A COMPREHENSIVE STUDY OF LIFETIME DISTRIBUTION MODELS FOR AIR CONDITIONING SYSTEMS USING FAILURE TIME DATA

S. Mythily[1], S.C. Premila[2], P. Manigandan[1], K. Geetha[1], D. Pachiyappan[3] and A. Radhika[1*]

•

[1]Department of Statistics, Periyar University, Salem-11, Tamil Nadu, India.
[2]Assistant Professor, Department of Mathematics, Saveetha Engineering College, Thandalam, Chennai.
[3]Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai.
mythilyshanmugam03@gmail.com, premilac@saveetha.ac.in, srimanigandan95@gmail.com, geethakmanju2002@gmail.com, d.pachiyappan321@gmail.com, radhisaran2004@gmail.com

Abstract

The purpose of this paper was to fit the model for the air conditioning system with failure time for the normal, gamma, and generalized gamma distributions and to identify the optimal distribution model for estimating the air conditioning system failure time. The performance of the normal, gamma, and generalized gamma distributions on air conditioning system failure data was investigated. To determine the optimal model for calculating the air conditioning system failure time, the secondary data was fitted to the normal, gamma, and generalized gamma distribution models. To evaluate the models performance, the Kolmogorov-Smirnov, and Akaike Information Criterion, were used. The outcomes of the research indicated that the gamma model had the lowest AIC and KS values, measuring 359.108 and 0.47624; the generalized gamma model came in second with AIC and KS values are 359.6058 and 0.32988; and the normal model had the highest AIC and KS values, measuring 3203.729 and 0.4298, respectively. Results show that the gamma model has better flexibility in handling real data over the conventional normal and generalized gamma distribution.

Keywords: Reliability measure, Normal, Gamma, Generalized Gamma distribution, Air Conditioning system, Failure time data.

I. Introduction

Normal, Gamma, Generalized Gamma distributions are a few of the distributions that are frequently used to represent lifetime data. Reliability is a key concept in engineering, manufacturing, and quality control that aims to design and produce products that operate reliably over time. It is the likelihood that a component, product, or system will perform on its intended function without failure under specified conditions and given time periods. When it comes to electrical products, aerospace, etc., reliability analysis will help save maintenance costs, minimize risk, and improve product quality. Dependability is critical to both consumer pleasure and their safety.

A new monitoring system for the intervals between events under exponential and gamma distributions was proposed by Shah et al. [1]. The three-parameter gamma-normal distribution was proposed by Lima et al. [2]. For two parameters, the probability density function is Pradhan and Kundu [3]. The probability density function for the three gamma distribution parameters is as follows Chen and Kotz [4]. The gamma distribution is widely used in quality and reliability engineering to fit product lifetimes since it may show growing, decreasing, and constant failure rates [5, 6]. The Gamma distributions are versatile and can be used to represent a variety of physical situations [7]. In the field of reliability, they are arguably the most often used statistical distributions [8]. Choosing the distribution that most closely matches the failure times is the analysis's task [9]. Reliability and other applications use the gamma distribution, one of the continuous probability distributions. Waiting times and service times are distributed using it [10]. According to Baran and Nemoda [11], the gamma distribution has also been demonstrated to be a suitable model in a variety of different application domains, such as ecosystems. Giving to Stacy [12], the three-parameter generalized gamma distribution is the most universal type of gamma distribution. Another kind of generalized gamma distribution was presented by Nadarajah and Gupta [13] and used to suit drought data. Using the exponentiated technique, Cordeiro et al. [14] constructed another generalization of Stacy's generalized gamma distribution and applied it to survival and life time analysis.

The exponentiated family was first presented by Ref. [15] and is a very versatile method for adding a new parameter to a continuous distribution. The exponential, gamma, Weibull, and Pareto distributions were all generalized using it, and the outcomes clearly show that this method improves the model's flexibility. This technique can be used to create generalized distributions, such as gamma, Weibull, and lognormal distributions [16], Mixtures of gamma distributions [17], generalized gamma and log gamma distributions [18].

