Zenodo – DOI 10.5281/zenodo.6352677

Weibull, κ-Weibull and 3-parameter extended Weibull distributions

Amelia Carolina Sparavigna

Department of Applied Science and Technology, Politecnico di Torino

Here we consider a function of κ-statistics, the κ-Weibull distribution. It is compared to the Weibull distribution. We also consider the 3-parameter extended Weibull according to Marshall–Olkin extended distributions. This and the κ-Weibull functions are compared.

Torino, March 14, 2022

Keywords: Weibull distribution, κ-Weibull distribution, κ-statistics.

Weibull distribution

The Weibull distribution is a continuous probability distribution. Its probability density function is given by:

$$f(x|\lambda,k)=\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^k} \quad \text{, per} \quad x\geq 0 \qquad (1)$$

If $x<0$, $f(x|\lambda,k)=0$.

In the distribution, $k>0$ is the *shape parameter* and $\lambda>0$ is the *scale parameter*. The Weibull distribution is related to a number of other probability distributions.

If $k=1$, we have the exponential distribution.

The Rayleigh distribution is given by $k=2$. $\lambda=\sqrt{2}\sigma$.

$$f(x|\sigma)=\frac{x}{\sigma^2}e^{-x^2/(2\sigma^2)} \qquad (2)$$

If the quantity x is a "time-to-failure", the Weibull distribution gives a distribution for which the failure rate is proportional to a power of time.

In the formalism of [1], the Weibull probability density function (pdf) is defined as:

$$f(t|B,C,D)=\frac{B}{C}\left(\frac{t-D}{C}\right)^{(B-1)}e^{-\left(\frac{t-D}{C}\right)^{B}} \qquad (3)$$

where $B>0,\ C>0,\ -\infty<D<\infty,\ t>D$.

Symbol t is representing the random variable. The distribution is suitable to analyse time-series, where t is the elapsed time.

Parameter *D* is the threshold, which is therefore representing the minimum value of time.

B is the shape parameter, which controls the overall shape of the probability density function. Its value usually ranges between 0.5 and 8.0 [1]. The Weibull distribution includes other useful distributions [1]. If $B=1$, we have the exponential distribution. For $B=2$, we have the Rayleigh distribution. For $B=2.5$ and $B=3.6$, the Weibull distribution approximates the lognormal distribution and the normal distribution respectively.

The scale parameter *C* changes the scale of the probability density function along the time axis (that is from days to months or from hours to days). It does not change the actual shape of the distribution [1]. Parameter *C* is known as the characteristic life. In [1], it is stressed that "No matter what the shape, 63.2% of the population fails by $t = C+D$ ". It is also told that "Some authors use $1/C$ instead of *C* as the scale parameter ".

Let us put $\alpha=B$, $\gamma=1/C$, $\tau=D$. Eq. (3) becomes:

$$f(t|\alpha,\gamma,\tau)=\alpha\,\gamma\,(\gamma\cdot(t-\tau))^{\alpha-1}e^{-(\gamma\cdot(t-\tau))^{\alpha}} \qquad (4)$$

Then, using $\beta=\gamma^{\alpha}$:

$$f(t|\alpha,\gamma,\tau)=\alpha\,\beta\cdot(t-\tau)^{\alpha-1}e^{-\beta(t-\tau)^{\alpha}} \qquad (5)$$

Then, in the same formalism of Eq.(1):

$$f(x|k,b)=b\,k\,x^{k-1}e^{-b\,x^{k}} \qquad (6)$$

In (6), $x=t-\tau$, $b=\beta$, $k=\alpha$. (6) is the form of the Weibull pdf used for applications in medical statistics and econometrics [2],[3].

κ-Weibull distribution

Let us consider the analogue of Weibull pdf in the κ statistics [4],[5].

The κ-Weibull probability distribution function (pdf) is described by:

$$f_\kappa(x|\alpha,\beta)=\frac{\alpha\,\beta\,x^{\alpha-1}}{\sqrt{1+\kappa^2\beta^2x^{2\alpha}}}\ \exp_\kappa(-\beta\,x^\alpha) \qquad (7)$$

where the κ-exponential is defined in the following manner:

$$\exp_\kappa(u)=\left(\sqrt{1+\kappa^2u^2}+\kappa\,u\right)^{1/\kappa} \qquad (8)$$

Parameters α,β are related to the shape and scale indexes of Weibull distribution, whereas κ is the index of κ-distribution, that is the statistical distribution introduced by G. Kaniadakis, Politecnico di Torino, in [4],[5]. Recently, the use of the distribution has been proposed in epidemiology [6],[7].

