International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

The Exponentiated-Epsilon Distribution: its Properties and Applications

Gongsin Isaac Esbond[1], Saporu W. O. Funmilayo[2]

[1]Department of Mathematical Sciences, University of Maiduguri, P.M.B. 1069, Bama Road, Maiduguri, Borno State, Nigeria

[2]National Mathematical Centre, Abuja-Lokoja Expressway, Kwali-Sheda, P. M. B. 118, Abuja, Nigeria

Abstract: ***A new lifetime distribution, called the Exponentiated-epsilon distribution, is generated. It is based on the Exponentiated-G family of distributions. It has a flexible density function that is skewed depicting its ability to model data generating processes with varying complexities. It has bathtub-shaped and J-shaped hazard rate function within its parameter space. It is applied to two differentreal-life datasets and provided good fit. Consequently, it holds a good potential for modeling real-life data generation processes.***

Keywords: epsilon distribution, exponentiated-G family of distributions, fatigue life data, fracture toughness data, hazard rate function, order statistics

1. Introduction

Generating probability distributions is a novel activity, with new distributions emerging almost on daily basis. This is a good development in statistical practice because it will create an avenue for the choice of better models among many to handle a problem situation. Nowadays, dynamic system modelers have a plethora of very similar probability density functions to fit to datasets, and then choose those densities that are adjudged the best by a combination of the many goodness-of-fit methods and model selection criteria.

There are several methods of generating probability density functions in use. Lai [7] provided a list of them; for example, probability integral transform (PIT); linear, inverse or log transformation; transformation of the cumulative distribution function or the survival function; adding a constant to existing hazard rate function; method of compounding; adding a frailty or tilt parameter; exponentiation and double exponentiation. These methods generate useful distributions that are applied in reliability analysis, dynamic system simulation and modeling. Although some of the newly generated probability distributions do not have closed-form expressions for parameter estimates in terms of the values of the random variable, moments and other important properties, the development in, and existence of, user friendly computer softwarehas made their applications easier.

In this study, a new probability density function called exponentiated epsilon distribution (later denoted by E-epsilon distribution) is introduced, some of its properties studied and applied to real life datasets.

2. Literature Review

One of the early works that propelled the generation of probability distributions through exponentiation was carried out in the first half of the nineteenth century. Gompertz[4] and Verhulst [16] introduced the popular Gompertz-Verhulst model that was used to compare known mortality tables and to represent population growth. A pioneering commendation to the method of exponentiation is credited to Lehmann [8]. His work stirred the introduction of two probability density function generators; namely Lehmann Type I and Lehmann Type II generators. Lehmann Type I was linked directly to $Exp^c(G)$, from which cumulative distribution functionand probability density functionare generated from the root distribution G. Likewise, Lehmann Type II has the generator $Exp^c(1-G)$.

There are many distributions and density functions generated following the introduction of this procedure. One of the earlier works is the exponentiated Weibull distribution credited to Mudholkar and Srivastava [11]. They showed that the distribution has a broad class of monotone failure rates than many distributions with bathtub shapes and unimodal failure rates. Its application in reliability and survival analysis were illustrated [12].

The two parameter exponentiated Pareto distribution was introduced[5]and suggested as a model for analyzing life-time datasets. This distribution is shown to have a decreasing, upward and downward bathtub shaped failure rates, depending on the value of the shape parameter. Nadarajah and Gupta [13] introduced a very flexible family of gamma distributions, called the exponentiated gamma distribution, with the gamma distribution as a special case. The distribution generalizes the standard gamma distribution in the same way the exponentiated exponential distribution generalizes the standard exponential distribution. Statistical properties such as the hazard rate function, moment generating function, skewness, kurtosis, Shannon entropy, asymptotic distribution of the extreme order statistic, were provided. Many other distributions generated by exponentiation, and their applications, are also found in literature; for example, generalized logistic[9], generalized log-normal [15], exponentiated Gumbel [1] and exponentiated exponential [6] distributions.

