

Article

Nonparametric Estimation of Extreme Quantiles with an Application to Longevity Risk

Catalina Bolancé  and Montserrat Guillen * 

Department Econometrics, Riskcenter-IREA, Universitat de Barcelona, 08007 Barcelona, Spain; bolance@ub.edu

* Correspondence: mguillen@ub.edu; Tel.: +34-934-037-039

Abstract: A new method to estimate longevity risk based on the kernel estimation of the extreme quantiles of truncated age-at-death distributions is proposed. Its theoretical properties are presented and a simulation study is reported. The flexible yet accurate estimation of extreme quantiles of age-at-death conditional on having survived a certain age is fundamental for evaluating the risk of lifetime insurance. Our proposal combines a parametric distributions with nonparametric sample information, leading to obtain an asymptotic unbiased estimator of extreme quantiles for alternative distributions with different right tail shape, i.e., heavy tail or exponential tail. A method for estimating the longevity risk of a continuous temporary annuity is also shown. We illustrate our proposal with an application to the official age-at-death statistics of the population in Spain.

Keywords: extreme quantile; kernel estimation; extreme value distribution; lifetime annuity



Citation: Bolancé, Catalina, and Montserrat Guillen. 2021. Nonparametric Estimation of Extreme Quantiles with an Application to Longevity Risk. *Risks* 9: 77. <https://doi.org/10.3390/risks9040077>

Academic Editor: Mogens Steffensen

Received: 4 March 2021

Accepted: 12 April 2021

Published: 15 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Longevity risk, the possibility that individuals die much later than the theoretical average, is covered by a range of insurance products. A primary example is lifetime annuities, which pay a guaranteed income even if individuals live longer than expected and exhaust their own savings. Measuring longevity risk in lifespan products of this type requires establishing assumptions or, at least, choosing a specific loss variable and a concrete risk measure, such as value-at-risk (VaR).

Here, we propose a flexible nonparametric estimator that is used to measure longevity risk with the (extreme) quantile of the age-at-death distribution. Our nonparametric approach is a generalization of the proposal by [Alemany et al. \(2013\)](#) that was based on only two given transformations. [Alemany et al. \(2013\)](#) proposed the transformed kernel estimation of the cumulative distribution function and the quantile. In order to transform the data, a parametric distribution function was proposed and then the bias and variance of the transformed kernel estimator were obtained. The general result established that with the transformed kernel estimation approach, the variance of the estimator diminishes at the cost of increasing its bias. [Alemany et al. \(2013\)](#) argued that, since the transformed variable follows a uniform distribution, then boundary bias emerges naturally. To reduce the bias, they proposed a second transformation based on the inverse Beta distribution. However, they did not study how this second transformation affects the bias and variance of the transformed kernel estimator. Here, we start from the initial results by [Alemany et al. \(2013\)](#) and analyze the role of the second transformation. Furthermore, we have deduced the properties of the estimator and have proved that the second transformation can increase or decrease the bias associated with the first transformation that will be able to be selected between alternative families of distributions. We argue that by doing so, we are able to reduce the model selection problem and, thus, eliminate the need to choose a specific parametric model which may be too restrictive for older ages, that is, “rare longevity events” corresponding to individuals that live much longer than the rest.

In Section 2, the literature on longevity risk is gathered. In Section 3, the estimator of the conditional quantile and its properties are shown. In Section 4, we describe our

proposed transformed kernel estimator of the cumulative distribution function and present interesting findings about its performance and about the smoothing parameter. Additionally, basic notions about the empirical distribution and kernel estimation are outlined. In Section 5, we estimate conditional quantiles for each of the years in the period 2011–2017 for the population aged 65 and over and apply these results to split the price of an annuity into two segments: that is, the price up to the longevity risk age and the price beyond that age. By so doing, we express the age risk measure in terms of an annuity price in monetary units. Our conclusions are outlined in Section 6. Proofs of the asymptotic properties of our estimator and the results of a simulation study analyzing finite sample properties are presented in the Appendix, along with a table showing the results of the application to the pricing of a single annuity.

2. Longevity Risk

Longevity risk, as a broad concept, has been widely studied at the aggregate level and has been associated with the forecasting of such demographic variables as life expectancy and mortality. Stallard (2006) differentiated between aggregate and individual longevity risk and proposed approaching the latter using individual survival functions, given a set of covariates related to the factors that affect mortality and life expectancy; however, implementing this proposal is far from straightforward as individual information on habits and health is generally unavailable. Denuit (2009) proposed estimating aggregate longevity risk from an extreme quantile of the life expectancy distribution which can be deduced from the Lee–Carter model (see Lee and Carter 1992). Basellini and Camarda (2019) proposed a parametric model to forecast the age-at-death of the adult population in Japan, considering ages from 30 to 110 and beyond. More recently, using compositional data, Shang and Haberman (2020) forecast the whole age-at-death distribution and used these results to obtain the t year survival probabilities at age x (${}_tP_x$), as well as evaluating the “single-premium temporary immediate annuity”. Here, we show that, using the forecast distribution, our nonparametric estimator can also be used to obtain a smooth estimation of these probabilities. In short, all these perspectives, in estimating longevity risk for populations at retirement age, convey just how important it is for public and private institutions alike to establish reserves for lifetime contracts and guarantee the sustainability of public and private pensions, and other health and long-term care systems.

Unlike the studies cited above, our aim is to obtain a longevity risk measure that can be expressed as a value in years and which corresponds to an advanced risk age, above which the probability of death is low or very low. In line with our definition, the higher this “longevity risk age” value, the greater the risk, since it can be interpreted directly as an age that a given fraction of individuals will survive. For example, if our quantile level is fixed at 99% and the measure of longevity risk equals 101 years, then this means that 1% of the population will survive beyond the age of 101.

Typically, the number of observations in this range of older ages is small and disperse, i.e., we have very few data. To address this, we propose an estimator for this extreme quantile that adapts to a situation in which statistical information is scarce in the right tail of the distribution. It should be noted, at this point, that in finance and insurance, the notion of the extreme quantile and VaR are used interchangeably.

Chen and Cummins (2010) argued that some events can be extremely rare and extreme value distributions are necessary to model the improvement in mortality rates. However, the use of these distributions on the whole domain of the variable to estimate the VaR of the age-at-death distribution can give longevity risk values of over 150 years. An alternative is to focus on the extreme ages to fit the shape of the tail of the distribution based on the Generalized Pareto Distribution (GPD). Coles et al. (2001) described the semiparametric method based on the Mean Residual Life plot to obtain a threshold and then estimate shape and scale parameters for the GPD. However, catastrophic events can cause major changes in the shape of the age-at-death distribution, so that distributions that have provided a good fit over a long period of time stop doing so. An example is found in the current COVID

pandemic. The proposed nonparametric estimator can easily be adapted to changing longevity scenarios.

Based on the shape parameter of the GPD, [Rootzén and Zholud \(2017\)](#) analysed if age-at-death could be a limit or not (see [Rootzén and Zholud 2018](#), for discussion). Here compare results obtained with semiparametric GPD for extreme ages, the Gompertz distribution, the generalized Gompertz distribution and our proposal to estimate extreme quantiles. We think that a nonparametric approach constitutes a suitable alternative; yet, given the small number of advanced ages observed, the direct implementation of basic nonparametric methods may not be efficient. Thus, we develop a procedure based on both parametric and kernel estimation methods to evaluate longevity risk using the cumulative distribution function of the age-at-death variable. We show that our proposal improves the inefficiency of using the empirical distribution or the classical kernel estimator, while coping with the bias of choosing possibly inadequate parametric distributions.

Specifically, we employ the conditional quantile because we estimate longevity risk from an extreme quantile of the age-at-death variable for a given population that has already survived to age a . This is the most common situation in the case of insurance products covering lifetime resources after retirement, which usually means after the age of 65. [Stallard \(2006\)](#) explained that at the individual level “longevity risk refers to the possibility of living longer than assumed in financial planning for the retirement of a single individual”. It is upon this idea that the conditional extreme quantile is based. Conditional extreme quantiles are convenient to evaluate the effect of longevity risk on the unitary price of a lifetime and temporary annuities, where the former consider individuals throughout their lifespan for the probabilities that, at a certain age, they live $t \geq 0$ more years and the latter depend on a threshold (the maturity) from which the probability of living is assumed to be zero (see [Bowers et al. 1997](#); [Dickson et al. 1997](#), for a review in relation to life insurance). As we show in the last section of this paper, relating maturity to the conditional extreme quantile allows us to quantify and interpret longevity risk in lifetime annuities.

We illustrate the methodology with official micro-data containing the dates of birth and death of the deceased population in Spain between 2011 and 2017. The study is conducted differentiating by gender. In general, women have greater longevity than men, although these differences may disappear with age in some countries. For example, drawing on data from England and Wales, [Mayhew and Smith \(2014\)](#) showed that the gap between the life expectancy of men and women tends to close. Earlier, basing their arguments on reliability theory, [Gavrilov and Gavrilova \(2001\)](#) explained why relative differences in the mortality rates of populations within the same species and, hence, differences in their life expectancies, tend to disappear with age. From an empirical perspective, [Waldron \(1993\)](#), [Gjonça et al. \(2005\)](#) and [Glei and Hirouchi \(2007\)](#) analyse the gender gap between life expectancies in different periods. These authors show that there are countries in which the gender longevity gap narrowed and others in which it widened, depending on the period under analysis (see also [Mayhew and Smith 2011](#)). The results presented here focus on Spain’s older population from 2011 to 2017 and show that the gap between male and female longevity risk exists and does not fade until a very advanced age is reached.

3. The Conditional Quantile and Longevity Risk

Let X be a continuous random variable with probability density function (pdf) f_X and cumulative distribution function (cdf) F_X . In this paper, X refers to the age-at-death of an individual in a given population. In such cases, X takes positive values. Given a subpopulation with an age of death equal to or greater than a , we propose a longevity risk measure obtained from the truncated cdf, which is equal to the conditional quantile (CQ) and is defined as:

$$CQ_{p_a}(X|X \geq a) = \inf \left\{ x : F_{X|X \geq a}(x|X \geq a) \geq p_a \right\} = x_{p_a}, \quad (1)$$

where a is a threshold and $p_a \in (0, 1)$ is the conditional confidence level. The truncated cdf is:

$$F_{X|X \geq a}(x|X \geq a) = \frac{F_X(x) - F_X(a)}{1 - F_X(a)}, \quad x \geq a. \quad (2)$$

To estimate the CQ defined in (1) the theoretical truncated cdf needs to be replaced by an estimator. Given that we need to focus on the quantile at confidence level $p_a = F_{X|X \geq a}(x_{p_a}|X \geq a)$, with this aim in mind, we can relate the probability associated with the truncated cdf with the probability associated with the untruncated cdf:

$$F_X(x) = F_{X|X \geq a}(x|X \geq a)[1 - F_X(a)] + F_X(a). \quad (3)$$

Thus, from expression (3) and given a known confidence level p_a :

$$p = p_a[1 - F_X(a)] + F_X(a), \quad (4)$$

where $p \geq p_a$. In addition to this empirical approach, alternative kernel estimators (KEs) of p are analysed, with the aim of finding the simplest and most efficient proposal.