The objective of this study was to evaluate and compare the performance of normal, gamma, and generalized gamma distributions in modeling the failure time of air conditioning systems. Specifically, the study aimed to identify the most optimal distribution model for accurately estimating the failure time of these systems by fitting secondary data to the three distribution models. The performance of each model was assessed using the Kolmogorov-Smirnov (KS) test and Akaike Information Criterion (AIC) to determine which model best represents the failure time data. Ultimately, the study sought to determine which distribution (normal, gamma, or generalized gamma) is the most effective for estimating the failure time of air conditioning systems.

The following is the presentation of the remainder of the paper. The normal, gamma, and generalized gamma distributions are introduced in Section 2. Applications of air conditioning system failure data in real-world settings are shown in Section 3. Section 4 presents the study's conclusion.

II. Methods

I. Normal Distribution

The Gaussian distribution is another name for the normal distribution. The two parameters of the normal distribution are the standard deviation (σ) and the mean (μ). The standard deviation (σ) can be used to show the variability of failure time, whereas the mean (μ) represents the average failure time. It is stated that the normal distribution's PDF equals

$$F(t) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$$

$$F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

In standard form of normal distribution

$$F(t) = \varphi\left(\frac{t-\mu}{\sigma}\right)$$

The Reliability function of Normal distribution.

$$R(t) = 1 - \varphi\left(\frac{t-\mu}{\sigma}\right)$$

And the failure rate is defined by the equation.

$$\lambda(t) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2} \quad 1 - \varphi\left(\frac{t-\mu}{\sigma}\right)\right)^{-1}$$

II. Gamma Distribution

The Gamma distribution has a continuous probability distribution. Its sole criteria are scale and shape. The scale parameter of the gamma distribution controlled the distribution's dispersion, and its form and scale parameter will represent the various failure behaviours. The gamma distribution reliability analysis is frequently used in industrial reliability analysis applications. The pdf of the gamma distribution shape parameter a = α and scale parameter b = $\frac{1}{\beta}$ is said to be

$$f(t,a,b) = \frac{b^a}{\Gamma(a)} \; t^{a-1}e^{-bt}$$

$$f(t,\alpha,\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \; t^{\alpha-1}e^{-\frac{t}{\beta}}$$

The Cumulative Distribution Function of the gamma distribution is: $F(t) = \int_0^t f(t)\, dt$

The Reliability function of Gamma Distribution. $R(t) = 1 - F\,(t)$
The failure rate is defined by.

$$h(t) = (f(t))/(R(t))$$

The two-parameter Weibull probability density function f(t) is given as

$$PDF{=}f(f) = \frac{b}{a} \times (\frac{f}{a})^{b-1} \times e^{-\left(\frac{f}{a}\right)^b}$$

$$CDF{=}F(f) = 1 - exp\left[-(\frac{f}{a})^b\right]$$

III. Generalized Gamma Distribution

We refer to the continuous probability density function as the generalised gamma. It was in 1962 when American mathematician Harbert Stacy created this distribution. The extension of the gamma distribution is known as the generalised gamma distribution. Shape parameter (k), Scale parameter (β), and another shape parameter (θ) are the three parameters in this distribution. The PDF of the generalized gamma distribution is known as

$$f(t) = \frac{\beta}{\Gamma(k).\theta}\left(\frac{t}{\theta}\right)^{k\beta-1} e^{-\frac{t}{\theta}^{\beta}}$$

Where $\theta > 0$ is scale parameter, $\beta > 0$ and $k > 0$ is shape parameters and $\Gamma(k)$ is gamma function X is defined as

$$\Gamma(x) = \int_0^{\infty} s^{x-1}.e^{-s}\,ds$$

The Reliability function of Generalized Gamma Distribution.

$$R(t) = \begin{cases} 1 - \Gamma I\left(\frac{1}{\lambda^2}; \frac{e^{\lambda\left(\frac{In(t)-\mu}{\sigma}\right)}}{\lambda^2}\right) & if\ \lambda > 0 \\ 1 - \phi\left(\frac{In(t)-\mu}{\sigma}\right) & if\ \lambda = 0 \\ \Gamma I\left(\frac{1}{\lambda^2}; \frac{e^{\lambda\left(\frac{In(t)-\mu}{\sigma}\right)}}{\lambda^2}\right) & if\ \lambda < 0 \end{cases}$$

Where $\phi(Z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Z} e^{-\frac{x^2}{2}}dx$

And $\Gamma(x)$ is the incomplete gamma function of k.
And x, is given by:

$$\Gamma I(k;x) = \frac{1}{\Gamma(k)}\int_0^x s^{k-1}e^{-s}\,ds$$

The Γ(x) is the gamma function of x.

