Imam Hussein's Arbaeen: Statistical Modeling of the Rate of Visitors Arriving in Karbala

Prof. Dr. Nadia Hashim Al-Noor
nadialnoor@uomustansiriyah.edu.iq

Prof. Dr. Faez Hassan Ali
faezhassan@uomustansiriyah.edu.iq

Mustansiriyah University / College of Science / Department of Mathematics

Abstract

By following the numerical indicators that accompanied the arbaeen visit of Imam Hussein (peace be upon him), in which the number of visitors exceeded the largest numbers of human gatherings that can occur on religious, national, sporting, and other occasions throughout the world, especially since the number of visitors for the year 2022 AD/1444 AH, as official statistics indicate, has reached more than twenty million (21198640), including (4012091) visitors from outside Iraq and from more than 150 countries around the world, it has become important to deal with these indicators with a mathematical model. In general, the mathematical model helps to explain a system or event, studying the effects of its various components, and predicting its behavior. Mathematical models may be dynamic systems, statistical models, differential equations, or theoretical models, and these types may overlap with a specific model that includes several abstract structures.

The current paper is concerned with employing four statistical models of extended exponential distribution in addition to its traditional version, and adopting multiple information criteria to determine the best model through which data related to the rate of visitors arriving to the holy Karbala Governorate to participate in performing the arbaeen pilgrimage rituals can be represented, and then extrapolating the most important relevant statistical characteristics. The results show that the best fit is the extended exponential distribution that is associated with the generator family of the truncated exponential Topp-Leone.

Keywords: Statistics of the Arbaeen Visit, Statistical Modeling, Statistical Measures.

Introduction

Every year ,the Hussein processions set out toward Karbala to renew the pledge with the Master of Martyrs, peace be upon him (PBUH), and his brother, Abu al-Fadl al-Abbas (PBUH) to continue their path and to remain steadfast in revolutionary approach to uphold truth, nullify falsehood, and fight all those who seek to falsify religion, distort its rulings, or control the people. Therefore, this great Hussein procession has been a target for tyrants to fight and oppose throughout the ages, seeking to bring down the curtain on this immortal path. Despite the harassment and the prohibition of visits, the "Scholars' Road", which wound through palm trees on the banks of the Euphrates River in the city of Najaf, was the route taken by clerics to reach the sacred shrines.

In Safar 1424 AH, the streets of Karbala were packed with visitors, the arbaeen pilgrimage recaptured its former glory, and the number of visitors continues to grow. On January 14/2012, the official institutions declared that the number of Karbala's visitors had exceeded 17 million, including about 600 thousand Arab and foreign visitors from more than 60 countries around the world. This is the world's biggest number of visitors in a short period of time, lasting no more than 20 days, beginning on 1 Safar Al-Khair 1433 AH, when people arrived at the holy shrines of Karbala, and ending on the day of the special visit on 20 Safar. This is an event that is considered the largest in many aspects, in terms of the number of those present in this short period, and in terms of the volume of free services provided by Iraq's people to all these visitors, including food, accommodation, medical services, etc. In addition, the presence of these millions of people is achieved by walking great distances of up to hundreds of kilometers, and other things that make Karbala unique among other cities in the world.

Many studies have focused on the historical review of the arbaeen pilgrimage, the reasons for its emergence, its importance and its impact as a civilizational and global starting point, in addition to documenting the epic journey of the family of Imam Hussein (PBUH) from Medina to Karbala, then to Kufa and Damascus, arriving at the return to Karbala and many others. In this regard, the interested can review (Al-Jumaili, 2023; Al-Muhanna, 2023; and Shams al-Din, 2023) at https://www.c-karbala.com/, among the series publications of the Karbala Center for Studies and Research of the Holy Shrine of Imam Hussein (PBUH), which gives special attention to the arbaeen million-person march.

