

A quasi exponential distribution

Abstract

In this paper, a quasi exponential distribution has been proposed and studied. Its statistical properties including moments, survival function, hazard rate function, mean residual life function and stochastic ordering have been discussed. Maximum likelihood estimation has been discussed for estimating the parameter of the distribution. Application of the distribution has been explained with a real lifetime data from biomedical science and its goodness of fit shows quite satisfactory over exponential distribution.

Keywords: exponential distribution, moments, hazard rate function, mean residual life function, stochastic ordering, parameter estimation, application.

Volume 7 Issue 2 - 2018

Rama Shanker,¹ Kamlesh Kumar Shukla,¹
 Shambhu Sharma,² Ravi Shanker³

¹Department of Statistics, Eritrea Institute of Technology, Eritrea

²Department of Mathematics, Dayalbagh Educational Institute, India

³Department of Mathematics, N.P University, India

Correspondence: Rama Shanker, Department of Statistics, College of Science, Eritrea Institute of Technology, Asmara, Eritrea, Email shankerrama2009@gmail.com

Received: February 23, 2018 | **Published:** March 09, 2018

Abbreviations: QED, quasi exponential distribution; PDF, probability density function; CDF, cumulative distribution function; MLE, maximum likelihood estimate; AIC, Akaike information criterion

Introduction

The exponential distribution having scale parameter θ (also known as rate parameter for failure, death, arrival etc) is defined by its probability density function (pdf) and cumulative distribution function (cdf)

$$f_1(x; \theta) = \theta e^{-\theta x}; x > 0, \theta > 0 \quad (1.1)$$

$$F_1(x; \theta) = 1 - e^{-\theta x}; x > 0, \theta > 0 \quad (1.2)$$

The exponential distribution provides a simple, elegant and closed form solution to many problems in lifetime testing and reliability studies. One reason for the wide popularity of the exponential distribution in reliability is that it is the limiting lifetime distribution of a series system of substantially similar components. Also the exponential distribution is important due to its memory loss property. However, the exponential distribution does not provide a significant fit for some real lifetime applications, where the failure rates are not constant. Epstein¹ has detailed study on exponential distribution and its role in life testing. Sato et al.² obtained a discrete exponential distribution and discussed its properties and applied the distribution to model defect count distribution in semi-conductor deposition equipment and defect count distribution per chips. Gupta & Kundu³ have introduced generalized exponential distribution and discussed its properties and applications. Most of the works done on exponential distribution, its extension, mixture and applications in different fields of knowledge are available in Ahsanullah & Hamedani.⁴ Although exponential distribution was the first one parameter lifetime distribution for modeling lifetime data from engineering and biomedical sciences, but due to one parameter it is not always a suitable lifetime distribution. Also the hazard rate function and the mean residual life function of exponential distribution are always constant.

The statistical modeling and analysis of lifetime data are crucial in almost all branches of knowledge especially engineering and biomedical sciences. The one parameter exponential distribution is popular in statistics literature for modeling lifetime data. It has been

observed that exponential distribution is not always a suitable model either due to theoretical or applied point of view. An attempt has been made to find a one parameter lifetime distribution which competes well with exponential distribution.

In this paper a quasi exponential distribution has been proposed. Its statistical properties including shapes of pdf for varying values of parameter, moments, hazard rate function, mean residual life function, stochastic ordering have been discussed. Estimation of parameter has been discussed using maximum likelihood estimation. Finally, application to a real lifetime data from biomedical science has been presented to test its goodness of fit over exponential distribution.

