Weibull, κ-Weibull and Other Probability Distributions

57 Pages Posted: 23 May 2022

Amelia Carolina Sparavigna

Polytechnic University of Turin

Date Written: April 13, 2022

Abstract

Here we will consider a function of κ-statistics, the κ-Weibull distribution, and compare it to the well-known Weibull distribution. The κ-Weibull will be also compared to the 3-parameter extended Weibull function, obtained according to the Marshall–Olkin extended distributions. The log-logistic distribution will be considered for comparison too, such as the exponentiated Weibull, the Burr and the q-Weibull distributions. The most important observation, coming from the proposed calculations, is that the κ-Weibull hazard function is strongly depending on the values of parameter κ, a parameter which is deeply influencing the behaviour of the tail of the probability distribution. As a consequence, the κ-Weibull function turns out to be quite relevant for generalizations of the Weibull approach to modeling failure times. Discussions about the Maximum Likelihood approach for Weibull, κ-Weibull and Burr distributions will be also given.

Keywords: Weibull Distribution, κ-Weibull Distribution, Marshall-Olkin Distribution, Log-Logistic Distribution, Exponentiated Weibull Distribution, Burr Distribution, q-Weibull Distribution, κ-Statistics, Modeling Failure Times, Maximum Likelihood

May 3, 2022

Weibull, κ-Weibull and other probability distributions

Amelia Carolina Sparavigna

Department of Applied Science and Technology, Politecnico di Torino

Here we will consider a function of κ-statistics, the κ-Weibull distribution, and compare it to the well-known Weibull distribution. The κ-Weibull will be also compared to the 3-parameter extended Weibull function, obtained according to the Marshall–Olkin extended distributions. The log-logistic distribution will be considered for comparison too, such as the exponentiated Weibull, the Burr and the q-Weibull distributions. The most important observation, coming from the proposed calculations, is that the κ-Weibull hazard function is strongly depending on the values of parameter κ, a parameter which is deeply influencing the behaviour of the tail of the probability distribution. As a consequence, the κ-Weibull function turns out to be quite relevant for generalizations of the Weibull approach to modeling failure times. Discussions about the Maximum Likelihood approach for Weibull, κ-Weibull and Burr distributions will be also given.

DOI 10.2139/ssrn.4076871

Keywords: Weibull distribution, κ-Weibull distribution, Marshall-Olkin distribution, log-logistic distribution, logistic distribution, exponentiated Weibull distribution, Burr distribution, *q*-Weibull distribution, Gumbel and Gompertz distributions, Gamma distribution, κ-statistics, modeling failure times.

Introduction

In Volume 28 of the "Methods in Experimental Physics", it is noted that, although there is no limit to the number of possible univariate probability distributions that can be theoretically proposed, only a limited number of them is occurring repeatedly in scientific work. Seven of these popular distributions are discussed by Laurent Hodges in the Chapter on univariate distributions [Hodges, 1994]: three are the discrete distributions (binomial, Pascal or negative binomial, Poisson distributions) and four are the continuous distributions (Gaussian or normal, log-normal, exponential, and Weibull distributions).

In introducing the Weibull distribution [Weibull, 1951], in [Hodges, 1994] it is observed that, for experiments characterized by continuous random variables, a best fit of the related probability distributions can be easily obtained by means of functions involving two or more parameters, because of a consequent greater freedom in fitting the experimental results. Actually, the Weibull distribution falls into this category of functions, being it a two- parameter generalization of the exponential distribution [Hodges, 1994). The two parameters of Weibull distribution are known as "scale" and "shape" parameters.

Besides being used to describe several empirical distributions [Weibull, 1951], the Weibull distribution is widely involved in modeling the failure times. This use has its origin in the great variety of shapes of probability curves that can be generated by different choices of scale and shape parameters [Hodges, 1994]. "Weibull distributions range from exponential distributions to curves resembling the normal distribution. The exponential distribution limit corresponds to a "memoryless" failure rate (the failure rate of an individual item is independent of its current age), while the other Weibull distributions correspond to distributions of failure times that are peaked at certain ages and skewed in different fashions" [Hodges, 1994].

Here, we will discuss the Weibull distribution. Then, we will consider its generalization in the form of the κ-Weibull distribution. This function has been obtained in the framework of the κ-calculus [Kaniadakis, 2013], a calculus which has its roots in special relativity and is used for statistical analyses involving power law tailed statistical distributions. The κ-Weibull distribution is featuring the κ-deformed exponential function, which is a continuous one parameter deformation of the Euler exponential function. Then, the Weibull distribution, the shape of which can be widely changed, can be further adjusted in its tail, by means of the κ-exponential. The consequence is that we have the possibility of a different investigation of failure times.

Besides a comparison of Weibull and κ-Weibull function, we consider also the 3-parameter extended Weibull according to Marshall–Olkin extended distributions. The log-logistic distribution will be considered for comparison too, such as the exponentiated Weibull and Burr distributions. As we will see in the following discussion, the most important observation that we can obtain from the proposed comparisons is that the κ-Weibull hazard function is strongly depending on the values of parameter κ, a parameter which is deeply influencing the behaviour of the tail of the probability distribution. For this reason, the κ-Weibull function turns out to be interesting for generalizations of the Weibull approach to modeling failure times.

1. Weibull distribution

The Weibull distribution is a continuous probability distribution. Its probability density function (pdf) 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 $x<0$, $f(x|\lambda,k)=0$.

The ratio x/λ is dimensionless. In the case that x indicates a position (dimension: length, L) or a time interval (dimension: time, T), λ has dimension L or T respectively. k is a dimensionless parameter. Function $f(x|\lambda,k)$ has the dimension of x, then integral $\int f(x|\lambda,k)dx$ is dimensionless.

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: $f(x|\lambda,1)=\frac{1}{\lambda}e^{-(x/\lambda)}$.

Let us note that, from (1), 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.

The cumulative and reliability functions are given by [Hodges, 1994]:

$$F(x|\lambda,k)=1-e^{-(x/\lambda)^k} \quad ; \quad R(x|\lambda,k)=e^{-(x/\lambda)^k}$$

The failure rate (hazard function) is given by [Hodges, 1994]:

$$f/R = h(x|\lambda,k)=\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \quad .$$

Mean, median and variance are (Γ is the Gamma function) [Hodges, 1994]:

$$\lambda\Gamma\left(1+\frac{1}{k}\right) \quad ; \quad \lambda(\ln 2)^{x/k} \quad ; \quad \lambda^2\Gamma\left(1+\frac{2}{k}\right)-\left[\lambda\Gamma\left(1+\frac{1}{k}\right)\right]^2$$

In the formalism of [NCSS], 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 $B>0,\ C>0,\ -\infty<D<\infty,\ t>D$.

When $D=0$, Eq. (3) becomes:

$$f(t|B,C,0)=\frac{B}{C}\left(\frac{t}{C}\right)^{(B-1)}e^{-\left(\frac{t}{C}\right)^{B}} \quad .$$

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. If $D=0$, it means that the threshold time is zero.

To compare with (1), let us pose $B=k$, $C=\lambda$, $t-D=x$, and we find again:

$$f(x|\lambda,k)=\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^{k}} \quad .$$

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 [NCSS]. The Weibull distribution includes other useful distributions [NCSS]. 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 log-normal 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 [NCSS]. Parameter *C* is known as the characteristic life. In [NCSS], 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:

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

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

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

Another formalism gives the Weibull distribution in the following form:

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

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 [Collett, 2015],[Cameron et Trivedi, 2005].

2. κ-Weibull distribution

Let us consider the analogue of Weibull pdf in the κ-statistics [Kaniadakis, 2002], [Kaniadakis, 2001]. The κ-Weibull probability distribution function (pdf) is described by [Kaniadakis et al., 2020] in the form:

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

where the κ-exponential is defined in the following manner (see Appendix for further discussion):

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

Parameters α,β are related to the shape and scale parameters of Weibull distribution, whereas κ is the index of κ-distribution, that is the statistical distribution introduced by Giorgio Kaniadakis, Politecnico di Torino, in [Kaniadakis, 2002],[Kaniadakis, 2001]. Recently, the use of the distribution has been proposed in epidemiology [Kaniadakis et al. 2020],[Sparavigna, 2021]. Let us note that κ is a dimensionless parameter. β has dimensions $[x^{-\alpha}]$.

Eq. 7 is the derivative of the lifetime distribution function [Kaniadakis et al., 2020]:

$$L_\kappa = 1 - \exp_\kappa(-\beta x^\alpha) \quad .$$

In [Hristopulos et al., 2015], we can find discussed and defined the κ-Weibull.

In the formalism of the given reference:

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

In [Hristopulos et al., 2015], authors had investigated the systems that obey the weakest-link scaling (WLS) principle. In this case, the system response, the material strength for instance, is controlled by the weakest link. “The Weibull model is the archetypical probability distribution for WLS systems and is widely used in reliability modeling”. In the Ref. [Hristopulos et al., 2015], a list of applications and related references are given. Models of fracture strength of brittle- and quasi-brittle materials, geologic media, waiting times between earthquakes, wind speed, annual hydrological maxima are mentioned.

In Eq.(9), x is the random variable. In the formalism of [NCSS], 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 $\beta = \gamma^\alpha$:

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

which is the expression of κ-Weibull used in [Kaniadakis et al., 2020].

When $\kappa \to 0$, we find the Weibull distribution as we can see in the following figure.

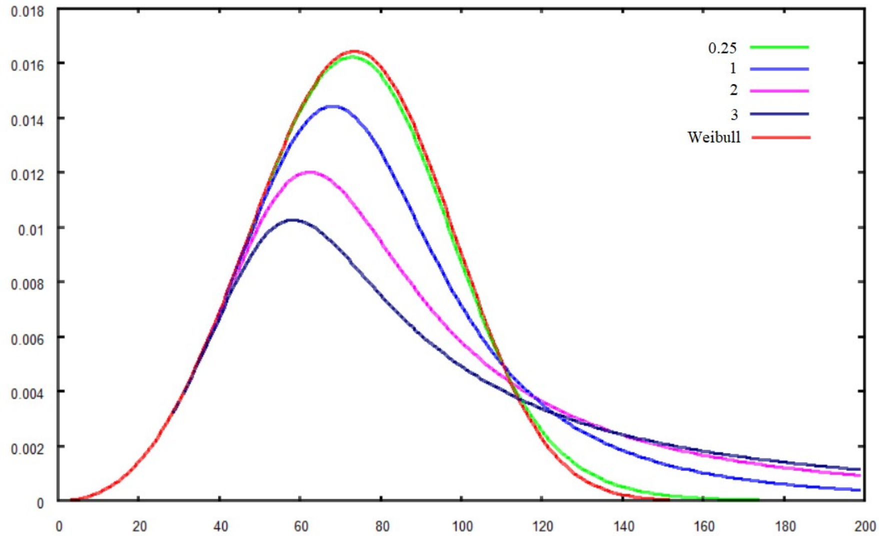


Figure 1 (a) – Comparing Weibull and κ-Weibull. The Weibull pdf is given in red. Parameters are $\alpha=3.5$, $\beta=2.0\times10^{-7}$, and $\tau=0$. The κ-Weibull curves have different κ values: 0.25, 1, 2 and 3. See also Ref. [Sparavigna, 2021].

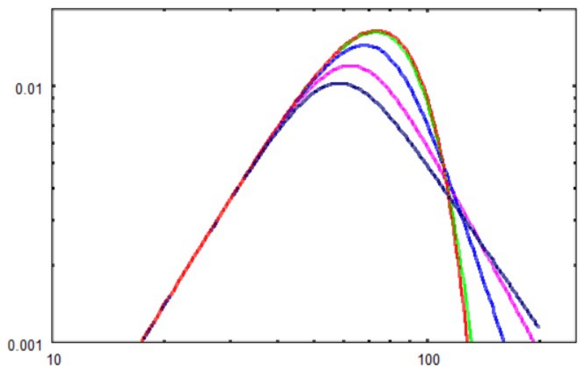


Figure 1 (b) – Comparing Weibull and κ-Weibull in a log-log graph.

Figure 1 shows the comparison of Weibull pdf with those of κ-Weibull for different κ values. 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. Note also the behaviour of the left tail, which is almost the same from $t=0$ to a value t_κ, determined by the parameter κ.

3. 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 [Razali et Salih, 2009]. It is also a rough manner to consider the fact that the set of population, involved by pandemic (see discussion in [Sparavigna, 2021]), changed for sure during the considered time period.

In the case that we have three peaks, (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.

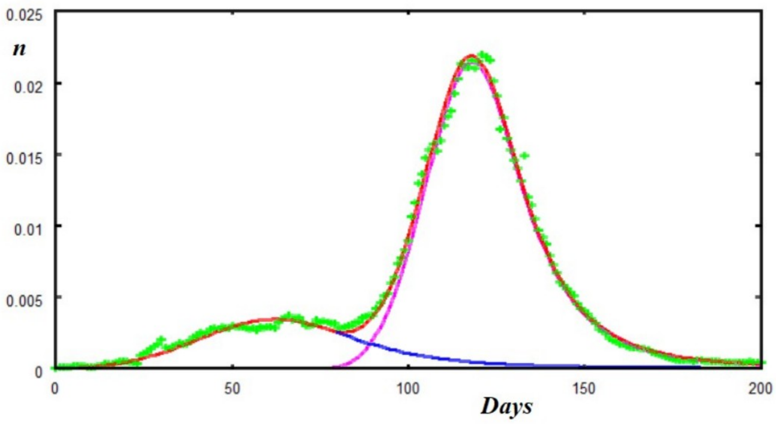


Figure 2 – An example of mixture density (see Ref. [Sparavigna, 2021]). The larger peak is related to the spread of the α-variant of Sars-Cov-2 infections. Data are concerning the population in London. The onset of the larger peak is corresponding to November 25, 2020.

4. A 3-parameter extended Weibull distribution

The Weibull distribution is based on two parameters, which are the shape and the scale. We can consider the threshold as a third parameter. For instance, at the web page of www.real-statistics.com, the author Charles Zaiontz. Eq. (3) is defined as describing a three-parameter Weibull distribution. However, let us add a third parameter in the different manner discussed in the following.

In [Caroni, 2010], we can find an approach, based on Marshall–Olkin extended distributions [Marshall et Olkin, 1997], to the Weibull distribution. In Ref. [Caroni, 2010], the 2-parameter Weibull appears as:

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

The 3-parameters extended distribution is given as:

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

In (16a), $x>0$, $\alpha,\beta,\lambda>0$, $\tilde{\alpha}=1-\alpha$. In the formalism of Eq. (1):

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

(16a) becomes:

$$f(x|\alpha,\lambda,k)=\frac{\alpha k(x/\lambda)^{k-1}\ e^{-(x/\lambda)^k}}{\lambda\left[1-\tilde{\alpha}\ e^{-(\lambda x)^k}\right]^2} \qquad \text{3-parameter Weibull} \qquad (16b)$$

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

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

It is clear that, from the point of view of dimensional analysis, it would be better to use the formalism (16b). However, from now on, and just for a convenience in writing the formulae, let us assume λ having dimensions $[x^{-1}]$.

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

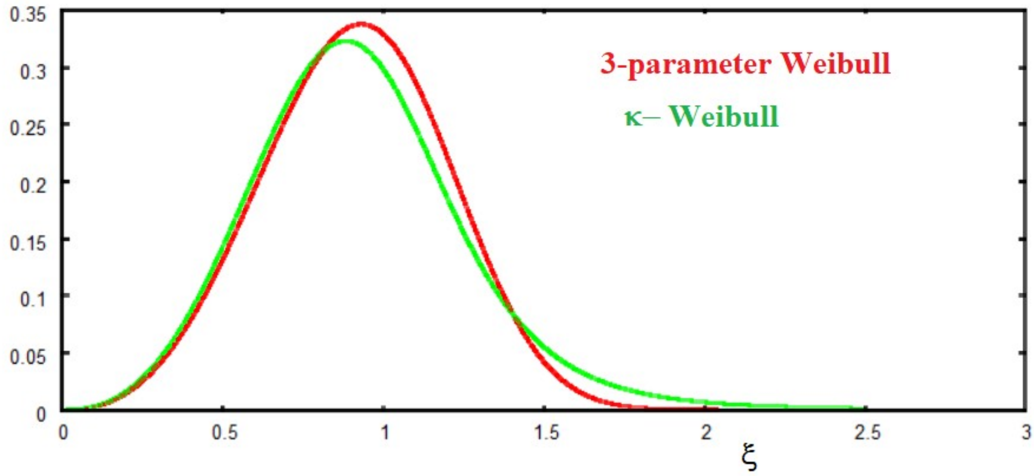


Figure 3: Comparing functions (16a) and (17). Parameters used for the calculation: $\beta=3.5$, $\lambda=0.25$, $\kappa=0.5$, $\alpha=1.1$.

