*Journal of Probability and Statistical Science*
*24(1), 1-15 Feb. 2026*

# Length-biased Sujit Distribution with Survival Data Analysis

O. Anu
Department of Statistics
Annamalai University, Tamil Nadu, India

P. Pandiyan
Department of Statistics
Annamalai University, Tamil Nadu, India

**ABSTRACT**

In this study, we introduce a new model called the Length-Biased Sujit Distribution (LBSD), derived from the original Sujit distribution. The proposed distribution is formulated by applying the concept of length-biased transformation to the Sujit distribution, making it more suitable for modeling lifetime data where longer durations have a higher chance of observation. Several statistical characteristics of the LBSD, such as moments, moment-generating function, reliability measures, entropy measures (Renyi and Tsallis entropies), Bonferroni and Lorenz curves, and order statistics, are derived and discussed. The parameters of the model are estimated using the Maximum Likelihood Estimation (MLE) method. Finally, the applicability of the proposed model is demonstrated using a real-life dataset on bone cancer, showing that the Length-Biased Sujit Distribution provides a superior fit compared to the original Sujit distribution.

***Keyword:*** Length biased, One parameter, Cancer data, Mathematical Properties, Order Statistics, Entropy, Pdf-Cdf-Survival-hazard Plot.

## 1. Introduction

In several applied disciplines such as medicine, engineering, insurance, and finance, lifetime data modeling plays a crucial role in understanding the behavior and characteristics of time-to-event phenomena. Numerous continuous probability distributions, including the exponential, Weibull, gamma, Lindley, and lognormal distributions, as well as their generalizations, have been employed to describe lifetime data. However, in many cases, traditional distributions may not adequately represent real-world data due to limitations in their shape flexibility and hazard rate behavior. To overcome these shortcomings, several one-parameter lifetime distributions such as Shanker, Aradhana, Devya, Sujatha, Akash, and Suja distributions were proposed by Shanker and collaborators between 2015 and 2017, offering improved modeling capabilities over the exponential and Lindley models.

The Sujit distribution was later introduced as a one-parameter lifetime distribution that represents a mixture of gamma and exponential distributions, providing a better fit for various types of lifetime data. Its mathematical tractability and flexibility in modeling increasing and decreasing hazard rate functions make it an appealing choice in reliability and survival analysis. However, real-world data often exhibit selection or ascertainment bias, where larger or longer observations are more likely to be included in the sample. To address this issue, Fisher (1934) first introduced the concept of weighted distributions, which was later generalized by Rao (1965) to model such biases effectively. Among these, length-biased distributions are a special class of weighted models that give greater probability to larger observations, making them particularly useful in lifetime data analysis.

Motivated by these considerations, the present study introduces a new model referred to as the Length-Biased Sujit Distribution (LBSD), derived from the original Sujit distribution. The proposed model accounts for the inherent length bias in lifetime data and extends the applicability of the Sujit distribution by providing greater flexibility in its shape and reliability structures. Several statistical properties of the LBSD, such as moments, moment-generating function, hazard rate, mean residual life, entropy measures (Renyi and Tsallis entropies), Bonferroni and Lorenz curves, stochastic ordering, and order statistics, are derived and

- Received October 13, 2025, in final form February 2026.
- O. Anu (corresponding author) is affiliated with Department of Statistics, Annamalai University, Tamil Nadu, India.
  anuabd0115@gmail.com

analysed in detail. Parameter estimation is carried out using the Maximum Likelihood Estimation (MLE) method. Finally, the efficiency and suitability of the proposed distribution are illustrated through an application to a real-life breast cancer dataset, demonstrating that the Length-Biased Sujit Distribution provides a superior fit compared to the original Sujit distribution.

The probability density function of Sujit distribution (PDF)

$$f(x,\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}dx \qquad (1)$$

The Cumulative distribution function of Sujit distribution (CDF)

$$F(x,\theta) = 1 - \left[1 + \frac{\theta x}{\theta+1}\right]e^{-\theta x} \qquad (2)$$

**The Length Biased Sujit Distribution (LBSD)**

The probability density function of the Length Biased Sujit Distribution is given by

$$f_l(x) = \frac{w(x)f(x)}{E\left(w(x^2)\right)}; x > 0,$$

$$f_l(x) = \frac{xf(x)}{E(x^2)}; x > 0$$

Where $w(x)$ be a non-negative weight function and $E\big(w(x)\big) = \int w(x)f(x)dx < \infty$.

