Eur. Actuar. J. (2015) 5:139–163
DOI 10.1007/s13385-015-0105-8

ORIGINAL RESEARCH PAPER

Goodness-of-fit tests and applications for left-truncated Weibull distributions to non-life insurance

Markus Kreer · Ayşe Kızılersü · Anthony W. Thomas · Alfredo D. Egídio dos Reis

Received: 26 November 2014 / Revised: 25 January 2015 / Accepted: 5 February 2015 /
Published online: 24 March 2015

Abstract In risk theory with application to insurance, the identification of the relevant distributions for both the counting and the claim size processes from given observations is of major importance. In some situations left-truncated distributions can be used to model, not only the single claim severity, but also the inter-arrival times between two consecutive claims. We show that left-truncated Weibull distributions are particularly relevant, especially for the claim severity distribution. For that, we first demonstrate how the parameters can be estimated consistently from the data, and then show how a Kolmogorov-Smirnov goodness-of-fit test can be set up using modified critical values. These critical values are universal to all left-truncated Weibull distributions, independent of the actual Weibull parameters. To illustrate our findings we analyse three applications using real insurance data, one from a Swiss excess of loss treaty over automobile insurance, another from an American private passenger automobile insurance and a third from earthquake inter-arrival times in California.

M. Kreer
CAMPUSERVICE GmbH, Servicegesellschaft der Johann Wolfgang Goethe-Universität Frankfurt, Rossertstrasse 2, 60323 Frankfurt am Main, Germany
e-mail: mkreer@web.de

A. Kızılersü (✉) · A. W. Thomas
CSSM, Department of Physics, University of Adelaide, Adelaide 5005, Australia
e-mail: ayse.kizilersu@adelaide.edu.au

A. D. Egídio dos Reis
Department of Management, ISEG and CEMAPRE, Universidade de Lisboa, Rua do Quelhas 6, 1200-781 Lisboa, Portugal
e-mail: alfredo@iseg.ulisboa.pt

Keywords Maximum likelihood estimation · KS goodness-of-fit tests · Left-truncated Weibull distribution · Sparre-Andersen risk model · Claim sizes · Claim inter-arrival times · Anti-conservative test

1 Introduction and motivation

There are many practical situations in which there exists a region where some particular random variable of interest is either not observable or its measurements are unreliable, flawed or must be excluded for some other reason. Examples of such situations occur in the insurance business, for instance in automobile insurance, actual reported claims are influenced by bonus hunger behaviour and therefore catastrophe claims may not be fully accounted. We can look at [18, Chapter 14], for some insight into estimation for modified data in insurance. When studying the properties of the distribution from which a random variable is sampled, it is necessary to only consider data in the region of reliability and relevance. The process of selecting data only in certain parts of the domain and adjusting the probability distribution accordingly is known as truncation; for a detailed discussion of this see [17, pp. 551, section 32.15]. Truncation is often done either in the right or in the left tail of a distribution. Here we are only concerned with left-truncation.

In actuarial science there are several situations where it is necessary to perform left-truncation on a sample data set. An obvious case is the excess-of-loss reinsurance treaty (abbreviated as X-Loss) with a predefined monetary limit, from the point of view of the reinsurer. We study an example of automobile insurance claims with an excess of CHF 100,000 as in [19], the claim size distribution is left-truncated with a threshold level of CHF 100,000 (this is typical for a reinsurance contract). Other situations in the insurance business like contracts where a franchise deductible applies, see [7, 18, Chapter 14], may also require left-truncated distributions. These are further examples where left-truncated distributions can be used in the claim size problem.

There are other random quantities appearing in actuarial problems that might also require a left-truncation, such as the inter-arrival time between two consecutive claims of the same portfolio. As an example of this we studied the inter-arrival times of earthquakes where the left-truncation makes sure that aftershocks are not counted as earthquakes of their own right. Therefore, reinsurance treaties contain an "hours clauses" which define earthquakes occurring up to 3 or 7 days after the original earthquake as aftershocks and not earthquakes of their own right [31].

The above examples involving both left-truncation and heavy-tails motivated our research, and relates to the well known actuarial risk model in particular with the view of extending the classical Cramér-Lundberg model to the Sparre-Andersen risk model with non-Erlang inter-arrival times.

The Sparre-Andersen collective risk model with applications to insurance considers modelling the surplus or reserve of an insurance portfolio as a function of the aggregate claims process, denoted as $\{S(t), t \geq 0\}$, where $S(t)$ represents the aggregate claims occurring up to time $t \geq 0$,

$$S(t) = \sum_{n=0}^{N(t)} \xi_n \,, \; \xi_0 \equiv 0 \,. \tag{1}$$

The total claim amount $S(t)$ consists of a random sum of independent identically distributed (positive) claims ξ_n. $\{N(t)\}_{t \geq 0}$ is a renewal process modelling the number of claims and it is assumed to be independent of the claim size sequence $\{\xi_n\}_{n=1}^{\infty}$. The $\{T_n\}_{n \geq 1}, T_0 = 0$, are the points of the renewal process whose inter-arrival times $W_n = T_n - T_{n-1}$ have finite mean, and which are associated with the time when a loss ξ_n occurs. In this notation $t \in [T_n, T_{n+1})$ iff $N(t) = n$. The model has been recently developed from the classical compound Poisson, as presented in [12] for instance. Another key reference is [11]. More general renewal models, where the distribution of W_n is Erlang(n), generalized Erlang(n) or even to Phase-type PH(n) can be found in [16] or [6].

There is a large amount of literature on the properties of the risk process Eq. (1) extending classical Cramér-Lundberg theory. Further work has been reported on the counterpart model, known as the dual risk model, which has applications in finance. For the latter see, for instance, the recent references [1, 28]. There is no literature considering applications of left-truncated Weibull to either the claim size ξ_n or the waiting time W_n, having evidence of those applications from real data brings challenging problems.

In this research we investigate how we can identify from given data samples the distributional properties of either the claim size ξ_n or the inter-arrival times W_n. We restrict our attention to Weibull distributions [33], which are a flexible generalisation of the exponential distribution, can be either light-tailed (for $\beta \geq 1$) or heavy-tailed (for $0 < \beta < 1$), and have generated a lot of interest recently. The domain of the two-parameter Weibull distribution is from zero to infinity but when unobserved or unreliable data is removed, the distribution becomes a truncated Weibull distribution. Of particular interest is the left-truncated Weibull distribution, in which case all data less than or equal to a certain value x_L is excluded from analysis.

When considering left-truncated data samples in the context of parametric distributions, the problem of estimating the parameters arises and in particular whether or not the "quality" of the estimated parameters becomes better with increasing sample size. The quality of the estimated parameters is described in statistics in terms of consistency and asymptotic efficiency. In addition, a goodness-of-fit measure, such as the Kolmogorov-Smirnov (KS) test, measures the level of confidence with which we assert that the left-truncated data is sampled from a predefined left-truncated parametric distribution. If the parameters of the distribution are estimated from the data sample itself then the standard critical values (CVs) of Kolmogorov-Smrinov test cannot be applied, as discussed in [9]. In this case the CVs will depend in general on the distribution and the method for obtaining its parameters.

Throughout this paper we use the term *complete* Weibull distribution to refer to an untruncated Weibull distribution. The literature on data analysis tends to focus either on complete or censored data, with much less attention paid to truncated data.

There are a few consequences of dealing with a truncated data set, as opposed to a complete set. One of the consequences is that the estimation of the Weibull parameters is more complicated and has no known analytic form. It is therefore necessary to check the conditions under which non-trivial solutions exist, an issue which is addressed in Sect. 2. Another consequence of having truncated data is that the critical values depend, in a small but measurable way, on the left-truncation value. If this effect is ignored it can lead to incorrect conclusions from hypothesis testing. We quantify the effect that using "conservative" and "anti-conservative" CVs have on the KS hypothesis testing pass rate in the results Sect. 4. As there is no closed-form solution to finding the CVs for a left-truncated Weibull distribution whose parameters are determined from the data, it is necessary to perform Monte-Carlo simulations to estimate their values.

Consider the random variable, say X, that may denote either the claim size or the inter-arrival times. We are interested in i.i.d. samples $X_1, X_2, \ldots, X_n$ satisfying $x_L < X_i$ for $i = 1, 2, \ldots, n$ for some positive x_L and some integer n. The cumulative distribution function (cdf) for the left-truncated Weibull random variable X with scale, shape, and left-truncation parameters, $\alpha > 0, \beta > 0$, and $x_L > 0$ respectively has been given by [34] and is[1]

$$F(x|\alpha, \beta, x_L) = 1 - \exp\left[\left(\frac{x_L}{\alpha}\right)^{\beta} - \left(\frac{x}{\alpha}\right)^{\beta}\right] \text{ for } x > x_L\,, \tag{2}$$

consequently the left-truncated probability density function (pdf) is

$$f(x|\alpha, \beta, x_L) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left[\left(\frac{x_L}{\alpha}\right)^{\beta} - \left(\frac{x}{\alpha}\right)^{\beta}\right] \text{ for } x > x_L\,. \tag{3}$$

In the limit $x_L \longrightarrow 0+$ the left-truncated Weibull cdf and pdf Eqs. (2) and (3) are the complete Weibull cdf and pdf, as expected.

