LIVING to 100
SOCIETY OF ACTUARIES
INTERNATIONAL SYMPOSIUM

Extreme Value Analysis of Mortality at the Oldest Ages: A Case Study Based on Individual Ages at Death

Samuel Gbari
Institut de statistique, biostatistique et sciences actuarielles (ISBA)
Université Catholique de Louvain (UCL)
Louvain-la-Neuve, Belgium
Michel Poulain
Institute for the Analysis of Change in Historical
and Contemporary Societies (IACCHOS)
Université Catholique de Louvain (UCL)
Louvain-la-Neuve, Belgium;
Estonian Institute for Population Studies
Tallinn University
Tallinn, Estonia
Luc Dal
Institut de démographie (DEMO)
Universit_e Catholique de Louvain (UCL)
Louvain-la-Neuve, Belgium
Michel Denuit
Institut de statistique, biostatistique et sciences actuarielles (ISBA)
Université Catholique de Louvain (UCL)
Louvain-la-Neuve, Belgium

Presented at the Living to 100 Symposium
Orlando, Fla.
January 4–6, 2017

Abstract

In this paper, the force of mortality at the oldest ages is studied using the statistical tools from extreme value theory. A unique database recording all individual ages at death above 95 for extinct cohorts born in Belgium between 1886 and 1904 is used to illustrate the relevance of the proposed approach. No leveling off in the force of mortality at the oldest ages is found, and the analysis supports the existence of an upper limit to human lifetime for these cohorts. Therefore, assuming that the force of mortality becomes ultimately constant—i.e., that the remaining lifetime tends to the negative exponential distribution as the attained age grows—is a conservative strategy for managing life annuities.

1 Introduction and Motivation

Insurers need reliable estimates of mortality at oldest ages to price life annuities and reverse mortgages, for instance. Gompertz, Makeham, logistic and other parametric models have often been used by actuaries to fit mortality statistics. The exponential increase in the force of mortality postulated by the Gompertz model is generally appropriate for adults and early old ages, but the mortality increases tend to slow down at older ages, typically around 85, so that this model fails to describe the end of the life table. Alternative models have been proposed to account for the mortality deceleration observed at older ages in industrialized countries. The plateau that appears at these advanced ages is attributed to the earlier selection of the more robust individuals in heterogeneous cohorts. This empirical finding supports the logistic model obtained by including a gamma distributed frailty coefficient accounting for the heterogeneity in the Gompertz model (Manton et al. 1986).

When the main interest is in mortality at the oldest ages, inference is conducted in an area of the sample where there is a very small amount of data. Moreover, extrapolation beyond the range of the data is often desirable, a procedure known as closure of the life table. In such a case, it is essential to have a good model for the longest lifetimes. Extreme value theory deals with statistical problems concerning the far tail of the probability distribution, being supported by strong theoretical arguments, and thus appears as the natural candidate for analyzing mortality at the oldest ages.

Models for extreme values have already been used to describe individual lifetimes but concentrating on minima. Brillinger (1961) was among the first authors who draw the attention of the actuarial community to the use of extreme value techniques in the analysis of mortality. Specifically, he pointed out that the Gompertz law naturally appears to describe lifetimes if the human body is considered as a series system made of a large number of components with independent and identically distributed failure times. The death then occurs when the first component fails that is, at the minimum of these failure times. The minimum of these failure times asymptotically obeys the Gompertz model.

By construction, the approximation of the lifetime distribution by an extreme value distribution such as the Gompertz law can be expected to hold only for small or moderate ages, but not for high ages, as pointed out by Aarssen and de Haan (1994). For high ages, another approach is thus needed, such as the one proposed in this paper. Notice that some authors argue that the Gompertz model may also apply to higher ages. For instance, Gavrilov and Gavrilova (2011) concluded from a study of several extinct cohorts that mortality at advanced ages obeys the Gompertz model up to ages 102 to 105, without noticeable deceleration. However, this study is subject to debate. See, e.g., Ouellette and Bourbeau (2014) for an opposite view.

The present paper adopts a distribution-free approach to analyze mortality at the oldest ages; only some weak regularity of the tail of the distribution function is assumed. This method is acceptable from both the theoretical and practical viewpoints. To this end, we recall the basic features of extreme value theory (EVT) that gives the theory for describing extremes of random phenomena. EVT, and its close link to limiting residual life distributions, offers a unified approach to the modeling of the right tail of a lifetime distribution. This method is not solely based on the data at hand but includes a probabilistic argument concerning the behavior of the extreme sample values. EVT has already been successfully

applied to solve various non-life-insurance problems; see, e.g., Cebrian et al. (2003) and the references therein. In this paper, we demonstrate that it also offers the appropriate framework for dealing with the highest ages at death. The analysis of residual lifetimes at high ages is in line with the peaks-over-threshold (POT) method used to study tail behavior exhibited by non-life-insurance losses. POT is based on the convergence of the distribution of exceedances over a threshold to the generalized Pareto (GP) distribution when the threshold increases. Transposed to the life insurance setting, this approach is justified by the convergence of the remaining lifetime distribution at high ages to the GP distribution.

The present paper is not the first attempt to use EVT techniques to model mortality at the oldest ages. EVT was even initially developed to deal with such problems, as can be seen from Gumbel (1937), Balkema and de Haan (1974) and Aarssen and de Haan (1994), for instance. Our approach is closely related to the pioneering study by Watts et al. (2006), where the usefulness of EVT techniques in life insurance data analysis is clearly established. Actuaries have also developed threshold life tables where the extrapolation of mortality to older ages (above the threshold) is done using EVT. We refer the readers to the works by Han (2005), Li et al. (2008, 2010) and Bravo et al. (2008, 2012). Compared with these previous studies, the originality of our paper relies on several facts:

- We work here with reliable individual exact ages at death, not with annual aggregated mortality data.

- We resort to advanced statistical tools to explore the trajectory of mortality near the end of the lifetimes, questioning the choice of the appropriate threshold in the POT approach above 95 years old, as well as the estimation method for the tail index.

- We conduct the analysis using cohort mortality statistics, and not period data, as the theoretical arguments supporting the use of EVT techniques are fulfilled in this case, as will become clear in the next sections.

EVT techniques have also been considered in other disciplines such as demography, as a natural tool to analyze mortality at the oldest ages. For instance, Hanayama and Sibuya (2015) estimated the ultimate age in the Japanese population by means of EVT techniques, based on aggregated mortality statistics for Japanese centenarians in the Human Mortality Database (HMD). The approximate linearity of the mean residual life function led to the use of the GP distribution to describe survival above a high age. The GP parameters are then estimated by binomial regression on yearly death counts, and a significantly negative estimated tail index resulted for male cohorts born after 1883 and female cohorts born after 1867. Despite obvious similarities, there are also important differences between this study and the one conducted in the present paper. First, we work with individual data. These data consists in exact ages at death for all residents of Belgium who were born there and died there at age 95 or older. Also, these authors arbitrarily selected the age range where mortality obeys the GP model, starting at age 100, and only considered maximum likelihood estimation techniques. We discuss these two choices here, and we stress their great impact on the final estimates.

Several authors, including Gampe (2010), have suggested that age-specific death probabilities may level off at some point above 100 years, rather than continuing to increase with

age. This proposal has been supported by the analysis of the mortality of supercentenarians (aged 110 and over) from many countries included in the International Database of Longevity (IDL). Notice that if the force of mortality tends to flatten at oldest ages, then remaining lifetimes become ultimately negative exponential. The analysis conducted in the present paper suggests that the remaining lifetimes display a lighter tail than the negative exponential and that there is an upper bound on the human lifetime, which actuaries call ultimate age. Our results are in line with those obtained by Poulain et al. (2001), who concluded that the annual death probabilities between ages 100 and 108 rise from 0.35 to 0.50 for Belgian females and from 0.42 to 0.55 for Belgian males. These values are similar to those obtained by Kannisto (1994) from a significantly larger study population. Importantly, we work here with individual exact ages at death and never aggregate them over intervals, as this may bias the analysis.

Notice that actuaries are more interested in closing the life table in an appropriate manner than in providing a final answer to the question of maximum life span. Indeed, the problem posed by the mortality pattern at oldest ages, with a still increasing, or plateauing, or even decreasing force of mortality after a maximum level has been reached, goes far beyond the actuarial expertise, as almost no data are available at ages above 115. Actuaries nevertheless need an appropriate model, supported by empirical evidence, to close the life table so that actuarial calculations can be performed, typically in life annuity or reverse mortgage portfolios. To this end, this paper uses a unique database recording the ages at death for 19 extinct cohorts born in Belgium between 1886 and 1904; restricting the analysis to these 19 extinct cohorts ensures that we deal with complete data. The survival of Belgian centenarians has been studied by Poulain et al. (2001). As explained by these authors, the National Register, a centralized, fully computerized population register for the whole population of Belgium that was established by law in 1985 (see Poulain 2010), minimizes the risk of errors in recorded ages at death. This is why we use data extracted from the Belgian National Register to conduct the present study. As ages are recorded up to the day, we can model lifetimes on a continuous scale and apply EVT techniques to complete data that are not censored by intervals.

The remainder of this paper is organized as follows. We start in Section 2 by describing the data used in the present study. Section 3 gives a brief, user-friendly description of EVT tools that are necessary to analyze mortality at the oldest ages. The definition of the generalized extreme value distribution is recalled, providing actuaries with a limit distribution for maxima subject to an appropriate standardization. Its connection to the generalized Pareto distribution is established by taking the limit of the conditional distribution for the exceedances over a threshold. The presentation is made in terms of demographic indicators, to make it more appealing for actuaries working in life insurance. In particular, the close links existing between EVT and limiting residual lifetime distributions are made explicit. Selecting the threshold above which the exceedances obey the generalized Pareto distribution is the key first step in actuarial application based on EVT. The analysis conducted here is based on tools proposed by Pickands (1975), Reiss and Thomas (1997) and McDonald et al. (2011). Section 4 is devoted to several applications of the model, such as the estimation of ultimate age and the prediction of the highest age at death recorded among a group of individuals. The final Section 5 summarizes the main findings of the paper and briefly concludes.