$$f_l(x) = \frac{xf(x)}{E(x^2)}; x > 0$$

Where,

$$E(x^2) = \int_0^\infty x\, f(x;\theta)dx \qquad (3)$$

$$= \int_0^\infty x\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}dx$$

$$= \frac{\theta^2}{\theta+1}\int_0^\infty x(1+x)e^{-\theta x}\,dx$$

$$= \frac{\theta^2}{\theta+1}\int_0^\infty x\,e^{-\theta x}dx + \int_0^\infty x^2e^{-\theta x}dx$$

$$= \frac{\theta^2}{\theta+1}\int_0^\infty x^{2-1}e^{-\theta x}dx + \int_0^\infty x^{3-1}e^{-\theta x}dx$$

$$= \frac{\theta^2}{\theta+1}\left[\frac{\Gamma 2}{\theta^2} + \frac{\Gamma 3}{\theta^3}\right]$$

$$= \frac{\theta^2}{\theta+1}\left[\frac{\theta\,\Gamma 2 + \Gamma 3}{\theta^3}\right]$$

$$= \frac{\theta^2}{\theta+1}\left[\frac{\theta+2}{\theta^3}\right]$$

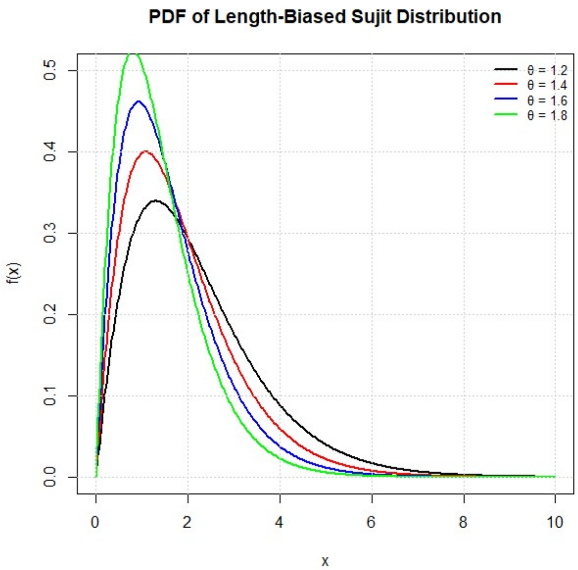
$$E(x^2) = \frac{\theta+2}{\theta(\theta+1)} \qquad (4)$$

Substitute (1) and (4) in equation (3), and we will get the required probability density function of Length-biased Sujit distribution as

$$f_l(x) = \frac{xf(x)}{E(x^2)}$$

$$= \frac{\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}dx}{\frac{\theta+2}{\theta(\theta+1)}}$$

$$f_l(x) = \frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta x}dx \qquad (5)$$




The cumulative distribution function (cdf) of the length-biased Sujit distribution (LBSD).

$F_l(x) = \int_0^x f_l(x)dx$ (6)

$F_l(x) = \frac{\theta^3}{\theta+2} x\,(1+x)e^{-\theta}\ dx$

$= \frac{\theta^3}{\theta+2}\int_0^x x(1+x)\ e^{-\theta x}\,dx$

put $x = \frac{z}{\theta}$, $x = \theta z, dx = \frac{dz}{\theta}$

when $x \to 0, t \to 0$, and $x \to x, t \to \theta x$

$= \frac{\theta^3}{\theta+2}\int_0^{\theta x} x\,e^{-\theta}\ \ dx\ \int_0^{\theta x} x^2\,e^{-\theta}\ \ dx$

$= \frac{\theta^3}{\theta+2}\int_0^{\theta x}\left(\frac{z}{\theta}\right)e^{-z}\ \frac{dz}{\theta} + \int_0^{\theta x}\left(\frac{z}{\theta}\right)^2\ e^{-z}\ \frac{dz}{\theta}$

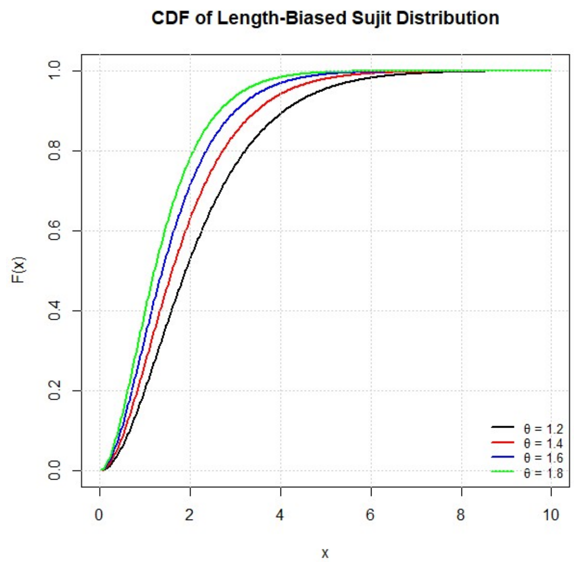
$= \frac{\theta^3}{\theta+2}\int_0^{\theta x}\frac{1}{\theta^2} z e^{-z} dz + \int_0^{\theta x}\frac{1}{\theta^3} z^2 e^{-z} dz$

$= \frac{\theta^3}{\theta+2}\frac{1}{\theta^2}\int_0^{\theta x} z^{2-1}e^{-z}dz + \frac{1}{\theta^3}\int_0^{\theta x} z^{3-1}e^{-z}dz$

$= \frac{\theta^3}{\theta+2}\left(\frac{1}{\theta^2}\gamma(2,\theta x) + \frac{1}{\theta^3}\gamma(3,\theta x)\right)$