3. The Exponentiated- Epsilon Probability Density Function

Let X be a random variable characterized by the epsilon distribution, then the cumulative distribution and probability density functions of X are given [3], respectively, by

Licensed Under Creative Commons Attribution CC BY

$$G(x) = 1 - \left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}} \quad (1)$$

and

$$g(x) = \lambda\left(\frac{d^2}{d^2-x^2}\right)\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}} \quad (2)$$

Based on Lehmann Type I family of exponentiated distributions, a random variable, X, is said to be distributed according to the exponentiated-epsilon (hereafter E-epsilon) distribution if it has a cumulative distribution function given by

$$F_X(x;\alpha,\lambda,d) = \left[1 - \left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha} \quad (3)$$

Its corresponding probability density function is given by

$$f_X(x;\alpha,\lambda,d) = \alpha\lambda\left(\frac{d^2}{d^2-x^2}\right)\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\left[1 - \left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1} \quad (4)$$

where $\alpha > 0,\ \lambda > 0,\ d > 0,\ 0 < x < d$.

Proposition

The E-epsilon probability density function given in equation (4) is unimodal.

Proof

We determine the limiting values of the density function in equation (4) at the extreme points in the range of the values of X. That is

$$\lim_{x\to 0} f_x(x;\alpha,\lambda,d) = \lim_{x\to 0}\alpha\lambda\left(\frac{d^2}{d^2-x^2}\right)\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\left[1-\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1}$$

$$= \alpha\lambda\cdot\left(\frac{d^2}{d^2-0^2}\right)\cdot\left(\frac{0+d}{d-0}\right)^{-\frac{\lambda d}{2}}\cdot\left[1-\left(\frac{0+d}{d-0}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1}$$

$$= \alpha\lambda\cdot 1\cdot 1\cdot[1-1]^{\alpha-1}$$

$$= 0$$

Also

$$_x(x;\alpha,\lambda,d) = \lim_{x\to d}\alpha\lambda\left(\frac{d^2}{d^2-x^2}\right)\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\left[1-\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1}$$

$$= \alpha\lambda\lim_{x\to d}\left(\frac{d^2}{d^2-x^2}\right)\cdot\lim_{x\to d}\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\cdot\lim_{x\to d}\left[1-\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1}$$

$$= \alpha\lambda\cdot\left(\frac{d^2}{d^2-d^2}\right)\cdot\left(\frac{d+d}{d-d}\right)^{-\frac{\lambda d}{2}}\cdot\left[1-\left(\frac{d+d}{d-d}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1}$$

$$= \alpha\lambda\cdot\infty\cdot 0\cdot[1-0]^{\alpha-1}$$

$$= 0$$

The function tends to zero at both the upper and lower limits in the range of values for which it is a true probability density function, this imply the E-epsilon density function is unimodal.

The plots of E-epsilon probability density function (4) for varying parameter values are presented in Figure 1 below.

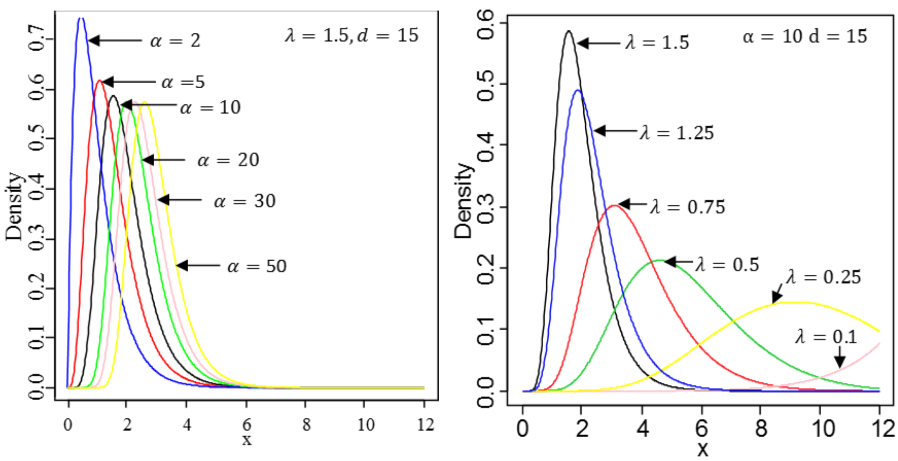

Figure 1: Density plots of E-epsilon distribution at varying values of shape parameters α and λ

4. Properties of the E-Epsilon Distribution

4.1 Quantile Function

For a random variable X characterized by the E-epsilon distribution function in equation (3), its quantile function is given by

$$q = d\,\frac{(\psi-1)}{(\psi+1)} \quad (5)$$

where $\psi = \left(1 - p^{\frac{1}{\alpha}}\right)^{-\frac{2}{\lambda d}}$ and $0 < p < 1$.

