International Journal of Mathematics and
Computer Science, **17**(2022), no. 1, 21–31

A Generalization of Length-biased Nakagami Distribution

Ibrahim Abdullahi[1], **Wikanda Phaphan**[2]

[1]Department of Mathematics and Statistics
Faculty of Science
Yobe State University
Damaturu, Nigeria

[2]Department of Applied Statistics
Faculty of Applied Science
King Mongkuts University of Technology North Bangkok
Bankok, Thailand

email: wikanda.p@sci.kmutnb.ac.th

(Received June 1, 2021, Revised June 28, 2021, Accepted July 1, 2021)

Abstract

The aims of this paper is to develop and establish the theoretical properties of a new survival distribution based on a preceding well-known distribution. The proposed distribution is called the length-biased Nakagami distribution. Also, we apply the proposed distribution to a set of actual data to show that the length-biased Nakagami distribution is appropriate for generating a survival model for data.

1 Introduction

The descriptive and analytic survival analyses are important and they are widely applied to various fields such as medical research and actuarial science. For instance, in cancer recurrence after surgery research, a researcher can examine further whether gene fusion was predictive of recurrence by studying

Key words and phrases: Nakagami distribution, length-biased, survival function, moment.
AMS (MOS) Subject Classifications: 62E15, 62E17, 62E10, 62N99.
ISSN 1814-0432, 2022, http://ijmcs.future-in-tech.net

the 5-year recurrence-free survival for patients, an interval estimation, a hazard ratio, and so forth.

In a study of general statistical probability, the function of the starting point which is appealing to study is the probability density function. However, in the case of a study of survival analysis, the functions we are interested in are survival and hazard functions. The properties of the two functions are related to the properties of the cumulative distribution function and probability density function.

Examples of survival distributions, which are applied to the lifetime data for survival analysis are log-normal distribution, gamma distribution, generalized gamma distribution, exponential distribution, Weibull distribution, length biased inverse Gaussian distribution, log-logistic distribution, Gompertz distribution, Birnbaum-Saunders distribution, Nakagami distribution, and so on. In this paper, we are interested in studying the Nakagami distribution.

The weighted distributions were presented by Fisher [1] and Rao [2] which consist of two types of weighted distributions: length biased and size biased distributions. The length biased distribution is defined as follows: Given a non-negative random variable X having the continuous probability density function $f(.)$ and a finite first moment $\mathrm{E}(.)$ exist. We say that a random variable Y is a length-biased random variable of X if Y has a probability density function $f_{LB}(x)$ given by

$$f_{LB}(x) = \frac{xf(x)}{E(X)}.$$

The length-biased distribution is appealing to researchers as it can be applied to reliability analysis, biology, and survival analysis (See, for instance, [3], [4], [5], [6], [7], [8], [10]). Gupta and Tripatni [9] compared the properties of length-biased distribution with the former distributions such as comparing the moment of distribution and the properties of data dispersion for two distributions, and so forth. In this paper, we are interested in studying the Nakagami distribution [11] for the development of the new distribution which is applied to the survival analysis, named “length-biased Nakagami distribution”.

2 Length-biased Nakagami Distribution

The Nakagami distribution (Nak), first appeared in 1960, proposed by Nakagami[11], is the intensity distribution due to rapid fading. The Nakagami distribution has two parameters; $\lambda \geq 0.5$ is the shape parameter and $\beta > 0$ is the scale parameter. The corresponding cumulative distribution function (cdf) is given by

$$F(x) = \frac{1}{\Gamma(\lambda)}\gamma\left(\lambda, \frac{\lambda}{\beta}x^2\right). \tag{2.1}$$

The probability density function (pdf) is given by

$$f(x) = \frac{2\lambda^\lambda}{\Gamma(\lambda)\beta^\lambda}x^{2\lambda-1}\exp\left(\frac{-\lambda}{\beta}x^2\right); \; x > 0, \tag{2.2}$$

and its mean is

$$E(X) = \int_0^\infty \frac{2\lambda^\lambda}{\Gamma(\lambda)\beta^\lambda}x^{2\lambda}\exp\left(\frac{-\lambda}{\beta}x^2\right)dx = \frac{\Gamma(\lambda+1/2)}{\Gamma(\lambda)}\left(\frac{\beta}{\lambda}\right)^{1/2}. \tag{2.3}$$

