'The life of man, solitary, poore, nasty, brutish, and short'

Discussion of the paper by Rootzén and Zholud

A. C. Davison

Ecole Polytechnique Fédérale de Lausanne

EPFL-FSB-MATH-STAT, Station 8,

1015 Lausanne, Switzerland

`Anthony.Davison@epfl.ch`

September 11, 2018

The title, taken from Thomas Hobbes's 1651 work *Leviathan*, is his idea of the state of mankind in the absence of strong government. Whether or not he was right, his first four adjectives are fortunately now much less applicable than in the seventeenth century, and by 'short' he probably meant ages much lower than those discussed in the paper.

I congratulate Holger Rootzén and Dmitrii Zholud on their very stimulating paper. Data on the exceptionally long-lived are widely available, but rarely is there information about how they were gathered and what biases might be expected as a result, so I applaud the care with which the authors address the sampling scheme and accommodate its effects. There is a lot to discuss, but I shall focus on inference for the upper limit to human life, and on possible explanations why the conclusions of the paper differ from those in other recent work.

Inference on the endpoint

A key target of such modelling is the estimation of the endpoint η to human life. Although the estimated shape parameter of the extremal model that the authors fit to their data is very close to zero, it may be misleading to assume that the distribution is exponential. Accordingly I downloaded the data from the International Database on Longevity (`http://www.supercentenarians.org`, accessed on 10 January 2018), fitted the extremal model with survivor function $\mathcal{F}(t) = (1 + \xi t/\sigma)_+^{-1/\xi}$ to the residual lifetimes t beyond $u = 110$ years without attempting to allow for differences between

nationalities or sexes, and computed the profile log likelihood $\ell_p(\eta)$ for the endpoint $\eta = u - \sigma/\xi$ of the extremal distribution. I used slightly different sampling intervals from those in the paper, found 40 individuals with validation level B and dropped them from the analysis, and retained all the US and Japanese individuals, but otherwise my analysis was the same. The resulting maximum likelihood estimates of the shape and scale parameters and their standard errors are $\hat{\xi} = 0.007$ (0.046) and $\hat{\sigma} = 1.45$ (0.084), giving $\hat{\eta} = \infty$. Standard large-sample likelihood theory indicates that the likelihood ratio statistic $2\{\ell_p(\hat{\eta}) - \ell_p(\eta)\}$ has an approximate χ_1^2 distribution, so a $(1 - \alpha) \times 100\%$ confidence set comprises those values of η for which $\ell_p(\eta) \geq \ell_p(\hat{\eta}) - \frac{1}{2}c_1(1 - \alpha)$, where $c_1(1 - \alpha)$ is the $1 - \alpha$ quantile of the χ_1^2 distribution. Figure 1 shows the profile log likelihood, with horizontal lines indicating the levels determining the 95%, 99% and 99.9% confidence regions, namely $(133, \infty]$, $(129, \infty]$ and $(126, \infty]$ years. In this case $\hat{\eta}$ lies on the boundary of the parameter space and the standard theory does not apply, but as a more appropriate theory would yield intervals with higher lower limits, those here can be regarded as conservative.

The results based on the full dataset disagree with those in certain other recent articles, which appear to establish that η is finite. For example, Einmahl *et al.* (2017) analyse data on the ages at death of all Dutch residents, born in the Netherlands, who died over the years 1986–2015 aged at least 92 years. Table 1 is taken from their paper, in which they write

> We sort the data according to the year of *death* not birth. In this way, we can compare recent years instead of (birth) years in the 19th century. We see the women/men who died in such a given year as a random sample from the imaginary population of all women/men who could have (been born and who) died in that given year.