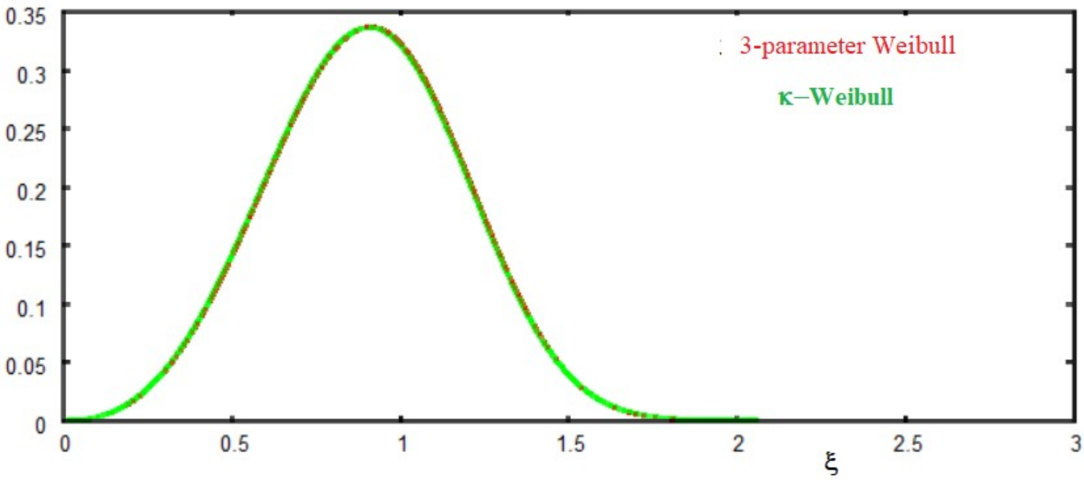


Fig. 4: In the case we change parameter κ in $\kappa=0.05$, with the same other parameters ($\beta=3.5$, $\lambda=0.25$, $\alpha=1.01$), the curves are indistinguishable.

We can note again the role of parameter κ in determining the tail of the function.

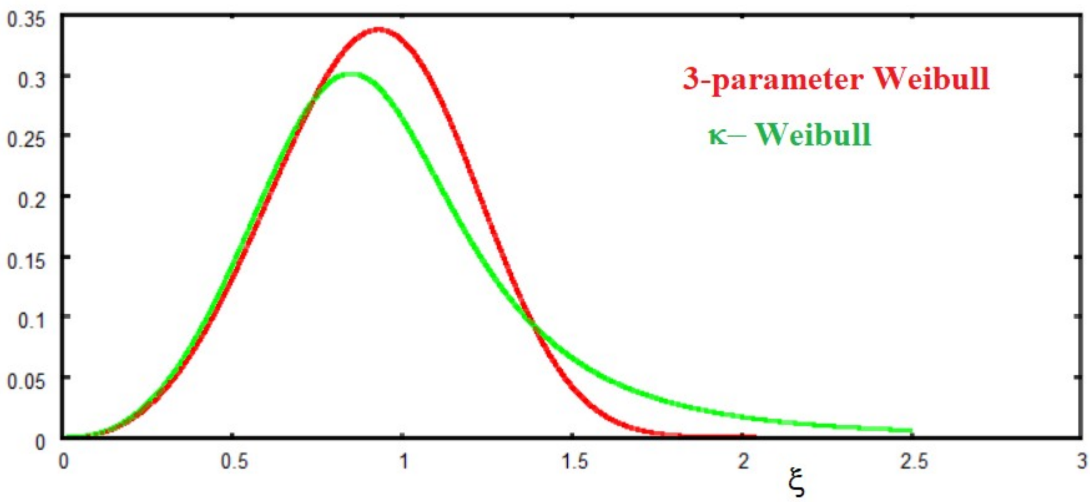


Fig. 5 : In the plot given above, parameters are $\beta=3.5$, $\lambda=0.25$, $\kappa=0.9$, $\alpha=1.1$.

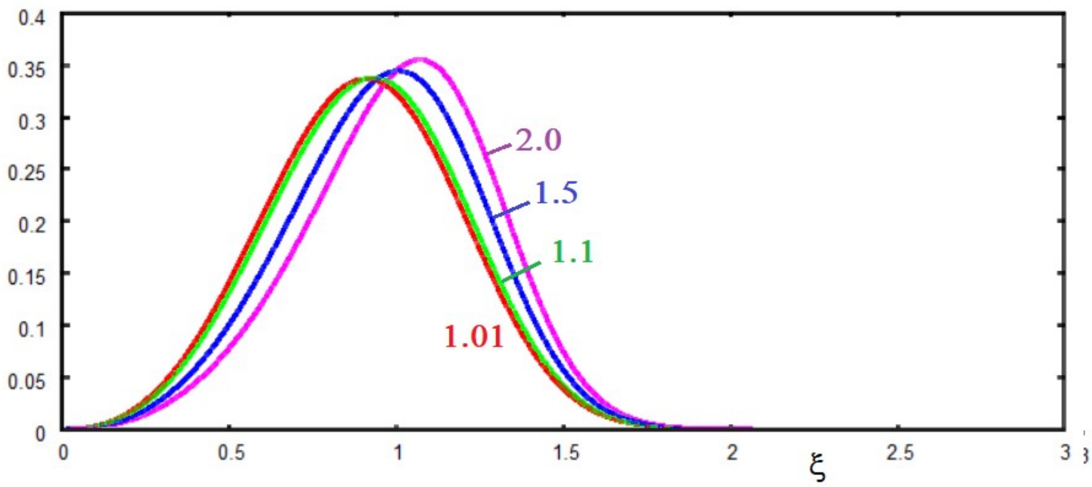


Fig. 6 : In the plot given above, the 3-parameter Weibull is given for parameters $\beta=3.5$, $\lambda=0.25$. Values of α . are 1.01 (red), 1.1 (green), 1.5 (blue) and 2.0 (violet).

Form the Figure 6, we can see that, in the 3-parameter Weibull, parameter α is changing the left part of the curve.

5. Cumulative κ-Weibull

In the following Figure, and in the formalism of (17), the cumulative function of κ-

Weibull, for three different values of parameter κ.

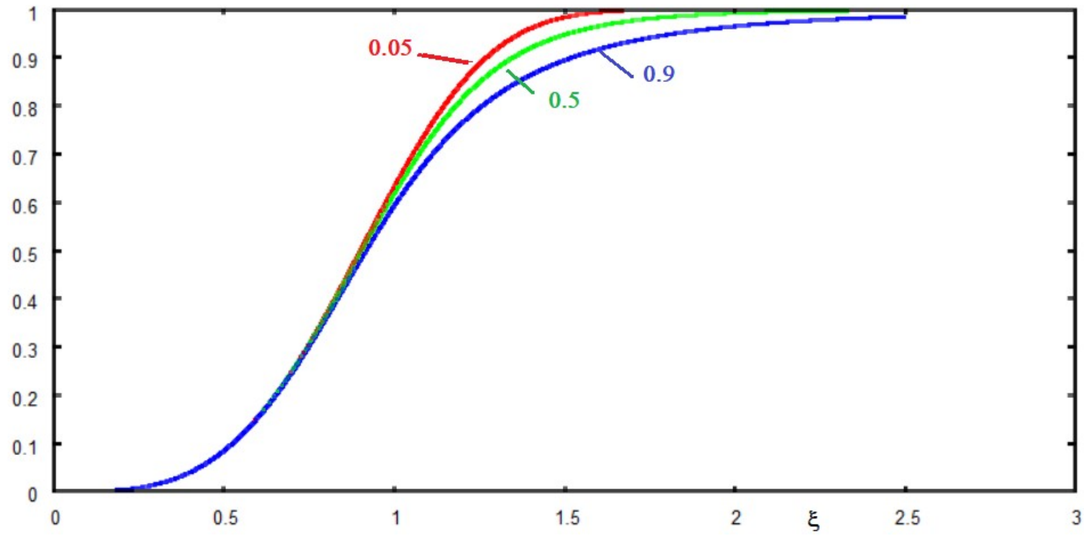


Fig. 7 : Cumulative function of κ-Weibull for parameters $\beta=3.5$, $\lambda=0.25$. Numbers in the image are referring to the values of parameter κ.

In the Figure 7, we have shown the cumulative function related to the κ-Weibull distribution. This function, let us call it $F(x)$, is defined as:

$$f(x)=\frac{dF(x)}{dx}$$

where $f(x)$ is the pdf. Every probability distribution supported on the real numbers, discrete or "mixed" as well as continuous, is uniquely identified by an upwards continuous monotonic increasing cumulative distribution function [Çakallı, 2015].

The cumulative function is linked to the reliability function.

6. Reliability Function (Weibull)

In the formalism of [NCSS], 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 that the reliability is:

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

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

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

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

7. Reliability Function (κ-Weibull)

Defining the function as in (20) and using the formalism of Eq. (17), we have that the reliability of κ-Weibull is depending on κ parameter as in the following figure.

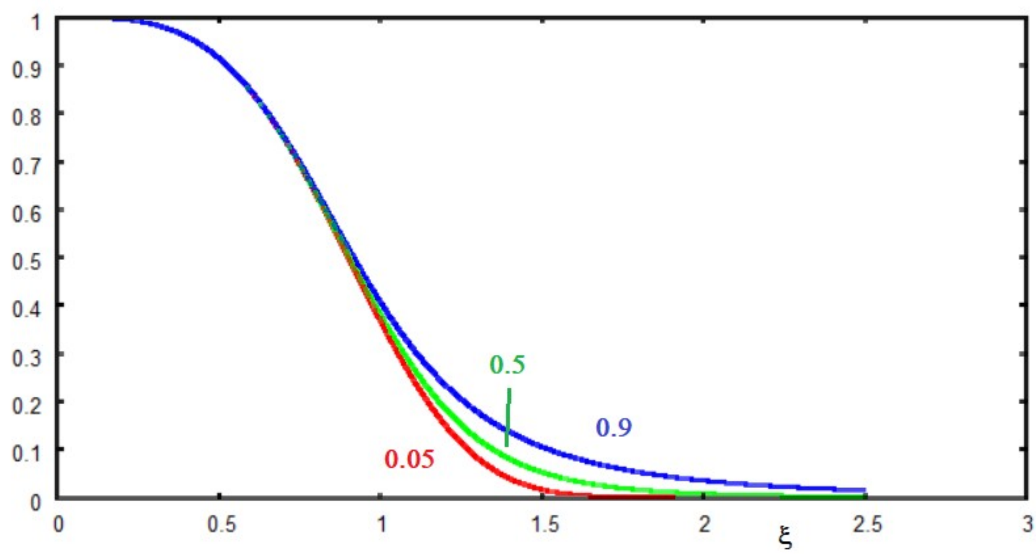


Fig. 8: Reliability function of κ-Weibull for parameters $\beta=3.5$, $\lambda=0.25$. Numbers in the image are referring to the values of parameter κ.

8. Hazards in reliability analysis

The hazard function is a conditional failure rate. In fact, it is conditional because an organism or device had actually survived until time t . In this manner, the function at year 10 only applies to organisms or devices (items) who were actually alive in year 10. It doesn't count those who died or failed in previous periods.

The hazard function are frequently associated with products and applications.

Usually, the hazard functions are featured as increasing, constant and decreasing function. An increasing hazard function indicates that items are more likely to fail with time. For instance, mechanical items subjected to stress or fatigue have an increased risk of failure over the lifetime of the product. A decreasing hazard function indicates failures that are more likely to occur early in the life of an item. An example is errors in a computer program; they are more likely near the release of a new software program, decreasing as time passes with improved releases. A constant hazard function indicates failures that are equally likely to occur at any time in the item's life.

Products exist having failure rates that follow a "bathtub" curve. The name of the curve is derived from the cross-sectional shape of a bathtub, which has steep sides and a flat bottom. These items have hazard rate which is high initially and low in the centre. Then the hazard is high again at the end of item's life. For this reason, a bathtub curve is widely used in reliability engineering and in the models of deterioration (Wikipedia)

In particular, a bathtub hazard function which comprises three parts:

1) The first part is a decreasing failure rate, known as early failures.

2) The second part is a constant failure rate, known as random failures.

3) The third part is an increasing failure rate, known as wear-out failures.

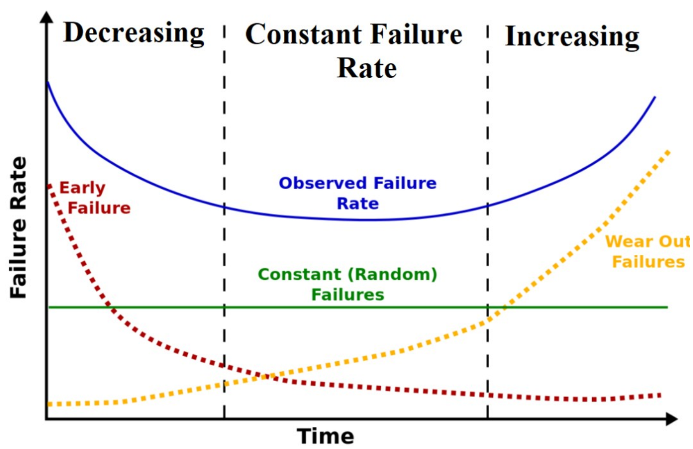


Fig. 9- Bathtub failure rate, Image Courtesy: Wikipedia, McSush

Many electronic consumer product life cycles strongly exhibit the bathtub curve [Lienig et Bruemmer, 2017]. In reliability engineering, the cumulative distribution function corresponding to a bathtub curve may be analysed using a Weibull chart [Lienig et Bruemmer, 2017].

9. Hazard Function (Weibull)

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

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

In the following figure, it is shown the behaviour of function $(\lambda x)^{(\beta-1)}$, for three different values of β . As discussed in [Kızılersü et al., 2018], it is helpful to visualise the differences between values of β in the hazard function by using a "bathtub" diagram, such as that given in the Figure 2 of [Kızılersü et al., 2018]. Hazard function for $\beta<1$ we have a "likely to fall at the start". When $\beta\sim 1$. "failure rate is fairly constant". When $\beta>1$, the "failure rate increases as the time goes by".

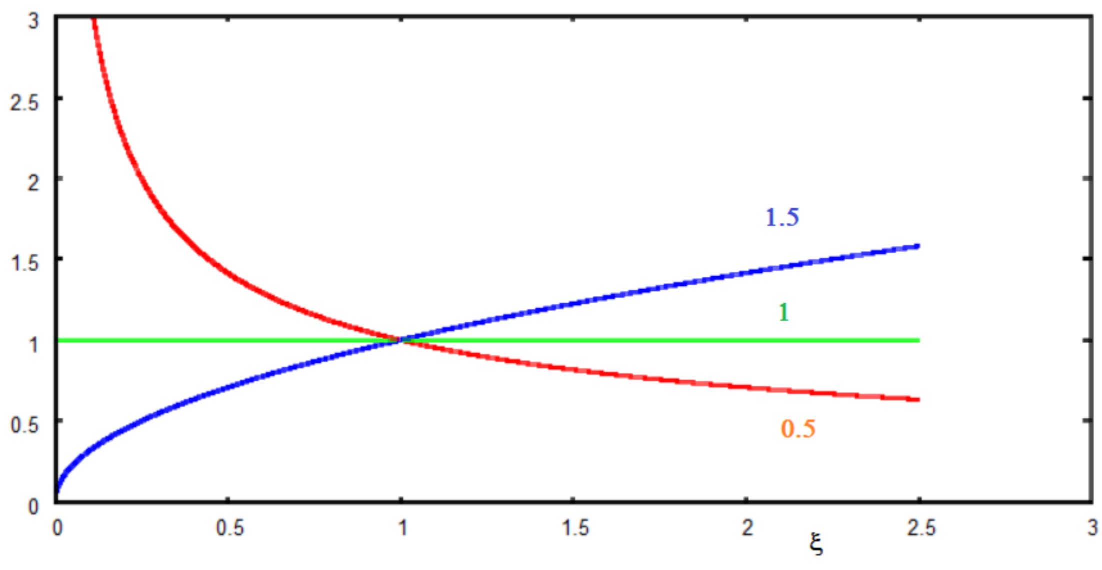


Fig. 10: Function $(\lambda x)^{(\beta-1)}$ for parameter $\lambda=0.25$. Numbers in the image are referring to the values of parameter β .

10. Hazard Function (κ-Weibull)

The hazard function of the κ-Weibull, is given by: $h_\kappa(x|\lambda,\beta)=\frac{f_\kappa(x|\lambda,\beta)}{R_\kappa}$. In the following figure, the hazard function is given for three different values of parameter κ. The formalism is the same of the Figures 3 to 7.

