birth in Brazil in 2010, were constructed by using data available in the Atlas of Human Development in Brazil.

**Keywords:** actuarial mathematics, actuarial science, annuities, life insurance, mortality table.

## Correspondence address

Filipe Costa de Souza

Universidade Federal de Pernambuco, Centro de Ciências Sociais Aplicadas, Departamento de Ciências Contábeis e Atuariais  
Avenida dos Funcionários, S/N – CEP 50670-901  
Cidade Universitária – Recife – PE – Brazil



## 1. INTRODUCTION

According to Pitacco, Denuit, Haberman and Olivieri (2009), actuarial science flourished in the mid-17<sup>th</sup> century, based on compound interest rate theory and probability theory, as well as observations on mortality. Also according to these authors, one of the first themes to be addressed by this new science was calculating the expected present value of annuities. Such interest arose because governments used to sell whole life annuities as a way to finance public enterprises.

Pitacco et al. (2009) also indicate that Jan de Witt, in 1671, was the precursor in the calculation of annuities using a hypothetical mortality table and a constant interest rate. However, as well emphasized by Haberman and Sibbet (1995), de Witt's contribution had little repercussion at the time. The same, in turn, cannot be said of Edmund Halley's work, in 1693, which, in addition to developing a mortality table by means of actual observations, also introduced a method for calculating the cost of annuities that reverberates up to the present day. The reader interested in historical aspects of actuarial science will benefit from the reading of Hald (1990) and Haberman and Sibbet (1995).

Thus, it can be noticed that actuarial science, since its inception, has taken as one of its central foci the study of annuities. To calculate the expected present value in an annuity for any given individual of age  $x$ , it is necessary to have access to a complete mortality table (having information from at least age  $x$ ) or a survival function representative of the population to which the person belongs. In some instances, it is also possible to generate a complete table by means of abridged life tables and use the resulting complete table to calculate the cost of desired annuity or insurance (Baili, Micheli, Montanari, & Capocaccia, 2005; Ibrahim, 2008).

However, when there is no data on the mortality probabilities (age after age) since age  $x$ , calculating the expected present value of desired annuity is compromised. The impossibility to obtain detailed mortality data for a sequence of ages may occur due to lack of historical records or little reliability of existing data or also because this is a new market that is still poorly explored, for instance.

The absence of complete mortality tables also hinders the calculation of longevity measures, as in the case of

the complete life expectancy. In this way, Cohen (2011) established upper and lower bounds on life expectancy at a given age  $x$ , knowing only detailed mortality data from age  $x + n$  (and, indeed, life expectancy at age  $x + n$ ), as well as the probability that a person aged  $x$  survives to at least age  $x + n$ .

This study aims to extend the results of Cohen (2011), by setting bounds (upper and lower) for the expected present value of annuities and life insurance policies subscribed to by a person of age  $x$ , both in the continuous and discrete cases. Based on the definition of these bounds, it might be possible to estimate the actuarial premium of these financial products lacking detailed knowledge of the mortality profile between age  $x$  and a given age  $x + n$ .

It is worth emphasizing that, although the absence of data for more advanced ages is more common, it is understood that rather atypical cases also deserve careful discussion. In practice, the data gap for younger ages could occur, for instance, in the so-called Special Regimes of Social Security (Regimes Próprios de Previdência Social – RPPS) – or also in pension funds – which have few employees in the early working ages (or even a sequence of early ages without employees) and those in need to estimate mortality tables and/or calculate the cost of annuities for retirement purposes. So, this study brings a theoretical contribution to the literature, by generalizing the results of Cohen (2011), with a potential practical applicability.

In order to achieve the goal outlined, the remainder of the text is organized as follows: section 2 presents the theoretical framework, focusing on the calculation of net single premiums for whole life annuities and life insurance, as well as an exposition of the results of Cohen (2011) for life expectancy. Section 3 shows the bounds, upper and lower, for the expected present value of annuities and insurance policies; and section 4 shows numerical results in an illustrative way, also promoting discussions on the theme. In turn, section 5 brings the final remarks of this study. It is also worth mentioning that, throughout the text, the standard actuarial notation is adopted. For more details, see Bowers, Gerber, Hickman, Jones and Nesbitt (1997).

