Mathematical Modelling and Analysis
Volume 22 Number 2, March 2017, 186–201
https://doi.org/10.3846/13926292.2017.1289481

Publisher: Taylor&Francis and VGTU
http://www.tandfonline.com/TMMA
ISSN: 1392-6292
eISSN: 1648-3510

The Power Muth Distribution*

Pedro Jodrá[a], Héctor Wladimir Gómez[b], María Dolores Jiménez-Gamero[c] and María Virtudes Alba-Fernández[d]

[a] *Dpto. de Métodos Estadísticos, Universidad de Zaragoza*
María de Luna 3, 50018 Zaragoza, Spain
[b] *Dpto. de Matemáticas, Universidad de Antofagasta*
601 Avda. Angamos, Antofagasta, Chile
[c] *Dpto. de Estadística e Investigación Operativa, Universidad de Sevilla*
Avda. Reina Mercedes s.n., 41012 Sevilla, Spain
[d] *Dpto. de Estadística e Investigación Operativa, Universidad de Jaén*
Paraje Las Lagunillas s.n., 23071 Jaén, Spain
E-mail(*corresp.*): pjodra@unizar.es
E-mail: hector.gomez@uantof.cl
E-mail: dolores@us.es
E-mail: mvalba@ujaen.es

Received September 29, 2016; revised January 24, 2017; published online March 1, 2017

Abstract. Muth introduced a probability distribution with application in reliability theory. We propose a new model from the Muth law. This paper studies its statistical properties, such as the computation of the moments, computer generation of pseudo-random data and the behavior of the failure rate function, among others. The estimation of parameters is carried out by the method of maximum likelihood and a Monte Carlo simulation study assesses the performance of this method. The practical usefulness of the new model is illustrated by means of two real data sets, showing that it may provide a better fit than other probability distributions.

Keywords: Muth distribution, survival analysis, estimation, simulation, data analysis.

AMS Subject Classification: 60E05; 62F10; 33B30.

1 Introduction

In the last decades, there exists an increasing interest in the development of new parametric distributions with the aim of providing flexible probability models

* Research in this paper has been partially funded by the research projects: CTM2015-68276-R MINECO/FEDER,UE (M.V. Alba-Fernández), Diputación General de Aragón –Grupo consolidado PDIE– (P. Jodrá), FONDECYT 1130495 (H.W. Gómez) and MTM2014-55966-P of the Spanish Ministry of Economy and Competitiveness (M.D. Jiménez-Gamero).

useful in many different areas. In 1977, Muth [25] introduced a continuous probability distribution in the context of reliability theory. However, this law has been overlooked in the literature until a recent paper by Jodrá et al. [20], where its main statistical properties were thoroughly studied. To be more precise, a random variable Y is said to have a Muth distribution if its probability density function (pdf) is given by

$$f_Y(y;\alpha) = (e^{\alpha y} - \alpha)\,\exp\Big\{\alpha y - \frac{1}{\alpha}\,(e^{\alpha y} - 1)\Big\}, \quad y > 0,$$

where $\alpha \in (0,1]$ is a shape parameter. The cumulative distribution function (cdf) of Y is the following

$$F_Y(y;\alpha) = 1 - \exp\Big\{\alpha y - \frac{1}{\alpha}\,(e^{\alpha y} - 1)\Big\}, \quad y > 0. \tag{1.1}$$

As a natural extension, Jodrá et al. [20] considered the scaled Muth distribution, which is defined as βY, with $\beta > 0$, and they showed the usefulness of the scaled Muth law for modelling rainfall data.

In this paper, we introduce a new probability distribution from the Muth law. Specifically, we define the so-called power Muth (PM) distribution by means of the transformation $X = \beta Y^{1/\gamma}$, where $\beta > 0$, $\gamma > 0$ and Y is the Muth law with parameter $\alpha = 1$. As it will be seen, the introduction of the new parameter γ leads to a rich class of probability distributions for non-negative random variables with a wide range of values for the asymmetry and kurtosis coefficients, increasing generalized failure rate as well as increasing or bathtub shape failure rate. The choice $\alpha = 1$ is adopted to avoid unnecessarily increasing the number of parameters since both α and γ are shape parameters. From the viewpoint of applications, the new family of distributions may provide a better fit than other probability distributions previously used to this end.

Using ordinary results related to the transformation of variables, it is easy to see that the pdf and cdf of X are given, respectively, by

$$f(x;\beta,\gamma) = \frac{\gamma}{\beta^\gamma}\,x^{\gamma-1}\left(e^{(x/\beta)^\gamma} - 1\right)\exp\left\{(x/\beta)^\gamma - \left(e^{(x/\beta)^\gamma} - 1\right)\right\},\ x > 0,$$

$$F(x;\beta,\gamma) = 1 - \exp\left\{(x/\beta)^\gamma - \left(e^{(x/\beta)^\gamma} - 1\right)\right\}, \quad x > 0, \tag{1.2}$$

where β is a scale parameter and γ is a shape parameter. Figure 1 displays the pdf of the PM distribution for different values of the parameters. Throughout this paper, the PM law with parameters β and γ is denoted by PM(β,γ).

It is interesting to note that $f(x;\beta,\gamma)$ is a well-defined pdf for any $\beta > 0$ and $\gamma > 0$. However, routine calculations show that

$$\lim_{x\to 0^+} f(x;\beta,\gamma) = \begin{cases} \infty, & \text{if } \gamma \in (0, 0.5), \\ 0.5/\beta, & \text{if } \gamma = 0.5, \\ 0, & \text{if } \gamma \in (0.5, \infty). \end{cases}$$

Commonly used pdf estimators, such as the histogram or kernel based estimators, always yield a finite pdf estimation. Therefore, from a practical point