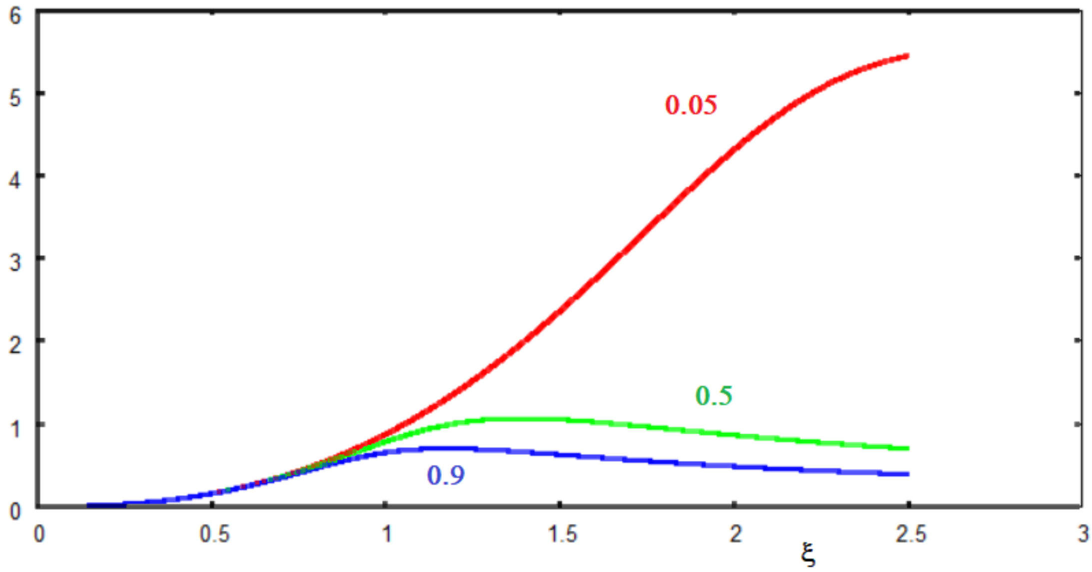


Fig. 11: Hazard function of κ-Weibull for parameters $\beta = 3.5$, $\lambda = 0.25$. Numbers in the image are referring to the values of parameter κ.

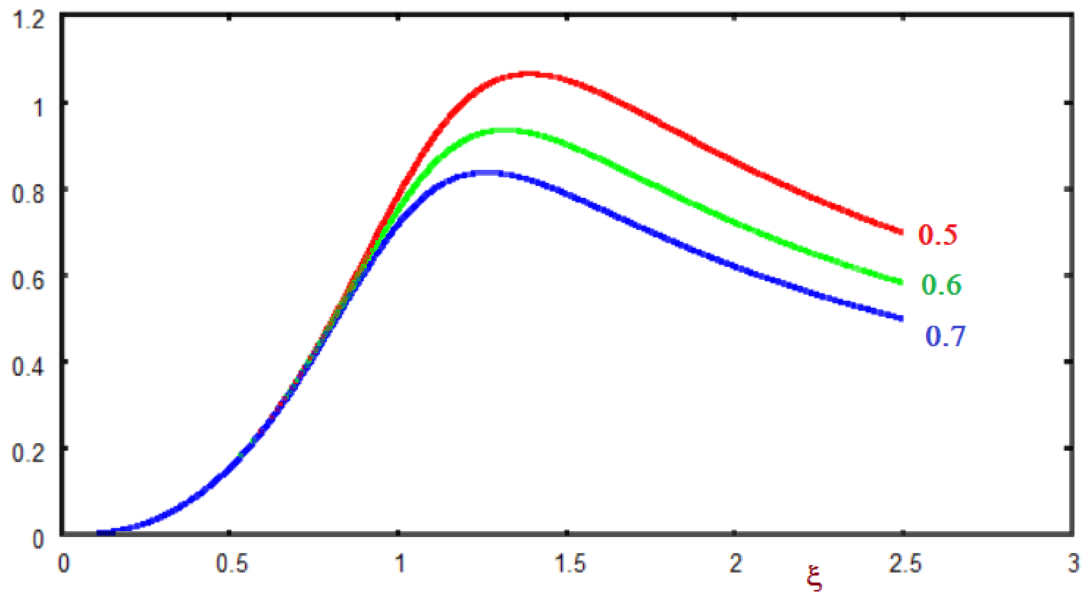


Fig. 12: Hazard function of κ-Weibull for parameters $\beta = 3.5$, $\lambda = 0.25$. Numbers in the image are referring to the values of parameter κ.

Let us remember that in [Kaniadakis et al., 2020], pdf is proposed as the following function:

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

In the same reference, the hazard function is given as:

$$h_\kappa(t|\alpha,\beta)=\frac{\alpha\beta t^{\alpha-1}}{\sqrt{1+\kappa^2\beta^2 t^{2\alpha}}} \qquad (22)$$

Eq. (22) in the formalism of (17) becomes: $h_\kappa(t|\beta,\lambda)=\frac{\beta\lambda(\lambda x)^{\beta-1}}{\sqrt{1+\kappa^2(\lambda x)^{2\alpha}}}$ (23)

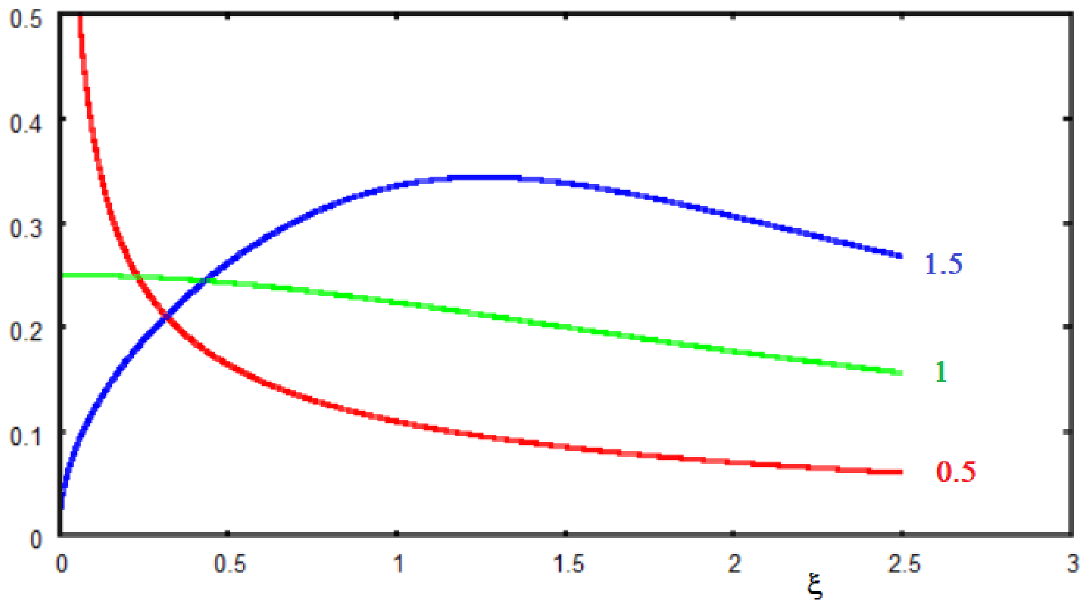


Fig. 13: Hazard function of κ-Weibull for parameters $\kappa=0.5$, $\lambda=0.25$. Numbers in the image are referring to the values of parameter β .

In the Figure 13, it is given the hazard function of κ-Weibull, where parameter κ has a value fixed to $\kappa=0.5$. As in the previous figures, $\lambda=0.25$. Numbers in Fig. 13

are referring to the values of parameter β. We can see that, due to the role of κ parameter in the tail of the distribution, the behaviour is different from that given in the Figure 10 for function $(\lambda x)^{(\beta-1)}$, referring to the Weibull distribution. Then, also a "bathtub" diagram is modified for the κ-Weibull.

In the following Figure, a real case is proposed as that in the Figure 2 of Ref. [Kaniadakis et al., 2020].

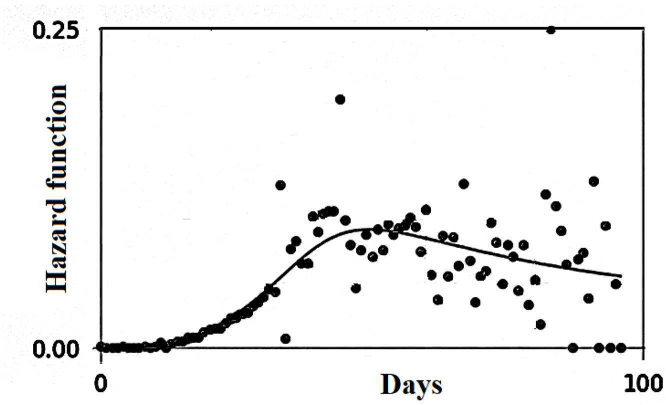


Fig. 14 : This plot is illustrating the Figure 2 of Ref. [Kaniadakis et al., 2020].

In the Figure 2 of [Kaniadakis et al., 2020], among other functions, the theoretical continuous curve and empirical dots are plotted for the κ-Weibull hazard function versus time (see [Kaniadakis et al., 2020] for parameters). The data are those concerning Covid-19 in China, at the beginning of the pandemic.

The hazard function is unimodal with a large tail, as in the Figure 12.

Let us consider Ref. [Kartsonaki, 2016]. Its told that "A Weibull distribution allows a monotonic (either continuously increasing or decreasing hazard) and a log-logistic distribution allows either a monotonic or a unimodal hazard function." The figures given above show unimodal behaviours for κ-Weibull. For this reason, a further comparison to the log-logistic function, mentioned in [Kartsonaki, 2016], is interesting.

11. Log-logistic function

In the formalism (16a) used before, the log-logistic function is given by:

$$f(x|\lambda,\beta)=\frac{\lambda\beta(\lambda x)^{(\beta-1)}}{[1+(\lambda x)^{\beta}]^2} \qquad (24)$$

In r-project.org, see please https://search.r-project.org/... html/Llogis.html .

In the following figure, this function is compared to the 3-parameter Weibull and the κ-Weibull.

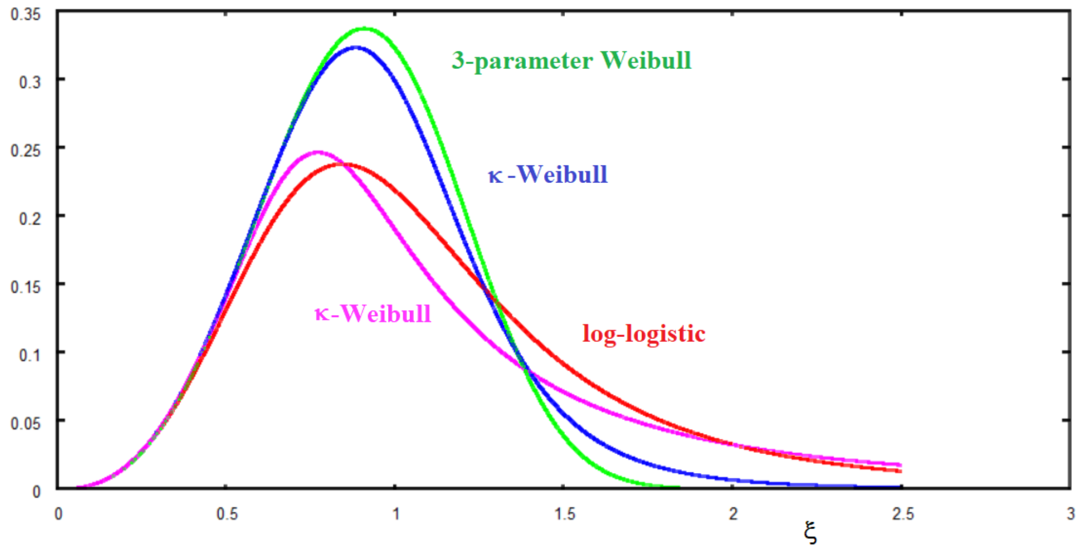


Fig. 15 : In the plot given above, we can see the log-logistic pdf in red colour, given for parameters $\beta=3.5$, $\lambda=0.25$. The same parameters are used for the 3-parameter Weibull (in green) , with $\alpha=1.01$. Two κ-Weibull are considered for $\kappa=0.5$ (blue) and $\kappa=2.0$ (violet).

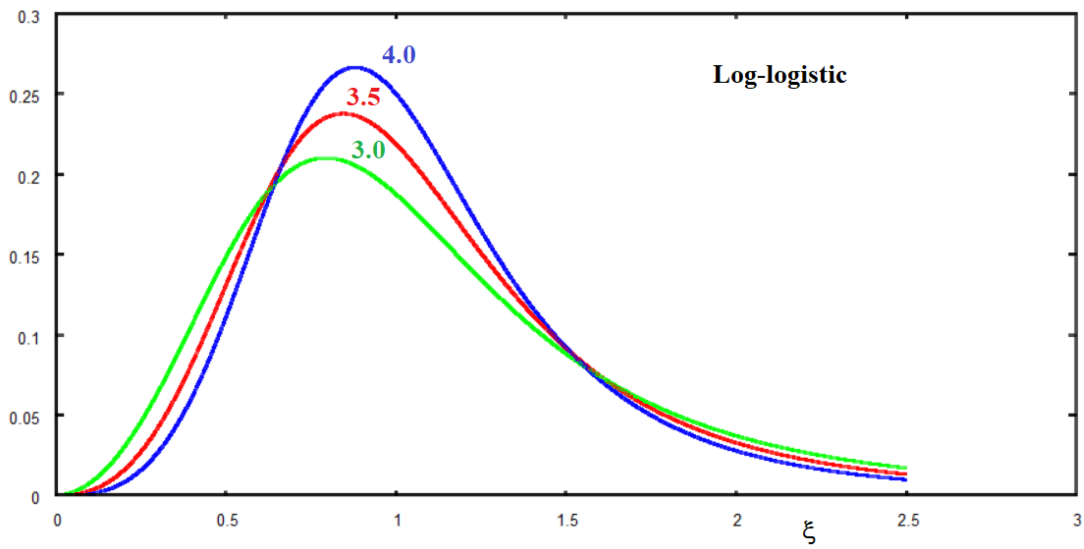


Fig. 16 : Log-logistic pdf for thee different values of parameter β (3.5 for the curve in red, 3.0 for the green curve and 4.0 for the blue one). The three curves are plotted with parameter $\lambda=0.25$.

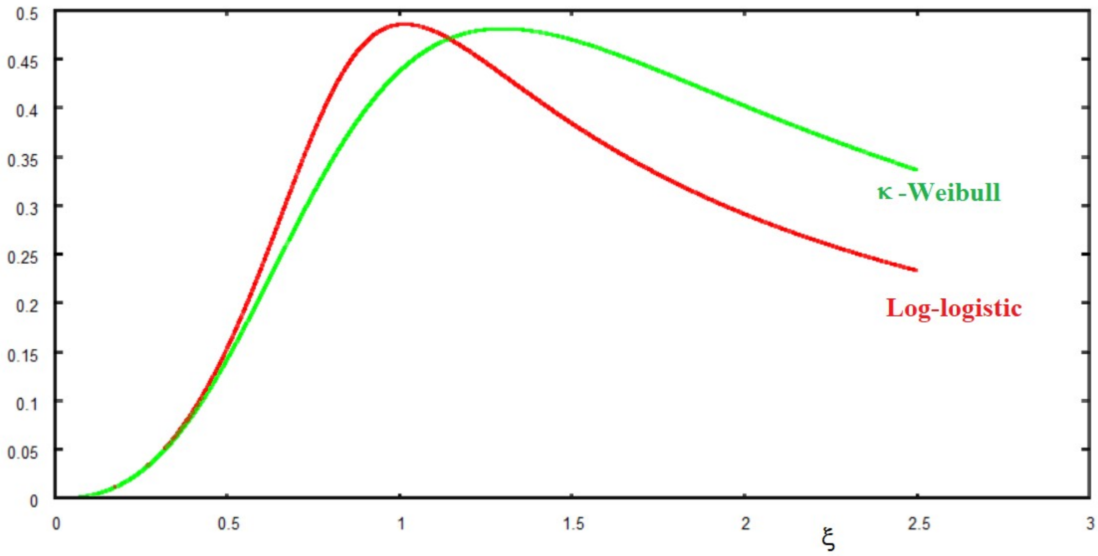


Fig. 17 : Hazard functions of log-logistic (red) and κ-Weibull (green). Parameters are $\beta=3.5$ and $\lambda=0.25$. The κ-Weibull is given for $\kappa=1.5$.

Both log-logistic and κ-Weibull functions have unimodal hazard functions, as shown in the Figure 17. Note that different behaviour exists for the decreasing risk of failure over the lifetime of the organism or device.

