



RESEARCH ARTICLE - MATHEMATICS

The Rank Set Sampling Estimation Method under Progressive Censoring Sample for Exponential-Rayleigh Distribution

Rihaam N. Shatti¹, Iden H. Al-Kinani², Fadhaa O. Sameer^{3*}

¹ Al Nisour University College, Baghdad, Iraq

²Department of Mathematics, College of Science for Women, University of Baghdad, Al-Jadriya, Baghdad, Iraq

^{3*}Tropical-Biological Research Unit, College of Science, University of Baghdad, Al-Jadriya, Baghdad, Iraq

* Corresponding author E-mail: fadhaa.sameer@sc.uobaghdad.edu.iq

| Article Info. | Abstract |
|------------------|---|
| Article history: | <p>The Exponential-Rayleigh (ERD) distribution has two scale parameters; these parameters were estimated under Progressive Censored samples and to create the suggested rank set. Sampling estimation technique adding the progressive censored sample for the rank set technique. The estimation of parameter values in the present study is based on the Newton-Raphson method. The Iraqi Ministry of Health and Environment's Al Karkh General Hospital provided the real COVID-19 data used in the current research. The investigation was carried out in 2020 during a period of 120 days, from May 4 to August 31. There were (n=785) patients admitted to the hospital, (n-m=697) patients who lived, and (m=88) patients who passed away. Following the determination of the (ERD) estimating parameter values for the rank set sampling estimation method, the survival function, hazard function, and probability density function were determined.</p> |
| Received | |
| 3 June 2024 | |
| Accepted | |
| 8 July 2024 | |
| Publishing | |
| 30 March 2025 | |

This is an open-access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>)

The official journal published by the College of Education at Mustansiriyah University

Keywords: Progressive censored Sample, COVID-19, Rank Set Sampling Estimation Method, Survival function, Exponential-Rayleigh distribution

1-Introduction

In (2021), Rasheed et al. proposed and introduced the Exponential-Rayleigh distribution, which they named the Exponential-Rayleigh distribution (ERD). This distribution combined the cumulative functions of the exponential distribution with one scale parameter and the Rayleigh distribution with one scale parameter[1][2].

In various reliability and life testing studies, researchers and experimenters might not receive all the information describing the failure times of experimental units; this is known as censoring data[3].

There are three types of censored samples: Interval, left, and Right-censored samples.

Additionally, there are three branches inside of the right-censored sample: progressive censoring sample, single Type-I, and single Type-II censoring sample[4].

It is also known as (Multiply Censored Sample) or (Randomly Censored Sample).

Progressive censored Sample contains special cases of complete sample and Singly (Type-I and Type-II). This type is characterized by the different entering times in the period of the experiment[5].

The duration of the study is fixed in many epidemiologic studies and clinical, certain patients pass away before the study's conclusion, providing precise survival times; other patients are lost to follow-up; and still other patients survive until the experiment's conclusion.

While the data were given as fuzzy numbers, Makhdoom, et, al. in 2016 calculated the parameters for Exponential distribution by Type-II censoring samples and approximation forms of Lindley, Tierney, and Kadane. The effectiveness of the various strategies is assessed using a Monte Carlo simulation [6].

In 2018, Singh, and Chaturvedi, were estimated the Rayleigh distribution's parameters by discussing different estimation methods, the methods were the Bayes estimation method, Moments method, (MLEM), and the Computational approach estimate technique. Next that, the progressive type-II censoring sample was used[5].

In 2020, Abadi and Al-Kanani were interested in estimating a new mixture distribution for progressive censored data and estimated one shape parameter and two scale parameters of it[4].

This research contains, section two includes Exponential Rayleigh distribution and its properties, section three about parameters estimation, section four consists of the description of data, section five includes numerical results and the last section is the conclusion.

