*IOSR Journal of Mathematics (IOSR-JM)*
*e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 5, Ser. 1 (Sept. – October. 2023), PP 34-38*
*www.iosrjournals.org*

# Goodness-Of-Fit Tests For The Suja Distribution And Its Application On Covid-19 Data In Saudi Arabia

Faten A. Momenkhan[1], Tamadur H. Alsumairi[1] and Hanaa Abu-Zinadah[1*]
[1] *University of Jeddah, College of Science, Department of Statistics, Jeddah, Saudi Arabia*

***Abstract:***
*In this paper, the goodness-of-fit tests for the Suja distribution with one parameter was studied based on a complete sample. The maximum likelihood method is utilized to obtain the estimator of the parameter. Also, simulation was carried out using different sample sizes and different values of the parameter. In addition, the goodness of fit test statistics was calculated for proposed models to find out the best of it for data COVID-19 in Saudi Arabia.*

***Key words:*** *Suja distribution, maximum likelihood, goodness-of-fit test, complete sampling.*

---

Date of Submission: 06-09-2023 Date of Acceptance: 16-09-2023

---

## I. Introduction

Lifetime data analysis and modeling are crucial because they are used by statisticians and researchers in practically every field of applied science, including engineering, medicine, biology, insurance and finance among others. In reliability analysis, the time until the relevant event occurs is referred to as the lifetime, survival or failure time. The event may be the failure of a piece of equipment, a person's death, development of disease symptoms and health code violation. According to Alshenawy (2020), the study of lifespan distributions with one parameter is still fascinating to researchers today. A new lifetime distribution with one parameter, the Suja distribution, was proposed by Shanker (2017). Its probability density function is as follows:

$$f(x,\theta) = \frac{\theta^5}{\theta^4+24}(1+x^4)e^{-\theta x} \quad ; x > 0;\ \theta > 0 \tag{1}$$

and its cumulative distribution function as:

$$F(x,\theta) = 1 - \left[1 + \frac{\theta^4 x^4 + 4\theta^3 x^3 + 12\theta^2 x^2 + 24\theta x}{\theta^4+24}\right] e^{-\theta x} \quad ; x > 0;\ \theta > 0 \tag{2}$$

Researchers are always interested in one-parameter lifespan distributions. Sharma et al. (2019) conducted a comparative investigation of the statistical properties and applicability of certain well-known one parameter lifetime distributions from the statistics literature to model lifetime data from biological research and engineering. Alshenawy (2020) has also introduced and explored the A distribution, a new one parameter distribution. On the other hand, Abu-Zinadah (2014) examined the goodness of fit for the exponentiated Gompertz distribution using type II censored sampling. Shankar (2019) also studied Chi-square goodness of fit testing to see whether it is reasonable to believe that a random sample of data comes from a given probability distribution, which was one of the subjects covered in an undergraduate engineering probability course. Moreover, Abu-Zinadah and Binkhamis (2020) developed goodness-of-fit tests for the beta Gompertz distribution with four parameters using a complete sample. Furthermore, Shanker (2017) suggested and investigated a new one parameter distribution named "Suja distribution" for modeling lifetime data. The parameter estimate has been studied using maximum likelihood estimation and the method of moments. Finally, the Suja distribution's goodness of fit was tested using real lifetime data, and the fit was compared to other one parameter lifetime distributions. Al-Omari et al. (2019) also proposed a new extension of the length-biased Suja distribution known as the power length-biased Suja distribution (PLBSD). The PLBSD is a distribution with two parameters. The PLBSD's use for modeling dependability data was demonstrated using real data to demonstrate the new model's performance. In the article, estimate the Suja distribution parameter using the maximum likelihood estimation and carry out by different complete samples size and different parameter values of Suja distribution. In addition, the goodness of fit test statistics calculate for proposed models to find out the best of it for data Coronavirus disease (COVID-19) in Saudi Arabia.