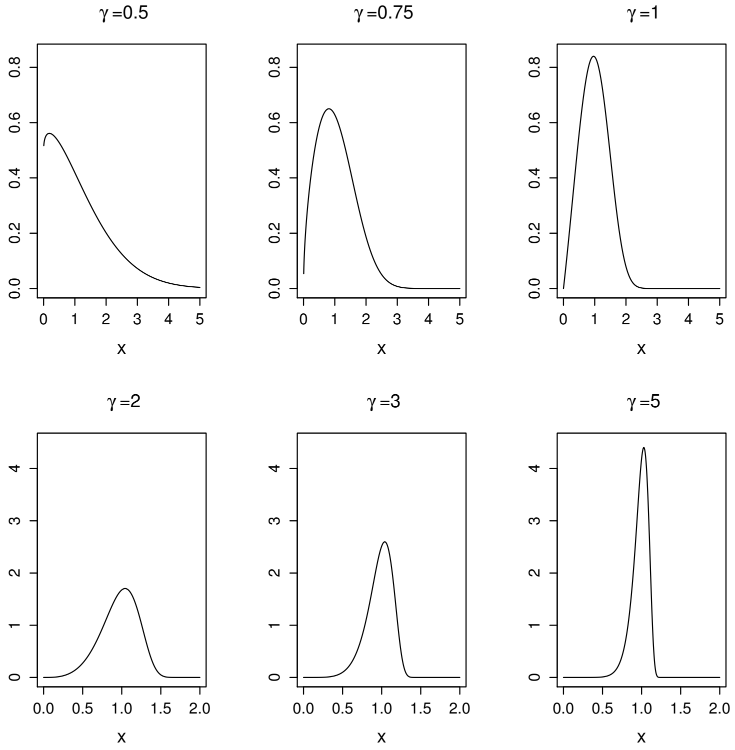

Figure 1. Pdf of the PM(β, γ) distribution for $\beta = 1$ and different values of γ.

of view, an unbounded pdf will not provide a reasonable fit for any data set. Because of this reason, we will only consider $\gamma \in [0.5, \infty)$. Nevertheless, most properties studied for the proposed family are valid for any $\gamma > 0$.

The remainder of this paper is organized as follows. Section 2 studies some statistical properties of the PM distribution. More precisely, the mode, the non-central moments and moments of the order statistics are expressed in terms of the generalized integro-exponential function. This function has also appeared in the expression of the moments of the order statistics of other distributions such as the Gompertz–Makeham distribution (see Jodrá [18]). The quantile function is written in terms of the Lambert W function. This function has found nice applications in probability and statistics for calculating certain statistical distances between some discrete distributions (see Adell and Jodrá [1]) and for the expression of the quantile function of other probability laws (see [16,17,19], among others). The behavior of the (generalized) failure rate function and the mean residual life is also described. The parameter estimation problem is considered in Section 3 and a simulation study is carried out to assess the performance of the maximum likelihood method. In Section 4, an application to two real data sets is presented to illustrate the practical usefulness of the proposed distribution.

2 Statistical properties

2.1 Mode

The mode of a continuous probability distribution, denoted by mode(X), is the value at which the pdf has its maximum. The next result shows that the PM(β,γ) distribution is unimodal for $\gamma \geq 0.5$. We first introduce some notation. Let

$$\Delta(\gamma) = \frac{\gamma - 1}{\gamma}, \quad g(z) = z\frac{(e^z - r_1)(e^z - r_2)}{e^z - 1},$$

where $r_1 = 2 - \varphi$, $r_1 = \varphi^2$ and φ stands for the golden ration. Recall that φ can be defined as the positive solution of the equation $x^2 - x - 1 = 0$, that is,

$$\varphi = \frac{1 + \sqrt{5}}{2} \approx 1.618034.$$

Therefore, $r_1 \approx 0.381966$ and $r_2 \approx 2.618034$. Clearly, $g(0) = g(r_2) = 0$, $g(z) < 0$, $\forall z \in (0, r_2)$, and $g(z) > 0$, $\forall z \in (r_2, \infty)$. Routine calculations show that g is a strictly convex function for $z \geq 0$ with a global minimum at $r_0 \in (0, r_2)$, specifically $r_0 \approx 0.226874$ and $g(r_0) \approx -1.059945$.

Proposition 1. *Let X be a random variable having a* PM(β,γ) *distribution with $\beta > 0$ and $\gamma \geq 0.5$. Then, X has a unique mode*

$$\text{mode}(X) = \beta\varsigma^{1/\gamma},$$

where: (a) *ς is the unique solution in $z \in (r_0, r_2)$ of the equation*

$$g(z) = \Delta(\gamma), \tag{2.1}$$

if $\gamma \in [0.5, 1)$, (b) *$\varsigma = r_2$, if $\gamma = 1$,* (c) *ς is the unique solution in $z \in (r_2, \infty)$ of Eq.* (2.1)*, if $\gamma > 1$.*

Proof. In order to obtain the mode of X we have to solve with respect to x the equation $(\partial/\partial x) f(x; \beta, \gamma) = 0$, which is equivalent to solve the following equation $(\partial/\partial x) \log f(x; \beta, \gamma) = 0$, where log is the natural logarithm. This is tantamount to solve with respect to z Eq. (2.1), with $z = (x/\beta)^\gamma$. Taking into account that $\Delta(\gamma)$ is a strictly increasing function of γ satisfying $\Delta(\gamma) \in [-1, 0]$ if $\gamma \in [0.5, 1]$ and $\Delta(\gamma) > 0$ if $\gamma > 1$, as well as the properties of $g(z)$ sketched above, the result follows. □

2.2 Moments

The moments can be written in terms of the generalized integro-exponential function, which is defined by the integral representation (see Milgram [24])

$$E_s^m(z) = \frac{1}{\Gamma(m+1)} \int_1^\infty (\log u)^m e^{-zu} u^{-s} du, \quad z \in (-\infty, \infty), \tag{2.2}$$

where $s \in (-\infty, \infty)$, $m > -1$ and Γ stands for the gamma function.