2-Exponential Rayleigh distribution (ERD)

Mohammed and Hussein's (2019) presentation of the exponential Rayleigh distribution relies on a combination of the tail (survival) function of the Rayleigh distribution and the tail (survival) function of the exponential distribution through the application of an application (minimum) [7]. Additionally, there is another approach to produce this distribution, which involves combining the exponential distribution function with the Rayleigh distributions' cumulative distribution function[8]. The cumulative distribution functions (*CDF*) of exponential Rayleigh distribution as follows:

$$F(t; \theta, \beta) = 1 - e^{-(\theta t + \frac{\beta}{2}t^2)} \quad (1)$$

Plots of different values for the two parameters (θ, β) are also displayed in the following figures.

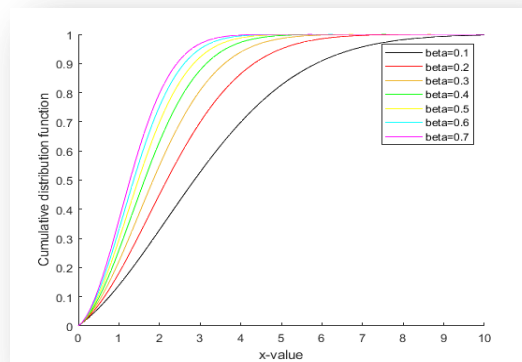


Fig.1 . Plot of the cumulative distribution function of *ER* distribution for $\theta = 0.1$ and $\beta = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7)$

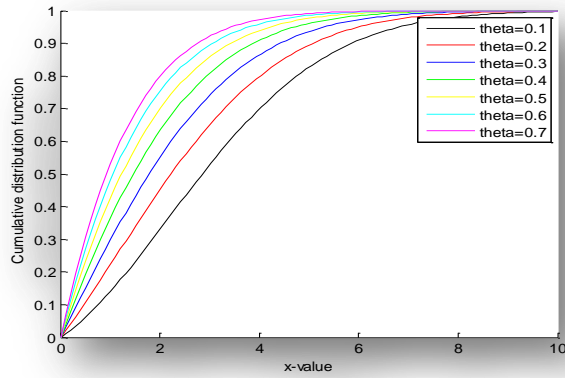


Fig.2 . Plot of the cumulative distribution function of *ER* distribution for $\beta = 0.1$ and $\theta = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7)$

For this distribution, the probability function of ERD is [9]:

$$f(t; \theta, \beta) = \begin{cases} (\theta + \beta t)e^{-(\theta t + \frac{\beta}{2}t^2)} & , \quad t > 0 \\ 0 & , \textit{otherwise} \end{cases} \quad (2)$$

Where the (θ, β) are the scale parameters of the ERD, and $\theta > 0, \beta > 0$

The Survival function and the hazard rate function is given respectively as[3]:

$$S(t; \theta, \beta) = e^{-(\theta t + \frac{\beta}{2}t^2)} \quad t > 0 \quad (3)$$

$$h(t; \theta, \beta) = \theta + \beta t \quad t > 0 \quad (4)$$

3-Parameters Estimation

In this section, the parameters of (ERD) will be estimate using the rank set sampling (RSS) for the Progressively Censoring sample.

In Australia, rank set sampling was proposed as a method for calculating pasture yield. This technique was not employed for a long time, but over the last 30 years, it has been widely used in research, making it indispensable in many areas [10].

Extensive methodological research has been conducted on ranked set sampling, a technique for data collection and analysis. Other related approaches that are currently active study topics have been developed by it [11].

In order to create the suggested Rank Set Sampling Estimator technique (RSSEM) in the Censoring sample, we incorporate the Progressive Censored sample for the Rank Set technique in this work.

Let X_1, X_2, \dots, X_n be a random sample which was from continuous probability density function $f(x)$ [12]. Let T_1 be the smallest of X_i , T_2 be the next order to magnitude and T_n be the largest of X_i .