2 Description of the Data Set

2.1 Data Source and Validation Process

As mentioned in the introduction to this paper, a national population register serves as the centralizing database in Belgium and provides official population figures. Statistics on births and deaths are available from this register according to basic demographic characteristics (e.g., age, gender, marital status). We have at our disposal the ages at death for every individual who died at age 95 or older, plus a list of survivors aged 95 and older at the end of the observation period. There are 46,666 observations relating to individuals born in Belgium who died after age 95 since 1981. The inviduals in the database are 22 percent males and 78 percent females. For each individual, we know the exact birth and death dates.

Data have been validated in a number of ways. For more details, we refer to Poulain et al. (2001), so that we can be confident in the reliability of the ages at death recorded in the database. In particular, all individuals who were not born in Belgium have been excluded from the data set, so that migrations can be neglected (as international mobility is almost nonexistent for native Belgian citizens above age 95), and date of birth is recorded from official certificates (and not reported by individuals). Notice that no exposures to risk need to be included in the analysis, as our method uses individual observations (and not aggregate ones—by attained age, for instance). Also, these individual data are known to be reliable, which is not always the case for aggregate mortality statistics at old ages.

Finally, in population theory, the complete cohort approach consists of observing a group of people born at the same time until all of them pass away, but this assumption is often difficult to validate in practice. Fortunately, in the Belgian case, the very scrupulous age validation performed allows us to consider the complete extinction of the different cohorts under study, as far as all newborns in Belgium who died in Belgium are concerned. Few belgian citizens born in Belgium died abroad above 95 years of age. Intense international cooperation made it possible to identify a limited number of centenarians born in Belgium who died abroad. For reliability and coverage reasons, this study did not consider these cases. This explains why the complete cohort approach has been adopted in the present study.

2.2 Descriptive Statistics

Basic descriptive statistics of the data sets are given in Table 2.1. The average age at death for females is greater than for males, as expected. However, the differences are rather modest, compared with the differences observed for the total lifetime from birth, except for the much higher number of females reaching age 95 compared with males. The highest ages at death observed in Belgium for these cohorts are 112.58 for females and 111.47 for males. Recall that Jeanne Calment died at the age of 122 years in 1997. After this record, only a single person, Sarah Knauss, has lived for 119 years (she died in 1993). Since then, three women lived more than 117 years, and three others are still alive at 116 years old.

Cohort-specific data are displayed in Table 2.2 for both females and males. We can read there the initial number of individuals included in the analysis, for each extinct cohort born between 1886 and 1904 (that is, the number L_{95} of individuals born in calender year

Table 2.1: Basic Descriptive Statistics for Observed Ages at Death Above 95, Cohorts 1886 to 1904

	Males	**Females**
Number of observations	10,050	36,616
Mean	97.35	97.75
Std. dev.	2.05	2.37
1st quartile	95.77	95.91
Median	96.79	97.12
3rd quartile	98.36	98.99
Maximum	111.47	112.58

$t \in \{1886, \ldots, 1904\}$ still alive at age 95), as well as the age at death of the last survivor for each cohort. These values are given separately for males and females. We can see there that L_{95} increases as time passes, as expected. Also, the number of individuals entering the study is larger for females than for males. The maxima m are relatively stable over cohorts and are higher for females than for males: For each cohort, the last survivor was a female. In addition, there is no visible trend in the shifted life expectancy at 95. It even remains stable around 97 for all male and female cohorts. The observation that e_{95} did not improve across the 19 birth cohorts is quite remarkable. This seems to suggest no potential gain of extreme longevity through generations at these advanced ages. We address this issue using a formal test in Section 3.

Notice that the value of L_{95} increases for successive cohorts, due to an increase in the number of births (which for boys starts roughly at 90,000 in 1886 and reaches about 100,000 in 1904, and for girls starts at about 85,000 and reaches about 95,000) and an improvement of survival up to age 95 for both males and females. Nevertheless, the life expectancy at 95 did not show any improvement; the different cohorts show similar figures, and no trend emerged (see Table 2.2). The method adopted is not appropriate for using a stationary population. In fact, we observe a 10 percent increase in the number of births, but this does not affect our results, as the survival above 95 is similar for all cohorts (see Section 3.5 for a formal test). We agree that a problem could have emerged if the number of births had not increased substantially and the survival above 95 had improved, which is not the case here.

Box plots are provided in Figure 2.1, and Figure 2.2 displays the corresponding histograms. For both genders, box plots reveal ages at death above median and third quartile that are spread on a wider range of values. The same conclusion is reached on the basis of the histogram, where we see that the right tail expands on old ages, whereas the majority of deaths are concentrated before age 100. This suggests that we deal with a skewed distribution with a significant proportion of values in the right tail that might be considered extreme.

3 Methodology

Now that the data basis has been presented, let us perform a statistical analysis using the tools from EVT.

Table 2.2: Initial Size for Each Cohort (L_{95}), Observed Highest Age at Death (m) and Mean Age at Death ($95 + \hat{e}_{95}$), Males and Females

	Males			Females		
t	L_{95}	m	$95 + \hat{e}_{95}$	L_{95}	m	$95 + \hat{e}_{95}$
1886	359	108.17	97.35	1045	107.78	97.70
1887	418	105.13	97.41	1130	110.45	97.64
1888	412	106.33	97.20	1242	110.32	97.68
1889	462	105.58	97.53	1208	110.16	97.69
1890	425	107.70	97.50	1311	112.58	97.60
1891	428	105.81	97.35	1368	109.72	97.70
1892	444	105.44	97.29	1510	110.89	97.72
1893	492	110.29	97.46	1544	107.75	97.63
1894	498	106.19	97.22	1760	107.41	97.71
1895	526	106.62	97.24	1852	109.38	97.86
1896	545	106.27	97.43	1894	109.79	97.85
1897	568	106.43	97.33	2009	109.85	97.72
1898	572	105.74	97.26	2149	110.89	97.71
1899	630	106.88	97.35	2301	111.60	97.72
1900	576	111.47	97.27	2460	111.70	97.78
1901	635	103.77	97.23	2787	110.36	97.89
1902	632	106.79	97.42	2829	112.36	97.82
1903	669	104.71	97.46	2980	109.96	97.81
1904	759	106.15	97.32	3237	110.18	97.74

Figure 2.1: Box Plots for the Observed Ages at Death Above 95, Cohorts 1886 to 1904 for Males (Left) and Females (Right)

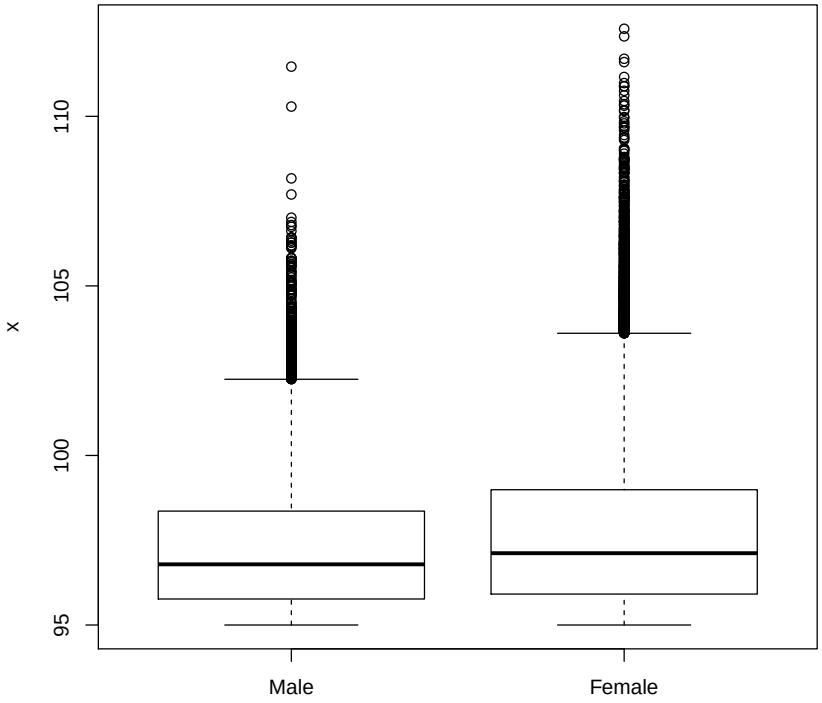
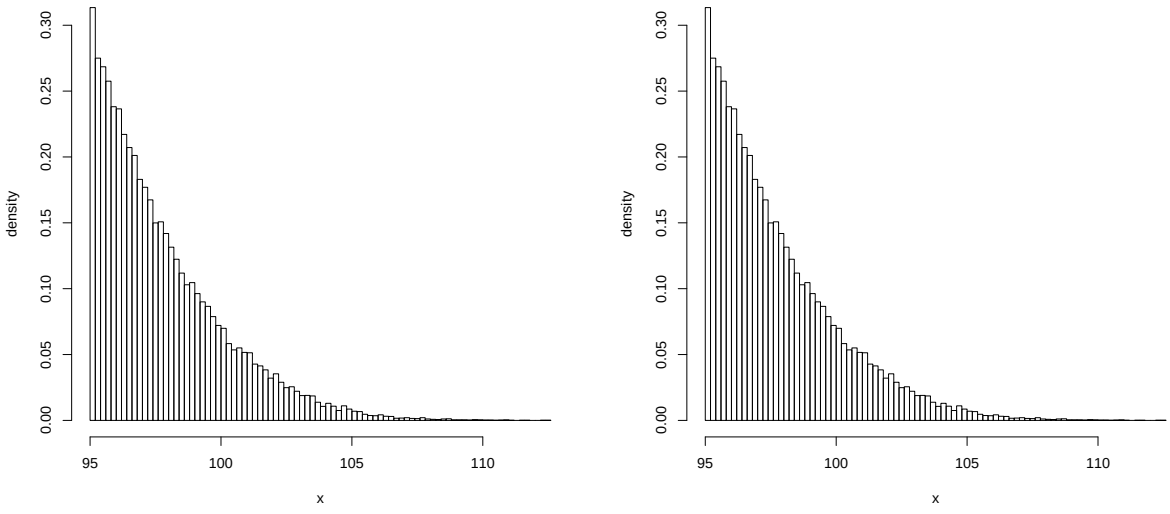


Figure 2.2: Histograms for the Observed Ages at Death Above 95 for Males (Left) and Females (Right), Cohorts 1886 to 1904



3.1 Extreme Value Theory (EVT) for Mortality at the Oldest Ages

Consider a sequence of independent individual lifetimes $T_1, T_2, T_3, \ldots$ with common distribution function

$$ {}_xq_0 = \mathrm{P}[T_i \leq x], \;\; x \geq 0, $$

for $i = 1, \ldots, n$, satisfying ${}_0q_0 = 0$. Here, T_i represents the total lifetime, from birth to death. It may also represent the remaining lifetime after a given initial age α (equal to 95 in our study), with common distribution function ${}_xq_\alpha$.