12. The logistic distribution

The log-logistic distribution is the probability distribution of a random variable whose logarithm has a logistic distribution. As we have seen, the log-logistic is given by: $f(x|\lambda,\beta)=\frac{\lambda\beta(\lambda x)^{(\beta-1)}}{[1+(\lambda x)^{\beta}]^2}$. However, how is the logistic distribution? In the same formalism, it is given as:

$$f(x|\lambda,\mu)=\frac{\lambda e^{-\lambda(x-\mu)}}{\left[1+e^{-\lambda(x-\mu)}\right]^2}$$

where $1/\lambda$ is the scale parameter and μ is the location parameter. We can also write the distribution in the following form:

$$f(x|\lambda,\mu)=\frac{\lambda}{\left[e^{\lambda(x-\mu)/2}+e^{-\lambda(x-\mu)/2}\right]^2}=\frac{\lambda}{4}\,\mathrm{sech}^2(\lambda(x-\mu)/2) \quad .$$

The cumulative function is given by:

$$F(x|\lambda,\mu)=\frac{1}{1+e^{-\lambda(x-\mu)}} = \frac{1}{2}+\frac{1}{2}\tanh\left(\lambda(x-\mu)/2\right)$$

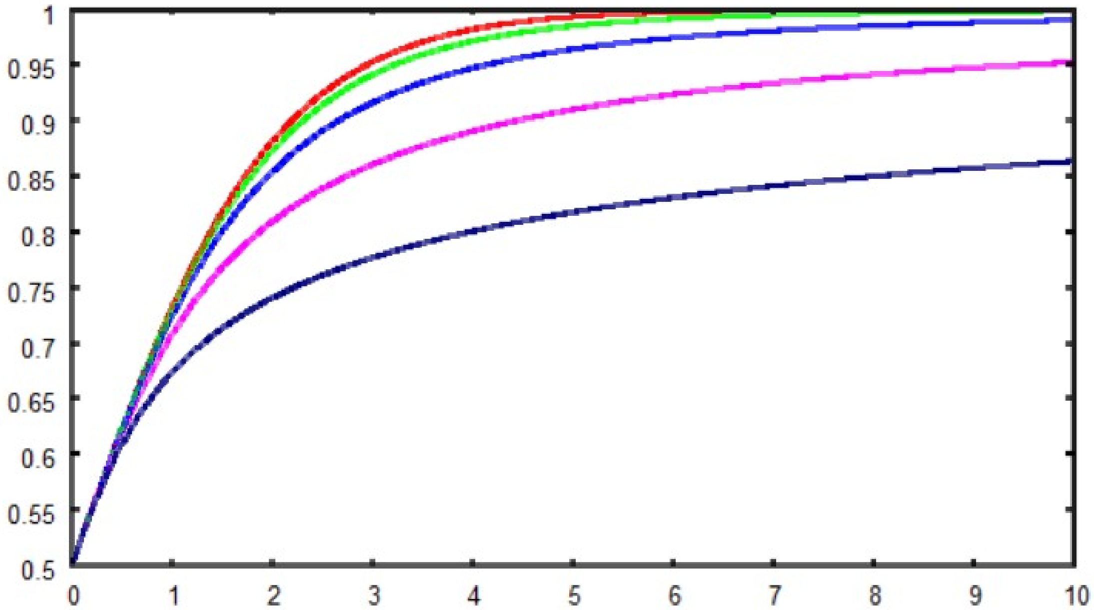


Fig.18 – The cumulative function of logistic distribution (in red), and of the κ-logistic for different values of parameter κ ($\kappa=0.25$ green, $\kappa=0.5$ blue, $\kappa=1.0$ violet, $\kappa=2.0$ dark blue). Parameters are $\lambda=1.0$, $\mu=0.0$.

13. κ-logistic distribution

In [Kaniadakis, 2021], the author is introducing new classes of statistical distributions, which are the κ-deformed version of known distributions such as the Generalized Gamma, Weibull, Logistic. In the given reference, the author categorize the distributions in five types.

Type I, corresponding to the Generalized Gamma distribution

Type II, Weibull distribution

Type III, Generalized Logistic distribution

Type IV, not defined

Type V, Exponential distribution

The κ-logistic is given by:

$$F_\kappa(x|\lambda,\mu)=\frac{1}{1+\exp_\kappa-\lambda(x-\mu)}$$

14. The κ-exponential to deform Gumbel and Gompertz distributions

As previously seen, the κ-exponential can be used to modify distributions which contains exponentials. Besides those mentioned in [Kaniadakis, 2021], we can add other distributions such as the Gumbel and Gomperts functions, which can be turned into κ-Gumbel and κ-Gompertz distributions.

The Gumbel distribution is used to model the distribution of the maximum (or the minimum) of a number of samples of various distributions. It is also known as the log-Weibull distribution and the double exponential distribution. The Gumbel distribution was proposed in [Gumbel, 1935, 1941]. The cumulative distribution is given, in the formalism that we are using here as:

$$F(x|\mu,\lambda)=\exp(-\exp(-\lambda(x-\mu)))$$

Then, we can define the cumulative function κ-Gumbel:

$$F_\kappa(x|\mu,\lambda)=\exp_\kappa(-\exp_\kappa(-\lambda(x-\mu)))$$

The Gompertz distribution is a continuous probability distribution, named after Benjamin Gompertz (1779 – 1865).

The cumulative function is:

$$F(x|b,\beta)=1-\exp(-\beta(\exp(bx)-1))$$

with two parameters b,β. The κ-deformed cumulative function is:

$$F_\kappa(x|b,\beta)=1-\exp_\kappa(-\beta(\exp_\kappa(bx)-1))$$

In [Nadarajah, 2006], it is proposed an exponentiated Gumbel distribution, in the framework of climate applications. In the introduction of the article, the authors tells that the Gumbel distribution is perhaps the most widely applied distribution for climate modeling. Application include the global warning problems, fllod frequency analysis, offshore modeling, rainfall modeling, and wind sped modeling.

In [Nadarajah, 2006], the cumulative $F(x|\mu,\lambda)=\exp(-\exp(-\lambda(x-\mu)))$ becomes:

$$F(x|\mu,\lambda)=1-[1-\exp(-\exp(-\lambda(x-\mu)))]^\alpha$$

We can have an exponentiated κ-Gumbel distribution:

$$F_\kappa(x|\mu,\lambda)=1-[1-\exp_\kappa(-\exp_\kappa(-\lambda(x-\mu)))]^\alpha$$

For plots see please https://doi.org/10.5281/zenodo.6425678

It is necessary to stress that we can have two different κ parameters, for instance, a κκ-Gumbel distribution which is given by:

$$F_{\kappa_2,\kappa_1}(x|\mu,\lambda)=\exp_{\kappa_2}(-\exp_{\kappa_1}(-\lambda(x-\mu)))$$

In the following figure, the cumulative function (upper panel) and the pdf (lower panel) of the Gumbel distribution (in red) are given with parameters $\mu=0.0,\lambda=1.0$. The same parameter are used for the κκ-Gumbel. The green curve has parameters $\kappa_1=0.05,\ \kappa_2=2.0$ and the blue curve has . $\kappa_1=2.0,\ \kappa_2=0.05$.

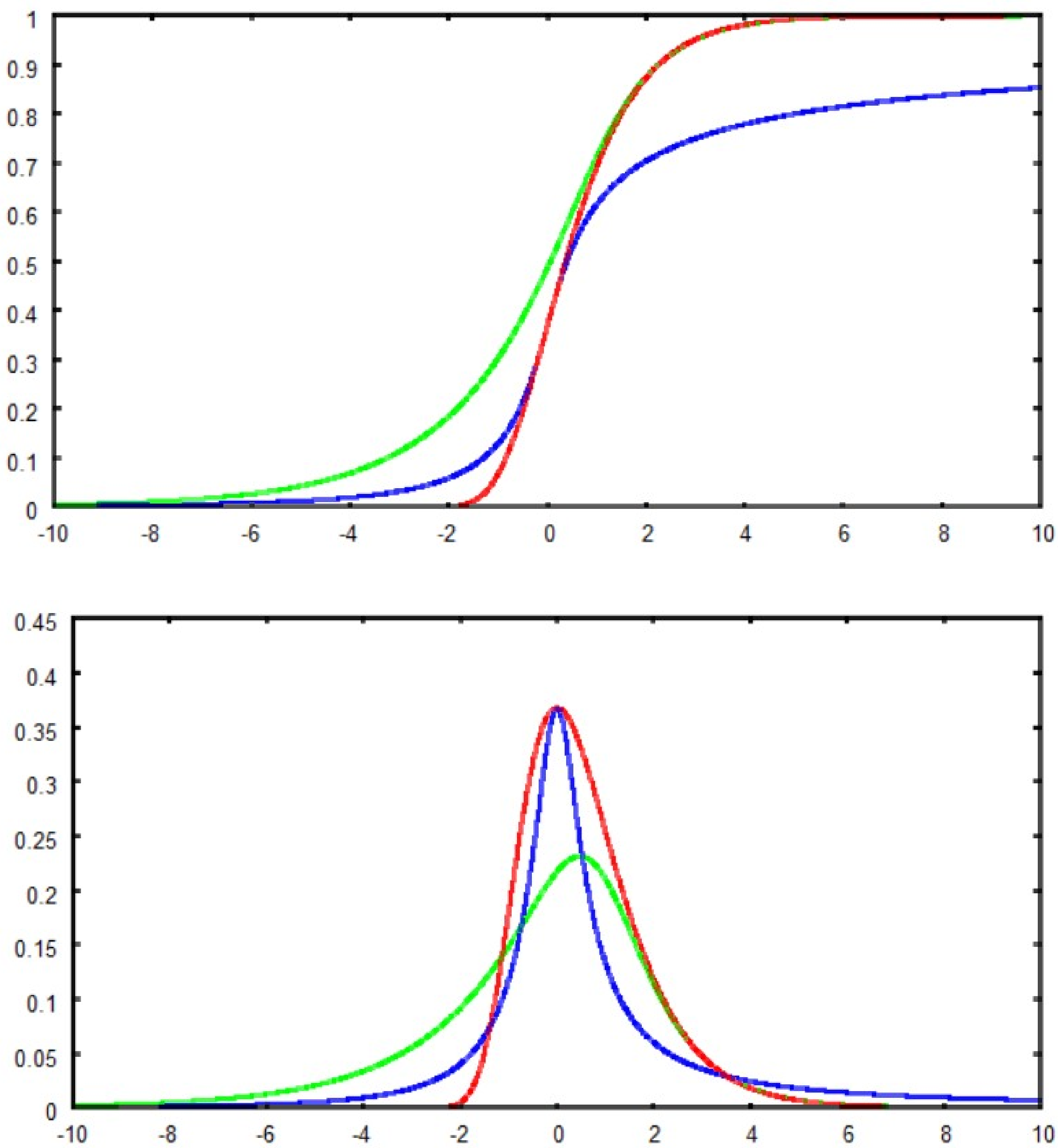


Fig. 19 – Gumbel (red) and κκ-Gumbel
(green, $\kappa_1 = 0.05$, $\kappa_2 = 2.0$; blue, $\kappa_1 = 2.0$, $\kappa_2 = 0.05$).

15. The Gamma distribution

We have previously mentioned the Kaniadakis Type I distributions, corresponding to the Generalized Gamma distribution [Kaniadakis, 2021]. Then, let us consider this distribution, which contains, as shown in [Rinne, 2008], the Weibull distribution as a specific case. In [Rinne, 2008], the four-parameter gamma distribution, also name generalized gamma distribution is written in the following form:

$$f(x|a,b,c,d)=\frac{c(x-a)^{cd-1}}{b^{cd}\Gamma(d)}\exp\left[-\left(\frac{x-a}{b}\right)^{c}\right]$$

Among the special cases of the generalized gamma distribution we can find three-parameter Weibull distribution ($\Gamma(1)=1$)

$$f(x|a,b,c,1)=\frac{c(x-a)^{c-1}}{b^{c}}\exp\left[-\left(\frac{x-a}{b}\right)^{c}\right]$$

that we can easily compare to (3):

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

Let us add reference to [Preda, 1985]. The article shows that the generalized gamma distribution can be obtained by means of the principle of maximum entropy. “The principle of maximum entropy was established, independently by Ingarden (1963), Jaynes (1957), Kullback and Leibler (1951). According to this principle we choose the probability distribution which maximizes the entropy compatible to some set of restrains. Using this principle, Jaynes (1957), Kampé de Fériet (1963), Ingarden and Kossakowski (1971), Preda (1982, 1984) obtained some usual probability distributions”.

16. The Student distribution

In a certain manner, the log-logistic distribution is linked to a distribution which is fundamental for statistical analyses. This distribution is the Student distribution. About it, let us just mention what is told in [Yang, 2017].

“The Student's t-test is a very powerful method for testing the null hypothesis to see if the means of two normally distributed samples are equal. This method was designed by W. S. Gosset in 1908 and he had to use a pen name ‘Student’ because of his employer's policy in publishing research results at that time. This is a powerful method for hypothesis testing using small-size samples. This test can also be used to test if the slope of the regression line is significantly different from 0. It has become one of the most popular methods for hypothesis testing. The theoretical basis of the t-test is the Student's t-distribution for a sample population with the unknown standard deviation σ, which of course can be estimated in terms of the sample variance S^2 from the sample data".

The Student's t-distribution (or simply the t-distribution) is symmetric and bell-shaped, like the normal distribution, but it has heavier tails. It means that the distribution is more prone to give values that fall far from its mean. The distribution is given by:

$$f(t|\nu)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left[1+\frac{t^2}{\nu}\right]^{-\frac{\nu+1}{2}} \qquad (25)$$

where ν is the number of degrees of freedom and Γ is the gamma function.

Let us assume $\nu=3$, we have:

$$f(t|3)=\frac{2}{\pi\sqrt{3}}\left[1+\frac{t^2}{\nu}\right]^{-2}=\frac{2}{\pi\sqrt{3}\left(1+\frac{t^2}{3}\right)^2} \qquad (26)$$

Just to play with distributions, let us consider $x=t$, $\lambda=1/\sqrt{3}$ (in this case, with dimension $[x^{-1}]$), we have :

$$f(x|\lambda,2)=\frac{2\lambda}{\pi\left(1+(\lambda x)^2\right)^2} \qquad (27)$$

Generalizing, by inserting $(\lambda x)^{(2-1)}$ instead of $1/\pi$, we remove the symmetry and have:

$$f(x|\lambda,2)=\frac{2\lambda(\lambda x)^{(2-1)}}{\left(1+(\lambda x)^2\right)^2} \qquad (28).$$

Therefore, in (28), we have the log-logistic for $\beta=2$.

About the Gamma function, here a remarkable link:

https://www.robertobigoni.it/Matematica/Trascendenti/f11/f11.htm

17. The paralogistic distribution

At the link docs.plm.automation.siemens, it is given the paralogistic distribution, that we compare to the log-logistic:

$$f(x|\lambda,\beta)=\frac{\lambda\beta(\lambda x)^{(\beta-1)}}{[1+(\lambda x)^{\beta}]^2} \quad \text{log-logistic} \quad (24)$$

$$f(x|\lambda,\beta)=\frac{\lambda\beta^2(\lambda x)^{\beta}}{(\lambda x)[1+(\lambda x)^{\beta}]^{\beta+1}} \quad \text{paralogistic} \quad (30)$$

18. Unimodal Hazard

Let us consider again the hazard function, to stress some facts.

In Ref.[Greenwich, 1992], it is stressed that a unimodal hazard rate function is mainly used for modelling "a failure rate that has a relatively high rate of failure in the middle of expected life time. This unimodal hazard rate function has two shape parameters. One of the parameters indicates the location (time) of the mode and the other controls the height of the mode. In effect, these two parameters index the class of unimodal hazard rate functions. The reliability function and the failure density function of the unimodal hazard rate function are relatively uncomplicated and mathematically tractable." Abstract of [Greenwich, 1992] also stresses that unimodal hazard rate functions are "particularly useful when the time of the peak failure rate is of prime interest. The failure distribution provides a practical way of estimating the peak failure time".

The peak is important, but it is also important the tail of the risk of failure.

19. Bathtub and Unimodal Hazard Curves

Let us consider Ref. [Lacey et Nguyen, 2015]. In this discussion it is told that the shape of the hazard function "allows us to see changes in risk over time for data modeled with lifetime distributions. For example, in the typical bathtub shape, the flat line at the bottom represents the useful lifetime. The longer that flat line is, the longer lifetime the distribution can model. A V-shaped bathtub with a very short flat line is more suitable to model shorter lifetimes, such as the lifetime of a mosquito, but a wider U-shaped bathtub with a longer flat line would be more suitable for modeling the lifetime of a human. On the other hand, parametric lifetime distributions with unimodal hazard shapes can be used to model situations such as survivability after surgery, where risk quickly increases due to the chances of complications such as infection, and then decreases as the patient recovers".