$= \frac{\theta^3}{\theta+2}\left(\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta^3}\right)$

$F_l(x) = \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}$ (7)




## 2. Reliability Analysis

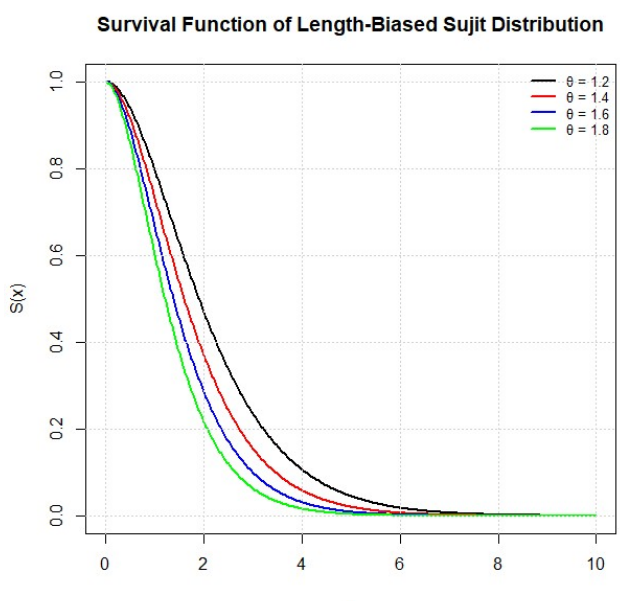
We will discuss the survival function, failure rate, reverse hazard rate and the Mills ratio of the Length-biased Sujit distribution (LBSD).

**Survival function**

The survival function of the Length-biased Sujit distribution is given by

$$S(x) = 1 - F_l(x;\theta)$$

$$S(x) = 1 - \left(\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}\right) \tag{8}$$



**Hazard function**

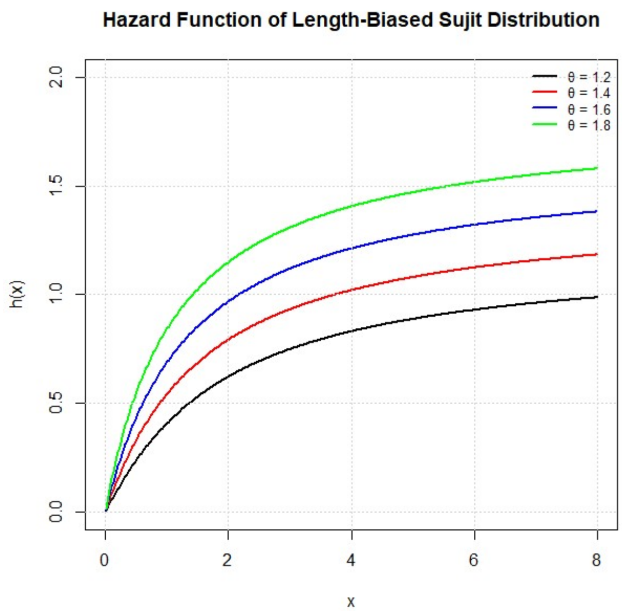
The hazard function is also known as the hazard rate is given by

$$h(x)=\frac{f(x)}{1-F(x)}$$

$$h(x)=\frac{f_l(x;\theta)}{1-F_l(x;\theta)}$$

$$h(x)=\frac{\frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta}\ dx}{1-\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}}$$

$$h(x)=\left(\frac{\theta^3\ \ x\,(1+x)e^{-\theta}\ dx}{\theta+2-\theta\ \ (2,\theta x)+\gamma(3,\theta x)}\right) \tag{9}$$




**The Reverse hazard rate**

The reverse hazard rate is given by

$$h_r(x)=\frac{f_l(x)}{F_l(x)}$$

$$h_r(x)=\frac{f_l(x;\theta)}{F_l(x;\theta)}$$

$$h_r(x)=\frac{\frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta x}dx}{\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}}$$

$$h_r(x)=\left(\frac{\theta^3\,x\,(1+x)e^{-\theta x}dx}{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}\right) \tag{10}$$

**Odds Rate function**

Odds Rate function of the Length biased Sujit distribution

$$O(x)=\frac{F_l(x)}{1-F_l(x)}$$

$$O(x)=\frac{F_l(x;\theta)}{1-F_l(x;\theta)}$$

$$O(x)=\left(\frac{\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}}{1-\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}}\right)$$

$$O(x)=\left(\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2-\theta\gamma(2,\theta x)+\gamma(3,\theta x)}\right) \tag{11}$$

**Cumulative hazard rate function**

Cumulative hazard rate function of the Length biased Sujit distribution

$H(x) = -\ln(1 - F_l(x)$

$H(x) = -\ln(1 - F_l(x;\theta)$

$$H(x) = -\ln\left(1 - \frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}\right) \quad (12)$$