## II. Maximum Likelihood Estimator

This section presents the maximum likelihood estimation of the Suja distribution's parameter($\theta$). If $x_1, x_2, \dots, x_n$is a random sample from Suja distribution, the log-likelihood function is L:

$$L = \left(\frac{\theta^5}{\theta^4+24}\right)^n e^{-n\theta\bar{x}} \prod_{i=1}^{n}(1 + x_i^4) \quad (3)$$

and the log-likelihood function is

$$\ln L = 5n \ln\theta - n \ln(\theta^4 + 24) - n\theta\bar{x} + \sum_{i=1}^{n} \ln(1 + x_i^4) \quad (4)$$

By differentiating both sides of the equation with respect to $\theta$ we get

$$\frac{d \ln L}{d\theta} = \frac{5n}{\theta} - \frac{n(4\theta^3)}{\theta^4+24} - n\bar{x} \quad (5)$$

The Maximum Likelihood equation can be written $\frac{d \ln L}{d\theta} = 0$

$$\bar{x}\theta^5 - \theta^4 + 24\,\bar{x}\theta - 120 = 0 \quad (6)$$

There isn't a closed form answer to this equation for $\theta$. As a result, the equation can be solved numerically using a method like the Newton-Raphson method to determine the Maximum Likelihood Estimate $\hat{\theta}_{MLE}$.

## III. Simulation Study

In this section, we compare the performance of the different estimators provided in the preceding sections by showing the results of several numerical experiments. We carry out in-depth Monte Carlo simulations to examine the performance of the various estimators, particularly in terms of their relative root mean square error (RRMSE) and absolute relative bias (ARBias) for various sample sizes and parameter values. The parameter values are selected as θ = 2.0, 2.5, 3.0, 3.5 and n = 10, 30, 50 and 100. The ARBias's and RRMSE's of estimators are calculated using 1000 replications, where

$$RRMSE(\hat{\theta}) = \frac{\sqrt{MSE(\hat{\theta}-\theta)}}{\theta} \quad \text{and} \quad ARBias(\hat{\theta}) = \left|\frac{(\hat{\theta}-\theta)}{\theta}\right|$$

Comparing the computational complexities of the different estimators, it is observed that when the parameter θ , we need some iterative techniques to compute MLE.

**Table (1): shows the simulated values of ARBias's and RMSE's of estimators of θ.**

| N | | $\theta = 2$ | $\theta = 2.5$ | $\theta = 3$ | $\theta = 3.5$ |
|---|---|---|---|---|---|
| 10 | $\hat{\theta}$ | 2.02 | 2.59054 | 3.07871 | 3.31057 |
| | RRMSE | 0.1238 | 0.15353 | 0.183757 | 0.2655 |
| | ARBias | 0.0100011 | 0.0362178 | 0.0262367 | 0.0541222 |
| 30 | $\hat{\theta}$ | 2.00355 | 2.52822 | 3.03659 | 3.54257 |
| | RRMSE | 0.0698142 | 0.0783569 | 0.0864441 | 0.121666 |
| | ARBias | 0.00177642 | 0.0112865 | 0.0121963 | 0.0121642 |
| 50 | $\hat{\theta}$ | 2.00462 | 2.50981 | 3.02961 | 3.54719 |
| | RRMSE | 0.0527543 | 0.0569595 | 0.0693762 | 0.833204 |
| | ARBias | 0.00231156 | 0.00392317 | 0.00986941 | 0.0134817 |
| 100 | $\hat{\theta}$ | 2.00257 | 2.51008 | 3.01281 | 3.51864 |
| | RRMSE | 0.0367069 | 0.0569818 | 0.0456125 | 0.057272 |
| | ARBias | 0.0012827 | 0.00403198 | 0.00426992 | 0.00532641 |

From Table (1): shows the following findings about the performance of parameter estimation for the Suja distribution:

1- Almost as the sample size increases, the estimates of the parameter of MLE is decrease.

2- The RRMSE's and ARBias's of the parameter estimations often decrease as sample size increases.

3- The results for the RRMSE and ARBias of the parameter estimates had small values. This indicates that the MLE were suitable estimators for the parameter of the Suja distribution.

## IV. Real Data Applications

This section compares the Suja distribution to other lifetime models, such as the Lindley and exponential distributions, in order to demonstrate the value of utilizing it as a lifetime model. Theta parameter estimates for the distributions using machine learning are obtained. To confirm the ML estimate of the θ parameter, goodness of fit tests using the Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC), and Hannan-Quinn information criterion (HQIC) are conducted, [See Whittaker and Furlow (2009)] for further information on the model selection process. They are defined as follows:

$$AIC = -2l(\hat{\theta}) + 2k$$
$$BIC = -2l(\hat{\theta}) + k\log(n)$$
$$CAIC = -2l(\hat{\theta}) + \frac{2kn}{n-k-1}$$
$$HQIC = -2l(\hat{\theta}) + 2k\log(\log(n))$$

Mathematica 13.0 was used to perform the computations presented for data analysis.