Proposition 2. *Let X have the* PM(β,γ) *distribution. Then,*

$$E[X^k] = e\,\beta^k \Gamma\left(\frac{k}{\gamma}+1\right) E_0^{\frac{k}{\gamma}-1}(1)\,, \qquad k=1,2,\ldots. \tag{2.3}$$

Proof. From Eq. (1.2), for any $k=1,2,\ldots$, we have

$$\begin{aligned} E[X^k] =& \int_0^\infty x^k dF(x;\beta,\gamma) \\ =& \int_0^\infty \frac{\gamma}{\beta^\gamma}\, x^{k+\gamma-1}\left(e^{(x/\beta)^\gamma}-1\right)\exp\left\{\left(\frac{x}{\beta}\right)^\gamma - \left(e^{(x/\beta)^\gamma}-1\right)\right\}dx \\ =& \; e\,\beta^k \int_1^\infty (\log u)^{k/\gamma} e^{-u}\,(u-1)du, \end{aligned}$$

where in the last equality we have made the change of variable $u=\exp\{(x/\beta)^\gamma\}$. Then, by virtue of Eq. (2.2), we get

$$E[X^k] = e\,\beta^k \Gamma\,(k/\gamma+1)\left\{E_{-1}^{k/\gamma}(1) - E_0^{k/\gamma}(1)\right\}.$$

Finally, by taking into account in the above equation the following recurrence formula (see Milgram [24, eqn 2.4])

$$(1-s)E_s^m(z) = zE_{s-1}^m(z) - E_s^{m-1}(z), \quad z>0, \quad s\neq 1, \quad m\geq 0, \tag{2.4}$$

where it is assumed $E_s^{-1}(z)=e^{-z}$, the desired result is obtained. □

The main mathematical properties of the generalized integro-exponential function have been studied by Ozalp and Bairamov [26], who provided an accurate algorithm to evaluate (2.2). As a consequence, from Eq. (2.3) we can compute efficiently the usual statistical measures involving $E[X^k]$. The behavior of some common measures can be graphically seen in the following figures, where we set $\beta=1$ since β is a scale parameter: Figure 2 displays the mean, standard deviation, asymmetry coefficient ($\kappa = E[(X-E[X])^3]/\sigma^3$) and kurtosis coefficient ($E[(X-E[X])^4]/\sigma^4$). As can be seen, the PM family has a wide range of values for the coefficients of asymmetry and kurtosis, which provides a great flexibility in modelling data. Moreover, for $\gamma\in[1.12,1.36]$ the asymmetry coefficient satisfies $|\kappa|\leq 0.1$, leading to quasi-symmetric densities. As suggested in Vargo et al. [30], a diagram of asymmetry coefficient versus kurtosis is also shown in Figure 3.

2.3 Quantile function

The PM distribution inherits the variate generation property from the Muth distribution, that is, its quantile function can also be written in closed form. Specifically, it can be expressed explicitly in terms of the Lambert W function (see Corless et al. [8] for a review of the theory and applications of W). In this regard, it is interesting to note that the Lambert W function is implemented in computer algebra systems such as Maple, Mathematica and Matlab and

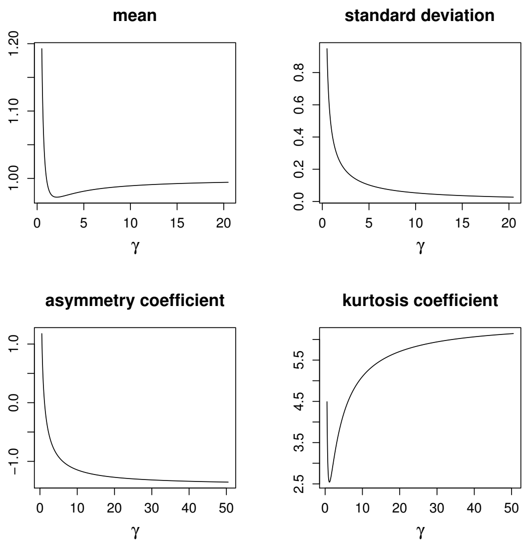

Figure 2. Mean, standard deviation, asymmetry and kurtosis as functions of γ ($\beta = 1$).

also in programming languages such as R [29]. Therefore, pseudo-random data from the PM model can be computer-generated in a straightforward manner by virtue of the following result.

Proposition 3. *The quantile function of the* PM(β, γ) *distribution is*

$$F^{-1}(u; \beta, \gamma) = \beta \left(\log \left\{ -W_{-1}\left(\frac{u-1}{e} \right) \right\} \right)^{1/\gamma}, \qquad 0 < u < 1, \tag{2.5}$$

where W_{-1} denotes the negative branch of the Lambert W function.

Proof. For any $0 < u < 1$, we have to solve with respect to x the equation $F(x; \beta, \gamma) = u$, which is equivalent to solve $F_Y((x/\beta)^\gamma; 1) = u$, with F_Y given by (1.1). Clearly, the solution of the latter equation is $x = \beta \left(F_Y^{-1}(u; 1) \right)^{1/\gamma}$. Then, taking into account the closed-form expression for F_Y^{-1} provided in Jodrá et al. [20, Corollary 2], the result in (2.5) is obtained. □

2.4 Failure rate function

The failure (or hazard) rate function of the PM(β, γ) distribution is

$$h(x; \beta, \gamma) = \frac{f(x; \beta, \gamma)}{1 - F(x; \beta, \gamma)} = \frac{\gamma}{\beta^\gamma} \left(e^{(x/\beta)^\gamma} - 1 \right) x^{\gamma - 1}, \qquad x > 0.$$

Proposition 4. (a) *If $\gamma \geq 1$ then $h(x; \beta, \gamma)$ is increasing in x for any $\beta > 0$.* (b) *If $0 < \gamma < 1$ then there exists an $x_0 = x_0(\beta, \gamma) > 0$ so that $h(x; \beta, \gamma)$ is (strictly) decreasing when $x < x_0$ and (strictly) increasing when $x > x_0$.*

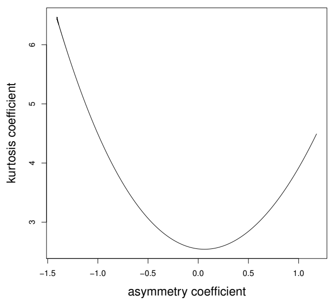

Figure 3. Graph of asymmetry coefficient versus kurtosis coefficient.