Define a sequence of maxima $M_n = \max\{T_1, T_2, \ldots, T_n\}$. In words, M_n represents the oldest age at death observed in a homogeneous group of n individuals subject to the same life table $x \mapsto {}_xq_0$. EVT studies the asymptotic behavior of M_n when n tends to infinity and provides results analogous to the central limit theorem for maxima (rather than sums), provided some mild technical conditions on $x \mapsto {}_xq_0$ are fulfilled. Of course, without further restriction, M_n obviously approaches the upper limit of the support

$$ \omega = \sup\{x \geq 0 | {}_xq_0 < 1\}, \text{ possibly infinite.} $$

This is easily seen from

$$ \mathrm{P}[M_n \leq x] = \left({}_xq_0\right)^n \to \left\{ \begin{array}{l} 0 \text{ if } x < \omega \\ 1 \text{ if } x \geq \omega \end{array} \right. \quad \text{as } n \to \infty. $$

Once M_n is appropriately centered and normalized, however, it may converge to some specific limit distribution. Precisely, if there exist sequences of real numbers $a_n > 0$ and $b_n \in \mathbb{R}$ such that the normalized sequence $(M_n - b_n)/a_n$ converges in distribution to H, i.e.,

$$ \lim_{n \to +\infty} \mathrm{P}\left[\frac{M_n - b_n}{a_n} \leq x\right] = \lim_{n \to +\infty} \left({}_{a_n x + b_n}q_0\right)^n = H(x) \tag{3.1} $$

for all points of continuity of H, then H is a generalized extreme value (GEV) distribution, i.e., $H = H_\xi$ given by

$$ H_\xi(x) = \left\{ \begin{array}{l} \exp\left(-(1+\xi x)_+^{-1/\xi}\right) \text{ if } \xi \neq 0, \\ \\ \exp\left(-\exp(-x)\right) \text{ if } \xi = 0, \end{array} \right. $$

where $y_+ = \max\{y, 0\}$ is the positive part of y. The definition domain of H_ξ is $(-1/\xi, +\infty)$ if $\xi > 0$, $(-\infty, -1/\xi)$ if $\xi < 0$, and the whole real line $\mathbb{R}$ if $\xi = 0$. Here, the parameter ξ controlling the right tail is called the tail index or the extreme value index. The three classical extreme value distributions are special cases of the GEV family: if $\xi > 0$, we have the Fréchet distribution, if $\xi < 0$, we have the Weibull distribution, and $\xi = 0$ gives the Gumbel distribution. When $\xi > 0$, we face lifetimes with heavy tails, which contradicts empirical evidence available for human lifetimes (in this case, forces of mortality decrease with attained age). Thus, the cases $\xi = 0$ and $\xi < 0$ are of interest for life insurance applications. Notice that if (3.1) holds with $\xi < 0$, then $\omega < \infty$, so that a negative value of ξ supports the existence of a finite ultimate age ω.

A sufficient condition for (3.1) to hold is

$$\lim_{x\to\omega} \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\mu_x}\right) = \xi \tag{3.2}$$

where

$$\mu_x = \frac{\frac{\mathrm{d}}{\mathrm{d}x}{}_x q_0}{{}_x p_0}$$

is the force of mortality at age x. Intuitively speaking, $\frac{1}{\mu_x}$ can be considered as the force of resistance to mortality, or force of vitality at age x. The resistance to mortality must thus stabilize when $\xi = 0$ or become ultimately linear. A negative ξ indicates that the resistance ultimately decreases at advanced ages. For $\xi < 0$, we have $\omega < \infty$, and condition (3.2) implies

$$\lim_{x\to\omega} \big((\omega - x)\mu_x\big) = -\frac{1}{\xi}.$$

3.2 Limiting Remaining Lifetime Distribution

Consider the remaining lifetime $T - x$ at age x, given $T > x$, with distribution function

$$s \mapsto {}_s q_x = \mathrm{P}[T - x \le s | T > x].$$

It may happen that for large attained age x, this conditional probability distribution stabilizes after a normalization—that is, there exists a positive function a such that

$$\lim_{x\to\omega} \mathrm{P}\left[\frac{T-x}{a(x)} > s \middle| T > x\right] = 1 - G(s), \;\; s > 0, \tag{3.3}$$

where G is a nondegenerate distribution function. It is possible to establish that only a limited class of distribution functions are eligible in (3.3), namely

$$\begin{aligned} G(s) = G_\xi(s) &= \ln H_\xi(s) \\ &= \begin{cases} 1 - (1+\xi s)_+^{-1/\xi} \text{ if } \xi \neq 0, \\ 1 - \exp(-s) \text{ if } \xi = 0. \end{cases} \end{aligned}$$

Here, the support is the half positive real line if $\xi \ge 0$ and $[0, -1/\xi]$ if $\xi < 0$. The related scale family known as the generalized Pareto (GP) distribution is then defined as

$$G_{\xi,\beta}(s) = G_\xi\left(\frac{s}{\beta}\right), \;\; \beta > 0.$$

At particular cases of the GP distribution functions $G_{\xi;\beta}$, we find some classical distributions, namely, the Pareto distribution when $\xi > 0$, the type II Pareto distribution when $\xi < 0$, and the negative exponential distribution when $\xi = 0$. Thus, when $\xi = 0$, the remaining lifetimes at high ages become ultimately negative exponentially distributed, so that the forces of mortality stabilize, in line with the empirical study conducted by Gampe (2010).

Clearly, extreme value analysis for maxima is thus closely connected with the study of residual lifetimes. It can be shown that (3.1) holds with $H = H_\xi$, if and only if (3.3) holds with $G = G_\xi$. In words, G_ξ describes the remaining lifetimes above sufficiently old ages if and only if H_ξ governs the sample maximum behavior, i.e., if F belongs to the domain of attraction of a GEV distribution.

For some appropriate function $\beta(\cdot)$, the approximation

$$ {}_sq_x \approx G_{\xi;\beta(x)}(s) \text{ for } s \geq 0 \tag{3.4}$$

holds for x large enough. The approximation (3.4) is justified by the Pickands–Balkema–de Haan theorem:

$$\lim_{x\to+\infty} \sup_{s\geq 0} \left| {}_sq_x - G_{\xi,\beta(x)}(s) \right| = 0 \tag{3.5}$$

is true provided that F satisfies some rather general technical conditions. In view of (3.4), the remaining lifetimes at age x can be treated as a random sample from the GP distribution, provided x is large enough.

If $\omega < \infty$ (i.e., $\xi < 0$), then a suitable transformation of the extreme value index ξ possesses an intuitive interpretation. Recall that the remaining life expectancy at age x, denoted as e_x, is defined as

$$e_x = \mathrm{E}\left[T - x | T > x\right].$$

Aarssen and de Haan (1994) established that for $\xi < 0$, so that an upper bound ω exists on the life span, (3.1) is equivalent to

$$\lim_{x\to\omega} \mathrm{E}\left[\frac{T-x}{\omega-x}\Big| T > x\right] = \lim_{x\to\omega} \frac{e_x}{\omega - x} = -\frac{\xi}{1-\xi} = \alpha$$

These authors call $\alpha = \alpha(\xi)$ the perseverance parameter and provide the following explanation. Consider an individual still alive at some advanced age x. The ratio $\frac{T-x}{\omega-x}$ represents the percentage of the actual remaining lifetime $T - x$ to the maximum remaining lifetime $\omega - x$. This percentage stabilizes, on average, when $x \to \omega$ and converges to α, which thus appears as the expected percentage of the maximum possible remaining lifetime effectively used by the individual.

3.3 Threshold Selection

To be in a position to apply the results presented in the preceding section, we first have to determine a threshold age $x^\star$ such that the approximation (3.4) is sufficiently accurate for $x \geq x^\star$. Thus, our aim is to determine the youngest age beyond which the GP distribution offers a reasonable approximation of the remaining lifetime distribution. Notice that Li et al. (2008) and Bravo et al. (2012) also used statistical procedures to select the threshold age $x^\star$ above which the GP behavior appears. This is done by modeling the mortality below the threshold, using a standard parametric model (Gompertz in these studies), and by shifting to the GP model above the threshold. The optimal threshold value is then obtained by a grid search, maximizing the likelihood of the composite model. However, this selection procedure heavily depends on the appropriateness of the parametric model describing mortality below

the threshold. This introduces an additional specification risk in the analysis of mortality at the oldest ages, whereas EVT is in essence a nonparametric approach to modeling the tails of a distribution. The flexible extreme value mixture model discussed in this section is also based on a composite model, but it replaces the parametric specification for younger ages with a flexible nonparametric one, avoiding this drawback.

To identify the optimal threshold value, we can use the following graphical tools from EVT.

3.3.1 Empirical Mean Excess Function Plot

It is easily checked that when a lifetime has the GP distribution function $G_{\xi,\beta}$, the remaining life expectancy is a linear function in the attained age x; that is,

$$e_x = \frac{\beta}{1-\xi} + \frac{\xi}{1-\xi}x \tag{3.6}$$

provided $x < \omega \Leftrightarrow \beta + x\xi > 0$. Hence, the idea is to determine, on the basis of the graph of the empirical version $\widehat{e}_x$ of e_x, a pivot age $x^\star$ such that $\widehat{e}_x$ becomes approximately linear for older ages.

