Dating Hypatia's birth: a probabilistic model

Canio Benedetto* , Stefano Isola†[a] and Lucio Russo‡[b]

[a]*Scuola di Scienze e Tecnologie, Università degli Studi di Camerino, Camerino, Italy*
[b]*Dipartimento di Matematica, Università degli Studi di Roma 'Tor Vergata', Roma, Italy*

Abstract

We propose a probabilistic approach as a dating methodology for events like the birth of a historical figure. The method is then applied to the controversial birth date of the Alexandrian scientist Hypatia, proving to be surprisingly effective.

Contents

*canio.benedetto@gmail.com
†stefano.isola@gmail.com
‡Corresponding author: russo@axp.mat.uniroma2.it

Jun 3 2016 09:40:48 PDT
Version 3 - Submitted to MEMOCS

1 Introduction

Although in historical investigation it may appear meaningless to do experiments on the basis of a pre-existing theory - and in particular it does not make sense to prove theorems of History - it can make perfect sense to use forms of reasoning typical of the exact sciences as an aid to increase the degree of reliability of a particular statement regarding a historical event. This paper deals with the problem of dating the birth of a historical figure when, on that event, one disposes of a set of indirect information, such as testimonials about various aspects of his/her life. The strategy is then based on the construction of a probability distribution for the birth date out of each testimony and subsequently combining the distributions so obtained in a sensible way. One might raise several objections to this program. According to Charles Sanders Peirce, a probability 'is the known ratio of frequency of a specific event to a generic event' (see [13]), but a birth is neither a specific event, nor a generic event, but an 'individual event'. Nevertheless, a probabilistic reasoning is used quite often in situations dealing with events which can be classified as 'individual'. In probabilistic forecasting, one tries to summarize what is known about future events with the assignation of a probability to each of a number of different outcomes which are often events of this kind. For instance, in sport betting a summary of bettors' opinions about the likely outcome of a race is produced in order to set bookmakers' pay-off rates. By the way, this type of observations lies at the basis of the theoretical formulation of the subjective approach in probability theory (see [5]). Although we do not endorse de Finetti's approach in all its implications, we embrace its severe criticism addressed to the exclusive use of the frequentist interpretation in the application of probability theory to concrete problems. In particular, we feel entitled to look at an 'individual' event of the historical past with a spirit similar to that with which one bets on a future outcome (this is a well known issue in the philosophy of probability, see, e.g., [3]). Plainly, as the information about an event like the birth of an historical figure is first extracted by a material drawn from various literary sources and then treated with mathematical tools, both our approach and goal are interdisciplinary in their essence.

2 A probabilistic method for combining testimonials

Let $X = [x_-, x_+] \subset \mathbb{Z}$ be the time interval which includes all possible birth dates of a given subject (*terminus ad quem*). X can be regarded as a set of mutually exclusive statements about a singular phenomenon (the birth of a given subject in a given year), only one of which is true, and can be made a probability space $(X, \mathcal{F}, P_0)$, with $\mathcal{F}$ the σ-algebra made of the $2^{|X|}$ events of interest and P_0 the uniform probability measure on $\mathcal{F}$ (reference measure): $P_0(A) = |A|/|X|$ (where $|A|$ denotes the number of elements of A). In the context of decision theory, the assignment of this probability space can be regarded as the expression of a basic state of knowledge, in absence of any information capable to discriminate among the possible statements on the given phenomenon, namely a situation in which it appears legitimate to apply Laplace's *principle of indifference*.

Now suppose to have k testimonials T_i, $i = 1, \ldots, k$, which in first approximation we may assume independent of each other, each providing some kind of information about the life of the subject, which can be translated into a probability distribution p_i on $\mathcal{F}$ so that $p_i(x)$ is the probability that the subject is born in the year $x \in X$ based on the information given by the testimony T_i, assumed true, along with supplementary information such as, e.g., life tables for the historical period considered. The precise criteria for the construction of these probability distributions depends on the kind of information carried by each testimonial and will be discussed case by case in the next section. Of course we shall take into account also of the possibility that some testimonials are false, not thereby producing any additional information. We model this possibility by assuming that the corresponding distributions equal the reference measure P_0.

The problem that we want to discuss in this section is the following: how to combine the distributions p_i in such a way to get a single probability distribution Q which somehow optimizes the available information? To address this question, let us observe that from the k testimonials taken together, each one with the possibility to be true or false, one gets $N = 2^k$ combinations, corresponding to as many binary words $\sigma_s = \sigma_s(1) \cdots \sigma_s(k) \in \{0, 1\}^k$, which can be ordered lexicographically according to $s = \sum_{i=1}^k \sigma_s(i) \cdot 2^{i-1} \in \{0, 1, \ldots, N-1\}$, and given by

$$P_s(\cdot) = \frac{\prod_{i=1}^k p_i^{\sigma_s(i)}(\cdot)}{\sum_{x \in X} \prod_{i=1}^k p_i^{\sigma_s(i)}(x)} \quad , \quad p_i^{\sigma_s(i)} = \begin{cases} p_i \ , & \sigma_s(i) = 1 \\ P_0 \ , & \sigma_s(i) = 0 \end{cases} \tag{2.1}$$

In particular, one readily verifies that P_0 is but the reference uniform measure.

Now, if Ω denotes the class of probability distributions $Q : X \to [0, 1]$, we look for a *pooling operator* $T : \Omega^N \to \Omega$ which combines the distributions P_s by weighing them in a sensible way. The simplest candidate has the general form of a linear combination

$$T(P_0, \ldots, P_{N-1}) = \sum_{s=0}^{N-1} w_s P_s; \quad w_s \geq 0, \quad \sum_{s=0}^{N-1} w_s = 1 \tag{2.2}$$

which, as we shall see, can also be obtained by minimizing some information theoretic function.

Remark 2.1 *The issue we are discussing here has been the object of a vast literature regarding the normative aspects of the formation of aggregate opinions in several contexts (see e.g. [7] and references therein). In particular, it has been shown by McConway (see [11]) that if one requires the existence of a function $F : [0, 1]^N \to [0, 1]$ such that*

$$T(P_0, \ldots, P_{N-1})(A) = F(P_0(A), \ldots, P_{N-1}(A)) \,, \quad \forall A \in \mathcal{F} \tag{2.3}$$

with $P_s(A) = \sum_{x \in A} P_s(x)$, then, whenever $|X| \geq 3$, F must necessarily have the form of a linear combination as in (2.2). The above condition implies in particular that the value of the combined distribution on coordinates depends only on the corresponding values on the coordinates of the distributions P_s, namely that the pooling operator commutes with marginalization.

However, some drawbacks of the linear pooling operator have also been highlighted. For example it does not 'preserve independency' in general: if $|X| \geq 5$, it is not true that $P_s(A \cap B) = P_s(A) P_s(B)$, $s = 0, \ldots, N-1$, entails

$$T(P_0, \ldots, P_{N-1})(A \cap B) = T(P_0, \ldots, P_{N-1})(A) T(P_0, \ldots, P_{N-1})(B)$$

unless $w_s = 1$ for some s and 0 for all others (see [10], [8]).

By the way, another form of the pooling operator considered in the literature to overcome the difficulties associated with the use of (2.2) is the log-linear combination

$$T(P_0, \ldots, P_{N-1}) = C \prod_{s=0}^{N-1} P_s^{w_s}; \quad w_s \geq 0, \quad \sum_{s=0}^{N-1} w_s = 1 \tag{2.4}$$

where C is a normalizing constant (see [7], [1]).
On the other hand, in our context the independence preservation property does not seem so desirable: the final distribution $T(P_0, \ldots, P_{N-1})$ relies on a set of information much wider than those associated to the single distributions P_s, and one can easily imagine how the alleged independence between two events can disappear as the information on them increases.

2.1 Optimization

The linear combination (2.2) can also be viewed as the marginal distribution[1] of $x \in X$ under the hypothesis that one of the distributions $P_0, \ldots, P_{N-1}$ is the 'true' one (without knowing which) (see [6]). In this perspective, (2.2) can be obtained by minimizing the expected loss of information due to the need to compromise, namely a function of the form

$$I(w, Q) = \sum_{s=0}^{N-1} w_s \, D(P_s \| Q) \geq 0 \tag{2.5}$$

where

$$D(P\|Q) = \sum_{x \in X} P(x) \log\left(\frac{P(x)}{Q(x)}\right) \tag{2.6}$$

is the *Kullback-Leibler divergence*, representing the information loss using the measure Q instead of P (see [9]). Note that the concavity of the logarithm and Jensen inequality yield

$$-\sum_x P(x) \log\left(\frac{P(x)}{Q(x)}\right) \leq \log \sum_x P(x) \frac{Q(x)}{P(x)} = 0$$

and therefore

$$D(P\|Q) \geq 0 \quad \text{and} \quad D(P\|Q) = 0 \Longleftrightarrow Q \equiv P \tag{2.7}$$

We have the following result.

Lemma 2.2 *Given a probability vector $w = (w_0, w_1, \ldots, w_{N-1})$, it holds*

$$\arg\min_{Q \in \Omega} I(w, Q) = Q_w \equiv \sum_s w_s P_s \tag{2.8}$$

Moreover,

$$I(w, Q_w) = H(\sum_s w_s P_s) - \sum_s w_s H(P_s) \tag{2.9}$$

where $H(Q) = -\sum_{x \in X} Q(x) \log Q(x)$ is the entropy of $Q \in \Omega$.

