

# Is human life limited or unlimited? (A discussion of the paper by Holger Rootzén and Dmitrii Zholud)

A. Ferreira<sup>1</sup>  · F. Huang<sup>2</sup>

Received: 28 February 2018 / Revised: 19 March 2018 / Accepted: 5 April 2018 /  
Published online: 17 May 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** We thank Holger Rootzén and Dmitrii Zholud’s for their stimulating work, that led us to further investigate the problem of best fitting the human life span distribution. Their assertion “human life is unlimited but short” is based on their conclusion that the exponential model is the best to fit. We provide results based on standard Extreme Value approaches, the Block Maxima and Peaks-over-Threshold, and on the whole data available in the IDL database (IDL A, B validations and GRG). We verify that negative values for the extreme value index are more likely, supporting the conclusion that models with finite endpoint seem better to fit to the human life span distribution.

**Keywords** Peaks-Over-Threshold (POT) · Block Maxima (BM) · Human lifespan distribution · Extreme value index and endpoint estimation

**AMS 2000 Subject Classifications** 62G32 · 62P05 · 62P10

---

✉ A. Ferreira  
anafh@tecnico.ulisboa.pt

F. Huang  
fei.huang@anu.edu.au

<sup>1</sup> Instituto Superior Técnico, da Universidade de Lisboa, 1049-001 Lisbon, Portugal

<sup>2</sup> Research School of Finance, Actuarial Studies & Statistics, Australian National University, Canberra, ACT 2601, Australia

## 1 Introduction

The existence of an upper limit to the length of human life is an important subject among biologists, demographers and actuaries for its potential contribution to understanding human aging behavior and for determining the appropriate endpoint of a life table. From a statistical point of view, it is natural to consider it as an application of Extreme Value Theory (EVT) to lifespan at higher ages.

A characterizing feature of EV statistics is that it provides a consistent way to extrapolate beyond the sample information (usually to higher values) under very general assumptions. Generalized Extreme Value (GEV) or Generalized Pareto (GP) distributions can be justified after assuming that maxima converge, under independence or some mild dependence conditions on the underlying sample. Hence no parametric model needs to be assumed beforehand, with the GEV or GP justified as appropriate for large sample sizes (i.e. asymptotically). That is, the well known maximum domain of attraction condition (Fisher and Tippett 1928; Gnedenko 1943) justifies the GEV model for the distribution of maxima, and the GP distribution to model excesses above a high threshold. These lead respectively to the widely used approaches Block Maxima (BM) and Peaks-over-Threshold (POT, Balkema and de Haan 1974; Pickands 1975). The two distributions GEV and GP share the same shape parameter, the extreme value index. If this parameter is negative, this implies that the underlying distribution verifying the maximum domain of attraction condition has a finite right endpoint. Although EVT provides methods to estimate the endpoint we concentrate our discussion on extreme value index estimation and on its sign.

The article by Holger Rootzén and Dmitrii Zholud on “Human life is unlimited - but short” (Rootzén and Zholud 2017), brought the subject back to the EVT community and, not less important, bridged the discussion with several other communities, especially demography and actuarial sciences. There has been a long debate on whether there is a fixed upper limit to the length of human life among researchers from different disciplines for many years. As summarized in Thatcher (1999), there are three possibilities:

The first is that there will be a definite, fixed highest attainable age. Aarssen and de Haan (1994) found a finite age limit using mortality data from the Netherlands. Their results also suggest differences of maximum age-length between men and women. Watts et al. (2006) analyzed the high age mortality using  $r$ -largest and POT approaches to EV modeling with Japanese and Canadian cohort mortality data. Their results suggest the existence of a finite upper bound on the life span distribution. Li et al. (2008) proposed a threshold life table model using EVT to extrapolate survival distributions to extreme ages. Their results were then applied to the valuation of a life annuity portfolio and to the estimation the highest attained ages of different birth cohorts. A recent study (Einmahl et al. 2017) also found compelling statistical evidence that there is an upper limit to the life span of men and women using death data of the Netherlands from 1986 to 2015.