The empirical version of the remaining life expectancy $\widehat{e}_x$, viewed as a function of attained age x, is displayed in Figure 3.1. These values have been obtained from the observed ages at death $\{t_1, t_2, \ldots, t_n\}$ represented in the database by

$$\widehat{e}_x = \frac{\sum_{i=1}^n t_i \mathrm{I}[t_i > x]}{\#\{t_i : t_i > x\}} - x = \frac{\sum_{i=1}^n (t_i - x)\mathrm{I}[t_i > x]}{\#\{t_i : t_i > x\}}$$

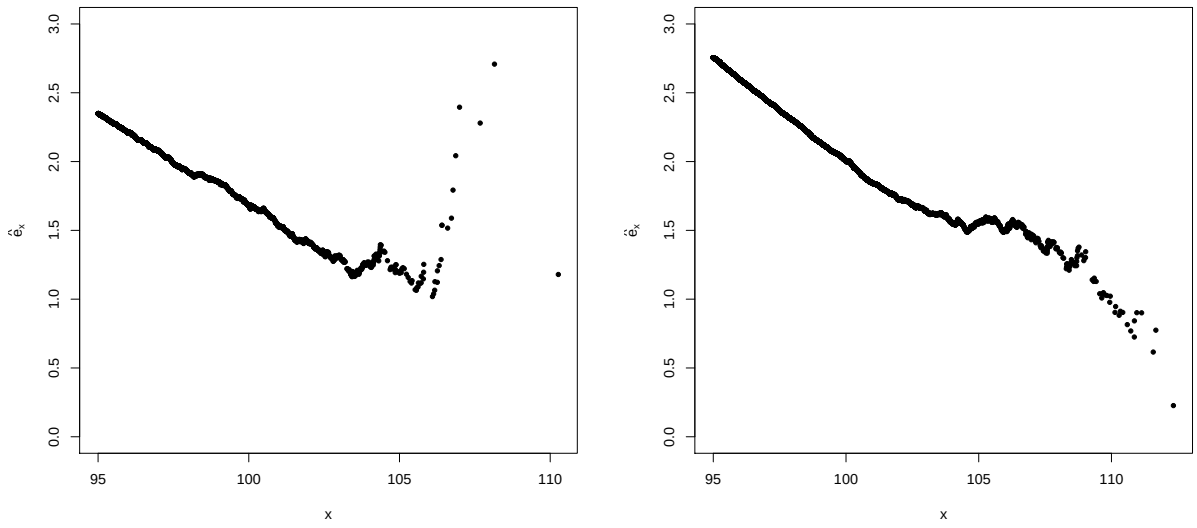
where $\mathrm{I}[A] = 1$ if the event A did occur and 0 otherwise, and where $\#B$ is the number of elements in the set B. As usual, $\widehat{e}_x$ is only evaluated at the observations, i.e., for $x \in \{t_1, t_2, \ldots, t_n\}$. Denoting the observed ages at death arranged in ascending order as $t_{[1]} \leq t_{[2]} \leq \ldots \leq t_{[n]}$, we have in this case

$$\widehat{e}_{t_{[k]}} = \frac{1}{n-k}\sum_{j=1}^{n-k}(t_{[k+j]} - t_{[k]}).$$

Clearly, if the remaining lifetimes obey the negative exponential distribution, then the mean excess function is constant. Consequently, the plot of e_x versus age x will be a horizontal line. Short-tailed distributions will show a downward trend, and an upward trend will be an indication of heavy-tailed behavior. The linear, decreasing shape of $\widehat{e}_x$ in Figure 3.1 contradicts the negative exponential behavior of the remaining lifetime. A downward trend is clearly visible in Figure 3.1, supporting short-tailed behavior. Therefore, a negative value of tail index ξ is expected. Notice that the apparent increasing trend visible in the far right part of Figure 3.1 for males, which conflicts with our intuition, is a feature of the data and can be attributable to the high volatility in the last ages at death observed in the population.

According to Poulain et al. (2001), the remaining life expectancies for Belgian male and female centenarians (i.e. e_{100}) were estimated to 1.68 and 1.97, respectively. This is in line with the values displayed in Figure 3.1.

Figure 3.1: Empirical Expected Remaining Lifetimes $\widehat{e}_x$ for Males (Left) and Females (Right)



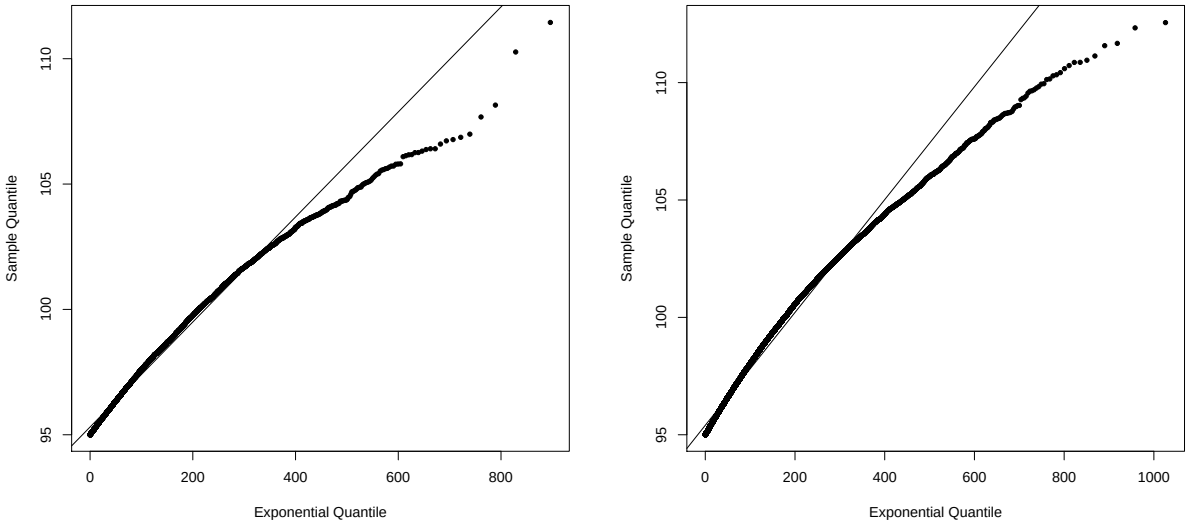
The comparison with negative exponentially distributed lifetimes can be performed with the help of an exponential Q-Q plot. Its interpretation is easy. If the data are an independent and identically distributed sample from a negative exponential distribution, the points should lie approximately along a straight line. A convex departure from the ideal shape indicates a heavier-tailed distribution in the sense that empirical quantiles grow faster than the theoretical ones. Concavity indicates a shorter-tailed distribution. It is easy to see a concave pattern of the exponential Q-Q plot in Figure 3.2, confirming the short-tailed behavior of the lifetime distribution.

Here it is tricky to detect the lowest threshold beyond which the mean excess function plot becomes approximately linear (see Figure 3.2). Consequently, we favor automatic selection procedures for the threshold suggested in the literature. We refer the reader to Scarrott et al. (2012) for a detailed review of threshold selection methods.

3.3.2 Pickands Method

In his pioneered paper establishing the theorem underlying the POT approach, Pickands (1975) also proposed a selection procedure for the threshold. Pickands (1975) defined $x^\star \equiv T_{[n-4M+1]}$, since the $4M$ largest observations intuitively contain information about the upper tail of the distribution function. We recall that $T_{[k]}$ is the kth-order statistics in ascending order. Specifically, Pickands (1975) suggested computing M in the following way. For each $l, l = 1, 2, \ldots, \lfloor n/4 \rfloor$, let $d_l = \sup_{0 \leq x < \infty} |\widehat{S}_l(x) - \widehat{\overline{G}}_l(x)|$, where $\widehat{S}_l(x)$ is the empirical upper tail, defined as

Figure 3.2: Exponential Q-Q Plot for Lifetimes for Males (Left) and Females (Right)



$$\hat{S}_l(x) = \frac{1}{4l} \sum_{m=1}^{4l-1} \mathrm{I}\Big[T_{[n-m+1]} - T_{[n-4l+1]} \geq x\Big] \tag{3.7}$$

and $\widehat{\overline{G}}(x)$ is the estimate of the survival function of the GP distribution with parameters ξ and β being estimated by the percentile method using the median and the third quartile. Notice that d_l is a kind of Kolmogorov-Smirnov test statistic. Finally, M is obtained as the smallest integer solution of

$$d_M = \min_{1 \leq l \leq \lfloor n/4 \rfloor} d_l. \tag{3.8}$$

Applying this method to our data sets, we obtain the results displayed in Table 3.1.

Table 3.1: Values of M, $4M$, and $T_{[n-4M+1]}$ Obtained by the Pickands Method for Both Males and Females

	Male	**Female**
M	2,302	3,159
$4M$	9,208	32,256
$T_{[n-4M+1]}$	95.24	95.41

3.3.3 Reiss-Thomas Method

Reiss and Thomas (1997) proposed an automatic selection procedure to choose the optimal value of the threshold $x^\star$. According to these authors, $x^\star \equiv T_{[n-k^\star+1]}$, where $k^\star$ minimizes

$$\frac{1}{k-1}\sum_{i\leq k} i^\delta \left(\widehat{\xi}_i - \widehat{\xi}_k\right)^2. \tag{3.9}$$

Here, $\widehat{\xi}_i$ is a shape parameter estimator based on the upper ordered statistics $T_{[n-i+1]}, \ldots, T_{[n]}$. The tuning parameter δ satisfies $0 \leq \delta \leq 0.5$. Neves and Alves (2004) recommend taking 0.35 as the value of δ for computations. In this setting, final values for $k^\star$ and $x^\star$ are recorded in Table 3.2.