A quasi exponential distribution

A quasi exponential distribution (QED) having scale parameter θ is defined by its pdf

$$f_2(x; \theta) = \frac{2\theta^{1/2}}{\Gamma\left(\frac{1}{2}\right)} e^{-\theta x^2}; x > 0, \theta > 0 \quad (2.1)$$

The survival (reliability) function of QED (2.1) can be obtained as

$$S(x; \theta) = P(X > x) = \frac{2\theta^{1/2}}{\Gamma\left(\frac{1}{2}\right)} \int_x^\infty e^{-\theta t^2} dt = \frac{\Gamma\left(\frac{1}{2}, \theta x^2\right)}{\Gamma\left(\frac{1}{2}\right)}; x > 0, \theta > 0 \quad (2.2)$$

Where the function $\Gamma(\alpha, z)$ is the upper incomplete gamma function defined as

$$\Gamma(\alpha, z) = \int_z^\infty e^{-y} y^{\alpha-1} dy; \alpha > 0, z \geq 0, \quad (2.3)$$

Thus the cdf of QED (2.1) can be defined as

$$F_2(x; \theta) = 1 - S(x; \theta) = 1 - \frac{\Gamma\left(\frac{1}{2}, \theta x^2\right)}{\Gamma\left(\frac{1}{2}\right)}; x > 0, \theta > 0 \quad (2.4)$$

The behavior of the pdf and cdf of QED and exponential distribution (ED) for varying values of parameter have been shown in figures 1 and 2 respectively.

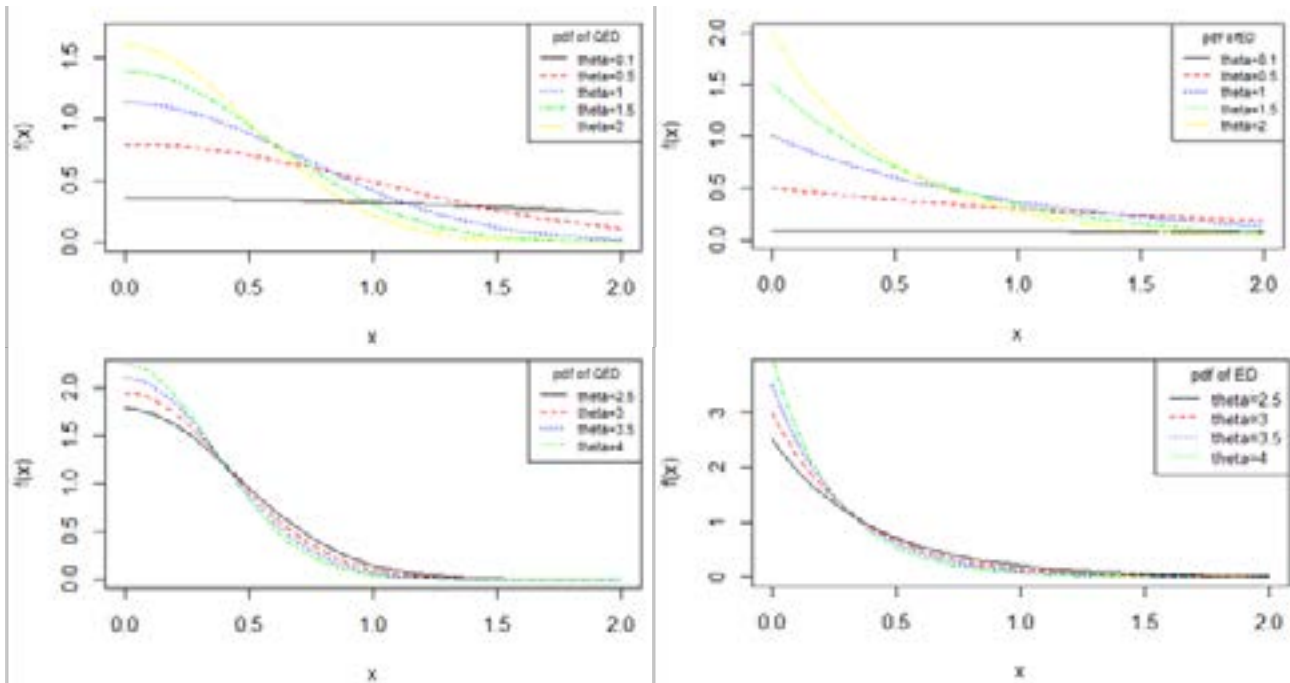


Figure 1 Behavior of the pdf of the QED and ED for varying values of parameter θ .