Note that the left-truncated Weibull pdf, Eq. (3), is continuously differentiable in the argument x and its two positive finite parameters α and β to any order and thus $f \in C^{\infty}((x_L, \infty) \times (0, \infty) \times (0, \infty))$. Furthermore, f and all its derivatives with respect to x, α, β vanish in the limit $x \rightarrow +\infty$, at least like $\exp\left[-(x/\alpha)^{\beta'}\right]$ for $\alpha > 0$ and any $\beta' \in (0, \beta)$.

Given a sample $X_1, X_2, \ldots, X_n$ i.i.d. satisfying $x_L < X_i$ for $i = 1, 2, \ldots, n$, one can test the hypothesis that they come from a left-truncated Weibull distribution using a standard Kolmogorov-Smirnov (KS) test provided the Weibull parameters α, β and truncation point x_L are known a priori [29]. By contrast if any of the Weibull parameters are unknown and must be estimated from the sample data itself, then the critical values of the KS test might depend on the true values of the Weibull distribution and the method for estimating them. In this case the critical values for the KS statistics can only be determined from a Monte-Carlo study and the results

[1] In our study the left-truncation parameter $x_L > 0$ is known and not to be confused with the unknown location parameter x_0 of the 3-parameter Weibull distribution. Here, we only consider 2-parameter Weibull distributions.

presented in tabular form. If the parameters of the distribution are known a priori then the CVs are larger than if they are unknown.

In most practical situations (and also our case), the left-truncation value $x_L > 0$ is known and only the scale α and/or shape β parameter(s) must be estimated. In this paper we present results for KS tests for two types of left-truncated Weibull distributions:

- Case I deals with unknown scale α and shape β parameters, and
- Case II deals with unknown scale α but known shape β parameter.

In both cases the parameters are estimated using the maximum likelihood estimation (MLE) procedure. Case II actually falls under the well-studied exponential family, see [5] for details.

It was noted in the 1950s by [10], that if the sample size is "small" the MLE of the parameter in a right-truncated data set does not always possess a non-trivial solution. Therefore there is no reason to assume that the MLE of the Weibull parameters will always possess a non-trivial solution. We address the issue of obtaining non-trivial parameter estimates in Sect. 2 and provide a method for determining whether a set of data is sampled from a left-truncated Weibull distribution with a known truncation value for the two cases outlined above. We also show that the maximum likelihood estimate of scale and shape parameters of a left-truncated Weibull distribution are consistent, asymptotically normal, and efficient.

We outline the KS test for a left-truncated Weibull distribution with a known left-truncation value when the parameters are estimated from the sample data in Sect. 3.

In Sect. 4 we develop a KS test with modified critical values to test the hypothesis that the sampled data set cannot be rejected as coming from a left-truncated Weibull distribution. The modified critical values are calculated from Monte Carlo simulations and presented in table form for a range of sample sizes and truncation values.[2]

In Sect. 5 we illustrate our findings by working applications of our anti-conservative KS tests on automobile claims data from both Switzerland and US and also for the inter-arrival times of earthquakes in California. Finally, we summarise our findings in Sect. 6.

2 Maximum likelihood estimation

It follows from Eq. (3) that the logarithm of the likelihood function for the left-truncated Weibull distribution is

[2] Due to a scaling property of the Weibull distribution, first noted for the complete Weibull sample by [32], the modified critical values are independent of various parameter combinations of α and β. Our results show that although the CVs depend predominantly on the sample size, there is also a slight dependence on the truncation value. This dependence leads us to two sets of critical values: one anti-conservative and one conservative test. The difference between using conservative and anti-conservative CVs leads to an error of the first kind less than 2%.

$$\begin{aligned}\ell(X_1, X_2, \ldots, X_n|\alpha, \beta, x_L) &\equiv \log\left(\prod_{i=1}^{n} f(X_i|\alpha, \beta, x_L)\right)\\ &= n\log\beta - n\beta\log\alpha + (\beta-1)\sum_{i=1}^{n}\log X_i - \sum_{i=1}^{n}\left(\frac{X_i}{\alpha}\right)^{\beta} + n\left(\frac{x_L}{\alpha}\right)^{\beta}.\end{aligned} \tag{4}$$

The maximization condition for α, $\partial\ell/\partial\alpha = 0$, yields

$$\alpha^{\beta} = \frac{x_L^{\beta}}{n}\sum_{i=1}^{n}\left[\left(\frac{X_i}{x_L}\right)^{\beta} - 1\right] \tag{5}$$

and similarly the maximization condition for β, $\partial\ell/\partial\beta = 0$ (after using Eq. (5) to eliminate α) is (see also [34])

$$0 = \frac{1}{\beta} - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i}{x_L}\right)^{\beta}\log\frac{X_i}{x_L}}{\frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{X_i}{x_L}\right)^{\beta} - 1\right]} + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}. \tag{6}$$

The solutions of Eqs. (5) and (6) are denoted by $\hat{\beta}_n = \hat{\beta}(X_1, \ldots, X_n|x_L)$ and $\hat{\alpha}_n = \hat{\alpha}(X_1, \ldots, X_n|x_L)$. For convenience we shall suppress the dependence on the sample $X_1, \ldots, X_n$ and the left-truncation value x_L and simply write $\hat{\alpha}_n$ and $\hat{\beta}_n$. Sometimes we might even drop the n and speak of $\hat{\alpha}$ and $\hat{\beta}$.

The necessary and sufficient conditions for the existence and uniqueness of $\hat{\alpha}$ and $\hat{\beta}$ are determined by the following lemma. This lemma was motivated by the work of [3, 13, 20, Example 6.1], all of which dealt with the complete (untruncated) Weibull distribution only. We prove this lemma for the left-truncated Weibull distribution.

Lemma 1 *For a left-truncated Weibull sample* $0 < x_L < X_i$ *for* $i = 1, 2, \ldots, n$ *define the function* $h : (0, \infty) \to \mathbb{R}$ *by*

$$h(\beta) = \frac{1}{\beta} - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i}{x_L}\right)^{\beta}\log\frac{X_i}{x_L}}{\frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{X_i}{x_L}\right)^{\beta} - 1\right]} + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}. \tag{7}$$

Then the following holds true:

1. *The function* $h(\cdot)$ *is monotonically decreasing.*
2. $\lim_{\beta\to+\infty} h(\beta) < 0$
3. $\lim_{\beta\to 0+} h(\beta) > 0$ *if and only if*

$$2\cdot\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\log^2\frac{X_i}{x_L} > 0\,. \tag{8}$$

Proof: The Proof is given in Appendix 1—"Proof of Lemma 1". □

From Lemma 1 we see immediately that the solution of the MLE problem is equivalent to finding a zero of $h(\cdot)$. This is possible if and only if Eq. (8) is satisfied, because $\lim_{\beta\to 0+} h(\beta)$ is positive, $\lim_{\beta\to+\infty} h(\beta)$ is negative and $h(\cdot)$ is a monotonically decreasing continuous function so the intermediate value theorem asserts the existence of a unique zero. This unique zero, when it exists, is our $\hat{\beta}$ and from Eq. (5) we obtain the corresponding $\hat{\alpha}$.

The next theorem deals with consistency, asymptotic normality and efficiency of the MLE method, i.e. the limiting behaviour as $n\to\infty$. This topic is discussed in [20] for a general case of pdfs satisfying certain regularity criteria. For the left-truncated Weibull pdf, Eq. (3), we adapt Theorem 5.1 of Sect. 6.5 (p. 463) of [20], to our notation for Case I. This result has been informally mentioned in [30] as a standard result in estimation theory, albeit for the complete Weibull distribution. For the remainder of this paper we denote the true value of α by α^0 and the true value of β by β^0.

Theorem 1 (Theorem 5.1 in Lehmann and Casella [20])

Let the sample $X_1,\ldots,X_n$ be i.i.d. each with a probability density function given by Eq. (3) *which satisfies conditions (A0)–(A2) and (A)–(D) given in* Appendix 1—"Proof of Theorem 1". *Then with probability tending to 1 as $n\to\infty$ there exist solutions of the likelihood equations such that*[3]

1. *The vector $(\hat{\alpha}_n,\hat{\beta}_n)$ is consistent for estimating (α^0,β^0).*
2. $\sqrt{n}\Big((\hat{\alpha}_n,\hat{\beta}_n)-(\alpha^0,\beta^0)\Big)$ *is asymptotically normal with vector mean zero and covariance matrix $[Z((\alpha^0,\beta^0))]^{-1}$ being the inverse of the Fisher information matrix*

$$Z(\alpha^0,\beta^0) = -\begin{bmatrix} \mathbb{E}\left(\frac{\partial^2}{\partial\alpha^2}\log f(X|\alpha^0,\beta^0,x_L)\right) & \mathbb{E}\left(\frac{\partial^2}{\partial\alpha\partial\beta}\log f(X|\alpha^0,\beta^0,x_L)\right) \\ \mathbb{E}\left(\frac{\partial^2}{\partial\beta\partial\alpha}\log f(X|\alpha^0,\beta^0,x_L)\right) & \mathbb{E}\left(\frac{\partial^2}{\partial\beta^2}\log f(X|,\alpha^0,\beta^0,x_L)\right) \end{bmatrix}$$

3. *$(\hat{\alpha}_n,\hat{\beta}_n)$ is asymptotically efficient in the sense that as $n\to\infty$*

$$\sqrt{n}(\hat{\alpha}_n-\alpha^0)\overset{\text{distrib.}}{\longrightarrow} N(0,[Z(\alpha^0,\beta^0)]^{-1}_{\alpha,\alpha})$$

$$\sqrt{n}(\hat{\beta}_n-\beta^0)\overset{\text{distrib.}}{\longrightarrow} N(0,[Z(\alpha^0,\beta^0)]^{-1}_{\beta,\beta})$$

Proof: This theorem is a specific case of the original result given in the general Theorem 5.1 of section 6.5 (p. 463) in [20], and adapted to our situation of left-

[3] Statements 1. and 3. actually follow for MLE from statement 2. but we have kept them in the theorem to be consistent with Lehmann & Casella [20]).

truncated Weibull distributions. Thus, we only need to check conditions (A0)–(A2) of section 6.3 and assumptions (A)–(D) from section 6.5 in [20] for the left-truncated Weibull pdf f. The details of how to do this are given in Appendix 1—"Proof of Theorem 1". The key ingredients are: the smoothness property of pdf f in x and its two parameters α, β, and the strong decay property as $x \to \infty$. □

An analogous result holds for Case II. Note that when the shape parameter β is known the left-truncated Weibull pdf Eq. (3) belongs to the exponential family. We can then use the results for left-truncated exponential distributions of [4] and the more general results of [5].