Table 3.2: Values of $k^\star$ and $x^\star$ for Both Males and Females With the Reiss-Thomas Method

	Male	**Female**
$k^\star$	9,959	4,108
$T_{[n-k^\star+1]}$	95.03	100.89

3.3.4 Flexible Extreme Value Mixture Model

As pointed out in Scarrott et al. (2012), more sophisticated threshold selection procedures have been developed in the last decade, such as the mixture model. Our goal is to approximate the upper tail of the lifetime distribution with the GP distribution without specifying any parametric form for the bulk of the distribution in order to avoid a specification risk. This differs from Li et al. (2008), who specified a Gompertz distribution below the threshold. Our purpose is achieved if the bulk of the distribution is simultaneously estimated nonparametrically. This model is precisely the one developed by MacDonald et al. (2011) with their flexible extreme value mixture distribution. Formally, if F is the common distribution of the $T_1, T_2, T_3, \ldots$, then it is expressed as

$$F(x|\lambda, u, \beta, \xi) = \begin{cases} (1-\phi_u)\frac{H(x|\lambda)}{H(u|\lambda)} & x \leq u \\ (1-\phi_u) + \phi_u G(x|u,\beta,\xi) & x > u \end{cases} \tag{3.10}$$

where ϕ_u is the probability of being above the optimal threshold $u \equiv x^\star$, H is a nonparametric kernel distribution that depends on a bandwidth parameter λ, and G is the GP distribution function. The likelihood of model (3.10) is expressed as

$$L(\theta) = L_K(\lambda, u) L_{PP}(u, \mu, \beta, \xi) \tag{3.11}$$

where $\theta = (\lambda, u, \mu, \beta, \xi)$ is the vector of parameters. L_K and L_{PP} respectively refer to the likelihood of the nonparametric and GP parts of the overall distribution.

The function `nlbckdengpd` from the R-package `evmix` has been used for estimation. The Epanechnikov kernel was specified for the nonparametric part, as suggested by Wang et al. (1998), since it is optimal in the mean square sense. To avoid issue of ties in kernel estimation, similar exact ages were uniformly spread over the same day, as suggested in Einmahl et al. (2008). Maximum likelihood estimation of the model requires initial values for parameters λ, u, β and ξ for numerical optimization. As there exists no rule of thumb for choosing these values, a reasonable grid of values for each parameter was fixed, and then the negative log likelihood was computed at each point of the Cartesian product of those grids in order to reasonably surround the parameters' space. Initial values were set as parameter values, which minimized the negative log likelihood. These values are in Table 3.3.

Table 3.3: Initial Values for Parameters of the Flexible Extreme Value Mixture Model for Both Males and Females

	Males	**Females**
u	99	100
ξ	–0.1	–0.1
β	2	2.2
λ	0.019	0.019

Estimation results are summarized in Table 3.4. To asses the impact of the age range considered for estimation, we also consider the restricted data set covering ages 98 and over. We start at age 98 because market statistics are available from the Belgian National Bank, acting as the insurance regulator, up to that age (the last category is open, gathering ages 99 and over).

Table 3.4: Estimated Values Parameters of the Flexible Extreme Value Mixture Model for Both Males and Females

	Males		**Females**	
x	≥ 98	≥ 95	≥ 98	≥ 95
u	98.89	97.82	99.99	100.01
$\widehat{\xi}$	–0.156	–0.140	–0.125	–0.125
$\widehat{\beta}$	2.146	2.218	2.251	2.248
$\widehat{\lambda}$	0.030	0.025	0.028	0.039

We see from Table 3.4 that the estimated parameters remain stable when the age range is modified. In particular, the optimal age $x^\star \equiv u$ is not affected when the initial age moves from 95 to 98.

3.4 Final Choice

Thresholds selected by all selection procedures are summarized in Table 3.5.

The GP distribution enjoys the convenient threshold stability property, which ensures that if the lifetime has distribution function $G_{\xi,\beta}$, then the remaining lifetime at any age x

Table 3.5: Thresholds Selected by All Three Automatic Selection Procedures for Both Males and Females

	Males		Females	
Age	≥ 98	≥ 95	≥ 98	≥ 95
Pickands	98.40	95.24	98.60	95.41
Reiss & Thomas	98.07	95.03	100.89	100.89
Mixture model	98.89	97.82	99.99	100.01

has distribution function $G_{\xi,\beta+\xi x}$. This basically says that, provided the lifetime obeys the GP model, the remaining lifetime at any attained age x is still GP with the same index ξ. Therefore, the threshold $x^\star$ was finally fixed at the maximum of those values in order to be in line with the property of threshold stability of the GP distribution. Henceforth, for the remainder of the paper, $x^\star$ is set to 98.89 for males and to 100.89 for females. This means that the extreme value region for female ages appeared much later than for male ages. It also clearly reflects the gender gap.

3.5 Maximum Likelihood (ML) Estimation of the GP Parameters

Now that the threshold age $x^\star$ has been selected, the GP parameters have to be estimated from the observed T_i exceeding $x^\star$. Several methods are available, including moment estimation, probability-weighted moment estimation and maximum likelihood estimation. Maximizing the likelihood function is used here, due to its optimal statistical properties, intuitive contents and general acceptance in the actuarial community. Alternative methods are considered in Section 3.6.

The GP likelihood function for observed remaining lifetimes at attained age $x^\star$ is given by

$$\mathcal{L}(\xi,\beta) = \prod_{i|t_i>x^\star} \left(\frac{1}{\beta}\left(1+\frac{\xi}{\beta}(t_i-x^\star)\right)^{-\frac{1}{\xi}-1} \right).$$

The log likelihood to be maximized is

$$L(\xi,\beta) = -\ln\beta \#\{t_i|t_i>x^\star\} - \left(1+\frac{1}{\xi}\right)\sum_{i|t_i>x^\star} \ln\left(1+\frac{\xi}{\beta}(t_i-x^\star)\right) \tag{3.12}$$

where $\#\{t_i|t_i>x^\star\}$ is the number of survivors at age $x^\star$, denoted as $L_{x^\star}$. This optimization problem requires numerical algorithms. There are different approaches to get starting values for the parameters ξ and β. A natural approach consists of using moment conditions (that is, we equate sample mean and sample variance to their theoretical expressions involving ξ and β), which gives

$$\widehat{\xi}^{MOM} = \frac{1}{2}\Big(1-\frac{\bar{x}^2}{s^2}\Big) \text{ and } \widehat{\beta}^{MOM} = \frac{1}{2}\bar{x}\Big(\frac{\bar{x}^2}{s^2}+1\Big), \tag{3.13}$$

where $\bar{x}$ and s^2 are the sample mean and variance of the observed $t_i - x^\star$. For $\xi < 0$, the use of (3.13) does not require any additional assumption, as all moments are finite.

Another convenient way to get starting values for the maximum likelihood estimates consists of fitting a straight line to the empirical remaining life expectancies, as a function of attained age x. The intercept and slope of a straight line fit to $\widehat{e}_x$ determine the estimations of ξ and β in accordance with (3.6).

It can be shown that for $\xi > -0.5$, the maximum likelihood (ML) regularity conditions are fulfilled. Hosking and Wallis (1987) proved that the ML estimators of the GP parameters $(\hat{\xi}, \hat{\beta})$ are asymptotically normally distributed with expected value (ξ, β) and approximate covariance matrix

$$\mathbf{\Sigma}_{(\widehat{\beta},\widehat{\xi})} = \frac{1}{n} \begin{pmatrix} (1+\xi)^2 & \beta(1+\xi) \\ \beta(1+\xi) & 2\beta^2(1+\xi) \end{pmatrix}. \tag{3.14}$$

This result allows us to obtain the standard errors for the ML estimators. For reasons of convergence of the algorithm, a reparametrization of the GP model with parameters $\sigma = \log(\beta)$ and $\alpha = 1/\xi$ is often used to maximize the likelihood.

ML estimation of GP parameters can be found in Table 3.6, together with standard errors. Since $\widehat{\xi} > -0.5$ in both cases, upper bounds of the 95% confidence interval for tail index ξ can be obtained from the asymptotic normality of the ML estimator $\widehat{\xi}$ recalled above. The upper bounds of the resulting confidence intervals are –0.065 for females and –0.102 for males. Both bounds are negative, supporting the assumption $\xi < 0$ for both genders.

Table 3.6: Maximum Likelihood Estimates for the GP Parameters β and ξ, With Standard Errors and Kolmogorov-Smirnov Goodness-of-Fit p-Values

Parameters	Male	Female
$x^\star$	98.89	100.89
$L_{x^\star}$	1,940	4,104
$\widehat{\xi}^{ML}$	–0.132	–0.092
s.e $(\widehat{\xi}^{ML})$	0.015	0.014
$\widehat{\beta}^{ML}$	2.098	2.019
s.e $(\widehat{\beta}^{ML})$	0.057	0.042
KS pvalue	0.828	0.951

To check whether the data comply with the GP distribution, we plotted the histogram and the GP density function (see Figure 3.3) and the GP Q-Q plot (see Figure 3.4). In the Q-Q plot, the empirical quantiles versus the estimated GP quantiles are represented. If the GP model fits, a linear pattern must be visible.

Both histogram and density function exhibit the same shape. Furthermore, there is a clear linear pattern in the Q-Q plot, which confirms that the GP model adequately describes the remaining lifetime distribution above $x^\star$.

Moreover, a formal Kolmogorov-Smirnov goodness-of-fit test has been performed to check the compliance of the data with the GP distributions. The resulting p-values are reported in Table 3.6. As they exceed all the usual confidence levels, the null hypothesis of a GP behavior cannot be rejected for the lifetime data under study.

The analysis performed in this section could theoretically be applied separately to each cohort. In practice, however, this is hardly possible, because of the much-restricted group sizes at older ages when considering cohorts in isolation. As a consequence, parameter

estimates would have a poor efficiency. This limitation is inherent to all studies devoted to mortality at old ages. The descriptive statistics reported in Table 2.2 do not reveal any trend across cohorts, based on observed cohort-specific highest ages at death and remaining life expectancies.

