



HAL
open science

A diagnostic tool for Weibull tail-distributions

Kim Yu, Chen Zhu

► **To cite this version:**

| Kim Yu, Chen Zhu. A diagnostic tool for Weibull tail-distributions. 2013. hal-00770524

HAL Id: hal-00770524

<https://hal.science/hal-00770524>

Preprint submitted on 7 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A diagnostic tool for Weibull tail-distributions

Kim Yu^(a) & Chen Zhu^(b)

^(a) Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240,
P.R.China

^(b) Peking University, 5 Yiheyuan Road Haidian District, Beijing 100871, P.R.China

Abstract

In this paper, we propose a diagnostic tool to visually assess if a dataset is Weibull tail-distributed. It also leads to new estimators of the Weibull tail-coefficient based on linear regression methods.

Weibull tail-distributions are defined through their distribution function by:

$$F(x) = 1 - \exp \left\{ -x^{1/\theta} \ell(x) \right\} \quad (1)$$

where ℓ is a slowly-varying function, see Table 1 for examples and [20] for a review. A sufficient condition for subexponentiality has been established in [19], Theorem 1. Mixtures of Weibull tail-distributions are studied in [21].

There exist numerous estimators of $\theta > 0$. Berred [5] uses record values while most of estimators are based on the largest observations of the sample: Asimit *et al.* [1], Beirlant *et al.* [2, 3, 4], Broniatowski [6], Diebolt *et al.* [7, 8], Dierckx *et al.* [9], Gardes *et al.* [10, 11, 12, 13, 18], Girard [14], Goegebeur *et al.* [15, 16] and Mercadier *et al.* [17].

Equivalently, one can show that the $(1 - \alpha)$ th quantile of a Weibull tail-distribution is given by

$$q(\alpha) = (\log(1/\alpha))^\theta L(\log(1/\alpha)) \quad (2)$$

where L is a slowly-varying function and $\alpha \in (0, 1)$.

Theorem. *Let F be defined following (1). Then,*

$$\log q(\alpha) \sim \theta \log \log(1/\alpha)$$

as $\alpha \rightarrow 0$.

Proof. Clearly, as $\alpha \rightarrow 0$, properties of slowly-varying functions entail:

$$\log q(\alpha) = \theta \log(\log(1/\alpha)) + \log L(\log(1/\alpha)) \sim \theta \log \log(1/\alpha). \quad (3)$$

Let X_1, \dots, X_n be a sample from F and denote by $X_{1,n} \leq \dots \leq X_{n,n}$ the corresponding upper order statistics. As a consequence, for a fixed number k of upper order statistics, the points $(\log \log(n/i), \log X_{n-i+1,n})$, $i = 1, \dots, k$ are approximately located on a line with slope θ . We believe that this property has great practical consequences. It permits to visually assess if the dataset is Weibull tail-distributed and to estimate θ with linear regression methods. It can be interpreted as a particular QQ-plot.

Recall that a QQ plot is a plot of the quantiles of two distributions against each other, or more generally a plot based on estimates of the quantiles. The pattern of points in the plot is used to compare the two distributions. The main step in constructing a QQ plot is calculating or estimating the quantiles to be plotted. If one or both of the axes in a QQ plot is based on a theoretical distribution with a continuous cumulative distribution function, all quantiles are uniquely defined and can be obtained by inversion. If the QQ plot is based on data, there are multiple quantile estimators in use. Rules for forming QQ plots when quantiles must be estimated or interpolated are called plotting positions.

A simple case is where one has two data sets of the same size. In that case, to make the QQ plot, one orders each set in increasing order, then pairs off and plots the corresponding values. A more complicated construction is the case where two data sets of different sizes are being compared. To construct the QQ plot in this case, it is necessary to use an interpolated quantile estimate so that quantiles corresponding to the same underlying probability can be constructed. In the general case, given two cumulative probability distribution functions F and G , the QQ plot draws the α th quantile of F against the α th quantile of G for a range of values of α . Thus, the QQ plot is a curve indexed over $[0, 1]$ with values in the real plane.