That means $T_1 < T_2 < \dots < T_n$ represent X_1, X_2, \dots, X_n when the latter are arranged in a sending order of magnitude[13]. Then T_1, T_2, \dots, T_n are denoted of i^{th} order statistic of a random sample X_1, X_2, \dots, X_n . Now easy to formulate the probability density function of any order statistic, called T_i in term of $f(x)$ and $F(x)$ as follows[14]:

$$g(y_i) = \begin{cases} \frac{m!}{(i-1)!(m-i)!} [F(t_i)]^{i-1} [1 - F(t_i)]^{m-i} f(t_i) & a < t_i < b \\ 0 & otherwise \end{cases} \quad (5)$$

Using the order statistic formula found in Equation (5), the (ER) distribution can be calculated as follows:

$$g(t_{(i)}) = \frac{m!}{(i-1)!(m-i)!} \left[1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{i-1} \left[e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{m-i+1} (\theta + \beta t_{(i)}) \quad (5)$$

After that finding the likelihood function for Eq. (6) by utilizing Progressive Censoring formula as following:

$$L = \frac{n!}{(n-m)!} \prod_{i=1}^n \{ [g(t_{(i)}, \theta, \beta)]^{\delta_i} [s(t_{(i)})]^{1-\delta_i} \}$$

where $\delta_i = \begin{cases} 1 & \text{patient who die} \\ 0 & \text{if patient still alive} \end{cases}$

$$L = \frac{n!}{(n-m)!} \prod_{i=1}^n \left\{ \frac{m!}{(i-1)!(m-i)!} \left(1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right)^{i-1} \left(e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right)^{m-i+1} (\theta + \beta t_{(i)}) \right\}^{\delta_i} \cdot \left[e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{1-\delta_i}$$

let $\frac{n!}{(n-m)!} = c$, $\frac{m!}{(i-1)!(m-i)!} = d$ then we get:

$$L = c \prod_{i=1}^n \left\{ d \left(1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right)^{i-1} \left(e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right)^{m-i+1} (\theta + \beta t_{(i)}) \right\}^{\delta_i} \cdot \left[e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{1-\delta_i}$$

$$L = c d^{n\delta_i} \prod_{i=1}^n \left[1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{\delta_i(i-1)} \prod_{i=1}^n \left[e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{\delta_i(m-i+1)} \cdot \prod_{i=1}^n [\theta + \beta t_{(i)}]^{\delta_i} \prod_{i=1}^n \left[e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right]^{1-\delta_i} \quad (6)$$

Applying the natural logarithm of each side of Eq.(7):

$$\begin{aligned} \ln L = & \ln c + n\delta_i \ln d + \sum_{i=1}^n \delta_i (i-1) \ln \left(1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)} \right) - \sum_{i=1}^n \delta_i (m-i+1) \left(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2 \right) \\ & + \sum_{i=1}^n \delta_i \ln(\theta + \beta t_{(i)}) - \sum_{i=1}^n (1 - \delta_i) \left(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2 \right) \end{aligned} \quad (7)$$

Deriving the Eq.(8) partially w.r.t θ , β and setting it at zero.

$$\frac{\partial \ln l}{\partial \theta} = \sum_{i=1}^n \frac{\delta_i(i-1)t_{(i)} e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}}{1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}} - \sum_{i=1}^n \delta_i(m - i + 1)t_{(i)} + \sum_{i=1}^n \frac{\delta_i}{(\theta + \beta t_{(i)})} - \sum_{i=1}^n (1 - \delta_i)t_{(i)} = 0$$

$$\frac{\partial \ln l}{\partial \beta} = \frac{1}{2} \sum_{i=1}^n \frac{\delta_i(i-1)t_{(i)}^2 e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}}{1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}} - \frac{1}{2} \sum_{i=1}^n \delta_i(m - i + 1)t_{(i)}^2 + \sum_{i=1}^n \frac{\delta_i t_{(i)}}{(\theta + \beta t_{(i)})} - \frac{1}{2} \sum_{i=1}^n (1 - \delta_i)t_{(i)}^2 = 0$$