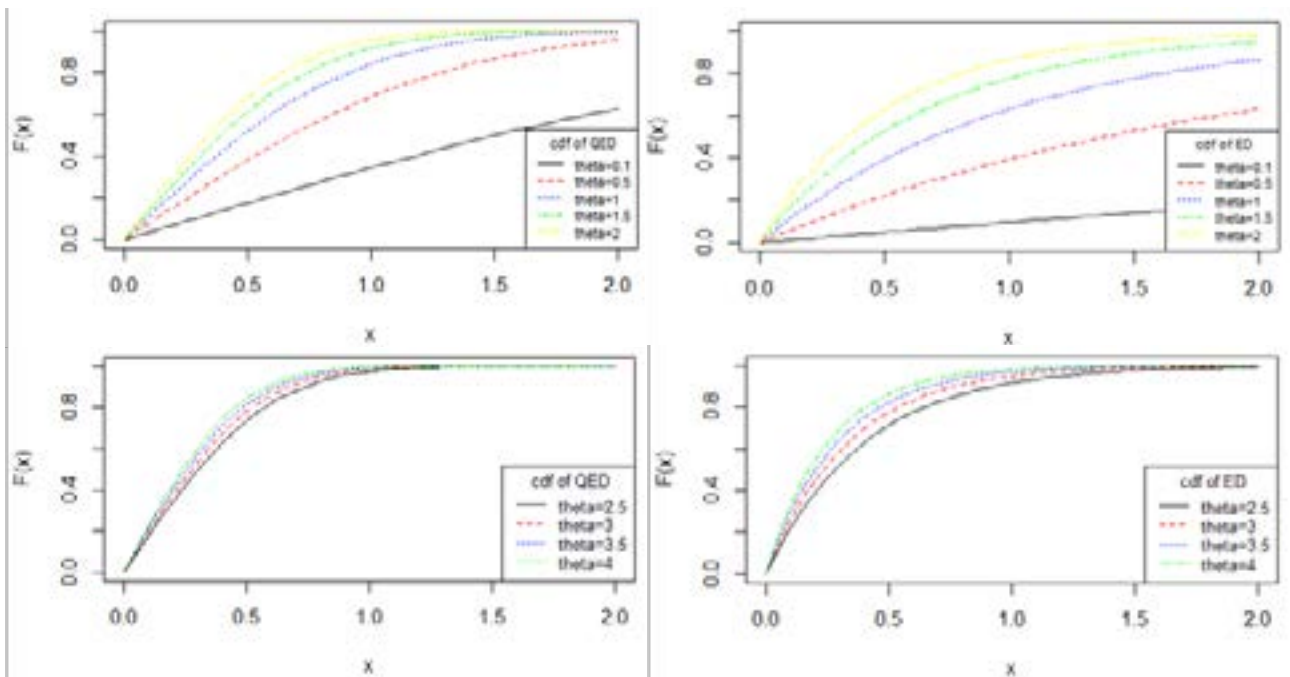


Figure 2 Behavior of the cdf of the QED and ED for varying values of parameter θ .

Moments

The r^{th} moment about origin of QED can be obtained as

$$\mu'_r = E(X^r) = \frac{2\theta^{1/2}}{\Gamma\left(\frac{1}{2}\right)} \int_0^\infty e^{-\theta x^2} x^r dx = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\theta^{r/2} \Gamma\left(\frac{1}{2}\right)}; r = 1, 2, 3, \dots \quad (3.1)$$

Taking $r = 1, 2, 3$, and 4 in (3.1), the first four moments about origin of QED are obtained as

$$\mu'_1 = \frac{1}{\theta^{1/2} \Gamma\left(\frac{1}{2}\right)}, \mu'_2 = \frac{1}{2\theta}, \mu'_3 = \frac{1}{\theta^{3/2} \Gamma\left(\frac{1}{2}\right)} \text{ and } \mu'_4 = \frac{3}{4\theta^2}.$$