The second possibility is that there is no definite, fixed upper limit, but as age increases, the probability of survival becomes so tiny that it can be ignored in practice. Rootzén and Zholud’s finding lies in this category.

The third possibility is that there is no definite, fixed upper limit to life and there is no age reached when the probability of further survival is negligible. Vaupel, quoted in Baringa (1991) argued that there was no evidence of a biological limit to life. Similar conclusions were made by other researchers independently, for example, see Gavrilov and Gavrilova (1991).

Rootzén and Zholud's results include several interesting statements. For example, they show the importance of taking truncation into account and draw the conclusion that excess human life length after age 110 follows an exponential distribution. Additionally, they show that there is no difference in survival at extreme ages between women and men, between earlier periods and later ones, between countries with very different lifestyles and genetic compositions, and between different ages. Those interesting and surprising facts call for novel demographic or biological theories of aging. Additionally, they shed light on determining the appropriate endpoint of a life table, which is crucial in evaluating low-frequency, high-severity losses in actuarial practices, such as life annuity portfolios and defined-benefit pension plans.

However, we found out that the chosen IDL data set constrained to validated A data - the choice made by Rootzén and Zholud - can substantially affect the results. Furthermore we argue that their truncation scheme can contribute to biased results. From our data analysis, negative values for the extreme value index are more likely, supporting the conclusion that models with finite endpoint are more reliable to fit to human life span.

The paper is organized as follows. In Section 2 we discuss issues raised in Rootzén and Zholud (2017), with special emphasis on the data choice and the truncation scheme. In Section 3, we show data analysis results based on standard Extreme Value approaches, the Block Maxima and Peaks-over-Threshold, and on the whole data available in the IDL database that includes the International Database on Longevity (IDL A,B validation; <http://www.supercentenarians.org/>) and the Gerontological Research Group (GRG; <http://www.grg.org/>) database.

## 2 Discussion

In Rootzén and Zholud's paper the force of mortality at extreme ages is estimated by modeling excess life length of supercentenarians with a truncated GP distribution. They defend the importance of taking truncation into account and draw the conclusion that excess human life length after age 110 follows an exponential distribution.

However, we find that Rootzén and Zholud's model is statistically just one possibility, and a further step now is to improve on the precision of the results. From our findings, understanding the quality of the data seems crucial. The authors choose validated A data for avoiding bias but, we ask, by contrast, whether choosing only validated A data may incorporate bias into the results.

The IDL contains thoroughly validated information on the lifetimes of individuals of extreme ages, which is claimed to be free of age ascertainment bias. There are two validation levels in the IDL database. Validation level A is fully validated, which means both an early life document and a document certifying the date of death were

available. Validation level B is not validated as thoroughly as validation level A, but individual cases have been fully checked. In Table 1 are the sample sizes of validated A and B data, for each available country. One can see that validation criteria are spread throughout countries, though with the northern European countries FIN, NOR and SWE very representative of validation B.

The GRG database is another independent data resource, which contains information on validated deceased supercentenarians. Verified supercentenarians in the GRG database have multiple reviewed documents certifying their age, including, at minimum, one early-life document; one mid-life document; and one late-life document. While Rootzén and Zholud in their paper claimed that the GRG is expected to be “age-biased”, we find that this needs further investigation to prove/disprove their judgment.

IDL and GRG cover deaths occurring in different time periods: IDL mainly contains data before 2007 and GRG mainly contains data after 2007 through to 2016 (though with a few early year cases). In the data analysis shown in Section 3, we include both IDL (validation levels A and B) and GRG data. It is seen that the conclusion of finite endpoint seems more plausible, supporting several other studies as mentioned in the Introduction.

