Hacettepe Journal of Mathematics and Statistics
Volume 45 (3) (2016), 851 – 868

The (P-A-L) extended Weibull distribution: A new generalization of the Weibull distribution

Bander Al-Zahrani[*], Pedro Rafael Diniz Marinho[†‡], A. A. Fattah[§], A-Hadi N. Ahmed [¶] and Gauss M. Cordeiro [‖]

Abstract

Recently, some attempts have been made to construct new families of models to extend well-known distributions and at the same time provide great flexibility in modeling data in practice. So, several classes by adding shape parameters to generate new models have been explored in the statistical literature. We propose a new generalization of the three-parameter extended Weibull distribution pioneered by Pappas et al. (2012) by using the generator by Marshall and Olkin (1997). The new model is called the (P-A-L) extended Weibull, where (P-A-L) denote the first letters of the scientists Pappas, Adamidis and Loukas.

Keywords: Extended Weibull distribution, Hazard rate function, Maximum likelihood estimation, Moment, Order statistic, Weibull distribution.

2000 AMS Classification: 62E15; 62E10; 62N99.

Received : 23.12.2014 *Accepted* : 02.04.2015 *Doi* : 10.15672/HJMS.20156110040

[*]Department of Statistics, King Abdulaziz University, Saudi Arabia.
Email: bmalzahrani@kau.edu.sa
[†]Department of Statistics, Federal University of Permanbuco, Recife, Brazil.
Email: pedro.rafael.marinho@gmail.com
[‡]Corresponding Author.
[§]Institute of Statistical Studies & Research, Cairo University, Egypt.
Email: ahmedzi1987@gmail.com
[¶]Institute of Statistical Studies & Research, Cairo University, Egypt.
Email: dr.hadi@cu.edu.eg
[‖]Department of Statistics, Federal University of Permanbuco, Recife, Brazil.
Email: gauss@de.ufpe.br

1. Introduction

For more than half a century the Weibull distribution has attracted the attention of statisticians working on theory and methods in various fields of applied researchers. Thousands of papers have been written on this distribution. It is of most interest to the theory because of its great number of special features and to practitioners because of its ability to fit to real data from various fields, ranging from life data to weather data or observations made in economics and business administration, hydrology, biology and engineering sciences. When modeling monotone hazard rates, the Weibull distribution may be an initial choice because of its negatively and positively skewed density shapes. However, this distribution does not provide a reasonable parametric fit for some practical applications where the underlying hazard rates may be bathtub or unimodal shapes. [13] introduced a new generalization of any distribution, which is derived by using the generator by [10]. In the literature, several generalizations of the Weibull distribution have been proposed such as those studied by [3], [19], [14] and [20].

The *extended Weibull* (EW) distribution with parameters $\alpha > 0$, $\beta > 0$ and $\nu > 0$ has probability density function (pdf) given by

$$g(t) = \frac{\frac{\nu\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{\left[1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]^{2}}. \tag{1.1}$$

The reliability function corresponding to (1.1) becomes

$$\overline{G}(t) = \frac{\nu\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}. \tag{1.2}$$

Let $G(t)$ be a baseline cumulative distribution function (cdf) with corresponding survival function $\overline{G}(t) = 1 - G(t)$, pdf $g(t) = dG(t)/dt$ and hazard rate function (hrf) $\lambda(t)$. [13] proposed the *(P-A-L) extended family* with the additional parameter $p > 1$, where the survival function $\overline{F}(t)$, cdf $F(t)$ and pdf $f(t)$ are given by (for $t > 0$)

$$\overline{F}(t) = \frac{\log\left[1-(1-p)\,\overline{G}(t)\right]}{\log(p)}, \tag{1.3}$$

$$F(t) = 1 - \frac{\log\left[1-(1-p)\,\overline{G}(t)\right]}{\log(p)} \tag{1.4}$$

and

$$f(t) = \frac{(p-1)\;g(t)}{\left[1-(1-p)\,\overline{G}(t)\right]\;\log(p)},$$

respectively.

Further, [13] studied the *(P-A-L) extended modified Weibull* distribution. In this paper, we take the EW distribution given by (1.1) as the baseline model to define a new four-parameter *(P-A-L) extended Weibull*, say the (P-A-L)EW distribution.

The rest of the paper is organized as follows. In Section 2, we provide the pdf and cdf of the new distribution and present some special models. In Section 3, we study some of its structural properties including moments, moment generating function (mgf), quantile and residual life functions. The mean deviations and two types of entropies are determined in Sections 4 and 5, respectively. Section 6 is devoted to the reliability function. In Section 7, we present the reliability function, hrf, cumulative hazard rate function (chrf) and mean residual lifetime function (mrlf). The order statistics and the minimum and maximum order statistics are investigated in Section 8. In Section 9, we obtain the maximum likelihood estimates (MLEs) of the model parameters. In Section 10, we apply a particle swarm optimization (PSO) method to estimate the parameters. In Section 11, we provide one application to real data in order to illustrate the potentiality of the new model. Concluding remarks are addressed in Section 12.

2. The (P-A-L) Extended Weibull Distribution

Combining (1.2) and (1.4), the cdf of the (P-A-L)EW distribution follows as

$$F(t) = 1 - \frac{1}{\log(p)} \log\left\{ \frac{1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}} \right\}. \tag{2.1}$$

By differentiating (2.1), the corresponding pdf reduces to

$$f(t) = \frac{1}{\log(p)} \left\{ \frac{(p-1)\,\frac{\nu\beta}{\alpha}\,\left(\frac{t}{\alpha}\right)^{\beta-1}\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{\left[1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]\left[1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]} \right\}. \tag{2.2}$$

Henceforth, we denote by $T\sim$(P-A-L)EW(α, β, ν, p) a random variable having pdf (2.2). It is clear that the new distribution is very flexible (as it can seen from Table 1). In fact, several distributions can be obtained as special cases of the new model for selected parameter values. These special cases include at least eleven distributions displayed in Figure 1: the (P-A-L) extended Rayleigh (P-A-L)ER, (P-A-L) extended Exponential (P-A-L)EE, (P-A-L) Weibull (P-A-L)W, (P-A-L) Rayleigh (P-A-L)R, (P-A-L) exponential (P-A-L)E, extended Weibull (EW) (Marshall and Olkin, 1997), extended Rayleigh (ER), extended exponential (EW), Weibull (W), Rayleigh (R) and exponential (E) distributions.