Proof. Eq. (2.8) can be obtained using the method of Lagrange multipliers. An alternative argument makes use of the easily derived 'parallelogram rule':

$$\sum_s w_s \, D(P_s\|Q) = \sum_s w_s \, D(P_s\|Q_w) + D(Q_w\|Q), \qquad \forall Q \in \Omega \tag{2.10}$$

From (2.7) we thus get $I(w, Q_w) \leq I(w, Q)$, $\forall Q \in \Omega$. The uniqueness of the minimum follows from the convexity of $D(P\|Q)$ w.r.t. Q. Finally, checking (2.9) is a simple exercise. ■

Remark 2.3 *It is worth mentioning that if we took $\sum_s w_s \, D(Q\|P_s)$ (instead of $\sum_s w_s \, D(P_s\|Q)$) as the function to be minimized (still varying Q with w fixed), then, instead of the 'arithmetic mean' (2.2), the 'optimal' distribution would have been the 'geometric mean' (2.4) (see also [1]).*

[1] In the sense that a marginal probability can be obtained by averaging conditional probabilities.

2.2 Allocating the weights

We have seen that for each probability vector w in the N-dimensional simplex $\{w_s \geq 0, \sum_{s=0}^{N-1} w_s = 1\}$ the distribution $Q_w = \sum_s w_s P_s$ is the 'optimal' one. We are now left with the problem of determining a sensible choice for w. This cannot be achieved by using the same criterion, in that by (2.7) $\inf_w I(w, Q_w) = 0$ and the minimum is realized whenever $w_s = 1$ for some s and 0 for all others.

A suitable expression for the weights w_s can be obtained by observing that the term $\sum_{x\in X}\prod_{i=1}^k p_i^{\sigma_s(i)}(x)$ is proportional to the probability of the event (in the product space $X^{[1,k]}$) in which the birth dates of k different subjects, with the i-th birth date distributed according to $p_i^{\sigma_s(i)}$, coincide, and thus it furnishes a measure of the degree of compatibility of the distributions p_i involved in the product associated to the word σ_s.

It appears thus natural to consider the weights

$$w_s = \frac{\sum_{x\in X}\prod_{i=1}^k p_i^{\sigma_s(i)}(x)}{\sum_{s=0}^{N-1}\sum_{x\in X}\prod_{i=1}^k p_i^{\sigma_s(i)}(x)} \tag{2.11}$$

which, once inserted in (2.2), yield the expression

$$T(P_0, \ldots, P_{N-1})(\,\cdot\,) = \frac{\sum_{s=0}^{N-1}\prod_{i=1}^k p_i^{\sigma_s(i)}(\,\cdot\,)}{\sum_{x\in X}\sum_{s=0}^{N-1}\prod_{i=1}^k p_i^{\sigma_s(i)}(x)} \tag{2.12}$$

Remark 2.4 *There are at least $k+1$ strictly positive coefficients w_s. They correspond to the words $\sigma_s^{(i)}$ with $\sigma_s^{(i)}(i) = 1$ for some $i \in \{1, \ldots, k\}$ and $\sigma_s^{(i)}(j) = 0$ for $j \neq i$, plus the one corresponding to the word 0^k, that is to the distributions $P_{s^{(i)}} \equiv p_i$, $i \in \{0, 1, \ldots, k\}$, where $p_0 \equiv P_0$.*

2.3 Weights as likelihoods

A somewhat complementary argument to justify the choice (2.11) for the coefficients w_s can be formulated in the language of probabilistic inference, showing that they can be interpreted as (normalized) *average likelihoods* associated to the various combinations corresponding to the words σ_s. More precisely, to each pair of 'hypotheses' of the form

$$D_i^e = \begin{cases} \{T_i \text{ true}\}\,, & e = 1, \\ \{T_i \text{ false}\}\,, & e = 0, \end{cases}$$

we associate its *likelihood, given the event in which the birth date is $x \in X$*, with the expression[2]

$$V(D_i^e|x) = \frac{\mathrm{P}(x|D_i^e)}{\mathrm{P}(x)} = \begin{cases} p_i(x)/p_0(x)\,, & e = 1, \\ 1\,, & e = 0, \end{cases} \tag{2.13}$$

with $i \in \{1, \ldots, k\}$ and $p_0 \equiv P_0$. In this way, the 'posterior probability' $\mathrm{P}(D_i^e|x)$, that is the probability of D_i^e in the light of the event in which the subject is born in the year $x \in X$, is given by the product of $V(D_i^e|x)$ with the 'prior probability' $\mathrm{P}(D_i^e)$, according to Bayes' formula.

[2]Here the symbol P denotes either the reference measure P_0, or any probability measure on X compatible with it.

If we now consider two pairs of 'hypotheses' $D_i^{e_i}$ and $D_j^{e_j}$, which we assume conditionally independent (without being necessarily independent), that is

$$\mathrm{P}(D_i^{e_i}, D_j^{e_j}|x) = \mathrm{P}(D_i^{e_i}|x)\mathrm{P}(D_j^{e_j}|x), \qquad e_i, e_j \in \{0,1\}$$

then we find

$$\begin{aligned} P(D_i^{e_i}, D_j^{e_j}|x) &= \frac{\mathrm{P}(x|D_i^{e_i}, D_j^{e_j})}{\mathrm{P}(x)} = \frac{\mathrm{P}(D_i^{e_i}, D_j^{e_j}|x)}{\mathrm{P}(D_i^{e_i}, D_j^{e_j})} = \frac{\mathrm{P}(D_i^{e_i}|x)\mathrm{P}(D_j^{e_j}|x)}{\mathrm{P}(D_i^{e_i}, D_j^{e_j})} \\ &= \frac{\mathrm{P}(D_i^{e_i})\mathrm{P}(D_j^{e_j})}{\mathrm{P}(D_i^{e_i}, D_j^{e_j})} \cdot V(D_i^{e_i}|x)V(D_j^{e_j}|x) \end{aligned}$$

More generally, given k testimonials T_i, to each of which there corresponds the pair of events D_i^e, and given a word $\sigma_s \in \{0,1\}^k$, if we assume the conditional independence of the events $(D_1^{\sigma_s(1)}, \ldots, D_k^{\sigma_s(k)})$, we get

$$V(D_1^{\sigma_s(1)}, \ldots, D_k^{\sigma_s(k)}|x) = \rho_s \prod_{i=1}^{k} V(D_i^{\sigma_s(i)}|x) \tag{2.14}$$

where

$$\rho_s = \frac{\prod_{i=1}^{k} \mathrm{P}(D_i^{\sigma_s(i)})}{\mathrm{P}(D_1^{\sigma_s(1)}, \ldots, D_k^{\sigma_s(k)})} \cdot \tag{2.15}$$

If, in addition, there is grounds to assume unconditional independence, i.e. $\rho_s = 1$, then (2.14) simply reduces to the product rule. Under this assumption, we can evaluate the *average likelihood* of the set of information $(D_1^{\sigma_s(1)}, \ldots, D_k^{\sigma_s(k)})$ with the expression

$$V_s = \frac{1}{|X|} \sum_{x \in X} V(D_1^{\sigma_s(1)}, \ldots, D_k^{\sigma_s(k)}|x) = |X|^{k-1} \sum_{x \in X} \prod_{i=1}^{k} p_i^{\sigma_s(i)}(x) \tag{2.16}$$

Comparing with (2.11) we see that

$$w_s = \frac{V_s}{\sum_{s=0}^{N-1} V_s} \tag{2.17}$$

In other words, within the hypotheses made so far, the allocation of the coefficients (2.11) corresponds to assigning to each distribution P_s a weight proportional to the average likelihood of the set of information from which it is constructed.

3 Hypatia

The method so far exposed is now applied to a particular dating process, the one of Hypatia's birth. This choice stems from the desire of studying a case both easy to handle and potentially useful in its results. The problem of dating Hypatia's birth is indeed open, in that there are different possible resolutions of the constraints imposed by the available data. According to the reconstruction given by Deakin, 'Hypatia's birth has been placed as early as 350 and as late as 375. Most authors settle for "around 370"' ([2], 51). Not many are the historical records concerning the birth of the Alexandrian scientist (far more are the ones about her infamous death), but they have the desirable feature to be independent from each other, as will be apparent in the sequel, so that the scheme discussed in the previous Section can be directly applied. The hope is to obtain something which is qualitatively significant when compared with the pre-existing proposals, based on a qualitative discussion of the sources, and quantitatively unambiguous. From each of the above mentioned historical records is extracted a probability distribution for the year of Hypatia's birth (the method is briefly discussed case by case), eventually all distributions are combined according the criteria outlined in the previous section.

3.1 Historical record 1 - Hypatia *floruit* between 395 and 408

The **Ὑπατία** entry of the *Suda Lexicon* reads as follows:

> **[Ὑπατία] ἤκμασεν ἐπί τῆς βασιλείας Ἀρκαδίου.**[3]

On the one hand, it's well known that Arcadius - first emperor of the Byzantine Empire - ruled from 395 to 408; on the other hand, it is not so straight-forward to associate a particular age or age interval to the Greek expression **ἤκμασεν**. Here is adopted the usual convention of the *floruit* indicating an age of approximately 40 years and so the following probability distribution $f(x)$ can be introduced (see Figure 1).