## 2. THEORETICAL FRAMEWORK

Suppose a complete mortality table that has as its maximum age, i.e. no individual is supposed to be alive at age  $\omega$ . Being  $q_x$  the probability that a person of age  $x$  dies along that age, this implies that  $q_{\omega-1} = 1$ . In addition,  ${}_t p_x$ , for  $0 \leq x \leq \omega$  a  $t \geq 0$ , indicates the probability that a person of age  $x$  survives to at least age  $x + t$ . In the continuous case,  ${}_t p_x$  is named as survival function. Of course, as a survival function,  ${}_t p_x$  is a non-increasing function of  $t$ , i.e. as  $t$  increases,  ${}_t p_x$  decreases or remains at the same value. In addition,  ${}_t p_x = 1$  and  ${}_t p_x = 0$  whenever  $t \geq \omega - x$ . In addition, being  $\mu_x(t) = \lim_{dt \rightarrow 0^+} \frac{d({}_t p_{x+t})}{dt}$  the force of mortality (or instantaneous mortality rate) at age  $x + t$ , for  $t > 0$ , then, the probability that a person of age  $x$  survives to at least age  $x + t$  and dies instantaneously thereafter is defined by  ${}_t p_x \mu_x(t) dt$  (Dickson, Hardy, & Waters, 2013). Finally,  $i \geq 0$  is the annual effective interest rate in a compound capitalization regime and  $\delta \geq 0$  is the instantaneous interest rate in the continuous capitalization regime, so that  $\ln(1 + i) = \delta$ . In this way, the financial decapitalization factor is defined as  $v = 1/(1 + i) = e^{-\delta}$ .

So, as well taught by Dickson, Hardy and Waters (2013), the expected present value of whole life continuous annuity subscribed to by a person aged  $x$  is given by:

$$\bar{a}_x = \int_0^{\omega-x} {}_t p_x \cdot e^{-\delta t} dt. \quad [1]$$

The discrete case, i.e. the net single premium for a whole life annuity-due subscribed to by a person aged  $x$  is defined as:

$$\ddot{a}_x = \sum_{t=0}^{\omega-x-1} {}_t p_x \cdot v^t. \quad [2]$$

Being  $\bar{A}_x$  the expected present value of a whole life insurance policy in the continuous case subscribed to by a person aged  $x$ , and being  $A_x$  the discrete counterpart, i.e. paying a monetary unit in the end of the contracting party's death year, then such values are formally defined as:

$$\bar{A}_x = \int_0^{\omega-x} {}_t p_x \cdot \mu_x(t) \cdot e^{-\delta t} dt \quad [3]$$

$$A_x = \sum_{t=0}^{\omega-x-1} {}_t p_x \cdot q_{x+t} \cdot v^t. \quad [4]$$

It is possible to associate the actuarial present value of a whole life insurance benefit (continuous and discrete) to the present value of the respective annuities as follows:

$$\bar{A}_x = 1 - \delta \cdot \bar{a}_x \quad [5]$$

$$A_x = 1 - (1 - v) \cdot \ddot{a}_x. \quad [6]$$

Further details about the relation between annuities and insurance are presented by Dickson, Hardy and Waters (2013).

To calculate both the value of  $\bar{a}_x$  and  $A_x$  (when  $i > 0$ ), it is necessary to know the entire survival history since age  $x$ . The same requirement would also be needed if we wished to calculate the complete life expectancy of a person aged  $x$ , which is defined as:

$$\dot{e}_x = \int_0^{\omega-x} {}_t p_x dt. \quad [7]$$

In this spirit, Cohen (2011) set bounds, upper and lower, on life expectancy of a person aged  $x$ ,  $\dot{e}_x$ , even lacking detailed knowledge of the mortality history between ages  $x$  and  $x + n$ , as long as the probability that an individual aged  $x$  survives to at least age  $x + n$  is known,  ${}_n p_x$ , as well as  $\dot{e}_{x+n}$ . Formally, according to the author, such bounds are defined as:

$${}_n p_x (n + \dot{e}_{x+n}) \leq \dot{e}_x \leq n + {}_n p_x \cdot \dot{e}_{x+n}. \quad [8]$$

So, using these bounds, a researcher could have better knowledge about, or even to approximate, the value of  $\dot{e}_x$ . An example of the application of this result is the study by Rabbi (2013), who calculated the upper and lower bounds for life expectancy in various age groups, as a way of analyzing trends in life expectancy in Bangladesh.

Besides the clear practical relevance, this result was also used by the author in the development of other theoretical results. Cohen (2015) showed sharper bounds for Markov and Chebyshev's inequalities when information on conditional expectation is available. According to the author, Markov's inequality provides an upper bound for the probability that a non-negative random variable takes a value greater than  $n$  (for  $x > 0$ ). Thus, for instance, being such a random variable the

time lived by a newborn up to her/his death, Markov's inequality would indicate that:

$${}_n p_0 \leq \frac{\dot{e}_0}{n}. \tag{9}$$

However, from (8), once  $\dot{e}_n$  (conditional expectation) is known, it is possible to refine inequality as follows:

$${}_n p_0 \leq \frac{\dot{e}_0}{n + \dot{e}_n}. \tag{10}$$

So, these results only reinforce the importance of the study by Cohen (2011), stimulating the extension of these ideas to other areas.