3 Kolmogorov-Smirnov goodness-of-fit test for hypothesis testing

3.1 General description of the test

We want to test the following null hypothesis H_0: The i.i.d. sample $X_1, \ldots, X_n$ satisfying $x_L < X_i$ for $i = 1, \ldots, n$ for some positive x_L and some integer n, is drawn from a left-truncated Weibull distribution F_0 as given in Eq. (2) with estimated parameters $(\hat{\alpha}, \hat{\beta})$ obtained from MLE as discussed in the previous section.[4] With the definition of the empirical distribution function $F_n(X)$ as the proportion of those values of the order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ that are smaller than $X \in (x_L, \infty)$, the Kolmogorov-Smirnov (KS) test statistic is given by (e.g. [17], sect. 30.49)

$$\begin{aligned} D_n &= \sup_{x_L < X < \infty} \{F_n(X) - F_0(X), F_0(X) - F_n(X)\} \\ &= \max_{1 \le j \le n} \left\{ \frac{j}{n} - F_0(X_{(j)}), F_0(X_{(j)}) - \frac{j-1}{n} \right\}. \end{aligned} \tag{9}$$

Here D_n is the KS distance which is compared with a critical value $D_{cv}(n, \rho, p_H)$, that depends on the sample size n, the truncation level ρ (the theoretical percentage removed from the untruncated distribution) and significance level p_H (in this paper we only consider $p_H = 0.05$). The hypothesis H_0 can not be rejected if $D_n < D_{cv}(n, \rho, p_H)$.

When parameters are estimated from a sample itself, the probability integral transformation of the sample variables destroys their independence [9], and the Kolmogorov-Smirnov argument in Eq. (9) leads to distribution-dependent critical values. If the parameters of a left-truncated Weibull distribution are estimated from the sample data itself then the KS distance becomes

$$\begin{aligned} D_n = \max_{1 \le j \le n} \Bigg\{ &\frac{j}{n} - 1 + \exp\left[\left(\frac{x_L}{\hat{\alpha}}\right)^{\hat{\beta}} - \left(\frac{X_{(j)}}{\hat{\alpha}}\right)^{\hat{\beta}} \right], \\ &1 - \exp\left[\left(\frac{x_L}{\hat{\alpha}}\right)^{\hat{\beta}} - \left(\frac{X_{(j)}}{\hat{\alpha}}\right)^{\hat{\beta}} \right] - \frac{j-1}{n} \Bigg\} \quad , \end{aligned}$$

[4] From this point onwards we will drop the index n and use $\hat{\alpha}$ and $\hat{\beta}$.

which can be written, using Eq. (20) for the left-truncated Weibull random variables X_j, as

$$D_n = \max_{1 \le j \le n} \left\{ \frac{j-n}{n} + \exp\left[\left(\eta^{\hat{\beta}/\beta^0} - \left(\eta + y_{(j)}\right)^{\hat{\beta}/\beta^0} \right) \left(\frac{\alpha^0}{\hat{\alpha}} \right)^{\hat{\beta}} \right], \right.$$
$$\left. \frac{n+1-j}{n} - \exp\left[\left(\eta^{\hat{\beta}/\beta^0} - \left(\eta + y_{(j)}\right)^{\hat{\beta}/\beta^0} \right) \left(\frac{\alpha^0}{\hat{\alpha}} \right)^{\hat{\beta}} \right] \right\} \tag{10}$$

where $y_{(j)}$ is the jth order statistic of a sample of standard exponential random variates as described in Appendix 2, (α^0, β^0) are the true parameters as introduced in the previous section and $\eta \equiv (x_L/\alpha^0)^{\beta^0}$ is the new transformed truncation parameter.

In the next subsection we argue why it is sufficient to study only Weibull distributions with parameter vector $(\alpha^0, \beta^0) = (1, 1)$ to obtain critical values for our KS test.

3.2 Why do we study Weibull distributions with $\alpha^0 = 1$, $\beta^0 = 1$?

Following [32] we denote for general Weibull distributions with any positive (α^0, β^0) the random variables $(\alpha^0/\hat{\alpha})^{\hat{\beta}}$ and $\hat{\beta}/\beta^0$ as *pivotal functions*. Note that the KS distance D_n in Eq. (10) depends on these pivotal functions, n and η alone. Consequently, D_n is "universal" for different combinations of (α^0, β^0) for same n and η, provided the following holds true

$$\left(\frac{\alpha^0}{\hat{\alpha}_{(\alpha^0,\beta^0)}} \right)^{\hat{\beta}_{(\alpha^0,\beta^0)}} \overset{distrib.}{=} \left(\frac{1}{\hat{\alpha}_{(1,1)}} \right)^{\hat{\beta}_{(1,1)}}$$
$$\left(\frac{\hat{\beta}_{(\alpha^0,\beta^0)}}{\beta^0} \right) \overset{distrib.}{=} \hat{\beta}_{(1,1)} \tag{11}$$

where $\hat{\alpha}_{(1,1)}$ and $\hat{\beta}_{(1,1)}$ are the MLE estimates originating from the simplest choice of a Weibull distribution with $(\alpha^0, \beta^0) = (1, 1)$. Likewise $\hat{\alpha} = \hat{\alpha}_{(\alpha^0,\beta^0)}$ and $\hat{\beta} = \hat{\beta}_{(\alpha^0,\beta^0)}$ are the MLE estimates originating from a Weibull distribution for arbitrary positive (α^0, β^0). The latter equality in distribution, Eq. (11), is demonstrated in Appendix 3.

For untruncated data where $x_L = 0$ (thus η and $\rho = 1 - e^{-\eta}$ both vanish) $D_{cv}(n, 0, 0.05)$ will only depend on n. This was observed by [32] and allowed [22] and [26] to produce confidence tables for in-sample KS tests with MLE equations solved for standard exponential random variates. Similarly, in Case II when the shape parameter β is known, Eq. (10) simplifies and becomes independent of truncation, x_L, (and also of the truncation level $\rho = 1 - e^{-\eta}$) because $\hat{\beta}/\beta^0 = 1$ and the η-terms cancel out. Only for the Case I we do need to investigate the dependence of $D_{cv}(n, \rho, 0.05)$ on the parameter η ($\eta = x_L$ if $\alpha^0 = \beta^0 = 1$) and n in greater detail.

4 Critical values for Kolmogorov-Smirnov test

4.1 The Monte Carlo simulation

Before presenting our results we outline the procedure used to generate critical values

1. Fix $n \in \{30, 50, 100, 200, 500, 1000\}$, $\alpha^0 = 1$, $\beta^0 = 1$ and $\eta = x_L \in \{0,\ 0.11,\ 0.24,\ 0.37,\ 0.51,\ 0.70,\ 0.92, 1.20,\ 1.61,\ 2.31\}$. As can be confirmed from Eq. (2) with $\alpha^0 = \beta^0 = 1$ these η-values correspond to average truncation rates $\rho = 0, 0.1, \ldots, 0.9$. Note that there is a one-to-one correspondence between η and ρ, namely $\rho = 1 - e^{-\eta}$.
2. For $i = 1, \ldots, n$, draw left-truncated Weibull random variates $x_L < X_i$ using Eq. (20).
3. Solve Eqs. (5)–(6) to obtain $\hat{\beta}_n$ and $\hat{\alpha}_n$.
4. Calculate $D_n(\eta)$ from Eq. (10) and store in a list.
5. For given n and η repeat steps 2 to 4 $N_{MC} = 1{,}000$ times.
6. Sort list of $D_n(\eta)$ in ascending order; for p_H =0.05 choose the value of $D_n(\eta)$ ranking between 950 and 951 as $D_{cv}(n, \rho, p_H = 0.05)$. The a-priori error of the quantile is $\sigma^2_{D_{cv}(n,\rho,p_H)} = p_H \cdot (1 - p_H)/(N_{MC} f(D_{cv}(n, \rho, p_H))^2)$ [24], where the density function $f(D_{cv}(n, \rho, p_H))$ is estimated from the sample using the Freedman-Diaconis rule [14] for the optimal bin-size of the histogram.
7. Repeat steps 2–6 for every n and η (ρ respectively), 100 times to obtain mean values and standard deviation for $D_{cv}(n, \rho, p_H)$.

4.2 The results

Figure 1 depicts the average critical values for Case I as a function of truncation level ρ for various values of n. Each of these plots show the same truncation dependence of the critical values namely as the truncation level, ρ, increases, the CV's decreases until they reach their minima then they increase with the increasing truncation level.