Theorem 2.1. *Define X as a random variable of the length-biased Nakagami (LBNak) distribution. Then the corresponding probability density function (pdf) of X is given by*

$$f_{LBNak}(x) = \frac{2\lambda^{\lambda+1/2}x^{2\lambda}\mathrm{e}^{-\frac{\lambda}{\beta}x^2}}{\Gamma(\lambda+1/2)\beta^{\lambda+1/2}} \tag{2.4}$$

Proof.
The probability density function of the initial length-biased distribution is given by

$$f_{LB}(x) = \frac{xf(x)}{E(X)}, \; x > 0. \tag{2.5}$$

Substituting equations (2.2) and (2.3) into equation (2.5), we obtain the new probability density function (pdf) of the length-biased Nakagami distribution

$$f_{LBNak}(x) = \frac{2\lambda^{\lambda+1/2}x^{2\lambda}\mathrm{e}^{-\frac{\lambda}{\beta}x^2}}{\Gamma(\lambda+1/2)\beta^{\lambda+1/2}}.$$

Theorem 2.2. *Let X be a random variable of the length-biased Nakagami distribution. The cumulative density function (cdf) of X is given by*

$$F_{LBNak}(x) = \frac{1}{\Gamma(\lambda+1/2)}\gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right). \tag{2.6}$$

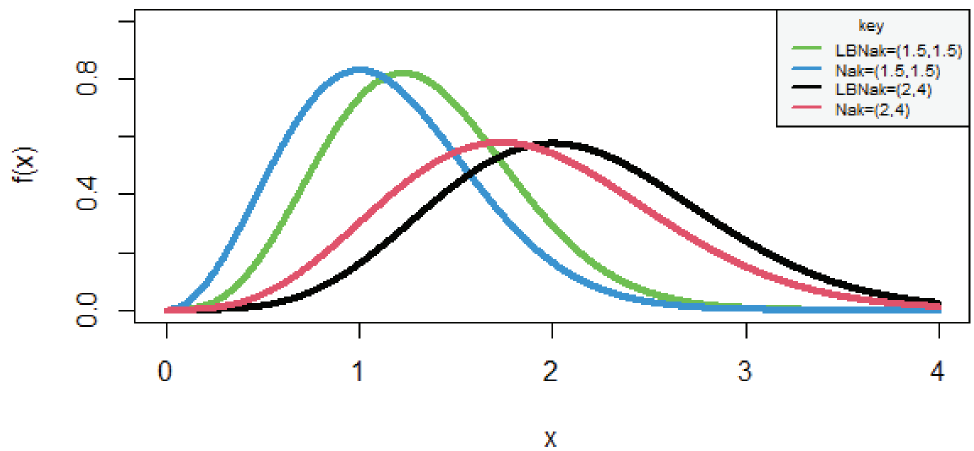

Figure 1: Probability density function of LBNak and Nak distribution

Proof.

$$F_{LBNak}(x) = \int_0^x \frac{2\lambda^{\lambda+1/2} t^{2\lambda} e^{-\frac{\lambda}{\beta}t^2}}{\Gamma(\lambda+1/2)\beta^{\lambda+1/2}} dt. \tag{2.7}$$

$$F_{LBNak}(x) = \frac{1}{\Gamma(\lambda+1/2)} \gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right), \tag{2.8}$$

where the lower incomplete gamma function in equation (2.8) is defined by

$$\gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right) = \int_0^{\frac{\lambda}{\beta}x^2} u^{(\lambda+1/2)-1} e^{-u} du.$$

Theorem 2.3. *Let X be a random variable of the length-biased Nakagami distribution with parameters λ and β. The survival function of X is*

$$S_{LBNak}(x) = 1 - \frac{1}{\Gamma(\lambda+1/2)} \gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right). \tag{2.9}$$