### 3. UPPER AND LOWER BOUNDS FOR ANNUITIES AND LIFE INSURANCE

In this section, the idea proposed by Cohen (2011) is extended to the case of continuous annuities, as set out in Proposition 1. In addition, the discrete case is also presented in Proposition 2. However, it is first necessary to define the present value of a given annuity-certain,

immediate, temporary per  $n$  years and paying one monetary unit per year, for both the continuous and the discrete (in advance) cases, as expressed in equations (11) and (12), respectively:

$$\bar{a}_{\overline{n}|} = \int_0^n e^{-\delta t} dt = \begin{cases} n, & \delta = 0 \\ \frac{1 - e^{-\delta n}}{\delta}, & \delta > 0. \end{cases} \tag{11}$$

$$\ddot{a}_{\overline{n}|} = \sum_{t=0}^{n-1} v^t = \begin{cases} n, & i = 0 \\ \frac{1 - v^n}{1 - v}, & i > 0. \end{cases} \tag{12}$$

**Proposition 1 (continuous case):** being  $x$  and  $x + n$  two ages such that  $0 \leq x \leq x + n < \omega$ , it is possible to state that

$$(\bar{a}_{\overline{n}|} + \bar{a}_{x+n} \cdot e^{-\delta n}) \cdot {}_n p_x \leq \bar{a}_x \leq \bar{a}_{\overline{n}|} + {}_n p_x \cdot e^{-\delta n} \cdot \bar{a}_{x+n} \tag{13}$$

This proposition sets upper and lower bounds on  $\bar{a}_x$  using known values of  $\bar{a}_{x+n}$  and  ${}_n p_x$ . The proof follows the same line as Cohen (2011).

**Proof:** it is known that

$$\begin{aligned} \bar{a}_x &= \int_0^{\omega-x} {}_t p_x \cdot e^{-\delta t} dt. \\ &= \int_0^n {}_t p_x \cdot e^{-\delta t} dt + \int_n^{\omega-x} {}_t p_x \cdot e^{-\delta t} dt \\ &= \int_0^n {}_t p_x \cdot e^{-\delta t} dt + {}_n p_x \cdot e^{-\delta n} \cdot \int_0^{\omega-x-n} {}_t p_{x+n} \cdot e^{-\delta t} dt. \end{aligned}$$

Since it is a probability, one must have that  ${}_t p_x \leq 1$ , for every  $t \geq 0$ , so

$$\int_0^n {}_t p_x \cdot e^{-\delta t} dt \leq \int_0^n e^{-\delta t} dt = \bar{a}_{\overline{n}|}.$$

Therefore,

$$\begin{aligned}\bar{a}_x &\leq \int_0^n e^{-\delta t} dt + {}_n p_x \cdot e^{-\delta n} \cdot \int_0^{\omega-x-n} {}_t p_{x+n} \cdot e^{-\delta t} dt \\ &= \bar{a}_{\bar{n}|} + {}_n p_x \cdot v^n \cdot \bar{a}_{x+n}.\end{aligned}$$

Now, it remains to prove the lower bound. Thus, return to the point where

$$\bar{a}_x = \int_0^n {}_t p_x \cdot e^{-\delta t} dt + {}_n p_x \cdot e^{-\delta n} \cdot \int_0^{\omega-x-n} {}_t p_{x+n} \cdot e^{-\delta t} dt.$$

Since  ${}_t p_x$  is a non-increasing function of  $t$ , so  ${}_n p_x \leq {}_t p_x$ , for  $0 \leq t \leq n$ . Indeed,

$$\int_0^n {}_t p_x \cdot e^{-\delta t} dt \geq \int_0^n {}_n p_x \cdot e^{-\delta t} dt = {}_n p_x \cdot \bar{a}_{\bar{n}|}$$

Therefore,

$$\bar{a}_x \geq (\bar{a}_{\bar{n}|} + \bar{a}_{x+n} \cdot e^{-\delta n}) \cdot {}_n p_x,$$

which ends the proof.

After this demonstration, two points deserve to be highlighted. First, Proposition 1 is, in fact, an extension of the findings by Cohen (2011). To verify this, it is enough to assume that  $\delta = 0$  in (13), from there it naturally follows that

$$(n + \dot{e}_{x+n}) \cdot {}_n p_x \leq \dot{e}_x \leq n + {}_n p_x \cdot \dot{e}_{x+n},$$

that is, the result of Cohen (2011) is a particular case of Proposition 1 when  $\delta = 0$ .