In [8], we can find discussed and defined the κ-Weibull. In the formalism of the given reference:

$$f_\kappa=\frac{m}{x_s}\left(\frac{x}{x_s}\right)^{m-1}\frac{\exp_\kappa\left(-[x/x_s]^m\right)}{\sqrt{1+\kappa^2(x/x_2)^{2m}}} \qquad (9)$$

In Eq.(9), x is the random variable. In the formalism of [1], with time and threshold, Eq.(9) becomes:

$$f_\kappa(t|B,C,D)=\frac{B}{C}\left(\frac{t-D}{C}\right)^{B-1}\frac{\exp_\kappa\{-[(t-D)/C]^B\}}{\sqrt{1+\kappa^2((t-D)/C)^{2B}}} \qquad (10)$$

Let us put $\alpha=B$, $\gamma=1/C$. $\tau=D$, (10) becomes:

$$f_\kappa(t|\alpha,\gamma,\tau)=\alpha\,\gamma\,\gamma^{\alpha-1}(t-\tau)^{\alpha-1}\frac{\exp_\kappa\{-\gamma^\alpha(t-\tau)^\alpha\}}{\sqrt{1+\kappa^2\gamma^{2\alpha}(t-\tau)^{2\alpha}}} \qquad (11)$$

Then, using $\beta=\gamma^{\alpha}$:

$$f_k(t|\alpha,\beta,\tau)=\frac{\alpha\beta(t-\tau)^{\alpha-1}}{\sqrt{1+\kappa^2\beta^2(t-\tau)^{2\alpha}}}\ \exp_\kappa(-\beta(t-\tau)^{\alpha}) \qquad (12)$$

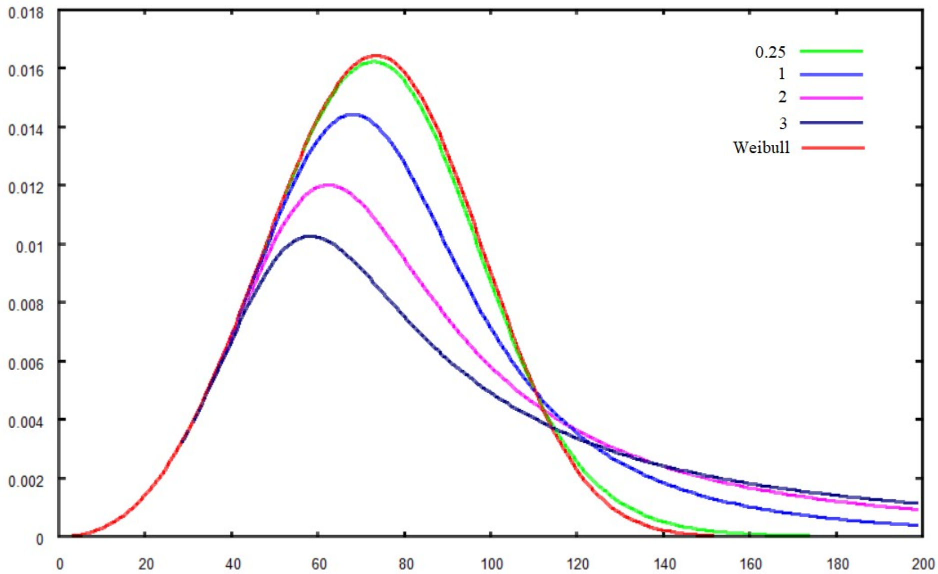


Figure 1 (a) – Comparing Weibull and κ-Weibull. The Weibull pdf is given in red. Parameters are $\alpha=3.5$, $\beta=2.0\times10^{-7}$, and $\tau=0$.

The κ-Weibull curves have different κ values: 0.25, 1, 2 and 3.