Given the importance of this event, using a statistical model to address its numerical indications has become imperative. Consequently, the current paper focuses on employing five statistical distributions to represent the rate of Asian visitors arriving in the holy city of Karbala to participate in the arbaeen pilgrimage rituals. The best-fitted distribution can then be used to extrapolate the most significant relevant statistical properties. The rest of the paper is organized as follows. Section 2 constructs the methodology. Sections 3 and 4 conduct the practical application and analyze the results. Section 5 presents the conclusions and direction for future studies.

Methodology

Statistical distributions are used to model data and examine its key features. Making decisions and drawing conclusions can be aided by useful information that is produced by an appropriate distribution. One of the most important distributions is the traditional exponential (E) distribution that has been used in various fields to match life data. The E distribution has a constant hazard function and a decreasing density function, which limits its use for some recent real-life applications. Because of this restriction, researchers have the chance to improve this distribution in a number of ways to increase its adaptability through constructing new extensions or generalizations. As a result, when more flexible distributions are needed, researchers use the new extension.

In the literature, based on employing the E distribution as a baseline, numerous extensions/generalizations related to different generator (G) families of probability distributions are introduced, including beta exponential (BE), exponentiated generalized exponential (EGE), and Weibull exponential (WE), that respectively related to the beta-G (Eugene et al., 2002), exponentiated generalized-G (Cordeiro et al., 2013), and Weibull–G (Bourguignon et al., 2014).

On the other side, truncated distributions are also a key research topic in both theoretical and applied contexts. It arises in problems of probabilistic modeling in various areas. Recently, the truncated exponential Topp-Leone exponential (TETLE) distribution was introduced by Al-Noor & Hilal, 2021 as a sub-model of the flexible truncated exponential Topp-Leone-G family. For more details about various truncated versions related to different distributions, see Matsushita et al., 2023; Al-Noor & Hadi, 2021; Almarashi et al., 2020; and Aldahlan et al., 2020. The extension versions of exponential distribution are characterized by the fact that their density and hazard functions take different shapes, unlike the exponential distribution. This reflects the fact that these distributions are more flexible in handling data in some recent applications. The following sub-section will shed light on the essential statistical functions of the TETLE distribution.

1. Model Description

The cumulative distribution function (CDF), probability density function (PDF), and hazard function (HF) for a random variable related to the three-parameter TETLE distribution are given by (Al-Noor & Hilal, 2021)

$$F(x)_{TETLE} = \frac{1}{1-e^{-\lambda}}\left(1-e^{-\lambda\left(1-e^{-2\theta x}\right)^{\alpha}}\right);\ x>0, \lambda, \alpha, \theta > 0 \qquad (1)$$

$$f(x)_{TETLE} = \frac{2\lambda\alpha\theta}{1-e^{-\lambda}}\, e^{-2\theta x} e^{-\lambda\left(1-e^{-2\theta x}\right)^{\alpha}}\left(1-e^{-2\theta x}\right)^{\alpha-1} \qquad (2)$$

and

$$H(x)_{TETLE} = \frac{2\lambda\alpha\theta\, e^{-2\theta x} e^{-\lambda\left(1-e^{-2\theta x}\right)^{\alpha}}\left(1-e^{-2\theta x}\right)^{\alpha-1}}{e^{-\lambda\left(1-e^{-2\theta x}\right)^{\alpha}} - e^{-\lambda}} \qquad (3)$$

where λ and θ are scale parameters and α represents the shape parameter.

Some possible shapes of the TETLE PDF are displayed in Figure 1.