**Mills Ratio**

And the Mills ratio of the Length-biased Sujit distribution is

Mills Ratio $= \frac{1}{h_r(x)}$

Mills Ratio $= \left(\frac{1}{\frac{\theta^3}{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}x\,(1+x)e^{-\theta x}dx}\right)$

$$\text{Mills Ratio} = \left(\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta^3}x\,(1+x)e^{-\theta x}dx\right) \quad (13)$$

**MOMENTS AND ASSOCIATED MEASURES**

Let X denote the random variable of Length-biased Sujit distribution with parameter $\theta$ then the $r^{th}$order moments $E(X^r)$of Length-biased Sujit distribution are obtained as

$E(X^r) = \mu_r' = \int_0^\infty x^r\, f_l(x)dx$

$= \int_0^\infty x^r \frac{\theta^3}{\theta+2} x\,(1+x)e^{-\theta x}dx$

$= \frac{\theta^3}{\theta+2}\int_0^\infty x^{r+1}\,(1+x)\,e^{-\theta x}\,dx$

$= \frac{\theta^3}{\theta+2}\int_0^\infty x^{r+1}\,e^{-\theta}\quad dx + \int_0^\infty x^{r+2}\,e^{-\theta}\quad dx$

$= \frac{\theta^3}{\theta+2}\int_0^\infty x^{(r+2)-1}\,e^{-\theta}\quad dx + \int_0^\infty x^{(r+3)-1}\,e^{-\theta x}\,dx$

$= \frac{\theta^3}{\theta+2}\left(\frac{\Gamma r+2}{\theta^{r+2}} + \frac{\Gamma r+3}{\theta^{r+3}}\right)$

$= \frac{\theta^3}{\theta+2}\left(\frac{\theta\,\Gamma r+2+\,\Gamma r+3}{\theta^{r+3}}\right)$

$$E(X^r) = \mu_r' = \frac{\theta\,\Gamma r+2+\,\Gamma r+3}{\theta^r(\theta+2)} \quad (14)$$

Put r = 1,2, in the equation, we will obtain the first raw moments of length-biased Sujit distribution, which is given by

If r=1,

$E(X^1) = \mu_1' = \frac{\theta\,\Gamma 1+2+\,\Gamma 1+3}{\theta(\theta+2)}$

$\mu_1' = \frac{\theta\,\Gamma 3+\,\Gamma 4}{\theta(\theta+2)}$

$\mu_1' = \frac{\theta\;2!+\,3!}{\theta(\theta+2)}$

If r=2,

$E(X^2) = \mu_2' = \frac{\theta\,\Gamma 2+2+\,\Gamma 2+3}{\theta(\theta+2)}$

$\mu_2' = \frac{\theta\,\Gamma 4+\,\Gamma 5}{\theta(\theta+2)}$

$\mu_2' = \frac{\theta\;3!+\,4!}{\theta(\theta+2)}$

Variance = $\mu_2' - (\mu_1')^2$

$= \left(\frac{\theta\;3!+\,4!}{\theta(\theta+2)} - \left(\frac{\theta\;2!+\,3!}{\theta(\theta+2)}\right)^2\right)$

$$= \left(\frac{\theta\ 3! + 4!}{\theta(\theta+2)} - \frac{\theta\ 2! + 3!^2}{\left(\theta(\theta+2)\right)^2}\right) \tag{15}$$

**Harmonic mean**

The Harmonic mean of the proposed model can be obtained as

$$H.M = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_l(x)dx$$

$$= \int_0^\infty \frac{1}{x}\frac{\theta^3}{\theta+2} x\ (1+x)e^{-\theta x}dx$$

$$= \frac{\theta^3}{\theta+2}\int_0^\infty (1+x)e^{-\theta x}dx$$

$$= \frac{\theta^3}{\theta+2}\int_0^\infty e^{-\theta x}\,dx + \int_0^\infty x\ e^{-\theta x}\,dx$$

$$= \frac{\theta^3}{\theta+2}\left(\frac{1}{\theta} + \frac{\Gamma 2}{\theta^2}\right)$$

$$= \frac{\theta^3}{\theta+2}\left(\frac{\theta + \Gamma 2}{\theta^2}\right)$$

$$= \frac{\theta^3}{\theta+2}\left(\frac{\theta+1}{\theta^2}\right)$$

$$= \frac{\theta(\theta+1)}{\theta+2} \tag{16}$$

**Moment Generating Function and Characteristics Function**

Suppose the random variable X follows Length biased Sujit distribution with parameters $\theta$, then the MGF of X can be obtained as:

$$M_X(t) = E(e^{tx})$$

$$= \int_0^\infty e^{tx} f_l(x;\theta)dx \tag{17}$$

Using Taylor's Series Expansion

$$M_X(t) = \int_0^\infty \left[1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots\right]$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!}\int_0^\infty x^j f_l(x;\theta)dx$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!}\mu_j'$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!}\left(\frac{\theta\,\Gamma j+2 + \Gamma j+3}{\theta^j(\theta+2)}\right) \tag{18}$$

$$= \frac{1}{\theta+2}\sum_{j=0}^{\infty} \frac{t^j}{j!\theta^j}(\theta\,\Gamma j+2 + \Gamma j+3)$$