If one ignores the truncation dependence and assumes that the critical values are independent of ρ then the pass rates will be affected, while for certain truncation values the pass rates are too high and for others they will be too low. For example,

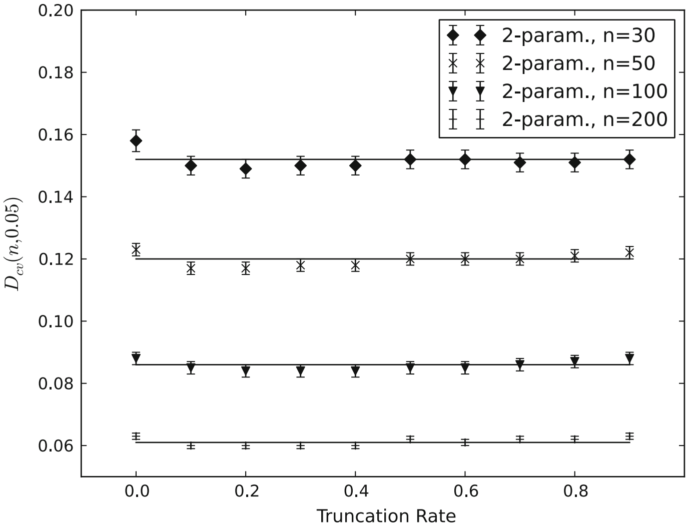

Fig. 1 Critical values $D_{cv}(n, \rho, 0.05)$ versus the truncation level ρ for various sample sizes n. The *straight line* corresponds to $D_{cv}^{(ave)}(n, 0.05)$ which is the average over all ρ values. The complete set which corresponds to $\rho = 0$, is the highest critical value

the consequence of using $D_{cv}^{(ave)}(n, 0.05)$ by forming an average of critical values over all possible truncation levels to perform hypothesis test on the data we find that 93.5% of the KS statistics are less than the critical value, i.e., the critical value is too small. We refer to this as the *anti-conservative* critical value. By contrast if we use the critical value for the complete case, $D_{cv}^{(0)}(n, 0.05)$, we find that in the worst case 97.5% of the KS statistics are less than this critical value, we refer to this as *conservative* critical value. The critical values for different values of n are given in Table 1.

In Case II, where the shape parameter β is known, i.e., $\hat{\beta} = \beta$, the critical values are *independent* of truncation as discussed in Section 3. Therefore $D_{cv}^{(ave)}(n, 0.05)$ is the *universal critical value* for a given n and for all truncation levels, ρ.

Table 1 summarises the critical values for Case I and Case II for a range of sample sizes n. It also gives an interpolation formula of $D_{cv}^{(ave)}(n, 0.05)$, which was motivated by Fig. 2 and Miller's formula [23] for distributions with known parameters.

The values of the complete sample, $D_{cv}^{(0)}(n, 0.05)$ in Table 1 agree with those in [8, 22] and [26] for Case I. The complete critical values for Case II agree with [8, 21] and [35]. As mentioned above, using these values of the untruncated sample for any truncated Weibull sample correspond to a conservative KS-test in which we accept the hypothesis (H_0) more often than we should.

Following [2] we also conducted a power-testing study. The test performs well for log-Cauchy, Pareto, log-double-exponential and log-logistic, namely one can

Table 1 Summary of critical values for Case I and Case II

Sample size n	Case I		Case II
	$D_{cv}^{(ave)}(n, 0.05)$ (anti-conservative)	$D_{cv}^{(0)}(n, 0.05)$ (conservative)	$D_{cv}^{(ave)}(n, 0.05)$
30	0.152 ± 0.003	0.157 ± 0.002	0.193 ± 0.001
50	0.120 ± 0.002	0.123 ± 0.001	0.151 ± 0.001
100	0.086 ± 0.001	0.088 ± 0.001	0.108 ± 0.001
200	0.061 ± 0.001	0.063 ± 0.001	0.076 ± 0.000
500	0.039 ± 0.001	0.040 ± 0.001	0.049 ± 0.000
1000	0.028 ± 0.001	0.028 ± 0.001	0.034 ± 0.000
$30 \leq n < \infty$	$\frac{0.886 \pm 0.004}{\sqrt{n}} - \frac{0.283 \pm 0.031}{n}$	$\frac{0.905 \pm 0.003}{\sqrt{n}} - \frac{0.246 \pm 0.026}{n}$	$\frac{1.094 \pm 0.001}{\sqrt{n}} - \frac{0.193 \pm 0.005}{n}$

rule out these distributions as candidates explaining the data set. On the other hand, we found that the power-testing does have problems ruling out χ^2-distributions with 1, 3 and 4 degrees of freedom and also log-normal distributions. The latter can be ruled out by a likelihood ratio test.

5 Applications to insurance data

In this section we apply our methods and show the importance of left-truncated Weibull distributions fitting to insurance data. The first and second examples deal with truncated claim-size distributions for ξ_n and the third example with the relevant inter-arrival times W_n for the counting process of claims.

5.1 Automotive data from Swiss insurance with excess of loss treaty

Table 2 represents a set of automotive data from a Swiss Insurance company as given in [19] with left-truncation of CHF 100,000. Results of our tests are given in Table 3[5]. Here and in the subsequent examples, we calculate the Kolmogorov-Smrinov distance D_n as in Eq. (9) and the anti-conservative critical value $D_{cv}^{(ave)}(n, 0.05)$ for given sample size n from the interpolation formula for Case I in Table 1. Errors are estimated using Theorem 1 statements 2. and 3. in conjunction with the findings of Appendix 4.

From our above analysis we may conclude with 95%-confidence level that the 33 automotive claims corresponding to the ξ_n from Eq. (1) with left-truncation of CHF 100,000 are Weibull-distributed with $\hat{\alpha} = 4910 \pm 2610$ CHF and $\hat{\beta} = 0.35 \pm 0.05$.

[5] In the table, the theoretical truncated percentage is estimated as $\approx 90\%$ using the relation $\rho = 1 - e^{-\eta}$ by assuming that the untruncated data set is Weibull. Since we do not have the complete data set and the sample size is small this value should be considered a very rough approximation.

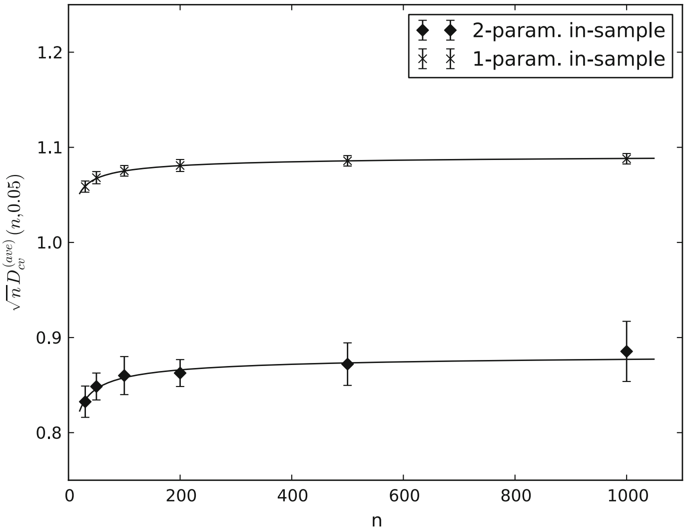

Fig. 2 Critical values $\sqrt{n}D_{cv}^{(ave)}(n, 0.05)$ as a function of various sample sizes n for Case I and Case II

Table 2 Automotive data with excess of loss treaty of CHF 100,000 from [19]

103,765	109,168	112,341	113,800	114,791	115,731
118,264	123,464	127,611	133,504	142,821	152,270
163,491	164,968	168,915	169,346	172,668	191,954
193,102	208,522	209,070	219,111	243,910	280,302
313,898	330,461	418,074	516,218	595,310	742,198
791,874	822,787	1,074,499			

5.2 Automotive claims data from US insurance

Next we investigate claims data "from a large midwestern US property and casualty insurer for private passenger automobile insurance" ([15]). The data set is much larger than in the previous example with altogether 6,773 events and provides the amout paid on a closed claim in US dollars for a certain business year. Each claim event contains as additional information the policy anonymized holder's state, sex, age and rating class. The latter is based on age, gender, marital status and use of vehicle. We analyse the data only with respect to gender and rating class (ignoring age and state). We present our analysis here for two rating classes, C1B and C71 and for both male (M) and female (F) policy holders respectively in Tables 4 and 5. For the other rating classes, fairly similar results hold. A symbol "-" in column x_L indicates a complete sample without truncation.