Let us now investigate more accurately whether mortality at old ages differs between cohorts. To this end, denote as T the age at death for an individual included in the study, belonging to cohort C. To keep reasonable sample sizes, we created four groups of consecutive cohorts. Each of the first three groups, G_1, G_2 and G_3, consists of five consecutive cohorts, whereas the last group, $G_4 = \{1901, \ldots, 1904\}$, contains the last four cohorts. Denote as

$$F_j(y) = \mathrm{P}[T \leq y | C \in G_j],\ j = 1, \ldots, 4 \tag{3.15}$$

the lifetime distribution for the jth group of cohorts. The unconditional lifetime distribution F is defined as $F(y) = \mathrm{P}[T \leq y]$. If both random variables T and C are independent, then both conditional F_j and unconditional F distributions coincide. Thus, we are interested in testing for

$$\begin{cases} H_0 : F_1 = F_2 = F_3 = F_4 = F \\ H_1 : F(y) \neq F_j(y) \text{ for some } j \text{ and } y. \end{cases} \tag{3.16}$$

Here, we use the test statistic of Cramer–Von Mises type proposed by Kiefer (1959). The p-values obtained for both male and female populations are respectively 0.853 and 0.7019, leading to the nonrejection of the null hypothesis at all usual confidence levels. No significant mortality trends across cohorts are detected at older ages in our data set, which supports the analysis based on the 19 cohorts. Notice that this homogeneity across cohorts is in line with the study by Fraga Alves et al. (2016) performed on supercentenarians, where no time trend has been detected.

3.6 Comparison With Alternative Estimators

Apart from the ML estimator, several other estimation procedures have been proposed in the literature for the tail index ξ. In this section, we compare ML estimation with some commonly used alternative estimators:

moment estimator denoted as $\widehat{\xi}^{MOM}$ and $\widehat{\beta}^{MOM}$, which have been presented in Section 3.5;

probability weighted moment estimator denoted $\widehat{\xi}^{PMOM}$ and $\widehat{\beta}^{PMOM}$, which are often used in EVT studies (see Beirlant et al. (2005a) for theoretical details);

extended Hill estimator applicable to the case $\xi < 0$, denoted as $\widehat{\xi}^{H}$ and $\widehat{\beta}^{H}$ and proposed by Beirlant et al. (2005b);

moment-type estimator denoted as $\widehat{\xi}^{M}$ and $\widehat{\beta}^{M}$ and proposed by Dekkers et al. (1989).

Results from these alternative estimators are summarized in Tables 3.7 and 3.8 for both females and males at optimal thresholds $x^\star$. We see there that the estimated tail index

Figure 3.3: Histogram and GP Density Function of the Remaining Lifetimes at Age $x^\star = 98.89$ for Males (Left) and $x^\star = 100.89$ for Females (Right)

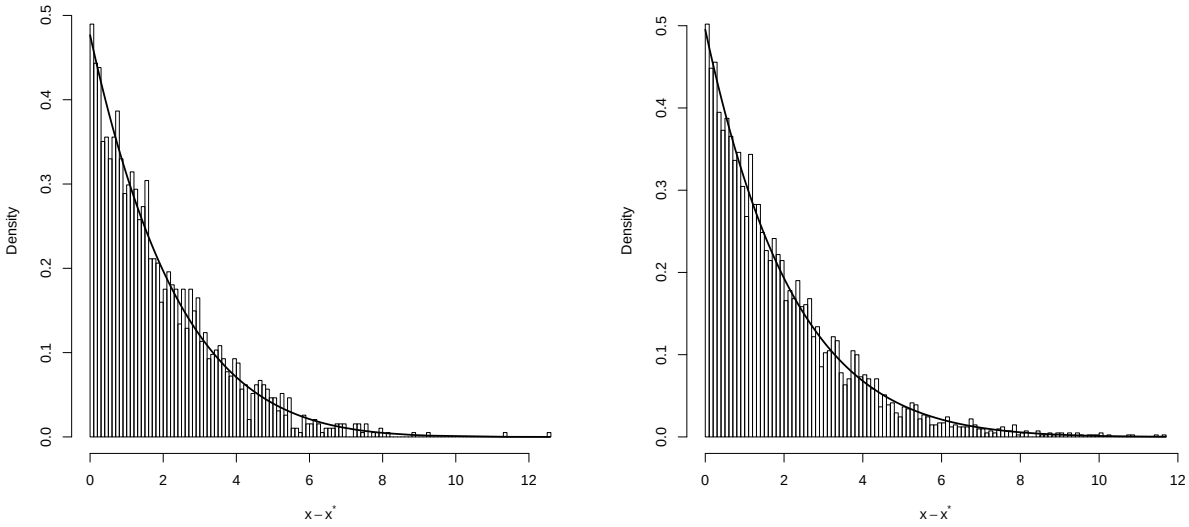
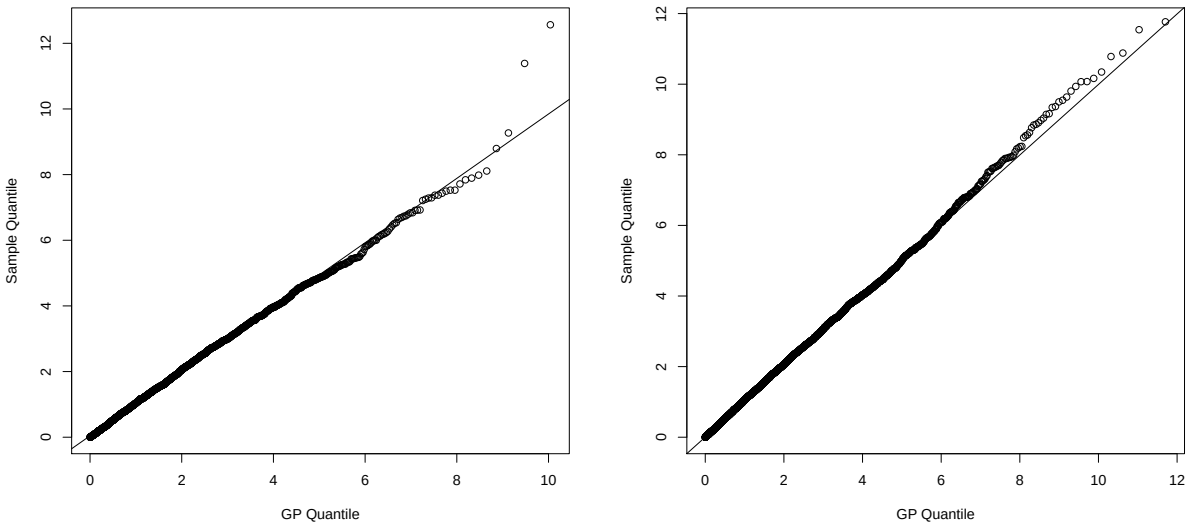


Figure 3.4: GP Q-Q Plots for the Remaining Lifetimes at Age $x^\star = 98.89$ for Males (Left) and $x^\star = 100.89$ for Females (Right)



for females is smaller than the one for males for all estimators, while the estimated scale parameter for females is higher than the one for males. Clearly, the estimation method affects the estimated tail index, which in turn may lead to different conclusions about, for instance, the value of the ultimate age. Despite these differences, all estimators support $\xi < 0$.

Table 3.7: Comparison of Estimation Results With Optimal Threshold $x^\star = 98.89$ for Males

	ML	**MOM**	**PMOM**	**M**	**H**
$\widehat{\xi}$	–0.132	–0.163	–0.171	–0.163	–0.155
$\widehat{\beta}$	2.098	2.160	2.175	2.128	2.113

Table 3.8: Comparison of Estimation Results With Optimal Threshold $x^\star = 100.89$ for Females

	ML	**MOM**	**PMOM**	**M**	**H**
$\widehat{\xi}$	–0.092	–0.097	–0.105	–0.097	–0.094
$\widehat{\beta}$	2.019	2.027	2.043	1.997	1.991

4 Some Applications of the Model

4.1 Ultimate Age ω

As the negativity of the tail index is supported by the data, this implies the existence of a finite ultimate age ω that can be estimated by

$$\widehat{\omega} = x^\star - \frac{\widehat{\beta}}{\widehat{\xi}}. \tag{4.1}$$

Estimated ultimate ages computed from (4.1) with ML estimates are displayed in Table 4.1. Simulated confidence intervals at 95% also are shown there, based on 10,000 simulations.

Table 4.1: ML Estimated Ultimate Ages with 95% Simulated Confidence Interval for Both Male and Female
Note: $\widehat{\omega}^-$ and $\widehat{\omega}^+$ are respectively lower and upper bounds.

	Males	**Females**
$\widehat{\omega}$	114.82	122.73
$\widehat{\omega}^-$	112.31	118.19
$\widehat{\omega}^+$	118.87	131.21

The female population has the highest estimated ultimate age. Estimations are in line with the highest ages at death recorded for females (112.58) and for males (111.47).

Estimated ultimate ages derived from alternative estimators also are considered. For moment type estimators $\widehat{\xi}^M$ and $\widehat{\xi}^H$, we estimate ω as proposed by Einmahl and Magnus (2008):

$$\widehat{\omega}_{k^\star} = T_{[n-k^\star]} \left(1 - M_n^{(1)}(k^\star) \frac{1 - \min(0, \widehat{\xi}_{k^\star})}{\widehat{\xi}_{k^\star}} \right) \tag{4.2}$$

with $k^\star = n - i^\star$ and

$$M_n^{(1)}(k) = \frac{1}{k} \sum_{i=0}^{k-1} \Big(\log(T_{[n-i]}) - \log(T_{[n-k]}) \Big).$$

Under some mild technical conditions that can be found in Dekkers et al. (1989),

$$\frac{\sqrt{k^\star}\,(\widehat{\omega} - \omega)}{T_{[n-k^\star]} M_n^{(1)}(k^\star) \left(1 - \widehat{\xi}\right)}$$

is normally distributed with mean 0 and variance

$$\frac{(1-\xi)^2(1-3\xi+4\xi^2)}{\xi^4(1-2\xi)(1-3\xi)(1-4\xi)}.$$

Therefore, simulated confidence intervals for ω can be provided. Results are summarized in Tables 4.2 and 4.3 for females and males, respectively. Notice how the estimated ultimate age is sensitive to the estimator used for computation.