To obtain the CQ, our strategy involves estimating $F_X(a)$, calculating the unconditional probability p , and deriving the quantile using the untruncated cdf. In practice, since $p \geq p_a$, this requires estimating an extreme quantile. Furthermore, this strategy is easily applicable if, as we propose in this paper, the quantile estimator is obtained by inverting an estimator of the cdf.

The probability p expressed in (4) is obtained by replacing $F_X(a)$ with the empirical distribution or a KE, i.e.,

$$\hat{p} = p_a[1 - \hat{F}_X(a)] + \hat{F}_X(a), \quad (5)$$

where \hat{F}_X may be the empirical distribution, the KE (see [Azzalini 1981](#)) or a transformed KE (TKE) (see [Alemany et al. 2013](#); [Bolancé et al. 2003](#); [Swanepoel and Van Graan 2005](#); [Wand et al. 1991](#), for a review on transformed kernel estimation).

Proposition 1. *The expected value of \hat{p} is:*

$$E(\hat{p}) = p + (1 - p_a) \text{Bias}[\hat{F}_X(a)]$$

and the variance is:

$$V(\hat{p}) = (1 - p_a)^2 V[\hat{F}_X(a)],$$

where $\text{Bias}[\hat{F}_X(a)]$ and $V[\hat{F}_X(a)]$ are those of the corresponding nonparametric estimator.

The proof of Proposition 1 is in Appendix A. This proposition shows how the bias and variance of \hat{p} are proportional to those of the estimator used for the unconditional cdf evaluated at a , $F_X(a)$.

Furthermore, given the properties of the KE (which we show in the following section), as $n \rightarrow \infty$, then $\text{Bias}[\hat{F}_X(a)] \rightarrow 0$ and, asymptotically,

$$\hat{p} \sim \text{Normal}\left(p, (1 - p_a)^2 V[\hat{F}_X(a)]\right)$$

and, therefore, $CQ_{p_a}(X|X \geq a)$ is normally distributed (see, for example, [Reiss 1981](#)).

The CQ provides us with a value of the age-at-death from which there is a probability $(1 - p_a)$ that this age is higher. This allows us to evaluate the risk for a life insurance product, with a confidence level p_a , once a given age has been reached.

4. Estimators of the Unconditional cdf

Let X_1, \dots, X_n be a sample of independent and identically distributed data observations of the continuous random variable X with cdf F_X . Now, we summarise the main properties associated with the empirical distribution and the KE. The empirical distribution is:

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x), \tag{6}$$

where $I(\cdot)$ is an indicator function that takes a value of 1 if the condition between parentheses is true and 0 otherwise. The bias of \hat{F}_n is zero and its variance is:

$$V[\hat{F}_n(x)] = F_X(x)[1 - F_X(x)]/n. \tag{7}$$

We note that the empirical distribution is only defined in the observed values in the sample.

The usual expression for the KE of a cdf is (see [Azzalini 1981](#); [Reiss 1981](#), for a first definition):

$$\hat{F}_X(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - X_i}{b}\right), \tag{8}$$

where the function K is a cdf associated with a kernel pdf k . Examples of such functions are the Epanechnikov and the Gaussian kernels ([Silverman 1986](#)). The parameter $b > 0$, known as the bandwidth or smoothing parameter, controls the smoothness of the estimation. Thus, the higher the value of b , the smoother the resulting function. However, obtaining an optimal value for this smoothing parameter is one of the greatest difficulties posed by kernel estimation.

Note that the classical kernel estimation of cdf defined in (8) bears many similarities to the expression of the well-known empirical distribution in (6). In (8), $K\left(\frac{x - X_i}{b}\right)$ has to be replaced by $I(X_i \leq x)$ to obtain (6). Focusing on the second-order properties, it has already been noted by [Reiss \(1981\)](#) and [Azzalini \(1981\)](#) that, when $b \rightarrow 0$ as $n \rightarrow \infty$, f_X is continuous and f'_X exists, the asymptotic bias and variance of (8) are (see also [Hill 1985](#)):

$$E[\hat{F}_X(x)] - F(x) = \frac{1}{2}b^2 f'_X(x) \int_{-1}^1 t^2 k(t) dt + o(b^2), \tag{9}$$

and

$$V[\hat{F}_X(x)] = \frac{F_X(x)[1 - F_X(x)]}{n} - f_X(x) \frac{b}{n} \int_{-1}^1 K(t)[1 - K(t)] dt + o\left(\frac{b}{n}\right). \tag{10}$$

The KE of the cdf has less variance than that of the empirical distribution, but it has some bias which tends to zero as $n \rightarrow \infty$.

When the distribution is right-skewed, at the extreme quantile with p near 1 the variance of the KE tends to the variance of the empirical quantile.

Let $T(\cdot)$ be a continuous and monotonic non-decreasing transformation, $Y = T(X)$ is the transformed random variable and $Y_i = T(X_i), i = 1 \dots n$, are the transformed data, the TKE of $F_X(x)$ is:

$$\begin{aligned} \hat{F}_X^T(x) = \hat{F}_{T(X)}[T(x)] &= \frac{1}{n} \sum_{i=1}^n K\left(\frac{T(x) - T(X_i)}{b}\right) \\ &= \frac{1}{n} \sum_{i=1}^n K\left(\frac{y - Y_i}{b}\right) = \hat{F}_Y(y). \end{aligned} \tag{11}$$

In practice, the estimator defined in (11) consists of transforming the data and obtaining the KE of the transformed variable. The smoothing parameter in the TKE is the same as that in the KE of the cdf associated with the transformed data. To calculate the KE, we need

to determine the kernel function K and the value of the smoothing parameter b . For the former, we use the Epanechnikov kernel ($K(t) = \frac{1}{4}(2-t)(t+1)^2$), although alternative kernels, such as the Gaussian, hardly affect the estimator properties. In the case of the latter, there are alternative methods that can be employed to calculate the smoothing parameter b . For instance, cross-validation and plug-in methods are common (see, for example, [Altman and Léger 1995](#); [Bowman et al. 1998](#)). However, they require considerable computational effort in large data sets and do not perform particularly well when there are parts of the distribution domain for which data are scarce.

Many studies propose transformations in the context of the transformed kernel estimation of the pdf (see [Bolancé 2010](#); [Bolancé et al. 2003, 2008](#); [Buch-Larsen et al. 2005](#); [Ruppert and Cline 1994](#); [Wand et al. 1991](#)); however, only a few analyse the transformed kernel estimation of the cdf and quantile (see [Alemany et al. 2013](#); [Swanepoel and Van Graan 2005](#)). These transformations can be classified into those that are a cdf and those that do not correspond to a specific cdf. However, if a cdf transformation is used, the theoretical properties of the estimator can be analyzed in terms of the differences between a true cdf and those used as the transformation.

Assume that $T(\cdot)$ is continuous and that its first two derivatives exist. Then, from Theorem 1 in ([Alemany et al. 2013](#)), it follows that ($'$ indicates first derivative):

$$E[\widehat{F}_X^T(x)] - F(x) = \frac{1}{T'(x)} \left(1 - \frac{\frac{f_X(x)}{T'(x)}}{\frac{f_X(x)}{T'(x)}} \right) \left(\frac{1}{2} f_X'(x) \int_{-1}^1 t^2 k(t) dt \right) b^2 + o(b^2) \quad (12)$$

and

$$V[\widehat{F}_X^T(x)] \approx \frac{F_X(x)[1 - F_X(x)]}{n} - \frac{f_X(x)}{T'(x)} \frac{b}{n} \int_{-1}^1 K(t)[1 - K(t)] dt + o\left(\frac{b}{n}\right), \quad (13)$$

where $o(\cdot)$ is an asymptotic term that tends to zero with its argument.

If $b \rightarrow 0$ when $n \rightarrow \infty$, then the TKE in (11) is a consistent estimator of $F_X(x)$. Furthermore, in the case where $F_X(x) = T(x)$, $f_X(x) = T'(x)$ and $f_X'(x) = T''(x)$, then the bias of TKE (12) is zero.

If the transformed random variable Y follows a Uniform(0,1) distribution then $F_Y''(y) = f_Y'(y) = 0$ and, therefore, any results based on a second-order approach like (12) and (13) cannot be obtained. We propose to choose T so that the distribution of the transformed variable is different from the Uniform(0,1) distribution and so that it can be easily estimated using the KE. [Alemany et al. \(2013\)](#), following ([Bolancé et al. 2008](#); [Buch-Larsen et al. 2005](#)), propose using a first transformation equal to the modified Champernowne distribution (see [Buch-Larsen et al. 2005](#)) and a second transformation equal to the inverse of a given *Beta*(3,3) distribution.

The proposal of ([Alemany et al. 2013](#)) is presented as a particular case of the TKE with double transformation. However, these authors do not analyze the influence of the second *Beta*(3,3) transformation on the resulting estimation. Here we generalize this TKE and obtain the theoretical properties of this new estimator of the cdf. These new results are shown below.

The procedure is based on the expression of bias of the KE shown below in (9) which squared and integrated is:

$$\int_0^\infty \left\{ E[\widehat{F}_X(x)] - F(x) \right\}^2 dx = \frac{1}{4} b^4 \int_0^\infty [f_X'(x)]^2 dx \left(\int_{-1}^1 t^2 k(t) dt \right)^2 + o(b^4). \quad (14)$$

Expression (14) depends on the shape of the density through $\int_0^\infty [f_X'(x)]^2 dx$. [Terrell \(1990\)](#) proved that the density that minimises this integral belongs to the Beta distribution family with pdf (see [Johnson et al. 1995](#)):

$$\frac{(x - \theta)^{\alpha-1}(\theta + D - x)^{\beta-1}}{B(\alpha, \beta)D^{\alpha+\beta-1}}, \theta \leq x \leq \theta + D, \tag{15}$$

where $\theta \geq 0$ is a location parameter, $D > 0$ is a scale parameter, $\alpha > 0$ and $\beta > 0$ are shape parameters, $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$ and $\Gamma(\cdot)$ is Euler’s Gamma function. Terrell (1990) proved that the pdf expressed in (15) with $\theta = -1, D = 2$ and $\alpha = \beta = 3$ is equal to the *Beta(3,3)* which takes the smallest value of $\int_0^\infty [f'_X(x)]^2 dx$ among all the distributions with domain $[-1, 1]$ and the same variance; therefore, given b and k , it takes the smallest value of the integrated bias defined in (14). Previously, Terrell and Scott (1985) deduced that when $\theta = -1/2, D = 1, \alpha = \beta = 2$ the density minimises $\int_0^\infty [f'_X(x)]^2 dx$ for all distributions with a domain $[-1/2, 1/2]$ and the same variance.