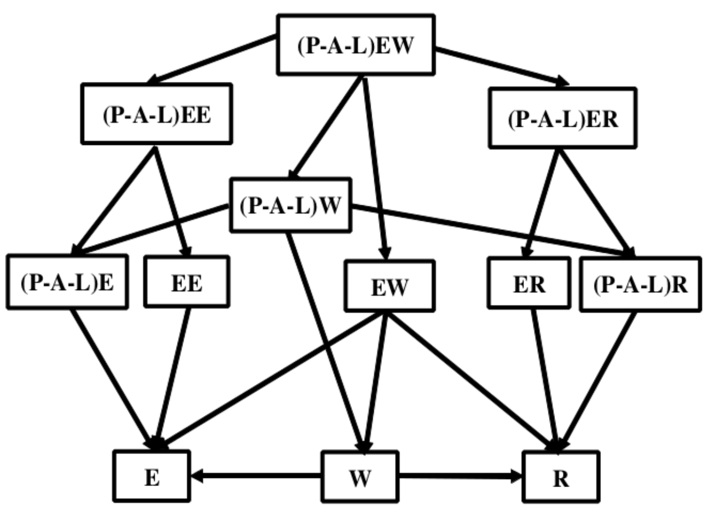
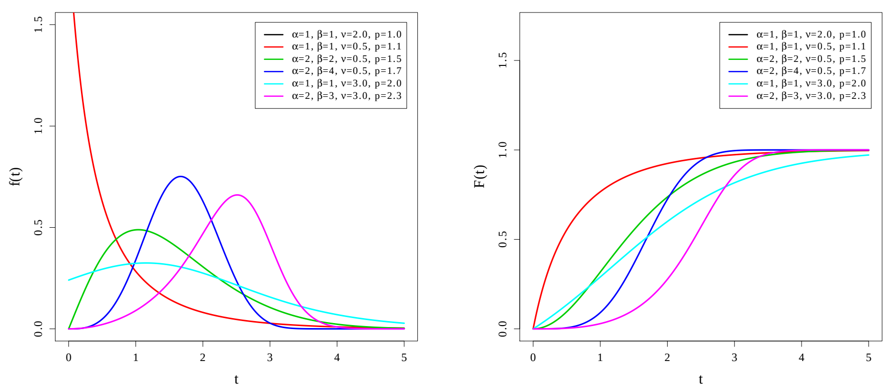


Figure 1. Sub-models of the (P-A-L)EW distribution.

Figures 2(a) and 2(b) display some of the possible shapes of the pdf and cdf of the new distribution, respectively, for different values of the parameters α, β, $\nu > 0$ and $p > 1$.

Table 1. Special cases of the (P-A-L)EW distribution.

Sub-Models	Parameters of (P-A-L)EWD				Cumulative distribution function
	α	β	ν	p	
(P-A-L)ER	-	2	-	-	$1-\frac{1}{\log p}\log\left\{\frac{1-(1-p\,\nu)\,e^{-\left(\frac{t}{\alpha}\right)^2}}{1-(1-\nu)\,e^{-\left(\frac{t}{\alpha}\right)^2}}\right\}$
(P-A-L)EE	-	1	-	-	$1-\frac{1}{\log p}\log\left\{\frac{1-(1-p\,\nu)\,e^{-\left(\frac{t}{\alpha}\right)}}{1-(1-\nu)\,e^{-\left(\frac{t}{\alpha}\right)}}\right\}$
(P-A-L)W	-	-	1	-	$1-\frac{1}{\log p}\log\left\{1-(1-p)\,e^{-\left(\frac{t}{\alpha}\right)^\beta}\right\}$
(P-A-L)R	-	2	1	-	$1-\frac{1}{\log p}\log\left\{1-(1-p)\,e^{-\left(\frac{t}{\alpha}\right)^2}\right\}$
(P-A-L)E	-	1	1	-	$1-\frac{1}{\log p}\log\left\{1-(1-p)\,e^{-\left(\frac{t}{\alpha}\right)}\right\}$
W	-	-	1	$p\to 1$	$F(t)=1-e^{-\left(\frac{t}{\alpha}\right)^\beta}$
R	-	2	1	$p\to 1$	$F(t)=1-e^{-\left(\frac{t}{\alpha}\right)^2}$
ED	-	1	1	$p\to 1$	$F(t)=1-e^{-\left(\frac{t}{\alpha}\right)}$



(a) (b)

Figure 2. The (P-A-L)EW for (a) densities and for (b) distributions.

3. Mathematical Properties

In this section, we derive some mathematical properties of the (P-A-L)EW distribution such as the quantile, median, random number generator, central and non-central moments and mgf.

3.1. Quantile Function. The quantile function (qf) is used to obtain the quantiles of a probability distribution. Consider $F_X : \mathbb{R} \to [0,1]$ a distribution function of the continuous random variable X. The pth quantile of $F(x)$ is given by the value of x such that

$$Q(u) = \inf\{\mathrm{x} \in \mathbb{R} : \mathrm{u} \le \mathrm{F(x)}\},$$

where $u \in (0,1)$. The qf $Q(u) = F^{-1}(u)$ of T comes by inverting (2.1) as

$$t = Q(u) = \left[\alpha^\beta \log\left(-\frac{\nu p^{u+1} - p^u - \nu p + p}{p^u - p}\right)\right]^{1/\beta}. \tag{3.1}$$

3.2. Central and Non-Central Moments. The rth non-central moment of T can be expressed as

$$E(T^r) = \mu_r = \frac{\alpha^r \nu(p-1)}{\log(p)} \Gamma\left(\frac{r}{\beta}+1\right) \sum_{i,j=0}^{\infty} \frac{(1-p\nu)^i \, (1-\nu)^j}{(1+i+j)^{\frac{r}{\beta}+1}}. \tag{3.2}$$

The nth central moment of T, say m_n, can be easily obtained from the non-central moments by (for $n \geq 1$)

$$m_n = E(T-\mu)^n = \sum_{r=0}^{n} \binom{n}{r} (-\mu)^{n-r} \, E(T^r).$$

Let $\alpha = 1.5$, $\beta = 1.3$, $\nu = 1.2$ and $p = 1.2$. We can easily check that equation (3.2) holds. The following script written in the `Julia` language implements equation (3.2) with $r = 2$. The `Julia` language can be obtained from `http://julialang.org/downloads/` (see [4]). So, we provide a numerical check for $i = 0, \ldots, 5000$ and $j = 0, \ldots, 5000$. The code follows below.

```
alpha = 1.5
beta = 1.3
nu = 1.2
p = 1.2
I = 5000
J = 5000
r=2 # Moment of order 2.
constant = alpha^r*nu*(p-1)/log(p)*gamma(r/beta+1)
sum_I = zeros(Float64,I+1,1)
sum_J = zeros(Float64,J+1,1)
for i = 0:I
    for j = 0:J
        numerator = (1-p*nu)^i * (1-nu)^j
        denominator = (1+i+j)^(r/beta+1)
        sum_J[j+1] = numerador/denominador
   end
   sum_I[i+1] = sum(sum_J)
end
constant*sum(sum_I) # The result is 3.66655262332183.
```

Thus, for the fixed parameters, $r = 2$ and using the above code, we obtain $E(T^2) = 3.6665526$. The same result follows by numerical integration of (2.2). Established algebraic expansions to determine the moments of T can be more efficient than computing these moments directly by using this numerical integration, which can be prone to rounding off errors among others.