Table 4.2: Alternative Estimation Results With Optimal Threshold $x^\star = 98.89$ for Males

	ML	**MOM**	**PMOM**	**M**	**H**
$\widehat{\omega}$	114.82	112.14	111.64	111.91	112.53
$\widehat{\omega}^-$	112.31	109.81	109.05	109.07	109.29
$\widehat{\omega}^+$	118.87	116.20	116.65	114.75	115.76

Table 4.3: Alternative Estimation Results With Optimal Threshold $x^\star = 100.89$ for Females

	ML	**MOM**	**PMOM**	**M**	**H**
$\widehat{\omega}$	122.73	121.86	120.30	121.41	122.14
$\widehat{\omega}^-$	118.19	117.43	115.80	115.80	116.09
$\widehat{\omega}^+$	131.21	130.22	129.53	127.03	128.20

4.2 Highest Age at Death M_n

As pointed out by Wilmoth and Robine (2003), the maximum age at death can be modeled statistically as an extreme value of a random sample. To study the behavior of the highest age at death, we use the GP approximation of the exceedances. Under the conditions that yield to the GP approximation $G_{\xi;\beta}$ for the remaining lifetime distribution at age $x^\star$, the number $L_{x^\star}$ of survivors at threshold age $x^\star$ is roughly Poisson with mean $\ell_{x^\star} = \mathrm{E}[L_{x^\star}]$ (as an approximation to the binomial distribution valid for large sizes and small success probabilities). As a consequence, the distribution function

$$\begin{aligned} s \mapsto \quad & \mathrm{P}[M_{[L_{x^\star}]} \leq s] \\ = \quad & \mathrm{P}[L_{x^\star} = 0] + \sum_{k=1}^{\infty} \mathrm{P}[L_{x^\star} = k, T_{[1]} - x^\star \leq s - x^\star, \ldots, T_{[k]} - x^\star \leq s - x^\star] \end{aligned}$$

can be approximated by

$$\begin{aligned} & \exp(-\ell_{x^\star}) + \sum_{k=1}^{\infty} \exp(-\ell_{x^\star}) \frac{\ell_{x^\star}^k}{k!} \left(1 - \left(1 + \xi \frac{s - x^\star}{\beta}\right)_+^{-1/\xi}\right)^k \\ = \quad & \exp\left(-\lambda \left(1 + \xi \frac{s - x^\star}{\beta}\right)_+^{-1/\xi}\right). \end{aligned}$$

Hence, we find the following approximation for the quantile at probability level $1 - \epsilon$ of $M_{[L_{x^\star}]}$:

$$x^\star + \frac{\beta}{\xi}\left(\left(-\frac{\ell_{x^\star}}{\ln(1-\epsilon)}\right)^{\xi} - 1\right)$$

Notice that if we define the location-scale family $H_{\xi;\mu,\psi}$ by

$$H_{\xi;\mu,\psi}(x) = H_\xi\left(\frac{x - \mu}{\psi}\right), \ \ \mu \in \mathbb{R}, \ \ \psi > 0,$$

then the distribution of the maximum $M_{L_{x^\star}}$ of the lifetimes of the $L_{x^\star}$ individuals reaching age $x^\star$ can be approached by a GEV distribution $H_{\xi;\mu,\psi}$, where

$$\mu = \beta\xi^{-1}(\ell_{x^\star}^{\xi} - 1) \text{ and } \psi = \beta\ell_{x^\star}^{\xi}.$$

The corresponding quantile function is given by

$$H_{\xi,\mu,\psi}^{-1}(\epsilon) = \begin{cases} \mu + \frac{\psi}{\xi}\Big((-\ln(\epsilon))^{-\xi} - 1\Big) & \text{if } \xi \neq 0 \\ \mu - \psi \ln(-\ln(u)) & \text{if } \xi = 0 \end{cases}$$

So that the median of the highest age at death is given by

$$H_{\xi,\mu,\psi}^{-1}(0.5) = \mu + \psi\frac{(\ln(2))^{-\xi} - 1}{\xi}.$$

This median will limit $\mu - \psi \ln(\ln 2)$ as $\xi \to 0$. If $\xi < 1$, then

$$\mathrm{E}[M_{L_{x^\star}}] = \mu + \frac{\psi}{\xi}\Big(\Gamma(1-\xi) - 1\Big)$$

will limit $\mu + \psi j$ as $\xi \to 0$, where j is the Euler constant (≈ 0.57721), which equals $\Gamma'(1)$, where Γ' is the digamma function, i.e., the derivative of the gamma function. If $\xi < 1/2$, then

$$\mathrm{V}[M_{L_{x^\star}}] = \frac{\psi^2\Big(\Gamma(1-2\xi) - \big((\Gamma(1-\xi)\big)^2\Big)}{\xi^2}$$

with limit $\psi^2\pi^2/6$ as $\xi \to 0$.

Notice that this suggests another estimation procedure for the tail index ξ, often referred to as the block maxima approach. See, for instance, Watts et al. (2006) for an application in life insurance. The idea is to fit the GEV distribution to maxima taken from independent blocks of data. Here, this involves recording the cohort-specific maxima and fitting the 19 gender-specific observations, using the GEV distribution $H_{\xi;\mu,\psi}$. Estimates are summarized in Table 4.4. They are obtained by using the R-package `ismev`, which provides tools for statistical modeling of extreme values.

Table 4.4: GEV Estimates and Their Standard Errors (in Parentheses) for Males and Females

	$\widehat{\mu}$	$\widehat{\beta}$	$\widehat{\xi}$
Males	105.83	1.323	–0.012
	(0.334)	(0.236)	(0.139)
Females	109.78	1.477	–0.434
	(0.375)	(0.279)	(0.170)

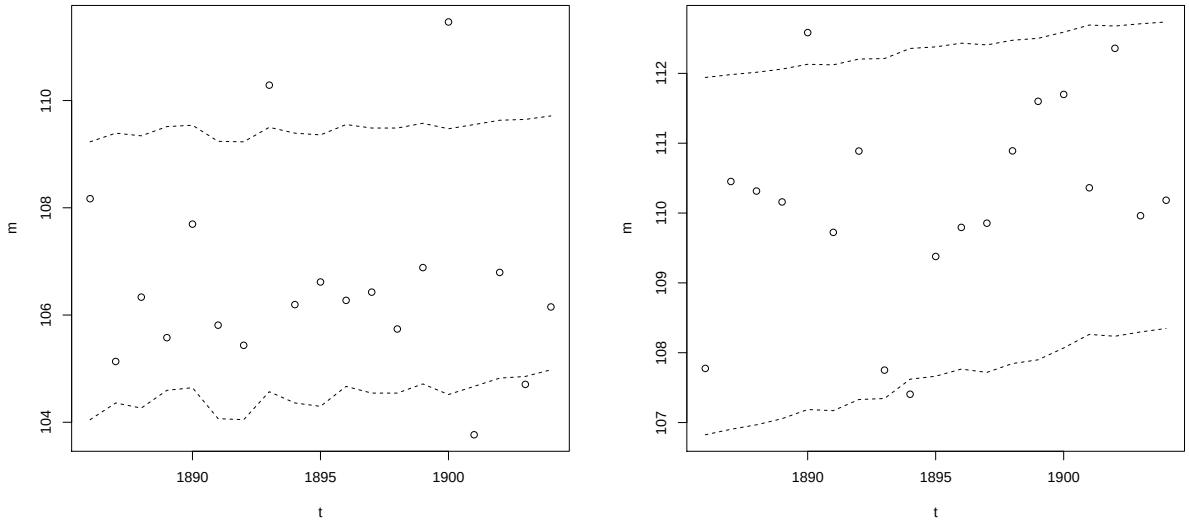
Theoretically, tail index estimates from the block maxima and POT approach should be very close. As one can see, results are slightly different, meaning that estimation is sensitive to the chosen approach. This difference may also be linked to the block definition we adopted for the block maxima approach. In fact, we simply consider cohort blocks, which makes more sense demographically. However, the block maxima only take into account the 19 observed cohort maxima, leading to more estimation instability.

Let us now compute the prediction intervals for the highest ages at death observed for each cohort and compare them with the observed, cohort-specific M_n. The result is visible in Figure 4.1, where we can see the observed cohort-specific maxima surrounded by 95% prediction intervals. Almost all cohort maxima are within the prediction interval.

4.3 Point Estimation of High Quantiles

Quantiles at usual probability levels can easily be estimated by their empirical counterparts. However, when we are interested in the very high quantiles, this approach is no longer

Figure 4.1: Observed Cohort-Specific Maxima Surrounded by 95% Prediction Intervals for Males (Left) and Females (Right)



valid, since estimation based on a limited number of large observations would be strongly inaccurate. Fortunately, the EVT approach offers an efficient alternative.

From (3.4), we see that (provided $x^\star$ is sufficiently large) a potential estimator for the remaining lifetime distribution ${}_sq_{x^\star}$ is $G_{\widehat{\xi};\widehat{\beta}}(s)$. Quantile estimators derived from this curve are conditional quantile estimators that indicate the potential survival beyond the threshold age $x^\star$ when it is attained. When estimates of the unconditional quantiles are of interest, relating the unconditional cumulative distribution function ${}_xq_0$ to $G_{\hat{\xi};\hat{\beta}}$ for $x^\star - \frac{\hat{\beta}}{\hat{\xi}} \geq x > x^\star$, through

$$\begin{aligned} {}_xq_0 &= 1 - {}_xp_0 \\ &= 1 - {}_{x^\star}p_{0x-x^\star}p_{x^\star} \\ &\approx 1 - {}_{x^\star}p_0\Big(1 - G_{\widehat{\xi};\widehat{\beta}}\big(x - x^\star\big)\Big) \end{aligned}$$

we can obtain high-level quantiles, i.e., solutions $t(\epsilon)$ to the equation

$${}_{t(\epsilon)}q_0 = 1 - \epsilon.$$

Provided the sample size is large enough, we can estimate ${}_{x^\star}p_0$ by its empirical counterpart so that estimated quantiles are estimated as

$$\widehat{t}(\epsilon) = x^\star + \frac{\widehat{\beta}}{\widehat{\xi}}\left(\left(\frac{L_{95}}{L_{x^\star}}(1-\epsilon)\right)^{-\widehat{\xi}} - 1\right) \tag{4.3}$$

where $L_{x^\star}$ and L_{95} are the number of survivors at the threshold age $x^\star$ and at age 95 (i.e., the total number of individuals under study for a given cohort), respectively.