The variance of QED can be obtained as

$$\mu_2 = \frac{\left(\Gamma\left(\frac{1}{2}\right)\right)^2 - 2}{2\theta\left(\Gamma\left(\frac{1}{2}\right)\right)^2}.$$

The third and fourth central moments are not being given here because their expressions are lengthy. However, they can be calculated using relationship $\mu_r = E(Y - \mu_1)^r = \sum_{k=0}^r \binom{r}{k} \mu_k' (-\mu_1')^{r-k}$ between central moments and moments about mean.

Hazard rate function and mean residual life function

Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. The hazard rate function (also known as the failure rate function) and the mean residual life function of X are respectively defined as

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \quad (4.1)$$

$$\text{and } m(x) = E[X - x | X > x] = \frac{1}{1 - F(x)} \int_x^\infty [1 - F(t)] dt$$

$$= \frac{1}{S(x)} \int_x^\infty t f(t) dt - x \quad (4.2)$$

The corresponding hazard rate function, $h(x)$ and the mean residual life function, $m(x)$ of QED (2.1) are obtained as

$$h(x) = \frac{2\theta^{1/2} e^{-\theta x^2}}{\Gamma\left(\frac{1}{2}, \theta x^2\right)} \quad (4.3)$$

and

$$\begin{aligned} m(x) &= \frac{1}{S(x; \theta)} \int_x^\infty t f_2(t; \theta) dt - x = \frac{\Gamma(1, \theta x^2)}{\theta^{1/2} \Gamma\left(\frac{1}{2}, \theta x^2\right)} - x \\ &= \frac{e^{-\theta x^2}}{\theta^{1/2} \Gamma\left(\frac{1}{2}, \theta x^2\right)} - x \quad (4.4) \end{aligned}$$

The hazard rate function and the mean residual life function of the QED show flexibility over exponential distribution because in case of exponential distribution these measures are always constant. The behavior of $h(x)$ and $m(x)$ of QED for varying values of parameter are shown in Figure 3 & Figure 4.

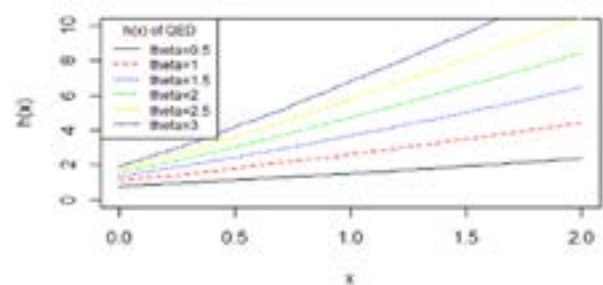


Figure 3 Behavior of the hazard rate function of QED for varying values of parameter θ .

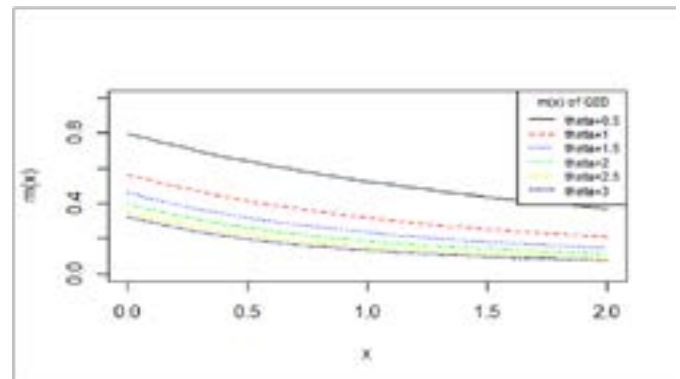


Figure 4 Behavior of the mean residual life function of QED for varying values of parameter θ .