Definition 1. Let $T(\cdot)$ be a first transformation coinciding with a cdf and $M(\cdot)$ be the *Beta(3,3)* cdf of the second transformation; then the *Beta Transformed Kernel Estimator (BTKE)* of the original cdf is defined as:

$$\begin{aligned} \widehat{F}_X^{BT}(x) &= \widehat{F}_{M^{-1}[T(x)]}\{M^{-1}[T(x)]\} \\ &= \frac{1}{n} \sum_{i=1}^n K\left(\frac{M^{-1}[T(x)] - M^{-1}[T(X_i)]}{b}\right), \end{aligned} \tag{16}$$

where the cdf associated with the *Beta(3,3)* is:

$$M(y) = \frac{3}{16}y^5 - \frac{5}{8}y^3 + \frac{15}{16}y + \frac{1}{2}, -1 \leq y \leq 1. \tag{17}$$

The BTKE has suitable properties to estimate the extreme quantile because, among other characteristics (see Bolancé et al. 2020), as we describe in Section 4.1, an optimal smoothing parameter based on the *Beta(3,3)* distribution can be found.

Theorem 1, whose proof is given in Appendix A, gives expressions for the asymptotic bias and variance of the BTKE in (16).

Theorem 1. Let $f_X(x)$ be the true pdf that is continuous and has at least one continuous derivative, $T(x)$ be a cdf with continuous pdf and at least two continuous non-null derivatives and, for $y = T(x)$, $m(y) = \frac{15}{16}(1 - y^2)^2, -1 < y < 1$ be the pdf of $M(y)$ defined in (17), then the asymptotic bias of BTKE of the true cdf $F_X(\cdot)$ is:

$$\begin{aligned} &E[\widehat{F}_X^{BT}(x)] - F_X(x) \\ &= \frac{1}{2}b^2 \left[\frac{m'\{M^{-1}[T(x)]\}}{m\{M^{-1}[T(x)]\}} f_X(x) + \frac{f'_X(x)}{T'(x)} m\{M^{-1}[T(x)]\} \left(1 - \frac{f_X(x)}{\frac{f'_X(x)}{T'(x)}}\right) \right] \int_{-1}^1 t^2 k(t) dt + o(b^2) \end{aligned} \tag{18}$$

and its variance is:

$$\begin{aligned} &V[\widehat{F}_X^{BT}(x)] \\ &= \frac{F_X(x)[1 - F_X(x)]}{n} - \frac{f_X(x)}{T'(x)} m\{M^{-1}[T(x)]\} \frac{b}{n} \int_{-1}^1 K(t)[1 - K(t)] dt + o\left(\frac{b}{n}\right). \end{aligned} \tag{19}$$

From the results of Theorem 1, we can analyze the bias and variance of the BTKE and obtain Proposition 2, whose proof is direct and, hence, is not given.

Proposition 2. If the generating process or the distribution of the data is known and the first transformation $T(\cdot)$ can be estimated so that $E[\widehat{T}(\cdot)] = F_X(\cdot)$ or, at the point evaluated x , $f_X(x) = E[\widehat{T}'(x)]$ and $f'_X(x) = E[\widehat{T}''(x)]$, the bias of the BTKE for a given smoothing parameter b and kernel k is:

$$\begin{aligned}
 & E\left[\widehat{F}_X^{BT}(x)\right] - F(x) \\
 &= \left(\frac{m'\{M^{-1}[T(x)]\}}{m\{M^{-1}[T(x)]\}} \frac{b^2}{2} \int_{-1}^1 t^2 k(t) dt\right) f_X(x) + o(b^2) \tag{20}
 \end{aligned}$$

and the variance is:

$$\begin{aligned}
 & V\left[\widehat{F}_X^{BT}(x)\right] \\
 &= \frac{F_X(x)[1 - F_X(x)]}{n} - m\{M^{-1}[T(x)]\} \frac{b}{n} \int_{-1}^1 K(t)[1 - K(t)] dt + o\left(\frac{b}{n}\right). \tag{21}
 \end{aligned}$$

From Proposition 2, we can see that in the best case our estimation has some bias that tends to zero as $n \rightarrow \infty$; however, this bias can be compensated by a low variance, especially in the most extreme quantiles. Alternatively, [Swanepoel and Van Graan \(2005\)](#) and [Alemamy et al. \(2013\)](#) proved that the TKE is unbiased when $T(\cdot) = F_X(\cdot)$; so, given the difficulties associated with the kernel estimation of the Uniform(0, 1) distribution function, the TKE based on a parametric or nonparametric cdf transformation cannot be used to estimate extreme CQs.

If the first transformation $T(\cdot)$ is not equal to the true $F_X(\cdot)$ or, at the extreme quantile evaluated, $x = F_X^{-1}(\alpha)$, $f_X(x) \neq T'(x)$ and $f'_X(x) \neq T''(x)$, we obtain that the sign of the bias in (19) depends on:

$$\left(1 - \frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}}\right). \tag{22}$$

For extreme quantiles, the value of (22) will be positive or negative depending on whether the distribution used as the first transformation has a heavier or lighter tail than the true distribution. In this way we obtain two different results:

- If $T(\cdot)$ is the cdf of a heavier tailed distribution than $F_X(\cdot)$ the bias is the sum of two negative terms. The results are greater than the CKE overestimation of the quantile.
- If $T(\cdot)$ is the cdf of a lighter tailed distribution than $F_X(\cdot)$ the bias is the sum of a negative term and a positive term associated with the positive sign of (22). In this case, we could overestimate or subestimate the quantile; even, the two terms can be compensated.

For the variance, if $T(\cdot)$ is the cdf of a lighter tailed distribution than $F_X(\cdot)$, so $\frac{f_X(x)}{T'(x)} \geq 1$ and the variance of BTKE will be smaller than the variance of KE. Alternatively, if $T(\cdot)$ is the cdf of a heavier tailed distribution than $F_X(\cdot)$, so $0 < \frac{f_X(x)}{T'(x)} \leq 1$ and the result will depend on this quotient, the less it is the greater the variance is.

4.1. Smoothing Parameter of BTKE

If we assume $F_X(\cdot) = T(\cdot)$, the transformed data in the BTKE follow the *Beta(3,3)* distribution; then, the optimal bandwidth can be obtained by minimizing the asymptotic MSE, that is equal to the sum of the variance and the squared bias of the KE of the transformed variable, eliminating the asymptotic order terms:

$$\begin{aligned}
 A - MSE\left[\widehat{F}_{M^{-1}[T(X)]}\left\{M^{-1}[T(x_p)]\right\}\right] &= \frac{1}{4} b^4 \left[m'\{M^{-1}[T(x_p)]\} \int_{-1}^1 K(t)[1 - K(t)] dt t^2 k(t) dt \right]^2 \\
 &+ \frac{F_{M^{-1}[T(X)]}\left\{M^{-1}[T(x_p)]\right\} \left[1 - F_{M^{-1}[T(X)]}\left\{M^{-1}[T(x_p)]\right\}\right]}{n} \\
 &- m\{M^{-1}[T(x_p)]\} \frac{b}{n} \int_{-1}^1 K(t)[1 - K(t)] dt. \tag{23}
 \end{aligned}$$

The resulting smoothing parameter is:

$$b_p = \left(\frac{m\{M^{-1}[T(x_p)]\} \int_{-1}^1 K(t)[1-K(t)]dt}{\left(|m'\{M^{-1}[T(x_p)]\} \int_{-1}^1 t^2 k(t)dt\right)^2} \right)^{\frac{1}{3}} n^{-\frac{1}{3}}, \quad (24)$$

where $|\cdot|$ is absolute value.

The smoothing parameter b_p in expression (24) minimises the MSE of the BTKE of the quantile x_p if the true cdf fulfills $F_X(x_p) = T(x_p)$. However, if $F_X(x_p) \neq T(x_p)$ then, as we show in Remark 1, the bandwidth expressed in (24) could be lower or higher than the optimal.

Remark 1. Let M be the Beta(3, 3) cdf and b_p^* the optimal bandwidth when $F_X(x_p) = T(x_p)$ and let x_p be the p th quantile of the random variable X , then:

- If T is heavier tailed than F_X then $T(x_p) \leq F_X(x_p)$, $M^{-1}[T(x_p)] \leq M^{-1}(p)$ and $b_p \geq b_p^*$.
- If F_X is heavier tailed than T then $T(x_p) \geq F_X(x_p)$, $M^{-1}[T(x_p)] \geq M^{-1}(p)$ and $b_p \leq b_p^*$.

A higher b_p provides a higher estimated quantile and vice versa. In practice, Remark 1 implies that the difference between F_X and T increases the bias when $T(x_p) < F_X(x_p)$ and increases the variance when $T(x_p) > F_X(x_p)$, although both tend to zero as $n \rightarrow \infty$.

In Appendix B, we present a simulation study comparing the MSE of the estimated CQ_{p_a} based on the empirical distribution, the KE and the BTKE. We simulated alternative distributions with different tail shapes: a long-tail with an exponential and subexponential decrease (Weibull and Lognormal, respectively) and a heavy tail with a power decrease (Pareto).

5. Data Analysis

We analyze the longevity risk of the population aged over 65 in Spain from 2011 to 2017. For this population, we fit the distribution of the age-at-death in months and we estimate extreme quantiles and conditional extreme quantiles with $p_a = 0.99$, $p_a = 0.995$ and $p_a = 0.999$, which we also denominate as VaR at $p_a \times 100$ confidence level. Since we have a large number of observations, we can estimate quantiles based on the BTKE with p_a very close to 1 quite accurately.

The data were provided by the Spanish National Institute of Statistics (INE—Instituto Nacional de Estadística) and contain information about the gender, the year and month of birth and the year and month of death of the population aged over 65. The information available allows us to obtain the age-at-death in months but, for ease of interpretation, the results are shown in years.

In Table 1 we present the mortality rates and the basic descriptive statistics for the population aged over 65, differentiated by gender. It can be seen that, in aggregate terms, for all years, female mortality rates are lower than those for men and, inversely, the average age-at-death for women is higher. However, if we focus on the maximum age-at-death values, the results are not so clear.

The first step in calculating the longevity risk is to fit alternative parametric distributions to our data. We obtain results for Lognormal, Weibull, Gamma, Generalised Pareto Distribution (GPD) and modified Champernowne, the best fit being provided by this last distribution. Not to confuse the reader, the parameters associated with the first transformation cdf, are not shown here, they are available from the authors. We observe that the three parameters of the modified Champernowne distribution endow this model with more flexibility near the minimum age (65 years) and in the tail of the distribution. However, if we use this distribution to estimate the CQs, in the more extreme cases, we obtain highly unlikely ages-at-death of over 150 years. For this reason, we use the BTKE to correct these unlikely values.

By way of illustration, and using the data analyzed for the year 2015, considering that we have more than 170,000 data items for men and women, in Figure 1 we plot the CQs at the 99.9% level that were estimated using the empirical distribution, the KE and the BTKE.

Table 1. Descriptive statistics of age-at-death (in years) for the population aged 65 or over in Spain (2011).