Again, we may conclude with 95%-confidence level that the left-truncated claim data corresponding to the ξ_n from Eq. (1) are Weibull-distributed with parameters $\hat{\alpha}$ and $\hat{\beta}$ nearly independent of the truncation paramter x_L, but depending on sex and rating class. If the sample size n is large enough, we obtain consistent parameter

Table 3 Claim size with ρ the truncation level

x_L [CHF]	n	$\hat{\alpha}$ [CHF]	$\hat{\beta}$	$\rho[\%]$	D_n	$D_{cv}^{(ave)}(n, 0.05)$	H_0
100,000	33	4910 ± 2610	0.35 ± 0.05	≈ 90	0.0807	0.1457	Accept

Table 4 Claim size with truncation level ρ for rating class C1B

Sex	x_L [USD]	n	$\hat{\alpha}$ [USD]	$\hat{\beta}$	$\rho[\%]$	D_n	$D_{cv}^{(ave)}$ (n, 0.05)	H_0
F	–	165	1,770 ± 150	0.96 ± 0.06	0	0.1165	0.0673	Reject
F	400	142	540 ± 100	0.54 ± 0.04	14	0.0624	0.0724	Accept
F	500	126	630 ± 100	0.57 ± 0.04	24	0.0720	0.0767	Accept
F	1,000	80	400 ± 100	0.50 ± 0.04	52	0.0635	0.0955	Accept
F	2,000	44	210 ± 100	0.44 ± 0.05	73	0.0870	0.1271	Accept
M	–	259	1,920 ± 130	0.94 ± 0.05	0	0.0829	0.0540	Reject
M	400	219	850 ± 100	0.60 ± 0.03	15	0.0493	0.0586	Accept
M	500	200	870 ± 110	0.61 ± 0.03	23	0.0532	0.0612	Accept
M	1,000	133	870 ± 130	0.60 ± 0.04	49	0.0379	0.0747	Accept
M	2,000	77	670 ± 140	0.56 ± 0.05	70	0.0618	0.0973	Accept

Table 5 Claim size with truncation level ρ for rating class C71

Sex	x_L [USD]	n	$\hat{\alpha}$ [USD]	$\hat{\beta}$	$\rho[\%]$	D_n	$D_{cv}^{(ave)}$ (n, 0.05)	H_0
F	–	415	1,880 ± 100	0.94 ± 0.04	0	0.0969	0.0428	Reject
F	400	358	530 ± 60	0.52 ± 0.02	14	0.0305	0.0460	Accept
F	500	329	370 ± 50	0.47 ± 0.02	21	0.0280	0.0480	Accept
F	1,000	209	460 ± 70	0.49 ± 0.03	50	0.0373	0.0599	Accept
F	2,000	113	820 ± 140	0.57 ± 0.04	73	0.0529	0.0808	Accept
M	–	714	1,700 ± 70	0.98 ± 0.03	0	0.0833	0.0328	Reject
M	400	602	600 ± 40	0.57 ± 0.02	16	0.0282	0.0356	Accept
M	500	543	550 ± 40	0.56 ± 0.02	24	0.0301	0.0375	Accept
M	1,000	334	700 ± 70	0.59 ± 0.03	53	0.0279	0.0476	Accept
M	2,000	175	590 ± 80	0.56 ± 0.03	75	0.0453	0.0654	Accept

estimates for truncation level from 10 to 75 %. Note also, that the complete (untruncated) data seem to be Weibull-distributed with $\beta = 1$, i.e. exponentially distributed. However, for complete data the KS goodness-of-fit test is never passed. We assume that the data are zero-inflated in this case which causes a significant bias of the true Weibull parameters.

Table 6 Waiting times between two subsequent earthquakes in CA

x_L [d]	n	$\hat{\alpha}$ [y]	$\hat{\beta}$	$\rho[\%]$	D_n	$D_{cv}^{(ave)}(n, 0.05)$	H_0
1	96	1.76 ± 0.26	0.73 ± 0.06	1	0.0753	0.0875	Accept
3	92	1.94 ± 0.27	0.79 ± 0.06	5	0.0800	0.0893	Accept
7	91	1.93 ± 0.27	0.79 ± 0.06	6	0.0805	0.0898	Accept
30	87	1.88 ± 0.27	0.78 ± 0.07	11	0.0827	0.0917	Accept

[*d*] indicates days and [*y*] represents year

5.3 Earthquake inter-arrival times in California

The inter-arrival times between two subsequent earthquakes in California from 25 March 1806 until 29 March 2014 are investigated. In our notation of ruin theory in Eq. (1) they correspond to W_n. The data was taken from http://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=1&d=1, i.e. from the US government Significant Earthquake database which contains earthquakes "of moderate damage (approximately USD 1 million or more), 10 or more deaths, magnitude 7.5 or greater, modified Mercalli intensity X or greater, or the earthquake generated a tsunami" [25].

When one is interested in the time between *independent* earthquakes, aftershocks (i.e. those earthquakes that follow the largest shock of an earthquake sequence) are usually excluded. Aftershocks can continue over a period of weeks, months, or even years. However, it is not entirely clear whether or not an earthquake has to be classified as aftershock. For this reason many reinsurance treaties contain an "hours clause" limiting the definition of a loss occurrence or event to all losses within a certain number of hours of the original earthquake. Common hours clauses provide for 72 or 168 hours (i.e. 3 or 7 days) [31]. These numbers will correspond to the relevant left-truncation here. As we do not observe in this data base any significant events whose timely difference is less than two days it is likely that these events have been classified as "aftershocks" in the spirit of an hours clause of 3 days. Nevertheless, we investigate left-truncation times of 1, 3, 7 and 30 days. Our analysis shows that for both, complete and left-truncated sets the distribution is Weibull. Also, the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$ are nearly independent of the truncation parameter x_L. The details of this analysis are given in Table 6.

6 Conclusion

In this paper we have shown the applicability of left-truncated Weibull distributions to insurance problems. For that we have proven that the maximum likelihood estimation of the scale α and shape β parameters (or the scale parameter only) from a sample of left-truncated Weibull distributed data set is consistent, asymptotically normal and efficient. While the critical values of the Kolmogorov-Smirnov test statistic differ from those when the Weibull parameters are known, they do not depend on the Weibull parameters themselves: there is a universality to the critical values. When the shape and scale parameter are unknown, our studies show that the critical values have a slight dependence on the truncation values. For practical applications, Table 1 provides some numerical values and an interpolation formula for both the "conservative" and the "anti-conservative" critical values for the cases in which the scale parameter α is unknown or both scale α and shape β parameters are unknown. The significance level varies between 93.5 and 97.5 % if the critical value is assumed to be independent of truncation depending on whether the anti-conservative or conservative values are used.

In the applications Sect. 5 we demonstrate the importance of left-truncating the data for actuarial problems. Both, claim distribution and inter-arrival times can be modelled by left-truncated Weibull distributions in certain circumstances.

Acknowledgments The authors are greatful to Ross Frick for useful comments and discussions. The 4th author gratefully acknowledges financial support from FCT - Fundação para a Ciência e a Tecnologia (Project reference PEst-OE/EGE/UI0491/2013).

Appendix 1: Proofs

Proof of Lemma 1

1. Starting from the Lemma 1 precondition $X_i > x_L > 0$ for all $i = 1, 2, \ldots, n$, we see from Eq. (7) with the new variables $\zeta_i \equiv X_i/x_L > 1$ that the first derivative of $h(\cdot)$ with respect to β is given by

$$h'(\beta) = -\frac{\left[\sum_{i=1}^{n}(\zeta_i^{\beta}-1)\right]^2 + \beta^2\sum_{i=1}^{n}(\zeta_i^{\beta}\log^2\zeta_i)\cdot\sum_{i=1}^{n}(\zeta_i^{\beta}-1) - \beta^2\left(\sum_{i=1}^{n}\zeta_i^{\beta}\log\zeta_i\right)^2}{\beta^2\left[\sum_{i=1}^{n}(\zeta_i^{\beta}-1)\right]^2}. \tag{12}$$

To prove that $h(\cdot)$ is monotonically decreasing, we need to show that $h'(\beta) < 0$ for $\beta > 0$. Thus we need to show that the numerator of Eq. (12) is positive. We prove this statement by mathematical induction over n

$$A(n): \left[\sum_{i=1}^{n}(\zeta_i^{\beta}-1)\right]^2 + \beta^2\sum_{i=1}^{n}(\zeta_i^{\beta}\log^2\zeta_i)\cdot\sum_{i=1}^{n}(\zeta_i^{\beta}-1) - \beta^2\left(\sum_{i=1}^{n}\zeta_i^{\beta}\log\zeta_i\right)^2 > 0 \tag{13}$$

For the base case $A(1)$ we have to show that for any $\zeta_1 > 1$ and $\beta > 0$

$$A(1): (\zeta_1^{\beta} - 1)^2 + \beta^2(\zeta_1^{\beta}\log^2\zeta_1)\cdot(\zeta_1^{\beta} - 1) - \beta^2(\zeta_1^{\beta}\log\zeta_1)^2 > 0 \tag{14}$$