**Daily cases of Covid 19 in Saudi Arabia**

In this section, we will study the data number of daily COVID-19 cases in Saudi Arabia will be compared with other known lifetime models. The data is taken from Sehhty website, which is a statistical site for Corona cases around the world that relies on international sources, with URL: https://sehhty.com/sa-covid/

**Data set 1:**

The number of daily COVID-19 cases in Saudi Arabia of Fourth Quarter of 2020 . These data of size 92 are

481, 419, 390, 379, 477, 468, 421, 407, 405, 323, 348, 474, 501, 472, 433, 359, 348, 381,385, 405, 401, 383, 395, 323, 357, 399, 416, 435, 398, 402, 374, 381, 473, 426, 450,436, 407, 363, 392, 471, 394, 311, 441, 349, 305, 301, 362, 290, 319, 286, 221, 224,231, 252, 326, 322, 302, 220, 217, 232, 263, 249, 230, 234, 190, 187, 209, 193, 159, 141, 168, 166, 139, 125, 142, 180, 181, 174, 158, 162, 168, 181, 177, 189, 178, 163, 154, 119, 149, 113, 140, 137

**Table (2): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia of Fourth Quarter of 2020.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.0165581 | 1137.02 | 1139.54 | 1138.04 | 1137.07 |
| Lindley | 0.00660137 | 1181.17 | 1183.69 | 1182.19 | 1181.22 |
| Exponential | 0.001 | 1328.59 | 1331.11 | 1329.61 | 1329.61 |

**Data set 2:**

The number of daily COVID-19 cases in Saudi Arabia of First Quarter of 2021 . These data of size 90 are

101, 82, 94, 104, 118, 108, 97, 110, 117, 140, 147, 175, 169, 173, 140, 176, 170, 226, 238, 212, 213, 197, 186, 213, 223, 216, 253, 267, 270, 261, 255, 310, 306, 303, 327, 286, 317, 356, 353, 369, 364, 353, 337, 322, 314, 322, 334, 327, 337, 325, 315, 327, 335, 353, 356, 346, 338, 322, 317, 302, 331, 375, 384, 382, 357, 351, 390, 386, 390, 360, 351, 348, 345, 354, 393, 381, 391, 382, 367, 404, 410, 466, 482, 510, 502, 531, 541, 556, 585, 590

**Table (3): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia of First Quarter of 2021.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.0162931 | 1121.13 | 1123.63 | 1122.14 | 1121.18 |
| Lindley | 0.00649608 | 1159.88 | 1162.38 | 1160.89 | 1159.92 |
| Exponential | 0.001 | 1300.63 | 1303.13 | 1301.64 | 1300.68 |

**Data set 3:**

The number of daily COVID-19 cases in Saudi Arabia of Second Quarter of 2021 . These data of size 91 are

728, 684, 673, 695, 792, 783, 902, 904, 878, 799, 842, 951, 929, 985, 964, 948, 916, 970, 1070, 1028, 1055, 1098, 1072, 953, 958, 1045, 1062, 1026, 1056, 1048, 937, 953, 999,1016, 1090, 1039, 997, 942, 986, 999, 1020, 1116, 927, 837, 825, 886, 1046, 1213, 1330,1136, 1142, 1067, 1157, 1389, 1320, 1183, 1215, 1106, 907, 1245, 1251, 1269, 1261, 1201, 1144, 984, 1161, 1261, 1274, 1286, 1175, 1077, 1017, 1109, 1269, 1239, 1309, 1236, 1153, 1079, 1212, 1479, 1253, 1255, 1312, 1301, 1218, 1318, 1567, 1486, 1534

**Table (4): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia of Second Quarter of 2021.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.00461793 | 1308.72 | 1311.23 | 1309.73 | 1308.76 |
| Lindley | 0.00184547 | 1388.34 | 1390.85 | 1389.35 | 1388.39 |
| Exponential | 0.001 | 1456.27 | 1458.78 | 1457.28 | 1456.31 |