Similarly, the Characteristics function of Length-biased Sujit Distribution can be obtained by

$$\phi_X(t) = M_X(it)$$

$$M_X(it) = \frac{1}{\theta+2}\sum_{j=0}^{\infty} \frac{t^j}{j!\theta^j}(\theta\,\Gamma j+2 + \Gamma j+3) \tag{19}$$

**Order Statistics**

Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be the order statistics of a random sample $X_1, X_2, \ldots, X_n$ drawn from the continuous population with probability density function $f_x(x)$ and cumulative density function with $F_x(x)$, then the probability density function of $r^{th}$ order statistics $X_{(r)}$ is given by

$$f_{X(r)}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x)[F_X(x)]^{r-1}[1-F_X(x)]^{n-r}$$

The probability density function of $r^{th}$ order statistics $X_{(r)}$ of Length-biased Sujit distribution is given by

$$= \frac{n!}{(r-1)!(n-r)!}\left(\frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta x}dx \times \left(\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}\right)^{r-1} \times \left(1-\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}\right)^{n-r}\right) \tag{20}$$

Therefore, the Probability density function of first Order Statistics $X_1$ of Length-biased Sujit distribution is can be obtained as

$$f_{X(1)}(x) = \frac{n(n-1)!}{(1-1)!(n-1)!}\left(\left(\frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta x}dx\right) \times \left(1-\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}\right)\right)^{n-r}$$

$$f_{X(n)}(x) = \frac{n!}{(n-1)!(1-1)!}\left(\left(\frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta x}dx\right) \times \left(\frac{\theta\gamma(2,\theta x)+\gamma(3,\theta x)}{\theta+2}\right)\right)^{n-1} \tag{21}$$

**Likelihood Ratio Test**

The likelihood-ratio test is a hypothesis test that involves comparing the goodness of fit of two competing statistical models, typically one found by maximization over the entire parameter space and another found after imposing some constraint, based on the ratio of their likelihoods. If the more constrained model (i.e., the null hypothesis) is supported by the observed data, the two likelihoods should not differ by more than sampling error. Thus the likelihood-ratio test tests whether this ratio is significantly different from one, or equivalently whether its natural logarithm is significantly different from zero**.**

Let $X_1, X2_{,\dots,}X_n$ be a random sample from the Length-biased Sujit distribution. To test the hypothesis

$H_0: f(x) = f(x;\theta)$ against $H_{1:} f(x) = f_l(x;\theta)$

In order to test whether the random sample of size n comes from the Length-biased Sujit distribution, the following test statistics is used by

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n}\frac{f_l(x;\theta)}{f(x;\theta)}$$

$$= \prod_{i=0}^{n}\left(\frac{\frac{\theta^3}{\theta+2}x\,(1+x)e^{-\theta x}dx}{\frac{\theta^2}{\theta+1}(1+x)e^{-\theta}\ dx}\right)$$

$$= \prod_{i=0}^{n}\left(\frac{\theta+1}{\theta^2}\right) \times \left(\frac{\theta^3}{\theta+2}\right)x_i$$

$$= \prod_{i=0}^{n}\frac{\theta(\theta+1)}{\theta+2}x_i$$

$$= \left(\frac{\theta(\theta+1)}{\theta+2}\right)^n \prod_{i=0}^{n}x_i$$

We should reject the null hypothesis, if

$\Delta = \left(\frac{\theta(\theta+1)}{\theta+2}\right)^n \prod_{i=0}^{n}x_i > k$ (or)

Equivalently, We shall reject the null hypothesis, if

$\Delta^* = \prod_{i=0}^{n}x_i > k\left(\frac{\theta(\theta+1)}{\theta+2}\right)^n$

$\Delta^* = \prod_{i=0}^{n}x_i > k^*\ where,$

$k^* = k\left(\frac{\theta(\theta+1)}{\theta+2}\right)^n$

Then

$p(\Delta^* > \lambda^*)$, where, $\lambda^* = \prod_{i=0}^{n}x_i$ is less than a specified level of significance, and $\prod_{i=0}^{n}x_i$ is the observed value of $\Delta^*$

**Maximum Likelihood Estimate and Fisher Information Measure**

The Maximum likelihood estimation (MLE) is a method of estimating the parameters of an assumed probability distribution, given some observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable. The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate. The logic of maximum likelihood is both intuitive and flexible, and as such the method has become a dominant means of statistical inference.

The Fisher information is a way of measuring the amount of information that an observable random variable *X* carries about an unknown parameter $\theta$ of a distribution that models *X*. Formally, it is the variance of the score, or the expected value of the observed information.