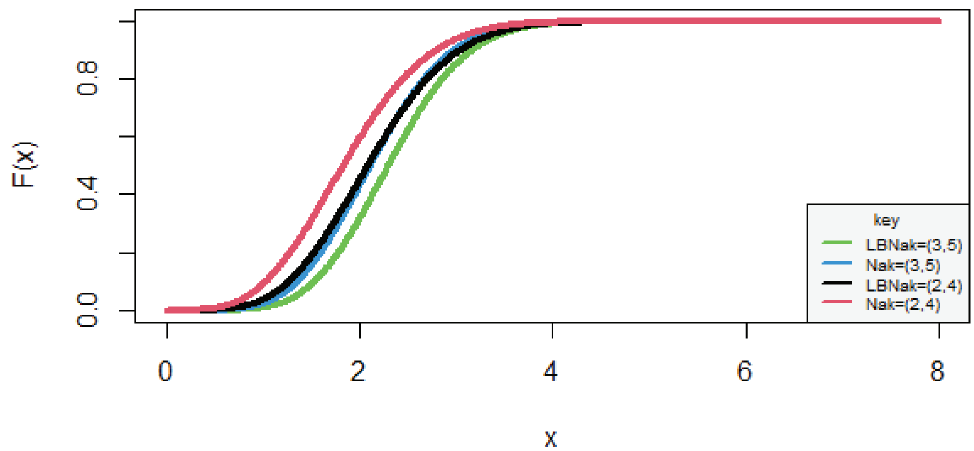

Figure 2: Cumulative density function of LBNak and Nak distribution

Proof.
Let X be a continuous random variable with a cumulative density function $F_{LBNak}(x)$ on the interval $[0, \infty)$. The survival function of X, can be written in the form

$$S_{LBNak}(x) = 1 - F_{LBNak}(x). \tag{2.10}$$

Substituting equation (2.8) into equation (2.10) yields the new survival function of the length-biased Nakagami distribution in equation (2.9).

Theorem 2.4. *Let X be a random variable of the length-biased Nakagami distribution with parameters λ and β. The hazard rate function of X is given by*

$$h_{LBNak}(x) = \frac{2\lambda^{\lambda+1/2} x^{2\lambda} \mathrm{e}^{-\frac{\lambda}{\beta}x^2}}{\beta^{\lambda+1/2}\Gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right)}, \tag{2.11}$$

where the upper incomplete gamma function in equation (2.11) is defined by

$$\Gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right) = \Gamma(\lambda+1/2) - \gamma\left((\lambda+1/2), \frac{\lambda}{\beta}x^2\right).$$

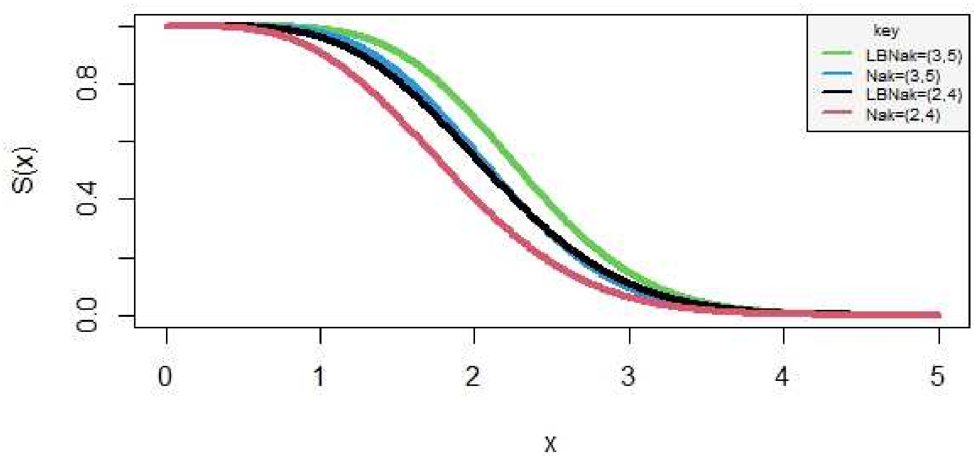

Figure 3: Survival function of LBNak and Nak distribution

Proof.
Let X be an absolutely continuous non-negative random variable with the probability density function $f_{LBNak}(x)$ and the survival function $S_{LBNak}(x)$, then the hazard rate of X can be defined by

$$h_{LBNak}(x) = \frac{f_{LBNak}(x)}{S_{LBNak}(x)} \tag{2.12}$$

Substituting equations (2.4) and (2.9) into equation (2.12) yields the new hazard rate function of LBNak distribution in equation (2.11).

3 Properties of the LBNak Distribution

In this section, we establish certain mathematical features of the length-biased Nakagami distribution.

3.1 Moment

Theorem 3.1. *Let X denote the random variable of the length-biased Nakagami distribution with two parameters λ and β. The r^{th} ordinary moment* *of the length-biased Nakagami distribution is obtained as follows*

$$\mu_r' = \frac{\beta^{r/2}\Gamma\left(\lambda + \frac{(r+1)}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right)\lambda^{r/2}}; \quad r = 1, 2, 3, \cdots, \tag{3.13}$$

where

$$\Gamma\left(\lambda + \frac{r+1}{2}\right) = \int_0^\infty u^{\left[\lambda + \frac{(r+1)}{2}\right]-1} \mathrm{e}^{-u} du.$$