4.2 Moments

The r^{th} moment of anE-epsilon distributed random variable X is given by

$$E(X^r) = \alpha\lambda d^2\sum_{j=0}^{\alpha-1}(-1)^j\binom{\kappa}{j}\int_0^d\left(\frac{x^r}{d^2-x^2}\right)\left(\frac{x+d}{d-x}\right)^{-\omega}\partial x \quad (6)$$

where $\kappa = \alpha - 1,\ \omega = \frac{\lambda d}{2}(j+1),\ r = 1, 2, 3, \cdots$

Although the expression for the moments of anE-epsilon distributed random variable are is explicit, it can be evaluated easily when the values of the parameters are specified. This requires that the parameters be first estimated for a given dataset.

4.3 Hazard Rate Function

The hazard rate function for a random variable X distributed according to E-epsilon distribution is given by

Licensed Under Creative Commons Attribution CC BY

$$HRF(x;\alpha,\lambda) = \alpha\lambda\left(\frac{d^2}{d^2-x^2}\right)\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\frac{A}{B} \qquad (7)$$

where $A = \left[1-\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha-1}$, $B = 1-\left[1-\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}}\right]^{\alpha}$

Plots of the E-epsilon hazard rate function atvarying shape parameter values are presented in Figure 1 below.

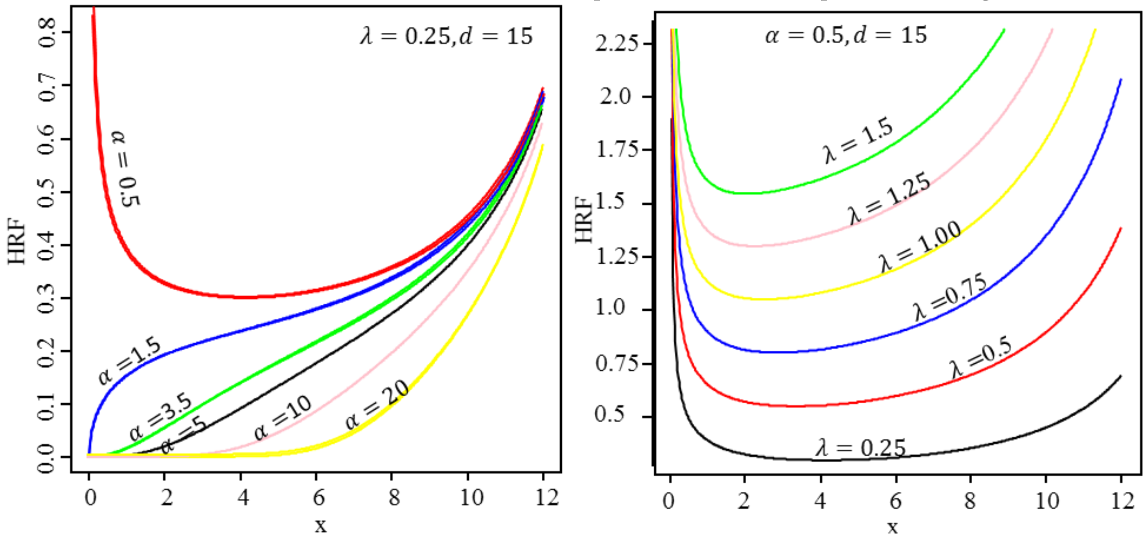

Figure 2: E-epsilon hazard rate function curves at varying values of shape parameters α and λ

4.4 Distribution of Order Statistics

Let $X_1, X_2, \ldots, X_n$ be n identically and independently distributed E-epsilon random variables, and $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be their ordered statistics. Then the distribution of the r^{th} ordered statistic is given by

$$h_{X_{(r)}}(x;\alpha,\lambda,d) = \alpha\lambda n\binom{n-1}{r-1}\left(\frac{d^2}{d^2-x^2}\right)\sum_{i=0}^{\tau}\sum_{j=0}^{\varphi}(-1)^{i+j}\binom{\tau}{i}\binom{\varphi}{j}\left(\frac{x+d}{d-x}\right)^{-\xi} \qquad (8)$$

where $\varphi = \alpha(i+r)-1$, $\tau = n-r$, $\xi = \frac{\lambda d}{2}(j+1)$, $r = 1, 2, \ldots, n$.