**Table 1** IDL records per type of validation and countries

Country	<i>Validated A</i>	<i>Validated B</i>	<i>Validated A &amp; B</i>
AUS	3	5	8
BEL	0	5	5
CAN	12	0	12
CAP	1	0	1
CHE	0	4	4
DEN	2	1	3
DEU	16	0	16
DNK	2	0	2
ESP	28	0	28
FIN	0	5	5
FRA	49	0	49
GBR	68	0	68
GER	2	0	2
GRE	2	0	2
ITA	44	0	44
JPN	78	0	78
NOR	0	9	9
SWE	0	8	8
UK	1	0	1
USA	323	0	323
Total	631	37	668

We end this section by pointing out two other aspects related to the modeling choice by Rootzén and Zholud that attracted our interest:

How crucial is the truncation scheme? Even Rootzén and Zholud admit that different results may be obtained without taking it into account. This scheme choice means inclusion in the model of information on the time interval that people died, and ultimately adopting a non-stationary EV model for parameter estimation. But this tends to contradict their observation that results show great stability in general, including between ages. Further note that from our data analysis below using standard EVT methods (hence not incorporating truncation) and using the same dataset as Rootzén and Zholud’s, similar results concluding for the exponential model are obtained.

Finally, note that the non-stationary model proposed by Rootzén and Zholud for the estimation of the extreme value index has a time constrained threshold. That is, the threshold  $t$  (age of 110, or more precisely in the model the times of achieving age 110) is constrained by  $e$  from the calendar time interval  $(b, e)$ , i.e. from

$$\frac{f(x)}{F(e - t) - F(b - t)_+}, \quad \text{with } t < e.$$

So, on the one hand the chosen model for estimation is not stable for a threshold tending to infinity, and only explains data up to 2007 since validated A data is only up to this year. But, on the other hand, the ultimately claimed exponential model to describe life span at higher ages indicates stability for any threshold converging to infinity, hence it satisfies threshold stability with unlimited endpoint. Therefore, some lack-of-fitting in the proposed model seems to be a remaining problem (also identified in our data analysis below).

We recall that by standard EV conditions on  $F$  (the target d.f.), it must hold that,

$$\frac{1 - F(t + xa(t))}{1 - F(t)} \sim (1 + \gamma x)_+^{-1/\gamma}, \quad \text{“for } t \text{ large”}, \quad \gamma \in \mathbb{R}, \quad (2.1)$$

with  $x > 0$ , and  $a(t) > 0$  a scaling normalizing function. The condition on the threshold, “ $t$  large” means for  $t$  approaching the right endpoint of the pertaining d.f.  $F$ , i.e.  $\sup\{x : F(x) < 1\} \in (0, \infty]$ . If the extreme value index  $\gamma$  is negative, the right endpoint of  $F$  and also of GEV and GP distributions are finite; if positive the endpoints of all equal infinity. The exponential model corresponding to  $\gamma = 0$ , is obtained by taking the limit in the right-hand side of Eq. 2.1, as  $\gamma$  approaches zero. Although the endpoint of the exponential distribution equals infinity, the underlying  $F$  may have finite or infinite endpoint. Property (2.1), with the GP survival function in the right-hand side, gives the characterizing threshold stability of the POT approach: the GP approximation remains valid with constant extreme value index  $\gamma$  while varying the high threshold  $t$ . For instance this stability property allows for model checking and optimizing estimation methods. For more on EV theory we refer e.g. to de Haan and Ferreira (2006).

### 3 Data analysis

Results on extreme value index estimation using standard EVT estimators and based on both BM and POT approaches are shown. The following data sets have been considered (cf. Table 2): (a) IDL together with GRG data, (b) IDL data with only validated A criteria, (c) IDL together with GRG only for women, and (d) IDL women data with only validated A criteria. Men seem to follow the same pattern as the other data sets but with greater uncertainty due to small sample sizes, and the results are omitted for this case.

The considered estimators are: for the POT approach MLE (maximum likelihood estimators), M-moment estimator (Dekkers et al. 1989) and PWM (probability weighted moment) and, for the BM approach MLE and PWM. Note that these last two methods give different estimators for POT and BM. The moment estimator is only defined for the POT approach. In Tables 2 and 3, estimates of the extreme value index obtained from all the estimators are given, the first table under the POT approach and the second under BM.

The POT method needs a threshold selection that was chosen after selecting a  $k$ -window with stable estimates, cf. Figs. 1, 2, 3 and 4. Our specific choices of the threshold can be found in Table 2. In Figs. 1–4, it is clearly seen that several other threshold choices could have been taken, while still reaching similar conclusions.