Proof.
The r^{th} ordinary moment of the length-biased Nakagami distribution is defined by

$$\mu_r' = E(X^r) = \int_0^\infty x^r f_{LBNak}(x) dx \tag{3.14}$$

Substituting the pdf equation (2.4) into equation (3.14), we get

$$\mu_r' = E(X^r) = \int_0^\infty \frac{2\lambda^{\lambda+1/2} x^{r+2\lambda} \mathrm{e}^{-\frac{\lambda}{\beta}x^2}}{\Gamma(\lambda+1/2)\beta^{\lambda+1/2}} dx. \tag{3.15}$$

Let

$$u = \frac{\lambda}{\beta}x^2 \Rightarrow x = \left(\frac{u\beta}{\lambda}\right)^{1/2}. \tag{3.16}$$

$$dx = \frac{\beta}{2\lambda}\left(\frac{\lambda}{u\beta}\right)^{1/2} du. \tag{3.17}$$

Substituting equation (3.16) and (3.17) into equation (3.15) yields

$$\mu_r' = \frac{\beta^{r/2}}{\Gamma(\lambda+1/2)\lambda^{r/2}} \int_0^\infty u^{\left[\lambda + \frac{(r+1)}{2}\right]-1} \mathrm{e}^{-u} du \tag{3.18}$$

$$\mu_r' = \frac{\beta^{r/2}\Gamma\left(\lambda + \frac{(r+1)}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right)\lambda^{r/2}}; \quad r = 1, 2, 3, \cdots, \tag{3.19}$$

where

$$\Gamma\left(\lambda+\frac{r+1}{2}\right)=\int_{0}^{\infty}u^{\left[\lambda+\frac{(r+1)}{2}\right]-1}\mathrm{e}^{-u}du.$$

The mean of LBNak distribution is obtained from equation (3.19) by setting $r=1$ as follows:

$$E(X)=\left(\frac{\beta}{\lambda}\right)^{1/2}\frac{\Gamma\left(\lambda+1\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)}. \tag{3.20}$$

The equation (3.21) is obtained by setting $r=2$ into equation (3.19)

$$E(X^2)=\left(\frac{\beta}{\lambda}\right)\frac{\Gamma\left(\lambda+\frac{3}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)}. \tag{3.21}$$

Using

$$Var(X)=E(X^2)-[E(X)]^2, \tag{3.22}$$

the variance of LBNak distribution is obtained by substituting equation (3.20) and (3.21) into equation (3.22)

$$Var(X)=\left(\frac{\beta}{\lambda}\right)\frac{\Gamma\left(\lambda+\frac{3}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)}-\left[\left(\frac{\beta}{\lambda}\right)^{1/2}\frac{\Gamma\left(\lambda+1\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)}\right]^2. \tag{3.23}$$

Proposition 3.2. *The moment generating function of the length-biased Nakagami distribution is given by*

$$M_X(t)=E(e^{tX})=\sum_{r=0}^{\infty}\frac{t^r}{r!}\left[\frac{\beta^{r/2}\Gamma\left(\lambda+\frac{(r+1)}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)\lambda^{r/2}}\right]. \tag{3.24}$$

4 Parameter Estimation

The maximum likelihood estimators of the parameters of the length-biased Nakagami (LBNak) distribution will be estimated in this section. Let $x_1,\cdots,x_n$

be a random sample of size n drawn from the density function in equation (2.4), then the likelihood function of LBNak distribution is given by:

$$\ell(\lambda, \beta) = \prod_{i=1}^{n} \left\{ \frac{2\lambda^{\lambda+1/2} x_i^{2\lambda} \mathrm{e}^{-\frac{\lambda}{\beta} x_i^2}}{\Gamma(\lambda + 1/2)\beta^{\lambda+1/2}} \right\}. \tag{4.25}$$

The log likelihood function in equation (4.26) is obtained by taking the log of equation (4.25).

$$\ln \ell(\vartheta) = n \ln(2) + n \left(\lambda + \frac{1}{2}\right) \ln(\lambda) - n \ln \left[\Gamma\left(\lambda + \frac{1}{2}\right)\right] - n \left(\lambda + \frac{1}{2}\right) \ln(\beta) + 2\lambda \sum_{i=1}^{n} \ln(x_i) - \left(\frac{\lambda}{\beta}\right) \sum_{i=1}^{n} x_i^2. \tag{4.26}$$

The maximum likelihood estimators can be obtained by differentiating equation (4.26) with respect to λ and β and solving for λ and β.