20. Survival and Reliability Analysis

In the Reference [NCSS], an overview is proposed. In the Introduction it is stressed that, when the analysis is concerning "a biological event associated with animals (including humans), it is usually called survival analysis". When the event concerns "machines in an industrial setting, it is usually called reliability analysis". It is also important to note that Ref. [NCSS] assumes the survival analysis, as emphasizing a nonparametric estimation approach (Kaplan-Meier estimation), while the reliability analysis is emphasizing a parametric approach (Weibull or lognormal estimation). However, 'reliability' and 'survival' are two terms referring to the same type of analysis.

Survival analysis is the study of the distribution of life times.

It means that the survival analysis is the study of "the elapsed time between an initiating event (birth, start of treatment, diagnosis, or start of operation) and a terminal event (death, relapse, cure, or machine failure). The data values are a mixture of complete (terminal event occurred) and censored (terminal event has not occurred) observations" [NCSS]. According to available data values, a survival analysis allows to have some statements about the survival distribution of the failure times.

Let t be the elapsed time until the occurrence of a specified event. The probability distribution of t may be specified using one of the basic functions that we have seen before. Let us note that, once one of these functions has been specified, the others may be derived using mathematical relationships [NCSS]. Basic functions are:

1) the Probability Density Function, $f(t)$, that is the probability that an event occurs at time t;

2) the Cumulative Distribution Function, $F(t)$, which is the probability that an individual survives until time t;

3) the Survival Function, $S(t)$ or Reliability function, $R(t)$. This is the probability that an individual survives beyond time t. As told in [NCSS], this function is the first usually studied. It may be estimated using the nonparametric Kaplan-Meier curve or one of the parametric distribution functions;

4) the Hazard Rate, $h(t)$, which is the probability that an individual at time t experiences the event in the next instant. It is a fundamental quantity in survival analysis. It is also known as the conditional failure rate in reliability. Ref. [NCSS] stresses that the empirical hazard rate may be used to identify the appropriate probability distribution of a particular mechanism, since each distribution has a different hazard rate function. "Some distributions have a hazard rate that decreases with time, others have a hazard rate that increases with time, some are constant, and some exhibit

all three behaviors at different points in time" [NCSS];

5) Cumulative Hazard Function, $H(t)$.

21. Nonparametric estimators

Two competing nonparametric estimators of the survival distribution, S(t), available, according to [NCSS]. The first is the Kaplan-Meier Product limit estimator and the second is the Nelson-Aalen estimator of the cumulative hazard function, $H(t)$.

The Kaplan-Meier Product limit estimator is defined as follows in the range of time values for which there are data [NCSS]:

$$\hat{S}(t) = \begin{cases} 1 & \text{if } t<t_i \\ \prod_{i:\ t_i \le t} \left[1-d_i/Y_i\right] & \text{if } t_1 \le t \end{cases} \tag{31}$$

In (31), d_i represents the number of failures at time t_i and Y_i and represents the number of individuals who are at risk at time t_i .

An alternative estimator is Nelson-Aalen Hazard Confidence Limit of $H(t)$:

$$\hat{H}(t) = \begin{cases} 0 & \text{if } t<t_1 \\ \sum_{i:\ t_i \le t} d_i/Y_i & \text{if } t_1 \le t \end{cases} \tag{32}$$

22. Parameter estimation by means of the maximum likelihood

The parameters, such as those in the Weibull reliability distribution, can be estimated with different methods. Two methods are based on the maximum likelihood or on the least squares regression to the probability plot. The probability plot method allows to have a nice "visual analysis of the goodness of fit of the distribution to the data" [NCSS]. A maximum likelihood estimation is usually preferred in statistical analyses because it allows to provides estimates of standard errors and confidence limits. In the Ref. [NCSS], it is noted that "there are situations in which maximum likelihood does not do as well as the regression approach". As an example it is proposed the estimation of the threshold parameter. The Reference suggests to find the threshold parameter by means of a regression and then apply the maximum likelihood estimation, with the threshold value assumed as a known quantity.

Let us consider shortly the maximum likelihood estimation (MLE) method. It allows to find the most "likely" values of the distribution parameters, from a set of data, by maximizing the value of a likelihood function. This function is based on the probability density function. Let us consider a generic pdf:

$$f(x|\theta_1, \theta_2, \ldots, \theta_k)$$

Here, x is representing the time-to-failure and $\alpha_1, \alpha_2, \ldots, \alpha_k$ are the parameters to be estimated. For the given data, the likelihood function is a product of of the pdf, so that:

$$L = \prod_{i=1}^{n} f(x_i|\alpha_1, \alpha_2, \ldots, \alpha_k)$$

(here, we use a formalism such as that of LINK , ReliaSoft).

In this function, n is the number of failure data points and is the *i-th* failure time. Mathematically, it is easier to manipulate not the function but its logarithm of it. This log-likelihood function then has the form:

$$\Lambda = \ln L = \sum_{i=1}^{n} \ln f(x_i|\alpha_1, \alpha_2, \ldots, \alpha_k) \quad (33)$$

Then, we have to find the values of parameters, for which it is resulting the highest value for function (33). It happens when:

$$\frac{\partial \Lambda}{\partial \alpha_i} = 0, \quad j = 1, 2, \ldots, k \quad (34)$$

As a consequence, we have a number of equations with an equal number of unknowns, the parameters. The solution can be relatively simple, in the case that if there are closed-form for the partial derivatives; otherwise, some numerical techniques can be used.

The process given above, can be easily illustrated in the case of the one-parameter exponential distribution. Since there is only one parameter, let us call it λ , there is only one equation to be solved. First, let us write the likelihood function:

$$L(\lambda, t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} f(t_i) = \prod_{i=1}^{n} \lambda e^{-\lambda t_i} = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} t_i\right)$$

Let us pass from the likelihood function to the log-likelihood function, which has has the form:

$$\Lambda = \ln L = n \ln \lambda - \lambda \sum_{i=1}^{n} t_i$$

Taking the derivative of this equation, we have:

$$\frac{\partial \Lambda}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i = 0$$

Then, we find the estimation of the parameter $\tilde{\lambda} = n / \left(\sum_{i=1}^{n} t_i\right)$. The tilde mark indicated that this is an estimation of the parameter. In the case of the one-parameter exponential function, we have a closed-form solution for a MLE (maximum likelihood estimator) value for its parameter.

Let us consider the Weibull distribution. The maximum likelihood approach is described in Ref. [Evans et al., 2019]. Here, let us use the notation of this reference, so that the Weibull distribution is given as $f(x) = a b^{-1} (x/b)^{a-1} \exp\left[-x^a / b^a\right]$.
The likelihood function is:

$$L(x_1, x_2, \ldots, x_n; a, b) = \prod_{i=1}^{n} (a/b)(x_i/b)^{a-1} \exp\left[-x_i^a / b^a\right] \quad (35)$$

Using the logarithm, we have: $\Lambda = n \ln a - n a \ln b + (a-1) \sum_{1}^{n} \ln x_i - \sum_{1}^{n} \frac{x_i^a}{b^a}$ [Yang et al., 2019]. Appling the partial derivations:

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a} - n \ln b + \sum_{1}^{n} \ln x_i - \sum_{1}^{n} (x_i/b)^a \ln(x_i/b) = 0 \qquad (36)$$

$$\frac{\partial \ln L}{\partial b} = -n\frac{a}{b} + \frac{a}{b^{a+1}}\sum_1^n x_i^a = 0 \qquad (37)$$

From the second equation: $b^a = \frac{1}{n}\sum_1^n x_i^a$. Eliminating b , and simplifying, we obtain:

$$\left[\frac{\sum_1^n x_i^a \ \ln x_i}{\sum_1^n x_i^a} - \frac{1}{a}\right] = \frac{1}{n}\sum_1^n \ln x_i \qquad (38)$$

Solving this equation, we have an estimate for the shape parameter, denoted by $\hat{a}$. Then, we have the other parameter:

$$\hat{b} = \sum_i^n x_i^{\hat{a}}/n \qquad (39)$$

The circumflex mark means that this is an estimate. The solution can be obtained by a numerical approach.

22.1 Excel's solver

See how to use Excel

https://www.real-statistics.com/distribution-fitting/distribution-fitting-via-maximum-likelihood/fitting-weibull-parameters-mle/ ,

archived here.

23. The maximum likelihood in the case of κ-Weibull

Before considering the estimation of Weibull parameters by means of a plot model, let us address the problem of the maximum likelihood in the case of the κ-Weibull. It hass been discussed in Ref. [Clementi et al, 2008].

In this reference, the probability density function is given in the form (7):

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

The reference is defining the complementary cumulative distribution function (ccdf) of the κ-statistics:

$$P_>(x|\alpha,\beta)=\exp_\kappa(-\beta\,x^\alpha) \qquad (40)$$

The complementary cumulative distribution function is also known as the tail distribution or exceedance, and is defined as:

$$P_>(x)=1-F(x)$$

where $F(x)$ is the cumulative function. We have already seen the ccdf: in survival analysis, it is called the survival function, while the term reliability function is common in engineering.

Ref. [Clementi et al, 2008] is giving the moments of κ-Weibull. As told in the reference, assuming that all observations are independent, the likelihood function is:

$$L=\prod_{i=1}^n f_\kappa(x_i|\alpha,\beta)=(\alpha\,\beta)^n\prod_{i=1}^n\frac{x_i^{\alpha-1}}{\sqrt{1+\kappa^2\beta^2x_i^{2\alpha}}}\ \exp_\kappa(-\beta\,x_i^\alpha) \qquad (41)$$

As told by the authors in the article, "obtaining explicit expressions for the ML estimators of the three parameters is difficult, making direct analytical solutions intractable, and one needs to use numerical optimization methods". However, it is possible to obtain an expression of parameter β as a function of the parameters κ,α :

$$\beta=\frac{1}{2\kappa}\left[\frac{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)}{\kappa+\alpha\,\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}\right]^{\alpha} \qquad (42)$$

The problem is reduced to find the other two parameters κ,α (see Clementi et al., 2008, for further discussions).

Let us consider the form (17) of κ-Weibull:

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

Assuming $\lambda=1$, and a dimensionless $\bar{x}$: $f_\kappa(\bar{x}|\beta)=\frac{\beta\bar{x}^{\beta-1}}{\sqrt{1+\kappa^2\bar{x}^{2\beta}}}\ \exp_\kappa(-\bar{x}^\beta)$.

The likelihood function and its logarithm become:

$$L=\prod_{i=1}^{n} f_\kappa(\bar{x}_i|\beta)=\beta^n\prod_{i=1}^{n}\frac{\bar{x}_i^{\beta-1}}{\sqrt{1+\kappa^2\bar{x}_i^{2\beta}}}\ \exp_\kappa(-\bar{x}_i^\beta)$$

$$\Lambda=n\ln\beta+(\beta-1)\sum_{i=1}^{n}\bar{x}_i-\frac{1}{2}\sum_{1}^{n}\ln(1+\kappa^2\bar{x}_i^{2\beta})+\sum_{i=1}^{n}\ln\left(\exp_\kappa(-\bar{x}_i^\beta)\right)$$

We could determine, numerically, parameter β , assuming κ as a fixed data.

24. A plot model for Weibull

Another manner to estimate the parameter is to use a plot; that is, the estimation procedure uses the data from the probability plot. Let us see how.

The Weibull cumulative distribution function is:

$$F(t)=1-e^{-\left(\frac{t-D}{C}\right)^B} \qquad (43)$$

Let us assume $D=0$, then $\ln(1-F(t))=-(t/C)^B$ and:

$$\ln(-\ln(1-F(t)))=-B\ln C+B\ln(t)$$

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

$$y=-B\ln C+Bx \qquad (44)$$

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

In [Hodges, 1994], this fact is stressed, in the following manner. Let us consider the cumulative distribution function $F(x|\beta,\alpha)$. We can easily derive:

$$\ln\ln\left(\frac{1}{1-F(x|\beta,\alpha)}\right)=\beta\ln(x/\alpha)=-\beta\ln\alpha+\beta\ln x \qquad (45)$$

This relation gives a straight line when ln ln [.] is plotted against ln x. Therefore, we have a useful method of determining whether the Weibull distribution is appropriate [Hodges, 1994] . Moreover, it is also possible to find the intercept and slope of the straight line, and , as a consequence, the parameters of the Weibull distribution. The plot will also identify an exponential function, if $\beta=1$.

In [Hodges, 1994], it is also proposed an example from Ref. [Berrettoni, 1964].

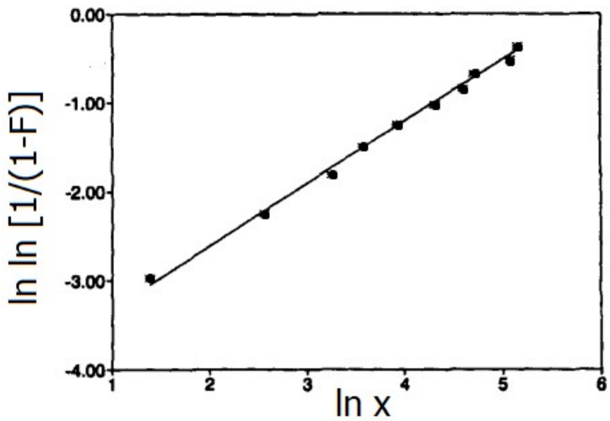


Fig. 20: An example from [20].

The straight line in the figure is from an experiment in [Berrettoni, 1964]. It is a plot of capacitor failure data. The random variable is hours to failure. The straight line corresponds to a Weibull distribution having parameters as $\beta=0.70$ and $\alpha=310$.

24.1 Regression - See the example at

https://www.real-statistics.com/distribution-fitting/fitting-weibull-regression/ ,

archived here.

24.2 MATLAB and curve fitting - "I have some data and I want to fit a Weibull distribution. What MATLAB functions can I use to do Weibull curve fitting?" Here an example.

25. The exponentiated Weibull distribution and others

In the reliability analysis, the Weibull distribution is commonly used. Then, it is not surprising that several generalizations of it have been proposed, in order to improve its performance. One of them is appearing in [Zacks, 1984]. In the given reference, complex mechanical systems, such as cars and air-planes are considered. These systems have two phases in their life, the second one - the wear-out phase - being characterized by a higher failure rate. Between the two phases there is a change point. The problem consists in the estimation of this point. In this framework, in [Zacks, 1984] we find introduced a new distribution, defined as "Weibull-exponential" distribution. It is similar to the exponentiated Weibull distribution, introduced in [Mudholkar et Srivastava, 1993].

The exponentiated Weibull family of probability distributions was introduced as an extension of the Weibull family, by adding a second shape parameter. The cumulative distribution function for the exponentiated Weibull distribution is (for threshold $D = 0$) :

$$F(t)=\left[1-e^{-(t/C)^B}\right]^E \quad \text{instead of} \quad F(t)=1-e^{-(t/C)^B}$$

Here B is the first shape parameter, and E is the second shape parameter.

The density function is given by:

$$f(t)=E\,\frac{B}{C}\,(t/C)^{B-1}\,e^{-(t/C)^B}\left[1-e^{-(t/C)^B}\right]^{E-1} \qquad (46)$$

Let us remember that the probability density is the derivative of the cumulative distribution function.

There are two important special cases of the exponentiated distribution. When $E = 1$, we have the Weibull distribution. If $B = 1$, we find the exponentiated exponential distribution, introduced by Gupta and Kundu [Kundu et Gupta, 1999]:

$$f(t)=\frac{E}{C}\ e^{-(t/C)}\left[1-e^{-(t/C)}\right]^{E-1} \qquad (47)$$

The 3-parameter Weibull (Marshall–Olkin based) and the exponentiated Weibull distributions are two of the many continuous modifications of the Weibull distribution. In Ref. [Almalki et Nadarajah, 2014], we can find the Inverse Weibull, the Log-Weibull, the Compound Weibull distributions, the Reflected Weibull, the Gamma Weibull, the Kies and Phani's modified Weibull distributions, the Exponentiated Weibull (that we have previously discussed), the Generalized Weibull, the Additive Weibull, the Extended Weibull (that is the Marshall–Olkin based one), the Power Lindley distribution, the Generalized power Weibull distribution, the Modified Weibull extension, the Beta Weibull distributions, the Odd Weibull, the Flexible Weibull extension, the Generalized modified Weibull, the Sarhan and Zaindin's modified Weibull distribution, the Kumaraswamy Weibull distribution, and the Almalki and Yuan's modified Weibull distribution.