**Data set 4:**

The number of daily COVID-19 cases in Saudi Arabia month of March and April of 2022 . These data of size 61 are

476, 407, 363, 283, 317, 279, 219, 18, 190, 178, 135, 145, 146, 129, 127, 97, 105, 98, 126, 133, 125, 110, 115, 99, 79, 98, 127, 106, 118, 107, 96, 88, 78, 104, 116, 108, 114,104, 95, 96, 130, 135, 110, 103, 105, 83, 93, 133, 143, 146, 130, 117, 91, 85, 109, 106, 98, 92, 99, 90, 99

**Table (5): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia month of March and April of 2022.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.0365226 | 675.043 | 677.154 | 675.87 | 675.111 |
| Lindley | 0.0145031 | 691.897 | 694.008 | 692.725 | 691.965 |
| Exponential | 0.01 | 730.851 | 732.962 | 731.678 | 730.919 |

**Data set 5:**

The number of daily COVID-19 cases in Saudi Arabia month of May and June of 2022 . These data of size 61 are

128, 102, 124, 159, 219, 234, 339, 565, 569, 642, 611, 559, 434, 431, 630, 621, 602, 545,540, 411, 467, 650, 557, 540, 516, 483, 408, 530, 686, 667, 569, 775, 662, 565, 652, 967, 952, 1029, 955, 932, 753, 905, 1188, 1152, 1033, 963, 945, 831, 930, 1232, 1143, 1082, 1002, 927,734, 827, 1076, 869, 759, 698, 625

**Table (6): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia month of May and June of 2022.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.00731397 | 872.018 | 874.129 | 872.845 | 872.086 |
| Lindley | 0.00292132 | 887.754 | 889.865 | 888.581 | 887.822 |
| Exponential | 0.001 | 928.148 | 930.259 | 928.975 | 928.216 |

**Data set 6:**

The number of daily COVID-19 cases in Saudi Arabia month of September and October of 2022 . These data of size 61 are

74, 66, 82, 114, 108, 102, 104, 86, 78, 109, 132, 125, 107, 108, 98, 62, 98, 131, 105, 98, 80, 77, 57, 96, 161, 130, 132, 127, 109, 77,89, 156, 155, 152, 137, 139, 87, 122, 226, 195, 203, 191, 165, 106, 139, 246, 240, 247, 241, 216, 129, 178, 310, 276, 326, 294, 237, 160, 195, 327, 259

**Table (7): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia month of September and Octoberof 2022.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.0332389 | 678.642 | 680.753 | 679.47 | 678.71 |
| Lindley | 0.0132077 | 701.4 | 703.511 | 702.227 | 701.468 |
| Exponential | 0.01 | 747.351 | 749.462 | 748.178 | 747.419 |

**Data set 7:**
The number of daily COVID-19 cases in Saudi Arabia month of January and February of 2023 . These data of size 59 are

39, 29, 28, 24, 20, 18, 23, 33, 29, 30, 27, 26, 14, 19, 31, 31, 35, 34, 34, 22, 30, 50, 39, 40, 35, 35, 21, 33, 55, 46, 50, 52, 42, 25, 40, 65, 58, 62, 54, 46, 30, 39, 70, 67, 67, 64, 49, 32, 43, 66, 58, 54, 42, 48, 43, 59, 101, 90, 76

**Table (8): The ML estimates of $\theta$ parameter and the value of AIC, BIC, HQIC and CAIC for the number of daily cases of COVID-19 in Saudi Arabia month of January and February of 2023.**

| Model | MLE($\hat{\theta}$) | AIC | BIC | HQIC | CAIC |
|---|---|---|---|---|---|
| Suja | 0.11697 | 501.0 | 503.078 | 501.811 | 501.07 |
| Lindley | 0.0457179 | 529.756 | 531.834 | 530.567 | 529.826 |
| Exponential | 0.05 | 607.696 | 609.774 | 608.507 | 607.767 |

Figures (1-7) show The models' Empirical Distribution and Estimated CDF for Data Set for the data number of daily COVID-19 cases in Saudi Arabia.