Second, for  $\delta > 0$ , and based on the relation between  $\bar{A}_x$  and  $\bar{a}_x$  presented in Equation (5), it is also possible to establish bounds for  $\bar{A}_x$  using an a priori knowledge about  $\bar{A}_{x+n}$  and  ${}_n p_x$ . In this way, we have:

$$e^{-\delta n}(1 - {}_n p_x) + {}_n p_x \cdot e^{-\delta n} \cdot \bar{A}_{x+n} \leq \bar{A}_x \leq 1 - {}_n p_x + {}_n p_x \cdot e^{-\delta n} \cdot \bar{A}_{x+n} \quad \boxed{14}$$

To develop this result, realize initially that replacing  $\bar{a}_x$  by the upper bound,  $\bar{a}_{\bar{n}|} + {}_n p_x \cdot e^{-\delta n} \cdot \bar{a}_{x+n}$ , in Equation (5), would imply that:

$$\bar{A}_x \geq 1 - \delta \left( \frac{1 - e^{-\delta n}}{\delta} + {}_n p_x \cdot e^{-\delta n} \cdot \bar{a}_{x+n} \right)$$

and how, also using Equation (5), we know that

$$\bar{a}_{x+n} = \frac{1 - \bar{A}_{x+n}}{\delta},$$

so,

$$\bar{A}_x \geq 1 - \delta \left[ \frac{1 - e^{-\delta n}}{\delta} + {}_n p_x \cdot e^{-\delta n} \cdot \left( \frac{1 - \bar{A}_{x+n}}{\delta} \right) \right]$$

$$= e^{-\delta n}(1 - {}_n p_x) + {}_n p_x \cdot e^{-\delta n} \cdot \bar{A}_{x+n},$$

thus establishing the lower bound for  $\bar{A}_x$ .

Similarly, replacing  $\bar{a}_x$  by the lower bound,  $(\bar{a}_{\bar{n}|} + \bar{a}_{x+n} \cdot e^{-\delta n}) \cdot {}_n p_x$ , in Equation (5), would imply that:

$$\bar{A}_x \leq 1 - \delta \left( \frac{1 - e^{-\delta n}}{\delta} + e^{-\delta n} \cdot \bar{a}_{x+n} \right) \cdot {}_n p_x$$

$$= 1 - \delta \left[ \frac{1 - e^{-\delta n}}{\delta} + e^{-\delta n} \cdot \left( \frac{1 - \bar{A}_{x+n}}{\delta} \right) \right] \cdot {}_n p_x$$

$$= 1 - {}_n p_x + {}_n p_x \cdot e^{-\delta n} \cdot \bar{A}_{x+n},$$

which, in turn, sets the upper bound for  $\bar{A}_x$ .

Once the result is known for the continuous case, the discrete case is developed in a similar way, as set out in Proposition 2.

**Proposition 2 (discrete case):** being  $x$  and  $x + n$  two integer ages, such that  $0 \leq x \leq x + n < \omega$ , it is possible to state that

$$(\ddot{a}_{\overline{n}|} + \ddot{a}_{x+n} \cdot v^n) \cdot {}_n p_x \leq \ddot{a}_x \leq \ddot{a}_{\overline{n}|} + {}_n p_x \cdot v^n \cdot \ddot{a}_{x+n} \tag{15}$$

**Proof:** it is known that

$$\begin{aligned} \ddot{a}_x &= \sum_{t=0}^{\omega-x-1} {}_t p_x \cdot v^t \\ &= \sum_{t=0}^{n-1} {}_t p_x \cdot v^t + \sum_{t=n}^{\omega-x-1} {}_t p_x \cdot v^t \\ &= \sum_{t=0}^{n-1} {}_t p_x \cdot v^t + {}_n p_x \cdot v^n \cdot \sum_{t=0}^{\omega-x-n-1} {}_t p_{x+n} \cdot v^t. \end{aligned}$$

Since it is a probability, one must have that  ${}_t p_x \leq 1$ , for  $t = 0, 1, \dots, n - 1$ , so

$$\sum_{t=0}^{n-1} {}_t p_x \cdot v^t \leq \sum_{t=0}^{n-1} v^t = \ddot{a}_{\overline{n}|}.$$

Therefore,

$$\begin{aligned} \ddot{a}_x &\leq \sum_{t=0}^{n-1} v^t + {}_n p_x \cdot v^n \cdot \sum_{t=0}^{\omega-x-n-1} {}_t p_{x+n} \cdot v^t \\ &= \ddot{a}_{\overline{n}|} + {}_n p_x \cdot v^n \cdot \ddot{a}_{x+n}. \end{aligned}$$