Einmahl *et al.* (2017) apply standard extreme-value methods to the deaths in each calendar year, estimate the shape parameter ξ and assess the evidence for a finite endpoint to the lifetime distribution. Their results for deaths of men and women, analysed separately for each of 30 calendar years, give strong evidence that the estimated shape parameter is negative, with most of their 60 estimates lying in the interval $(-0.2, -0.1)$, and they conclude that although as a whole the age distribution is moving to the right, its upper endpoint is not. Their average estimate is $\hat{\eta} \approx 114.1$ years for men, with upper one-sided 95% confidence bound roughly 119.6 years, and corresponding values 115.7 and 120.3 years for women; their highest estimates for men and women are 123.7 and 124.7 years. As these estimates are based on the 1500 and 1000 highest ages attained annually by women and men, they are rather precise—with 95% intervals very different from those in Figure 1. This is hardly surprising, since the oldest woman in the Dutch data dies aged 115 years, while three of those contributing to Figure 1 are older than
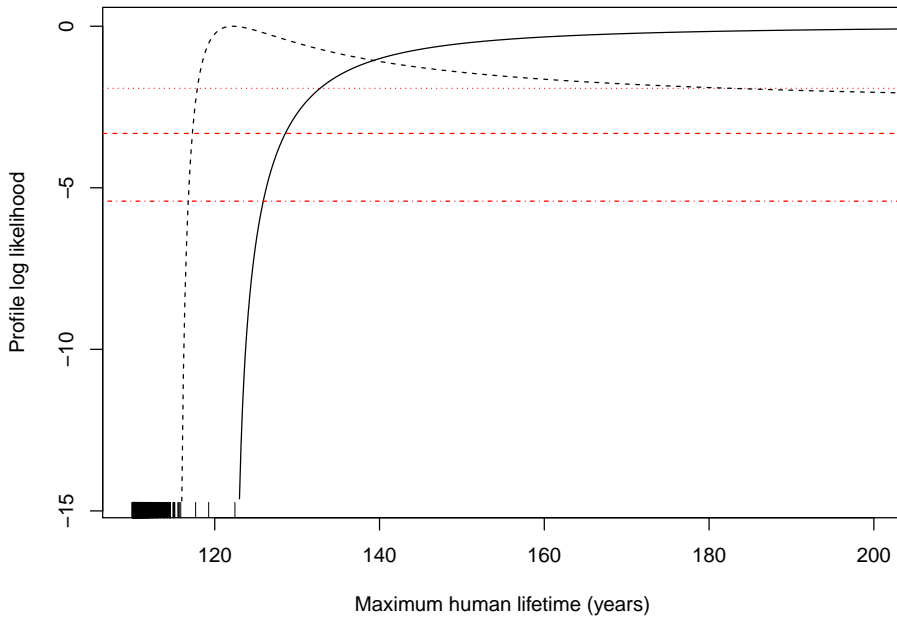
Figure 1: Profile log likelihood (solid) for the maximum attainable age, based on fitting an extremal model to 621 records in the International Database on Longevity. The horizontal lines indicate 95%, 99% and 99.9% confidence regions, given by the sets for which the profile log likelihood exceeds the dotted, dashed and dot-dashed horizontal lines, respectively. The rug at the lower left shows the observed lifetimes, the largest, for Jeanne Calment, is 122 years and 164 days. The dashed line is the profile log likelihood when the three largest observations are dropped.

116. If these three individuals are dropped from their data, then the maximum likelihood estimates for my fit of the Rootzén and Zholud model are $\hat{\xi} = -0.13$ (0.05) and $\hat{\sigma} = 1.57$ (0.097). Figure 1 shows the corresponding profile log likelihood, which gives $\hat{\eta} = 122.2$ years, with 95% confidence interval $(118, 182)$ years. As so often in extremal problems, the inference depends very strongly on the highest observations—which, in this case, are well separated from the others. In view of the care with which these observations were checked, one should not simply ignore them, but detailed validation is essential, as stressed by Rootzén and Zholud. The data on US deaths from the Human Mortality Database (http://www.mortality.org) used by Feifel *et al.* (2017), for example, are adjusted using actuarial techniques and seem to have some unrealistically high excess ages.

Table 1: Summary of the data used in Einmahl *et al.* (2017) on deaths of Dutch residents, born in the Netherlands, from 1986–2015. Each cell contains the number of women/men who died in the given age categories, in five-year periods.

	Age at death	Year of death					
	(years)	86–90	91–95	96–00	01–05	06–10	11–15
Women	113–115	0	0	0	2	0	0
	110–112	2	2	3	2	6	3
	107–109	20	17	25	27	28	53
	104–106	97	148	178	198	286	393
	101–103	604	945	1038	1251	1506	1961
	98–100	2383	3313	3992	4650	5487	6914
	95–97	6914	9122	11013	12560	13924	16635
	92–94	11647	14538	17623	19210	20819	24730
Men	110–112	0	0	0	1	0	1
	107–109	6	8	6	7	3	7
	104–106	47	50	32	38	31	57
	101–103	246	265	254	247	270	359
	98–100	1022	1154	1040	1107	1196	1554
	95–97	3060	3196	3275	3349	3752	4793
	92–94	5369	5684	5914	6374	7268	9634