III. Results and Discussions

I. Application

The air conditioning system failure time datasets were used to compare the sub-models with some standard models. We employed information criteria and goodness of fit techniques, which include the Kolmogorov-Smirnov (Ks), the Akaike information criterion (AIC), and the log-likelihood function assessed at the MLEs. The models used by the competitors were gamma, normal, and generalized gamma distributions. Consequently, the fit is better for the one with lower information criterion values.

$$AIC = -2LL + 2p$$

$$KS\ (D) = max\ \mid Fn(f) - F(f) \mid$$

II. Application to Failure Time Data

This section uses the dataset provided by Linhart and Zucchini [19] to assess the applicability of the proposed normal, gamma, and generalized gamma distributions. The dataset shows the frequency of cooling system failures in aeroplanes. This dataset was used in the research of several authors, including [20, 21]. Failure measurements are initially crisp values; however, for the sake of illustration, we treat data as uncertain sample values for some air condition systems and the data is given as: 23, 261, 87, 7, 121, 125, 14, 62, 47, 225, 71, 74, 246, 21, 42, 20, 5, 12, 120, 11, 13, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 55, 95.

Table 1: *Results of the scale and form parameters produced by the air conditioning system's failure time*

Distributions	Shape parameter	Scale parameter
Normal	59.14286	68.32584
Gamma	0.867236	68.19695
Generalized Gamma	1.215989	39.98209

Table 1 compares three distributions Normal, Gamma, and Generalized Gamma based on their shape and scale parameters. The Normal distribution, with a shape parameter of 59.14286 and a scale parameter of 68.32584, exhibits a symmetric spread. In contrast, the Gamma and Generalized Gamma distributions, with shape parameters of 0.867236 and 1.215989, respectively, and scale parameters of 68.19695 and 39.98209, exhibit skewness. These results highlight the differences in distributional properties, where the Normal distribution follows a symmetric pattern, while the Gamma-based distributions show asymmetry.

Table 2: *Comparison of distribution models for the air condition system failure time*

Distributions	Loglik	AIC	LRT	p-value
Normal	-197.0129	398.0259	0.54327	0.8961
Gamma	-177.554	359.108	0.47624	0.4901
Generalized Gamma	-176.8029	359.6058	1.161419	0.2812

Table 2 evaluates the fit of different distributions for failure time data using Log-likelihood (Loglik), Akaike Information Criterion (AIC), Likelihood Ratio Test (LRT), and p-values. The Normal distribution has the highest AIC (398.0259), indicating the poorest model fit. In contrast, the Gamma and Generalized Gamma distributions exhibit lower AIC values (359.108 and 359.6058, respectively), suggesting better fits. Among these, the Gamma model has the lowest AIC, making it the most suitable choice for modeling failure time data. The LRT values and p-values suggest that the improvement from the Gamma to the Generalized Gamma model is not statistically significant, reinforcing the Gamma model's adequacy.

Table 3: *Result of the KS test for the failure time of the air condition system.*

Distribution	Statistics KS	D-critic	Hypothesis
Normal	0.2158753	0.429882	It Cannot be rejected
Gamma	0.1622605	0.29832	It Cannot be rejected
Generalized Gamma	0.913989	0.32988	It Cannot be rejected