Now, it remains to prove the lower bound. Thus, return to the point where

$$\ddot{a}_x = \sum_{t=0}^{n-1} {}_t p_x \cdot v^t + {}_n p_x \cdot v^n \cdot \sum_{t=0}^{\omega-x-n-1} {}_t p_{x+n} \cdot v^t.$$

Since  ${}_t p_x$  is a non-increasing function of  $t$ , so  ${}_n p_x \leq {}_t p_x$  for  $t = 0, 1, \dots, n - 1$ . Indeed,

$$\sum_{t=0}^{n-1} {}_t p_x \cdot v^t \geq \sum_{t=0}^{n-1} v^t \cdot {}_n p_x = {}_n p_x \cdot \ddot{a}_{\overline{n}|}.$$

Therefore,

$$\ddot{a}_x \geq (\ddot{a}_{\overline{n}|} + \ddot{a}_{x+n} \cdot v^n) \cdot {}_n p_x,$$

which ends the proof.

Again, for  $i > 0$ , and using Equation (6), the bounds are set for  $A_x$  by means of prior knowledge about  $A_{x+n}$  and  ${}_n p_x$ . In this way, we have:

$$v^n(1 - {}_n p_x) + {}_n p_x \cdot v^n \cdot A_{x+n} \leq A_x \leq 1 - {}_n p_x + {}_n p_x \cdot v^n \cdot A_{x+n} \tag{16}$$

The development of this result is omitted, since it follows the same argument introduced for the continuous case and, therefore, it can be easily replicated by the reader.

## 4. APPLICATION AND DISCUSSIONS

This section suggests applications and provides numerical examples about the results demonstrated in section 3. In addition, the effects of variation in the interest rate and the data gap (i.e. the value of  $n$ ) are discussed, as well as the retangularization of the survival curve.

Initially, for illustrative purposes, intervals for the values of  $\ddot{a}_{20}$  and  $A_{20}$  are calculated by considering the complete availability of data from the ages of 20, 30, 40, 50, 60, and 70 years, i.e. the intervals were calculated by means of different values for the data gap. In addition, still

for illustrative purposes, the central point of the interval is considered an estimate of the actuarial fair value of the respective financial products and, based on this estimate, the error (estimated value less actual value, which uses the complete data since 20 years) is also computed. For the calculations, the 'IBGE 2015 mortality table' was considered for both sexes – extrapolated for ages over 80 years, and an effective interest rate of 3% per year. The results for  $\ddot{a}_{20}$  are summarized in Table 1, while the results for  $A_{20}$  are set out in Table 2.

**Table 1**

*Bounds, estimates, and errors for  $\ddot{a}_{20}$  based on complete data on mortality from the ages of 20, 30, 40, 50, 60, and 70 years*

Age	Data gaps (n) - in years	Lower bound	Upper bound	Central value	Length of the interval	Error	Error percentage
20	0	27.21419	27.21419	27.21419	0.00000	0.00000	0.00000%
30	10	27.13380	27.27153	27.20267	0.13773	-0.01152	0.04234%
40	20	26.88249	27.42508	27.15379	0.54259	-0.06040	0.22194%
50	30	26.24663	27.65945	26.95304	1.41283	-0.26115	0.95960%
60	40	24.72109	28.00480	26.36294	3.28371	-0.85124	3.12794%
70	50	21.47332	28.50447	24.98890	7.03115	-2.22529	8.17695%

Source: Prepared by the author.

**Table 2**

*Bounds, estimates, and errors for  $A_{20}$  based on complete data on mortality from the ages of 20, 30, 40, 50, 60, and 70 years*

Age	Data gaps (n) - in years	Lower bound	Upper bound	Central value	Length of the interval	Error	Error percentage
20	0	0.20735	0.20735	0.20735	0.00000	0.00000	0.00000%
30	10	0.20568	0.20969	0.20768	0.00401	0.00033	0.16183%
40	20	0.20121	0.21701	0.20911	0.01580	0.00175	0.84840%
50	30	0.19438	0.23553	0.21496	0.04115	0.00760	3.66823%
60	40	0.18551	0.27382	0.22966	0.08831	0.02231	10.7621%
70	50	0.16977	0.37456	0.27216	0.20479	0.06481	31.2578%

Source: Prepared by the author.

It is noticed in Table 1 that the estimate of  $\ddot{a}_{20}$  using the central value of the interval for the age of 30 years underestimates the actual value of the annuity concerned by less than 0.05%. In turn, if we used data referring to 70 years, where there is a 50-year data gap, the value of  $\ddot{a}_{20}$  would be underestimated by just over 8%. In turn, in Table 2, the estimate of  $A_{20}$  using the central value of the interval for the age of 30 years might overestimate the actual value of the annuity concerned by only 0.16%. However, if data referring to 70 years were used, the value

would be overestimated by over 31%.