Mixtures

One possible explanation for the conflicting conclusions is that it is unreasonable to extrapolate from the very old persons in the lower rows of Table 1 to supercentenarians—perhaps we are dealing with a mixture distribution. Only 22 persons in Table 1 live to over 110 years, and although it would be foolhardy to base an extremal analysis on them, it would be valuable to contrast them with the data of Rootzén and Zholud. The full Dutch data were unavailable, but a very crude reality check is as follows. If the probability that a living supercentenarian dies in the next year is $1/2$, then her probability of dying in the next three years is $7/8$. Working down from the penultimate to the youngest age-groups in Table 1, and for all columns combined, the corresponding ratios for the women are 0.9, 0.89, 0.87, 0.83, 0.75, 0.66, and 0.51, and for the men they are 0.95, 0.87, 0.85, 0.79, 0.70 and 0.57. The ratios for the highest age groups, where the numbers are fairly stable across the columns, are strikingly close to the value 0.875 predicted by Rootzén and Zholud; the others cannot be validly compared, both because the ages are not similar to those of supercentenarians, and because of non-stationarity, as

discussed below. Authors such as Thatcher (1999) have suggested that a logistic hazard function may be appropriate at high ages. This could show a strong increase, as implied by the negative shape parameter of Einmahl *et al.* (2017), rising to a constant plateau corresponding to the residual lifetime at great age having an exponential distribution, as suggested by Rootzén and Zholud; such a hazard function could stem from a mixture, perhaps due to the effects of frailty.

Non-stationarity

Another possible explanation for the different conclusions is non-stationarity, which implies that the targets of inference are not the same. To appreciate this, let u represent a threshold for extreme old age and consider individuals who die aged $t+u$ in an interval $\mathcal{C} = [0, c]$ of calendar time. Any such individual must have passed u before time c, so suppose that individual lifetimes pass u, and are included in the sample, at rate $\nu(x)$, for $x \in (-\infty, c]$. Thus ν can depend on the ease with which individuals' ages can be verified—for example, if all the birth records for some period were lost, ν would be zero u years later—but it cannot depend on t. Suppose also that the excess lifetime $t > 0$ above u has density function $f(t)$, independent of when the threshold was breached. If individuals are independent, then deaths at age $t+u$ will occur at calendar time $s = x+t$ as a Poisson process of rate $\lambda(s, t) = \nu(s-t)f(t)$, so the death rate within $\mathcal{C}$ at age $t+u$ is

$$\int_0^c \nu(s-t)f(t)\,\mathrm{d}s = \int_{-t}^{c-t} \nu(x)f(t)\,\mathrm{d}x = f(t)\nu_{\mathcal{C}}(t),$$

say, and the distribution of lifetimes for those individuals dying in $\mathcal{C}$ has density

$$g(t) = \frac{f(t)\nu_{\mathcal{C}}(t)}{\int_0^\infty f(t)\nu_{\mathcal{C}}(t)\,\mathrm{d}t}, \quad t > 0.$$

If ν is constant, corresponding to individual lifetimes exceeding u at a steady rate, then $\nu_{\mathcal{C}}(t)$ is constant and $g(t) = f(t)$, but if $\nu(x)$ increases because of factors such as population growth and medical advances, then $\nu_{\mathcal{C}}(t)$ will decrease with t, and then g will have a shorter tail than f. The marked growth across the lower rows of Table 1 suggests that $\nu(x)$ is indeed increasing; the numbers in the top two rows are more or less constant, but seem too small for robust conclusions. Moreover $\nu_{\mathcal{C}}$ depends on $\mathcal{C}$ and thus will change over time.

Under non-stationarity a column of a mortality table such as Table 1 provides estimates not of $f(t)$ but of $g(t)$, whose upper tail is likely to be shorter. This need not imply that f and g have different shape parameters. Indeed, the Einmahl *et al.* (2017) results are based on the highest 1500 ages; as these appear to correspond to births across a decade or so, large changes in $\nu_{\mathcal{C}}$ seem unlikely. However, estimation will require more care even than usual: if, for example, $\nu(x) \propto \exp(ax)$ for $a > 0$, it is easy to check

5

that $g(t) \propto f(t) \exp(-at)$, so if f is exponential, its scale parameter reduces from σ to $\sigma/(1 + a\sigma)$.