3.3. Residual life function. Given that a component survives up to time $x \geq 0$, the residual life is the period beyond x until the time of failure and it is defined by the expectation of the conditional random variable $T|T > x$. In reliability, it is well-known that the mrlf and the ratio of two consecutive moments of residual life determine the distribution uniquely (see [9]). Therefore, we obtain the rth order moment of the residual life by

$$m_r(x) = E[(T-x)^r \mid T > x] = \frac{1}{\overline{F}(x)} \int_x^{\infty} (t-x)^r \, f(t) dt. \tag{3.3}$$

Applying the binomial expansion for $(T-x)^r$ and substituting $\overline{F}(x)$ given by (1.3) into equation (3.3), the rth moment of the residual life of T is

$$m_r(x) = \frac{\nu(p-1)}{\overline{F}(x)\,\log(p)} \sum_{i=0}^{r} \sum_{j,k=0}^{\infty} \binom{r}{i} \frac{(1-p\nu)^j\,(1-\nu)^k \alpha^i\,(-x)^{r-i}}{(j+k+1)^{\frac{i}{\beta}+1}} \times\Gamma\left(\frac{i}{\beta}+1, (j+k+1)\left(\frac{x}{\alpha}\right)^{\beta}\right), \tag{3.4}$$

where $\Gamma(a,y)=\int_y^{\infty} x^{a-1}\,\mathrm{e}^{-\mathrm{x}}\mathrm{dx}$ is the upper incomplete gamma function.

Another important characteristic for T is the mrlf obtained by setting $r=1$ in equation (3.4). It represents the mean lifetime left for an item of age x. Whereas the hrf at x provides information on a random variable T about a small interval after x, the mrlf at x considers information about the whole remaining interval (x,∞).

We obtain the mrlf of T as

$$m(x) = -x + \frac{\nu\alpha(p-1)}{\overline{F}(x)\,\log(p)} \sum_{j,k=0}^{\infty} \frac{(1-p\nu)^j\,(1-\nu)^k}{(j+k+1)^{\frac{1}{\beta}+1}} \times \Gamma\left(\frac{1}{\beta}+1, (j+k+1)\left(\frac{x}{\alpha}\right)^{\beta}\right).$$

3.4. Reversed residual life function. The waiting time since failure is the waiting time elapsed since the failure of an item on condition that this failure had occurred in $[0,x]$. The rth order moment of the reversed residual life function (rrlf) is given by $M_r(x)=E[(x-T)^r \mid T<x]=\frac{1}{F(x)}\int_0^x (x-t)^r\,f(t)dt$.

Following similar algebra as before, we obtain

$$M_r(x) = \frac{\nu(p-1)}{F(x)\,\log(p)} \sum_{i=0}^{r} \sum_{j,k=0}^{\infty} \frac{(-1)^i\,\alpha^i\,\binom{r}{i}\,(1-p\nu)^j\,(1-\nu)^k\,x^{r-i}}{(j+k+1)^{\frac{i}{\beta}+1}} \times\gamma\left(\frac{i}{\beta}+1, (j+k+1)\left(\frac{x}{\alpha}\right)^{\beta}\right),$$

where $\gamma(a,y)=\int_0^{y} x^{a-1}\,\mathrm{e}^{-\mathrm{x}}\mathrm{dx}$ is the lower incomplete gamma function.

Then, the mean reversed residual life of T becomes

$$M(x) = x - \frac{\nu\alpha(p-1)}{F(x)\log(p)} \sum_{j,k=0}^{\infty} \frac{(1-p\nu)^j\,(1-\nu)^k}{(j+k+1)^{\frac{i}{\beta}+1}} \times \gamma\left(\frac{i}{\beta}+1, (j+k+1)\left[\frac{x}{\alpha}\right]^{\beta}\right), \tag{3.5}$$

where $M(x)$ represents the mean time elapsed since the failure of T given that it fails at or before x.

4. Mean Deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. These are known as the mean deviations about the mean and the median – defined by $D_1(T)=\int_0^{\infty}|t-\mu|\,f(t)dt$ and $D_2(T)=\int_0^{\infty}|t-M|\,f(t)dt$, respectively, where $\mu=E(T)$ is the mean and $M=Q(0.5)$ is the median given by (3.1).

The measures $D_1(T)$ and $D_2(T)$ can be expressed as $D_1(T) = 2\mu F(\mu) - 2Z(\mu)$ and $D_2(T) = \mu - 2Z(M)$, where $Z(x) = \int_0^x t\, f(t)dt$ is the first incomplete mean of T. This integral can be determined from (2.2) by

$$Z(x) = \frac{\nu\alpha(p-1)}{\log(p)} \sum_{i,j=0}^{\infty} \frac{(1-p\nu)^i(1-\nu)^j}{(i+j+1)^{\frac{1}{\beta}+1}} \Gamma\left(\frac{1}{\beta}+1, (i+j+1)\left(\frac{x}{\alpha}\right)^{\beta}\right). \tag{4.1}$$

Thus, the mean deviations $D_1(T)$ and $D_2(T)$ can be obtained from (4.1).

Important applications of (4.1) refer to the Bonferroni and Lorenz curves to study income and poverty, but also in other fields such as reliability, demography, medicine and insurance. For given probability p, they are given by $B(p) = Z(q)/(p\mu)$ and $L(p) = Z(q)/\mu$, respectively, where $q = Q(p)$ comes directly from (3.1).