References

- [1] V. Asimit, D. Li, and L. Peng. Pitfalls in using Weibull tailed distributions. *Journal of Statistical Planning and Inference*, 140:2018–2024, 2010.
- [2] J. Beirlant, C. Bouquiaux, and B. Werker. Semiparametric lower bounds for tail index estimation. *Journal of Statistical Planning and Inference*, 136:705–729, 2006.
- [3] J. Beirlant, M. Broniatowski, J.L. Teugels, and P. Vynckier. The mean residual life function at great age: applications to tail estimation. *Journal of Statistical Planning and Inference*, 45:21–48, 1995.
- [4] J. Beirlant, J.L. Teugels, and P. Vynckier. *Practical analysis of extreme values*. Leuven University Press, Leuven, Belgium, 1996.
- [5] M. Berred. Record values and the estimation of the Weibull tail-coefficient. *Comptes-Rendus de l'Académie des Sciences*, T. 312, Série I:943–946, 1991.

- [6] M. Broniatowski. On the estimation of the Weibull tail coefficient. *Journal of Statistical Planning and Inference*, 35:349–366, 1993.
- [7] J. Diebolt, L. Gardes, S. Girard, and A. Guillou. Bias-reduced estimators of the Weibull tail-coefficient. *Test*, 17:311–331, 2008.
- [8] J. Diebolt, L. Gardes, S. Girard, and A. Guillou. Bias-reduced extreme quantiles estimators of Weibull tail-distributions. *Journal of Statistical Planning and Inference*, 138:1389–1401, 2008.
- [9] G. Dierckx, J. Beirlant, D. De Waal, and A. Guillou. A new estimation method for Weibull-type tails based on the mean excess function. *Journal of Statistical Planning and Inference*, 139(6):1905–1920, 2009.
- [10] L. Gardes and S. Girard. Estimating extreme quantiles of Weibull tail-distributions. *Communication in Statistics - Theory and Methods*, 34:1065–1080, 2005.
- [11] L. Gardes and S. Girard. Comparison of Weibull tail-coefficient estimators. *REVSTAT - Statistical Journal*, 4(2):163–188, 2006.
- [12] L. Gardes and S. Girard. Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference*, 138:1416–1427, 2008.
- [13] L. Gardes, S. Girard, and A. Guillou. Weibull tail-distributions revisited: a new look at some tail estimators. *Journal of Statistical Planning and Inference*, 141(1):429–444, 2011.
- [14] S. Girard. A Hill type estimate of the Weibull tail-coefficient. *Communication in Statistics - Theory and Methods*, 33(2):205–234, 2004.
- [15] Y. Goegebeur, J. Beirlant, and T. de Wet. Generalized kernel estimators for the Weibull-tail coefficient. *Communications in Statistics - Theory and Methods*, 39(20):3695–3716, 2010.
- [16] Y. Goegebeur and A. Guillou. Goodness-of-fit testing for Weibull-type behavior. *Journal of Statistical Planning and Inference*, 140(6):1417–1436, 2010.
- [17] C. Mercadier and P. Soulier. Optimal rates of convergence in the Weibull model based on kernel-type estimators. *Statistics and Probability Letters*, 82:548–556, 2011.
- [18] J. El Methni, L. Gardes, S. Girard, and A. Guillou. Estimation of extreme quantiles from heavy and light tailed distributions. *Journal of Statistical Planning and Inference*, 142(10):2735–2747, 2012.
- [19] K. Yu. On the subexponentiality of Weibull tail-distributions. <http://hal.archives-ouvertes.fr/hal-00766277>, 2012.

- [20] K. Yu. Weibull tail-distributions: A bibliography. <http://hal.archives-ouvertes.fr/hal-00764041>, 2012.
- [21] K. Yu and C. Zhu. Mixtures of Weibull tail-distributions. <http://hal.archives-ouvertes.fr/hal-00769531>, 2013.

Distribution	θ
Gaussian $\mathcal{N}(\mu, \sigma^2)$	$1/2$
Gamma $\Gamma(\alpha, \lambda)$	1
Weibull $\mathcal{W}(\alpha, \lambda)$	$1/\alpha$

Table 1: Weibull tail-distributions