As a way of visualizing the effects that a variation in the interest rate would produce on the results, using the same mortality table, intervals were constructed for the value of  $\ddot{a}_{20}$  and  $A_{20}$  considering data from 30 years and varying the value of interest rate. The results are detailed in tables 3 and 4, respectively. In these latter, given the choices made, it is observed that, in both cases, as the interest rate increases, the error also tends to increase.

**Table 3**

*Bounds, estimates, and erros for  $\ddot{a}_{20}$  based on data from the age of 30 years in the face of various interest rates*

Interest rate	Lower bound	Upper bound	Central value	Length of the interval	Error	Error percentage
1%	43.27120	43.42116	43.34618	0.14996	-0.01012	0.02335%
2%	33.71009	33.85372	33.78191	0.14363	-0.01086	0.03214%
3%	27.13380	27.27153	27.20267	0.13773	-0.01152	0.04234%
4%	22.46578	22.59801	22.53189	0.13223	-0.01211	0.05371%
5%	19.05272	19.17982	19.11627	0.12710	-0.01263	0.06604%

**Source:** Prepared by the author.

**Table 4**

*Bounds, estimates, and erros for  $A_{20}$  based on data from the age of 30 years in the face of various interest rates*

Interest rate	Lower bound	Upper bound	Central value	Length of the interval	Error	Error percentage
1%	0.57008	0.57157	0.57083	0.00148	0.00010	0.01756%
2%	0.33620	0.33901	0.33761	0.00281	0.00021	0.06312%
3%	0.20568	0.20969	0.20768	0.00401	0.00033	0.16183%
4%	0.13084	0.13593	0.13338	0.00508	0.00046	0.35037%
5%	0.08667	0.09272	0.08970	0.00605	0.00060	0.67511%

**Source:** Prepared by the author.

At this point, two aspects deserve to be highlighted: first, it would only be possible to calculate the error of having the actual values of financial products; but, surely, such values would not be available in situations involving incomplete mortality data. Second, the error would depend on the value used as an estimate. If, for instance, a researcher decided to conservatively use values greater than the central value of the interval as a way of analyzing the expected present value of annuities and life insurance policies, then the results on error might be different. Therefore, it is useful to evaluate how the length of the interval produced behaves in face of variations in the parameters, as the next subsection does.

### 4.1 The Length of Intervals

Once some illustrative results have been shown, it is worth formally discussing the impact of certain parameters on the length of the interval (upper bound minus lower bound) produced. Being  $T_a$  the length of the interval formed by the bounds in (15), so  $T_a$  is defined by the expression:

$$T_a = \ddot{a}_{\overline{n}|} (1 - {}_n p_x). \tag{17}$$

Using Equation (17), it is noticed that, everything else being constant,  $T_a$  increases with increases in  $n$ , since  $\ddot{a}_{\overline{n}|}$  is an increasing function of  $n$  and  ${}_n p_x$  is a non-increasing function of  $n$  (see example in Table 1); just as we observe that  $T_a$  decreases with increases in  $i$ , since  $\ddot{a}_{\overline{n}|}$  is a decreasing function of  $i$  (Faro, 2006). This result is shown in Table 3. Finally,  $T_a$  decreases with increases in  ${}_n p_x$ , since, as stated,

${}_n p_x$  is a non-increasing function of  $n$ .

This last result refers to the phenomenon of rectangularization of the survival curve. The rectangularization process, as well shown by Wilmoth and Horiuchi (1999), is characterized by high survival rates in childhood and adulthood, and rapid mortality in advanced ages. Thus, in a hypothetical case where  ${}_n p_x = 1$ , then the lower and upper bounds in (15) are equal and, indeed,  $\ddot{a}_x = \ddot{a}_{\overline{n}|} + \ddot{a}_{x+n} \cdot v^n$ . This fact reinforces the result that the more rectangular the survival curve, the smaller distance between the lower and upper bounds of  $\ddot{a}_{20}$ , i.e. the smaller interval is produced.

Just as in the discussion of the bounds of an annuity, it is also relevant to analyze the bounds for a life insurance. For this product, being  $T_A$  the length of the interval formed by the bounds in (16), then  $T_A$  is defined by:

$$T_A = (1 - v^n) \cdot (1 - {}_n p_x). \tag{18}$$

For (18), it is observed that  $T_A$  increases with increases in  $n$  (see Table 2) and decreases with increases in  ${}_n p_x$ , similarly to what happens with  $T_a$ . However,  $T_A$  increases with increases in  $i$ , since  $(1 - v^n)$  is an increasing function of  $i$  (see Table 4).