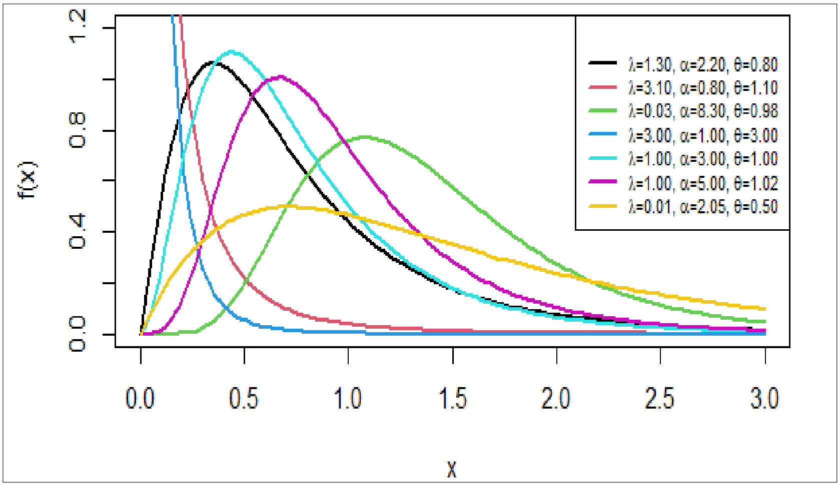

Figure 1. TETLE PDF with different parameter’s values

The TETLE distribution includes adding two extra parameters, one of which is a shape and the other is a scale, to the baseline one-scale parameter E distribution, whose essential functions CDF, PDF, and HF are given respectively by.

$$F(x)_E = 1 - e^{-\lambda x}\ ;\ x > 0, \lambda > 0 \tag{4}$$

$$f(x)_E = \lambda e^{-\lambda x} \tag{5}$$

and

$$H(x)_E = \lambda \tag{6}$$

2. Parameter Estimation

The maximum likelihood (ML) methodology is the most often used parameter estimation method due to its desirable features such as consistency, asymptotic efficiency, and invariance, see Dey et al., 2019 and Afify et al., 2024. Accordingly, the ML method is considered here to estimate the unknown parameters of the considered distributions.

Let x1,x2,…,xn be a random sample of size n from the distribution with PDF f(x;β), where β represents the general vector of the unknown parameter(s). The natural logarithm likelihood (NLL) function is.

$$\Delta = \ln L(x_1, x_2, \dots, x_n; \beta) = \sum_{i=1}^{n} \ln f(x_i; \beta) \tag{7}$$

The ML estimates $(\hat{\beta})$ of β determined by the values of β that maximize(7)with respect to β.

Now, the NLL corresponding to the TETLE distribution can be obtained from (2) as:

$$\Delta_{TETLE} = \sum_{i=1}^{n} n\ln\left(\frac{2\lambda\alpha\theta}{1-e^{-\lambda}}\right) - 2\theta\sum_{i=1}^{n} x_i - \lambda\sum_{i=1}^{n}\left(1-e^{-2\theta x_i}\right)^{\alpha} + (\alpha-1)\sum_{i=1}^{n}\ln\left(1-e^{-2\theta x_i}\right) \quad (8)$$

After taking the derivative of (8) concerning the relevant parameters, the ML estimates of λ,α, and θ can be obtained numerically, i.e., by solving the three nonlinear differential equationsTETLE) $\frac{\partial\Delta_{TETLE}}{\partial\lambda} = 0$, $\frac{\partial\Delta_{TETLE}}{\partial\alpha} = 0$, and $\frac{\partial\Delta_{TETLE}}{\partial\theta} = 0$.

The numerical technique is needed because the ML estimators are not in closed form.

The NLL corresponding to the E distribution can be obtained from (5) as:

$$\Delta_E = n\ln(\lambda) - \lambda\sum_{i=1}^{n} x_i$$

After taking the derivative of (9) concerning the parameter λ, i.e. ,$\frac{\partial\Delta_E}{\partial\lambda}$, ML estimate of λ $(\hat{\lambda})$ can be obtained directly by solving . Thus, where

3. Statistical Measures

For the evaluation process, the following common statistical information criteria are employed (see Al-Noor & Hilal, 2021) with the negative estimated NLL function evaluated at ML estimates $(-\hat{\Delta})$.,

Akaike: $A = 2m - 2\hat{\Delta}$;

Consistent Akaike: $CA = \frac{2nm}{n-m-1} - 2\hat{\Delta}$;

Hannan and Quinn: $HQ = 2m\ln(\ln(n)) - 2\hat{\Delta}$; and

Bayesian: $B = m\ln(n) - 2\hat{\Delta}$, where n: sample size,m: number of parameters to be estimated.