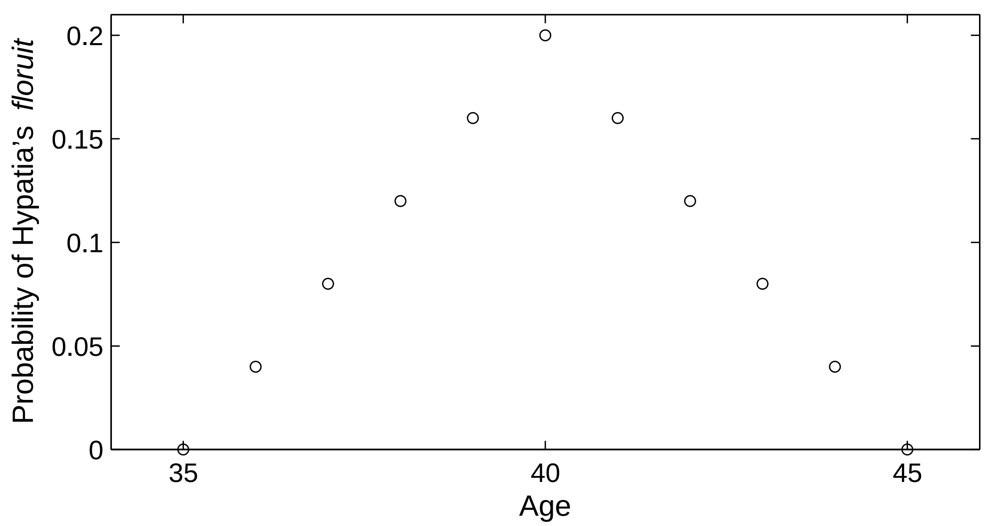

Figure 1: The probability distribution $f(x)$.

$\Upsilon_f(\xi)$, the probability distribution for the year of Hypatia's birth deduced from this historical record, is obtained by simply averaging the fourteen copies of the triangular $f(x)$, each placed centered around one of the years between 355 and 368 (the temporal limits of Arcadius' Empire, shifted by 40 years in the past - this being the value corresponding to the peak of $f(x)$) (see Figure 2).

[3][Hypatia] flourished under the emperor Arcadius.

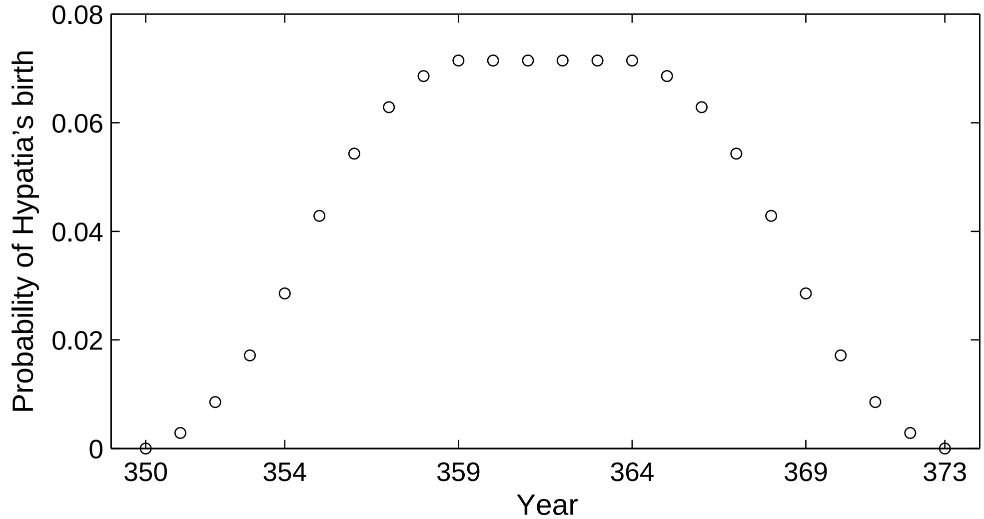

Figure 2: The probability distribution $\Upsilon_f(\xi)$.

3.2 Historical record 2 - Hypatia was *intellectually active* in 415

It is well documented how the martyrdom of Hypatia by the hand of a mob of Christian fanatics was moved by the envy that many felt both for her extraordinary intelligence and freedom of thought and her great political influence, being a woman.
For example, the above mentioned **Ὑπατία** entry of the *Suda Lexicon* states:

> **Τοῦτο δὲ πέπονθε διὰ φθόνον καὶ τὴν ὑπερβάλλουσαν σοφίαν, καὶ μάλιστα εἰς τὰ περὶ ἀστρονομίαν.**[4]

Socrates Scholasticus, in his **Εκκλησιαστική Ιστορία**, reports as follows:

> On account of the self-possession and ease of manner, which she had acquired in consequence of the cultivation of her mind, she not infrequently appeared in public in presence of the magistrates. Neither did she feel abashed in coming to an assembly of men. For all men on account of her extraordinary dignity and virtue admired her the more. Yet even she fell a victim to the political jealousy which at that time prevailed. For as she had frequent interviews with Orestes, it was calumniously reported among the Christian populace, that it was she who prevented Orestes from being reconciled to the bishop.[5]

Because of these and similar historical records, it seems reasonable to make 415 account for a year of intellectual activity in Hypatia's life.
To extract from this information a probability distribution for the year of birth, it is necessary to have in hand the one describing the probability of being intellectually active at all age, which can be calculated once one knows the probability of being alive at all age and to be active at all age (being alive), by simply multiplying them. In this work, the former is obtained from a 1974 mortality table (Italian males)[6], the latter is the result of an arbitrary model that seems reasonable: some obvious

[4]She suffered this [the violent death] because of the envy for her extraordinary wisdom, especially in the field of astronomy.

[5][15], Book VII, Chapter 15.

[6]All data are taken from `mortality.org`

and less obvious approximations are involved in this process, what follows tries to account for them. The choice operated for the first distribution could raise two main objections: one about the advisability in using a contemporary data set in a work about an Alexandrian scholar of the 4th century AD and one about the particular choice of the mortality table's itself.
Before addressing both these hypothetical objections, $a(x)$, the probability of being alive at all age is derived from the mentioned data set[7] (see Figure 3).

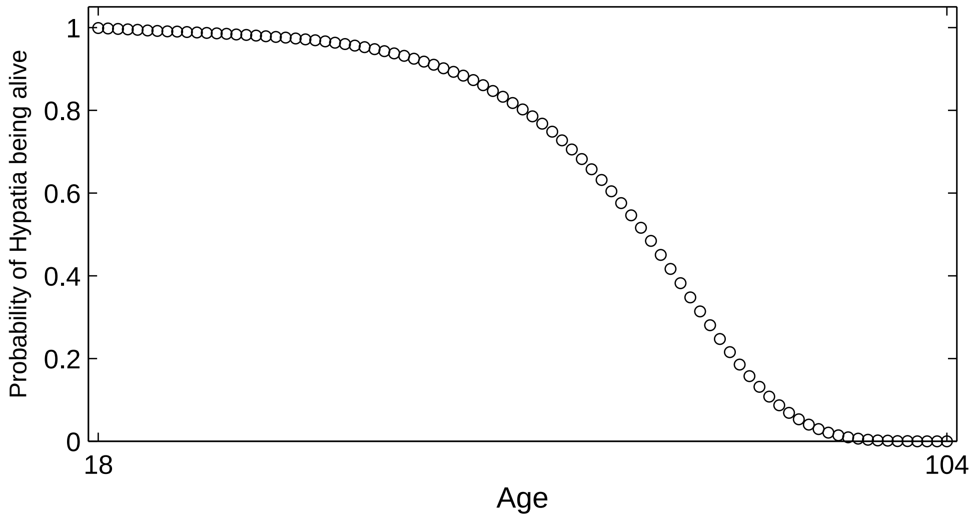

Figure 3: The probability distribution $a(x)$.

The particular choice is motivated by the evidence that the life expectancy inferable from it is in great agreement with a control value calculated for this purpose: the average lifespan of the first one hundred intellectuals to be 'well-dated' on *The Oxford Classical Dictionary*[8] (see Table 1).
Taking into account all the 'ancient intellectuals' and not only the ones lived in the 3rd/4th century AD is necessary to obtain a sample that is as statistically significant as possible, to use as an approximation for their mortality distribution the particular one here chosen seems to be justified by quantitative evidences.
$a_a(x)$ is used as a model of the probability of being intellectually active, being alive (see Figure 4).

Merging the two distributions $a(x)$ and $a_a(x)$ as explained, the one of being active at all age is calculated and - knowing that Hypatia was so in 415 - $\Upsilon_a(\xi)$, the probability distribution for the year of Hypatia's birth deduced from this historical record, is obtained in a straightforward manner (see Figure 5).

[7]The probability of dying before the age of eighteen is excluded from the original data, knowing that the subject in question has been intellectually active.

[8]The choice of the first one hundred intellectuals was made to have a random sample: random must also be considered eventual errors and omissions. An attempt at excluding all the cases of violent death was made as well.