5. Rényi and Shannon Entropies

The entropy of a random variable T with density function $f(t)$ is a measure of variation of the uncertainty. One of the popular entropy measure is the Rényi entropy given by

$$I_R(\eta) = \frac{1}{1-\eta} \log[\int_{\Re} f^{\eta}(t)dt],$$

where $\eta > 0,\ \eta \neq 1$. The quantity $f^{\eta}(t)$ for T reduces to

$$f^{\eta}(t) = \frac{1}{[\log(p)]^{\eta}} \left\{ \frac{(p-1)^{\eta}\ \left(\frac{\nu\beta}{\alpha}\right)^{\eta}\ \left(\frac{t}{\alpha}\right)^{\eta\beta-\eta}\ \mathrm{e}^{-\eta\left(\frac{t}{\alpha}\right)^{\beta}}}{\left[1-(1-p\ \nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]^{\eta}\left[1-(1-\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]^{\eta}} \right\}. \tag{5.1}$$

Using the power series in equation (5.1), we can write

$$\begin{aligned} f^{\eta}(t) &= \frac{(p-1)^{\eta}}{[\log(p)]^{\eta}} \left(\frac{\nu\beta}{\alpha}\right)^{\eta} \sum_{i,j=0}^{\infty} \frac{\Gamma(\eta+i)\Gamma(\eta+j)}{[\Gamma(\eta)]^2\, j!} (1-p\ \nu)^i (1-\nu)^j\ \left(\frac{t}{\alpha}\right)^{\eta\beta-\eta} \\ &\times\ \mathrm{e}^{-(\mathrm{i+j}+\eta)\left(\frac{t}{\alpha}\right)^{\beta}}. \end{aligned}$$

Then, after some calculations, $I_R(\eta)$ reduces to

$$\begin{aligned} I_R(\eta) &= \frac{1}{1-\eta} \log\left[\frac{(p-1)^{\eta}}{[\log(p)]^{\eta}} \nu^{\eta} \left(\frac{\beta}{\alpha}\right)^{\eta-1}\right] \\ &+ \frac{1}{1-\eta} \log\left[\sum_{i,j=0}^{\infty} \frac{\Gamma(\eta+i)\Gamma(\eta+j)}{[\Gamma(\eta)]^2\, j!} \frac{(1-p\ \nu)^i(1-\nu)^j}{(i+j+\eta)^{\eta-\frac{\eta-1}{\beta}}} \Gamma\left(\eta - \frac{\eta-1}{\beta}\right)\right]. \end{aligned}$$

The Shannon entropy, which is defined by $E\{-\log[f(T)]\}$, can be derived numerically from $\lim_{\eta\to 1} I_R(\eta)$.

6. Reliability Function

In the context of reliability, the stress-strength model describes the life of a component which has a random strength T_1 that is subjected to a random stress T_2. The component fails at the instant when the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $T_1 > T_2$. Hence, $R = \Pr(T_2 < T_1)$ is a measure of component reliability. It has many applications especially in engineering concepts such as strength failure and system collapse. Now, we obtain the reliability R when T_1 and T_2 have independent (P-A-L)EW($\alpha, \beta, \nu_1, p_1$) and (P-A-L)EW($\alpha, \beta, \nu_2, p_2$) distributions with the same shape parameter β and scale parameter α. The reliability R is defined by $R = \int_0^{\infty} f_1(t)\, F_2(t)dx$.

By using the power series $\log(1+x)=\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}x^i$, $F_2(t)$ can be written as

$$F_2(t)=1-\frac{1}{\log(p_2)}\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}\frac{\nu_2^i(p_2-1)^i\,\mathrm{e}^{-\mathrm{i}\left(\frac{t}{\alpha}\right)^{\beta}}}{\left[1-(1-\nu_2)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]^i}.$$

By expanding $f_1(t)$ and $F_2(t)$, we can write $f_1(t)\,F_2(t)$ as

$$\begin{aligned} f_1(t)F_2(t) &= f_1(t)-\frac{1}{\log(p_1+p_2)}\sum_{i=1}^{\infty}\sum_{j,k,l=0}^{\infty}\frac{(-1)^{i+1}\Gamma(i+l)}{\Gamma(i+1)\,l!}\frac{\nu_1\beta,\nu_2^i}{\alpha}(p_1-1)\\ &\quad\times(p_2-1)^i\,(1-p_1\,\nu_1)^j(1-\nu_1)^k(1-\nu_1)^l\left(\frac{t}{\alpha}\right)^{\beta-1}\mathrm{e}^{-(\mathrm{i+j+k+l})\left(\frac{t}{\alpha}\right)^{\beta}}. \end{aligned} \tag{6.1}$$

Inserting (6.1) into the general expression for R and, after some algebra, we obtain

$$\begin{aligned} R &= 1-\frac{1}{\log(p_1+p_2)}\sum_{i=1}^{\infty}\sum_{j,k,l=0}^{\infty}\frac{(-1)^{i+1}\Gamma(i+l)\,\nu_1\,(\nu_2)^i}{\Gamma(i+1)\,l!\,(i+j+k+l)}\\ &\quad\times(p_1-1)\,(p_2-1)^i\,(1-p_1\nu_1)^j\,(1-\nu_1)^k(1-\nu_1)^l. \end{aligned}$$

7. Reliability Analysis

Here, we present the reliability function, hrf, chrf and mrlf of T.

7.1. Survival function. The (P-A-L)EW distribution can be a useful characterization of lifetime data analysis for a given system. Its survival function is

$$\overline{F}(t)=\frac{1}{\log(p)}\log\left\{\frac{1-(1-p\,\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right\}.$$

Figure 3 illustrates the survival behavior of the new distribution for some parameter values.

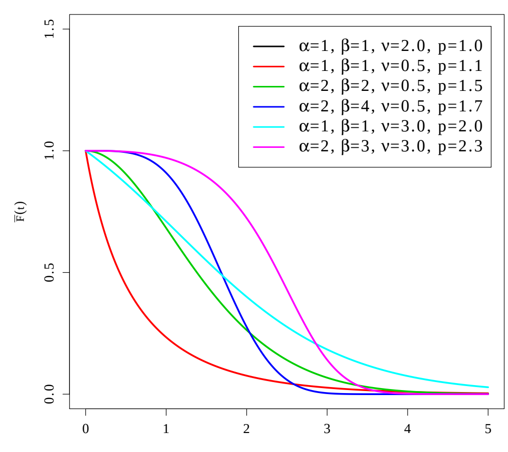


Figure 3. The survival function of the (P-A-L)EW distribution.

7.2. Hazard rate function. The hrf of T is given by

$$h(t)=\frac{(p-1)\frac{\nu\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\left[1-(1-\mathrm{p}\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]^{-1}}{\left[1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]\,\log\left\{\frac{1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right\}}.$$

We note that $h(t)$ can be constant, increasing, or decreasing depending on the parameter values. For example, if $p\to 1$, $\nu=1$ and $\beta=1$, then $h(t)=\frac{1}{\alpha}$ is constant, whereas if $p\to 1$ and $\nu=1$, then $h(t)=\frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}$, which is increasing for $\beta>1$ and decreasing for $\beta<1$. Figure 4 displays some plots of the hrf of T.