### 4.2 Defining Bounds for $e'_{40}$ and $e'_{60}$ Using the Atlas of Human Development in Brazil

This subsection presents an actual application of the results of Cohen (2011) and, consequently, of this

research. This is the calculation of bounds for  $\overset{\circ}{e}_{40}$  and  $\overset{\circ}{e}_{60}$  using information from the Atlas of Human Development in Brazil (Programa das Nações Unidas para o Desenvolvimento Humano, 2013a).

Life expectancy at birth is the main longevity measure and it is commonly used as an indicator of human development (Mayhew & Smith, 2015). Due to their importance, data on life expectancy at birth are typically set out in research on demography, public health, etc. However, when there is no detailed data on mortality, such availability does not occur as easily for life expectancy at other ages. Mathers, Stevens, Boerma, White and Tobias (2015) point out that life expectancy at the age of 60 years, for instance, is a relevant indicator of longevity for older people and knowing it for a given population is key to public planning in social security and health, among other areas.

Originally using the result of Cohen (2011), it has been found that we can estimate upper and lower bounds for complete life expectancy at age  $x$ , only knowing the probability that a person aged  $x$  is alive at age  $x + n$  and the complete life expectancy at age  $x + n$ . Of course, using the same argument, bounds for  $\overset{\circ}{e}_{x+n}$  are constructed by means of prior knowledge about  ${}_n p_x$  and  $\overset{\circ}{e}_x$ . That is, it

follows directly from (8) that:

$$\frac{\overset{\circ}{e}_x - n}{n p_x} \leq \overset{\circ}{e}_{x+n} \leq \frac{\overset{\circ}{e}_x}{n p_x} - n. \quad \boxed{19}$$

The Atlas of Human Development in Brazil (Programa das Nações Unidas para o Desenvolvimento Humano, 2013b) is a tool for accessing the Municipal Human Development Index, for the years 1991, 2000, and 2010 and it is available at Programa das Nações Unidas para o Desenvolvimento Humano (2013a). Using this tool, it is possible, for instance, for the mentioned years, to obtain information about life expectancy at birth, the probability that a newborn survives until the age of 40 years, and the probability that a newborn survives until the age of 60 years, for all Brazilian municipalities. It is worth noticing that, in 2010, Brazil had 5,565 municipalities. So, it is possible to construct intervals for life expectancy values at ages 40 and 60 years for Brazilian municipalities in the respective years.

Illustratively, Table 5 displays the bounds for  $\overset{\circ}{e}_{40}$  and  $\overset{\circ}{e}_{60}$  for the 10 municipalities with the highest life expectancy at birth in Brazil, in the year 2010. Curiously, all these municipalities belong to the state of Santa Catarina. Data were collected in the Atlas of Human Development in Brazil Programa das Nações Unidas para o Desenvolvimento Humano, 2013a) on December 22, 2017.

**Table 5**

*Bounds for  $\overset{\circ}{e}_{40}$  and  $\overset{\circ}{e}_{60}$  for the 10 municipalities with the highest life expectancy at birth in Brazil, in the year 2010*

Municipality	Lower bound for $\overset{\circ}{e}_{40}$	Upper bound for $\overset{\circ}{e}_{40}$	Length of the interval for $\overset{\circ}{e}_{60}$	Lower bound for $\overset{\circ}{e}_{60}$	Upper bound for $\overset{\circ}{e}_{60}$	Length of the interval for $\overset{\circ}{e}_{60}$
Blumenau	40.29197	42.00209	1.710115	21.15297	29.24194	8.08897
Brusque	40.29197	42.00209	1.710115	21.15297	29.24194	8.08897
Balneário Camboriú	40.26272	41.96414	1.701418	21.11829	29.16865	8.050357
Rio do Sul	40.22294	41.89395	1.671007	21.05918	28.95553	7.896345
Rancho Queimado	40.20629	41.88164	1.675349	21.04131	28.95303	7.911715
Rio do Oeste	40.14591	41.83429	1.688379	20.98346	28.96442	7.980965
Joaçaba	40.07924	41.78501	1.705766	20.91652	28.97459	8.058076
Iomerê	40.07924	41.78501	1.705766	20.91652	28.97459	8.058076
Porto União	40.06881	41.77458	1.705766	20.90992	28.98344	8.073519
Nova Trento	40.06881	41.77458	1.705766	20.90992	28.98344	8.073519

**Source:** Prepared by the author.

Taking, for instance, the case of Blumenau and if the central value of the interval (nearly 41.15 years) was used as an estimate for  $\overset{\circ}{e}_{40}$ ; even without knowing the actual value of this life expectancy, it is certain that, in this context, the error caused by using such estimate would be, at most, 312 days, i.e. a little more than 10 months.