The distribution with smaller values of these information measures is preferred for modeling data, see Alizadeh et al., 2018. Further, as a goodness of fit test, the Kolmogorov-Smirnov (KS) test is considered for testing the hypothesis H_1 vs.H_2, where

H_0: The data follow the specified distribution with the CDF, F(x).

H_1: The data do not follow the specified distribution with the CDF, F(x).

The KS test statistics can be computed for the specified distribution by

The KS test statistics can be computed for the specified distribution by

$D = \max_{1\le i\le n}\left\{\left|\frac{i}{n} - F\left(x_{(i)}\right)\right|, \left|F\left(x_{(i)}\right) - \frac{i-1}{n}\right|\right\}$ where $F\left(x_{(i)}\right)$ is the CDF of the ascending ordered observations, see equations and (1) and (4).

The higher KS p-values indicate a better fit model, see Alizadehet al., 2018.

Application

In this section, a dataset of Asian tourists who traveled to Karbala to take part in the arbaeen pilgrimage ceremonies has been modeled using the TE-TLE distribution. The traditional E distribution, along with three-parameter BE, EGE, and WE, are included to expand the comparison process.

Given the importance of using continuous data to ensure consistency with the type of approved distributions, visitor rates for the year 2024 per 100 visitors are adopted, which are calculated based on data from the annual statistical bulletin related to the previous two years, 2023 and 2024, given that the total number of Asian visitors was 3261940 and 3446325, respectively.

Semi-Annual Scientific Journal
c-karbala.com
Vol.4, 4th year , issue 1

The visitor numbers in 2024 within 28 countries, available in Statistical Bulletin / Arbaeen Pilgrimage / Karbala Center for Studies and Research, are:

"77427, 960, 19, 37667, 31, 458, 13967, 3157795, 834, 16428, 22902, 110, 1644, 63272, 377, 29430, 8279, 583, 176, 63, 74, 252, 67, 2144, 97, 38, 11220, 11".

The TETLE, BE, EGE, WE, and E parameters are estimated under the ML method, and all numerical results are performed using the common statistical software R.

Results and Discussion

In Table 1, the ML estimates and the negative estimated NLL function $(-\hat{\Delta})$, corresponding to each distribution are listed.

Based on the obtained $(-\hat{\Delta})$,, the values of four statistical information criteria: A, CA, B, and HQ corresponding to each considered distribution are calculated and shown in Table 2.

Table 1. ML estimates and negative estimated log-likelihood function

Dist.	$\hat{\lambda}$	$\hat{a}$	$\hat{\theta}$	$-\hat{\Delta}$
TETLE	6.6074293	0.3233866	0.0090455	-19.53845
BE	0.1637833	5.3711278	0.0089323	-11.13806
EGE	0.3386432	0.1570403	0.1551506	-10.57274
WE	0.3057558	0.0660267	0.3868847	-19.51885
E	0.2725101	---	---	64.40223

Table 2. ML estimates and negative estimated log-likelihood function

Dist.	A	CA	B	HQ	KS p-value
TETLE	-33.0769	-32.0769	-29.0802	-31.8550	0.6383
BE	-16.2761	-15.2761	-12.2795	-15.0543	0.0691
EGE	-15.1454	-14.1454	-11.1488	-13.9236	0.0435
WE	-33.0377	-32.0377	-9.04108	-31.8158	0.6357
E	130.8045	130.9583	132.1367	131.2117	1.56e-12

The outcomes of Table 2 indicate that the TETLE has the lowest values of all information criteria and the largest p-values compared to other competitive distributions, making it the most appropriate distribution to represent the considered data.

Furthermore, the probability-probability (P-P) plots in Figure 2 show how well the considered probability distributions fit the Asian visitors' rate data.