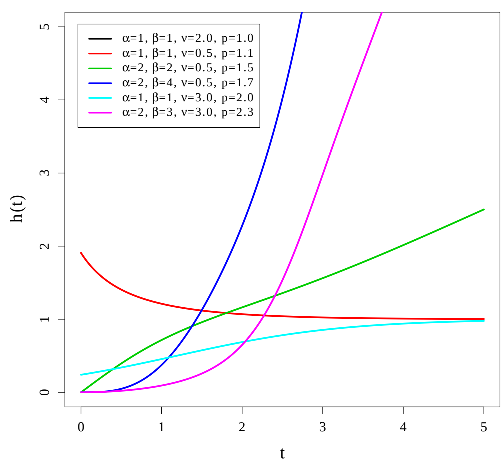


Figure 4. The hrf of the (P-A-L)EW distribution

7.3. Cumulative hazard rate function. Many generalized Weibull models have been proposed in reliability literature through the relationship between the reliability function $R(t)$ and the chrf $H(t)$, which is a non–decreasing function of t, given by $H(t)=-\log[R(t)]$. The chrf of T becomes

$$H(t)=\int_0^t h(u)du=\log\left(\log p\right)-\log\left\{\log\left[\frac{1-(1-p\,\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right]\right\},$$

where $H(t)$ is the total number of failures or deaths over an interval of time. Figure 5 illustrates the behavior of the chrf of T for some parameter values.

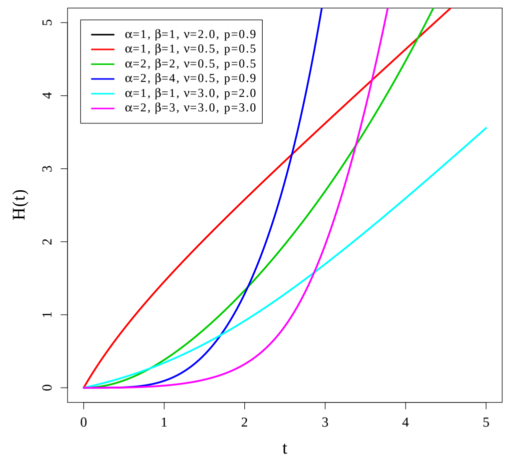


Figure 5. Cumulative Hazard Rate Function.

7.4. Mean residual lifetime function. The additional life time given that the component has survived up to time t is the rlf of the component. Then, the expectation of the random variable T_t represents the remaining lifetime reduces to

$$m(t) = E(T_t) = E(T - t \mid T > t) = \frac{\int_t^{\infty} R(u)du}{R(t)}\,.$$

The mrlf and the hrf are important since they characterize uniquely the corresponding lifetime distribution. We obtain

$$\begin{aligned} m(t) \quad &= \quad -t + \frac{\alpha}{\log\left[\frac{1-(1-p\,\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right]} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{1}{\beta}+1,(j+1)\left(\frac{t}{\alpha}\right)^{\beta}\right)}{(i+1)^{\frac{1}{\beta}+1}} \\ &\quad \times \left\{(1-\nu)^{i+1} - (1-p\nu)^{i-1}\right\}. \end{aligned}$$

7.5. Order Statistics. Let $T_1, \ldots, T_n$ denote n independent random variables from a distribution function $F(t)$ with pdf $f(t)$, and $T_{(1)}, \ldots, T_{(n)}$ denote the order sample arrangement. So, the pdf of $T_{(j)}$ is given by

$$f_{T_{(j)}}(t) = \frac{n!}{(j-1)!(n-j)!}\, f(t)\, F(t)^{j-1}\, [1 - F(t)]^{n-j} \quad \text{for} \;\; j = 1, \ldots, n.$$

Using equations (2.1) and (2.2), the pdf of $T_{(j)}$ becomes

$$\begin{aligned} f_{T_{(j)}}(t) &= \frac{n!}{(j-1)!(n-j)!} \times \left\{1 - \frac{1}{\log p} \log\left[\frac{1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right]\right\}^{j-1} \\ &\times \frac{(p-1)\frac{\nu\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{[\log(p)]^{n-j+1}\left[1-(1-p\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]\left[1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right]} \\ &\times \left\{\log\left[\frac{1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right]\right\}^{n-j}. \end{aligned}$$

Therefore, the pdf's of the smallest order statistic $T_{(1)}$ and of the largest order statistic $T_{(n)}$ are easily obtained from the last equation with $i = 1$ and $i = n$, respectively. Then, the minimum and maximum order statistics can be derived for some special models of the new distribution. For example, for the (P-A-L)ER ($\beta = 2$), (P-A-L)EE model ($\beta = 1$), (P-A-L)W ($\nu = 1$), (P-A-L)R ($\nu = 1$ and $\beta = 2$), (P-A-L)E ($\nu = 1$ and $\beta = 1$) and EW ($p \to 1$) distributions, among others.

The pdf's of the $(k+1)$th and kth ordered statistics from the (P-A-L)EW model obey the relationship

$$f_{T_{(k+1)}}(t) = \left(\frac{n-k}{k}\right) \frac{1 - \frac{1}{\log(p)} \log\left[\frac{1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right]}{\frac{1}{\log(p)} \log\left[\frac{1-(1-p\,\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}{1-(1-\nu)\,\mathrm{e}^{-\left(\frac{t}{\alpha}\right)^{\beta}}}\right]}\, f_{T_{(k)}}(t).$$

8. Estimation of the Parameters

Inference can be carried out in three different ways: point estimation, interval estimation and hypothesis testing. Several approaches for parameter point estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimates (MLEs) enjoy desirable properties and can be used in constructing confidence intervals and also in test-statistics. Large sample theory for these estimates delivers simple approximations that work well in finite samples. Statisticians often seek to approximate quantities such as the density of a test-statistic that depend on the sample size in order to obtain better approximate distributions. The resulting approximation for the MLEs in distribution theory is easily handled either analytically or numerically. In this section, we use the method of likelihood to estimate the model parameters and use them to obtain confidence intervals for the unknown parameters.

8.1. Maximum Likelihood Estimation. Let $t_1, \ldots, t_n$ be a sample of size n from the (P-A-L)EW distribution. Let $\theta = (\alpha, \beta, \nu, p)^T$ be the parameter vector. Then, the log-likelihood function $\ell = \ell(\theta)$ is given by

$$\begin{aligned} \ell &= n\log[\log(p)] + n\log(p-1) - n\log\left(\frac{\nu\beta}{\alpha}\right) + (\beta-1)\sum_{i=1}^{n}\log\left(\frac{t_i}{\alpha}\right) \\ &- \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} + \sum_{i=1}^{n}\log\left[1-(1-p\nu)\mathrm{e}^{-\left(\frac{t_i}{\alpha}\right)^{\beta}}\right] - \sum_{i=1}^{n}\log\left[1-(1-\nu)\mathrm{e}^{-\left(\frac{t_i}{\alpha}\right)^{\beta}}\right]. \end{aligned}$$