It may be instructive to use this framework to derive the likelihood used by Rootzén and Zholud. As it is rare for an individual to die on calendar date $s = x + t$ at the very old age $u + t$, we can approximate the corresponding events $(s_j, t_j) = (x_j + t_j, t_j)$ for $j = 1, \ldots, n$ by a Poisson process observed for $(x + t, t) \in \mathcal{R}$, where $\mathcal{R} = \mathcal{C} \times \mathcal{T}$; typically $\mathcal{T} = [0, \infty)$. The corresponding likelihood,

$$\exp\left\{-\int_{\mathcal{R}} \lambda(s, t)\, \mathrm{d}s\mathrm{d}t\right\} \times \prod_j \lambda(s_j, t_j), \tag{1}$$

reduces to

$$\exp\left\{-\int_{\mathcal{T}} f(t)\nu_{\mathcal{C}}(t)\, \mathrm{d}t\right\} \times \prod_{j=1}^n \nu(x_j)f(t_j),$$

but this is not useful for inference without assumptions on ν.

In order to eliminate ν, consider the independent Poisson processes that fall into the sets $\mathcal{R}_x = \{(x + t, t) : (x + t, t) \in \mathcal{R}, t \geq 0\}$, $x \leq c$. If this makes you queasy, suppose that each x represents a day, over which $\nu(x)$ is constant and a conditional density is well-defined. The $\mathcal{R}_x$ partition $\mathcal{R}$. In terms of the survivor function $\mathcal{F}$ corresponding to f, the intensity function for events in $\mathcal{R}_x$ equals

$$\int_{-x}^{c-x} \nu(x)f(t)\, \mathrm{d}t \;=\; \nu(x)\{\mathcal{F}(-x) - \mathcal{F}(c - x)\}, \quad x < 0, \tag{2}$$

$$\int_0^{c-x} \nu(x)f(t)\, \mathrm{d}t \;=\; \nu(x)\{1 - \mathcal{F}(c - x)\}, \quad 0 < x < c. \tag{3}$$

To appreciate this, consider Figure 2, in which the individual starting from $x_A < 0$ does not die in $\mathcal{R}$, that starting from $x_B < 0$ is aged $u - x_B > 0$ at calendar time 0 and would be aged $u + c - x_B$ at calendar time c if she did not die within $\mathcal{R}$, and that entering $\mathcal{R}$ at calendar time $x_C \in (0, c)$ would be aged $u + c - x_C$ at calendar time c if she did not die within $\mathcal{R}$; individuals B and C correspond to (2) and (3) respectively. We can merge these two expressions in the single formula $\nu(x)[\mathcal{F}\{\max(0, -x)\} - \mathcal{F}(c - x)]$.

Now if N_x is the number of events in $\mathcal{R}_x$, then (1) can be rewritten as

$$\prod_x \left[\frac{\Lambda(\mathcal{R}_x)^{n_x}}{n_x!} \exp\{-\Lambda(\mathcal{R}_x)\} \times n_x! \prod_{(s_j, t_j) \in \mathcal{R}_x} \frac{\lambda(s_j, t_j)}{\Lambda(\mathcal{R}_x)}\right], \tag{4}$$

where for each x the first term corresponds to the probability of observing n_x events in $\mathcal{R}_x$, a Poisson probability, and the second term is the conditional density for their positions given that $N_x = n_x$; note that as $\lambda(s, t) = \nu(s - t)f(t)$,

$$\Lambda(\mathcal{R}_x) = \int_{x + t \in \mathcal{C}} \lambda(x + t, t)\, \mathrm{d}t = \nu(x) \int_{\max(0, -x)}^{c-x} f(t)\mathrm{d}t.$$

If the x intervals are short enough, then each of the $\mathcal{R}_x$ contains at most one event, so the product of the second terms in (4) reduces to