Now, we can put $\frac{\partial \ln l}{\partial \theta}$ as a function $f(\theta)$ and put $\frac{\partial \ln l}{\partial \beta}$ as a function $f(\beta)$

$$f(\theta) = \sum_{i=1}^n \frac{\delta_i(i-1)t_{(i)} e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}}{1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}} - \sum_{i=1}^n \delta_i(m - i + 1)t_{(i)} + \sum_{i=1}^n \frac{\delta_i}{(\theta + \beta t_{(i)})} - \sum_{i=1}^n (1 - \delta_i)t_{(i)} \quad (8)$$

$$f(\beta) = \frac{1}{2} \sum_{i=1}^n \frac{\delta_i(i-1)t_{(i)}^2 e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}}{1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}} - \frac{1}{2} \sum_{i=1}^n \delta_i(m - i + 1)t_{(i)}^2 + \sum_{i=1}^n \frac{\delta_i t_{(i)}}{(\theta + \beta t_{(i)})} - \frac{1}{2} \sum_{i=1}^n (1 - \delta_i)t_{(i)}^2 \quad (9)$$

It should be noted that solving equations (9) and (10) with the Newton-Raphson approach is challenging due to the iterative nature of determining $(\hat{\theta}, \hat{\beta})$ values.

$$\theta_{k+1} = \theta_k - \frac{f(\theta_k)}{\hat{f}(\theta_k)}, \text{ also } \beta_{k+1} = \beta_k - \frac{f(\beta_k)}{\hat{f}(\beta_k)},$$

$$\text{Where } \hat{f}(\theta_k) = - \sum_{i=1}^n \frac{\delta_i(i-1)t_{(i)}^2 e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}}{\left(1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}\right)^2} - \sum_{i=1}^n \frac{\delta_i}{(\theta + \beta t_{(i)})^2}$$

$$\text{Where } \hat{f}(\beta_k) = - \frac{1}{4} \sum_{i=1}^n \frac{\delta_i(i-1)t_{(i)}^4 e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}}{\left(1 - e^{-(\theta t_{(i)} + \frac{\beta}{2} t_{(i)}^2)}\right)^2} - \sum_{i=1}^n \frac{\delta_i t_{(i)}^2}{(\theta + \beta t_{(i)})^2}$$

The error term formulation's is given by:

$$\epsilon_{k+1}(\theta) = \theta_{k+1} - \theta_k$$

$$\epsilon_{k+1}(\beta) = \beta_{k+1} - \beta_k$$

The assumed initial values are θ_k and β_k .

4-Description of data

From AL-Karkh General Hospital in the Iraq, real data was collected. The period of our study from 4-5-2020 to 31-8-2020, in days, is equal to (120). There were (1058) patients who signed in the hospital, six cases have been dismissed which consist of:

- a) There were 26 prisoners.

- b) There were 29 patients whose exit status was unknown
- c) 35 patients were moved to different hospitals.
- d) Two patients escaped from the hospital.
- e) People with non-positive swabs were 48
- f) There were 133 patients discharged under their care.

Then the total number of patients becomes 785. and 88 of them died during the period of study.

Scientific studies are made up of statistical tests, there are two types of these tests are said to be parametric tests and nonparametric tests[15]. For scientific research, statistical tests are important because they enable a researcher to analyze the data, also draw general conclusions, and are useful in summarizing the results[16].

We'll utilize Chi-square testing in this study; it's a kind of non-parametric test with significant applications that include the following[17]:

1. It is used to check the independence between variables.
2. It is used to test the goodness of fit (The goodness of fit utilized to determine whether the set of data follows from a suggested distribution[18]).

The Chi-square test's alternative and null hypotheses are

H_0 : If the real data are distributed as (ERD).

H_1 : If the data aren't distributed as (ERD).

This test's formula is provided by:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{O_i}$$

In which the class i observations are covered by O_i , the predictable rate in class i is denoted by E_i and k denotes to the whole number of classes.

Now, compute

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{O_i}$$

$$\begin{aligned} \chi^2 = & \frac{(15-9.26)^2}{9.26} + \frac{(12-10.808)^2}{10.808} + \frac{(13-11.573)^2}{11.573} + \frac{(22-31.671)^2}{31.671} + \frac{(7-11.007)^2}{11.007} + \frac{(3-9.948)^2}{9.948} + \frac{(7-8.602)^2}{8.602} \\ & + \frac{(3-7.138)^2}{7.138} + \frac{(2-5.696)^2}{5.696} + \frac{(4-4.379)^2}{4.379} \end{aligned}$$

$$\chi^2 = 15.32008$$

The null hypothesis was accepted and the alternative hypothesis was rejected because (21.67) is the number of tabulated values with the degree of freedom (9) is more than the calculated value (15.32008). So the data are distributed as (ERD).