Jun 3 2016 09:40:48 PDT
Version 3 - Submitted to MEMOCS

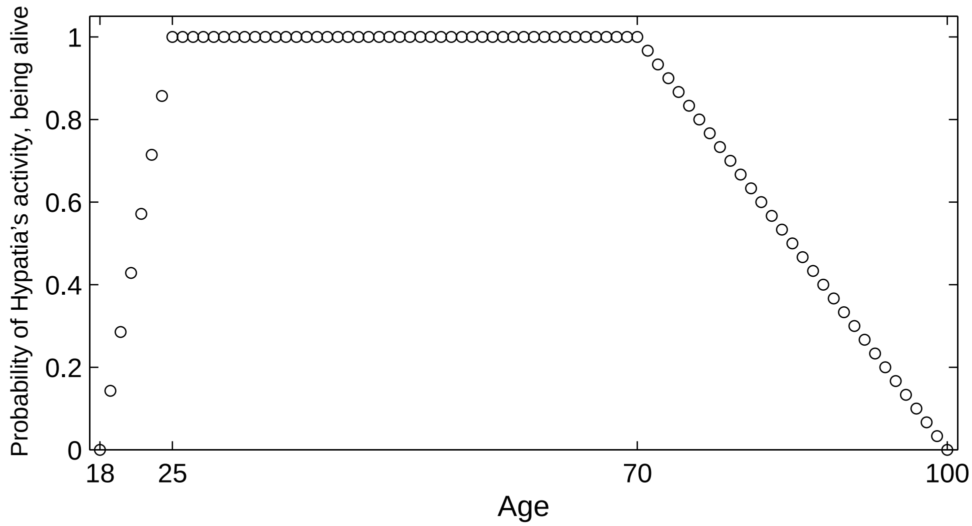

Figure 4: The probability distribution $a_a(x)$.

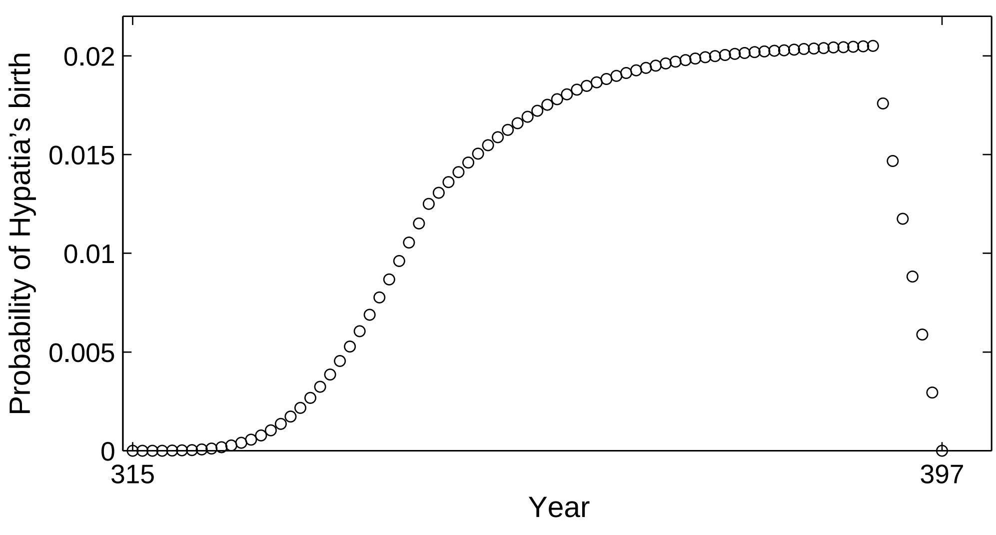

Figure 5: The probability distribution $\Upsilon_a(\xi)$.

3.3 Historical record 3 - Hypatia died *old*

In his **Χρονογραφία** (XIV.12), John Malalas wrote:

> **Κατ᾽ ἐκεῖνον δὲ τὸν καιρὸν παρρησίαν λαβόντες ὑπὸ τοῦ ἐπισκόπου οἱ Ἀλεξανδρεῖς ἔκαυσαν φρυγάνοις αὐθεντήσαντες Ὑπατίαν τὴν περιβόητον φιλόσοφον, περὶ ἧς μεγάλα ἐφέρετο ἦν δὲ παλαιὰ γυνή.**[9]

Based on the statistics for the average lifespan of 'ancient intellectuals' formerly introduced and on the authoritative opinion of historians like Maria Dzielska[10], the distribution $o(x)$ is used as a model for the probability of being defined as an 'old woman' at the various ages, in 415 (see Figure 6).

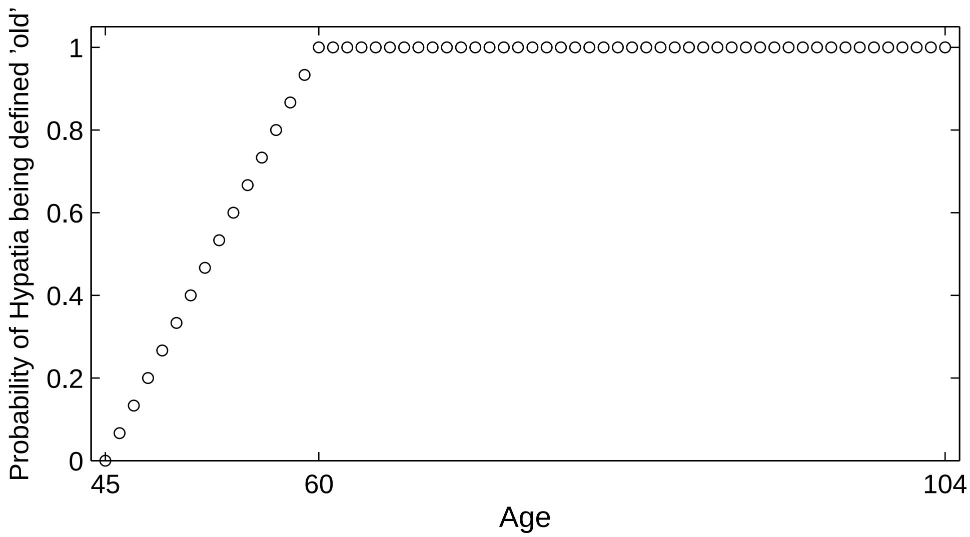

Figure 6: The probability distribution $o(x)$.

$\Upsilon_o(\xi)$, the probability distribution for the year of Hypatia's birth deduced from this historical record, is obtained in a straightforward manner (see Figure 7).

[9] At that time the Alexandrians, given free rein by their bishop, seized and burnt on a pyre of brushwood Hypatia the famous philosopher, who had a great reputation and who was an old woman.

[10] 'John Malalas argues persuasively that at the time of her ghastly death Hypatia was an elderly woman - not twenty-five years old (as Kingsley wants), nor even forty-five, as popularly assumed. Following Malalas, some scholars, including Wolf, correctly argue that Hypatia was born around 355 and was about sixty when she died' (see [4])

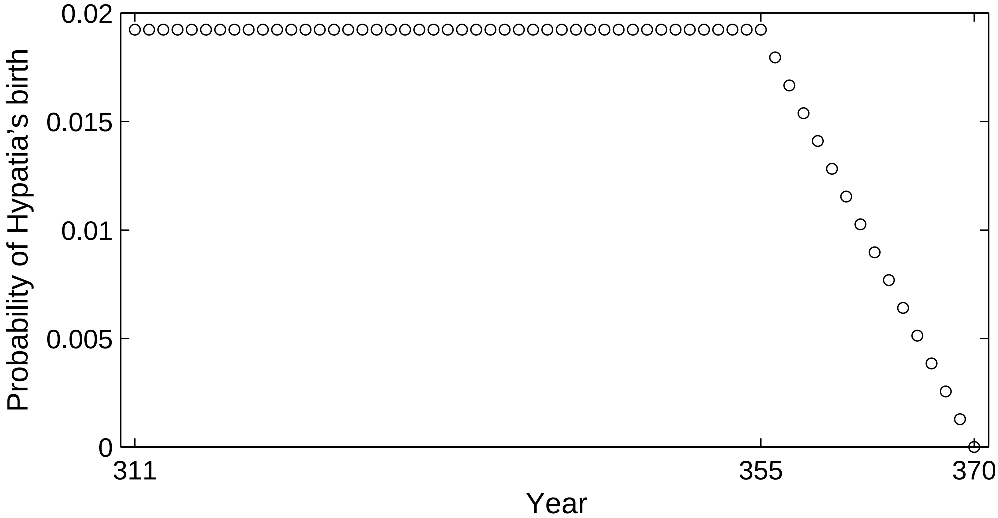

Figure 7: The probability distribution $\Upsilon_o(\xi)$.

3.4 Historical record 4 - Hypatia, *daughter* of Theon

Theon of Alexandria, best known for allowing the transmission of Euclid's Elements to the present day, was Hypatia's father. By knowing his birth year, one could think of deducing a probability distribution for the year of Hypatia's birth: sadly, this is unknown as well. Therefore it is necessary to calculate a probability distribution for the year of Theon's birth first. To serve this purpose, two main historical records come in handy:

- Theon was intellectually active between 364 and 377[11];
- Hypatia overhauled the third book of Theon's *Commentary* on the *Almagest*[12].

It is clear what the effect of the second historical record is: it acts making less probable that Theon became Hypatia's father being 'old' and, at the same time, less probable that he stopped being intellectually active at a young age, being active while her daughter was active as well.
The following *events* are defined:

- F_i, Theon becomes father at the age of i;
- $A_i^{T/I}$, Theon/Hypatia is intellectually active at the age of i;
- C, Theon is able to collaborate with Hypatia (they both are intellectually active);
- $B_k^{T/I}$, Theon/Hypatia begins being intellectually active at the age of k;
- $S_k^{T/I}$, Theon/Hypatia stops being intellectually active at the age of k.