For comparison, let us write Weibull, exponentiated Weibull and κ-Weibull:

$$f(t|B,C)=\frac{B}{C}(t/C)^{(B-1)}e^{-(t/C)^B} \quad \text{(Weibull)},$$

$$f(t|B,C,E)=E\ \frac{B}{C}\ (t/C)^{B-1}\ e^{-(t/C)^B}\left[1-e^{-(t/C)^B}\right]^{E-1} \quad \text{(exponentiated Weibull)}$$

$$f_\kappa(t|B,C)=\frac{B}{C}(t/C)^{B-1}\ \frac{\exp_\kappa\{-(t/C)^B\}}{\sqrt{1+\kappa^2(t/C)^{2B}}} \quad \text{(κ-Weibull)},$$

In the different formalism (16a) previously used:

$$f(x|\beta,\lambda)=\beta\lambda(\lambda x)^{\beta-1}\ \exp(-\lambda^{\beta}x^{\beta})$$

$$f(x|\beta,\lambda,\epsilon)=\epsilon\beta\lambda(\lambda x)^{\beta-1}\ \exp(-\lambda^{\beta}x^{\beta})\left[1-\exp(-\lambda^{\beta}x^{\beta})\right]^{\epsilon-1}$$

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

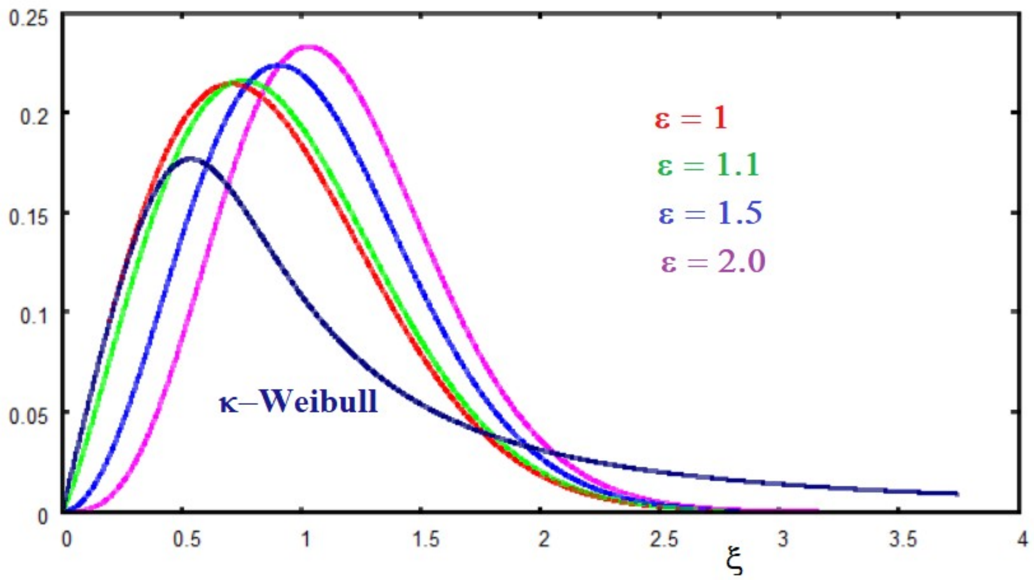


Fig. 21 : In the plot given above, we can see the Weibull pdf in red colour, given for parameters $\beta=2.0$, $\lambda=0.25$. The other curves are the exponentiated Weibull for the same β,λ parameters for and $\epsilon=1.1,\ 1.5$ and 2.0 . The κ-Weibull, plotted in dark blue, is given also for comparison, with the same β,λ parameters and $\kappa=2.0$.

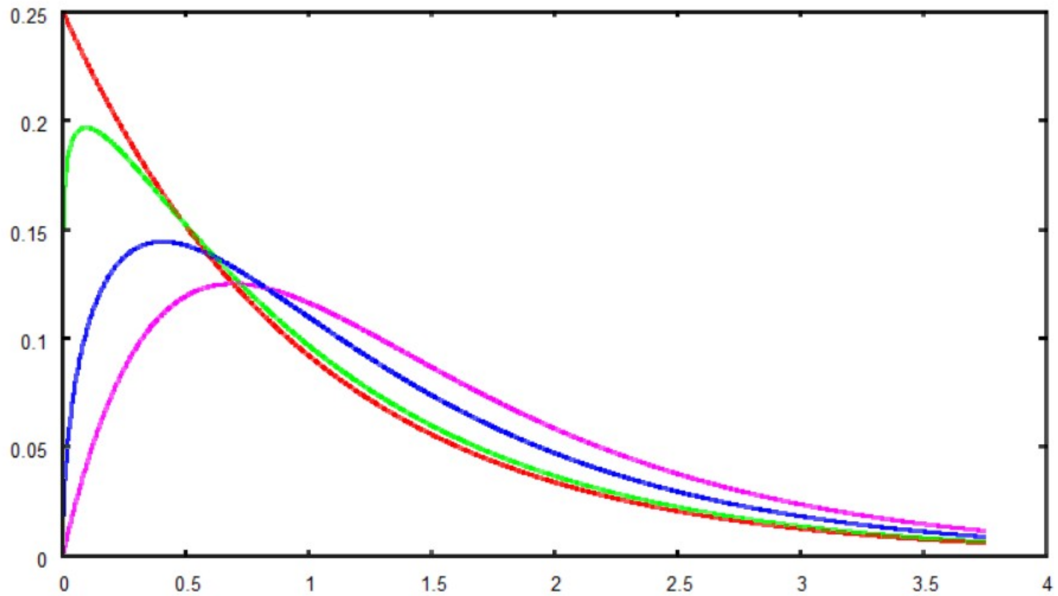


Fig. 22: Here the Weibull pdf in red colour, given for parameters $\beta=1.0$, $\lambda=0.25$ .and $\epsilon=1.0$ The other curves are the exponentiated Weibull for the same β,λ parameters for and $\epsilon=1.1,\ 1.5$ and 2.0 . In this case, being $\beta=1.0$,the exponentiated Weibul is the exponentiated exponential.

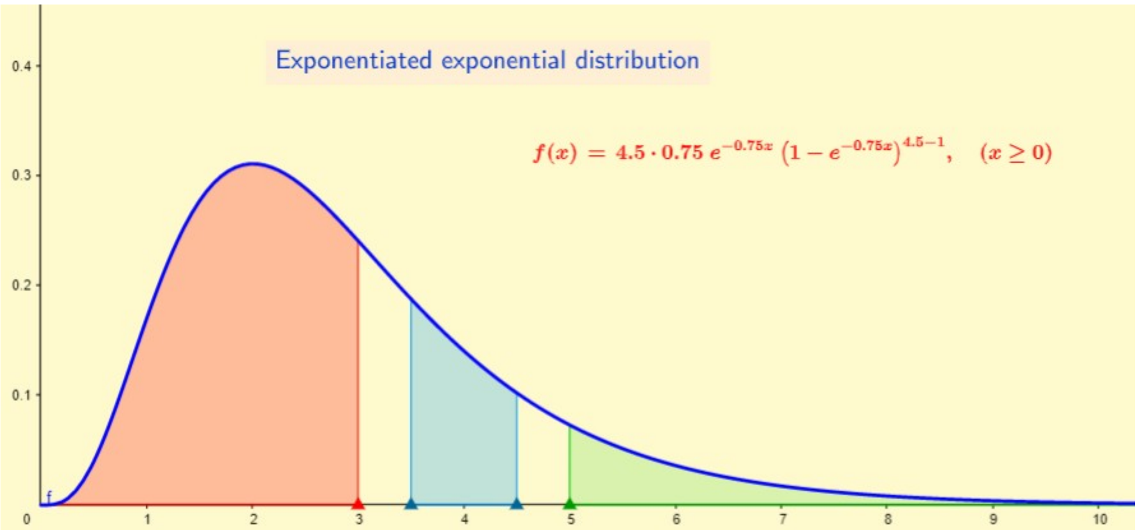


Fig. 23 – Here the exponentiated exponential distribution, in a screenshot of an application at the web site of Geogebra, https://www.geogebra.org/m/rGXMGrV4, author of the page Manoel Wallace. Many thanks to the author for this applications, where parameters cha ne changed to see how the distribution modify itself.

26. Burr Type XII distribution

Previously, we have compared the Weibull distribution and the κ-Weibull. We have also considered the comparison of κ-Weibull and log-logistic. Now, let us consider the Burr Type XII distribution, or simply Burr distribution [Burr, 1942]. It is also known as the Singh–Maddala distribution [Singh et Maddala, 1976]. The Burr distribution is sometimes called the "generalized log-logistic distribution". It is used for modeling household income.

Let us write the Burr pdf in the following form, and the other previously used:

$$f(x|\beta,\lambda,\epsilon)=\epsilon\,\beta\,\lambda(\lambda x)^{\beta-1}\left[1+\lambda^{\beta}x^{\beta}\right]^{-\epsilon-1} \quad \text{(Burr)} \qquad (48)$$

$$f(x|\beta,\lambda)=\beta\,\lambda(\lambda x)^{\beta-1}\ \exp(-\lambda^{\beta}x^{\beta}) \qquad \text{(Weibull)}$$

$$f(x|\beta,\lambda,\epsilon)=\epsilon\,\beta\,\lambda(\lambda x)^{\beta-1}\ \exp(-\lambda^{\beta}x^{\beta})\left[1-\exp(-\lambda^{\beta}x^{\beta})\right]^{\epsilon-1}$$

(exponentiated Weibull)

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

Let us note that <,when $\epsilon=1$, the Burr distribution is the log-logistic distribution.

In the following figure, we can see the comparison.

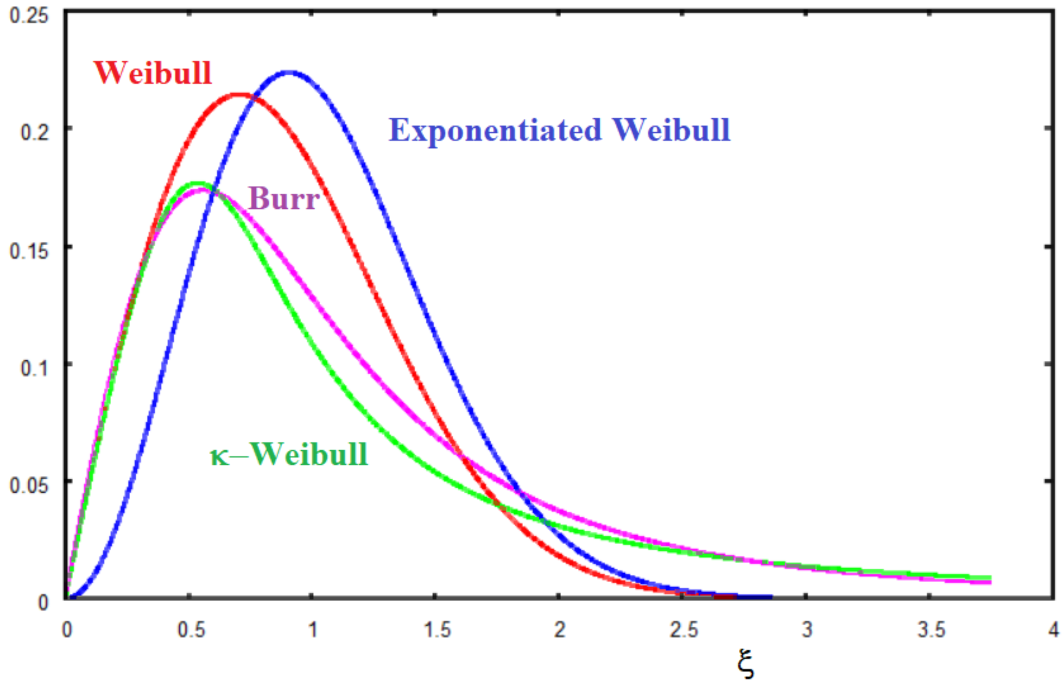


Fig. 24 : In the plot, we used parameters $\beta=2.0$, $\lambda=0.25$. In red, it is given the Weibull pdf , The other curves are the exponentiated Weibull (blue) with $\epsilon=1.5$. The κ-Weibull, plotted in green, is given for $\kappa=2.0$. The pink curve is the Burr distribution, with $\epsilon=1.5$.

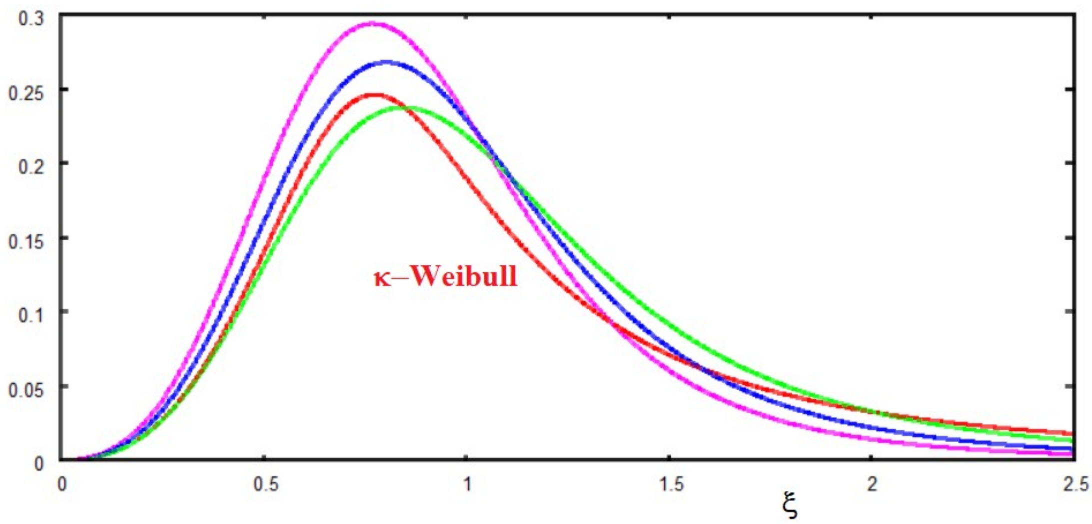


Fig. 24:5 In the plot, we used parameters $\beta=3.5$, $\lambda=0.25$. In red, it is given the κ-Weibull. The other curves are Bull distributions (green, $\epsilon=1.0$, blue, $\epsilon=1.25$, pink $\epsilon=1.5$).

27. The *q*-Weibull

Let us consider the q-Weibull [Picoli et al., 2003], in the formalism (16a) here used ($q\neq 1$):

$$f(x|\beta,\lambda,q)=(2-q)\beta\lambda(\lambda x)^{\beta-1}\left[1+(q-1)\lambda^{\beta}x^{\beta}\right]^{1/(1-q)} \quad (49)$$

In https://en.wikipedia.org/wiki/Q-Weibull_distribution , it appears the q-exponential. Let us compare (39) to the Burr distribution:

$$f(x|\beta,\lambda,\epsilon)=\epsilon\beta\lambda(\lambda x)^{\beta-1}\left[1+\lambda^{\beta}x^{\beta}\right]^{-\epsilon-1}$$

Posing $q=(2+\epsilon)/(1+\epsilon)$, we have $(2-q)=\epsilon/(1+\epsilon)$, $(q-1)=1/(1+\epsilon)$. Then:

$$f(x|\beta,\lambda,\epsilon)=\frac{\epsilon}{1+\epsilon}\beta\lambda(\lambda x)^{\beta-1}\left[1+\frac{1}{1+\epsilon}\lambda^{\beta}x^{\beta}\right]^{-\epsilon-1}$$

Assuming $1/(1+\epsilon)=\chi^{\beta}$, we have:

$$f(x|\beta,\lambda,\epsilon)=\epsilon\chi^{\beta}\beta\lambda(\lambda x)^{\beta-1}\left[1+\chi^{\beta}\lambda^{\beta}x^{\beta}\right]^{-\epsilon-1}$$

$$=\epsilon\beta\chi\lambda(\chi\lambda x)^{\beta-1}\left[1+\chi^{\beta}\lambda^{\beta}x^{\beta}\right]^{-\epsilon-1}$$

Therefore, we find the Burr distribution, when we change the scale in $\Lambda=\chi\lambda$:

$$f(x|\beta,\Lambda,\epsilon)=\epsilon\beta\Lambda(\Lambda x)^{\beta-1}\left[1+\Lambda^{\beta}x^{\beta}\right]^{-\epsilon-1} \quad (50)$$

28. Maximum likelihood and Burr distribution

Let us consider the Burr distribution. An approach to solve the maximum likelihood problem for complete and censored data is given in Refs. [Hakim et al., 2021], [Saei et al., 2019]. In these references, the Burr distribution is written as $f(x)=k\,c\,x^{c-1}\left(1+x^{c}\right)^{-(k+1)}$.