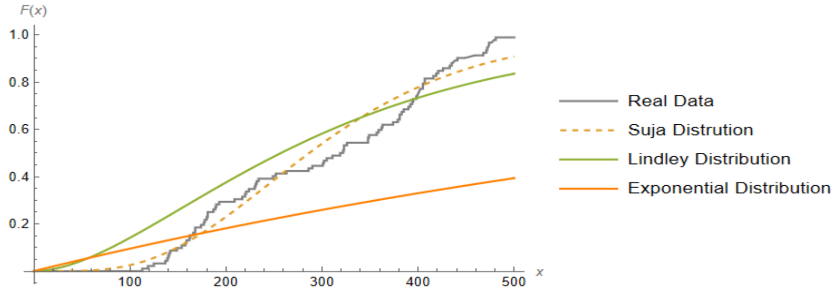


Figure 1: The models' Empirical Distribution and Estimated CDF for Data Set 1.

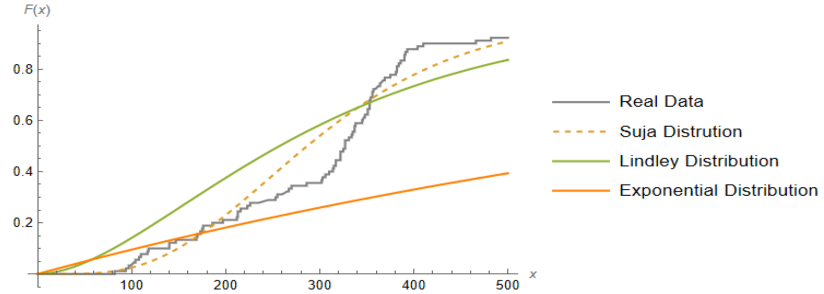


Figure 2: The models' Empirical Distribution and Estimated CDF for Data Set 2.

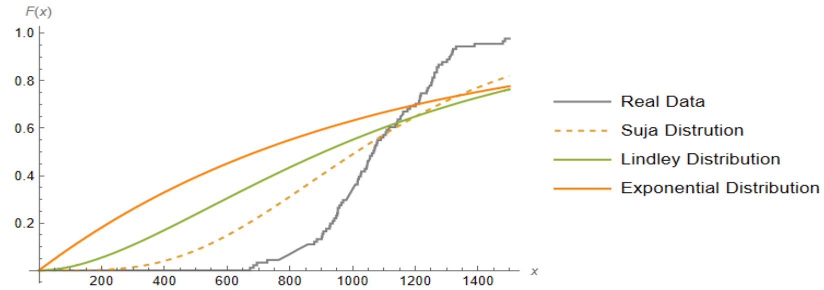


Figure 3: The models' Empirical Distribution and Estimated CDF for Data Set 3.

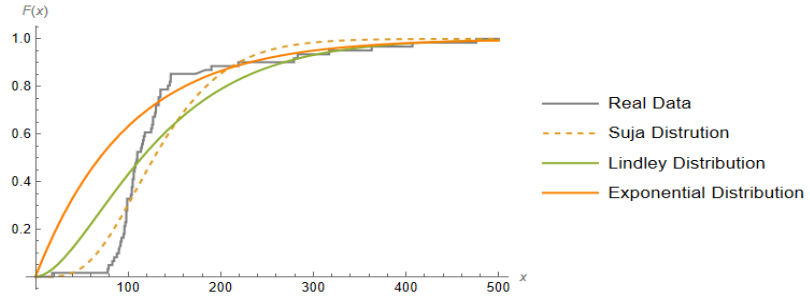


Figure 4: The models' Empirical Distribution and Estimated CDF for Data Set 4.

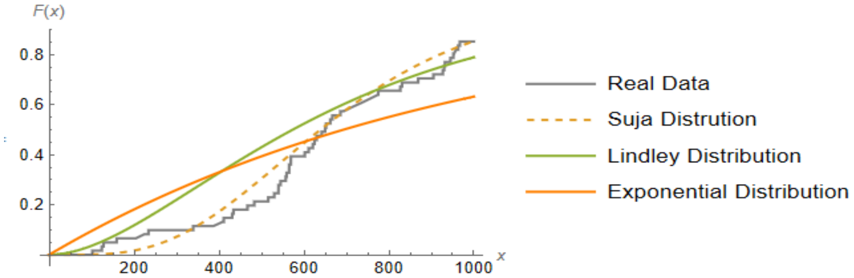


Figure 5: The models' Empirical Distribution and Estimated CDF for Data Set 5.