Table 3 presents the Kolmogorov-Smirnov (KS) test results for evaluating the fit of different distributions to the failure time data. The results indicate that the Gamma distribution has the lowest KS statistic (0.1622605), suggesting it provides the best fit. In contrast, the Normal and Generalized Gamma distributions have higher KS statistics, making them less favorable in comparison. The hypothesis test results confirm that none of the distributions can be rejected, meaning all three models provide an acceptable fit to the data. Additionally, the Gamma distribution has the lowest D-critical value (0.29832), further supporting its suitability for modeling failure time data. Furthermore, the CDFM estimation method is identified as the most effective approach for estimating the parameters of the Normal, Gamma, and Generalized Gamma distributions. The p-values from the KS test confirm that all three distributions can adequately describe the failure time data. Furthermore, the sample histogram and the fitted PDFs of the Normal, Gamma, and Generalized Gamma distributions using the parameter estimate findings in Table 3 are displayed in Figure 1.

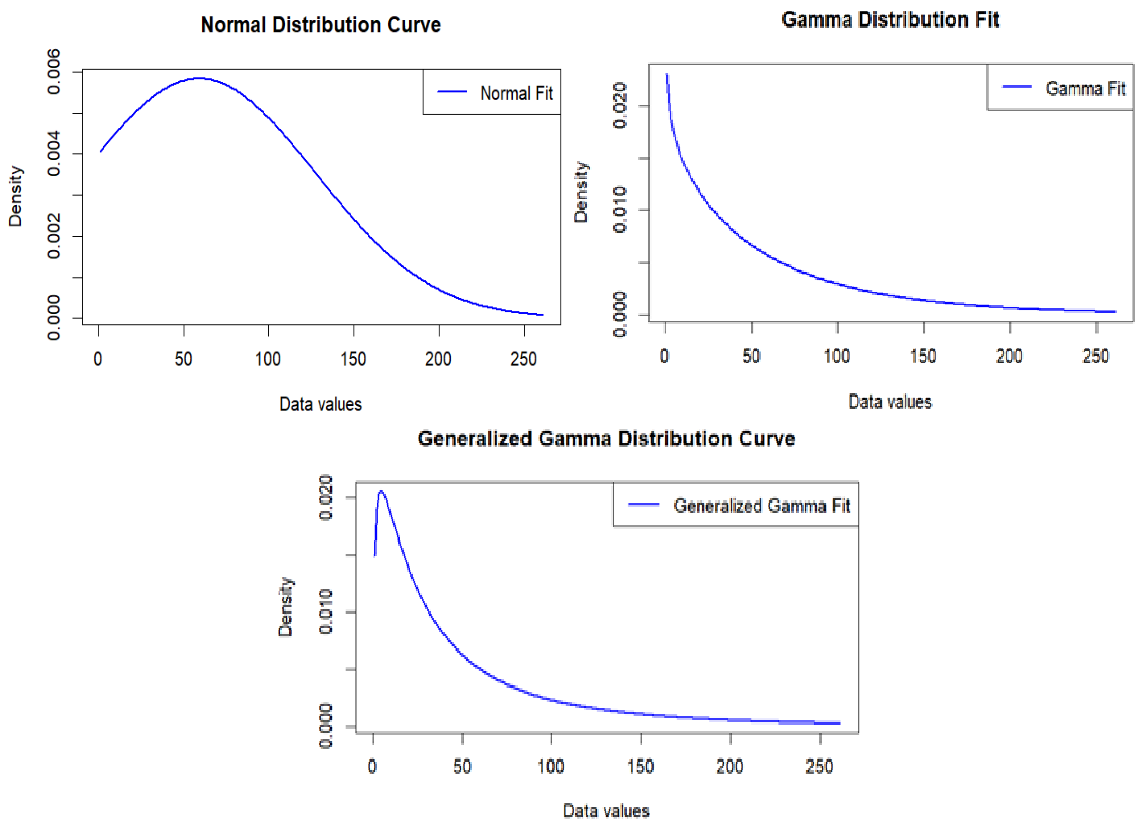

Figure 1: *Graphical presentation of failure rate functions of Normal, Gamma, Generalized Gamma Distribution*