$$\prod_x \prod_{(x_j + t_j, t_j) \in \mathcal{R}_x} \frac{\nu(x_j)f(t_j)}{\nu(x_j) \int_{\max(0, -x_j)}^{c-x_j} f(t)\, \mathrm{d}t} = \prod_j \frac{f(t_j)}{\mathcal{F}\{\max(0, -x_j)\} - \mathcal{F}(c - x_j)}, \tag{5}$$
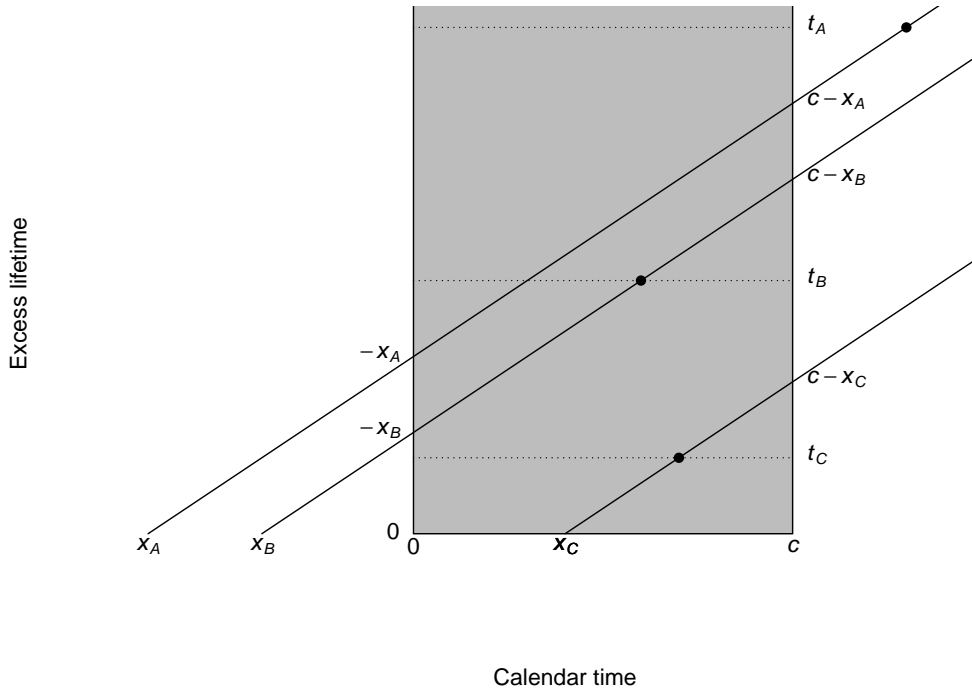
6

Figure 2: Lexis diagram illustrating likelihood contributions from individuals A, B and C (Gampe, 2010). The shaded region corresponds to $\mathcal{R} = \mathcal{C} \times \mathcal{T}$. Individual A passes age u at calendar time x_A but then dies outside $\mathcal{R}$, so makes no contribution to the likelihood. Individual B passes age u at calendar time $x_B < 0$ and could be observed to die within $\mathcal{R}$ between excess ages $-x_B$ and $c - x_B$. Individual C enters $\mathcal{R}$ at time $x_C \in (0, c)$ and could be observed to die at excess ages in the interval $(0, c - x_C)$.

because void $\mathcal{R}_x$ contribute factors of unity to the conditional likelihood, as would be the case for individual A in Figure 2. Expression (5) is a partial likelihood in the sense of Cox (1975); clearly it could be extended to account for censoring or more complex but uninformative sampling schemes. For example, it would be slightly more informative to use $\mathcal{F}(c - x_A)/\mathcal{F}(-x_A)$ as a likelihood contribution from individual A; this corresponds to the excess lifetime of A exceeding $c - x_A$, given that it exceeded $-x_A$.

As the conditional likelihood (5) eliminates $\nu(x)$, the vagaries of events such as the 2003 heatwave, which killed many elderly people in southern Europe, do not affect the conclusions. Exact removal of ν is only possible with data on individual births and deaths, though approximations are possible if data are not too heavily grouped.

Closing

Apart from our intrinsic interest in the possible span of a human life, the societal implications of demographic changes of the elderly make this an important topic, and it is appropriate that serious statistical attention is paid to it. But the long and the short of any statistical study is that high-quality data, including information about the sampling scheme, are essential, so it is disappointing that the database used by Rootzén and Zholud seems not to have been updated since 2010. Stronger conclusions would be possible from more comprehensive data.

Acknowledgement

References

Cox, D. R. (1975) Partial likelihood. *Biometrika* **62**, 269–76.

Einmahl, J. J., Einmahl, J. H. J. and de Haan, L. F. M. (2017) Limits to human life span through extreme value theory. Tilburg University Center for Economic Research Discussion Paper 2017-051, `https://pure.uvt.nl/portal/files/19635604/2017_051.pdf`.

Feifel, J., Genz, M. and Pauly, M. (2017) Who wants to live forever? An analysis of the maximum lifespan in the US. Preprint: Institute of Statistics, Ulm University.

Gampe, J. (2010) Human mortality beyond age 110. In *Supercentenarians*, eds H. Maier, J. Gampe, B. Jeune, J. W. Vaupel and J.-M. Robine, pp. 219–230. Berlin: Springer.

Rootzén, H. and Zholud, D. (2017) Human life is unlimited — but short. *Extremes* **20**, 713–728.

Thatcher, A. R. (1999) The long-term pattern of adult mortality and the highest attained age (with Discussion). *Journal of the Royal Statistical Society series A* **162**, 5–43.