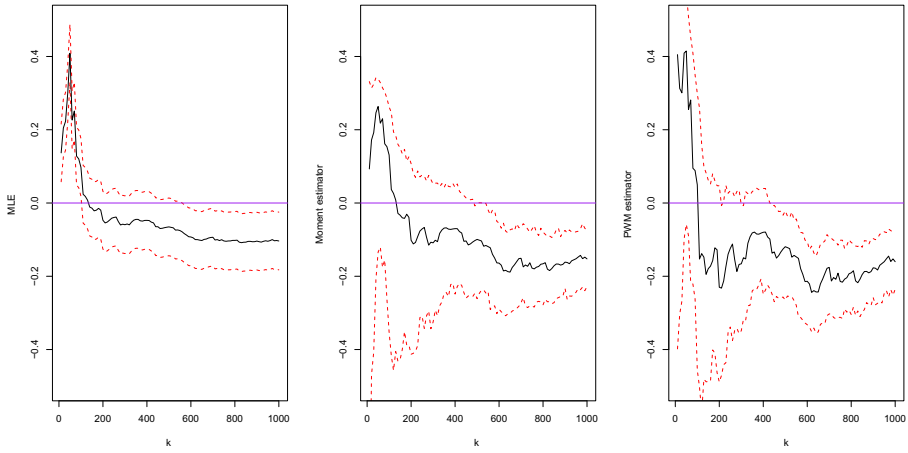
The BM method was applied with two different choices of blocks: year of birth and year of death. For the first choice, the highest attained age of all people who were born in the same year was selected. The year of birth (block) ranges from 1842 up to 1904 (with missing years in between, mostly at earlier years). In total we have 47 blocks. For the second choice of blocks, the highest attained age of all people who died in the same year was selected. The year of death (block) ranges from 1953 up to 2016 (with missing years in between, mostly at earlier years). In total we have 46 blocks. BM estimates were obtained using the *R* package *extRemes* from Gilleland and Katz (2016).

**Table 2** POT estimates and 95% confidence intervals for the extreme value index

<i>Data sets</i>	<i>Sample sizes</i>	<i>Threshold <math>k</math></i>	$\hat{\gamma}_{MLE}$	$\hat{\gamma}_M$	$\hat{\gamma}_{PWM}$
<i>MEN and WOMEN</i>					
IDL(A+B)+GRG	1015	400	−0.047 (−0.126,0.031)	−0.070 (−0.249,0.042)	−0.080 (−0.249,0.040)
IDL(A)	631	250	0.000 (−0.042,0.042)	0.004 (−0.210,0.125)	−0.017 (−0.192,0.134)
<i>WOMEN</i>					
IDL(A+B)+GRG	919	300	−0.044 (−0.109,0.021)	−0.072 (−0.265,0.060)	−0.100 (−0.273,0.051)
IDL(A)	565	300	−0.008 (−0.049,0.033)	−0.015 (−0.209,0.122)	−0.062 (−0.224,0.085)