$L(x) = \prod_{i=1}^{n} f_l\,(x)$

$L(x) = \prod_{i=1}^{n} \frac{\theta^3}{\theta+2} x_i\,(1+x)e^{-\theta x_i}dx$

$$L(x) = \frac{\theta^{3n}}{(\theta+2)^n} \prod_{i=1}^{n} x_i(1+x)e^{-\theta x_i} \quad (22)$$

The log likelihood function is given by

$= n\,log(\theta^3) - n\log(\theta+2) + 2\sum_{i=1}^{n}\log x_i(1+x_i) - \theta\sum_{i=1}^{n} x_i$

For the purpose of obtaining the confidence interval we use the asymptotic normality results. We have that if $\hat{\lambda} = (\hat{\theta})$ denotes the MLE of $\lambda = (\theta)$ We can state the results as follows

$\sqrt{n}\,(\hat{\lambda}-\lambda) \rightarrow N_2(0, I^{-1}(\lambda))$

Where, $I(\lambda)$ is Fisher's Information Matrix. i.e.,

$I(\lambda) = -\frac{1}{n}\left[E\left[\frac{\partial^2 logL}{\partial\theta^2}\right]\right]$

Where,

$\left[\frac{\partial^2 logL}{\partial\theta^2}\right] = E\left[\frac{\partial}{\partial\theta}\left(\frac{\partial logL}{\partial\theta}\right)\right]$

$$\left[\frac{\partial^2 logL}{\partial\theta^2}\right] = \frac{3n}{\theta} = \theta(0) - \frac{3n(1)}{\theta^2} = \frac{-3n}{\theta^2} \quad (23)$$

**Bonferroni and Lorenz Curves**

The Bonferroni and Lorenz curves are used in economics in order to study income, etc., but they are used in other fields like demography, insurance, medicine, and reliability. The Bonferroni and Lorenz curves are given by

$B(p) = \frac{1}{p\mu_1'}\int_0^q x_i f_l(x)dx$

$B(p) = \frac{1}{p\mu_1'}\int_0^q x_i f_l(x;\theta)dx$ and

$L(p) = \frac{1}{\mu_1'}\int_0^q x_i f_l(x;\theta)dx$

Where, $q = F^{-1}(p);\ q\epsilon[0,1]$ and $\mu = (x)$

Hence, the Bonferroni and Lorenz curves of our distribution are given by,

$\mu = \frac{2\theta+6}{\theta(\theta+2)}$

$B(p) = \frac{1}{P\left[\frac{\theta(\theta+2)}{2\theta+6}\right]}\int_0^q \frac{\theta^3}{\theta+2} x\,(1+x)e^{-\theta x}dx$

$= \frac{\theta[\theta+2]}{P[2\theta+6]} \times \frac{\theta^3}{\theta+2} x\,(1+x)e^{-\theta x}dx$

$= \frac{\theta^4}{P[2\theta+6]}\int_0^\infty x\,(1+x)e^{-\theta x}dx$

$= \frac{\theta^4}{P[2\theta+6]} \int_0^q x\, e^{-\theta x}\, dx + \int_0^q x^2\, e^{-\theta x}\, dx$

Put, $x = \frac{t}{\theta}, \qquad \theta q = t, \quad dx = \frac{1}{\theta} dt$

when $x \to 0, t \to 0$, and $x \to x, t \to \theta q$

$= \frac{\theta^4}{P[2\theta+6]} \int_0^{\theta q} \left[ \left(\frac{t}{\theta}\right)^1 + \left(\frac{t}{\theta}\right)^2 \right] e^{-t} \frac{1}{\theta} dt$

$= \frac{\theta^4}{P[2\theta+6]} \int_0^{\theta q} \left[ \frac{t}{\theta^2} + \frac{t^2}{\theta^3} \right]$

$= \frac{\theta^4}{P[2\theta+6]} \int_0^{\theta q} \left[ \frac{\theta t + t^2}{\theta^3} \right]$

$= \frac{\theta}{P[2\theta+6]} \int_0^{\theta q} [\theta t + t^2]\, e^{-t} dt$

$= \frac{\theta}{P[2\theta+6]} \int_0^{\theta q} t^{2-1} e^{-t} dt + \int_0^{\theta q} t^{3-1} e^{-t} dt$

$$B(p) = \frac{\theta}{P[2\theta+6]} [\theta \gamma(2, \theta q) + \gamma(3, \theta q)] \qquad (24)$$

**Lorenz Curves**

The Lorenz curves are used in economics in order to study income, etc., but they are used in other fields like demography, insurance, medicine, and reliability. The Lorenz curves are given by

$L(p) = B(p)$

$$L(p) = \frac{\theta}{P[2\theta+6]} [\theta \gamma(2, \theta q) + \gamma(3, \theta q)] \qquad (25)$$

**Entropies**

Entropy is important in different areas such as probability and economics, communication theory, physics, and statistics. Entropies are applied to quantify a system's diversity, uncertainty, or randomness. An indicator of the uncertainty's variation is the entropy of a random variable X.