Proof. Let $\tau = (x/\beta)^\gamma$. The first derivative of h can be written as follows

$$h'(x;\beta,\gamma) = \frac{\partial}{\partial x}\, h(x;\beta,\gamma) = \frac{\gamma}{\beta^\gamma}\,(\gamma-1)\,(e^\tau-1)\,x^{\gamma-2} + \frac{\gamma^2}{\beta^{2\gamma}}\,e^\tau x^{2(\gamma-1)}.$$

Part (a) is evident from the expression of h'. To prove part (b), notice that

$$h'(x;\beta,\gamma) = \frac{\gamma}{\beta^\gamma}\,(1-\gamma)\,(e^\tau-1)\,\{c(x)-1\}x^{\gamma-2}$$

with $c(x) = \dfrac{\gamma}{\beta^\gamma(1-\gamma)}\,\dfrac{e^\tau}{e^\tau-1}\,x^\gamma$. Since for any $x > 0$, $\beta > 0$ and $0 < \gamma < 1$, we have

$$c'(x) = \frac{\partial}{\partial x}\, c(x) = \frac{\gamma^2}{\beta^\gamma(1-\gamma)}\,\frac{e^\tau}{(e^\tau-1)^2}(e^\tau-\tau-1)x^{\gamma-1} > 0,$$

$\lim_{x\to 0} c(x) = 0$, $\lim_{x\to\infty} c(x) = +\infty$ and $\dfrac{\gamma}{\beta^\gamma}\,(1-\gamma)\,(e^\tau-1)\,x^{\gamma-2} > 0$, then there exists an $x_0 = x_0(\beta,\gamma) > 0$ so that $h'(x;\beta,\gamma) < 0$ for $x < x_0$ and $h'(x;\beta,\gamma) > 0$ for $x > x_0$. This proves part (b). □

Therefore, the PM(β,γ) distribution has either increasing failure rate (IFR) when $\gamma \geq 1$ or bathtub-shaped failure rate when $0 < \gamma < 1$.

2.5 Generalized failure rate

Lariviere and Porteus [22] introduced the concept of generalized failure rate (GFR) of a continuous non-negative random variable and they showed that the distributions with increasing GFR (IGFR) have useful applications in operations management (see also Lariviere [21]). The GFR of a random variable X is defined by $g(x) = xh(x)$, where $h(x)$ denotes the failure rate function. It is said that X is IGFR if $g(x)$ is non-decreasing. The next proposition shows that the PM distribution satisfies this desirable property.

Proposition 5. *The* PM(β,γ) *distribution is IGFR.*

Proof. The first derivative of the GFR of the PM distribution can be written as follows

$$g'(x;\beta,\gamma) = \frac{\partial}{\partial x}\, g(x;\beta,\gamma) = \frac{\gamma^2}{\beta^\gamma}\, x^{\gamma-1}\left(\left(\frac{1}{\beta^\gamma}\, x^\gamma + 1\right) e^{(x/\beta)^\gamma} - 1\right)$$

and clearly $g'(x;\beta,\gamma) > 0$ since $x > 0$, $\beta > 0$ and $\gamma > 0$. Hence, $g(x;\beta,\gamma)$ is a non-decreasing function in x, which implies the result. □

2.6 Mean residual life

Another important reliability measure for non-negative random variables is the mean residual life, which is defined as $E(X - x|X > x)$. It is well-known that if a random variable has IFR, then its mean residual life is decreasing, which occurs when $\gamma \geq 1$ for the PM(β,γ) law as a consequence of Proposition 4. An analytical expression for the mean residual life is given in the following result.

Proposition 6. *The mean residual life of the* PM(β,γ) *distribution is*

$$\begin{aligned} r(x;\beta,\gamma) &= \frac{\beta}{\gamma}\,\{\log u(x)\}^{\frac{1}{\gamma}-1}\, I(x) && (2.6)\\ &= \frac{\beta}{\gamma}\,\Gamma\left(\tfrac{1}{\gamma}\right)\{\log u(x)\}^{\frac{1}{\gamma}-1}\exp\{u(x)\}\, E_0^{\frac{1}{\gamma}-1}(u(x)) && (2.7)\end{aligned}$$

with

$$I(x) = \int_0^\infty \{\log(1+z)\}^{\frac{1}{\gamma}-1}\exp\{-zu(x)\}\,dz, \quad u(x) = \exp\{(x/\beta)^\gamma\}.$$

Proof. Let $S(x;\beta,\gamma) = 1 - F(x;\beta,\gamma)$. The mean residual life is

$$\begin{aligned} r(x;\beta,\gamma) &= E(X-x|X>x) = \int_x^\infty \frac{S(y;\beta,\gamma)}{S(x;\beta,\gamma)}\,dy\\ &= \frac{1}{S(x;\beta,\gamma)}\int_x^\infty \exp\left\{\left(\frac{y}{\beta}\right)^\gamma - \left(e^{(y/\beta)^\gamma} - 1\right)\right\}dy\\ &= \frac{\beta}{\gamma\, u(x)}\int_{u(x)}^\infty \log(t)^{\frac{1}{\gamma}-1}\exp\{u(x) - t\}\,dt,\end{aligned}$$

where in the last equality we have made the change of variable $t = \exp\{(y/\beta)^\gamma\}$. Then, the result in (2.6) follows by making the change of variable $v = t - u(x)$. The result in (2.7) is obtained from (2.6) by taking into account (2.2). □

Figure 4 displays the mean residual life of the PM(β,γ) distribution for $\beta = 1$ and several values of γ. As can be seen, for $\gamma \in (0,1)$ the mean residual life is first increasing up to a point and then decreasing, whereas for $\gamma \geq 1$ it is always decreasing.

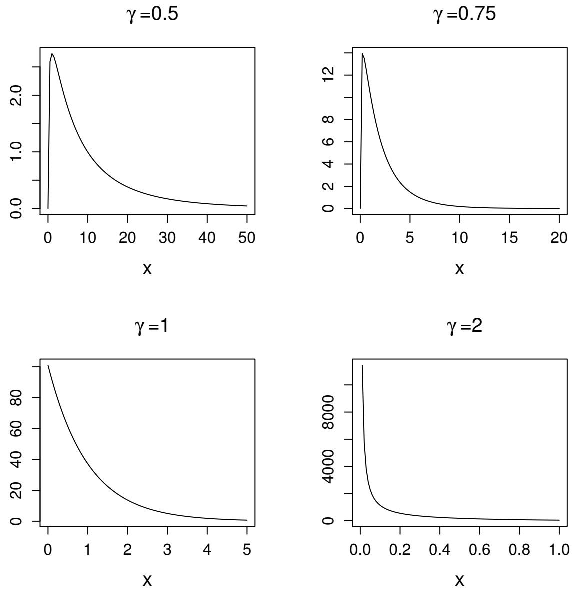

Figure 4. Mean residual life for $\beta = 1$ and $\gamma = 0.5, 0.75, 1, 2$.