Gender	Year	Population	Deaths	Rate	Min.	Mean	Max.	Variance
Men	2011	3,768,467	157,757	0.042	65	81.34	111.17	62.22
	2012	3,717,697	164,281	0.044	65	81.62	112.17	62.89
	2013	3,676,272	159,705	0.043	65	81.64	109.08	63.93
	2014	3,613,455	162,066	0.045	65	81.80	109.33	65.47
	2015	3,565,325	173,472	0.049	65	82.09	109.75	65.92
	2016	3,510,560	169,414	0.048	65	82.13	114.75	67.21
	2017	3,449,614	175,086	0.051	65	82.38	109.50	68.20
Women	2011	4,995,737	168,547	0.034	65	85.58	111.33	60.62
	2012	4,940,008	177,790	0.036	65	85.84	111.42	59.59
	2013	4,897,713	171,411	0.035	65	85.87	111.08	60.78
	2014	4,828,972	175,202	0.036	65	86.03	112.83	60.96
	2015	4,770,536	189,889	0.040	65	86.36	111.17	60.35
	2016	4,711,636	182,392	0.039	65	86.39	114.50	61.78
	2017	4,995,737	190,600	0.038	65	86.64	116.17	62.06

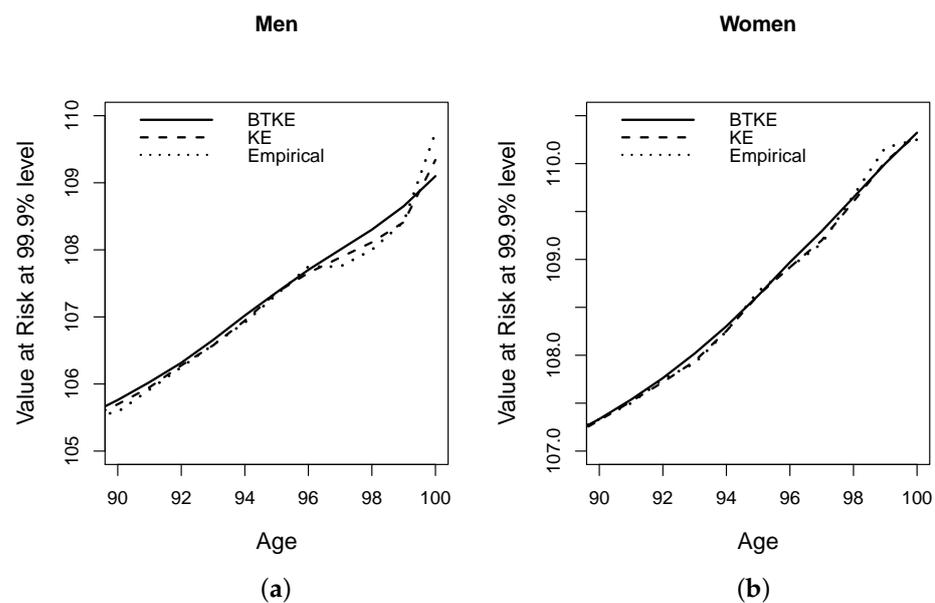


Figure 1. CQ_{p_a} for the year 2015 of the age-at-death variable for the population in Spain estimated from the Beta Transformed Kernel Estimator (BTKE), kernel estimator (KE) and empirical distribution, for men (a) and women (b).

Figure 1 emphasises the smoothed behavior of the estimation based on the BTKE compared to the respective shapes of those based on the empirical distribution and the KE, which are more strongly influenced by the scarce data observed for the largest values of the variable; this captures, perfectly, how the variance of the BTKE is lower than that of the KE, which is very similar to the variance of the empirical distribution.

The Gompertz and the Gompertz-G distributions (see Alizadeh et al. 2017) improve the fit of the modified Champernowne, the improvement is greater for Gompertz-G. The Gompertz has traditionally been used in demography for fitting age-at-death distributions (see, for example, Preston et al. 2001); the Gompertz-G is a generalization where a first cdf G is applied to the data. A particular case is G equal to the log-logistic distribution. In fact, the modified Champernowne distribution defined in expression (A1) with $c = 0$ is

the log-logistic distribution. Therefore, there is a parallelism between our BTKE and the Gompertz-G distribution; both start from similar cdf transformations.

Using the inverse of the estimated cdf, given a probability $\hat{p} = p_a[1 - \hat{F}_X(a)] + \hat{F}_X(a)$, we have estimated the VaR for a grid of values for a between the ages of 65 and 100 and at 99, 99.5 and 99.9% confidence levels, for each year from 2011 to 2017. To compare the results obtained with the BTKE, we consider GPD using the semiparametric estimates described in (Coles et al. 2001), we also consider Gompertz and Gompertz-G, with parameters estimated by maximum likelihood. To carry on this comparison, we estimate $CQ_{p_a}(X)$ and we obtain its empirical significance level, which should be similar to the theoretical $(1 - p_a)$. The quotients between the empirical vs. the theoretical significance levels are shown in Tables A4 and A5 of Appendix C. The results in Table A4 indicate that the BTKE provides the best fit. Results with GPD do not provide good fits. Some years the shape parameter takes a negative value (2013, 2014 and 2015 for men and 2011, 2012 and 2015 for women). In the other years, the parameter is positive although lower than 0.25. In general, the GPD procedure tends to underestimate the conditional quantiles more extreme (quotients greater than 1). In contrast, in Table A5 we show as the Gompertz distribution under-estimates the longevity risk (quotients greater than 1) while the Gompertz-G distribution greatly over-estimates the risk in most cases for men (quotients lower than 1) and provides highly unstable results for women. Some values of the quotient are equal to zero. They are obtained with the BTKE and the Gompertz-G parametric distribution and they indicate that there are no observations larger than the estimated $CQ_{p_a}(X)$. With BTKE this only occurs in the most extreme quantiles for men, i.e., in years 2012, 2013 and 2014, when $a = 100$ and confidence level 0.999. With Gompertz-G this also occurs for men but in lower confidence levels.

A summary of the $CQ_{p_a}(X)$ obtained with the BTKE is shown in Figure 2 and in Tables 2 and 3. Specifically, in Figure 2 we plot these VaRs between 2011 and 2017, differentiating by gender. In general, we observe that the VaR has a slight tendency to increase over the years, disappearing for men beyond age 80, approximately, at confidence levels 99% and 99.5%, and beyond age 90 at confidence level 99.9%. Furthermore, CQ for women are higher than for men in almost all values of a and p_a , although these differences tend to narrow with increasing age. In order to analyse the statistical significance of these differences during the period 2011–2017, in Tables 2 and 3 we present the results for $a = 65, 75, 85, 95$ with the bootstrap confidence intervals (CIs) at 90% based on a Normal distribution. We generate 1000 bootstrap samples with replacement and test how the estimated values in this sample follow a Normal distribution.

Focusing on Tables 2 and 3, where we show the estimated CQ_{p_a} , differences between men and women are observed in all cases, except for the most extreme age and quantile, i.e., $a = 95$ and $p_a = 0.999$.

To analyze how the longevity risk increases for men and women the year 2011 is taken as our reference. Specifically, for each value of a and p_a , we compare the upper limits of that year with the lower limits of the other years. In this way, we can determine whether or not the CIs overlap and, thus, conclude whether the VaR increases with respect to 2011 (no overlap of CIs) or not (overlap of CIs). In Tables 2 and 3, some of the CQ_{p_a} values corresponding to the intervals do not overlap and they indicate that the longevity risk increases more quickly for women than it does for men during the period under consideration (2011–2017).

Table 2. Results of estimated CQ_{p_a} and 90% confidence interval based on Beta Transformed Kernel Estimator (BTKE) for the age-at-death of the population aged over 65 in Spain separately for men and women and conditional ages of the deceased population ($a = 65, 75$), from 2011 to 2017.

p_a	Year	$a = 65$		$a = 75$	
		Men	Women	Men	Women
0.990	2011	98.49 (98.40;98.59)	101.23 (101.15;101.31)	99.08 (98.97;99.18)	101.45 (101.37;101.53)
	2012	98.57 (98.47;98.66)	101.33 (101.25;101.41)	99.11 (99.01;99.21)	101.53 (101.45;101.61)
	2013	98.54 (98.44;98.63)	101.40 (101.31;101.48)	99.14 (99.04;99.25)	101.61 (101.53;101.69)
	2014	98.56 (98.47;98.65)	101.50 (101.42;101.57)	99.14 (99.04;99.24)	101.72 (101.64;101.79)
	2015	98.74 (98.64;98.83)	101.78 (101.70;101.85)	99.29 (99.19;99.38)	101.98 (101.90;102.05)
	2016	98.62 (98.53;98.71)	101.76 (101.68;101.84)	99.19 (99.09;99.29)	101.97 (101.89;102.05)
	2017	98.75 (98.66;98.84)	101.92 (101.84;101.99)	99.31 (99.21;99.40)	102.14 (102.06;102.21)
	0.995	2011	99.99 (99.87;100.10)	102.59 (102.48;102.69)	100.49 (100.37;100.61)
2012		100.04 (99.93;100.15)	102.70 (102.60;102.79)	100.49 (100.37;100.61)	102.89 (102.79;102.99)
2013		100.09 (99.98;100.20)	102.81 (102.71;102.91)	100.56 (100.45;100.68)	103.00 (102.90;103.11)
2014		100.10 (99.99;100.21)	102.85 (102.75;102.94)	100.62 (100.50;100.74)	103.04 (102.95;103.14)
2015		100.24 (100.14;100.35)	103.11 (103.02;103.21)	100.75 (100.63;100.86)	103.31 (103.21;103.40)
2016		100.14 (100.03;100.26)	103.13 (103.04;103.23)	100.66 (100.54;100.78)	103.33 (103.23;103.43)
2017		100.26 (100.15;100.37)	103.28 (103.19;103.38)	100.76 (100.64;100.88)	103.47 (103.38;103.57)
0.999		2011	102.89 (102.69;103.09)	105.31 (105.13;105.50)	103.27 (103.05;103.49)
	2012	102.90 (102.70;103.10)	105.44 (105.26;105.61)	103.30 (103.09;103.52)	105.59 (105.41;105.77)
	2013	103.06 (102.86;103.26)	105.40 (105.21;105.58)	103.46 (103.24;103.68)	105.56 (105.37;105.75)
	2014	103.16 (102.96;103.35)	105.63 (105.47;105.80)	103.57 (103.36;103.78)	105.82 (105.65;106.00)
	2015	103.26 (103.06;103.45)	105.89 (105.72;106.05)	103.68 (103.47;103.89)	106.03 (105.86;106.20)
	2016	103.24 (103.04;103.43)	105.89 (105.72;106.05)	103.67 (103.46;103.88)	106.03 (105.86;106.20)
	2017	103.19 (102.99;103.38)	106.01 (105.84;106.18)	103.63 (103.42;103.85)	106.16 (105.98;106.33)

Table 3. Results of estimated CQ_{p_a} and 90% confidence interval based on Beta Transformed Kernel Estimator (BTKE) for the age-at-death of the population aged over 65 in Spain separately for men and women and conditional ages of the deceased population ($a = 85, 95$), from 2011 to 2017.