Then, the MLE of θ can be derived from the derivatives of ℓ. They should satisfy the following equations:

$$\begin{aligned}\frac{\partial \ell}{\partial \alpha} &= -\frac{n\beta}{\alpha} + \frac{\beta}{\alpha}\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} + \frac{\beta}{\alpha}\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} \mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}} \\ &\quad \times \left[\frac{1-p\nu}{\left[1-(1-p\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]} + \frac{(1-\nu)}{\left[1-(1-\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]}\right] = 0,\end{aligned}$$

$$\begin{aligned}\frac{\partial \ell}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^{n}\left[1-\left(\frac{t_i}{\alpha}\right)^{\beta}\right]\log\left(\frac{t_i}{\alpha}\right)^{\beta} - \sum_{i=1}^{n}\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}\log\left(\frac{\mathrm{t_i}}{\alpha}\right) \\ &\quad \times \left[\frac{(1-p\nu)}{\left[1-(1-p\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]} + \frac{(1-\nu)}{\left[1-(1-\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]}\right] = 0,\end{aligned}$$

$$\frac{\partial \ell}{\partial \nu} = \frac{n}{\nu} - p\sum_{i=1}^{n}\frac{\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}}{\left[1-(1-p\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]} - p\sum_{i=1}^{n}\frac{\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}}{\left[1-(1-\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]} = 0$$

and

$$\frac{\partial \ell}{\partial p} = \frac{-n}{p\log p} + \frac{-n}{p-1} - \nu\sum_{i=1}^{n}\frac{\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}}{\left[1-(1-p\nu)\mathrm{e}^{-\left(\frac{\mathrm{t_i}}{\alpha}\right)^{\beta}}\right]} = 0.$$

These equations cannot be solved analytically, and statistical softwares are required to solve them numerically. To solve these equations, it is usually more convenient to use nonlinear optimization algorithms such as quasi-Newton algorithm to numerically maximize the log-likelihood function. For interval estimation of the parameters, we obtain the 3×3 observed information matrix $J(\theta) = \{\frac{\partial^2 \ell}{\partial r\, s}\}$ (for $r, s = \alpha, \beta, \nu, p$), whose elements can be computed numerically.

Under standard regularity conditions when $n \to \infty$, the distribution of the MLE can be approximated by a multivariate normal $N_4(0, J(\widehat{\theta})^{-1})$ distribution to construct approximate confidence intervals for the parameters. Here, $J(\widehat{\theta})$ is the total observed information matrix evaluated at $\widehat{\theta}$. The method of the re-sampling bootstrap can be used for correcting the biases of the MLEs of the model parameters. Good interval estimates may also be obtained through the bootstrap percentile method.

9. Particle Swarm Optimization

In computer science, the particle swarm optimization (PSO) is a computational method for optimization of parametric and multiparametric functions. The PSO algorithm is a meta-heuristic which has been providing good solutions for problems of optimization global functions with box-constrained. The use of meta-heuristic methods such as PSO has proved to be useful for maximizing complicated log-likelihood functions without the need for early kick functions as the BFGS, L-BFGS-B, Nelder-Mead and simulated annealing methods. As in most heuristic methods that are inspired by biological phenomena, the PSO is inspired by the behavior of flying birds. The philosophical idea of the PSO algorithm is based on the collective behavior of birds (particle) in search of food (point of global optimal). The PSO technique was first defined by [6] in a paper published in the Proceedings of the IEEE International Conference on Neural Networks IV.

A modification of the PSO algorithm was proposed by [16] published in the Proceedings of IEEE International Conference on Evolutionary Computation. Further details on the philosophy of the PSO method are given in the book Swarm Intelligence (see [8]).

The PSO optimizes a problem by having a population of candidate solutions and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. The movement of the particles in the search space is randomized. Each iteration of the PSO algorithm, there is a leader particle, which is the particle that minimizes the objective function in the respective iteration. The remaining particles arranged in the search region will follow the leader particle randomly and sweep the area around this leading particle. In this local search process, another particle may become the new leader particle and the other particles will follow the new leader randomly. Each particle arranged in the search region has a velocity vector and position vector and its movement in the search region is given by changes in these vectors. The PSO algorithm is presented below, where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the objective function to be minimized, S is the number of particles in the swarm (set of feasible points, i.e. search region), each having particle a vector position $x_i \in \mathbb{R}^n$ in the search-space and a vector velocity defined by $v_i \in \mathbb{R}^n$. Let p_i be the best known position of particle i and g the best position of all particles.

(1) For each particle $i = 1, \ldots, S$ do:
  - Initialize the particle's position with a uniformly distributed random vector: $x_i \sim U(b_{lo}, b_{up})$, where b_{lo} and b_{up} are the lower and upper boundaries of the search-space.
  - Initialize the particle's best known position to its initial position: $p_i \hookleftarrow x_i$.
  - If $f(p_i) < f(g)$ update the swarm's best known position: $g \hookleftarrow p_i$.
  - Initialize the particle's velocity: $v_i \sim U(-|b_{up} - b_{lo}|, |b_{up} - b_{lo}|)$.

(2) Until a termination criterion is met (e.g. number of iterations performed, or a solution with adequate objective function value is found), repeat:
  - For each particle $i = 1, \ldots, S$ do:
    - Pick random numbers: $r_p, r_g \sim U(0, 1)$.
    - For each dimension $d = 1, \ldots, n$ do:
      * Update the particle's velocity: $v_{i,d} \hookleftarrow \omega\, v_{i,d} + \varphi_p r_p (p_{i,d} - x_{i,d}) + \varphi_g r_g (g_d - x_{i,d})$.
    - Update the particle's position: $x_i \hookleftarrow x_i + v_i$
    - If $f(x_i) < f(p_i)$ do:
      * Update the particle's best known position: $p_i \hookleftarrow x_i$
      * If $f(p_i) < f(g)$ update the swarm's best known position: $g \hookleftarrow p_i$.

(3) Now g holds the best found solution.

The parameter ω is called inertia coefficient and as the name implies controls the inertia of each particle arranged in the search region. The quantities ω_p and ω_g control the acceleration of each particle and are called acceleration coefficients.

10. Application

We consider an application using the (P-A-L)EW distribution. We use the `AdequacyModel` script version 1.0.8 available for the programming language `R`. The script is currently maintained by one of the authors of this paper and more information can be obtained f rom http://cran.rstudio.com/web/packages/AdequacyModel/index.html. The package is distributed under the terms of the licenses GNU General Public License (GPL-2 or GPL-3).

The application take into account the data relating to the percentage of body fat determined by underwater weighing and various body circumference measurements for 250 men. For details about the data set, see http://lib.stat.cmu.edu/datasets/.

Table 2. Descriptive statistics.