Stochastic ordering

Stochastic ordering of positive continuous random variables is an important tool for judging their comparative behavior.⁶ A random variable X is said to be smaller than a random variable Y in the

- stochastic order ($X \leq_{st} Y$) if $F_X(x) \geq F_Y(x)$ for all x
- hazard rate order ($X \leq_{hr} Y$) if $h_X(x) \geq h_Y(x)$ for all x
- mean residual life order ($X \leq_{mrl} Y$) if $m_X(x) \leq m_Y(x)$ for all x
- likelihood ratio order ($X \leq_{lr} Y$) if $\frac{f_X(x)}{f_Y(x)}$ decreases in x .

The following results due to Shaked & Shanthikumar⁵ are well known for establishing stochastic ordering of distributions

$$\begin{aligned} X \leq_{lr} Y &\Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y \\ &\Downarrow \\ X &\leq_{st} Y \end{aligned}$$

The QED is ordered with respect to the strongest 'likelihood ratio ordering' as established in the following theorem

Theorem: Let $X \rightarrow \text{QED}(\theta_1, \alpha_1)$ and $Y \rightarrow \text{QED}(\theta_2, \alpha_2)$. If $\theta_1 > \theta_2$, then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Proof: We have

$$\frac{f_X(x; \theta_1, \alpha_1)}{f_Y(x; \theta_2, \alpha_2)} = \left(\frac{\theta_1}{\theta_2}\right)^{1/2} e^{-(\theta_1 - \theta_2)x^2}; x > 0$$

$$\text{Now, } \ln \frac{f_X(x; \theta_1, \alpha_1)}{f_Y(x; \theta_2, \alpha_2)} = \frac{1}{2} \ln \left(\frac{\theta_1}{\theta_2} \right) - (\theta_1 - \theta_2) x^2.$$

$$\text{This gives } \frac{d}{dx} \left\{ \ln \frac{f_X(x; \theta_1, \alpha_1)}{f_Y(x; \theta_2, \alpha_2)} \right\} = -2(\theta_1 - \theta_2) x.$$

Thus if $\hat{\theta}_1 > \hat{\theta}_2$, $\frac{d}{dx} \left\{ \ln \frac{f_X(x; \theta_1, \alpha_1)}{f_Y(x; \theta_2, \alpha_2)} \right\} < 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$. This shows flexibility of QED over exponential distribution.

Estimation of parameter

In this section, the estimation of parameters of QED has been discussed using method of maximum likelihood. Let (x_1, x_2, \dots, x_n) be a random sample of size n from the QED (2.1). The natural log likelihood function of QED (2.1) can be given by

$$\ln L = n \left[\ln 2 + \frac{1}{2} \ln \theta - \ln \Gamma \left(\frac{1}{2} \right) \right] - \theta \sum_{i=1}^n x_i^2.$$

The maximum likelihood estimate (MLE) $\hat{\theta}$ of the parameter θ of QED (2.1) is the solutions of log likelihood equation

$$\frac{d \ln L}{d \theta} = \frac{n}{2 \theta} - \sum_{i=1}^n x_i^2 = 0, \text{ which gives } \hat{\theta} = \frac{n}{2 \sum_{i=1}^n x_i^2}.$$

Note that the method of moment estimate $\tilde{\theta}$ of the parameter θ is also the same as the MLE of the parameter.

Goodness of fit

The goodness of fit of the QED has been explained with a real dataset from medical science. The following data represents the lifetime's data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by Gross & Clark.⁶

1.1	1.4	1.3	1.7	1.9	1.8	1.6
2.2	1.7	2.7	4.1	1.8	1.5	1.2
1.4	3.0	1.7	2.3	1.6	2.0	

For this data set, QED has been fitted along with one parameter exponential distribution.