p_a	Year	$a = 85$		$a = 95$	
		Men	Women	Men	Women
0.990	2011	100.72 (100.60;100.84)	102.35 (102.26;102.45)	104.27 (103.99;104.55)	105.25 (105.08;105.42)
	2012	100.64 (100.52;100.76)	102.39 (102.30;102.49)	104.57 (104.29;104.85)	105.36 (105.20;105.53)
	2013	100.68 (100.56;100.80)	102.51 (102.41;102.60)	104.47 (104.19;104.75)	105.30 (105.13;105.48)
	2014	100.68 (100.56;100.79)	102.52 (102.43;102.61)	104.67 (104.40;104.93)	105.51 (105.35;105.66)
	2015	100.73 (100.62;100.85)	102.72 (102.63;102.81)	104.73 (104.47;105.00)	105.70 (105.54;105.85)
	2016	100.59 (100.48;100.71)	102.71 (102.61;102.80)	104.61 (104.34;104.89)	105.65 (105.49;105.81)
	2017	100.62 (100.50;100.74)	102.84 (102.75;102.92)	104.52 (104.25;104.78)	105.65 (105.49;105.81)
	0.995	2011	101.98 (101.82;102.15)	103.66 (103.53;103.78)	105.33 (104.96;105.71)
2012		101.86 (101.71;102.02)	103.66 (103.55;103.78)	105.95 (105.57;106.32)	106.37 (106.15;106.58)
2013		101.95 (101.79;102.11)	103.70 (103.58;103.83)	105.32 (104.93;105.70)	106.31 (106.09;106.54)
2014		102.01 (101.85;102.17)	103.78 (103.67;103.89)	105.58 (105.21;105.94)	106.71 (106.51;106.92)
2015		102.04 (101.88;102.20)	104.07 (103.95;104.18)	105.58 (105.22;105.95)	106.68 (106.48;106.88)
2016		101.96 (101.80;102.12)	104.00 (103.88;104.11)	105.66 (105.28;106.04)	106.63 (106.43;106.84)
2017		101.91 (101.76;102.07)	104.10 (103.99;104.22)	105.71 (105.35;106.08)	106.62 (106.42;106.83)
0.999		2011	104.34 (104.06;104.63)	106.19 (105.98;106.41)	108.47 (107.61;109.32)
	2012	104.66 (104.37;104.94)	106.22 (106.01;106.44)	109.40 (108.56;110.25)	108.51 (108.07;108.94)
	2013	104.46 (104.17;104.75)	106.18 (105.96;106.40)	107.16 (106.29;108.03)	108.09 (107.65;108.52)
	2014	104.65 (104.37;104.92)	106.55 (106.35;106.75)	107.58 (106.75;108.42)	109.18 (108.76;109.60)
	2015	104.78 (104.50;105.05)	106.57 (106.37;106.77)	107.37 (106.53;108.20)	108.62 (108.21;109.04)
	2016	104.68 (104.40;104.96)	106.56 (106.36;106.76)	107.80 (107.00;108.61)	108.89 (108.47;109.30)
	2017	104.69 (104.42;104.96)	106.62 (106.41;106.82)	107.43 (106.62;108.23)	108.51 (108.10;108.92)

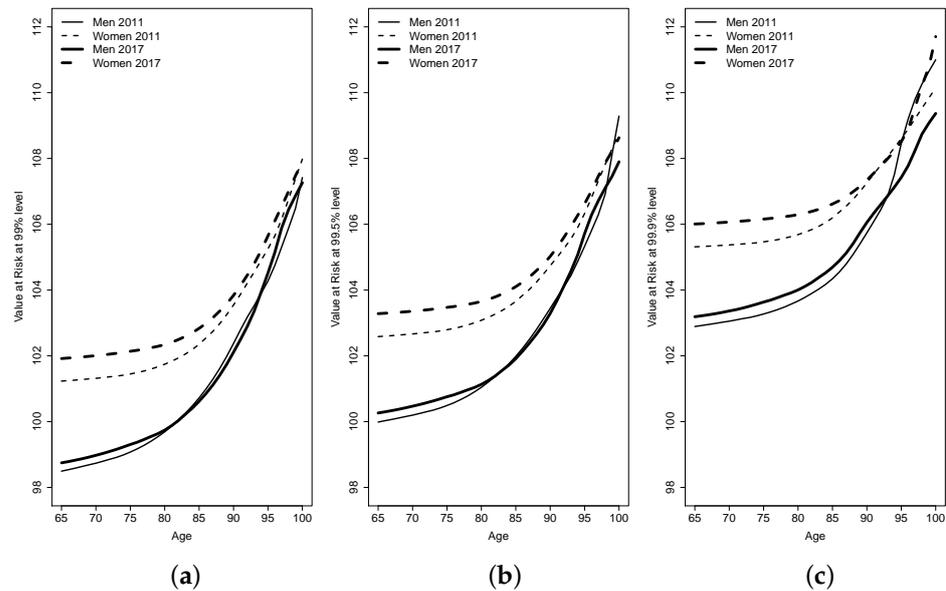


Figure 2. CQ_{p_a} of year 2011 (thin lines) and 2017 (thick lines) of the age-at-death variable estimated from Beta Transformed Kernel Estimator (BTKE), (a) $p_a = 0.99$, (b) $p_a = 0.995$ and (c) $p_a = 0.999$.

Application to the Annuity Longevity Risk

The actuarial present value of a lifetime annuity for a person of age a is:

$$\rho_a = \int_0^\infty {}_tP_a e^{-rt} dt, \tag{25}$$

where ${}_tP_a$ measures the probability of an individual of age a surviving t time longer, r is the risk-free rate and e^{-rt} is the financial discount. We know that the value of ${}_tP_a$ can be expressed as a function of the age-at-death distribution $F_X(\cdot)$ (see [Preston et al. 2001](#), pp. 53–54):

$${}_tP_a = \frac{1 - F_X(a + t)}{1 - F_X(a)} = \frac{P(X > a + t)}{P(X > a)}.$$

So, with the BTKE we can obtain smoothed and consistent estimations of ${}_tP_a, \forall t \geq 0$ and, therefore, calculate the value of single the premium expressed in (25). In [Figure 3](#), we show these estimated probabilities for $a = 75, 85, 95, 100$ with the data for 2017. The values of ρ_a calculated using these estimated ${}_tP_a$ and with $r = 0.2$ are shown in [Table 4](#).

Table 4. Estimates of a single annuity premium (ρ_a) for $r = 0.2$ and year 2017, at $a = 75, 85, 95, 100$. The probabilities ${}_tP_a$ are estimated with the Beta Transformed Kernel Estimator (BTKE).

	$a = 75$	$a = 85$	$a = 95$	$a = 100$
Men	4.8854	4.7064	4.2620	4.1138
Women	4.9168	4.7938	4.3975	4.2183

As an alternative to the lifetime annuities, the temporary annuities (see [Shang and Haberman 2020](#), for the justification for this type of financial contract) depend on a given maturity or a given age from which the probability of living is assumed to be zero and which we call w , being, a priori, the same for men and women (see [Chen et al. 2018](#), for an example of an application in solvency risk). So, in the integral expressed in (25) we can finally replace the ∞ by $w - a$.

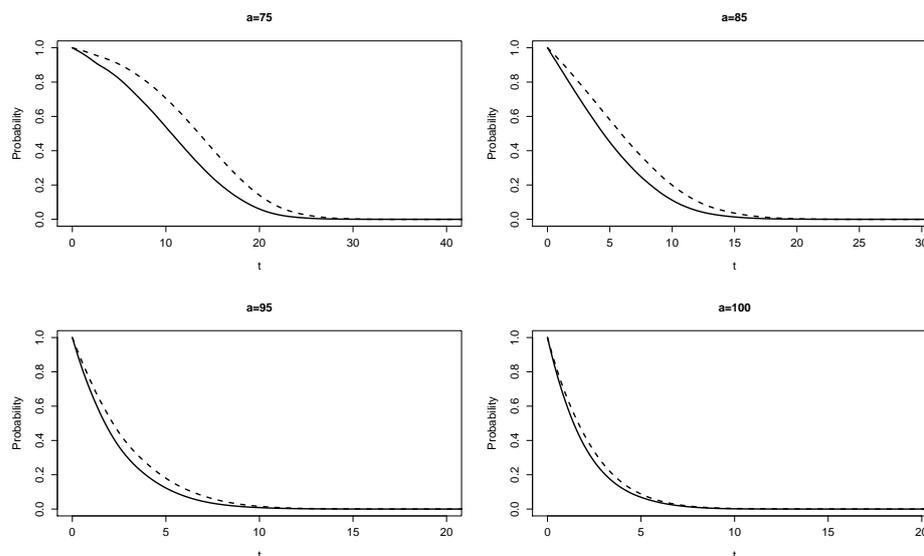


Figure 3. ${}_tP_a$ for the year 2017 estimated from Beta Transformed Kernel Estimator (BTKE), for men (solid line) and women (dashed line), with $a = 75, 85, 95, 100$.

Usually, the value of w is determined by the mortality rate or its projection. It normally corresponds to an age at which all individuals have died. Considering that longevity risk is the estimated CQ x_{p_a} at a given confidence level near 1, the integral (25) can be divided into two parts, from 0 to $x_{p_a} - a$ and from longevity risk $x_{p_a} - a$ to ∞ , i.e.,

$$\rho_a = \int_0^{x_{p_a}-a} {}_tP_a e^{-rt} dt + \int_{x_{p_a}-a}^{\infty} {}_tP_a e^{-rt} dt. \tag{26}$$

Expression (26) is convenient because it splits the lifetime annuity into two segments; only a $(1 - p_a)$ fraction of the participants will reach age x_{p_a} (we call this term “tail premium”), while a p_a fraction will not reach this age. For example, for age 75 with $p_a = 0.99$, in Table 2 we observe that for men $x_{p_a} = 99.31$ and the single premium in Table 4 is split as $4.8854 = 4.8503 + 0.0351$; for women $x_{p_a} = 102.14$ and this separation is $4.9168 = 4.8959 + 0.0209$. These results indicate that when considering longevity risk, the maturity for women should be longer than for men and, on the contrary, the tail premium should be lower. The reason is that the tail premium starts later for women than it does for men and, therefore, covers a shorter period with higher mortality rates. The opposite is true for men, i.e., maturity should be shorter and the tail premium should be higher. Table A6 in Appendix D shows the results of the split premiums for each age (a) and quantile level p_a .

Our approach helps identify the tail in terms of a percent of the whole distribution. An alternative approach would be to fix the advanced age, i.e., 100 years and estimate the proportion of cases above this bound or the annuity from this age. The difference between our method and this alternative is that, when age is fixed, the size of the tails are not comparable.

6. Conclusions

We have shown the utility of the kernel estimator for estimating longevity risk. Specifically, we have demonstrated that the beta transformed kernel estimator allows us to estimate conditional extreme quantiles efficiently when using a large database. Moreover, this estimator can be easily implemented with real or projected data. Finally, we obtained smoothed nonparametric estimations of the ${}_tP_a, \forall t \geq 0$.

Using the BTKE, we have estimated longevity risk using the conditional extreme quantile of the age-at-death and we conclude that it is higher for women than it is for men in almost all the cases analyzed ($a = 65, 75, 85, 95$ and $p_a = 0.99, 0.995, 0.999$) in Spain.