When $r = 1$, then equation (8) yields the distribution of the minimum of a sample of n observations from the E-epsilon distribution, given by

$$h_{X_{(1)}}(x;\alpha,\lambda,d) = \alpha\lambda n\left(\frac{d^2}{d^2-x^2}\right)\sum_{i=0}^{\tau}\sum_{j=0}^{\varphi}(-1)^{i+j}\binom{\tau}{i}\binom{\varphi}{j}\left(\frac{x+d}{d-x}\right)^{-\xi} \qquad (9)$$

where $\varphi = \alpha(i+1)-1$, $\xi = \frac{\lambda d}{2}(j+1)$, $\tau = n-1$.

And when $r = n$, we obtain from equation (8) the distribution of the largest observation in a sample of size n from the E-epsilon distribution, given by

$$h_{X_{(n)}}(x;\alpha,\lambda,d) = \alpha\lambda n\left(\frac{d^2}{d^2-x^2}\right)\sum_{i=0}^{\alpha n-1}(-1)^{i}\binom{\alpha n-1}{i}\left(\frac{x+d}{d-x}\right)^{-\frac{\lambda d}{2}(i+1)} \qquad (10)$$

5. Applications

5.1 The Data

Two real life datasets are used to illustrate the practical applications of the E-epsilon probability distribution expressed in equation (4). The first data is thefracture toughness of Alumina (Al_2O_3) (in MPa/m^2) obtained from [14]. The second dataset is the fatigue life (to the nearest thousand cycles) of 67 specimens of Alloy T7987 that failed before having accumulated 300 thousand cycles of testing. The data is obtained from [10]. These datasets are given in the Appendix.

5.2 Parameter Estimation

5.2.1 Fracture Toughness of AluminaDataset

The probability density function of the E-epsilon distribution in equation (4) was fitted to the fracture dataset using the fitdistrplus [2] package implemented in the R statistical programming language. The results are presented in Table 1 below. The plot of the density function at the estimated parameter values superimposed on the histogram for the data are presented in Figure 4 below.

Table 1: E-epsilon fit of fracture toughness of Alumina (Al_2O_3) (in MPa/m^2)

Distribution	α (se)	λ (se)	d (se)	LL value	KS (CV)	Remark
E-epsilon	13.178 (1.814)	0.588 (0.136)	6.930 (1.840)	-172.588	0.1014 (0.1247)	Good fit

se= standard error, LL = loglikelihood, KS = Kolmogorov-Smirnov stat. value, CV = critical value

Licensed Under Creative Commons Attribution CC BY

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

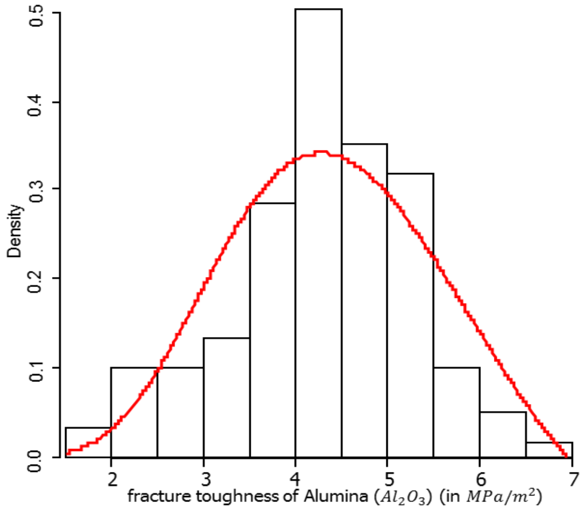

Figure 4: E-epsilondistribution plots of fracture toughness of Alumina

5.2.2 Fatigue Life Specimens of Alloy T7987Dataset

The density function was fitted to the fatigue time dataset (in thousand cycles) for specimens of Alloy T7987. The estimate of the parameters are given in Table 2 and the density plot at the estimated parameter values superimposed on the data histogram are presented in Figure5.