Statistics	Real data sets
	Body Fat (%)
Mean	19.3012
Median	19.2500
Mode	22.5000
Variance	67.7355
Skewness	0.1953
Kurtosis	-0.3815
Maximum	47.5000
Minimum	3.0000
n	250

In order to determine the shape of the most appropriate hazard function for modeling, graphical analysis data may be used. In this context, the total time in test (TTT) plot proposed by [1] is very useful. Let T be a random variable with non-negative values which represents the survival time. The TTT curve is obtained by constructing the plot of $G(r/n) = [(\sum_{i=1}^{r} T_{i:n}) + (n-r)T_{r:n}]/(\sum_{i=1}^{n} T_{i:n})$ versus r/n, for $r = 1, \ldots, n$ and $T_{i:n}$ $(i = 1, \ldots, n)$ are the order statistics of the sample (see [11]). The plots can be easily obtained using the function `TTT` of the script `AdequacyModel`. For more details on this function, see `help(TTT)`. The TTT plot for the current data is displayed in Figure 6, which is concave and according to [1] provides evidence that the monotonic hrf is adequate.

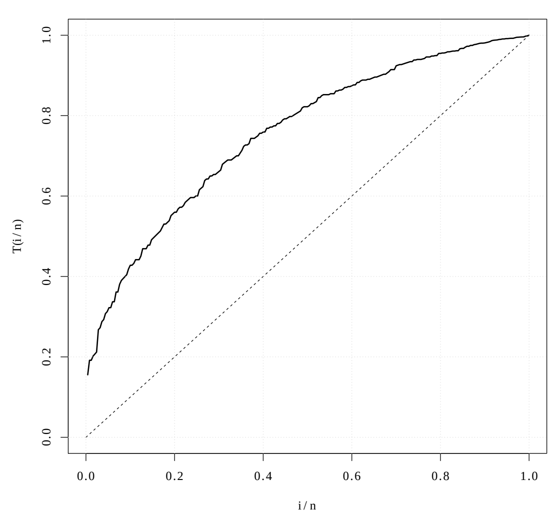


Figure 6. The TTT plot for percentage of body fat.

Figure 7 displays the estimated density to the data obtained in a nonparametric manner using kernel density estimation with the Gaussian filter. Let $X_1, \ldots, X_n$ be a random vector of independent and identically distributed random variables, when each random variable follows an unknown pdf f. The kernel density estimator is given by

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{10.1}$$

where $K(\cdot)$ is the kernel function usually symmetrical and $\int_{-\infty}^{\infty} K(x)dx = 1$. Here, $h > 0$ is a smoothing parameter known in literature as bandwidth. Numerous kernel functions are adopted in the literature. The normal standard distribution is the most widely used because it has convenient mathematical properties. [17] demonstrated that for the K standard normal, the bandwidth ideal is $h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}} \approx 1.06\,\hat{\sigma}\,n^{-1/5}$, where $\hat{\sigma}$ is the standard deviation of the sample.

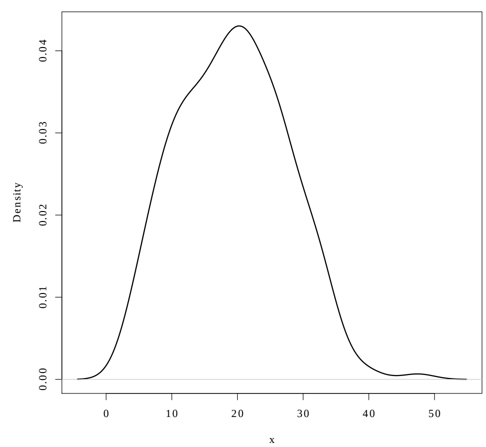


Figure 7. Gaussian kernel density estimation for percentage of body fat.

In order to verify which distribution fits better these data, we consider the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics described by [5]. Chen and Balakrishnan (see [5]) constructed the Cramér-von Mises and Anderson-Darling corrected statistics based on the suggestions from [18]. We use these statistics, where we have a random sample $(x_1, \ldots, x_n)$ with empirical distribution function $F_n(x)$ and we want to test if the sample comes from a special distribution. The Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics are, respectively, given by

$$\begin{aligned} W^* &= \left\{ n\int_{-\infty}^{+\infty} \{F_n(x) - F(x;\widehat{\theta}_n)\}^2 dF(x;\widehat{\theta}_n) \right\} \left(1 + \frac{0.5}{n}\right) = W^2\left(1 + \frac{0.5}{n}\right), \\ A^* &= \left\{ n\int_{-\infty}^{+\infty} \frac{\{F_n(x) - F(x;\widehat{\theta}_n)\}^2}{\{F(x;\widehat{\theta})(1 - F(x;\widehat{\theta}_n))\}} dF(x;\widehat{\theta}_n) \right\} \left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right) \\ &= A^2\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right), \end{aligned}$$

where $F_n(x)$ is the empirical distribution function, $F(x;\hat{\theta}_n)$ is the postulated distribution function evaluated at the MLE $\hat{\theta}_n$ of θ. Note that the statistics W^* and A^* are given by the differences of $F_n(x)$ and $F(x;\hat{\theta}_n)$. Thus, the lower are the statistics W^* and A^* more evidence we have that $F(x;\hat{\theta}_n)$ generates the sample. The details to compute the statistics W^* and A^* are given by Chen and Balakrishnan.

The `goodness.fit` function provides various adequacy of fit statistics, among them, the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics proposed by [5], Consistent Akaike Information Criterion (CAIC) defined by [2], Bayesian Information Criterion (BIC) defined by [15] and Hannan-Quinn Information Criterion (HQIC) given by [7]. These statistics are used to assess the adequacy of the fit of the distributions considered in the two real data sets.

The PSO methodology was used for the improvement of the MLEs. Initially, we use the Nelder-Mead method to maximize the log-likelihood function of the models under study using the `goodness.fit` function of the script `AdequacyModel`. After obtaining convergence using the Nelder-Mead method (see [12]), we use the PSO method as an attempt to obtain best candidates for global maximums of their log-likelihood functions for the compared models. We consider $S = 550$ (550 particles) and 500 iterations as stopping criterion. We choose as optimal candidates for the estimates, those MLEs calculated by the PSO method when ℓ (the maximized log-likelihood function for the current model) is higher than the log-likelihood function evaluated at the estimates computed by the Nelder-Mead method. Figure 8 displays the fitted densities to the current data. The MLEs used in Figure 8 are highlighted in Table 3. It is noted in Table 4 that the proposed distribution provides the best fit to the data.

Table 3. MLEs obtained by Nelder-Mead and PSO methods.