Furthermore, the longevity risk age increases across the period 2011–2017, especially in the case of women.

The method presented in this paper cannot be used with incomplete cohort data, because information on all ages of death is needed. Longevity cohort analysis, with still-living individuals, would require projections of mortality and, therefore, a parametric or a model approach would be better than the model-free kernels.

In short, here, we have proposed a way to measure longevity risk in insurance portfolios. This can be achieved by splitting the unitary lifetime annuity according to a conditional extreme quantile of the age-at-death distribution in the portfolio that is associated with the fact that individuals may reach a certain age with a low probability.

As a potential future research outline, it would be interesting to consider conventional risk factors as the alcohol consumption status or smoking status, alongside health issues (for example, diabetes or cardiac risk). The proposed methodology standardizes longevity risk comparison for annuity portfolios in insurance companies.

Author Contributions: Conceptualization, C.B. and M.G.; methodology, C.B. and M.G.; software, C.B. and M.G.; validation, C.B. and M.G.; formal analysis, C.B. and M.G.; investigation, C.B. and M.G.; resources, C.B. and M.G.; data curation, C.B. and M.G.; writing—original draft preparation, C.B. and M.G.; writing—review and editing, C.B. and M.G.; visualization, C.B. and M.G.; supervision, C.B. and M.G.; project administration, C.B. and M.G.; funding acquisition, C.B. and M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Spanish Ministry of Science and Innovation grant PID2019-105986GB-C21 and the BBVA Foundation grants for Big Data.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: www.ub.edu/rfa/R/BTKE, accessed on 12 April 2021.

Acknowledgments: The authors thank the comments and suggestions by two anonymous reviewers. M.G. thanks ICREA Academia.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Proofs

Proof of Proposition 1. We prove by means of simple calculations that:

$$\begin{aligned}
 E(\hat{p}) &= p_a - p_a E[\hat{F}_X(a)] + E[\hat{F}_X(a)] \\
 &= p_a - p_a \left\{ F_X(a) + \text{Bias}[\hat{F}_X(a)] \right\} + \left\{ F_X(a) + \text{Bias}[\hat{F}_X(a)] \right\} \\
 &= p_a [1 - F_X(a)] + F_X(a) + (1 - p_a) \text{Bias}[\hat{F}_X(a)] \\
 &= p + (1 - p_a) \text{Bias}[\hat{F}_X(a)]
 \end{aligned}$$

and

$$\begin{aligned}
 V\{\hat{p}\} &= V\left[p_a - p_a \hat{F}_X(a) + \hat{F}_X(a)\right] \\
 &= V\left[\hat{F}_X(a) - p_a \hat{F}_X(a)\right] = V\left[(1 - p_a) \hat{F}_X(a)\right] \\
 &= (1 - p_a)^2 V\left[\hat{F}_X(a)\right].
 \end{aligned}$$

□

Proof of Theorem 1. The proof of Theorem 1 is direct if we consider that the BTKE is the KE of the double transformed variable, i.e.,

$$\widehat{F}_X^{BT}(x) = \widehat{F}_{M^{-1}[T(X)]} \left\{ M^{-1}[T(x)] \right\}.$$

We obtain expressions (19) by replacing X with the transformed random variable $M^{-1}[T(X)]$ in the bias and variance of the KE, which will be defined in (9) and (10), respectively, and by deriving with respect to x , knowing that:

$$f_{M^{-1}[T(x)]} \left\{ M^{-1}[T(x)] \right\} = \frac{f_X(x)}{(M^{-1})'[T(x)]T'(x)}$$

and

$$(M^{-1})'[T(x)] = \frac{1}{m\{M^{-1}[T(x)]\}},$$

where $m(y) = M'(y) = \frac{15}{16}(1 - y^2)^2, -1 < y < 1$, is the pdf of the $Beta(3, 3)$. □

Appendix B. Simulation Study

The simulation study conducted can be summarised as follows. We compare the MSE of extreme CQ estimations based on the empirical distribution, KE and BTKE. We generate 2000 samples of size $n = 500$ and $n = 5000$ from each distribution in Table A1 (Weibull, Lognormal and two mixtures of Lognormal and Pareto models, 70–30% and 30–70%, respectively). In all cases, we consider two different combinations of parameters related to the tail shapes. The Weibull (We) and Lognormal (Ln) are long-tail distributions and, furthermore, Ln is a subexponential distribution. The composite Lognormal–Pareto (Ln-Pa) models are heavy-tailed distributions.

Table A1. Distributions in the simulation study.

Distribution	cdf	Parameters
We	$1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma}$	$(\lambda, \gamma) \in \{(1, 0.75), (1, 1.5)\}$
Ln	$\int_{-\infty}^{\log x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$	$(\mu, \sigma) \in \{(0, 0.5), (0, 1)\}$
Ln-Pa	$\alpha \int_{-\infty}^{\log x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$	$(\alpha, \mu, \sigma, \lambda, \rho, c) \in \{(0.7, 0, 1, 1, 1, -1), (0.7, 0, 1, 1, 1.1, -1)\}$
Ln-Pa	$+(1 - \alpha) \left[1 - \left(\frac{x-c}{\lambda}\right)^{-\rho} \right]$	$(\alpha, \mu, \sigma, \lambda, \rho, c) \in \{(0.3, 0, 1, 1, 1, -1), (0.3, 0, 1, 1, 1.1, -1)\}$

Considering the maximum distribution, for those with an exponential tail, we calculate CQ with $a = 0.5, 1$ and for those with a heavy tail we calculate it with $a = 1, 2$. The conditional confidence levels used are $p_a = 0.95, 0.99, 0.995$. In Table A2, the true values of these conditional quantiles CQ_{p_a} defined in (1) are shown; for each a the values were ordered from the lowest to the highest.

The results of the simulation study are obtained using two alternative cdfs as first transformation T . The first alternative is the Lognormal distribution, that has tail with exponential decrease. The second alternative is the cdf of the modified Champernowne distribution proposed by Alemany et al. (2013), that has tail with power decrease and is defined as:

$$T(x) = \frac{(x + c)^\delta - c^\delta}{(x + c)^\delta + (M + c)^\delta - 2c^\delta}, \tag{A1}$$

with parameters $\delta, M > 0$ and $c \geq 0$, and then applying an optimal second transformation. Note that when $c = 0$, the modified Champernowne distribution is also known as a log-logistic distribution.

Table A3 contains the results of a simulation study for each distribution analysed, each first transformation T and each value of a , using $p_a = 0.95, 0.99$ for sample size $n = 500$ and $p_a = 0.95, 0.99, 0.995$ for sample size $n = 5000$. In this table, we present the

ratio between the MSE of the BTKE and KE compared to the MSE of the empirical-based estimation. These results show that the BTKE performs better than KE when the aim is to estimate an extreme CQ of a heavy-tailed distribution.

Table A2. Theoretical values of CQ_{p_a} for each distribution in Table A1, given a and p_a .

	Distribution	$p_a = 0.950$	$p_a = 0.990$	$p_a = 0.995$
$a = 0.5$	We(1, 1.5)	2.239	2.908	3.173
	Ln(0, 0.5)	3.340	4.372	4.851
	We(1, 0.75)	5.498	9.008	10.644
	Ln(0, 1)	8.662	15.465	19.281
$a = 1$	We(1, 1.5)	2.518	3.155	3.411
	Ln(0, 0.5)	4.599	5.772	6.317
	We(1, 0.75)	6.341	9.957	11.631
	Ln(0, 1)	11.156	19.115	23.532
	Ln-Pa (0.7, 0, 1, 1, 1.1, -1)	11.248	41.655	78.235
	Ln-Pa (0.7, 0, 1, 1, 1, -1)	13.017	59.189	119.030
	Ln-Pa (0.3, 0, 1, 1, 1.1, -1)	20.902	92.338	174.248
	Ln-Pa (0.3, 0, 1, 1, 1, -1)	27.162	139.003	279.000
$a = 2$	Ln-Pa (0.7, 0, 1, 1, 1.1, -1)	18.094	73.965	139.512
	Ln-Pa (0.7, 0, 1, 1, 1, -1)	22.298	109.790	220.508
	Ln-Pa (0.3, 0, 1, 1, 1.1, -1)	33.922	149.225	281.095
	Ln-Pa (0.3, 0, 1, 1, 1, -1)	44.732	227.337	455.672

With sample size $n = 500$ and $p_a = 0.95$, the BTKE performs best for Ln distributions for $a = 0.5, 1$ and for Ln-Pa distributions for $a = 2$; when $p_a = 0.99$, in general, the BTKE performs best for Ln distributions and for Ln-Pa distributions, with the exception of Ln-Pa (0.7, 0, 1, 1, 1.1, -1) for $a = 2$, which, when $n = 5000$, also become the lowest ratio. In general, when the sample size is $n = 5000$, the Weibull results for the BTKE improve but they are still worse than those of the KE; the results for the Weibull are the worst, given that the right tail of this distribution is much shorter than that of the Champernowne and Lognormal and, for the Weibull, the BTKE tends to overestimate the conditional extreme quantiles analyzed. For the Ln and Ln-Pa distributions, the BTKE results are the best, with the exception of three cases that are influenced by the specific sample selection and where the ratio is very near to 1.

If the results obtained with alternative choices of T (Lognormal and Champernowne cdfs) are compared, we observe that Lognormal first transformation works in general when an extreme quantile of a heavy-tailed distribution is estimated. This result can be deduced from Theorem 1 if, for a given extreme value x in the right tail of the distribution, we observe that in this particular case

$$\left(1 - \frac{f_X(x_p) / f'_X(x_p)}{T'(x_p) / T''(x_p)} \right) \leq 0$$

and, therefore, the bias of the BTKE is the sum of a negative term and a positive term. In this case, we could overestimate or underestimate the quantile but the two terms can even be compensated. Furthermore, the variance of the BTKE is smaller than the variance of CKE. Furthermore, from Theorem 1 we can deduce that, if the distribution used in the first transformation is heavier-tailed than the true distribution, the bias for a given extreme value x in the right tail is the sum of two negative terms and BTKE tends to overestimate the conditional extreme quantiles.

From the simulation study, we conclude that the BTKE provides the best approach for heavy-tailed distributions. Indeed, this method overcomes the difficulties involved in estimating extreme quantiles when the classical nonparametric method is inefficient and a parametric distribution provides unrealistic results when the aim is to estimate the extreme quantile.

Table A3. Ratios between the MSE (Mean Squared Error) of Beta Transformed Kernel Estimator (BTKE) and Kernel Estimator (KE) of the CQ_{p_a} and MSE of the estimation based on empirical distribution (Ln-Pa distributions with parameters $(\alpha, \sigma, \lambda, \rho)$, given $\mu = 0$ and $c = -1$).