The survival probability to age $x^\star - \frac{\hat{\beta}}{\hat{\xi}} \geq x > x^\star$ can be estimated by

$$\widehat{{}_xp_0} = \frac{L_{x^\star}}{L_{95}} \left(1 - G_{\widehat{\xi};\widehat{\beta}}(x - x^\star)\right).$$

Notice that for $x^\star - \frac{\hat{\beta}}{\hat{\xi}} \geq x > x^\star$, we have

$$\widehat{\mu}_x = \frac{1}{\widehat{\beta} + \widehat{\xi}(x - x^\star)}.$$

5 Discussion

Mortality at oldest ages is difficult to study because of data scarcity. In this paper, a unique Belgian database of individual lifetimes allowed us to accurately study this mortality using tools from extreme value theory. Maximum likelihood estimation gives a negative estimated tail index, suggesting the existence of an ultimate age. We found an ultimate age of 114.82 for males and 122.73 for females. Those results are consistent with observed data and demographic experience in the sense that they are higher than observed highest ages in the data set (111.47 for males and 112.58 for females) and approximately equal to the worldwide maximum age records. Interestingly, the obtained ultimate age for females is close to Jeanne Calment's record of 122.42 (122 years and 164 days); she was born in Arles, France, on February 21, 1875, and died at the same place on August 4, 1997. The values obtained in this paper are also in line with the study by Dong et al. (2016), who conclude that the human life span has a natural limit of about 115 years, with occasional outliers like Jeanne Calment.

Alternative estimators for the tail index were also considered. Results clearly show the sensitivity to the selected estimation procedure.

Contrarily to previous studies by Robine et al. (2005) and Gampe (2010), we do not conclude that the force of mortality approaches a constant level (0.7 in their studies). This may be explained by the following reasons:

- The IDL gathers information from different populations, whereas we conduct our analysis on a single population.

- Younger ages are included in the analysis.

Notice also that the EVT model assumes homogeneity, in contrast to frailty models recognizing the heterogeneity present in the population with respect to mortality. In that context, the flattening of the mortality curve at old ages is attributed to mortality selection. The EVT approach could have nevertheless led to the conclusion that the tail index is 0, meaning that the remaining lifetimes become negative exponentially distributed above a sufficiently high threshold, which approximately conforms with the conclusion drawn from frailty models (with Gompertz baseline and gamma frailty, for instance). The empirical findings seem, however, to contradict this conclusion, as data support a negative tail index, which results in a lighter tail than with the negative exponential distribution, and an increasing force of mortality at the oldest ages, tending to infinity at ultimate age.

As a by-product of our analysis, we also conclude that an ultimate age exists. However, we do not claim that this limit must be interpreted in the biological or demographic sense, induced by genes or another natural mechanism. Rather, the ultimate age that has been obtained in the present paper serves as a working upper bound on a policyholder's lifetime when actuarial calculations need to be performed. From a practical point of view, closing the life table by assuming that the missing q_x are approximately constant beyond the last available age—i.e., assuming that the remaining lifetimes ultimately conform to the negative exponential distribution—is equally efficient and even conservative for products exposed to longevity risk.

Acknowledgments

The authors wish to thank Professor Johan Segers for fruitful discussions, as well as two anonymous referees for useful suggestions that helped us to improve previous versions of our work.

References

Aarssen, K., and L. de Haan. 1994. On the Maximal Life Span of Humans. *Mathematical Population Studies* 4: 259–281.

Balkema, A., and L. de Haan. 1974. Residual Life Time at Great Age. *Annals of Probability* 2: 792–804.

Beirlant, J., J. Goegebeur, J. Teugels, J. Segers, D. D. Waal and C. Ferro. 2005a. *Statistics of Extremes: Theory and Applications.* New York: Wiley.

Beirlant, J., G. Dierckx and A. Guillou. 2005b. Estimation of the Extreme-Value Index and Generalized Quantile Plots. *Bernoulli* 11: 949–970.

Bravo, J., E. Coelho and M. Magalhaes. 2008. Mortality and Longevity Projections for the Oldest-Old in Portugal. In *Proceedings of the European Population Conference* available at `http://epc2008.princeton.edu/papers/80105`

Bravo, J., and P. Real. 2012. Modeling Longevity Risk Using Extreme Value Theory: An Empirical Investigation Using Portuguese and Spanish Population Data. Technical Report.

Brillinger, D. R. 1961. A Justification of Some Common Laws of Mortality. *Transactions of the Society of Actuaries* 13: 116–119.

Cebrian, A. C., M. Denuit, and P. Lambert. 2003. Generalized Pareto Fit to the Society of Actuaries' Large Claims Database. *North American Actuarial Journal* 7: 18–36.

Dekkers, A., J. Einmahl and L. de Haan. 1989. A Moment Estimator for the Index of an Extreme-Value Distribution. *Annals of Statistics* 17: 1833–1855.

Dong, X., B. Milholland and J. Vijg. 2016. Evidence for a Limit to Human Lifespan.

Nature 538: 257–259.

Einmahl, J. H., and J. R. Magnus. 2008. Records in Athletics Through Extreme-Value Theory. *Journal of the American Statistical Association* 103: 1382–1391.

Einmahl, J. H., and S. G. Smeets. 2011. Ultimate 100-m World Records Through Extreme-Value Theory. *Statistica Neerlandica* 65: 32–42.

Frage Alves, I., C. Neves and P. Rosario. 2017. A General Estimator for the Right Endpoint With an Application to Supercentenarian Women's Records. *Extremes* 20, 199–237.

Gampe, J. 2010. Human Mortality Beyond Age 110. In *Supercentenarians*, pp. 219–230. Berlin: Springer.

Gavrilov, L. A., and N. S. Gavrilova. 2011. Mortality Measurement at Advanced Ages: A Study of the Social Security Administration Death Master File. *North American Actuarial Journal* 15: 432–447.

Gumbel, E. J. 1937. *La Durée Extrême de la Vie Humaine.* Paris: Hermann.

Han, Z. 2005. Living to 100 and Beyond: An Extreme Value Study. In *Living to 100 and Beyond Symposium Monograph.* Chicago: Society of Actuaries.

Hanayama, N., and M. Sibuya. 2015. Estimating the Upper Limit of Lifetime Probability Distribution, Based on Data of Japanese Centenarians. *Journal of Gerontology: Biological Sciences*, 1–8.

Hosking, J., and J. Wallis. 1987. Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics* 29: 339–349.

Kannisto, V. 1994. *Development of Oldest-Old Mortality 1950–1990: Evidence From 28 Developed Countries.* Monographs on Population Aging 1. Odense, Denmark: Odense University Press.

Kiefer, J. 1956. K-Sample Analogues of the Kolmogorov-Smirnov and Cramer–Von Mises Tests. *Annals of Mathematical Statistics* 30: 420–447.

Li, J., A. Ng and W. Chan. 2010. Modeling Old-Age Mortality Risk for the Populations of Australia and New Zealand: An Extreme Value Approach. *Mathematics and Computers in Simulation* 81: 1325–1333.

Li, J. S. H., M. R. Hardy and K. S. Tan. 2008. Threshold Life Tables and Their Applications. *North American Actuarial Journal* 12: 99–115.

MacDonald, A., C. J. Scarrott, D. Lee, B. Darlow, M. Reale and G. Russell. 2011. A Elexible Extreme Value Mixture mMdel. *Computational Statistics and Data Analysis* 55: 2137–2157.

Manton, K. G., E. Stallard and J. W. Vaupel. 1986. Alternative Models for the Heterogeneity of Mortality Risks Among the Aged. *Journal of the American Statistical*

Association 81: 635–644.

Neves, C. and M. I. F. Alves. 2004. Reiss and Thomas' Automatic Selection of the Number of Extremes. *Computational Statistics and Data Analysis* 47: 689–704.

Ouellette, N., and R. Bourbeau. 2014. The Trajectory of Mortality at Ages 100 and Beyond: An Analysis of Individual Level Data in Canada. Paper presented at the Annual Meeting of the Population Association of America, Boston, May 1–2, 2014.

Pickands, J. 1975. Statistical Inference Using Extreme Order Statistics. *Annals of Statistics* 3: 119–131.

Poulain M. 2010. Le registre de population belge. In *Histoire de la population de la Belgique et de ses territoires*, pp. 83–116.

Poulain, M., D. Chambre and M. Foulon. 2001. Survival Among Belgian Centenarians (1870–1894 Cohorts). *Population* 13: 117–138.

Reiss, R.-D., and M. Thomas. 1997. *Statistical Analysis of Extreme Values, with Applications to Insurance, Finance, Hydrology and Other Fields.* Basel: Birkhauser.

Robine, J. M., A. Cournil, J. Gampe and J. W. Vaupel. 2005. IDL, the International Database on Longevity. *Living to 100 and Beyond Symposium Monograph.* Chicago: Society of Actuaries.

Scarrott, C., and A. MacDonald. 2012. A Review of Extreme Value Threshold Estimation and Uncertainty Quantification. *REVSTAT-Statistical Journal* 10: 33–60.

Wang J.-L., H.-G. Muller and B. W. Capra. 1998. Analysis of Old-Oldest Mortality: Lifetables Revisited. *Annals of Statistics* 26: 126–163.

Watts, K., D. Dupuis and B. Jones. 2006. An Extreme Value Analysis of Advanced Age Mortality Data. *North American Actuarial Journal* 10: 162–178.

Wilmoth, J. R., and J. M. Robine. 2003. The World Trend in Maximum Life Span. *Population and Development Review* 29: 239–257.