**Table 3** BM estimates and 95% confidence intervals for the extreme value index

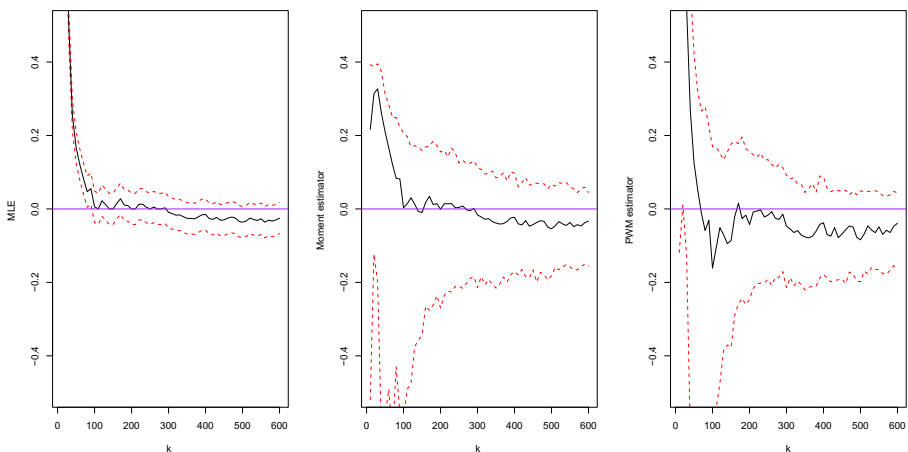
<i>Data sets</i>	Blocks - year of birth		Blocks - year of death	
	<i>Sample sizes</i>	$\hat{\gamma}_{MLE}$	<i>Sample sizes</i>	$\hat{\gamma}_{MLE}$
<i>MEN and WOMEN</i>				
IDL(A+B)+GRG	47	-0.104 (-0.263,0.055)	46	-0.062 (-0.253,0.128)
IDL(A)	35	0.007 (-0.207,0.222)	35	0.027 (-0.199,0.253)
				$\hat{\gamma}_{PWM}$
				-0.132 (-0.383,0.073)
				0.023 (-0.273,0.269)
<i>WOMEN</i>				
IDL(A+B)+GRG	44	-0.095 (-0.251,0.060)	44	-0.038 (-0.238,0.161)
IDL(A)	33	0.025 (-0.193,0.244)	33	0.077 (-0.171,0.326)
				0.088 (-0.215,0.320)



**Fig. 1** IDL+GRG data ( $n = 1015$ ): extreme value index estimation with 95% bootstrap confidence intervals

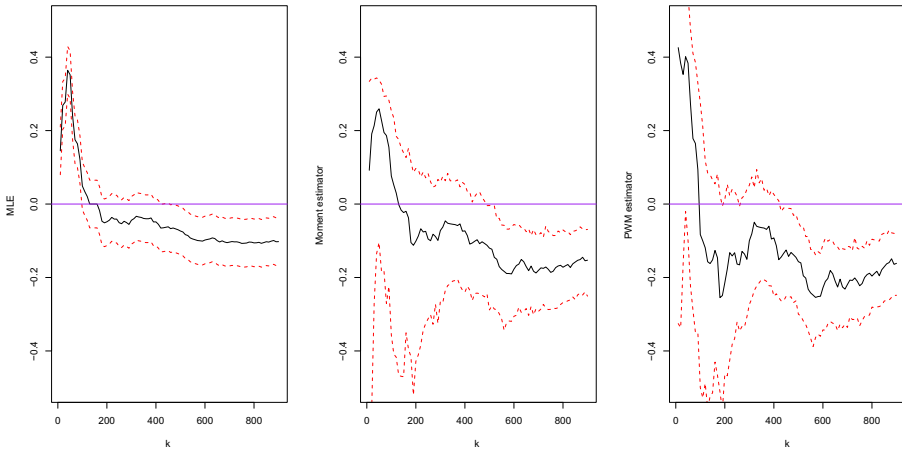
Standard GP and GEV fitting tests provided by *extRemes* with the considered samples, gave generally very reasonable results.

Summarizing our results: the effect of using data not restricted to IDL A validation for rejecting null (or even positive) extreme value index using the POT and BM approaches is clearly seen. It is a fact that the hypothesis  $H_0 : \gamma = 0$  is not rejected from the bootstrap 95% confidence intervals for several reasonable threshold choices in the POT approach and also from confidence intervals obtained using the BM method. But *all* point estimates using the data not restricted to A validation criteria are negative, and the corresponding confidence intervals cover more negative than positive values. This indicates that negative values for the extreme value index



**Fig. 2** IDL data with validation A criteria ( $n = 631$ ): extreme value index estimation with 95% bootstrap confidence intervals