<i>a</i>	Distribution	<i>n</i> = 500						<i>n</i> = 5000								
		<i>p_a</i> = 0.950		<i>p_a</i> = 0.990		<i>p_a</i> = 0.950		<i>p_a</i> = 0.990		<i>p_a</i> = 0.995						
		CKE	BTKE	CKE	BTKE	CKE	BTKE	CKE	BTKE	CKE	BTKE					
		Ch	Ln	Ch	Ln	Ch	Ln	Ch	Ln	Ch	Ln					
0.5	We(1, 1.5)	0.87	0.99	1.52	0.85	1.21	1.48	0.94	0.99	1.15	0.91	1.03	1.16	0.91	1.02	1.08
	We(1, 0.75)	0.89	1.27	2.45	0.90	1.43	1.76	0.97	1.03	2.71	0.97	1.11	1.73	0.96	1.13	1.49
	Ln(0, 0.5)	0.95	0.92	0.94	0.96	0.76	0.86	0.99	0.98	0.97	0.99	0.95	0.88	1.00	0.92	0.83
	Ln(0, 1)	1.00	0.84	0.87	1.05	0.63	0.77	1.00	0.97	0.97	1.01	0.88	0.93	1.02	0.83	0.90
1	We(1, 1.5)	0.82	1.06	1.47	0.73	1.10	1.16	0.92	0.97	1.15	0.91	1.01	1.07	0.91	1.04	1.04
	We(1, 0.75)	0.91	1.16	1.43	0.92	1.22	1.13	0.97	1.04	1.18	0.98	1.10	0.17	0.97	1.14	1.09
	Ln(0, 0.5)	0.99	0.93	0.96	0.99	0.79	0.87	1.00	0.99	0.99	1.00	0.95	0.97	1.00	0.93	0.96
	Ln(0, 1)	1.02	0.84	0.89	1.00	0.57	0.71	1.01	0.96	0.98	1.01	0.88	0.93	1.00	0.83	0.89
	Ln-Pa (0.7, 1, 1, 1.1)	0.90	1.01	0.91	1.00	0.74	0.53	0.97	0.94	0.92	0.97	0.89	0.85	0.99	0.95	0.79
	Ln-Pa (0.7, 1, 1, 1)	0.86	0.93	0.84	1.00	0.75	0.52	0.97	0.95	0.92	0.98	0.92	0.84	0.99	0.91	0.66
	Ln-Pa (0.3, 1, 1, 1.1)	0.90	0.89	0.82	1.00	0.82	0.54	0.98	0.94	0.90	0.98	0.91	0.85	0.96	0.94	0.81
2	Ln-Pa (0.3, 1, 1, 1)	0.86	0.92	0.84	1.00	1.05	0.72	0.97	0.91	0.89	0.97	0.95	0.86	1.00	1.05	0.85
	Ln-Pa (0.7, 1, 1, 1.1)	0.95	0.78	0.66	0.99	2.70	59.93	0.98	0.92	0.89	0.97	0.86	0.72	1.00	0.88	0.67
	Ln-Pa (0.7, 1, 1, 1)	0.94	0.75	0.63	1.00	0.95	49.04	0.98	0.90	0.87	0.99	0.93	0.69	1.00	0.82	0.74
	Ln-Pa (0.3, 1, 1, 1.1)	0.96	0.72	0.72	1.00	0.46	0.46	0.98	0.89	0.86	0.98	0.97	0.84	0.99	1.01	0.79
	Ln-Pa (0.3, 1, 1, 1)	0.94	0.39	0.78	1.00	0.70	0.70	0.98	0.90	0.87	0.99	1.00	0.84	1.00	0.90	0.66

BTKE transformations: Champernowne (Ch) or Lognormal (Ln).

Appendix C. Estimated Conditional Significant Levels

Table A4. Quotients between empirical and theoretical $(1 - p_a)$ significance levels of estimated CQ_{p_a} for the age-at-death of the population aged over 65 in Spain, separately for men and women, from 2011 to 2017, using a Beta Transformed Kernel Estimator (BTKE) or a Generalized Pareto Distribution (GPD).

	<i>p_a</i>	Year	BTKE for Each <i>a</i>					GPD for Each <i>a</i>				
			65	75	85	95	100	65	75	85	95	100
Men	0.990	2011	1.00	1.01	0.98	0.90	1.03	0.97	0.97	0.98	1.96	14.52
		2012	1.00	0.99	0.99	0.98	0.93	1.00	0.99	0.99	1.96	14.19
		2013	0.98	1.00	0.98	0.98	0.94	0.98	0.97	0.98	2.22	15.33
		2014	1.00	1.00	0.96	0.97	0.82	0.96	1.00	0.96	2.33	16.43
		2015	1.01	0.99	1.00	1.01	0.92	0.97	0.99	1.00	1.43	10.42
		2016	0.99	0.98	0.97	0.91	0.98	0.99	0.98	1.00	0.91	3.04
		2017	1.01	0.99	0.97	0.97	0.69	0.98	0.99	0.97	0.97	1.88
	0.995	2011	0.99	1.01	0.97	1.04	1.03	0.99	0.97	0.97	3.92	29.05
		2012	0.99	1.00	0.95	1.03	0.93	0.99	0.96	0.95	3.92	28.37
		2013	0.97	0.99	0.94	0.99	0.71	0.97	0.99	0.94	4.45	30.66
		2014	0.98	0.98	0.96	0.92	0.93	0.98	0.98	1.00	4.66	32.87
		2015	1.01	1.01	0.96	0.95	1.02	0.96	0.97	0.99	2.85	20.84
		2016	1.00	1.00	0.97	1.04	1.08	0.96	0.96	0.97	0.92	6.07
		2017	0.98	0.97	1.01	1.04	0.99	0.98	0.97	0.96	0.96	3.76
0.999	2011	1.00	0.99	0.92	0.87	1.29	0.94	0.99	2.07	19.59	145.24	
	2012	0.97	0.96	0.97	0.96	0.00	0.97	0.96	2.05	19.58	141.86	
	2013	0.98	1.00	0.97	1.20	0.00	0.98	1.04	2.21	22.25	153.30	
	2014	1.00	1.01	0.99	0.99	0.00	1.00	1.12	2.29	23.28	164.34	
	2015	0.93	0.95	1.00	0.84	1.02	0.93	0.95	1.48	14.27	104.19	
	2016	1.00	0.96	0.91	0.98	1.08	0.95	0.96	0.95	3.92	30.37	
	2017	0.95	1.00	0.91	0.60	0.99	0.98	0.93	0.99	2.29	18.81	

Table A4. Cont.

	p_a	Year	BTKE for Each a					GPD for Each a				
			65	75	85	95	100	65	75	85	95	100
Women	0.990	2011	1.00	0.99	0.96	0.92	1.02	0.97	0.99	0.96	1.95	11.38
		2012	1.02	0.99	1.00	0.96	0.97	0.98	0.99	0.96	0.96	2.13
		2013	1.00	0.98	0.98	0.98	0.90	1.00	0.98	0.98	0.98	1.91
		2014	1.00	1.00	0.97	0.95	1.00	0.97	1.00	0.97	0.95	2.77
		2015	0.99	1.01	0.98	0.98	0.90	0.99	0.96	0.98	0.98	2.47
		2016	0.98	1.00	0.99	0.98	0.93	0.98	1.00	0.99	0.98	0.96
		2017	0.98	1.00	0.98	0.98	0.98	0.98	1.00	0.98	0.98	2.42
	0.995	2011	0.97	1.00	1.00	0.99	0.85	0.97	1.00	0.95	3.89	22.75
		2012	1.00	1.00	1.00	0.95	0.94	1.00	0.95	0.95	0.95	4.25
		2013	1.00	0.98	0.98	1.01	0.90	1.00	0.98	0.98	0.89	3.82
		2014	0.97	0.99	0.98	1.02	0.86	0.97	0.99	0.98	1.02	5.53
		2015	0.99	0.99	0.99	0.95	0.95	0.99	0.99	0.99	0.95	4.94
		2016	0.99	1.00	1.01	0.96	1.01	0.99	1.00	0.95	0.96	1.16
		2017	0.98	0.99	0.98	0.93	0.91	0.98	0.99	0.98	0.99	4.83
	0.999	2011	0.97	0.94	1.01	0.84	0.66	2.06	2.29	3.63	19.46	113.77
		2012	1.00	0.97	0.97	1.00	0.89	0.90	0.99	0.97	3.80	21.26
		2013	0.99	0.96	0.94	0.81	0.90	0.99	0.96	0.94	3.47	19.11
		2014	0.98	0.98	1.00	0.95	1.14	0.98	0.98	0.98	5.11	27.65
		2015	0.98	0.96	0.97	1.05	0.95	0.98	0.96	0.97	4.74	24.71
		2016	0.99	0.94	0.99	1.06	0.75	0.99	0.94	0.99	1.06	5.78
		2017	0.99	1.00	0.93	0.90	0.91	0.99	0.96	1.00	4.35	24.16

The tail shape parameter of the GPD was between 0 and 0.5. So, there are first and second finite moments.

Table A5. Quotients between empirical and theoretical $(1 - p_a)$ significance levels of estimated CQ_{p_a} for the age-at-death of the population aged over 65 in Spain separately for men and women, from 2011 to 2017, using Gompertz and Gompertz-G distributions.

	p_a	Year	Gompertz for Each a					Gompertz-G for Each a				
			65	75	85	95	100	65	75	85	95	100
Men	0.990	2011	1.29	1.33	1.57	2.74	3.47	0.32	0.27	0.14	0.14	0.51
		2012	1.25	1.27	1.50	2.57	4.77	0.66	0.53	0.34	0.42	0.93
		2013	1.17	1.26	1.49	2.84	3.18	0.40	0.35	0.25	0.14	0.24
		2014	1.07	1.15	1.29	2.81	3.61	0.44	0.38	0.28	0.23	0.23
		2015	1.08	1.15	1.27	2.59	3.47	0.36	0.36	0.33	0.41	0.51
		2016	0.95	1.01	1.09	2.38	4.01	0.30	0.28	0.20	0.21	0.33
		2017	0.94	0.95	1.01	1.85	4.26	0.24	0.20	0.12	0.06	0.00
	0.995	2011	1.45	1.52	1.83	2.70	4.63	0.22	0.18	0.11	0.21	0.26
		2012	1.43	1.43	1.70	3.18	7.44	0.49	0.42	0.28	0.45	1.40
		2013	1.41	1.48	1.67	3.42	3.54	0.32	0.29	0.15	0.07	0.00
		2014	1.25	1.30	1.56	3.17	4.66	0.35	0.32	0.23	0.23	0.00
		2015	1.22	1.29	1.48	3.30	4.70	0.33	0.33	0.32	0.42	0.20
		2016	1.09	1.13	1.35	2.86	4.77	0.25	0.23	0.15	0.20	0.43
		2017	1.07	1.09	1.19	2.50	5.74	0.17	0.15	0.10	0.05	0.00
	0.999	2011	2.10	2.16	2.90	5.20	10.28	0.10	0.08	0.11	0.17	0.00
		2012	1.96	2.09	2.53	6.90	16.28	0.35	0.33	0.30	0.96	2.33
		2013	2.04	2.18	2.83	4.96	7.08	0.13	0.11	0.03	0.00	0.00
		2014	1.82	2.06	2.50	5.12	8.16	0.20	0.18	0.11	0.00	0.00
		2015	1.83	2.01	2.34	5.60	6.13	0.28	0.29	0.22	0.14	0.00
		2016	1.68	1.74	2.15	5.18	9.76	0.15	0.14	0.10	0.28	2.17
		2017	1.48	1.59	1.86	5.18	4.95	0.08	0.04	0.03	0.00	0.00