2.7 Moments of order statistics

This section shows that the moments of the order statistics can be expressed in closed form in terms of the generalized integro-exponential function.

Let $X_1, \ldots, X_n$ be a random sample of size n from the PM(β, γ) distribution. Let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ be the order statistics obtained by arranging X_i, $i = 1, \ldots, n$, in non-decreasing order of magnitude. For any $n = 1, 2, \ldots$ and $k = 1, 2, \ldots$, the kth moment of $X_{r:n}$, $r = 1, \ldots, n$, can be computed using the well-known formula (see Arnold et al. [4, p. 108])

$$E[X_{r:n}^k] = r\binom{n}{r}\int_0^\infty x^k \left\{F(x;\beta,\gamma)\right\}^{r-1}\left\{1 - F(x;\beta,\gamma)\right\}^{n-r} dF(x;\beta,\gamma), \quad (2.8)$$

where F is defined by (1.2). We state the following.

Proposition 7. *For any $n = 1, 2, \ldots$, the moments of the smallest order statistic $X_{1:n}$ are given by*

$$E[X_{1:n}^k] = e^n \beta^k \Gamma\left(\frac{k}{\gamma} + 1\right) E_{1-n}^{\frac{k}{\gamma}-1}(n), \qquad k = 1, 2, \ldots.$$

Proof. From Eqs. (1.2) and (2.8), together with the change of variable $u = \exp\{(x/\beta)^\gamma\}$, we have

$$\begin{aligned} E[X_{1:n}^k] &= n e^n \beta^k \left\{\int_1^\infty (\log u)^{k/\gamma} u^n e^{-nu} du - \int_1^\infty (\log u)^{k/\gamma} u^{n-1} e^{-nu} du\right\} \\ &= n e^n \beta^k \Gamma\left(\frac{k}{\gamma} + 1\right)\left\{E_{-n}^{k/\gamma}(n) - E_{1-n}^{k/\gamma}(n)\right\}, \end{aligned}$$

where in the last equality we have used Eq. (2.2). The result is obtained by taking into account in the above equation the recurrence formula (2.4). □

The expression of $E[X_{1:n}^k]$ in Proposition 7 can be used to evaluate $E[X_{r:n}^k]$, for $r = 2, \ldots, n$ and $k = 1, 2, \ldots$, computing only the moments of the smallest order statistic in samples of size $r = n - r + 1, \ldots, n$. With this aim, the following well-known formula can be applied (see Arnold et al. [4, p. 113])

$$E[X_{r:n}^k] = \sum_{j=n-r+1}^{n} (-1)^{j-(n-r+1)} \binom{n}{j} \binom{j-1}{n-r} E[X_{1:j}^k], \qquad r = 2, \ldots, n.$$

3 Parameter estimation

This section considers the parameter estimation problem. Subsection 3.1 describes the maximum likelihood (ML) method. Subsection 3.2 gives some practical advice to calculate the ML estimators. Finally, some simulation results are presented in Subsection 3.3.

3.1 The maximum likelihood method

Let $X_1, \ldots, X_n$ be a random sample of size n from the PM(β, γ) distribution with unknown parameters β and γ. Denote by x_1, x_2, …, x_n the observed values of the sample. From the likelihood function, $L(\beta, \gamma) = \prod_{i=1}^{n} f(x_i; \beta, \gamma)$, the log-likelihood function can be written as follows

$$\begin{aligned} \log L(\beta, \gamma) = n(\log \gamma - \gamma \log \beta) + (\gamma - 1) \sum_{i=1}^{n} \log x_i + \frac{1}{\beta^\gamma} \sum_{i=1}^{n} x_i^\gamma \\ + \sum_{i=1}^{n} \log \left\{ e^{(x_i/\beta)^\gamma} - 1 \right\} - \sum_{i=1}^{n} \left(e^{(x_i/\beta)^\gamma} - 1 \right). \end{aligned} \tag{3.1}$$

The ML estimates of β, γ are the values $\hat{\beta}, \hat{\gamma}$ that maximize Eq. (3.1). The system of partial derivatives of $\log L(\beta, \gamma)$ set equal to zero is the following

$$\begin{aligned} \frac{\partial}{\partial \beta} \log L(\beta, \gamma) &= -\frac{\gamma}{\beta^{\gamma+1}} \sum_{i=1}^{n} \frac{x_i^\gamma e^{(x_i/\beta)^\gamma}}{e^{(x_i/\beta)^\gamma} - 1} + \frac{\gamma}{\beta^{\gamma+1}} \sum_{i=1}^{n} x_i^\gamma e^{(x_i/\beta)^\gamma} \\ &\quad - \frac{\gamma}{\beta^{\gamma+1}} \sum_{i=1}^{n} x_i^\gamma - \frac{n\gamma}{\beta} = 0, \\ \frac{\partial}{\partial \gamma} \log L(\beta, \gamma) &= \frac{1}{\beta^\gamma} \sum_{i=1}^{n} \frac{x_i^\gamma \log\left(\frac{x_i}{\beta}\right) e^{(x_i/\beta)^\gamma}}{e^{(x_i/\beta)^\gamma} - 1} - \frac{1}{\beta^\gamma} \sum_{i=1}^{n} x_i^\gamma \log\left(\frac{x_i}{\beta}\right) e^{(x_i/\beta)^\gamma} \\ &\quad + \frac{1}{\beta^\gamma} \sum_{i=1}^{n} x_i^\gamma \log\left(\frac{x_i}{\beta}\right) + \sum_{i=1}^{n} \log\left(\frac{x_i}{\beta}\right) + \frac{n}{\gamma} = 0. \end{aligned} \tag{3.2}$$

Clearly, this system does not have an explicit solution, so to get $\hat{\beta}$ and $\hat{\gamma}$ it is preferable to maximize (3.1).