5-Numerical Results

The following parameters' values were estimated using Matlab programming (version 2021):

The estimate values were $\hat{\theta} = 0.00964$, $\hat{\beta} = 0.00558$ while the beginning values, which are chosen at random, were $\theta_0 = 0.0058$, $\beta_0 = 0.0058$

By Substituting the values of the lifetime and estimate values of parameters in Eq.1 , Eq.2 , Eq.3. And Eq.4 then obtain the following table.

Table1 show the estimation values of survival function, the probability function, distribution function and hazard function for each lifetime.

| lifetime (T) | $\hat{S}(t)$ | $\hat{f}(t)$ | $\hat{F}(t)$ | $\hat{h}(t)$ |
|--------------|--------------|--------------|--------------|--------------|
| (1) | (0.98764) | (0.01504) | (0.01236) | (0.01523) |
| (2) | (0.97001) | (0.02018) | (0.02999) | (0.02082) |
| (3) | (0.94738) | (0.02499) | (0.05262) | (0.02639) |
| (4) | (0.92014) | (0.02942) | (0.07986) | (0.03197) |
| (5) | (0.88871) | (0.03337) | (0.11129) | (0.03755) |
| (6) | (0.85357) | (0.03681) | (0.14643) | (0.04313) |
| (7) | (0.81526) | (0.03971) | (0.18474) | (0.04871) |
| (8) | (0.77433) | (0.04204) | (0.22567) | (0.05429) |
| (9) | (0.73137) | (0.04379) | (0.26863) | (0.05987) |
| (10) | (0.68695) | (0.04496) | (0.31305) | (0.06545) |
| (11) | (0.64163) | (0.04558) | (0.35837) | (0.07104) |
| (12) | (0.59597) | (0.04566) | (0.40403) | (0.07662) |
| (13) | (0.55048) | (0.04525) | (0.44953) | (0.08219) |
| (15) | (0.46185) | (0.04312) | (0.53816) | (0.09336) |
| (16) | (0.41951) | (0.041505) | (0.58049) | (0.09894) |
| (17) | (0.37893) | (0.03961) | (0.62107) | (0.10452) |
| (18) | (0.34037) | (0.03748) | (0.65963) | (0.1101) |

In table 1 the results can be seen:

The values of $\hat{f}(t)$ were increased till $t=12$; latter, they reduced from 13 to 18. An rise in failure times as in results, the values of $\hat{F}(t)$ were increased, $\hat{S}(t)$ were decreased and increased values of $\hat{h}(t)$. For $\hat{S}(t)$, the average of the square error (MSE) is 0.031229.

6-Conclusion

The progressive censored sample for the rank set method can be added in this study to create the suggested rank set sampling estimation method (RSSEM), which can be utilized to estimation the two scale parameters of the (ERD). Since the tabulated value exceeds the value determined by the Chi-square test, the real data is distributed as (ERD). Because of the increase of failure times, $\hat{h}(t)$ are increasing, $\hat{S}(t)$ are decreasing, and the $\hat{F}(t)$ are increasing.