Table A5. Cont.

p_a	Year	Gompertz for Each a					Gompertz-G for Each a				
		65	75	85	95	100	65	75	85	95	100
Women 0.990	2011	1.33	1.37	1.50	2.69	4.26	1.13	1.16	1.06	1.52	2.43
	2012	1.40	1.42	1.60	2.88	4.72	0.50	0.47	0.39	0.48	0.59
	2013	1.39	1.42	1.56	2.84	4.12	1.04	1.06	1.05	1.48	2.30
	2014	1.37	1.42	1.56	2.98	4.87	0.86	0.89	0.81	1.26	2.39
	2015	1.44	1.48	1.66	3.29	5.30	1.44	1.48	1.66	3.39	5.94
	2016	1.37	1.42	1.57	3.07	5.33	0.59	0.58	0.51	0.69	1.08
	2017	1.37	1.41	1.53	2.95	5.22	1.23	1.27	1.36	2.22	3.53
0.995	2011	1.55	1.60	1.84	3.58	6.30	1.20	1.21	1.21	1.53	3.15
	2012	1.70	1.73	2.00	3.76	6.32	0.46	0.44	0.37	0.48	0.83
	2013	1.70	1.76	2.04	3.62	6.27	1.16	1.23	1.16	1.84	2.57
	2014	1.66	1.78	1.95	3.83	7.24	0.93	0.95	0.89	1.53	2.96
	2015	1.83	1.84	2.06	4.58	6.89	1.83	1.84	2.06	4.77	7.84
	2016	1.71	1.74	1.98	3.79	6.99	0.57	0.57	0.52	0.78	1.31
	2017	1.74	1.77	1.95	3.78	7.06	1.46	1.50	1.61	2.69	4.01
0.999	2011	2.71	2.72	3.34	6.11	14.10	1.71	1.70	1.49	3.14	4.59
	2012	2.93	2.96	3.57	7.55	13.29	0.42	0.40	0.36	0.69	0.89
	2013	2.88	2.84	3.31	7.05	9.56	1.51	1.59	1.59	2.71	4.48
	2014	2.92	2.95	3.47	7.64	15.96	1.20	1.28	1.35	2.74	6.27
	2015	3.20	3.37	4.12	8.53	15.21	3.33	3.53	4.28	9.80	18.06
	2016	2.97	3.15	3.44	7.75	14.07	0.62	0.60	0.58	1.06	2.01
	2017	2.97	3.16	3.46	7.75	15.27	2.33	2.31	2.56	4.35	5.70

Appendix D. Split Lifetime Annuities Premiums

Table A6. Temporary annuity premium assuming $w = CQ_{p_a}$ (upper table) and “tail premium”, for $r = 0.2$ and year 2017. The probabilities ${}_tP_a$ were estimated with the Beta Transformed Kernel Estimator (BTKE).

P_a		Temporary Annuity Premium			
		$a = 75$	$a = 85$	$a = 95$	$a = 100$
0.990	Men	4.8503	4.5314	3.7946	3.3886
	Women	4.8959	4.6740	3.9961	3.5580
0.995	Men	4.8590	4.5736	3.9088	3.4867
	Women	4.9007	4.7025	4.0777	3.6583
0.999	Men	4.8708	4.6315	4.0247	3.6779
	Women	4.9074	4.7396	4.1900	3.9485
P_a		Tail Premium			
		$a = 75$	$a = 85$	$a = 95$	$a = 100$
0.990	Men	0.0351	0.1750	0.4674	0.7252
	Women	0.0209	0.1198	0.4014	0.6602
0.995	Men	0.0264	0.1328	0.3532	0.6272
	Women	0.0161	0.0914	0.3198	0.5600
0.999	Men	0.0146	0.0749	0.2373	0.4360
	Women	0.0093	0.0542	0.2075	0.2697

References

Alemany, Ramon, Catalina Bolancé, and Montserrat Guillen. 2013. A nonparametric approach to calculating value-at-risk. *Insurance: Mathematics and Economics* 52: 255–62. [CrossRef]

Alizadeh, Morad, Gauss M. Cordeiro, Luis G. Bastos Pinho, and Indranil Ghosh. 2017. The Gompertz-G family of distributions. *Journal of Statistical Theory and Practice* 11: 179–207. [CrossRef]

- Altman, Naomi, and Christian Léger. 1995. Bandwidth selection for kernel distribution function estimation. *Journal of Statistical Planning and Inference* 46: 195–214. [\[CrossRef\]](#)
- Azzalini, Adelchi. 1981. A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika* 68: 326–28. [\[CrossRef\]](#)
- Basellini, Ugofilippo, and Carlo Giovanni Camarda. 2019. Modelling and forecasting adult age-at-death distributions. *Population Studies* 73: 119–38. [\[CrossRef\]](#)
- Bolancé, Catalina. 2010. Optimal inverse Beta(3,3) transformation in kernel density estimation. *SORT-Statistics and Operations Research Transactions* 34: 223–37.
- Bolancé, Catalina, Montserrat Guillen, and Jens Perch Nielsen. 2003. Kernel density estimation of actuarial loss functions. *Insurance: Mathematics and Economics* 32: 19–36. [\[CrossRef\]](#)
- Bolancé, Catalina, Montserrat Guillen, and Jens Perch Nielsen. 2008. Inverse Beta transformation in kernel density estimation. *Statistics & Probability Letters* 78: 1757–64.
- Bolancé, Catalina, Montserrat Guillen, and Albert Pitarque. 2020. Sarmanov Distribution with Beta Marginals: An Application to Motor Insurance Pricing. *Mathematics* 8: 2020. [\[CrossRef\]](#)
- Bowers, Newton L., Hans U. Gerber, James C. Hickman, Donald A. Jones, and Cecil J. Nesbitt. 1997. *Actuarial Mathematics*, 2nd ed. Schaumburg: Society of Actuaries.
- Bowman, Adrian, Peter Hall, and Tania Prvan. 1998. Bandwidth selection for smoothing of distribution function. *Biometrika* 85: 799–808. [\[CrossRef\]](#)
- Buch-Larsen, Tine, Montserrat Guillen, Jens Perch Nielsen, and Catalina Bolancé. 2005. Kernel density estimation for heavy-tailed distributions using the Champernowne transformation. *Statistics* 39: 503–18. [\[CrossRef\]](#)
- Chen, An, Montserrat Guillen, and Elena Vigna. 2018. Solvency requirement in a unisex mortality model. *ASTIN Bulletin* 48: 1219–43. [\[CrossRef\]](#)
- Chen, Hua, and J. David Cummins. 2010. Longevity bond premiums: The extreme value approach and risk cubic pricing. *Insurance: Mathematics and Economics* 46: 150–61. [\[CrossRef\]](#)
- Coles, Stuart, Joanna Bawa, Lesley Trenner, and Pat Dorazio. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Berlin: Springer, vol. 208.
- Denuit, Michel. 2009. An index for longevity risk transfer. *Journal of Computational and Applied Mathematics* 230: 411–17. [\[CrossRef\]](#)
- Dickson, David, Mary Hardy, and Howard Waters. 1997. *Actuarial Mathematics for Life Contingent Risks*, 2nd ed. Cambridge: Cambridge University Press.
- Gavrilov, Leonid A., and Natalia S. Gavrilova. 2001. The reliability theory of aging and longevity. *Journal of Theoretical Biology* 213: 527–45. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gjonca, Arjan, Cecilia Tomassini, Barbara Toson, and Steve Smallwood. 2005. Sex differences in mortality, a comparison of the united kingdom and other developed countries. *Health Statistics Quarterly* 26: 6–16.
- Glei, Dana A., and Shiro Horiuchi. 2007. The narrowing sex differential in life expectancy in high-income populations: Effects of differences in the age pattern of mortality. *Population Studies* 61: 141–59. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hill, Peter D. 1985. Kernel estimation of a distribution function. *Communications in Statistics-Theory and Methods* 14: 605–20.
- Johnson, Norman L., Samuel Kotz, and Narayanaswamy Balakrishnan. 1995. *Continuous Univariate Distributions*. New York: John Wiley & Sons, Inc., vol. 1.
- Lee, Ronald D., and Lawrence R. Carter. 1992. Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association* 87: 659–71.
- Mayhew, Leslie, and David Smith. 2011. Human survival at older ages and the implications for longevity bond pricing. *North American Actuarial Journal* 15: 248–65. [\[CrossRef\]](#)
- Mayhew, Leslie, and David Smith. 2014. Gender convergence in human survival and the postponement of death. *North American Actuarial Journal* 18: 194–16. [\[CrossRef\]](#)
- Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford and Blackwell: John Wiley & Sons.
- Reiss, Rolf-Dieter. 1981. Nonparametric estimation of smooth distribution functions. *Scandinavian Journal of Statistics* 8: 116–19.
- Rootzén, Holger, and Dimitri Zholud. 2017. Human life is unlimited-but short. *Extremes* 20: 713–28. [\[CrossRef\]](#)
- Rootzén, Holger, and Dimitri Zholud. 2018. Rejoinder to discussion of the paper “human life is unlimited-but short”. *Extremes* 21: 415–24. [\[CrossRef\]](#)
- Ruppert, David, and Daren B. H. Cline. 1994. Bias reduction in kernel density estimation by smoothed empirical transformation. *Annals of Statistics* 22: 185–210. [\[CrossRef\]](#)
- Shang, Han Lin, and Steven Haberman. 2020. Forecasting age distribution of death counts: An application to annuity pricing. *Annals of Actuarial Science* 14: 150–69. [\[CrossRef\]](#)
- Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall/CRC Finance Series.
- Stallard, Eric. 2006. Demographic issues in longevity risk analysis. *The Journal of Risk and Insurance* 73: 575–609. [\[CrossRef\]](#)
- Swanepoel, Jan W. H., and François C. Van Graan. 2005. A new kernel distribution function estimator based on a nonparametric transformation of the data. *Scandinavian Journal of Statistics* 32: 551–62. [\[CrossRef\]](#)

-
- Terrell, George R. 1990. The maximal smoothing principle in density estimation. *Journal of the American Statistical Association* 85: 270–77. [[CrossRef](#)]
- Terrell, George R., and David W. Scott. 1985. Oversmoothed nonparametric density estimates. *Journal of the American Statistical Association* 80: 209–14. [[CrossRef](#)]
- Waldron, Ingrid. 1993. Recent trends in sex mortality ratios for adults in developed countries. *Social Science & Medicine* 36: 451–62.
- Wand, Matthew P., Stephen Marron, and David Ruppert. 1991. Transformations in density estimation. *Journal of the American Statistical Association* 86: 343–61. [[CrossRef](#)]