The probability of Theon becoming father at the various ages is, in a first approximation, described by the model distribution $F(x)$ (see Figure 8).

[11] In the *Little Commentary* on Ptolemy's *Handy Tables*, Theon mention some astronomical observations that can be dated with certainty: the two solar eclipses of 15th June and 26th November 364 and an astral conjunction in 377. It is reasonable to assume that he was also active in the interval between these two years.

[12] As referred by Theon in the very same *Commentary*.

Jun 3 2016 09:40:48 PDT
Version 3 - Submitted to MEMOCS

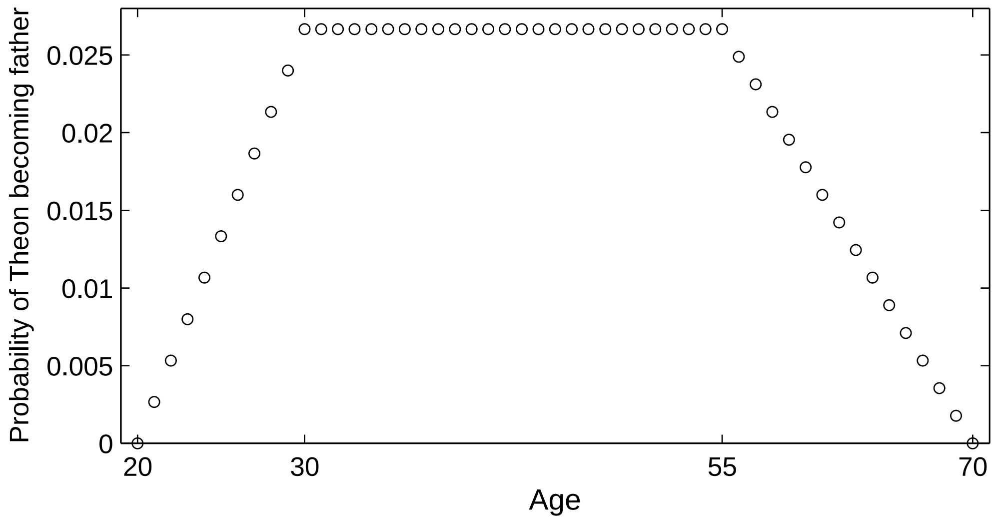

Figure 8: The probability distribution $F(x)$.

The probability distribution of Theon/Hypatia being active at all ages is the one obtained in Subsection 3.2.
The probability of both Theon and Hypatia beginning their intellectual activity at all ages is, in a first approximation, described by the model distribution $B(x)$ (see Figure 9).

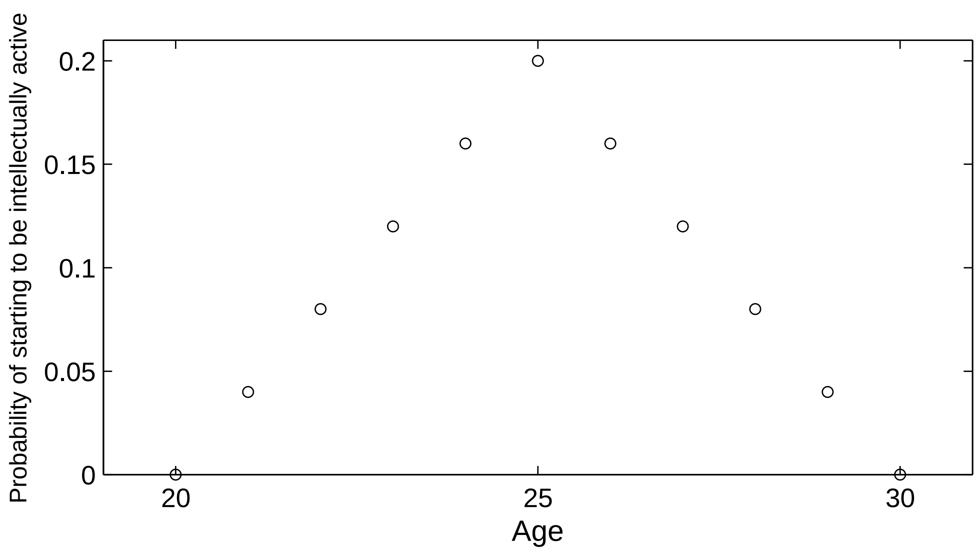

Figure 9: The probability distribution $B(x)$.

The probability of both Theon and Hypatia ending their intellectual activity at all ages is, in a first approximation, described by the model distribution $S(x)$ (see Figure 10).

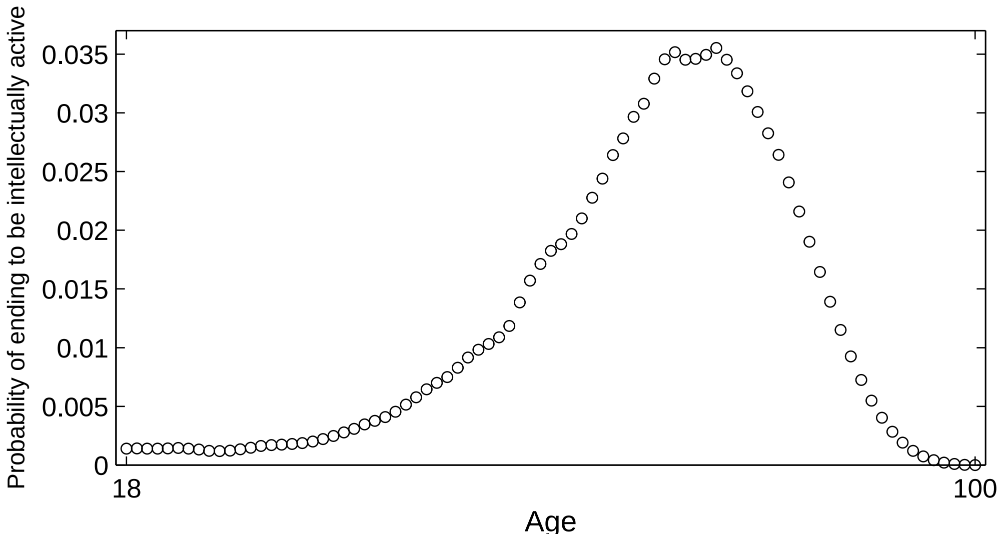

Figure 10: The probability distribution $S(x)$.

The probability of the C event is therefore:

$$P(C) = \sum_i \sum_k P(A^T_{i+k} \bigcap F_i \bigcap B^I_k)$$

By definition of conditional probability:

$$\sum_i \sum_k P(A^T_{i+k} \bigcap F_i \bigcap B^I_k) = \sum_i \sum_k P(A^T_{i+k} \bigcap I^I_k | F_i) \cdot P(F_i)$$

and since the beginning of the active life of Hypatia does not depend on the father's activity, the following simplification can be operated:

$$\sum_i \sum_k P(A^T_{i+k} \bigcap B^I_k | F_i) \cdot P(F_i) = \sum_i \sum_k P(B^I_k) \cdot P(A^T_{i+k} | F_i) \cdot P(F_i)$$

Without committing a large error, it is possible to confuse the probability of being active at the age of $i + k$ having had a daughter at the age of i, $P(A^T_{i+k} | F_i)$, with the one of being active at the age of $i + k$ having been alive at the age of i (V_i[13]), $P(A^T_{i+k} | V_i)$:

$$P(A^T_{i+k} | F_i) \approx P(A^T_{i+k} | V_i) = \frac{P(A^T_{i+k})}{P(V_i)}$$

In the end, the following equation can be written:

$$P(C) = \sum_i \sum_k P(B^I_k) \cdot \frac{P(A^T_{i+k})}{P(V_i)} \cdot P(F_i)$$

[13] V_i is obtained from the above mentioned 1974 Italian males mortality data set.

Based on the idea previously exposed, the next step is to calculate $P(F_i|C)$ and $P(S^T_k|C)$ (and so $P(A^T_i|C) = 1 - \sum_k P(S^T_k|C)$):

$$P(F_i|C) = \frac{P(F_i \bigcap C)}{P(C)} = \frac{\sum_k P(B^I_k) \cdot \frac{P(A^T_{i+k})}{P(V_i)} \cdot P(F_i)}{\sum_i \sum_k P(B^I_k) \cdot \frac{P(A^T_{i+k})}{P(V_i)} \cdot P(F_i)}$$

$$P(S^T_k|C) = \frac{P(S_k \bigcap C)}{P(C)} = \frac{\sum_{\substack{i,j \\ i+j \leq k}} P(S^T_k) \cdot P(F_i) \cdot P(B^I_j)}{\sum_i \sum_j P(B^I_j) \cdot \frac{P(A^T_{i+j})}{P(V_i)} \cdot P(F_i)}$$

$A^T_C(x)$ is the probability distribution of Theon being active at all ages, conditioned to the C event (see Figure 11).