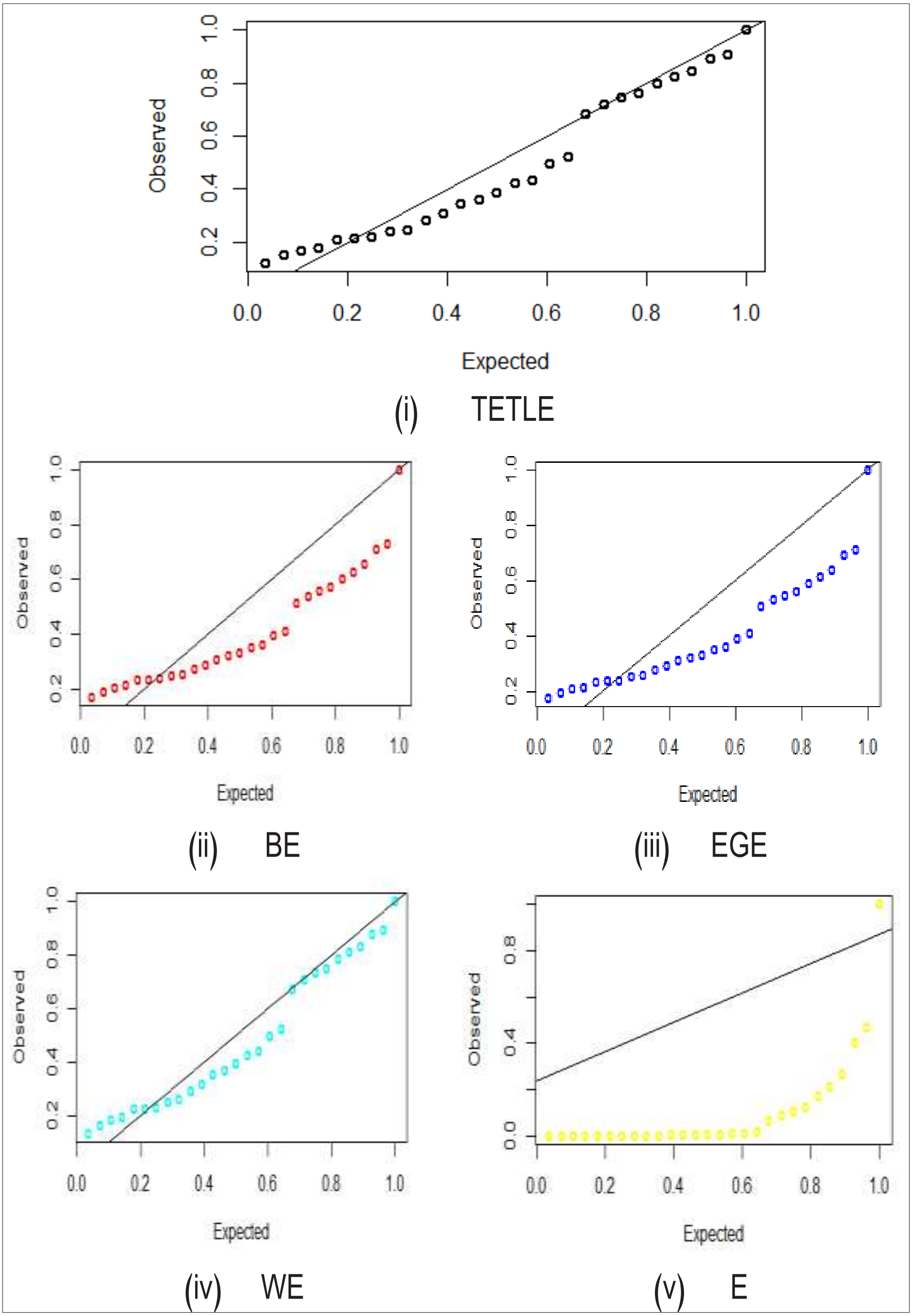

Figure 2. P-P plots of the considered distributions

From numerical and graphical illustrations, TETLE is the best distribution, and the three-parameter extended versions (BE, EGE, and WE) of the E distribution fit better than the one-parameter E distribution. Therefore, the most significant and pertinent statistical properties of the Asian visitors' rate can be derived from the statistical features of the TETLE distribution.