**Shannon Entropy**

$S_\lambda = -\int_0^\infty f_l\,(x) log\big(f_l(x)\big) dx$

$$S_\lambda = -\int_0^\infty \frac{\theta^3}{\theta+2} x\,(1+x) e^{-\theta}\; dx \;\; log\left( \frac{\theta^3}{\theta+2} x\,(1+x) e^{-\theta x} dx \right) \qquad (26)$$

**Renyi Entropy**

The Renyi entropy is significant as a diversity index. The Renyi entropy is also important in quantum information. It can be used as a measure of entanglement for a given probability distribution. Renyi entropy is given by

$R_\lambda = \frac{1}{1-\lambda} log \int_0^\infty \big(f(x)\big)^\lambda dx\,;\ \lambda > 0, \lambda \neq 1$

$R_\lambda = \frac{1}{1-\lambda} log \int_0^\infty \big(f(x;\theta)\big)^\lambda dx$

$R_\lambda = \frac{1}{1-\lambda} log \int_0^\infty \left( \frac{\theta^3}{\theta+2} x\,(1+x) e^{-\theta x} dx \right)^\lambda dx$

$$R_\lambda = \frac{1}{1-\lambda} log \left(\frac{\theta^3}{\theta+2}\right)^\lambda \int_0^\infty x^\lambda (1+x)^\lambda\, e^{-\lambda \theta x} dx \qquad (27)$$

Using binomial expansion

$(1 + x + x^4)^\lambda = \sum \binom{n}{r} a^{n-r} d^r$

$= \sum_{r=0}^n \binom{n}{r} a^{n-r} (b+c)^r$

$$= \sum_{r=0}^{n} \binom{n}{r} a^{n-r} \sum_{s=0}^{n} \binom{r}{s} b^{r-s} c^{s}$$

$$R_\lambda = \frac{1}{1-\lambda} log\left(\left(\frac{\theta^3}{\theta+2}\right)\right)^\lambda = \sum_{i=0}^{\lambda} \binom{\lambda}{i} x^{i} (1+x)^{\lambda-i} e^{-\lambda\theta x} dx$$

$$R_\lambda = \frac{1}{1-\lambda} log\left(\left(\frac{\theta^3}{\theta+2}\right)\right)^\lambda = \sum_{i=0}^{\lambda} \binom{\lambda}{i} x^{\lambda i} (1+x)^{\lambda-i} e^{-\lambda\theta x} dx$$

$$R_\lambda = \frac{1}{1-\lambda} log\left(\left(\frac{\theta^3}{\theta+2}\right)\right)^\lambda = \sum_{j=1}^{i} \binom{i}{j} x^{(i-j)} \; e^{-\lambda\theta} \; dx$$

$$R_\lambda = \frac{1}{1-\lambda} log\left(\left(\frac{\theta^3}{\theta+2}\right)\right)^\lambda \sum_{i=1}^{\lambda} \sum_{j=0}^{i} \binom{\lambda}{i} \binom{i}{j} \int_0^\infty x^{(\lambda+i-j+1)-1} e^{-\lambda\theta x} dx$$

$$R_\lambda = \frac{1}{1-\lambda} log\left(\left(\frac{\theta^3}{\theta+2}\right)\right)^\lambda \sum_{i=1}^{\lambda} \sum_{j=0}^{i} \binom{\lambda}{i} \binom{i}{j} \frac{\Gamma x^{(\lambda+i-j+1)}}{(\theta)^{(\lambda+i-j+1)}} e^{-\lambda\theta x} dx$$

$$R_\lambda = \frac{1}{1-\lambda} log\left(\left(\frac{\theta^3}{\theta+2}\right)\right)^\lambda \sum_{i=1}^{\lambda} \sum_{j=0}^{i} \binom{\lambda}{i} \binom{i}{j} \left(\frac{1}{\theta^\lambda}\right)^{(\lambda+i-j+1)} \Gamma\lambda + i - j + 1 \qquad (28)$$

**Tsallis Entropy**

The Boltzmann-Gibbs (B-G) statistical property generalization initiated by Tsallis has received a great deal of attention. This B-G statistic was first introduced as the mathematical expansion of Tsallis entropy (Tsallis, 1988) for continuous random variables; this generalization of B-G was introduced in order to suggest. Which is defined as

$$T_\lambda = \frac{1}{\lambda-1}\left(1 - \int_0^\infty \left(f_l(x;\theta)\right)^\lambda dx\right) \lambda > 0, \lambda \neq 1$$

$$T_\lambda = \frac{1}{\lambda-1}\left(1 - \int_0^\infty \left(\frac{\theta^3}{\theta+2}\right) x(1+x)\, e^{-\theta x}\right)^\lambda dx$$

$$T_\lambda = \frac{1}{\lambda-1}\left(1 - \left(\frac{\theta^3}{\theta+2}\right)^\lambda \int_0^\infty x^\lambda (1+x)^\lambda e^{-\lambda\theta x} dx\right) \qquad (29)$$