In the expressions (48) of the Burr distribution given above, it means to assume $\lambda=1$:

$$f(x|\beta,\lambda,\epsilon)=\epsilon\,\beta\,x^{\beta-1}\left[1+x^{\beta}\right]^{-\epsilon-1}$$

The likelihood function is:

$$L(x_1,x_2,\ldots,x_n;\beta,\epsilon)=\prod_{i=1}^{n}\epsilon\,\beta\,x_i^{\beta-1}\left(1+x_i^{\beta}\right)^{-(\epsilon+1)}$$

Using the logarithm, we have:

$$\Lambda=n\ln\beta+n\ln\epsilon+(\beta-1)\sum_{1}^{n}\ln x_i-(\epsilon-1)\sum_{1}^{n}\ln\left(1+x_i^{\beta}\right)$$

Appling the partial derivation:

$$\frac{\partial\ln L}{\partial\epsilon}=\frac{n}{\epsilon}-\sum_{1}^{n}\ln\left(1+x_i^{\beta}\right)=0$$

Therefore:

$$\hat{\epsilon}=\frac{n}{\sum_{i=1}^{n}\ln\left(1+x_i^{\hat{\beta}}\right)} \qquad (51)$$

In this relation, we have parameter $\hat{\beta}$; then we need the second equation, coming from the other partial derivative $\partial\ln L/\partial\beta=0$, and after substitution:

$$\frac{n}{\hat{\beta}}+\sum_{i=1}^{n}\ln x_i-\left[\frac{n}{\sum_{i=1}^{n}\ln\left(1+x_i^{\hat{\beta}}\right)}+1\right]\sum_{i=1}^{n}\frac{x_i^{\hat{\beta}}\ln x_i}{1+x_i^{\hat{\beta}}}=0 \qquad (52)$$

Equations (51) and (52) can be solved by numerical method.

29. Cumulative Burr functions

In [Hakim et al., 2021], we can find also the list of all the cumulative Burr functions.

The Burr $f(x)=k\,c\,x^{c-1}\left(1+x^{c}\right)^{-(k+1)}$ has the cumulative $F(x)=1-\left(1+x^{c}\right)^{-k}$.

Here the list [Hakim et al., 2021]:

I) $F(x)=x$, interval $(0,1)$

II) $F(x)=\left(1+e^{-x}\right)^{-k}$, interval $(-\infty,\infty)$

III) $F(x)=\left(1+x^{-c}\right)^{-k}$, interval $(0,\infty)$

IV) $F(x)=\{1+\left[x^{-1}(c-x)\right]^{1/c}\}^{-k}$, interval $(0,c)$

V) $F(x)=\left(1+c\,e^{-\tan(x)}\right)^{-k}$, interval $(-\pi/2,\pi/2)$

VI) $F(x)=\left(1+c\,e^{-r\sinh(x)}\right)^{-k}$, interval $(-\infty,\infty)$

VII) $F(x)=2^{-k}\left(1+\tanh(x)\right)^{k}$, interval $(-\infty,\infty)$

VIII) $F(x)=\left[2\pi^{-1}\arctan(e^{x})\right]^{-k}$, interval $(-\infty,\infty)$

IX) $F(x)=1-2\{2+c[(1+e^{x})^{k}-1]\}^{-1}$, interval $(-\infty,\infty)$

X) $F(x)=\left(1-e^{-x^{2}}\right)^{k}$, interval $(0,\infty)$

XI) $F(x)=\left[x-(2\pi)^{-1}\sin(2\pi x)\right]^{k}$, interval $(0,1)$

XII) $F(x)=1-\left(1-x^{c}\right)^{-k}$, interval $(0,\infty)$

In the list given above, x is dimensionless.

The reference tells that the “most important distribution in Burr system is Burr Type XII distribution which has two positive parameters *k* and c. Burr Type XII distribution is discussed in more detail by Burr and has gained special attention [Nasir et Al-Anber, 2012]". As given by the reference, the distribution can be used in the various fields of sciences, including reliability analysis [Lewis, 1981], [Zimmer et al., 1998], life testing [Gupta et al., 1996], survival analysis [Ghitany et Al-Awadhi, 2002], actuaries [Burnecki et al., 2004], economics [McDonald, 2008], forestry [Lindsay et al., 1996], hydrology [Shao et al., 2004], and meteorology [Usta et Kantar, 2012]. "Because of its popularity, Burr Type XII distribution is commonly known as Burr distribution".

30. The origin of Weibull distribution (Particle size distribution)

Let us continue to propose further information about background and application of the Weibull distribution. In [Jonasz et Fournier, 2007], we can find that it was involved in the particle size distribution.

Reference [Jonasz et Fournier, 2007] tells that the Weibull function was introduced in order to describe the size distribution of a particle population formed by fragmentation (crushing). The distribution was formulated by Tenchov and Yanev (1986), in the following form:

$$n(D) = n_0 \frac{c}{b} \left(\frac{D-D_0}{b} \right)^{c-1} \exp\left[-\left(\frac{D-D_0}{b} \right)^{c} \right] \quad (53)$$

In the formula (53), n_0 is the particle concentration (scale) factor with a dimension of number of particle per unit of volume. D_0 is the smallest particle diameter, related to the extent of fragmentation. In [Jonasz et Fournier, 2007], it is specified that, as in the case of the log-normal and modified gamma distributions, the Weibull distribution yields a finite total number of particles.

Tenchov and Yanev (1986) give also the value of the peak:

$$D_{peak} = D_0 + b \left(\frac{c-1}{c} \right)^{1/c} \quad (54)$$

and other quantities related to the distribution [Jonasz et Fournier, 2007].

Ref. [Jonasz et Fournier, 2007] explains that "the Weibull distribution is a result of a random fragmentation process where the probability of splitting a particle into fragments depends on the particle size. If that probability is independent of the particle size, the log-normal size distribution results. Both distributions can be made quite similar, and the experimental errors may prevent one form reaching a definite conclusion on which one is a better fit to experimental data", as shown by [Tenchov et Yanev, 1986].

Brown and Wohletz [Brown et Wohletz, 1995] have demonstrated that "the Weibull distribution arises naturally as a consequence of the fragmentation process being fractal" [Jonasz et Fournier, 2007]. "Fragmentation process is a cleavage of bonds between the component particles that can also be proven to lead to the Weibull distribution" [Tenchov et Yanev, 1986],[Jonasz et Fournier, 2007]. According to the derivation of Brown and Wohletz, exponent c in (41) is related to the three-dimensional fractal dimension d describing the fragmentation process as follows [Jonasz et Fournier, 2007]:

$$c-1=\frac{d}{3}$$

31. Rosin and Rammler distributions

In [Jonasz et Fournier, 2007], it is told that Brown and Wohletz, in [Brown et Wohletz, 1995], also show that the Weibull distribution is linked to the Rosin-Rammler distribution (Rosin and Rammler, 1993, [Rosin et Rammler, 1933]) that has been used, in particular, for describing the size distribution of fragments in coal processing and in geology. The Rosin-Rammler distribution is a cumulative distribution expressed as follows [Brown et Wohletz, 1995]:

$$N(D)=N_T \exp\left[-\left(\frac{D}{D_0}\right)^r\right] \quad (55)$$

$N(D)$ is the number of particles with sizes greater that D. N_T is the total number of particles, r is a dimensionless constant, and D_0 is related to the average particle diameter. Note that the simple fitting of this distribution to experimental data, to avoid using a direct non-linear fitting process, requires taking double logarithms which greatly reduces the effect the features of the data set may have on the values of the fit parameters [Jonasz et Fournier, 2007].

As historical background, in [Brown et Wohletz, 1995] we can find told that, in 1933 [Rosin et Rammler, 1933], Rosin and Rammler proposed the use of an empirical distribution for the description of particle sizes, which they obtained from data describing the crushing of coal and other materials. In 1939 [Weibull, 1939], Weibull proposed the same distribution, obtained from the study of the fracture of materials under repetitive stress. The description proposed was empirically based, until Austin et al. [Austin et al., 1972] derived it to describe batch grinding in 1972. Then, Peterson et al. (1985), Brown (1989) and Wohletz et al. (1989) independently rederived the distribution.

32. Size distribution of fly ash

In [Weibull, 1951], among the first proposed applications of the distribution, we can find Waloddi Weibull evaluating the distribution to the size of fly ash. The fly ash, also known as flue-ash or simply ash, is a fine powder consisting mostly of particles that are produced as a byproduct in coal-fired power stations. Let us observed that this complex anthropogenic material is a product which has pozzolanic properties; this means that it reacts with lime to form cementitious compounds [Paya et al., 2001]. In [Xu et Shi, 2018], due to a good performance, it can be involved in materials as alternative to ordinary Portland cement. In fact, the fly has has several other potential applications. In [Yao et al., 2015], we can find a review about the generation, physicochemical properties and hazards of coal fly ash and the applications, including use in the soil amelioration, ceramic industry, catalysis, depth separation, zeolite synthesis, etc. [Yao et al., 2015].

The fly ash is a coal combustion product, composed of the particulates that are driven out of combustion chambers with the flue gases. In modern coal-fired power plants, the particulate is generally captured by electrostatic precipitators or other particle filtration devices, before the flue gases reach the chimneys. The composition of fly ash is varying considerably, according to the burned coal; however all fly ash includes substantial amounts of silicon dioxide (SiO_2), aluminium oxide (Al_2O_3) and calcium oxide (CaO), which are the main mineral compounds in coal-bearing rock strata. Since the particulate is rapidly formed while suspended in the exhaust gases, the fly ash is made of particles, generally spherical in shape, ranging in size from 0.5 μm to 300 μm.

In [Sparavigna, 2017], it has been proposed a measurement of the particle size, by means of an image segmentation.

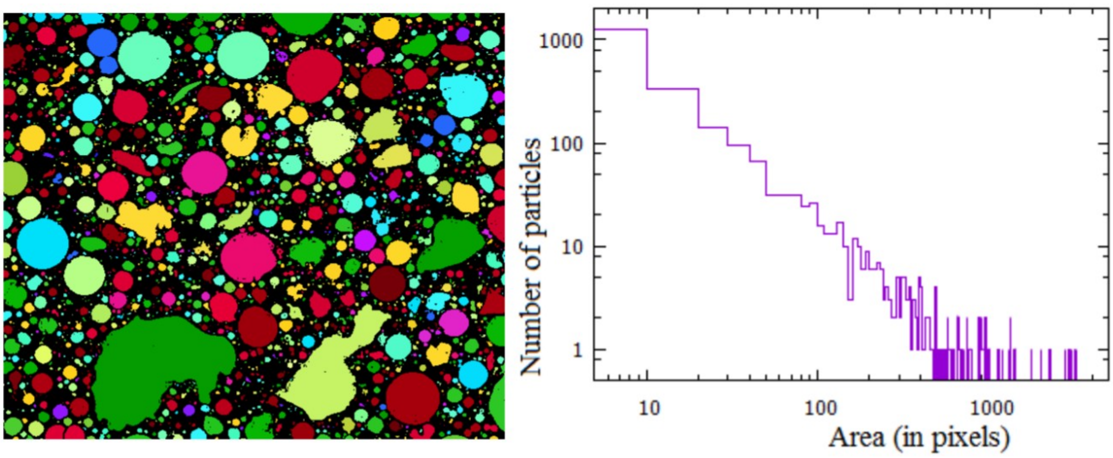


Fig. 26 – The result of the image segmentation of particles from Ref. [Sparavigna, 2017].

33. The statistical distribution of the strength of fibres

From Ref. [Pickering et Murray, 1999], we can read that "an accurate model for the strength of such a composite will rely on accurately characterising the strength of the fibres. The strength of a particular type of carbon fibre cannot be fully described by a single value. Carbon fibres are brittle and have strengths, as described by Griffith theory [Griffith, 1920], determined by the flaws which occur along the length of the fibres. … The strength of a particular fibre depends on the variation of flaws or more precisely, the worst flaw that exists along its length". The strengths of fibres are statistically distributed, and the statistical is usually described by means of the two-para#meter Weibull equation.

In the Eq. 3 of Ref. [Pickering et Murray, 1999], we find that the strengths for lengths L_1, L_2 are linked by the relation:

$$\sigma_2 / \sigma_1 = \left(L_1 / L_2 \right)^{1/w} \qquad (56)$$

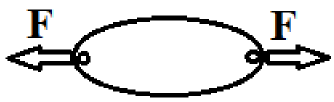
In this equation, w is the shape parameter of the Weibull model. The shape parameter is also defined as the Weibull modulus.

The equation given above contains the principle of 'weak-link scaling' (or 'weakest-link scaling' (WLS) in [Hristopulos et al., 2015]). A plot of the logarithm of characteristic strength versus the logarithm of length should give a straight line if weak link scaling is observed. From such a plot, the Weibull modulus can be obtained from the reciprocal of the gradient [Pickering et Murray, 1999].

34. The weak-link effect

In Section 2 we have introduced the κ-Weibull distribution and mentioned [Hristopulos et al., 2015] and, as previously told, the weakest-link scaling (WLS) principle.

Let us add that the "weakest link hypothesis" is an hypothesis which is implying that the strength of a structure depends on the most important defect (flaw) in the material under the application of a specific load [Chasiotis et Knauss, 2003]. In its simple form, the weak-link effect can be expressed in the following manner, as proposed in [Morton et Hearle, 2008]. "Suppose that we could determine the strength at every point along the length of a fibre. We should find that it varied from point to point, ... If a gradually increasing load is applied to this whole specimen, it will break at its weakest point, giving a strength S_1 , but if the specimen is tested in two half-lengths, each will break at its own weakest place, one giving the value S_1 , and the other a value S_2 , which is necessarily greater than S_1 . The mean strength, measured on half-lengths, is the mean of S_1 and S_2 , and must therefore be greater than the strength measured on the whole length". Going to quarter-lengths, we get four values. The mean strength from these values is greater still. “This increase will continue until at very short lengths the mean strength tends to the value S_0, which gives equal areas of the curve above and below the line $S=S_0$, since each small element will break at its own value of strength”.



35. The weak link in a chain

Let us consider the discussion in the Ref. [Gustafsson, 2014]. Let us start with a link in a chain, on which we apply the tensile force ***F.*** The Weibull model tells that the cumulative probability distribution function for the strength of the link is:

$$S(F) = \begin{cases} 0 & \text{if } F<0 \\ 1-e^{-(F/F_o)^m} & \text{if } F\geq 0 \end{cases} \quad (57)$$

$S(F)$ is the probability that the link will fail if loaded from zero up to the load F. In the expression given above, F_o and m are parameters which define the properties of the link. The Weibull weakest link model is used for analyses of the strength of structural elements, that is the strength of a chain made up of several links.

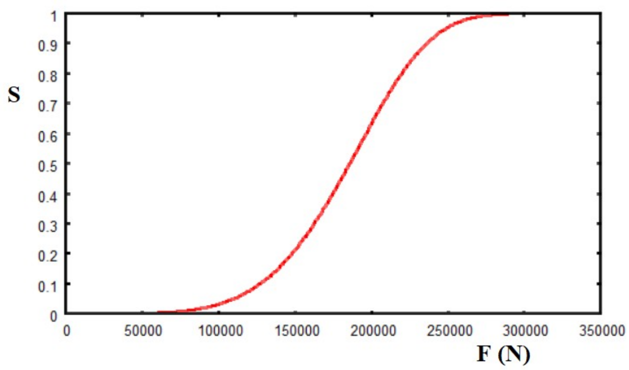


Fig. 27 - $S(F)$ for $m=5$, $F_o=200\,kN$.

From this plot, we can find that a failure probability, for instance equal to $S=0.21$, is corresponding to $F=150\,kN$. It means that the 21 % of links have a strength less than or equal to $F=150\,kN$ [Gustafsson, 2014]. We can calculate the corresponding distribution as dS/dF .

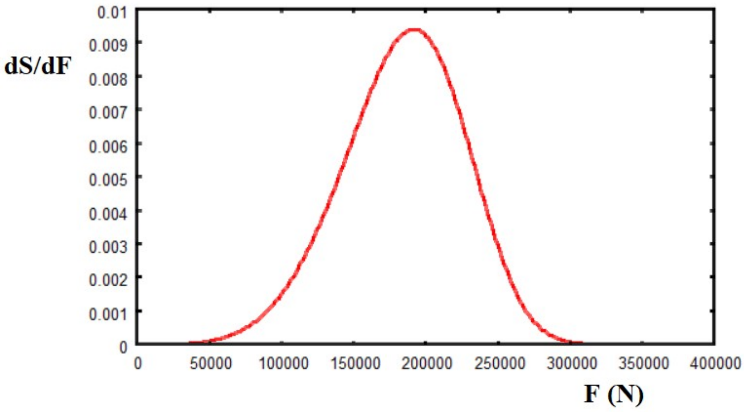


Fig. 28 - $dS(F)/dF$ for $m=5$, $F_o=200\,kN$.

The plot shows a failure probability density $dS/dF=0.0061\ kN^{-1}$ for $F=150\,kN$. The 0.61% of the links fail for each $1\,kN$ increase of the load, when the load is close to $150\,kN$ [Gustafsson, 2014]. Note that the same probability density, 0.61%, is also observed for $225\,kN$.