Table 2: E-epsilon fit of Fatigue Life Data of 67 Specimens of Alloy T7987

Distribution	α (se)	λ (se)	d (se)	LL Value	KS (CV)	Remark
E-epsilon	34.34 (2.080)	0.0238 (0.0052)	514.1 (40.2)	- 347.1	0.0546 (0.1662)	Good fit

se= standard error, LL = loglikelihood, KS = Kolmogorov-Smirnov, CV = critical value

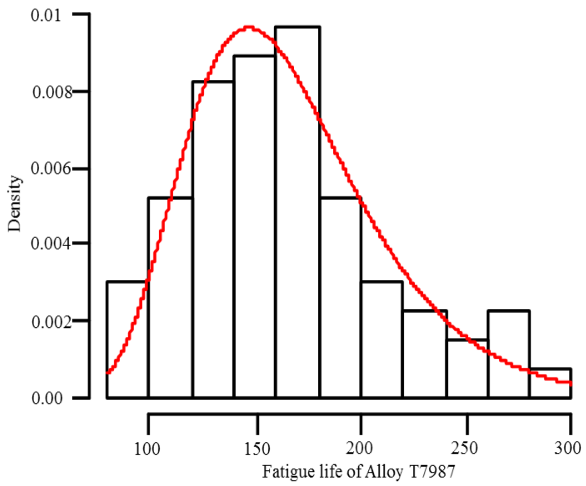

Figure 5: E-epsilon distribution plots of Alloy T7987 fatigue life (in thousand cycles)

6. Discussion

The plots of the density function of the E-epsilon distribution shown in Figure 1 suggest quite clearly that it has density function within its parameter space that can be used as a model for lifetime datasets. The bathtub and J-shaped curves of its hazard rate function shown in Figure 2 indicate clearly that it can serve as a more realistic hazard rate function model. The finite range of a random variable described according to the E-epsilon distribution, in conjunction with the bathtub and J-shapes of its explicit hazard rate function shows it is an attractive model for studying systems with finite lifetimes, for example life expectancy of humans and other biological organisms.
The results of fitting the two datasets are quite good and precise estimates of the parameters were obtained for both datasets. The average fatigue life for the specimen of Alloy T7987 is estimated as 1.6545×10^5 cycles with absolute error < 4.4e-02 These affirms the usability of the E-epsilon distribution for lifetime data analysis. Furthermore, the density plots at the estimated parameter values presented in Figures 4 and 5 atteststo the results of the goodness-of-fit tests in Tables 1 and 2, respectively.

7. Conclusion

A new distribution called the E-epsilon distribution, is introduced. It has several shapes within its parameter space that is capable of describing many data generating processes. In particular, its hazard rate function shapes suggest that it can be used as a more realistic hazard function model in life sciences. Its application to fracture toughness and fatigue lifetime (in thousand cycles) datasets produced good fit. All these suggest that this distribution has an inherent great potential for real life applications.

References

[1] Cooray, K. (2010) Generalized Gumbel Distribution. *Journal of Applied Statistics*. Volume 37, Issue 1, pp 171 – 179. https//doi.org,10.1080.02664760802698995
[2] Delignette-Muller, M. L. and Dutang, C. (2015), "fitdistrplus: An R Package for Fitting Distributions," *Journal of Statistical Software*, Volume 64, Issue 4. http://www.jstatsoft.org/
[3] Dombi, J., J´on´as, T. &T´oth, Z., E. (2018) The Epsilon Probability Distribution and its Application in Reliability Theory. *Acta PolytechnicaHungarica*, Vol. 15, No. 1, pp 197 – 216.
[4] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Statistical Society London*, vol. 115, 513 – 585.
[5] Gupta, R. C., Gupta, P. L. & Gupta, R. D. (1998) Modeling Failure Time Data by Lehmann Alternatives *Communications in. Statistical Theory and Methods.* 27, 887 – 904.