Simplifying Eq. (14) leads to

$$\begin{aligned} e^{\beta \log \zeta_1} - 2 + e^{-\beta \log \zeta_1} &> (\beta \log \zeta_1)^2 \\ 2\,(\cosh[y] - 1) &> y^2 \\ \sum_{k=2}^{\infty} \frac{y^{2k}}{(2k)!} &> 0 \end{aligned} \tag{15}$$

where $y = \beta \log \zeta_1$ and the base case $A(1)$ is verified. Next we do the inductive step $A(n) \rightarrow A(n+1)$. $A(n+1)$ reads as

$$\begin{aligned}
&\left[\sum_{i=1}^{n+1}(\zeta_i^{\beta} - 1)\right]^2 + \beta^2 \left\{ \sum_{i=1}^{n+1} \zeta_i^{\beta} \log^2 \zeta_i \cdot \sum_{i=1}^{n+1} (\zeta_i^{\beta} - 1) - \left(\sum_{i=1}^{n+1} \zeta_i^{\beta} \log \zeta_i \right)^2 \right\} \\
&= \left[\sum_{i=1}^{n} (\zeta_i^{\beta} - 1) + (\zeta_{n+1}^{\beta} - 1) \right]^2 + \\
&\quad + \beta^2 \left\{ \left[\sum_{i=1}^{n} \zeta_i^{\beta} \log^2 \zeta_i + \zeta_{n+1}^{\beta} \log^2 \zeta_{n+1} \right] \cdot \left[\sum_{i=1}^{n} (\zeta_i^{\beta} - 1) + (\zeta_{n+1}^{\beta} - 1) \right] \right. \\
&\quad \left. - \left(\sum_{i=1}^{n} \zeta_i^{\beta} \log \zeta_i + \zeta_{n+1}^{\beta} \log \zeta_{n+1} \right)^2 \right\} \\
&= \left[\sum_{i=1}^{n} (\zeta_i^{\beta} - 1) \right]^2 + 2 \sum_{i=1}^{n} (\zeta_i^{\beta} - 1) \cdot (\zeta_{n+1}^{\beta} - 1) + (\zeta_{n+1}^{\beta} - 1)^2 \\
&\quad + \beta^2 \left\{ \sum_{i=1}^{n} (\zeta_i^{\beta} \log^2 \zeta_i) \cdot \sum_{i=1}^{n} (\zeta_i^{\beta} - 1) + \sum_{i=1}^{n} (\zeta_i^{\beta} \log^2 \zeta_i) \cdot (\zeta_{n+1}^{\beta} - 1) \right. \\
&\quad + (\zeta_{n+1}^{\beta} \log^2 \zeta_{n+1}) \cdot \sum_{i=1}^{n} (\zeta_i^{\beta} - 1) + (\zeta_{n+1}^{\beta} \log^2 \zeta_{n+1}) \cdot (\zeta_{n+1}^{\beta} - 1) \\
&\quad \left. - \left(\sum_{i=1}^{n} \zeta_i^{\beta} \log \zeta_i \right)^2 - 2 \sum_{i=1}^{n} (\zeta_i^{\beta} \log \zeta_i) \cdot (\zeta_{n+1}^{\beta} \log \zeta_{n+1}) - (\zeta_{n+1}^{\beta} \log \zeta_{n+1})^2 \right\} \\
&> 2 \sum_{i=1}^{n} (\zeta_i^{\beta} - 1) \cdot (\zeta_{n+1}^{\beta} - 1) \\
&\quad + \beta^2 \left\{ \sum_{i=1}^{n} (\zeta_i^{\beta} \log^2 \zeta_i) \cdot (\zeta_{n+1}^{\beta} - 1) + (\zeta_{n+1}^{\beta} \log^2 \zeta_{n+1}) \sum_{i=1}^{n} (\zeta_i^{\beta} - 1) \right. \\
&\quad \left. - 2 \sum_{i=1}^{n} (\zeta_i^{\beta} \log \zeta_i) \cdot (\zeta_{n+1}^{\beta} \log \zeta_{n+1}) \right\} \geq 0
\end{aligned} \tag{16}$$

where we have used the Base Case $A(1)$, Eq. (14), and the induction assumption $A(n)$, Eq. (13), to arrive at Eq. (16). We shall prove this inequality Eq. (16)

again by mathematical induction. Rewriting the inequality in terms of new variables $z_i \equiv \zeta_i^\beta > 1$ the induction statment $B(n)$ reads as:

$$
\begin{aligned}
B(n)\colon 2\sum_{i=1}^{n}(z_i-1)\cdot(z_{n+1}-1)+\sum_{i=1}^{n}(z_i\log^2 z_i)\cdot(z_{n+1}-1)\\
+\sum_{i=1}^{n}(z_i-1)\cdot(z_{n+1}\log^2 z_{n+1})-2\sum_{i=1}^{n}(z_i\log z_i)\cdot(z_{n+1}\log z_{n+1})\geq 0\,.
\end{aligned}
\tag{17}
$$

The Base Case $B(1)$ is given by Proposition 1 below. The inductive step $B(n)\to B(n+1)$ is done by some simple algebraic mainipulations. We write $B(n+1)$ as:

$$
\begin{aligned}
&2\sum_{i=1}^{n+1}(z_i-1)\cdot(z_{n+2}-1)+\sum_{i=1}^{n+1}(z_i\log^2 z_i)\cdot(z_{n+2}-1)\\
&\quad+\sum_{i=1}^{n+1}(z_i-1)\cdot(z_{n+2}\log^2 z_{n+2})-2\sum_{i=1}^{n+1}(z_i\log z_i)\cdot(z_{n+2}\log z_{n+2})\\
&=\Bigg\{2\sum_{i=1}^{n}(z_i-1)\cdot(z_{n+2}-1)+\sum_{i=1}^{n}(z_i\log^2 z_i)\cdot(z_{n+2}-1)\\
&\quad+\sum_{i=1}^{n}(z_i-1)\cdot(z_{n+2}\log^2 z_{n+2})-2\sum_{i=1}^{n}(z_i\log z_i)\cdot(z_{n+2}\log z_{n+2})\Bigg\}\\
&\quad+\Bigg\{2(z_{n+1}-1)\cdot(z_{n+2}-1)+(z_{n+1}\log^2 z_{n+1})\cdot(z_{n+2}-1)\\
&\quad+(z_{n+1}-1)\cdot(z_{n+2}\log^2 z_{n+2})-2(z_{n+1}\log z_{n+1})\cdot(z_{n+2}\log z_{n+2})\Bigg\}\\
&\geq 0
\end{aligned}
\tag{18}
$$

Here, all terms in the first curly brackets are non-negative by the induction assumption $B(n)$, Eq. (17), (with some arbitrary number $z_{n+2} > 1$ playing the role of $z_{n+1} > 1$), and likewise the terms in the second curly brackets, constituting the Base Case $B(1)$, Eq. (19), (with two arbitrary numbers $z_{n+1}, z_{n+2} > 1$, playing the roles of $z_1, z_2 > 1$).

2. Note that the order statistic $X_{(n)} = \max\{x_L, X_1, \ldots, X_n\}$. Rewrite Eq. (7) as

$$\lim_{\beta\to+\infty} h(\beta) = \lim_{\beta\to+\infty}\left(\frac{1}{\beta} - \frac{\left(\frac{X_{(n)}}{x_L}\right)^{\beta}\sum_{i=1}^{n}\left(\frac{X_i}{X_{(n)}}\right)^{\beta}\log\frac{X_i}{x_L}}{\left(\frac{X_{(n)}}{x_L}\right)^{\beta}\sum_{i=1}^{n}\left[\left(\frac{X_i}{X_{(n)}}\right)^{\beta} - \left(\frac{x_L}{X_{(n)}}\right)^{\beta}\right]} + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}\right)$$
$$= 0 - \frac{1\cdot\log\frac{X_{(n)}}{x_L}}{1-0} + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}$$
$$= -\log\frac{X_{(n)}}{x_L} + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L} < 0$$

because the geometric mean of n different real numbers is smaller than their largest number.

3. Rewrite Eq. (7) as

$$h(\beta) = \frac{\sum_{i=1}^{n}\left[\left(\frac{X_i}{x_L}\right)^{\beta} - 1\right] - \beta\cdot\sum_{i=1}^{n}\left(\frac{X_i}{x_L}\right)^{\beta}\log\frac{X_i}{x_L}}{\beta\cdot\sum_{i=1}^{n}\left[\left(\frac{X_i}{x_L}\right)^{\beta} - 1\right]} + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}$$

and apply L'Hospital's rule for $\beta \to 0+$ twice to obtain

$$\lim_{\beta\to 0+} h(\beta) = \frac{2\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\log^2\frac{X_i}{x_L}}{\frac{2}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_L}}$$

and the proof is finished.

□

Proposition 1 *For two real numbers* $z_1, z_2 \geq 1$ *we have*

$$F(z_1, z_2) = (z_1 - 1)\cdot\left[z_2\cdot\left(1 + \log^2 z_2\right) - 1\right] + (z_2 - 1)\cdot\left[z_1\cdot\left(1 + \log^2 z_1\right) - 1\right] - 2(z_1\log z_1)(z_2\log z_2) \geq 0\,. \tag{19}$$

Proof We start from the following inequality

$$(\log x - \log y)^2 = \log^2 x - 2\log x\log y + \log^2 y \geq 0$$

By the monotonicity of the Riemann integral we have

$$\int_1^{z1}\int_1^{z2} dxdy\,(\log^2 x - 2\log x\log y + \log^2 y) \geq 0$$
$$\Leftrightarrow 2(z_1 - 1)\cdot(z_2 - 1) + (z_1\log^2 z_1)\cdot(z_2 - 1) + (z_1 - 1)\cdot(z_2\log^2 z_2) - 2(z_1\log z_1)\cdot(z_2\log z_2) \geq 0$$

which is the desired inequality, Eq. (19). □

Proof of Theorem 1: checking the assumptions of [20]

For the proof we apply [20], Theorem 5.1 of section 6.5 (p. 463). Thus, we only need to check conditions (A0)–(A2) of section 6.3 and assumptions (A)–(D) from Sect. 6.5. Let us define as in this reference the general parameter vector $\boldsymbol{\theta} \equiv (\alpha, \beta)$ and the true parameter vector of the distribution as $\boldsymbol{\theta}^0 \equiv (\alpha^0, \beta^0)$. Then the calculations for this are as follows:

Conditions

(A0): requires that the distributions of the observations are distinct, i.e. for different sets of parameters $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$ the corresponding pdf's are different. This is readily checked because for $\boldsymbol{\theta} = (\alpha, \beta)$ and $\boldsymbol{\theta}' = (\alpha', \beta')$ we see that $f(X|\alpha, \beta, x_L) \neq f(X|\alpha', \beta', x_L)$ almost everywhere in X.