3.2 Some practical considerations

As noted in the previous subsection, in order to obtain the ML estimates the following optimization problem is solved

$$\begin{aligned} &\max \ \log L(\beta,\gamma), \\ &\text{s.t.} \quad \beta > 0,\ \gamma \geq 0.5, \end{aligned} \tag{3.3}$$

where $\log L(\beta,\gamma)$ is given in (3.1). In our simulations, problem (3.3) was solved by using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, since (3.3) has linear constraints in the parametric space. The BFGS algorithm is available, for example, in the function `constrOptim` of the R language [29]. A starting point in the parametric space together with the gradient function of $\log L(\beta,\gamma)$ must be supplied; the latter is given in (3.2). To determine a suitable starting point, from the simulation experiments carried out, we propose to solve with respect to β and γ the following system of equations

$$F^{-1}(u_i;\beta,\gamma) = q_i, \quad F^{-1}(u_j;\beta,\gamma) = q_j,$$

for $(i,j) = (1,2),\ (1,3),\ (2,3)$ with $u_1 = 1/4$, $u_2 = 1/2$, $u_3 = 3/4$ and the sample quartiles q_i, $i = 1,2,3$. Then, take as starting point the solution with the highest value of $\log L(\beta,\gamma)$.

3.3 Numerical results

The performance of the ML method was assessed via a Monte Carlo simulation study. To this end, we generated $N = 10,000$ random samples of different sizes n for several values of β and γ. Pseudo-random data from the PM distribution were computer-generated by means of (2.5). In the simulation study, the following quantities were calculated for the simulated estimates $\hat{\beta}_j$, $j = 1,\ldots,N$.
(i) The mean: $\bar{\beta} = (1/N)\sum_{j=1}^{N}\hat{\beta}_j$.
(ii) The bias: $\text{Bias}(\hat{\beta}) = \bar{\beta} - \beta$.
(iii) The variance: $\text{Var}(\hat{\beta}) = (1/N)\sum_{j=1}^{N}\hat{\beta}_j^2 - \bar{\beta}^2$.
(iv) The mean-square error: $\text{MSE}(\hat{\beta}) = (1/N)\sum_{j=1}^{N}(\hat{\beta}_j - \beta)^2$.
The analogous quantities for γ were also calculated.

Table 1 presents some simulation results where the true values of the parameters are $\beta = 10$ and $\gamma = 0.5, 1, 3$. The bias, variance and mean-square errors are multiplied by 10^3.

Looking at this table, it can be concluded that the ML method provides acceptable estimates of the parameters. As expected, the bias, variance and mean-square error decrease as n increases.

4 Real data analysis

In this section, we consider two real data sets previously analyzed in the literature. The results of fitting the PM distribution to both sets are compared to the ones provided by other probability models formerly used to this end.

Table 1. ML parameter estimates.

	$\beta = 10$				$\gamma = 0.5$			
n	$\bar{\beta}$	Bias($\hat{\beta}$)	Var($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\gamma}$	Bias($\hat{\gamma}$)	Var($\hat{\gamma}$)	MSE($\hat{\gamma}$)
50	10.0891	89.1242	1859.7786	1867.7217	0.5148	14.8965	3.6490	3.8710
100	10.0465	46.5860	933.8444	936.01470	0.5075	7.5661	1.7267	1.7839
200	10.0188	18.8566	464.9652	465.32082	0.5039	3.9197	0.8351	0.8504
500	9.9996	-0.3416	184.5562	184.5563	0.5013	1.3391	0.3252	0.3270
1000	10.0035	3.5800	98.4787	98.4915	0.5007	0.7701	0.1662	0.1668
	$\beta = 10$				$\gamma = 1$			
n	$\bar{\beta}$	Bias($\hat{\beta}$)	Var($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\gamma}$	Bias($\hat{\gamma}$)	Var($\hat{\gamma}$)	MSE($\hat{\gamma}$)
50	9.9984	-1.5960	465.7295	465.7320	1.0277	27.7550	15.1706	15.9409
100	10.0057	5.7468	237.1530	237.1860	1.0152	15.2591	7.0338	7.2666
200	9.9991	-0.8303	116.2340	116.2347	1.0066	6.6421	3.3466	3.3907
500	10.0011	1.1851	46.7260	46.7274	1.0030	3.0313	1.3098	1.3189
1000	9.9996	-0.3374	23.1173	23.1174	1.0014	1.4786	0.6479	0.6501
	$\beta = 10$				$\gamma = 3$			
n	$\bar{\beta}$	Bias($\hat{\beta}$)	Var($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\gamma}$	Bias($\hat{\gamma}$)	Var($\hat{\gamma}$)	MSE($\hat{\gamma}$)
50	9.9990	-0.9099	51.6289	51.6297	3.0912	91.2156	132.4955	140.8158
100	10.0017	1.7169	25.5372	25.5402	3.0443	44.3853	61.4293	63.3993
200	10.0011	1.1989	12.8004	12.8018	3.0221	22.1501	29.6718	30.1624
500	10.0001	0.1717	5.2081	5.2082	3.0089	8.9863	11.5848	11.6656
1000	10.0003	0.3491	2.6170	2.6172	3.0038	3.8816	5.9737	5.9888

4.1 Data set 1

The first real data set corresponds to the breaking stress of carbon fibres (in Gba) and it was studied in Cordeiro and Lemonte [7]. The $n = 66$ data values are: 3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 3.56, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 1.57, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.89, 2.88, 2.82, 2.05, 3.65, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.35, 2.55, 2.59, 2.03, 1.61, 2.12, 3.15, 1.08, 2.56, 1.80, 2.53. In [7], these data were fitted by using the Birnbaum–Saunders (BS) and β-Birnbaum–Saunders (β-BS) distributions.

We fitted the PM distribution. Table 2 shows the ML estimated parameters. The correlation coefficient between the theoretical and the empirical cumulative probabilities is 0.9935. The results of the PM fitting were compared to the ones provided by the BS and β-BS distributions. As in [7], we calculated the Akaike information criterion AIC (see Akaike [2]) and the Bayesian information criterion BIC (see Schwarz [27]), which are defined as $\mathrm{AIC} = 2r - 2\log L$ and $\mathrm{BIC} = -2\log L + r\log n$, respectively, where r is the number of parameters and L denotes the maximized value of the likelihood function. The AIC and BIC values for each fitted distribution are given in Table 2 and the lowest values were obtained for the PM model.