$$\frac{\partial \ln \ell(\vartheta)}{\partial \lambda} = n \ln(\lambda) + \frac{n}{\lambda}\left(\lambda + \frac{1}{2}\right) - \frac{n\Gamma'\left(\lambda + \frac{1}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right)} - n \ln(\beta) + 2 \sum_{i=0}^{n} \ln(x_i) - \frac{\sum_{i=0}^{n} x_i^2}{\beta} = 0, \tag{4.27}$$

and

$$\frac{\partial \ln \ell(\vartheta)}{\partial \beta} = \left(\frac{\lambda}{\beta^2}\right) \sum_{i=0}^{n} x_i^2 - \frac{n}{\beta}\left(\lambda + \frac{1}{2}\right) = 0. \tag{4.28}$$

Because these equations are nonlinear equations, they can not be analytically solved but can be numerically solved through software such as R-language or Mathematica. This paper suggests the “nlminb” function in the R-language for the maximum likelihood estimation via the quasi-Newton method. In the calling sequence for using the “nlminb” function, we find the initial parameters λ and β for a given sample that maximizes the likelihood function, likelihood function to be maximized, and parameter limits.

5 Application to Survival Data Set

In this section, we apply the proposed distribution to a survival data set which was taken from [12]. The data provides information on the number of

Table 1: The MLE of the model parameters for the heart attack data, and AIC measure

Fitting Dist.	Estimate parameters		AIC
	λ	β	
Nak	3.0	0.5	1421.27
LBNak	0.5	0.5	239.8179

months the patient survived after suffering heart attacks. We estimate the parameters of the length-biased Nakagami distribution and Nakagami distributions by maximizing the likelihood function using the "nlminb" function and calculating the AIC Statistics. The results in Table 1 indicate that the AIC statistic of the LBNak distribution is smaller than the Nak distribution. Therefore, the LBNak distribution fits the heart attack data better than the Nak distribution.

6 Conclusion

We developed a new survival distribution, called the length-biased Nakagami (LBNak) distribution, based on the preceding well-known distributions. Moreover, we established theoretical properties such as cumulative density function, survival function, hazard function, r^{th} ordinary moment, and moment generating function. Furthermore, we applied the proposed distribution to a set of actual data to show that the length-biased Nakagami distribution is more appropriate to certain data than the original distribution for generating survival models.

Acknowledgment. The authors would like to express their gratitude to King Mongkuts University of Technology North Bangkok and Yobe State University for supporting this research through a grant.

References

[1] Ronald A. Fisher, The effect of methods of ascertainment upon the estimation of frequencies, *Annals of eugenics*, **6,** no. 1, (1934), 13–25.

[2] C. Radhakrishna Rao, On discrete distributions arising out of methods of ascertainment, *Sankhyā: The Indian Journal of Statistics, Series A*, (1965), 311–324.

[3] Marvin Zelen, Manning Feinleib, On the theory of screening for chronic diseases, *Biometrika*, **56,** no. 3, (1969), 601–614.

[4] Saul Blumenthal, Proportional sampling in life length studies, *Technometrics*, **9,** no. 2, (1967), 205–218.

[5] R. L. Scheaffer, Size-biased sampling, *Technometrics*, **14,** no. 3, (1972), 635–644.

[6] Ganapati P. Patil, Calyampudi R. Rao, Weighted distributions and size-biased sampling with applications to wildlife populations and human families, *Biometrics*, (1978), 179–189.

[7] Richard Simon, Length biased sampling in etiologic studies, *American Journal of Epidemiology*, **111,** no. 4, (1980), 444–452.

[8] Avital Cnaan, Survival models with two phases and length biased sampling, *Communications in Statistics-Theory and Methods*, **14,** no. 4, (1985), 861–886.

[9] Ram C. Tripatni, Ramesh C. Gupta, A comparison between the ordinary and the length-biased modified power series distributions with applications, *Communications in Statistics-Theory and Methods*, **16,** no. 4, (1987), 1195–1206.

[10] Teerawat Simmachan, Kamon Budsaba, Andrei Volodin, On two-sided length biased inverse gaussian distribution, *Chiang Mai Journal of Science*, **45,** no. 7, (2018), 2826–2837.

[11] Minoru Nakagami. The m-distributiona general formula of intensity distribution of rapid fading. In *Statistical methods in radio wave propagation*, Elsevier, 1960, 3–36.

[12] Chanaphun Chananet, Wikanda Phaphan, On the new weight parameter of the mixture Pareto distribution and its application to real data, *Applied Science and Engineering Progress*, **15,** (2022): in press.