(A1): requires that all distributions have a common support, which is true, since $X \in (x_L, \infty)$.

(A2): requires the observations $X_1, \ldots, X_n$ are i.i.d. with a probability density $f(X|\boldsymbol{\theta}^0, x_L)$, which follows from our assumption.

Assumptions

(A): There exists an open subset ω of Ω containing the true parameter point $\boldsymbol{\theta}^0$ such that for almost all X the pdf $f(X|\boldsymbol{\theta}, x_L)$ admits all third derivatives $(\partial^3/\partial\theta_j\partial\theta_k\partial\theta_l)f(X|\boldsymbol{\theta}, x_L)$ for all $\boldsymbol{\theta}$ in ω. This condition is also readily checked because the left-truncated Weibull distribution is contineously differentiable with respect to its parameters $0 < \theta_j < \infty$, with $j = 1, 2$ as mentioned in the first section.

(B): This condition requires the first and second logarithmic derivatives of f satisfy

$$\mathbb{E}_{\boldsymbol{\theta}}\left[\frac{\partial}{\partial\theta_j}\log f(X|\boldsymbol{\theta}, x_L)\right] = 0 \,\text{for}\, j = 1, 2$$

and

$$\begin{aligned} I_{jk}(\boldsymbol{\theta}) &= \mathbb{E}_{\boldsymbol{\theta}}\left[\frac{\partial}{\partial\theta_j}\log f(X|\boldsymbol{\theta}, x_L) \cdot \frac{\partial}{\partial\theta_k}\log f(X|\boldsymbol{\theta}, x_L)\right] \\ &= -\mathbb{E}_{\theta}\left[\frac{\partial^2}{\partial\theta_j\partial\theta_k}\log f(X|\boldsymbol{\theta}, x_L)\right] \text{for}\, j, k = 1, 2 \end{aligned}$$

Here, the expectation operator $\mathbb{E}_{\theta}[\cdot]$ denotes the expectation over the absolute continuous probability measure $f(X|\boldsymbol{\theta}, x_L)dX$. These two conditions are readily verified as the left-truncated Weibull distribution functions are continously differentiable and in $C^{\infty}((x_L,\infty)\times(0,\infty)\times(0,\infty))$. Thus integration and differentiation can be interchanged and integration by parts leads to the desired result since the integral of f over the integration domain (x_L,∞) is 1 by normalisation and thus any of the derivatives vanishes.

(C): All $I_{jk}(\boldsymbol{\theta})$ defined in Assumption (B) are finite and the 2×2 matrix $I(\boldsymbol{\theta})$ is positive definite for all $\boldsymbol{\theta}$ in ω: the relevant integrals can be computed explicitly and are finite, due to the asymptotic condition for $x\to\infty$ we have $f(x|\alpha,\beta,x_L)=O(\exp[-(x/\alpha)^{\beta'}])$ for $\alpha>0$ and any $\beta'\in(0,\beta)$. Hence also the $I_{jk}(\boldsymbol{\theta})$ are finite. Thus the matrix I_{jk} is well-defined and as a covariance matrix by construction positive definite. That the functions ${}^{\partial\ell(X_1,\ldots,X_n|\boldsymbol{\theta},x_L)}/_{\partial\theta_1}$ and ${}^{\partial\ell(X_1,\ldots,X_n|\boldsymbol{\theta},x_L)}/_{\partial\theta_2}$ are affinely linear independent with probability 1 can be seen immediately by explicit computation.

(D): The absolute values of all third derivatives $\left|(\partial^3/\partial\theta_j\partial\theta_k\partial\theta_l)\log f(X|\boldsymbol{\theta},x_L)\right|$ again can be bounded by integrable functions and their expectations $\mathbb{E}_{\theta}[\cdot]$ can be computed and are finite. This is a consequence of the properties of the left-truncated Weibull distributions, namely as $x\to\infty$ we have $f(x|\alpha,\beta,x_L)=O(\exp[-(x/\alpha)^{\beta'}])$ for $\alpha>0$ and any $\beta'\in(0,\beta)$.

Appendix 2: Truncated Weibull random variates and their representation by exponential variates

Let $u_i\in(0,1)$ denote the standard uniform random variable. Then from the cdf in Eq. (2) we obtain a left truncated Weibull distributed random variable X_i

$$X_i=\alpha\cdot\left[\left(\frac{x_L}{\alpha}\right)^{\beta}+\log\frac{1}{u_i}\right]^{1/\beta}=\alpha\cdot\left[\left(\frac{x_L}{\alpha}\right)^{\beta}+y_i\right]^{1/\beta}=\alpha\cdot[\eta+y_i]^{1/\beta} \qquad (20)$$

where y_i is a standard exponential random variate and $\eta\equiv(x_L/\alpha)^{\beta}$.

Appendix 3: Universality of pivotal functions

As in [32] we wish to demonstrate that for a given sample size n and the truncation parameter $\eta>0$ the pivotal functions $(\alpha^0/\hat{\alpha})^{\hat{\beta}}$ and $\hat{\beta}/\beta^0$ are distributed independently with the choice of (α^0,β^0) and have the same distribution as $(1/\hat{\alpha}_{(1,1)})^{\hat{\beta}_{(1,1)}}$ and $\hat{\beta}_{(1,1)}$ respectively for the same n and $\eta>0$, see Eq. (11). Our preference of $(\alpha^0,\beta^0)=(1,1)$ is the simplest choice for the Weibull distribution, i.e. a standard exponential distribution. Demonstrating universality of the KS distance can be achieved by confirming the equality in distribution of the pivotal functions, namely Eq. (11). For this purpose we show that the MLE equations Eqs. (5) and (6) can be written in terms of the pivotal functions, $\hat{\beta}/\beta^0$, $(\hat{\alpha}/\alpha^0)^{\hat{\beta}}$, the truncation parameter

$\eta = (x_L/\alpha^0)^{\beta^0}$ and n. Here $(\hat{\alpha}, \hat{\beta})$ denote the solutions of the MLE equations and (α^0, β^0) are the true values.

Making use of Eq. (20) to re-write the MLE equation for $\hat{\alpha}$, Eq. (5), becomes

$$\left(\frac{\hat{\alpha}}{\alpha^0}\right)^{\hat{\beta}} = \frac{1}{n}\sum_{i=1}^{n}\left[(\eta + y_i)^{\hat{\beta}/\beta^0} - \eta^{\hat{\beta}/\beta^0}\right], \tag{21}$$

where the y_i are standard exponential random variates. Note that Eq. (21) is just the inverse of the first pivotal function $(\alpha^0/\hat{\alpha})^{\hat{\beta}}$. If we can demonstrate that $\hat{\beta}/\beta^0$ is distributed as $\hat{\beta}_{(1,1)}$ (for fixed n and η), then we can conclude that $(\alpha^0/\hat{\alpha})^{\hat{\beta}}$ is distributed as $1/(\hat{\alpha}_{(1,1)})^{\hat{\beta}_{(1,1)}}$ because the right-hand side of Eq. (21) only depends on n, η, the sample of standard exponential random variables $y_1, \ldots, y_n$ and $\hat{\beta}_{(1,1)}$, and hence the first line in Eq. (11) is shown.

Let us now show that $\hat{\beta}/\beta^0$ is distributed as $\hat{\beta}_{(1,1)}$. Using Eq. (20) and performing some algebraic manipulations we can rewrite the MLE equation for $\hat{\beta}$, Eq. (6), in terms of a new random variable $\xi \equiv \hat{\beta}/\beta^0$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\log\,(\eta + y_i)^{\xi} = {} & \frac{1}{n}\sum_{i=1}^{n}\frac{(\eta + y_i)^{\xi}}{\frac{1}{n}\sum_{i=1}^{n}(\eta + y_i)^{\xi} - \eta^{\xi}}\cdot\log\frac{(\eta + y_i)^{\xi}}{\frac{1}{n}\sum_{i=1}^{n}(\eta + y_i)^{\xi} - \eta^{\xi}} \\ & + \log\frac{1}{n}\sum_{i=1}^{n}\left[(\eta + y_i)^{\xi} - \eta^{\xi}\right] - 1 \\ & + \frac{\eta^{\xi}}{\frac{1}{n}\sum_{i=1}^{n}(\eta + y_i)^{\xi} - \eta^{\xi}}\cdot\log\frac{\eta^{\xi}}{\frac{1}{n}\sum_{i=1}^{n}(\eta + y_i)^{\xi} - \eta^{\xi}}\end{aligned} \tag{22}$$

Solving Eq. (22)[6] for the unknown ξ, which is actually the second pivotal function, we see that $\xi = \xi(y_1, \ldots y_n|\eta, n)$ is a uniquely defined random variable depending on the sample of standard exponential random variates $y_1, \ldots, y_n$, sample size n and parameter η only - but not on the original parameters (α^0, β^0). Next we consider the special case of Weibull distributions with $(\alpha^0, \beta^0) = (1, 1)$ and conclude that the unique solution of Eq. (22), $\xi = \hat{\beta}_{(1,1)} = \xi(y_1, \ldots y_n|\eta, n)$ depends only on sample size n, parameter η and the same sample of standard exponential random variables $y_1, \ldots, y_n$. Thus our claim and the second line in Eq. (11) is shown.