Conclusions

This paper concerned with finding a suitable statistical distribution for the data related to the occurrence of an important events, namely the arbaeen pilgrimage to Imam Hussein (PBUH), due to its profound spiritual and religious atmosphere, from which believers derive tremendous moral energy, filled with the pleasure of worship, the sweetness of faith, and the purification and refinement of the soul. This is achieved by striving, not for the sake of this world, but rather for the struggle of the soul, which is the greatest struggle, as the Messenger of Allah (may Allah bless him and his family and grant them peace) called it. Five probability distributions are employed to model the data of Asian tourists who came to Karbala to participate in the arbaeen pilgrimage ceremonies. The ML method is used to estimate the three parameters of TETLE, BE, EGE, WE, and one-parameter E. Numerous statistical measures are used to select the best-fitted distribution. According to the results, TETLE distribution is the best fit, and the WE distribution is the second. Third and fourth orders, respectively, were filled by the extended distributions BE and EGE. Thus, the importance of the extended versions of the exponential distribution is confirmed, which have greatly contributed to increasing its flexibility compared to its traditional version. For future studies, the TETLE distribution may be applied to other arbaeen pilgrimage data.

References

1. Afify, A. Z., Hussein, E. A., Al-Mofleh, H., Alsultan, R., & Aljohani, H. M. (2024). The Flexible Fréchet Distribution: Properties, Inference, and Medical Applications. Journal of Mathematics, 2024(1), 2600560.

2. Aldahlan, M. A., Jamal, F., Chesneau, C., Elgarhy, M., & Elbatal, I. (2020). The truncated Cauchy power family of distributions with inference and applications. Entropy, 22(3), 346.

3. Alizadeh, M., Rasekhi, M., Yousof, H. M., & Hamedani, G. G. (2018). The transmuted Weibull-G family of distributions. Hacettepe Journal of Mathematics and Statistics, 47(6), 1671-1689.

4. Al-Jumaili, R. K. S. (2023). The environment of Karbala from the perspective of arbaeen visitors: A survey study of a selected sample of visitors to Imam Hussein, peace be upon him, during the arbaeen. Karbala Center for Studies and Research's series of studies, in Arabic, 1-142.

5. Almarashi, A. M., Elgarhy, M., Jamal, F., & Chesneau, C. (2020). The exponentiated truncated inverse Weibull-generated family of distributions with applications. Symmetry, 12(4), 650.

6. Al-Muhanna, S. M. S. W. (2023). The peak of hostility in what happened to Imam Hussein (peace be upon him) on the day of Ashura: an authentic study. Karbala Center for Studies and Research's series of studies, in Arabic, 1-300.

7. Al-Noor, N. H., & Hadi, H. H. (2021, May). Properties and applications of truncated exponential Marshall Olkin Weibull distribution. In Journal of Physics: Conference Series (Vol. 1879, No. 3, p. 032024). IOP Publishing.

8. Al-Noor, N. H., & Hilal, O. A. (2021, May). Truncated exponential Topp Leone exponential distribution: properties and applications. In Journal of Physics: Conference Series (Vol. 1879, No. 3, p. 032039). IOP Publishing.

9. Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. Journal of data science, 12(1), 53-68.

10. Cordeiro, G. M., Ortega, E. M., & da Cunha, D. C. (2013). The exponentiated generalized class of distributions. Journal of data science, 11(1), 1-27.

11. Dey, S., Nassar, M., Kumar, D., Alzaatreh, A., & Tahir, M. H. (2019). A new lifetime distribution with decreasing and upside-down bathtub-shaped hazard rate function. Statistica, 79(4), 399-426.

12. Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. Communications in Statistics-Theory and methods, 31(4), 497-512.

13. Matsushita, R., Brom, P., Nagata, M., & Da Silva, S. (2023). Retrodicting with the truncated Lévy flight. Communications in Nonlinear Science and Numerical Simulation, 116, 106900.

14. Shams al-Din, M. M. (2023). The revolution of AI-Hussein, Karbala Center for Studies and Research's series of studies, 1-249.

15. Statistical Bulletin/Arbaeen Pilgrimage/Karbala Center for Studies and Research/General Secretariat of the Holy Shrine of Imam Hussain/ https://c-karbala.com/ar/sdfagdhfj