Licensed Under Creative Commons Attribution CC BY

[6] Gupta, R. D. & Kundu. D (1999) Generalized Exponential Distribution. *Australian & New Zealand Journal of Statistics*. 41, 173 – 188.
[7] Lai, C. D. (2014). Generalized Weibull Distributions. *Springer Briefs in Statistics*, DOI:10.1007/978-3-642-39106-4_2
[8] Lehmann, E.L. (1953). The power of rank tests. *The Annals of Mathematical Statistics* 24, 23–43. www.jstor.org.
[9] Masoom, A. M., Pal, M & Woo, J. (2007) Some Exponentiated Distributions. *Korean Communications in Statistics*. Vol. 14, No. 1, pp 93 – 109.
[10] Meeker, W. Q, & Escobar, L, A, (1998) *Statistical Methods for Reliability Data*. New York, Wiley and Sons. pp 149.
[11] Mudholkar, G. S. & Srivastava, D. K. (1993) Exponentiated Weibull family for analyzing bathtub failure rate data. *IEEE Transactions in Reliability*, 42, 299 – 302.
[12] Mudholkar, G. S., Srivastava, D. K. &Freiner, M. (1995) the Exponentiated Weibull family: A reanalysis of the bus-motor failure data. *Technometrics*, 37, 436 – 445.
[13] Nadarajah, S. & Gupta, A. (2007) The Exponentiated Gamma Distribution with Application to Drought Data. *Calcutta Statistical Association Bulletin, Sage Journals*.
[14] Nadarajah, S. &Kotz, S. (2008) Strength modeling using Weibull distributions. *Journal of Mechanical Science and Technology*, 22, 1247 – 1254. http://dx.doi.org/10.1007/s12206-008-0426-5
[15] Singh, B., Sharma, K. K., Rathi, S. & Singh, G. (2012) A generalized log-normal distribution and its goodness-of-fit to censored data. *Computational Statistics*. Vol. 27, Issue 1, pp 51 – 67. DOI: 10.1007/s00180-011-0233-9.
[16] Verhulst, P. F. (1838). Notice sur la loi la population suit dans son accroissement. Correspondence Mathetmatique et physique. *Publee L. A. J. Quetelet*, vol. 10, 115 – 121.

Appendix

Fracture toughness of Alumina (Al_2O_3) (in MPa/m^2) obtained from Nadarajah & Kotz (2008)

5.50, 5.00, 4.90, 6.40, 5.10, 5.20, 5.20, 5.00, 4.70, 4.00, 4.50, 4.20, 4.10, 4.56, 5.01, 4.70, 3.13, 3.12, 2.68, 2.77, 2.70, 2.36, 4.38, 5.73, 4.35, 6.81, 1.91, 2.66, 2.61, 1.68, 2.04, 2.08, 2.13, 3.80, 3.73, 3.71, 3.28, 3.90, 4.00, 3.80, 4.10, 3.90, 4.05, 4.00, 3.95, 4.00, 4.50, 4.50, 4.20, 4.55, 4.65, 4.10, 4.25, 4.30, 4.50, 4.70, 5.15, 4.30, 4.50, 4.90, 5.00, 5.35, 5.15, 5.25, 5.80, 5.85, 5.90, 5.75, 6.25, 6.05, 5.90, 3.60, 4.10, 4.50, 5.30, 4.85, 5.30, 5.45, 5.10, 5.30, 5.20, 5.30, 5.25, 4.75, 4.50, 4.20, 4.00, 4.15, 4.25, 4.30, 3.75, 3.95, 3.51, 4.13, 5.40, 5.00, 2.10, 4.60, 3.20, 2.50, 4.10, 3.50, 3.20, 3.30, 4.60, 4.30, 4.30, 4.50, 5.50, 4.60, 4.90, 4.30, 3.00, 3.40, 3.70, 4.40, 4.90, 4.90, 5.00.

Fatigue life (to the nearest thousand cycles) of 67 Specimens of Alloy T7987 that failed before having accumulated 300 thousand cycles of testing obtained from Meeker & Escobar (1998, pp 149).

94 96 99 99 104 108112 114 117 117 118 121 121 123 129 131 133 135 136 139 139 140 141 141 143 144 149 149 152 153 159 159 159 159 162 168 168 169 170 170 171 172 173 176 177 180 180 184 187 188 189 190 196 197 203 205 211 213 224 226 227 256 257 269 271 274 291

Licensed Under Creative Commons Attribution CC BY