The ML estimates, values of $-2 \ln L$, Akaike Information criteria (AIC), and K-S statistics of the fitted distributions are presented in Table 1. The AIC and K-S Statistics are computed using the following formulae:

$$\text{AIC} = -2 \ln L + 2k \text{ and } K-S = \sup_x |F_n(x) - F_0(x)|,$$

Where k = the number of parameters, n = the sample size, $F_n(x)$ is the empirical (sample) cumulative distribution function, and $F_0(x)$ is the theoretical cumulative distribution function. The best distribution is the distribution corresponding to lower values of $-2 \ln L$, AIC, and K-S statistics. It is clear from the goodness of fit in Table 1 that QED gives better fit than exponential distribution and hence it can be considered an important lifetime distribution to model lifetime data from medical science.

Table 1 Summary of ML estimates, $-2 \ln L$, AIC, and K-S of the fitted distributions

Model	ML estimate	S.E	- 2 log L	AIC	K-S
QED	0.122518	0.03874	57.16	59.16	0.461
Exponential	0.526314	0.11768	65.67	67.67	0.471

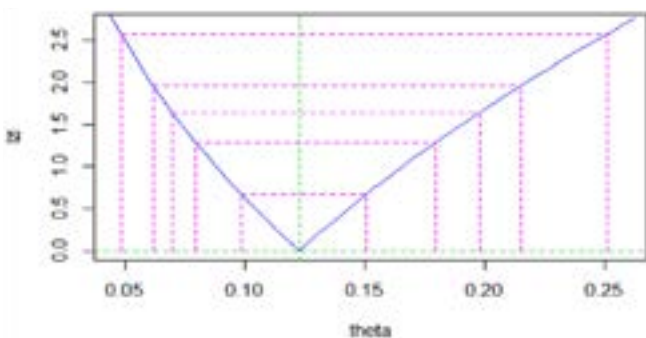


Figure 5 Profiles of the likelihood estimates of $\hat{\theta}$ of QED for the considered dataset.

The profiles of the likelihood estimate $\hat{\theta}$ of QED for the given dataset is presented in Figure 5. The fitted plots of the considered distributions for the given dataset are presented in Figure 6.

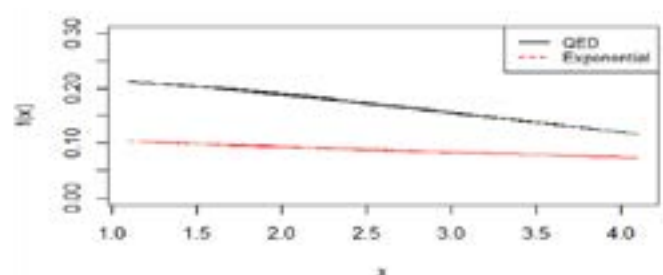


Figure 6 Fitted plots of the quasi exponential and exponential distributions for the given dataset.

Concluding remarks

This paper proposes a quasi exponential distribution for modeling real lifetime data from medical science. Its statistical properties including moments, survival function, hazard rate function, mean residual life function and stochastic ordering have been studied.

The estimation of parameter has been discussed using maximum likelihood. Profile of the likelihood estimate of the parameter for QED has been shown graphically. Application of the distribution has been explained with a real lifetime data from biomedical science and its goodness of fit shows quite satisfactory over exponential distribution.

Acknowledgements

None.

Conflicts of interest

Authors declares there is no conflict of interest.

References

1. Epstein B. The exponential distribution and its role in life testing. *Industrial Quality control*. 1958;15:2–7.
2. Sato H, Ikota M, Aritoshi S, et al. A new defect distribution in meteorology with consistent discrete exponential formula and its applications. *IEEE Trans Semicond Manufactur*. 1999;12(4):409–418.
3. Gupta RD, Kundu D. Generalized Exponential Distribution, *Austalian & New Zealand Journal of Statistics*. 1999;41(2):173–188.
4. Ahsanullah M, Hamedani GG. Exponential Distribution-Theory and Methods. New York: Nova Science Publishers, Inc.; 2010.
5. Shaked M, Shanthikumar JG. Stochastic Orders and Their Applications. New York: Academic Press; 1994.
6. Gross AJ, Clark VA. Survival Distributions: Reliability Applications in the Biometrical Sciences. New York: John Wiley; 1975.