[6] Note that $\xi = \hat{\beta}/\beta^0$ is the unique solution to the MLE equation Eq. (22) which is equivalent to the original MLE Eq. (6). For the existence of the unique solution Lemma 1 requires the inequality Eq. (8) to be satisfied. In our notation using Eq. (20) this means $2\cdot\left(\frac{1}{n}\sum_{i=1}^{n}\log\,(1 + \eta^{-1}y_i)\right)^2 > \frac{1}{n}\sum_{i=1}^{n}\log^2\,(1 + \eta^{-1}y_i)$, which depends only on sample size n, parameter η and some sample of standard exponential random variables $y_1, \ldots, y_n$.

Appendix 4: Computing the fisher information matrix

The covariance matrix in Theorem 1, $Z(\alpha,\beta)^{-1}$ is derived from the logarithm of the pdf $f(\cdot)$ from Eq. (3). One readily obtains the second derivatives necessary for the calculation of the Fisher information matrix $Z(\alpha,\beta)$, writing as usual $\eta=(x_L/\alpha)^{\beta}$,

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha^2}\log f(X|\alpha,\beta,x_L) &= -\frac{\beta}{\alpha^2}\left\{1+(\beta+1)\left[\eta-\left(\frac{X}{\alpha}\right)^{\beta}\right]\right\}\\
\frac{\partial^2}{\partial\alpha\partial\beta}\log f(X|\alpha,\beta,x_L) &= -\frac{1}{\alpha}-\frac{1}{\alpha}\left[\eta-\left(\frac{X}{\alpha}\right)^{\beta}\right]\\
&\quad -\frac{1}{\alpha}\left[\eta\log\eta-\left(\frac{X}{\alpha}\right)^{\beta}\log\left(\frac{X}{\alpha}\right)^{\beta}\right]\\
\frac{\partial^2}{\partial\beta^2}\log f(X|\alpha,\beta,x_L) &= -\frac{1}{\beta^2}+\frac{1}{\beta^2}\left\{\eta[\log\eta]^2-\left(\frac{X}{\alpha}\right)^{\beta}\left[\log\left(\frac{X}{\alpha}\right)^{\beta}\right]^2\right\}
\end{aligned}$$

To evaluate all relevant expectations $\mathbb{E}(\cdot)$ in the definition of the Fisher information matrix, the following integrals are needed

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{X}{\alpha}\right)^{\beta}\right] &= \int_{x_L}^{\infty} dx\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{\left(\frac{x_L}{\alpha}\right)^{\beta}-\left(\frac{x}{\alpha}\right)^{\beta}}\cdot\left(\frac{x}{\alpha}\right)^{\beta}\\
&= e^{\eta}\int_{\eta}^{\infty} dze^{-z}\cdot z\\
&= 1+\eta\\
\mathbb{E}\left[\left(\frac{X}{\alpha}\right)^{\beta}\log\left(\frac{X}{\alpha}\right)^{\beta}\right] &= 1+\eta\log\eta+[\log\eta+e^{\eta}E_1(\eta)]\\
\mathbb{E}\left[\left(\frac{X}{\alpha}\right)^{\beta}\left\{\log\left(\frac{X}{\alpha}\right)^{\beta}\right\}^2\right] &= \eta(\log\eta)^2+2[\log\eta+e^{\eta}E_1(\eta)]\\
&\quad+\left[(\log\eta)^2+2e^{\eta}E_2(\eta)\right]
\end{aligned}$$

where we have used the functions

$$\begin{aligned}
E_1(s) &= \int_s^{\infty} dy\frac{e^{-y}}{y}\\
E_2(s) &= \int_s^{\infty} dye^{-y}\frac{\log y}{y}
\end{aligned}$$

In the limit $\eta\longrightarrow 0+$ we recover the covariance matrix for the untruncated system as given by [27] (Eq. 11.17)

$$Z(\alpha,\beta)^{-1} = \begin{pmatrix} 1.1087 \cdot \dfrac{\alpha^2}{\beta^2} & 0.2570 \cdot \alpha \\ 0.2570 \cdot \alpha & 0.6079 \cdot \beta^2 \end{pmatrix}$$

Note that in our representation of $Z(\alpha,\beta)^{-1}$ the factor $1/n$ is displayed separately. For an estimate of the error in the MLE estimation of $\hat{\alpha}$ and $\hat{\beta}$, one may apply this expression with the hat-ed parameters rather than the unknown "true" parameters in conjunction with Theorem 1, statements 2 and 3.

References

1. Afonso LB, Cardoso RM, Egídio dos Reis AD (2013) Dividend problems in the dual risk model. Insur Math Econ 53(3):906–918
2. Aho M, Bain LJ, Engelhardt M (1985) Goodness-of-fit tests for the Weibull distribution with unknown-parameters and heavy censoring. J Stat Comput Simul 21(3–4):213–225
3. Balakrishnan N, Kateri M (2008) On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. Stat Probab Lett 78(17):2971–2975
4. Bar-Lev SK (1984) Large sample properties of the mle and mcle for the natural parameter of a truncated exponential family. Annal Inst Stat Math 36(2):217–222
5. Barndorff-Nielsen O (1978) Information and exponential families in statistical theory. Wiley, New York
6. Bergel AI, Egídio dos Reis AD (2013) On a Sparre-Andersen risk model with PH(n) interclaim times. Tech. rep., CEMAPRE Working paper, http://cemapre.iseg.ulisboa.pt/archive/preprints/617
7. Buchanan R, Priest C (2006) Deductible, encyclopedia of actuarial science, 1st edn. Wiley, New York
8. Chandra M, Singpurwalla ND, Stephens MA (1981) Kolmogorov statistics for tests of fit for the extreme-value and Weibull-distributions. J Am Stat Assoc 76(375):729–731
9. David FN, Johnson NL (1948) The probability integral transformation when parameters are estimated from the sample. Biometrika 35(1–2):182–190
10. Deemer WL, Votaw DF (1955) Estimation of parameters of truncated or censored exponential distributions. Annal Math Stat 26(3):498–504
11. Dickson DCM (2005) Insurance risk and ruin. Cambridge University Press, New York
12. Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling extremal events for insurance and finance, vol 33. Springer, Berlin
13. Farnum NR, Booth P (1997) Uniqueness of maximum likelihood estimators of the 2-parameter Weibull distribution. IEEE Trans Reliab 46(4):523–525
14. Freedman D, Diaconis P (1981) On the histogram as a density estimator—L_2 theory. Z Wahrscheinlichkeit 57(4):453–476
15. Frees EW (2009) Regression modelling with actuarial and financial applications (Wisconsin School of Business). http://instruction.bus.wisc.edu/jfrees/jfreesbooks/Regression
16. Ji L, Zhang C (2012) Analysis of the multiple roots of the Lundberg fundamental equation in the ph(n) risk model. Appl Stoch Model Bus Ind 28(1):73–90
17. Kendall MG, Stuwart A (1979) The advanced theory of statistics, vol 2. C. Griffin, London
18. Klugman SA, Panjer HH, Willmot GE (2012) Loss models: from data to decisions, 4th edn. Wiley, New York
19. Klüppelberg C, Villasenor JA (1993) Estimation of distribution tails—a semiparametric approach. Blätter der DGVFM 21(2):213–235
20. Lehmann EL, Casella G (1998) Theory of point estimation. Springer, New York
21. Lilliefors H (1969) On the Kolmogorov-Simirnov test for the exponential distribution with mean unknown. J Am Stat Assoc 64(325):387–389
22. Littell RC, McClave JT, Offen WW (1979) Goodness-of-fit tests for the 2 parameter Weibull distribution. Commun Stat Part B Simul Comput 8(3):257–269

23. Miller LH (1956) Table of percentage points of Kolmogorov statistics. J Am Stat Assoc 51(273):111–121
24. Mosteller F (1946) On some useful inefficient statistics. Annal Math Stat 17(4):377–408
25. National Geophysical Data Center (2015) The significant earthquake database. (n.d.). http://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=1&d=1
26. Parsons FG, Wirsching PH (1982) A Kolmogorov Simirnov goodness-of-fit test for the 2-parameter Weibull distribution when the parameters are estimated from the data. Microelect Reliab 22(2):163–167
27. Rinne H (2009) The Weibull distribution: a handbook. CRC Press, Boca Raton
28. Rodríguez-Martínez EV, Cardoso RMR, Egídio dos Reis AD (2015) Some advances on the Erlang(n) dual risk model. ASTIN Bull 45(1):127–150
29. Shorack GR, Wellner JA (2009) Empirical processes with applications to statistics. SIAM, Philadelphia
30. Smith RL (1985) Maximum-likelihood estimation in a class of nonregular cases. Biometrika 72(1):67–90
31. Stroock L (2011) Stroock special bulletin: reinsurance implications of the Japanese earthquake. http://www.stroock.com/SiteFiles/Pub1091
32. Thoman DR, Bain LJ, Antle CE (1969) Inferences on the parameters of the Weibull distribution. Technometrics 11(3):445–460
33. Weibull W (1951) A statistical distribution function of wide applicability. J Appl Mech Trans Asme 18(3):293–297
34. Wingo DR (1989) The left-truncated Weibull distribution: theory and computation. Statist Papers 30:39–48
35. Woodruff BW, Moore AH, Dunne EJ, Cortes R (1983) A modified Kolmogorov-Simirnov test for Weibull-distributions with unknown location and scale-parameters. IEEE Trans Reliab 32(2):209–213