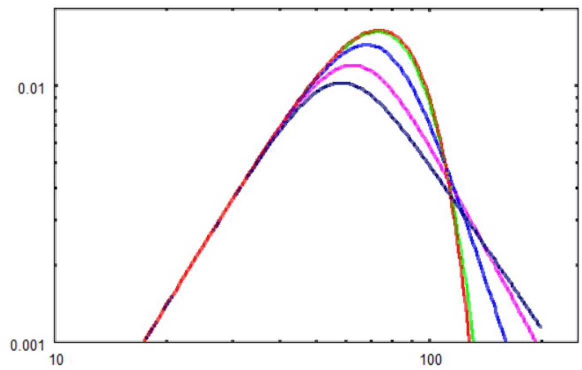


Figure 1 (b) – Comparing Weibull and κ-Weibull in a log-log graph.

Figure 1 shows the comparison of Weibull pdf with that of κ-Weibull. We can see that the value of κ parameter is strongly affecting the tail of the distribution. Increasing the value the tail becomes a “long” tail, that is, a portion of the distribution having many occurrences far from the head of the distribution.

Mixture density

In the case that the distribution is showing two peaks, a mixture of Weibull or κ-Weibull can be considered, in the form:

$$f = f_1 + f_2 = \xi f_{\kappa_1}(t|\alpha_1, \beta_1, \tau_1) + (1-\xi) f_{\kappa_2}(t|\alpha_2, \beta_2, \tau_2) \qquad (13)$$

Parameter ξ, the mixing parameter, is ranging from zero to 1. It is used to generalize the addition of peaks, as proposed for the Weibull distribution [9]. It is also a rough manner to consider the fact that the set of population, involved by pandemic, changed for sure during the considered time period (we will further discuss this point).

In the case that we have three peaks, then (13) becomes:

$$f = f_1 + f_2 + f_3 = \xi_1 f_{\kappa_1}(t|\alpha_1, \beta_1, \tau_1) + \xi_2 f_{\kappa_2}(t|\alpha_2, \beta_2, \tau_2) + \xi_3 f_{\kappa_3}(t|\alpha_3, \beta_3, T\tau_3) \qquad (14)$$

In (14), we must have $\xi_1 + \xi_2 + \xi_3 = 1$.

Being a finite sum, the mixture is known as a finite mixture, and the density is the "mixture density". Usually, “mixture densities” can be used to model a statistical population with subpopulations. Each component is related to a subpopulations, and its weight is proportional to the given subpopulation in the overall population.

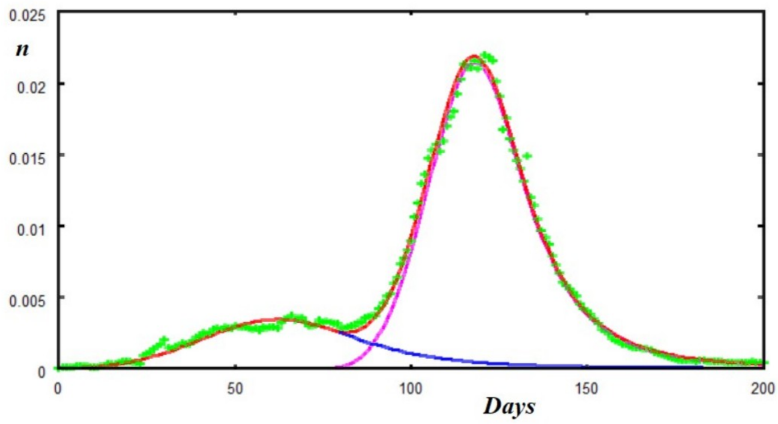


Figure 2 – An example of mixture density (see Ref. [7]).

3-parameter extended Weibull distribution

In [10], we can find an approach, based on Marshall–Olkin extended distributions [11], to the Weibull distribution. In Ref. [10], the 2-parameter Weibull appears as:

$$f(x|\beta,\lambda)=\beta\,\lambda^{\beta}\,x^{\beta-1}e^{-(\lambda x)^{\beta}} \quad , \quad x\geq 0 \qquad (15)$$

The 3-parameters extended distribution is given as:

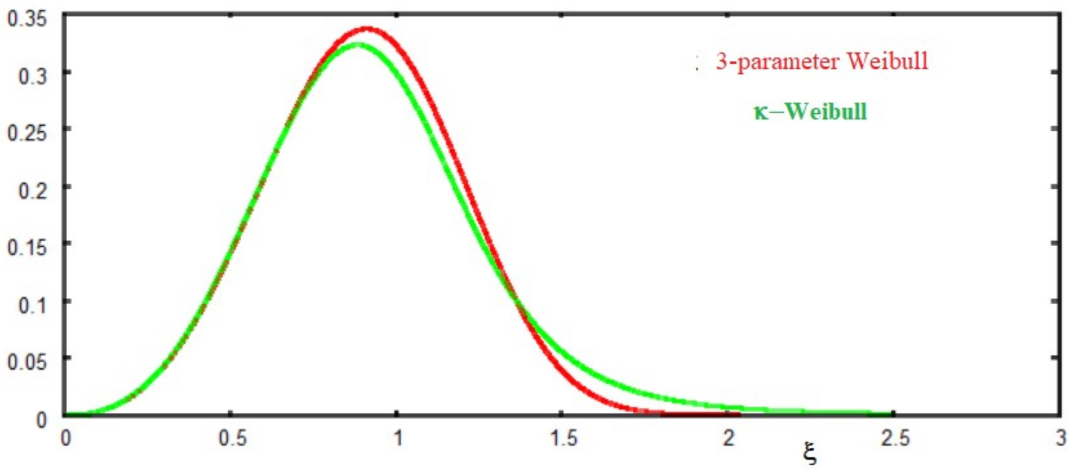
$$f(x|\alpha,\beta,\lambda)=\frac{\alpha\,\beta\,\lambda(\lambda x)^{\beta-1}\,e^{-(\lambda x)^{\beta}}}{\left[1-\tilde{\alpha}\,e^{-(\lambda x)^{\beta}}\right]^{2}} \qquad (16)$$

In (16), $x>0$, $\alpha,\beta,\lambda>0$, $\tilde{\alpha}=1-\alpha$.

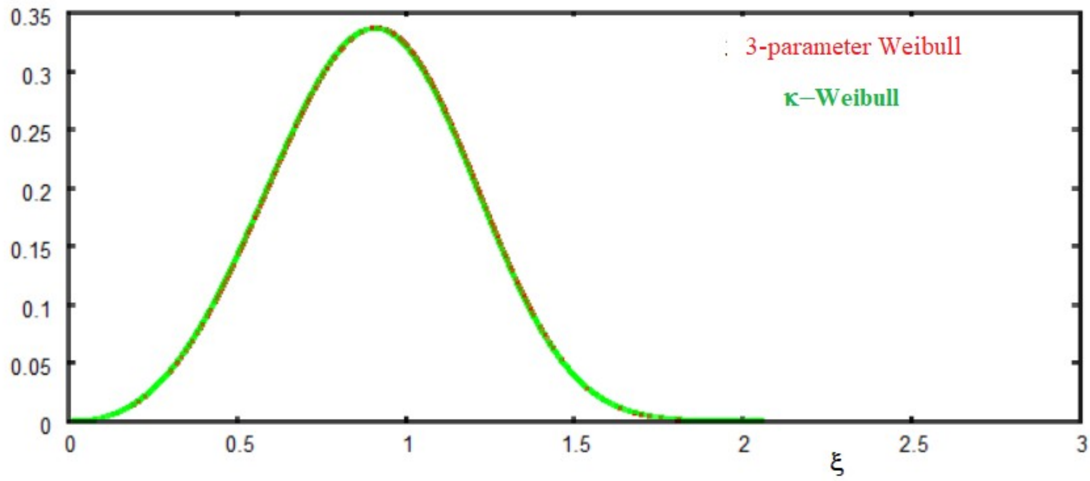
Let us compare to κ-Weibull. Here we rewrite Eq.(7) in the same formalism as (16):

$$f_{\kappa}(x|\beta,\lambda)=\frac{\beta\,\lambda(\lambda x)^{\beta-1}}{\sqrt{1+\kappa^{2}\lambda^{2\beta}x^{2\beta}}}\ \exp_{\kappa}(-\lambda^{\beta}x^{\beta}) \qquad (17).$$

Let us compare (16) and (17). In the following figure, $\xi=\lambda x$.



Parameters used for the calculation: $\beta=3.5$, $\lambda=0.25$, $\kappa=0.5$, $\alpha=1.01$.