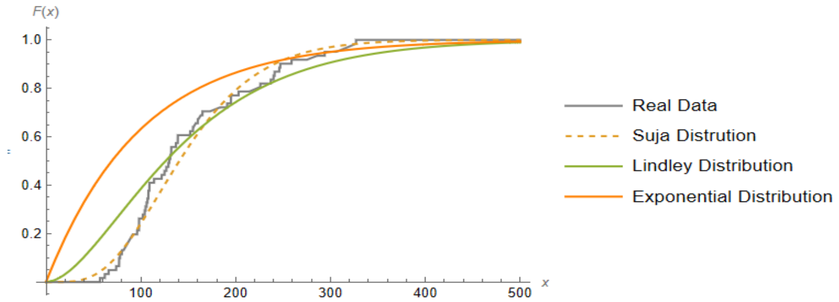


Figure 6: The models' Empirical Distribution and Estimated CDF for Data Set 6.

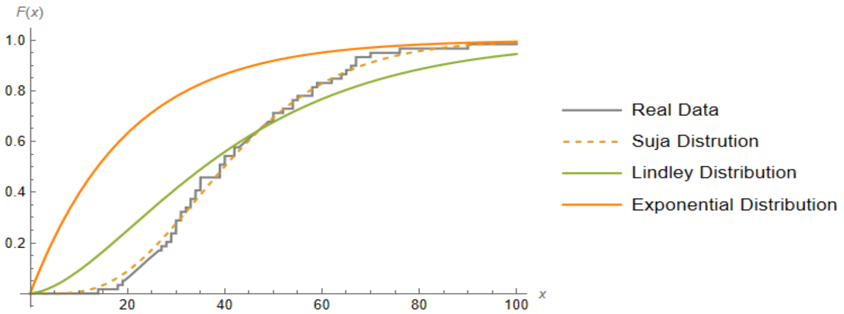


Figure 7: The models' Empirical Distribution and Estimated CDF for Data Set 7.

## V. Conclusion

From the data number of daily COVID-19 cases in Saudi Arabia, the following observations can be concluded:

1- The ML estimator of theta parameter for distributions is good estimate based on the AIC, BIC, HQIC and CAIC measures.

2- In all cases, the AIC has the smallest results which makes it the best measure for these data.

3- The Suja distribution is better than the other distributions in modeling lifetime data based on the goodness of fit measures.

4- When studying the data, it is preferable that the period does not exceed three months.

## References

[1]. Abu-Zinadah, H. H. (2014). Goodness-Of-Fit Tests For The Exponentiated Gompertz Distribution. International Journal Of Innovation In Science And Mathematics, 2(4), 2347–9051.
[2]. Abu-Zinadah, H., & Binkhamis, A. (2020). Goodness-Of-Fit Tests For The Beta Gompertz Distribution. Thermal Science, 24(Suppl. 1), 69-81.
[3]. Al-Omari, A. I., Alhyasat, K., Ibrahim, K., & Abu Bakar, M. A. (2019). Power Length-Biased Suja Distribution: Properties And Application. Electronic Journal Of Applied Statistical Analysis, 12(2), 429-452.
[4]. Alshenawy, R. (2020). A New One Parameter Distribution: Properties And Estimation With Applications To Complete And Type II Censored Data. Journal Of Taibah University For Science, 14(1), 11-18.
[5]. Shanker, R. (2017). Suja Distribution And Its Application. International Journal Of Probability And Statistics, 6(2), 11-19.
[6]. Shankar, P. M. (2019). Pedagogy Of Chi-Square Goodness Of Fit Test For Continuous Distributions. Computer Applications In Engineering Education, 27(3), 679-689.
[7]. Sharma, V., Shanker, R., & Shanker, R. (2019). On Some One Parameter Lifetime Distributions And Their Applications. Annals Of Biostatistics & Biometric Applications, 3(2), 2641-6336.
[8]. Whittaker, T. A., & Furlow, C. F. (2009). The Comparison Of Model Selection Criteria When Selecting Among Competing Hierarchical Linear Models. Journal Of Modern Applied Statistical Methods, 8(1), 173-193.