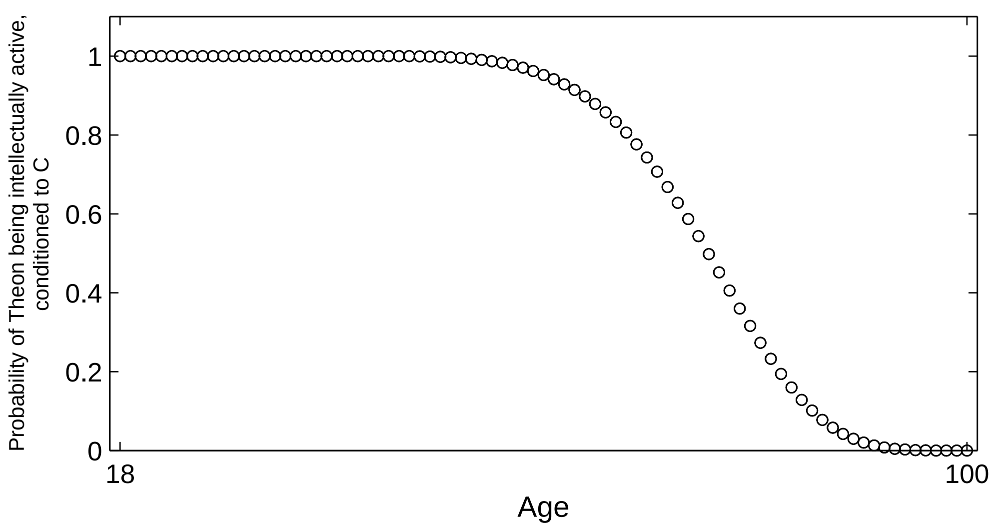

Figure 11: The probability distribution $A^T_C(x)$.

Having in mind the two years in which Theon was surely active (364 and 377), two distributions $364(\xi)$ and $377(\xi)$ for the Theon's year of birth are deduced as previously shown in Subsection 3.2 (see Figure 12). Then, following the instructions introduced in the theoretical section of this work, a single distribution $\Theta(\xi)$ is obtained (see Figure 13).
Finally, in order to calculate $\Upsilon_d(\xi)$, the probability distribution for the year of Hypatia's birth deduced from the historical record *Hypatia daughter of Theon*, the probability of the various events 'between father and daughter there's an age difference of i years' conditioned to the C event must be known. This is indeed the above calculated $P(F_i|C)$, now written as the function $F_C(x)$ (see Figure 14), so that $\Upsilon_d(\xi)$ is calculated in a straightforward manner (see Figure 15).

$$\Upsilon_d(\xi) = \sum_x \Theta(\xi) \cdot F_C(\xi - \xi)$$ [14]

[14] The sum is made on the whole domain of $\Theta(\xi)$

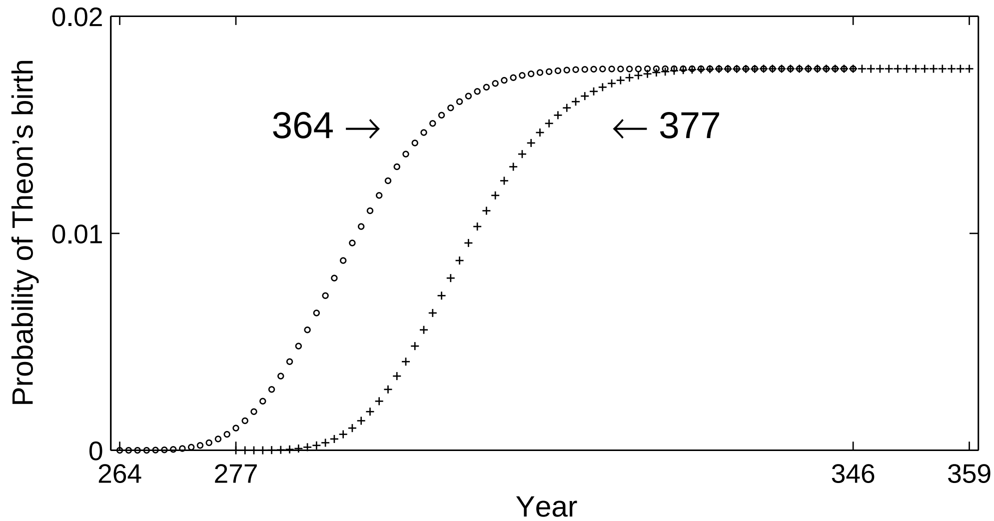

Figure 12: The probability distributions $364(\xi)$ and $377(\xi)$.

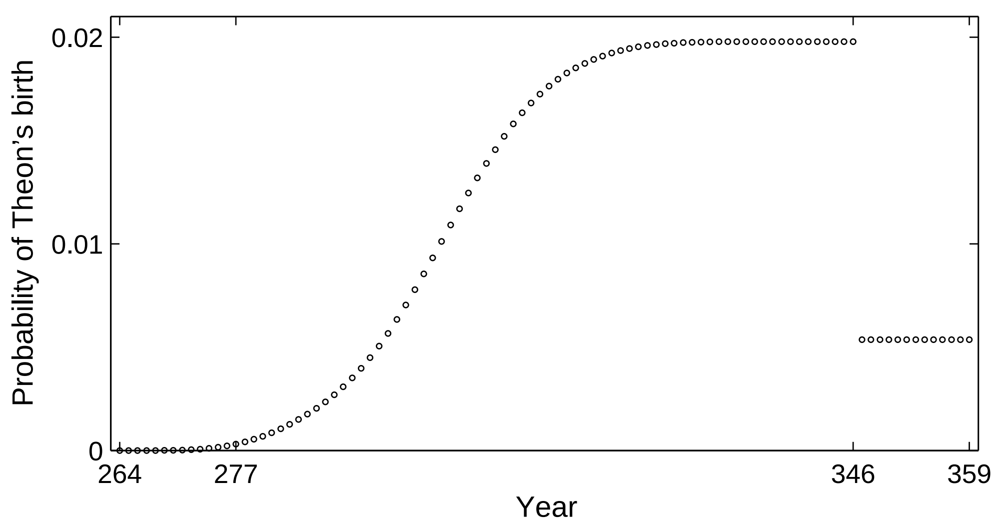

Figure 13: The probability distribution $\Theta(\xi)$.

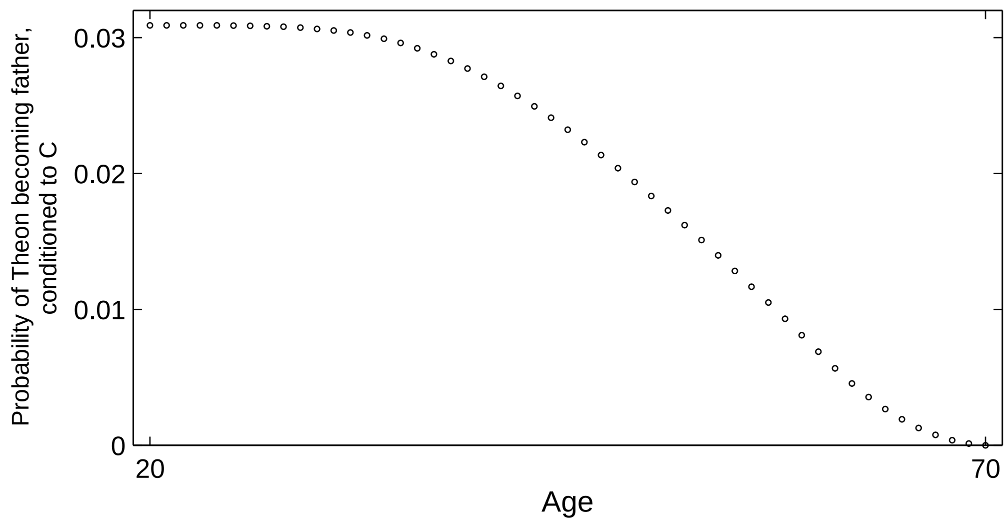

Figure 14: The probability distribution $F_C(x)$.

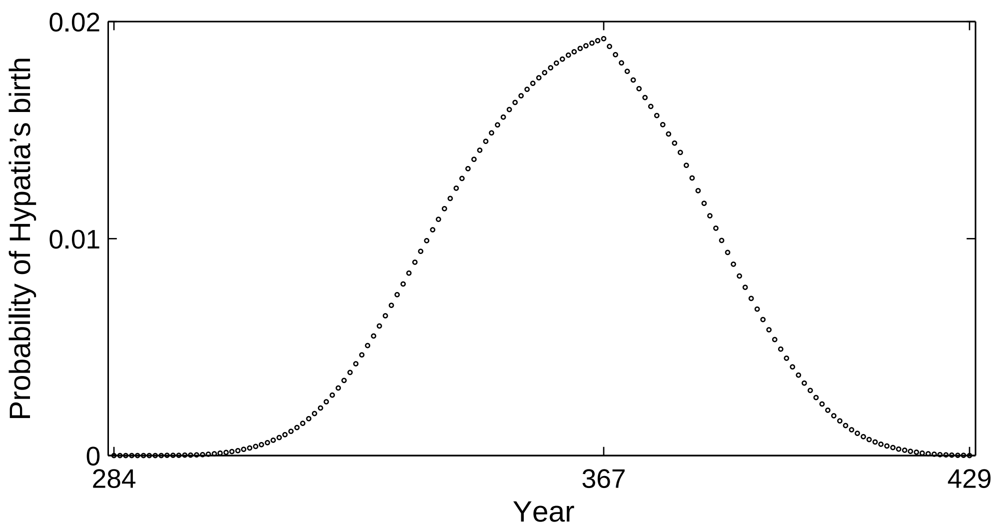

Figure 15: The probability distribution $\Upsilon_d(\xi)$.