Thus, using these intervals could provide relevant information for research on the heterogeneity of the Brazilian population's mortality and promote discussions on the adoption of a minimum retirement age, particularly in the current political-economic context that the country has experienced (Souza, 2018).

## 5. FINAL REMARKS

In this study, and inspired by the work of Cohen (2011), the upper and lower bounds were demonstrated for the expected present value of annuities and life insurance policies. In addition, through the analysis of bounds for continuous annuities, it may be seen that the results shown are an extension of those provided by Cohen (2011), and this extended the discussion to an actuarial context.

For both annuities and life insurance, the length of constructed intervals increases with increases in the data gap and it decreases as the survival curve becomes more rectangular. The interest rate, in turn, has a different effect

on the intervals, increasing them when the interest rate increases, in the case of insurance plans and decreasing them in the case of annuities.

Besides the theoretical contribution, it is worth pointing out that the definition of these intervals may have a significant potential for applicability in markets where there is no historical record of mortality for certain ages or where data are unreliable or even in new markets that are still poorly explored. And, in addition to be used for pricing these financial products, they can provide information for research on demography, public health, social security, and public planning, for instance.

## REFERENCES

- Baili, P., Micheli, A., Montanari, A., & Capocaccia, R. (2005). Comparison of four methods for estimating complete life tables from abridged life tables using mortality data supplied to EURO-CARE-3. *Mathematical Population Studies*, 12, 183-198.
- Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., & Nesbitt, C. J. (1997). *Actuarial mathematics* (2nd ed.). Schaumburg, IL: Society of Actuaries.
- Cohen, J. E. (2011). Life expectancy: lower and upper bounds from surviving fractions and remaining life expectancy. *Demographic Research*, 24, 251-256.
- Cohen, J. E. (2015). Markov's inequality and Chebyshev's inequality for tail probabilities: a sharper image. *The American Statistician*, 69(1), 5-7.
- Dickson, D. C., Hardy, M. R., & Waters, H. R. (2013). *Actuarial mathematics for life contingences risks* (2nd ed.). Cambridge: Cambridge University Press.
- Faro, C. (2006). *Fundamentos da matemática financeira: uma introdução ao cálculo financeiro e à análise de investimento de risco*. São Paulo: Saraiva.
- Haberman, S., & Sibbett, T. A. (1995). The history of actuarial science. In S. Haberman & T. A. Sibbett (Ed.), *History of actuarial science* (Vol. 1, pp. xix-lxxxiii). London: William Pickering.
- Hald, A. (1990). *A history of probability and statistics and their applications before 1750*. New York: John Wiley & Sons.
- Ibrahim, R. I. (2008). Expanding an abridged life table using the Heligman-Pollard model. *Matematika*, 24(1), 1-10.
- Mathers, C. D., Stevens, G. A., Boerma, T., White, R. A., & Tobias, M. I. (2015). Causes of international increases in older age life expectancy. *Lancet*, 385, 540-548.
- Mayhew, L., & Smith, D. (2015). On the decomposition of life expectancy and limits to life. *Population Studies*, 69(1), 73-89.
- Pitacco, E., Denuit, M., Haberman, S., & Olivieri, A. (2009). *Modelling longevity dynamics for pension and annuity business*. Oxford: Oxford University Press.
- Programa das Nações Unidas para o Desenvolvimento Humano. (2013a). Atlas do Desenvolvimento Humano no Brasil (homepage). Retrieved from <http://www.atlasbrasil.org.br/2013/>
- Programa das Nações Unidas para o Desenvolvimento Humano. (2013b). *O Índice de Desenvolvimento Humano Municipal Brasileiro*. Brasília, DF: PNDU/IPEA/FJD. (Série Atlas do Desenvolvimento Humano no Brasil 2013).
- Rabbi, A. M. F. (2013). Lower and upper bounds of life expectancies and total expected longevity of a developing country. *South East Asia Journal of Public Health*, 3(2), 52-57.
- Souza, F. C. (2018). A heterogeneidade da mortalidade da população brasileira e aspectos distributivos na previdência social: uma análise atuarial da proposta de idade mínima de aposentadoria. *Administração Pública e Gestão Social*, 10(1), 2-11.
- Wilmoth, J. T., & Horiuchi, S. (1999). Rectangularization revisited: variability of age at death within human population. *Demography*, 36(4), 475-495.