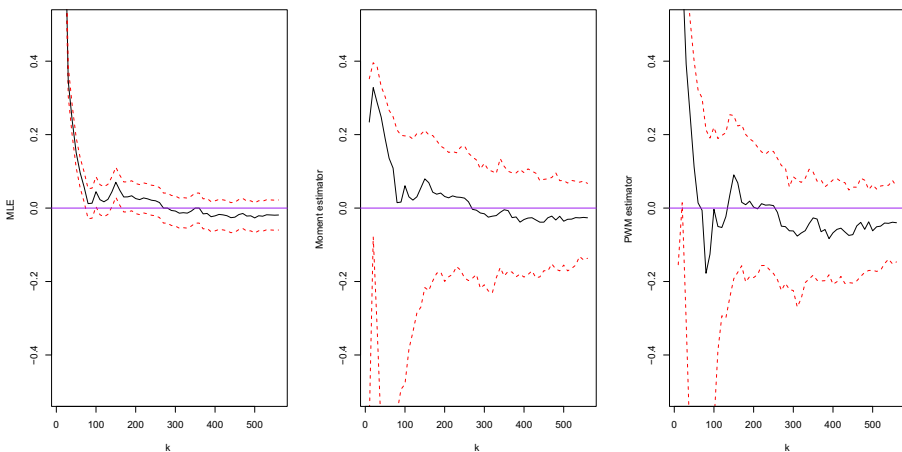


**Fig. 3** IDL+GRG Women data ( $n = 918$ ): extreme value index estimation with 95% bootstrap confidence intervals

are more likely, and so models with finite endpoint should fit better to the distribution of human lifespan.

Nonetheless when restricted to validated A data, the results show that the exponential model is very reliable.

Finally, one characteristic seen in Figs. 1–4 is the typical tendency of ML estimator to give higher values than the other estimators. This behavior has been confirmed asymptotically by Dombry and Ferreira (2018), i.e. that MLE tends to be more biased than PWM, though with lower variance.



**Fig. 4** IDL Women data with validation A criteria ( $n = 565$ ): extreme value index estimation with 95% bootstrap confidence intervals

**Acknowledgements** We thank Laurens de Haan, Ross Maller and Sidney Resnick for their careful reading and comments, and Thomas Mikosch for the opportunity to contribute to this discussion.

This research was partially funded by FCT - Fundação para a Ciência e a Tecnologia, Portugal, UID/MAT/00006/2013 and UID/Multi/04621/2013.

## References

- Aarssen, K., de Haan, L.: On the maximal life span of humans. *Math. Popul. Stud.* **4**(4), 259–81 (1994)
- Balkema, A.A., de Haan, L.: Residual life time at great age. *Ann. Probab.* **2**, 792–804 (1974)
- Baringa, M.: How long is the human life-span? *Science* **254**, 936–938 (1991)
- Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L.: A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17**, 1833–1855 (1989)
- Dombry, C., Ferreira, A.: Maximum likelihood estimators based on the block maxima method. *Bernoulli*: to appear (2018)
- de Haan, L., Ferreira, A.: *Extreme Value Theory: An Introduction*. Springer, Boston (2006)
- Einmahl, J., Einmahl, J., de Haan, L.: Limits to Human Life Span Through Extreme Value Theory. Center Discussion Paper Series No. 2017-051. CentER, Center for Economic Research, Tilburg (2017)
- Fisher, R.A., Tippett, L.H.C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Philos. Soc.* **24**, 180–190 (1928)
- Gavrilov, L.A., Gavrilova, N.S.: *The Biology of Life Span: A Quantitative Approach*. Harwood, New York (1991)
- Gilleland, E., Katz, R.W.: extRemes 2.0: an extreme value analysis package in R. *J. Stat. Softw.* **72**, 1–39 (2016)
- Gnedenko, B.V.: Sur la distribution limite du terme maximum d’une série aléatoire. *Ann. Math.* **44**, 423–453 (1943)
- Li, J., Hardy, M., Tan, K.S.: Threshold life tables and their applications. *North Am. Actuarial J.* **12**(2), 99–115 (2008)
- Pickands, J.: Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119–131 (1975)
- Rootzén, H., Zholud, D.: Human life is unlimited – but short. *Extremes* **20**, 713–728 (2017)
- Thatcher, A.R.: The long-term pattern of adult mortality and the highest attained age. *J. R. Stat. Soc. Ser. A* **162**(Part 1), 5–43 (1999)
- Watts, K.A., Dupuis, D.J., Jones, B.L.: An extreme value analysis of advanced age mortality data. *North Am. Actuarial J.* **10**(4), 162–178 (2006)