Jun 3 2016 09:40:48 PDT
Version 3 - Submitted to MEMOCS

3.5 Historical record 5 - Hypatia, *teacher* of Synesius

Synesius of Cyrene, Neo-Platonic philosopher and bishop of Ptolemais, was a student of Hypatia, as proven by a close correspondence between the two.
For instance, from his deathbed Synesius wrote:

> **Τῇ φιλοσόφῳ.**
> **Κλινοπετὴς ὑπηγόρευσα τὴν ἐπιστολήν, ἣν ὑγιαίνουσα κομίσαιο, μῆτερ καὶ ἀδελφὴ καὶ διδάσκαλε καὶ διὰ πάντων τούτων εὐεργετικὴ καὶ πᾶν ὅ τι τίμιον καὶ πρᾶγμα καὶ ὄνομα.**[15]

The distribution $T(x)$ is introduced as a model to describe the probability of a difference of x years of age between teacher and pupil (see Figure 16).

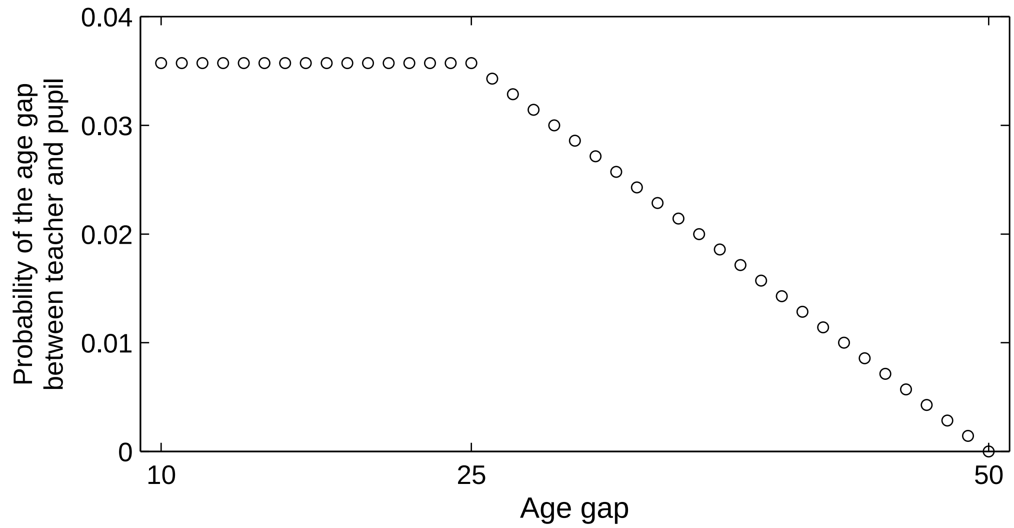

Figure 16: The probability distribution $T(x)$.

$\Upsilon_t(\xi)$, the probability distribution for the year of Hypatia's birth deduced from this historical record, is obtained in a straightforward manner by taking 370 as the year of birth of Synesius[16] (see Figure 17).

[15] To the Philosopher. I am dictating this letter to you from my bed, but may you receive it in good health, mother, sister, teacher, and withal benefactress, and whatsoever is honored in name and deed. (*Incipit* of Letter 16 as in [14])
[16] See, for example, [12].

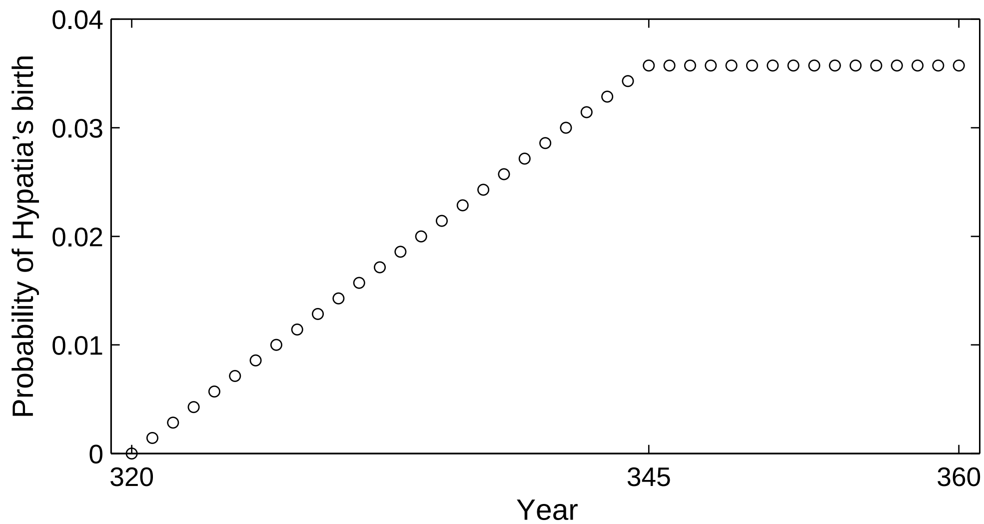

Figure 17: The probability distribution $\Upsilon_t(\xi)$.

3.6 One distribution

Combining the five probability distributions above deduced for the year of Hypatia's birth, one final distribution - $\Upsilon(\xi)$ - can be obtained following the rules introduced in the theoretical section of this work. $\Upsilon(\xi)$ can be compared to the one given by the simple arithmetic average of the various distributions resulting from every possible combination of testimonials considered true at the same time - $\Upsilon_A(\xi)$ (see Figure 18).

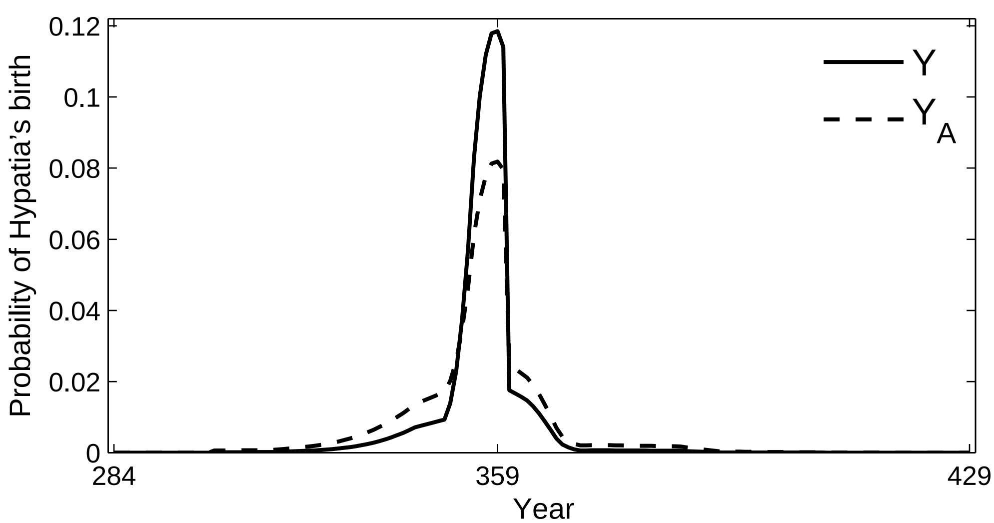

Figure 18: The probability distributions $\Upsilon(\xi)$ and $\Upsilon_A(\xi)$.

Therefore, the most probable year for the birth of Hypatia is 355 ($\sim$ 14.5%) with a total probability of the interval [350, 360] of about 90%.

4 Conclusions

The probabilistic dating model set out in this work, structured in three different steps, could be summarized making use of a culinary analogy. The first step being represented by the collection of enough raw ingredients (testimonials) to be refined, 'cooked' in the second step (turned into probability distributions) and - finally, in the third step - put together following a recipe (the one provided in Section 2), so that they blend well (as a single probability distribution).
Its application to the particular case of Hypatia proved to be very satisfactory in that the final probability distribution shows a marked peak, making it possible to give a quite accurate datum. More specifically, the datum so obtained contradicts the prevalent opinion (cf. above Section 3) but is in perfect agreement with a minority school of thought developed around the issue. In this regard, we have already mentioned the authoritative opinion of Maria Dzielska, which deems that Hypatia died about 60 years old, having - consequently - born around the year 355. A similar opinion is expressed by Deakin ([2], 52).
Future applications appear to be far-reaching, as the method could serve not only in cases strictly analogous to the one here presented but also in dating any event provided with a sufficient number of testimonials able to be turned into probability distributions.

References

[1] A. E. Abbas, *A Kullback-Leibler view of linear and log-linear pools*, Decision Analysis *6* (2009), pp. 23–37.

[2] M. A. B. Deakin, *Hypatia of Alexandria*, Prometheus Books, New York, 2007.

[3] *Philosophy of probability* (in *Philosophical Studies Series: vol. 56*), J. Dubucs, ed., Springer Science+Business Media B.V., Dordrecht, 2013 (2nd ed.).

[4] M. Dzielska, *Hypatia of Alexandria, adapted from the unpublished manuscript*, F. Lyra (Transl.), Harvard University Press, Cambridge, 1995.