Distributions		Estimates				ℓ
(P-A-L)EW	PSO	1.8571	0.7700	63.4424	37.5844	871.0364
	Nelder-Mead	**19.6993**	**2.5831**	**0.2865**	**28.0810**	**874.7802**
Kw-W	PSO	**71.3501**	**77.4079**	**0.1635**	**25.1942**	**888.7122**
	Nelder-Mead	0.6960	2.0492	3.3057	0.0314	875.8679
Exp-W	PSO	45.30062	69.5828	55.3411	-	875.8749
	Nelder-Mead	**0.0418**	**3.0356**	**0.7436**	**-**	**870.6432**
Weibull	PSO	**21.7567**	**2.5373**	**-**	**-**	**876.4216**
	Nelder-Mead	21.7552	2.5371	-	-	876.1854
Gamma	PSO	**5.8060**	**0.3036**	**-**	**-**	**888.5930**
	Nelder-Mead	4.6090	0.2388	-	-	884.6877

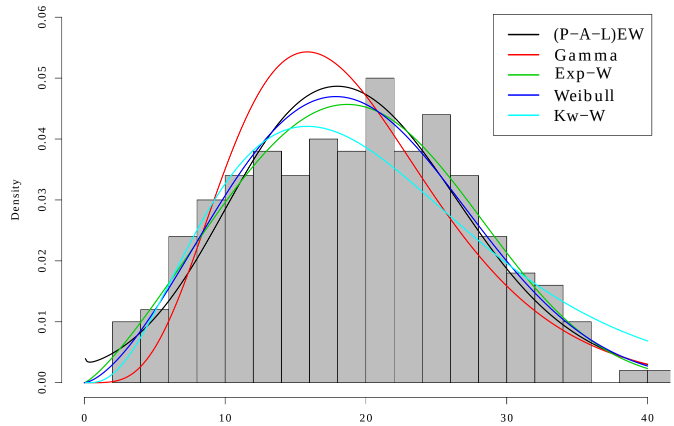


Figure 8. Fitted densities to the percentage of body fat data.

Table 4. Statistics of adequacy to adjust.

Distributions	AIC	CAIC	BIC	HQIC	A^*	W^*
(P-A-L)EW	1757.560	1757.724	1771.646	1763.230	**0.1192**	**0.0144**
Kw-W	1785.429	1785.592	1799.515	1791.098	1.8205	0.3005
Exp-W	1757.750	1757.847	1768.314	1762.002	0.2477	0.0334
Weibull	1756.843	1756.892	1763.886	1759.678	0.4357	0.0668
Gamma	1781.186	1781.235	1788.229	1784.021	1.9548	0.3233

11. Concluding Remarks

The idea of generating new extended models from classic ones has been of great interest among researchers in the past decade. A new four-parameter generalization of the Weibull model, called the (P-A-L) extended Weibull, (P-A-L)EW for short, distribution is defined and some of its mathematical properties studied. They include moments, generating, quantile, reliability and residual life functions, mean deviations and two types of entropies. Many well-known distributions emerge as special cases of the proposed distribution by using special parameter values. We use maximum likelihood and a particle swarm optimization method to estimate the model parameters. By means of a real data set, we prove that this model has the capability to provide consistent estimates from the considered estimation methods.

Acknowledgment

This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under Grant no. (to be specified soon). The authors acknowledge with thanks DSR technical and financial support. We also are grateful to FACEPE and CNPq for financial support.

References

[1] Magne Vollan Aarset, *How to identify a bathtub hazard rate.*, IEEE Transactions on Reliability **36** (1987), no. 1, 106–108.

[2] Hirotugu Akaike, *A new look at the statistical model identification*, Automatic Control, IEEE Transactions on **19** (1974), no. 6, 716–723.

[3] Saad J Almalki and Saralees Nadarajah, *Modifications of the weibull distribution: A review*, Reliability Engineering & System Safety **124** (2014), 32–55.

[4] Jeff Bezanson, Stefan Karpinski, Viral B Shah, and Alan Edelman, *Julia: A fast dynamic language for technical computing*, arXiv preprint arXiv:1209.5145 (2012).

[5] Gemai Chen and N Balakrishnan, *A general purpose approximate goodness-of-fit test*, Journal of Quality Technology **27** (1995), no. 2, 154–161.

[6] Russ C Eberhart and James Kennedy, *A new optimizer using particle swarm theory*, Proceedings of the sixth international symposium on micro machine and human science, vol. 1, New York, NY, 1995, pp. 39–43.

[7] Edward J Hannan and Barry G Quinn, *The determination of the order of an autoregression*, Journal of the Royal Statistical Society. Series B (Methodological) (1979), 190–195.

[8] James Kennedy, James F Kennedy, and Russell C Eberhart, *Swarm intelligence*, Morgan Kaufmann, 2001.

[9] Pushpa Lata Gupta and Ramesh C Gupta, *On the moments of residual life in reliability and some characterization results*, Communications in Statistics-Theory and Methods **12** (1983), no. 4, 449–461.

[10] Albert W Marshall and Ingram Olkin, *A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families*, Biometrika **84** (1997), no. 3, 641–652.

[11] Govind S Mudholkar and Alan D Hutson, *The exponentiated Weibull family: some properties and a flood data application*, Communications in Statistics–Theory and Methods **25** (1996), no. 12, 3059–3083.

[12] John A Nelder and Roger Mead, *A simplex method for function minimization*, The Computer Journal **7** (1965), no. 4, 308–313.

[13] Vasileios Pappas, Konstantinos Adamidis, and Sotirios Loukas, *A family of lifetime distributions*, International Journal of Quality, Statistics, and Reliability **2012** (2012).

[14] Manoel Santos-Neto, Marcelo Bourguignon, Luz M Zea, Abraão DC Nascimento, and Gauss M Cordeiro, *The marshall-olkin extended weibull family of distributions*, Journal of Statistical Distributions and Applications **1** (2014), no. 1, 9.

[15] Gideon Schwarz et al., *Estimating the dimension of a model*, The Annals of Statistics **6** (1978), no. 2, 461–464.

[16] Yuhui Shi and Russell Eberhart, *A modified particle swarm optimizer*, Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, IEEE, 1998, pp. 69–73.

[17] Bernard W Silverman, *Density estimation for statistics and data analysis*, vol. 26, CRC press, 1986.

[18] Michael A Stephens, *Tests based on EDF statistics*, Goodness-of-Fit Techniques, RB D'Agostino and MS Stephens, Eds. Marcel Dekker (1986).

[19] Ronghua Wang, Naijun Sha, Beiqing Gu, and Xiaoling Xu, *Statistical analysis of a weibull extension with bathtub-shaped failure rate function*, Advances in Statistics **2014** (2014).

[20] Tieling Zhang and Min Xie, *Failure data analysis with extended weibull distribution*, Communications in Statistics-Simulation and Computation **36** (2007), no. 3, 579–592.