References

- [1] S. A. Zain and I. Hassan Hussein, "Estimate for Survival and Related Functions of Weighted Rayleigh Distribution.," *Ibn AL- Haitham Journal For Pure and Applied Sciences*, vol. 34, no. 1, 2021.
- [2] S. A. Zain and I. H. Hussein, "Some Aspects of Weighted Rayleigh Distribution," *Ibn AL- Haitham Journal For Pure and Applied Sciences*, vol. 33, no. 2, p. 107, 2020.
- [3] R. N. Shatti and I. H. Al-Kinani, "Estimating the Parameters of Exponential-Rayleigh Distribution under Type-I Censored Data," *Baghdad Science Journal*, vol. 21, no. 1, pp. 146–150, 2024.
- [4] Q. S. Abadi and I. H. AL-Kanani, "Progressively censoring data to estimate the survival function in the new mixture distribution," *Journal of Critical Reviews*, vol. 7, no. 1, pp. 527–530, 2020.
- [5] A. Chaturvedi, S. K. Singh, and U. Singh, "Statistical inferences of type-II progressively hybrid censored fuzzy data with rayleigh distribution," *Austrian Journal of Statistics*, vol. 47, no. 3, pp. 40–62, 2018.
- [6] I. Makhdoom, P. Nasiri, and A. Pak, "Estimating the Parameter of Exponential Distribution under Type II Censoring From Fuzzy Data," *Journal of Modern Applied Statistical Methods*, vol. 15, no. 2, pp. 495–509, 2016.
- [7] M. J. Mohammed and I. H. Hussein, "Some estimation methods for new mixture distribution with simulation and application[c]," *IOP Conference Series: Materials Science and Engineering*, vol. 571, no. 1, 2019.
- [8] M. A. Qasim and A. R. Al-dubaicy, "Some Bayes estimators for Exponentiated Rayleigh Distribution under Doubly Type II Censored Data," *Mustansiriyah Journal of Pure and Applied Sciences*, vol. 1, no. 1, pp. 30–39, 2023.

- [9] R. N. Shatti and I. H. Al-Kinani, "Estimating the Parameters of Exponential-Rayleigh Distribution for Progressively Censoring Data with S- Function about COVID-19," *Baghdad Science Journal*, vol. 21, no. 2, pp. 496–503, 2024.
- [10] N. S. Karam and H. M. Abdawon, "Stress-Strength Reliability Bayesian of Generalized Exponential-Poisson Distribution for Complete Data," *Mustansiriyah Journal of Pure and Applied Sciences*, vol. 1, no. 1, pp. 91–107, 2023.
- [11] M. A. W. Mahmoud and M. G. M. Ghazal, "Estimations from the exponentiated rayleigh distribution based on generalized Type-II hybrid censored data," *Journal of the Egyptian Mathematical Society*, vol. 25, no. 1, pp. 71–78, 2017.
- [12] M. Imran, L. A. Al-Essa, M. H. Tahir, C. Chesneau, S. Shakoor, and F. Jamal, "An extended exponential model: Estimation using ranked set sampling with applications," *Journal of Radiation Research and Applied Sciences*, vol. 16, no. 4, p. 100642, 2023.
- [13] B. atiya Kalaf, N. Abdul Ameer Jabar, and U. Y. Madaki, "Truncated Inverse Generalized Rayleigh Distribution and Some Properties," *Ibn AL-Haitham Journal For Pure and Applied Sciences*, vol. 36, no. 4, pp. 414–428, 2023.
- [14] G. K. Vishwakarma, S. M. Zeeshan, and C. N. Bouza, *Ratio and product type exponential estimators for population mean using ranked set sampling*, vol. 38, no. 3. Elsevier Inc., 2017.
- [15] A. S. Hassan, R. S. Elshaarawy, and H. F. Nagy, "Parameter estimation of exponentiated exponential distribution under selective ranked set sampling," *Statistics in Transition New Series*, vol. 23, no. 4, pp. 37–58, 2022.
- [16] S. T. Nihan, "Karl Pearsons chi-square tests," *Educational Research and Reviews*, vol. 15, no. 9, pp. 575–580, 2020.
- [17] R. Rana and R. Singhal, "Chi-square test and its application in hypothesis testing," *Journal of the Practice of Cardiovascular Sciences*, vol. 1, no. 1, p. 69, 2015.
- [18] M. A. H. Sabry and E. M. Almetwally, "Estimation of the Exponential Pareto Distribution's Parameters under Ranked and Double Ranked Set Sampling Designs," *Pakistan Journal of Statistics and Operation Research*, vol. 17, no. 1, pp. 169–184, 2021.