We applied the following goodness-of-fit tests based on the empirical cdf: the

Table 2. Breaking stress of carbon fibres: Model, ML estimates, AIC and BIC values.

Model	ML estimates	AIC	BIC
$PM(\beta,\gamma)$	$\hat{\beta}=2.810$, $\hat{\gamma}=1.394$	176.11	180.49
β-$BS(\alpha,\beta,a,b)$	$\hat{\alpha}=1.044$, $\hat{\beta}=57.600$, $\hat{a}=0.193$, $\hat{b}=1876.732$	190.71	199.47
$BS(\alpha,\beta)$	$\hat{\alpha}=0.437$, $\hat{\beta}=2.515$	204.38	208.75

Cramér–von Mises statistic W^2, the Watson statistic U^2, the Anderson–Darling statistic A^2 and the Kolmogorov–Smirnov statistic D. A detailed definition together with simple formulae for computing these statistics can be found in D'Agostino and Stephens [9, Chapter 4]. We also applied two tests based on the empirical characteristic function [14, 15] by using the integral transformation, as proposed in Meintanis et al. [23], taking as weight functions: the standard normal law, FC_1, and the pdf $w(t) = \{1 - \cos(t)\}/\pi t^2$, which is the choice recommended in Epps and Pulley [10] (see also Section 4 in [15]), FC_2. To get the corresponding p-values we applied a parametric bootstrap generating 10,000 bootstrap samples (see Stute et al. [28] and Babu and Rao [5] for full details). The results are shown in Table 3. The apparent inconsistency of the p-values can be explained from the results in Janssen [13], which assert that the global power function of any nonparametric test is flat on balls of alternatives except for alternatives coming from a finite dimensional subspace. In other words, each test has a high power against a set of alternatives. This fact justifies the difference in the p-values.

Table 3. Breaking stress of carbon fibres: Goodness-of-fit tests.

	W^2	U^2	A^2	D	FC_1	FC_2
p-value:	0.1733	0.1512	0.9693	0.2138	0.9655	0.9653

From all the above results, it can be concluded that the PM distribution provides a better fit than the BS and β-BS probability models.

4.2 Data set 2

The second real data set corresponds to the stress-rupture life of Kevlar 49/epoxy strands which were subjected to constant sustained pressure at the 70% stress level until all had failed. The failure times are given in hours. The $n = 49$ values are: 1051, 1337, 1389, 1921, 1942, 2322, 3629, 4006, 4012, 4063, 4921, 5445, 5620, 5817, 5905, 5956, 6068, 6121, 6473, 7501, 7886, 8108, 8546, 8666, 8831, 9106, 9711, 9806, 10205, 10396, 10861, 11026, 11214, 11362, 11604, 11608, 11745, 11762, 11895, 12044, 13520, 13670, 14110, 14496, 15395, 16179, 17092, 17568, 17568. This data set can be found in Andrews and Herzberg [3, pp. 181–186].

We fitted the PM distribution to the above data set. Table 4 shows the

ML estimated parameters. The correlation coefficient between the theoretical and the empirical cumulative probabilities is 0.9947. Table 5 shows the p-values associated with the aforementioned goodness-of-fit tests. Moreover, the PM fitting was compared with the ones obtained by using other distributions such as: BS, Gamma (G), log-logistic (LL), power Lindley (PL) (see Ghitany et al. [11]) and Weibull (W). Table 4 shows the corresponding ML estimated parameters together with the AIC and BIC values. Other two-parameter distributions were fitted to the data set, which have not been included in Table 5, such as the inverse Gaussian, Lomax, paralogistic, Gumbel and generalized exponential (see Gupta and Kundu [12]). Neither of them improved the PM fitting. Moreover, in VedoVatto et al. [31] (see also Cooray and Ananda [6]), the following distributions were fitted to the above data set: Nadarajah–Haghighi (NH), beta NH, exponentiated NH, exponentiated generalized NH, Kumaraswamy NH, Marshall-Olkin NH, modified NH, Zografos–Balakrishnan NH, exponentiated exponential and generalized power Weibull. As can be seen in [31], the AIC and BIC values obtained for these distributions are greater than the ones provided by the PM fitting.

Table 4. Stress-rupture life: Model, ML estimates, AIC and BIC values.

Model	ML estimates	AIC	BIC
PM(β, γ)	$\hat{\beta} = 8603.034$, $\hat{\gamma} = 0.850$	963.85	963.96
BS(α, β)	$\hat{\alpha} = 6806.278$, $\hat{\beta} = 0.754$	980.87	984.65
G(α, β)	$\hat{\alpha} = 0.00031$, $\hat{\beta} = 2.779$	970.27	974.05
LL(α, β)	$\hat{\alpha} = 7986.578$, $\hat{\beta} = 0.381$	978.07	981.86
PL(α, β)	$\hat{\alpha} = 0.000017$, $\hat{\beta} = 1.279$	968.75	972.53
W(α, β)	$\hat{\alpha} = 9906.049$, $\hat{\beta} = 2.015$	965.69	969.48

Table 5. Stress-rupture life: Goodness-of-fit tests.

	W^2	U^2	A^2	D	FC_1	FC_2
p-value:	0.4577	0.4519	0.1674	0.6858	0.1688	0.1663

Overall, from the above results, it can be concluded that the PM distribution may be an interesting alternative to the other models under consideration.

Acknowledgements

The authors are very grateful to the anonymous referees for their careful reading of the manuscript and valuable suggestions, which helped to notably improve the paper.

References

[1] J.A. Adell and P. Jodrá. Exact Kolmogorov and total variation distances between some familiar discrete distributions. *J. Inequal. Appl.*, art. ID 64307:1–8, 2006.

[2] H. Akaike. A new look at statistical model identification. *IEEE Trans. Autom. Control*, **19**(6):716–723, 1974. https://doi.org/10.1109/TAC.1974.1100705.