IV. Conclusion

This study explored the challenge of selecting the most suitable distribution among the Normal, Gamma, and Generalized Gamma distributions for modeling failure time data. By analyzing their density function characteristics under different parameter values, we assessed their performance using graphical methods and various goodness-of-fit tests, including the Kolmogorov-Smirnov (KS) test, and Akaike Information Criterion (AIC). Our findings reveal that determining the best-fitting Gamma distribution becomes difficult when multiple models closely fit the data. However, the results suggest that CDFM estimations provide the most accurate fit for the dataset. Additionally, we estimated the reliability and failure rate of the air conditioning system using CDFM calculations, emphasizing the importance of valid parameter estimates in ensuring accurate predictions. The study highlights that different estimation methods can significantly influence reliability and failure rate outcomes. Moreover, as long as the data conforms to a Gamma distribution, the proposed estimation technique can be effectively applied to failure time analysis. Future research can further explore the asymptotic properties of these estimators and extend the method to other distributions. Ongoing work is focused on expanding this approach to broader applications in reliability analysis.

References

[1] Shah, M. T. Azam, M. Aslam, M. and Sherazi, U. (2021). Time between events control charts for gamma distribution. *Quality and Reliability Engineering International*, 37(2): 785-803.

[2] Maria do Carmo, S. L. and Cordeiro, G. M. (2015). A new extension of the normal distribution. *Journal of data Science*, 13(2): 385-408.

[3] Pradhan, B. and Kundu, D. (2011). Bayes estimation and prediction of the two-parameter gamma distribution. *Journal of Statistical Computation and Simulation*, 81(9): 1187-1198.

[4] Chen, W. W. and Kotz, S. (2013). The Riemannian structure of the three-parameter Gamma distribution, *Applied Mathematics*, 4(3): 28865.

[5] Piao, C. H. E. N. and Zhi-Sheng, Y. E. (2018). A systematic look at the gamma process capability indices. *European Journal of Operational Research*, 265(2): 589-597.

[6] Wang, B. X. and Wu, F. (2018). Inference on the gamma distribution. *Technometrics*, 60(2): 235-244.

[7] Zografos, K. and Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical methodology*, 6(4): 344-362.

[8] Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, 53(12): 4433-4440.

[9] Lawless, J. F. Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York, 2003.

[10] Whitt, W. (2000). The impact of a heavy-tailed service-time distribution upon the M/GI/s waiting-time distribution. *Queueing Systems*, 36(1): 71-87.

[11] Baran, S. and Nemoda, D. (2016). Censored and shifted gamma distribution-based EMOS model for probabilistic quantitative precipitation forecasting. *Environmetrics,* 27(5): 280-292.

[12] Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of mathematical statistics*, 1187-1192.

[13] Nadarajah, S. and Gupta, A. K. (2007). A generalized gamma distribution with application to drought data. *Mathematics and Computers in Simulation*, 74(1): 1-7.

[14] Cordeiro, G. M. Ortega, E. M. and Silva, G. O. (2011). The exponentiated generalized gamma distribution with application to lifetime data. *Journal of statistical computation and simulation,*

81(7): 827-842.

[15] Gupta, R. C. Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4): 887-904.

[16] Yang, L. Cheng, J. Luo, Y. Zhou, T. Zhang, X. Shi, L. and Xu, Y. (2025). Domain adaptation via gamma, Weibull, and lognormal distributions for fault detection in chemical and energy processes. *The Canadian Journal of Chemical Engineering*, 103(1): 359-372.

[17] Wiper, M. Insua, D. R. and Ruggeri, F. (2001). Mixtures of gamma distributions with applications. *Journal of computational and graphical statistics,* 10(3): 440-454.

[18] Lawless, J. F. (1980). Inference in the generalized gamma and log gamma distributions. *Technometrics,* 22(3): 409-419.

[19] Linhart, H. and Zucchini, W. Model selection. John Wiley & Sons, 1986.

[20] Magalhães, T. M. Gómez, Y. M. Gallardo, D. I. and Venegas, O. (2020). Bias reduction for the Marshall-Olkin extended family of distributions with application to an airplane's air conditioning system and precipitation data. *Symmetry*, 12(5): 851.

[21] Khan, M. S. and King, R. (2016). New generalized inverse Weibull distribution for lifetime modeling. *Communications for Statistical Applications and Methods*, 23(2): 147-161.