[5] B. de Finetti, *Sul significato soggettivo della probabilità*, Fundamenta Mathematicae *T. XVII* (1931), pp. 298–329.

[6] C. Genest, K. J. McConway, *Allocating the weights in the linear opinion pool*, Journal of Forecasting *9* (1990), pp. 53–73.

[7] C. Genest, J. V. Zidek, *Combining probability distributions: a critique and annotated bibliografy*, Statistical Science *1* (1986), pp. 114–148.

[8] C. Genest, C. G. Wagner, *Further evidence against independence preservation in expert judgement synthesis*, Aequationes Mathematicae *32* (1987), pp. 74–86.

[9] S. Kullback, R. A. Leibler, *On information and sufficiency*, Annals of Mathematical Statistics *22* (1951), pp. 79-86.

[10] K. Lehrer, C. G. Wagner, *Probability amalgamation and the independence issue: a reply to Laddaga*, Synthese *55* (1983), pp. 339–346.

[11] K. J. McConway, *Marginalization and linear opinion pools*, Journal of the American Statistical Association *76* (1981), pp. 410–414.

[12] *The Oxford Classical Dictionary*, S. Hornblower, A. Spawforth and E. Eidinow, eds., 4th ed., online version available at http://www.oxfordreference.com/view/10.1093/acref/9780199545568.001.0001/acref-9780199545568

[13] C. S. Peirce, *On the Logic of drawing History from Ancient Documents especially from Testimonies*, in *The Essential Peirce: Selected Philosophical Writings*, Peirce Edition Project, eds., Vol. 2, Indiana University Press, Bloomington, 1998.

[14] *The Letters of Synesius of Cyrene*, A. FitzGerald (Transl.), Oxford University Press, Oxford, 1926.

[15] *A Select Library of Nicene and Post-Nicene Fathers of the Christian Church. Second series*, P. Schaff and H. Wace, eds., Vol. 2, The Christian literature company, New York, 1890.

Accius, Lucius	[170-*c*.86 BC] ~84 years
Adrianus (Hadrianus), of Tyre	[*c.* AD 113-93] ~80 years
Aelian (Claudius Aelianus)	[AD 165/170-230/5] ~65 years
Aeschines	[*c*.397-*c*.322 BC] ~65 years
Aeschylus	[524/5-456/5 BC] ~70 years
Agathocles (2), (of Cyzicus)	[*c*.275/265-*c*.200/190 BC] ~75 years
Alexander of Tralles	[AD 525-605] 80 years
Alexis	[*c*.375-*c*.275 BC] ~100 years
Ammianus Marcellinus	[*c.* AD 330-395] ~65 years
Anaxagoras	[*probably* 500–428 BC] ~72 years
Anaximenes (2), of Lampsacus	[*c*.380-320 BC] ~60 years
Andocides	[*c*.440–*c*.390 BC] ~50 years
Androtion	[*c*.410–340 BC] ~70 years
Antiphon	[*c*.480-411 BC] ~69 years
Apollonius of Citium	[*c*.90-15 BC?] ~75 years
Arcesilaus	[316/5-242/1 BC] ~74 years
Aristarchus of Samothrace	[*c*.216-144 BC] ~72 years
Aristophanes of Byzantium	[*c*.257-180 BC] ~77 years
Aristotle	[384-322 BC] 62 years
Arius	[*c.* AD 260–336] ~76 years
Arrian (Lucius Flavius Arrianus)	[AD *c*.86-160] ~74 years
Aspasius	[*c.* AD 100-50] ~50 years
Athanasius	[*c.* AD 295–373] ~78 years
Atticus	[*c.* AD 150-200] ~50 years
Augustine, St	[AD 354-430] 76 years
Bacchius, of Tanagra	[275–200 BC?] ~75 years
Bacchylides	[*c*.520-450 BC] ~70 years
Basil of Caesarea	[*c*.AD 330–79] ~49 years
Bion of Borysthenes	[*c*.335-*c*.245 BC] ~90 years
Carneades	[214/3–129/8 BC] ~85 years
Cassius (1)	[31 BC–AD 37] 68 years
Cassius Longinus	[*c*.AD 213–73] ~60 years
Cato (Censorius)	[234–149 BC] 85 years
Chrysippus, of Soli	[*c*.280–207 BC] ~73 years
Chrysostom, John	[*c.* AD 354–407] ~53 years
Cinesias	[*c*.450–390 BC] ~60 years
Claudius Atticus Herodes (2), Tiberius	[*c*.AD 101–77] ~76 years
Cleanthes of Assos	[331–232 BC] 99 years
Clitomachus	[187/6–110/9 BC] ~77 years
Colotes (RE 1), of Lampsacus	[*c*.325–260 BC] ~65 years
Cornelius (RE 157) Fronto, Marcus	[*c.* AD 95–*c*.166] ~71 years
Crantor of Soli *in Cilicia*	[*c*.335–275 BC] ~60 years
Crates (2)	[*c*.368/365–288/285 BC] ~80 years
Demades	[*c*.380-319 BC] ~61 years
Demochares	[*c*.360-275 BC] ~85 years
Democritus, (of Abdera)	[*c*.460-370 BC] ~90 years
Demosthenes (2)	[384-322 BC] 62 years
Dinarchus	[*c*.360-*c*.290 BC] ~70 years
Dio Cocceianus	[*c*.40/50–110/120 BC] ~70 years
Diodorus (3) of Agyrium, *Sicily*	[*c*.90-30 BC] ~60 years
Diogenes (3) (of Babylon)	[*c*.240–152 BC] ~88 years
Diogenes (2), the Cynic	[*c*.412/403–*c*.324/321 BC] ~85 years
Duris	[*c*.340–*c*.260 BC] ~80 years
Empedocles	[*c*.492–432 BC] ~60 years
Ennius, Quintus	[239–169 BC] 70 years
Ennodius, Magnus Felix	[AD 473/4 –521] ~48 years
Ephorus, of Cyme	[*c*.405–330 BC] ~75 years
Epicurus	[341–270 BC] 71 years
Epiphanius	[*c.* AD 315–403] ~88 years
Erasistratus	[about 315–240 BC] ~75 years
Eratosthenes, of Cyrene	[*c*.285–194 BC] ~91 years
Eubulus (1)	[*c*.405–*c*.335 BC] ~70 years
Euclides (1), of Megara	[*c*.450–380 BC] ~70 years
Euripides	[probably 480s-407/6 BC] ~78 years

Eusebius, of Caesarea	[*c.* AD 260–339] ∼79 years
Evagrius Scholasticus	[*c.* AD 535–*c.*600] ∼65 years
Favorinus	[*c.* AD 85-155] ∼70 years
Fenestella	[52 BC–AD 19 or 35 BC–AD 36] 71 years
Galen, of Pergamum	[AD 129–216] 87 years
Gorgias (1), of Leontini	[*c.* 485–*c.*380 BC] ∼105 years
Gregory (2) of Nazianzus	[AD 329–89] 60 years
Gregory (3) of Nyssa	[*c.* AD 330-95] ∼65 years
Gregory (4) Thaumaturgus	[*c.* AD 213–*c.*275] ∼62 years
Hecataeus (2), of Abdera	[*c.* 360–290 BC] ∼70 years
Hegesippus (1)	[*c.* 390–*c.* 325 BC] ∼65 years
Hellanicus (1) of Lesbos	[*c.* 480–395 BC] ∼85 years
Hellanicus (2)	[*c.* 230/20-160/50 BC] ∼70 years
Herophilus, of Chalcedon	[*c.* 330-260 BC] ∼70 years
Hieronymus (2), of Rhodes	[*c.* 290-230 BC] ∼60 years
Himerius	[*c.* AD 310-*c.*390] ∼80 years
Horace (Quintus Horatius Flaccus)	[65-8 BC] 57 years
Idomeneus (2)	[*c.*325-*c.*270 BC] ∼55 years
Irenaeus	[*c.* AD 130-*c.*202] ∼72 years
Isaeus (1)	[*c.* 420-340s BC] ∼75 years
Isocrates	[436-338 BC] 98 years
Ister	[*c.*250-200 BC] ∼50 years
Jerome (Eusebius Hieronymus)	[*c.* AD 347-420] ∼73 years
Laberius, Decimus	[*c.*106–43? BC] ∼63 years
Libanius	[AD 314–*c.*393] ∼63 years
Livius Andronicus, Lucius	[*c.*280/70-200 BC] ∼75 years
Livy (Titus Livius)	[59 BC-AD 17 or 64 BC-AD 12] 76 years
Lucilius (1), Gaius	[probably 180 -102/1 BC] ∼75 years
Lucretius (Titus Lucretius Carus)	[*c.*94-55 or 51? BC] ∼41 years
Lyco	[*c.*300/298-226/4 BC] ∼74 years
Lycurgus (3)	[*c.*390-*c.*325/4 BC] ∼65 years
Lydus	[AD 490-*c.*560] ∼70 years
Lysias	[459/8-*c.*380 or *c.*445-*c.*380] ∼72 years
Malalas	[*c.* AD 480-*c.*570] ∼90 years
Mantias	[*c.*165-85 BC] ∼80 years
Megasthenes	[*c.*350-290 BC] ∼60 years
AVERAGE LIFESPAN OF THE 'ANCIENT INTELLECTUAL'	**∼71.7 years**

Table 1: Average lifespan of the 'ancient intellectual'