[3] D.F. Andrews and A.M. Herzberg. *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. Springer-Verlag, New York, 1985. https://doi.org/10.1007/978-1-4612-5098-2.

[4] B.C. Arnold, N. Balakrishnan and H.N. Nagaraja. *A First Course in Order Statistics*. John Wiley and Sons, Inc., New York, 1992.

[5] G.J. Babu and C.R. Rao. Goodness-of-fit tests when parameters are estimated. *Sankhyā*, **66**(1):63–74, 2004.

[6] K. Cooray and M.M.A. Ananda. A generalization of the half-normal distribution with applications to lifetime data. *Comm. Statist. Theory Methods*, **37**(9):1323–1337, 2008. https://doi.org/10.1080/03610920701826088.

[7] G.M. Cordeiro and A.J. Lemonte. The β-Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. *Comput. Stat. Data Anal.*, **55**(3):1445–1461, 2011. https://doi.org/10.1016/j.csda.2010.10.007.

[8] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey and D.E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, **5**(1):329–359, 1996. https://doi.org/10.1007/BF02124750.

[9] R.B. D'Agostino and M.A. Stephens (Eds.). *Goodness-of-Fit-Techniques*. Marcel Dekker, New York, 1986.

[10] T.W. Epps and L.B. Pulley. A test for normality based on the empirical characteristic function. *Biometrika*, **70**(3):723–726, 1983. https://doi.org/10.1093/biomet/70.3.723.

[11] M.E. Ghitany, D.K. Al-Mutairi, N. Balakrishnan and L.J. Al-Enezi. Power Lindley distribution and associated inference. *Comput. Stat. Data Anal.*, **64**:20–33, 2013. https://doi.org/10.1016/j.csda.2013.02.026.

[12] R.D. Gupta and D. Kundu. Generalized exponential distribution: Existing results and some recent developments. *J. Statist. Plann. Inference*, **137**(11):3537–3547, 2007. https://doi.org/10.1016/j.jspi.2007.03.030.

[13] A. Janssen. Global power functions of goodness of fit tests. *Ann. Statist.*, **28**(1):239–253, 2000. https://doi.org/10.1214/aos/1016120371.

[14] M.D. Jiménez-Gamero, M.V. Alba-Fernández, P. Jodrá and I. Barranco-Chamorro. An approximation to the null distribution of a class of Cramér-von Mises statistics. *Math. Comput. Simulation*, **118**:258–272, 2015. https://doi.org/10.1016/j.matcom.2014.11.011.

[15] M.D. Jiménez-Gamero, M.V. Alba-Fernández, J. Muñoz-García and Y. Chalco-Cano. Goodness-of-fit tests based on empirical characteristic functions. *Comput. Stat. Data Anal.*, **53**(12):3957–3971, 2009. https://doi.org/10.1016/j.csda.2009.06.001.

[16] P. Jodrá. A closed-form expression for the quantile function of the Gompertz–Makeham distribution. *Math. Comput. Simulation*, **79**(10):3069–3075, 2009. https://doi.org/10.1016/j.matcom.2009.02.002.

[17] P. Jodrá. Computer generation of random variables with Lindley or Poisson–Lindley distribution via the Lambert W function. *Math. Comput. Simul.*, **81**(4):851–859, 2010. https://doi.org/10.1016/j.matcom.2010.09.006.

[18] P. Jodrá. On order statistics from the Gompertz–Makeham distribution and the Lambert W function. *Math. Model. Anal.*, **18**(3):432–445, 2013. https://doi.org/10.3846/13926292.2013.807316.

[19] P. Jodrá and M.D. Jiménez-Gamero. A note on the Log-Lindley distribution. *Insur. Math. Econ.*, **71**:189–194, 2016. https://doi.org/10.1016/j.insmatheco.2016.09.005.

[20] P. Jodrá, M.D. Jiménez-Gamero and M.V. Alba-Fernández. On the Muth distribution. *Math. Model. Anal.*, **20**(3):291–310, 2015. https://doi.org/10.3846/13926292.2015.1048540.

[21] M.A. Lariviere. A note on probability distributions with increasing generalized failure rates. *Oper. Res.*, **54**(3):602–604, 2006. https://doi.org/10.1287/opre.1060.0282.

[22] M.A. Lariviere and E.L. Porteus. Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing & Service Operations Management*, **3**(4):293–305, 2001. https://doi.org/10.1287/msom.3.4.293.9971.

[23] S.G. Meintanis, M.D. Jiménez-Gamero and M.V. Alba-Fernández. A class of goodness-of-fit tests based on transformation. *Comm. Statist. Theory Methods*, **43**(8):1708–1735, 2014. https://doi.org/10.1080/03610926.2012.673673.

[24] M.S. Milgram. The generalized integro-exponential function. *Math. Comp.*, **44**(170):443–458, 1985. https://doi.org/10.1090/S0025-5718-1985-0777276-4.

[25] E.J. Muth. Reliability models with positive memory derived from the mean residual life function. In C.P. Tsokos and I. Shimi(Eds.), *The Theory and Applications of Reliability*, volume 2, pp. 401–435, New York, 1977. Academic Press, Inc.

[26] N. Ozalp and E. Bairamov. Uniform convergence and computation of the generalized exponential integrals. *J. Math. Chem.*, **49**(2):520–530, 2011. https://doi.org/10.1007/s10910-010-9756-5.

[27] G. Schwarz. Estimating the dimension of a model. *Ann. Statist*, **6**(2):461–464, 1978. https://doi.org/10.1214/aos/1176344136.

[28] W. Stute, W. González Manteiga and M. Presedo Quindimil. Bootstrap based goodness-of-fit tests. *Metrika*, **40**(1):243–256, 1993. https://doi.org/10.1007/BF02613687.

[29] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org, 2016.

[30] E. Vargo, R. Pasupathy and L. Leemis. Moment-ratio diagrams for univariate distributions. *J. Qual. Technol.*, **42**(3):276–286, 2010.

[31] T. VedoVatto, A.D.C. Nascimento, W.R. Miranda Filho, M.C.S. Lima, L.G.B. Pinho and G.M. Cordeiro. Some computational and theoretical aspects of the exponentiated generalized Nadarajah–Haghighi distribution. *Chil. J. Stat.*,To appear.