In the case we change parameter κ in $\kappa=0.05$, with the same other parameters ($\beta=3.5$, $\lambda=0.25$, $\alpha=1.01$), the curves are indistinguishable.
We can note again the role of parameter κ in determining the tail of the function.

Weibull distribution for pseudorandom numbers

Routine RNWIB can be used to generate pseudorandom numbers, starting from a Weibull distribution, with shape parameter A and unit scale parameter, so that:

$$f(x)=A\,x^{A-1}e^{-x^{A}} \quad , \quad x\geq 0 \qquad (18)$$

https://help.imsl.com/fortran/6.0/stat/default.htm?turl=rnwib.htm

Reliability Function (Weibull)

In the formalism of [1], we have seen before that the Weibull pdf is:

$$f(t|B,C,D)=\frac{B}{C}\left(\frac{t-D}{C}\right)^{(B-1)}e^{-\left(\frac{t-D}{C}\right)^{B}}$$

where $B>0,\ C>0,\ -\infty<D<\infty,\ t>D$.

The reliability (or survivorship) function, $R(t)$, is giving the probability of surviving beyond the time t . For the Weibull pdf, we have:

$$R(t)=e^{-\left(\frac{t-D}{C}\right)^B} \qquad (19)$$

The reliability function is one minus the cumulative distribution function. That is:

$$R(t)=1-F(t) \qquad (20)$$

dove $F(t|B,C,D)=\int_{-\infty}^{t} f(t'|B,C,D)\,dt'$.

Hazard Function (Weibull)

The hazard function represents the instantaneous failure rate. The rate is given by the function:

$$h(t)=\frac{f(t)}{R(t)}=\frac{B}{C}\left(\frac{t-D}{C}\right)^{B-1} \qquad (21)$$

A plot model for Weibull

The cumulative distribution function is:

$$F(t)=1-e^{-\left(\frac{t-D}{C}\right)^B}$$

Let us assume $D=0$:

$$\ln(1-F(t))=-(t/C)^B \quad \text{then:} \quad \ln(-\ln(1-F(t)))=-B\ln C+B\ln(t)$$

Let us introduce: $y=\ln(-\ln(1-F(t)))$, $x=\ln t$, we have:

$$y=-B\ln C+Bx$$

In this manner, the plot is that of a straight line.

References

[1] NCSS Statistical Software NCSS.com, Chapter 550. Distribution (Weibull) Fitting. https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Distribution-Weibull-Fitting.pdf

[2] Collett, D. (2015). Modelling survival data in medical research (3rd ed.). Boca Raton: Chapman and Hall / CRC. ISBN 978-1439856789.

[3] Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. p. 584. ISBN 978-0-521-84805-3.

[4] Kaniadakis, G. (2002). Statistical mechanics in the context of special relativity. Physical review E, 66(5), 056125.

[5] Kaniadakis, G. (2001). Non-linear kinetics underlying generalized statistics. Physica A: Statistical mechanics and its applications, 296(3-4), 405-425.

[6] Kaniadakis, G., Baldi, M. M., Deisboeck, T. S., Grisolia, G., Hristopulos, D. T., Scarfone, A. M., Sparavigna, A. C., Wada, T. and Lucia, U., 2020. The κ-statistics approach to epidemiology. Scientific Reports, 10(1), pp.1-14. https://www.nature.com/articles/s41598-020-76673-3

[7] Sparavigna, A. C. (2021). A decomposition of waves in time series of data related to Covid-19, applied to study the role of Alpha variant in the spread of infection. Zenodo. https://doi.org/10.5281/zenodo.5745008

[8] Hristopulos, D. T., Petrakis, M. P., and Kaniadakis, G. (2015). Weakest-link scaling and extreme events in finite-sized systems. Entropy, 17(3), 1103-1122.

[9] Razali, A. M., & Salih, A. A. (2009). Combining two Weibull distributions using a mixing parameter. European Journal of Scientific Research, 31(2), 296-305.

[10] Caroni, C. (2010). Testing for the Marshall–Olkin extended form of the Weibull distribution. Statistical Papers, 51(2), 325-336.

[11] Marshall, A. W., and Olkin, I. (1997) A new method of adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika 84, 641–652.