Using binomial expansion

$$= \sum_{i=0}^{\lambda} \binom{\lambda}{i} (1+x)^{\lambda-i}$$

$$T_\lambda = \frac{1}{1-\lambda}\left(1 - \left(\frac{\theta^3}{\theta+2}\right)^\lambda \sum_{i=1}^{\lambda} \sum_{j=0}^{i} \binom{\lambda}{i} \binom{i}{j} \frac{\Gamma x^{(\lambda+i-j+1)}}{(\theta)^{(\lambda+i+3j+1)}}\right)$$

$$T_\lambda = \frac{1}{1-\lambda}\left(1 - \left(\frac{\theta^3}{\theta+2}\right)^\lambda \sum_{i=1}^{\lambda} \sum_{j=0}^{i} \binom{\lambda}{i} \binom{i}{j} \left(\frac{1}{\theta^\lambda}\right)^{(\lambda+i-j+1)} \Gamma\lambda + i - j + 1\right) \qquad (30)$$

## 3. Data Analysis

In this study, we utilize the breast cancer dataset reported by Lee [9]. The data were collected at a large hospital between 1929 and 1938 and include the survival times of 121 breast cancer patients. The observations are as follows:
0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0 and 154.0.

| **Distribution** | **MLE (θ)** | **Std. Error** | **-2 Log-Likelihood** | **AIC** | **AICC** | **BIC** |
|---|---|---|---|---|---|---|
| Fuyi | 0.050 | 0.007 | 1345.2 | 1347.2 | 1347.3 | 1351.0 |
| Shreekant | 0.072 | 0.009 | 1312.8 | 1314.8 | 1314.9 | 1318.5 |
| Sujit | 0.085 | 0.010 | 1298.5 | 1300.5 | 1300.6 | 1304.2 |
| Length-biased Sujit | 0.092 | 0.011 | 1270.3 | 1272.3 | 1272.4 | 1276.0 |

## 4. Conclusion

In this paper, a new one-parameter lifetime distribution called the **Length-Biased Sujit Distribution (LBSD)** has been proposed as an extension of the original Sujit distribution. Several important statistical properties of the proposed model, including its probability density function, cumulative distribution function, survival and hazard rate functions, moments, and moment-generating function, have been derived and discussed in detail. The parameters of the distribution were estimated using the Maximum Likelihood Estimation (MLE) method. The LBSD is particularly suitable for modeling lifetime data arising from engineering, medical, and biological studies where longer lifetimes are more likely to be observed.

An application to a real-life breast cancer dataset demonstrates that the proposed Length-Biased Sujit Distribution provides a better fit compared to the original Sujit distribution, showing greater flexibility and improved reliability characteristics. Hence, the LBSD can be considered a more appropriate model for analyzing lifetime data affected by length bias.

**Reference**

[1]. Poovitha, R., and Pandiyan, P. (2025). Stochastic model for new extended Distribution with properties and application to Survival Time of Cancer Patients. *Indian Journal of Natural Science*, 16 (90), 94890-94901.

[2]. Sakthivel, M., and Pandiyan, P. (2024). A Stochastic Model for Length Biased Uma Distribution with Properties and its applications. *Indian Journal of Natural Sciences*, 15(84), 75906-75922.

[3]. Sakthivel, M., and Pandiyan, P. (2024). A Stochastic Model for Length biased Loai Distribution with Properties and its Applications. *International journal of Mathematics and Computer Research*, 12 (2320-7167), 4177-4195.

[4]. Kavitha, R., and Pandiyan, P. (2021). EXPECTED TIME USING PARETO DISTIRIBUTION THROUGH SHOCK MODEL FOR LUNGS CANCER. *Journal of Xi' an University of Architecture & Technology.*, XIII (11), 287-294.

[5]. Subramanian. C., and, Shenbagaraja, R. (2020). Length Biased Quasi-sujatha distribution with properties and applications to Blader Cancer Data. *Journal of Information and Computational Science*, 10(1), 493-506.

[6]. Aafaq, A. R., and Subramanian, C. (2019). Length-Biased Erlang-Truncated Exponential Distribution with Life Time Data. *International Journal of Information and Computational Sciences*, 9(8), 340-355.

[7]. Rather, A., and Subramanian., C. (2018). Length Biased Sushila Distribution. *Universal Review*, 7(XII), 1010-1023.

[8]. Elizabeth, T, Sangeetha, and Pandiyan, P. (2021). A STOCHASTIC MODEL FOR CHRONIC KIDNEY DISEASE. *Advances in Mathematics Scientific journal*, 10(10), 3363-3372.

[9]. Anu, O., and Pandiyan, P. (2025). A new generalization of fuyi distribution and its applications to the cancer data. *International Journal of Statistics and Applied Mathematics*, 10(7), 95-106.

[10]. Anu, O., and Pandiyan, P. (2025). A new extended Remkan distribution and its application to the cancer data. *Brazilian journal of Biometrics*, 43(800), 1-15.

[11]. Latha, J., Vijayakumar, M., and Shibu, D, S. (2025). A New Length Biased Exponential - Exponential Distribution and its Applications to Cancer Data. *Reliability: Theory and Applications*, 20(2), 834-843.

[12]. Vijayakumar, M., Anwar Shameem, Z, A,. and Bindu Punathumparambath. (2020). Length biased Rani distribution with survival data analysis. *High Technology Letters*, 26(6), 559-573.