Let us suppose to consider n links. The probability that a link can carry F is: $e^{-(F/F_o)^m}$. The probability that, in the chain with n links, the $i-$link can carry a load F_i is [Gustafsson, 2014]:

$$e^{-(F_1/F_o)^m}\,e^{-(F_2/F_o)^m}\cdots e^{-(F_n/F_o)^m}=e^{-\sum_{i=1}^{n}(F_i/F_o)^m} \qquad (58)$$

In the case that $F_i=F$, the failure probability is:

$$1-e^{-\sum_{i=1}^{n}(F_i/F_o)^m}=1-e^{-n(F/F_o)^m}=1-e^{-(F/\tilde{F}_o)^m} \qquad (59)$$

where $\tilde{F}_o=F_o/n^{1/m}$. In the case of proportional loading, $F_i=F_M\lambda_i$, we find:

$$1-e^{-\sum_{i=1}^{n}(F_M\lambda_i/F_o)^m}=1-e^{-n(F_M/F_o)^m\sum_{i=1}^{n}\lambda_i^m}=1-e^{-(F_M/\tilde{F}_o)^m} \qquad (60)$$

where $\widetilde{F}_o = F_o / \left(\sum_{i=1}^{m} \lambda_i^m \right)^{1/m}$. The distributions for one link and for a chain with n links differ for a scaling factor [Gustafsson, 2014]. Lecture [Gustafsson, 2014] continues with bulk materials.

36. Weibull distribution for pseudorandom numbers

Besides the reliability analysis, the evaluation of the size distribution and the strength of fibres, the Weibull function can be used for numerical calculus. Routine RNWIB can be used to generate pseudorandom numbers, starting from a Weibull distribution, with shape parameter A and unit scale parameter, so that:

$$f(x) = A x^{A-1} e^{-x^A} \quad , \quad x \geq 0 \qquad (61)$$

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

Appendix – κ-exponential

Giorgio Kaniadakis, Politecnico di Torino, derived the κ-exponential from the energy–momentum relation, determined by special relativity. The energy–momentum relation is the relativistic dispersion relation, relating total energy (relativistic energy) to invariant mass (the rest mass) and momentum. It can be written as the following equation:

$$E^2 = (pc)^2 + (m_o c^2)^2$$

In this equation, E is the total energy, m_o the invariant mass, p the momentum and constant c the speed of light. The dispersion relation assumes the special relativity case of flat spacetime.

Let us write the relation in the following form:

$$E^2 - p^2 c^2 = m_o^2 c^4 \text{ , then } \frac{E^2}{m_o^2 c^4} - \frac{p^2 c^2}{m_o^2 c^4} = 1$$

$$\left(\frac{E}{m_o c^2}-\frac{p\,c}{m_o c^2}\right)\left(\frac{E}{m_o c^2}+\frac{p\,c}{m_o c^2}\right)=1$$

Being $E=\sqrt{p^2 c^2+m_o^2 c^4}$, we can write $E=m_o c^2\sqrt{1+\left(\frac{p}{m_o c}\right)^2}$.

$$\left(\sqrt{1+\left(\frac{p}{m_o c}\right)^2}-\frac{p}{m_o c}\right)\left(\sqrt{1+\left(\frac{p}{m_o c}\right)^2}+\frac{p}{m_o c}\right) = 1$$

Let us write: $\kappa\, q=\frac{p}{m_o c}$, the previous relation becomes:

$$\left(\sqrt{1+\kappa^2 q^2}-\kappa\, q\right)\left(\sqrt{1+\kappa^2 q^2}+\kappa\, q\right) = 1$$

$$\left(\sqrt{1+\kappa^2 q^2}-\kappa\, q\right)^{1/\kappa}\ \left(\sqrt{1+\kappa^2 q^2}+\kappa\, q\right)^{1/\kappa} = 1$$

This is the constituent equation of the κ-exponential:

$$\exp_\kappa(-q)\ \exp_\kappa(q) = 1$$

References

1. Almalki, S. J., & Nadarajah, S. (2014). Modifications of the Weibull distribution: A review. Reliability Engineering & System Safety, 124, 32-55.
2. Austin, L. G., Luckie, P. T., & Klimpel, R. R. (1972). Solutions of the batch grinding equation leading to Rosin-Rammler distributions. Trans. AIME, 252(1), 87-94.
3. Berrettoni, J. N. (1964). Practical applications of the Weibull distribution. Industrial Quality Control (August), 71-78.
4. Brown, W. K. (1989). A theory of sequential fragmentation and its astronomical

applications. Journal of Astrophysics and Astronomy, 10(1), 89-112.

5. Brown, W. K., & Wohletz, K. H. (1995). Derivation of the Weibull distribution based on physical principles and its connection to the Rosin–Rammler and lognormal distributions. Journal of Applied Physics, 78(4), 2758-2763.
6. Burnecki, K., Härdle, W., & Weron, R. (2004). An introduction to simulation of risk processes. Encyclopedia of Actuarial Science', Wiley, Chichester, 1564-1570.
7. Burr, I. W. (1942). Cumulative frequency functions. Annals of Mathematical Statistics. 13 (2): 215–232. doi:10.1214/aoms/1177731607. JSTOR 2235756.
8. Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. p. 584. ISBN 978-0-521-84805-3.
9. Caroni, C. (2010). Testing for the Marshall–Olkin extended form of the Weibull distribution. Statistical Papers, 51(2), 325-336.
10. Çakallı, Hüseyin (2015). Upward and Downward Statistical Continuities. Filomat. 29 (10): 2265–2273. doi:10.2298/FIL1510265C. JSTOR 24898386. S2CID 58907979.
11. Chasiotis, I. & Knauss, W. G. (2003). Experimentation at the Micron and Submicron Scale, in Comprehensive Structural Integrity, Editors: I. Milne, R.O. Ritchie, B. Karihaloo, Pergamon, Pages 41-87, ISBN 9780080437491, https://doi.org/10.1016/B0-08-043749-4/08038-1.
12. Clementi, F., Di Matteo, T., Gallegati, M., & Kaniadakis, G. (2008). The κ-generalized distribution: A new descriptive model for the size distribution of incomes. Physica A: Statistical Mechanics and its Applications, 387(13), 3201-3208.
13. Collett, D. (2015). Modelling survival data in medical research (3rd ed.). Boca Raton: Chapman and Hall / CRC. ISBN 978-1439856789.
14. Evans, J., Kretschmann, D., & Green, D. (2019). Procedures for estimation of Weibull parameters. Gen. Tech. Rep. FPL-GTR-264. Madison, WI: US Department of Agriculture, Forest Service, Forest Products Laboratory. 17 p., 264, 1-17.
15. Ghitany, M. E., & Al-Awadhi, S. (2002). Maximum likelihood estimation of Burr XII distribution parameters under random censoring. Journal of Applied Statistics, 29(7), 955-965.
16. Greenwich, M. (1992). A unimodal hazard rate function and its failure distribution. Statistical Papers 33, 187–202. https://doi.org/10.1007/BF02925324
17. Griffith, A. A. (1920). The phenomenon of rupture and flow in solids. Phil. Trans. Royal Soc. London, A, 221, 163-198.
18. Gumbel, E. J. (1935), Les valeurs extrêmes des distributions statistiques.

Annales de l'Institut Henri Poincaré, 5 (2): 115–158

19. Gumbel E. J. (1941). The return period of flood flows. The Annals of Mathematical Statistics, 12, 163–190.
20. Gupta, P. L., Gupta, R. C., & Lvin, S. J. (1996). Analysis of failure time data by Burr distribution. Communications in Statistics -Theory and Methods, 25(9), 2013-2024.
21. Gustafsson, Per Johan (2014). Lecture Notes on some probabilistic strength calculation models. Report TVSM-7161, Division of Structural Mechanics, Lund University, 2014.
22. Hakim, A. R., Fithriani, I., & Novita, M. (2021). Properties of Burr distribution and its application to heavy-tailed survival time data. In Journal of Physics: Conference Series (Vol. 1725, No. 1, p. 012016). IOP Publishing.
23. Hodges, L. (1994). Common Univariate Distributions, in Methods in Experimental Physics, J. L. Stanford and S. B. Vardeman, Editors, Academic Press, Volume 28, Pages 35-61. ISSN 0076-695X, ISBN 9780124759732, https://doi.org/10.1016/S0076-695X(08)60252-5.
24. Hristopulos, D. T., Petrakis, M. P., & Kaniadakis, G. (2015). Weakest-link scaling and extreme events in finite-sized systems. Entropy, 17(3), 1103-1122.
25. Jonasz, M., R. Fournier, G. R. (2007). Chapter 5 - The particle size distribution, In Light Scattering by Particles in Water, Editors Miroslaw Jonasz and Georges R. Fournier, Academic Press, Pages 267-445, ISBN 9780123887511, https://doi.org/10.1016/B978-012388751-1/50005-3.
26. Kaniadakis, G. (2001). Non-linear kinetics underlying generalized statistics. Physica A: Statistical mechanics and its applications, 296(3-4), 405-425.
27. Kaniadakis, G. (2002). Statistical mechanics in the context of special relativity. Physical review E, 66(5), 056125.
28. Kaniadakis G. (2013). Theoretical Foundations and Mathematical Formalism of the Power-Law Tailed Statistical Distributions. Entropy, Volume 15, Issue 10. Pages 3983-4010. https://doi.org/10.3390/e15103983
29. 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
30. Kaniadakis, G. (2021). New power-law tailed distributions emerging in κ-statistics (a). Europhysics Letters, 133(1), 10002.
31. Kartsonaki, C. (2016). Survival analysis, Diagnostic Histopathology, Volume 22, Issue 7, Pages 263-270, ISSN 1756-2317, DOI https://doi.org/10.1016/j.mpdhp.2016.06.005.

32. Kızılersü, A., Kreer, M. and Thomas, A.W. (2018), The Weibull distribution. Significance, 15, 10-11. https://doi.org/10.1111/j.1740-9713.2018.01123.x

33. Kundu, D., & Gupta, R. D. (1999). Generalized exponential distribution. Aust NZJ Stat, 41, 173-188.

34. Lacey, D., & Nguyen, A. (2015). Bathtub and Unimodal Hazard Flexibility Classification of Parametric Lifetime Distributions. Rose-Hulman Undergraduate Mathematics Journal, 16(2), 6.

35. Lewis A. W. (1981)- The Burr Distribution as a General Parametric Family in Survivorship and Reliability Theory Applications PhD Dissertation (Chapel Hill: Department of Biostatistics, University of North Carolina at Chapel Hill)

36. Lienig, J., and Bruemmer, H. (2017). Fundamentals of Electronic Systems Design. Springer International Publishing. p. 54. doi:10.1007/978-3-319-55840-0. ISBN 978-3-319-55839-4.

37. Lindsay, S. R., Wood, G. R., & Woollons, R. C. (1996). Modelling the diameter distribution of forest stands using the Burr distribution. Journal of Applied Statistics, 23(6), 609-620.

38. 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.

39. McDonald, J. B. (2008). Some generalized functions for the size distribution of income. In Modeling income distributions and Lorenz curves (pp. 37-55). Springer, New York, NY.

40. Morton, W. E., & Hearle, J. W. S. (2008). The effects of variability. In Physical Properties of Textile Fibres (Fourth Edition). Woodhead Publishing Series in Textiles. ISBN 978-1-84569-220-9

41. Mudholkar, G.S., & Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure-ratedata. IEEE Transactions on Reliability. 42 (2): 299–302. doi:10.1109/24.229504.

42. Nadarajah, S. (2006). The exponentiated Gumbel distribution with climate application. Environmetrics: The official journal of the International Environmetrics Society, 17(1), 13-23.

43. Nasir, S. A., & Al-Anber, N. J. (2012). A Comparison of the Bayesian and other Methods for Estimation of Reliability Function for Burr-xii Distribution. Journal of Mathematics and statistics, 8(1), 42-48.

44. 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

45. Paya, J., Borrachero, M. V., Monzo, J., Peris-Mora, E., & Amahjour, F. (2001). Enhanced conductivity measurement techniques for evaluation of fly ash pozzolanic activity. Cement and Concrete Research, 31(1), 41-49.
46. Peterson, T. W., Scotto, M. V., & Sarofim, A. F. (1985). Comparison of comminution data with analytical solutions of the fragmentation equation. Powder technology, 45(1), 87-93.
47. Pickering, K. L., & Murray, T. L. (1999). Weak link scaling analysis of high-strength carbon fibre. Composites Part A: Applied Science and Manufacturing, 30(8), 1017-1021.
48. Picoli, S. Jr., Mendes, R. S., & Malacarne, L. C. (2003). q-exponential, Weibull, and q-Weibull distributions: an empirical analysis. Physica A: Statistical Mechanics and Its Applications. 324 (3): 678–688. arXiv:cond-mat/0301552.
49. Preda, V. (1985). The generalized gamma distribution and the principle of maximum entropy. Proceedings of the Seventh Conference on Probability Theory: August 29–September 4, 1982, Brasov, Romania, edited by , Berlin, Boston: De Gruyter, 2020, pp. 569-572. Available De Gruyter https://doi.org/10.1515/9783112314036-072
50. Razali, A. M., & Salih, A. A. (2009). Combining two Weibull distributions using a mixing parameter. European Journal of Scientific Research, 31(2), 296-305.
51. Rinne. H. (2008). Related distributions from: The Weibull Distribution, A Handbook CRC Press, Accessed on: 13 Apr 2022, Available at https://www.routledgehandbooks.com/doi/10.1201/9781420087444.ch3
52. Rosin, P., & Rammler, E. (1933). Regularities in the distribution of cement particles. J Inst Fuel, 7, 29-33.
53. Rosin, P., & Rammler, E. (1933). Application of Rosin's distribution in size-frequency analysis of clastic rocks. Journal of Sedimentary Research, 34(3), 483-502.
54. Saei, S., Mohammadi, M., Fekriseri, M., & Jenab, K. (2019). A computational method for estimating Burr XII parameters with complete and multiple censored data. arXiv preprint arXiv:1901.09299. DOI 10.48550/arXiv.1901.09299
55. Shao, Q., Wong, H., Xia, J., & Ip, W. C. (2004). Models for extremes using the extended three-parameter Burr XII system with application to flood frequency analysis/Modèles d'extrêmes utilisant le système Burr XII étendu à trois paramètres et application à l'analyse fréquentielle des crues. Hydrological Sciences Journal, 49(4).
56. Singh, S., & Maddala, G. (1976). A Function for the Size Distribution of Incomes. Econometrica. 44 (5): 963–970. doi:10.2307/1911538. JSTOR 1911538.

57. Sparavigna, A. C. (2017). Measuring the particles in fly ash by means of an image segmentation. Philica, 2017(1105).

58. 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

59. Tenchov, B. G., & Yanev, T. K. (1986). Weibull distribution of particle sizes obtained by uniform random fragmentation. Journal of colloid and interface science, 111(1), 1-7.

60. Usta, I., & Kantar, Y. M. (2012). Analysis of some flexible families of distributions for estimation of wind speed distributions. Applied Energy, 89(1), 355-367.

61. Weibull, W. (1939). A statistical theory of strength of materials. The Royal Swedish Institute of Engineering Research (Ingenors Vetenskaps Akadiens Handlingar), Proc. No. 151, Stockholm.

62. Weibull, Waloddi (1951). A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, American Society of Mechanical Engineers, 1951. Volume 18, Pages 293-297. https://hal.archives-ouvertes.fr/hal-03112318

63. Wohletz, K. H., Sheridan, M. F., & Brown, W. K. (1989). Particle size distributions and the sequential fragmentation/transport theory applied to volcanic ash. Journal of Geophysical Research: Solid Earth, 94(B11), 15703-15721.

64. Xu, G., & Shi, X. (2018). Characteristics and applications of fly ash as a sustainable construction material: A state-of-the-art review. Resources, Conservation and Recycling, 136, 95-109.

65. Yang, X. S. (2017). Engineering mathematics with examples and applications. Academic Press. ISBN 9780128099025 012809902X

66. Yang, F., Ren, H., & Hu, Z. (2019). Maximum likelihood estimation for three-parameter Weibull distribution using evolutionary strategy. Mathematical Problems in Engineering, 2019.

67. Yao, Z. T., Ji, X. S., Sarker, P. K., Tang, J. H., Ge, L. Q., Xia, M. S., & Xi, Y. Q. (2015). A comprehensive review on the applications of coal fly ash. Earth-Science Reviews, 141, 105-121.

68. Zacks, S. (1984). Estimating the Shift to Wear-Out of Systems Having Exponential-Weibull Life Distributions. Operations Research, Vol. 32, No. 3, Reliability and Maintainability (May - Jun., 1984), pp. 741-749

69. Zimmer, W. J., Keats, J. B., & Wang, F. K. (1998). The Burr XII distribution in reliability analysis. Journal of quality technology, 30(4), 386-394.