EXTENDED INVERSE GAMMA DISTRIBUTION: DEFINITION, PROPERTIES, AND APPLICATIONS

Talha Arslan

Abstract

In this study, the inverse Gamma (IGam) distribution is extended to have an α-monotone density, and the resulting distribution is called α-monotone inverse Gamma distribution (αIGam). The density function of the αIGam distribution is obtained as a scale-mixture between the IGam and Uniform(0,1) distributions. Some properties of the αIGam and its sub-models are expressed. In parameter estimations of the αIGam distribution, the maximum likelihood (ML) method is used, and a small Monte-Carlo simulation study is conducted to show the performances of the ML estimates of the parameters. In the application part of the study, real data sets, which include the survival time from two groups of patients suffering from head and neck cancer disease, are modeled by using the αIGam distribution. The modeling capability of the αIGam distribution is compared with its rivals by using the well-known criteria such as $\log L$, Bayesian information criterion, root mean squares error, and coefficient of determination. Results show that the αIGam distribution is preferable over the IGam distribution and its rivals in modeling data from the patients which were treated only radiotherapy. It can be concluded that αIGam distribution can be considered as an alternative to its rivals in modeling the lifetime data.

Key words: alpha-monotone density, maximum likelihood, scale-mixture extension

JEL Code: C13, C15

Introduction

Statistical distributions are used for modeling the data in various areas of science such as chemistry, medicine, engineering and so on. For example, the gamma, Weibull and inverse Gaussian distributions are commonly used for modeling the lifetime data in reliability studies. In this context, Lin et al. (1989) considered the distribution of random variable $X = Z^{-1}$, where Z follows the gamma distribution, as an alternative to the log-normal and inverse Gaussian distributions in modeling lifetime data.

The distribution of X is called inverse gamma (IGam) distribution and has the probability density function (pdf)

$$f_X(x;\beta,\sigma)=\frac{\sigma^\beta}{\Gamma(\beta)}x^{-\beta-1}e^{-\sigma x^{-1}};\ \ x>0 \tag{1}$$

and cumulative distribution function (cdf)

$$F_X(x;\beta,\sigma)=\Gamma(\sigma x^{-1},\beta). \tag{2}$$

Here, $\beta>0$ and $\sigma>0$ are shape and scale parameters, respectively. Also, $\Gamma(\cdot)$ is the gamma function and $\Gamma(\cdot,\cdot)$ denotes the upper incomplete gamma function defined as

$$\Gamma(w,a)=\frac{1}{\Gamma(a)}\int_w^\infty u^{a-1}e^{-u}du.$$

Hereinafter, $X\sim\text{IGam}(\beta,\sigma)$ is used for representing a random variable X having the pdf given in (1). The r-th moment of the IGam distribution is

$$E[X^r]=\sigma^r\frac{\Gamma(\beta-r)}{\Gamma(\beta)};\ \beta>r. \tag{3}$$

Lin et al. (1989) stated that even though the IGam distribution becomes inverse exponential distribution when $\beta=1$, the IGam distribution may not be superior over the inverse exponential distribution when simplicity taken into account. Recently, Mead (2015) introduced generalized version of the IGam distribution that includes many distributions as a sub-model. In the literature, several methods are proposed to improve distribution's modeling performance or obtain new distribution having better modeling capacity than the exist ones. Lee et al. (2013) provided methods used for generating families of univariate continuous distribution. In literature, there exist many studies including generalized/extended version of the popular statistical distribution; see for example Arslan (2023) and references therein.

Let we have a component and X be a random variable including its lifetime under optimum conditions. It is clear that the lifetime of the component will be reduce under a stress. Therefore, random variable T including its lifetime under stress can be defined as

$$T=X\times Y^{\frac{1}{\alpha}}. \tag{4}$$

Here, X and Y are independent random variables having distributions on $\mathbb{R}^+$ and Uniform(0,1), respectively. Jones (2020) stated that the random variable T has α-monotone density; therefore, such distribution can be called α-monotone distribution.

In this study, the inverse Gamma (IGam) distribution is extended to have an α-monotone density, and the resulting distribution is called α-monotone inverse Gamma distribution (αIGam). Some properties of the αIGam and its sub-models are expressed. The maximum likelihood (ML) method is considered for estimating the parameters of the αIGam distribution, and a small Monte-Carlo simulation study is conducted to show the performances of the ML estimates of the parameters. In the application section, modeling capability of the αIGam distribution is compared with its rivals by using the information criteria values ($\log L$, Bayesian information criteria (BIC)) and goodness-of-fit statistics (root mean squares error (RMSE) and coefficient of determination (R^2)).

1 The α-monotone inverse Gamma distribution

In this section, the pdf and cdf of the αIGam distribution is given. Also, some tractable properties of the αIGam distribution and its sub-models are provided. Then, the ML method is used for estimating the parameters α, β and σ of the αIGam distribution.

1.1 The pdf and cdf of the αIGam distribution

Let X be a random variable having pdf given in (1) and random variable Y follows the Uniform(0,1) distribution. Then, random variable T defined in (4) has the pdf

$$f_T(t;\alpha,\beta,\sigma) = \alpha\sigma^{-\alpha}\frac{\Gamma(\alpha+\beta)}{\Gamma(\beta)}t^{\alpha-1}\Gamma(t^{-1};\alpha+\beta,\sigma);\;\; t>0 \tag{5}$$

and the cdf

$$F_T(t;\alpha,\beta,\sigma) = \sigma^{-\alpha}\frac{\Gamma(\alpha+\beta)}{\Gamma(\beta)}t^{\alpha}\Gamma(t^{-1};\alpha+\beta,\sigma) + \Gamma(\sigma x^{-1},\beta) \tag{6}$$

where $\alpha > 0$ and $\beta > 0$ are the shape parameters, and $\sigma > 0$ is the scale parameters. Here,

$$\Gamma(t^{-1};\alpha+\beta,\sigma) = \frac{\sigma^{\alpha+\beta}}{\Gamma(\alpha+\beta)}\int_0^{t^{-1}} u^{\alpha+\beta-1}e^{-\sigma u}\,du.$$

Hereinafter, $T \sim \alpha\mathrm{IGam}(\alpha,\beta,\sigma)$ is used for representing a random variable T having the pdf given in (5). The plots for the pdf and cdf of the αIGam distribution for certain values of the parameters are shown in Figure 1.

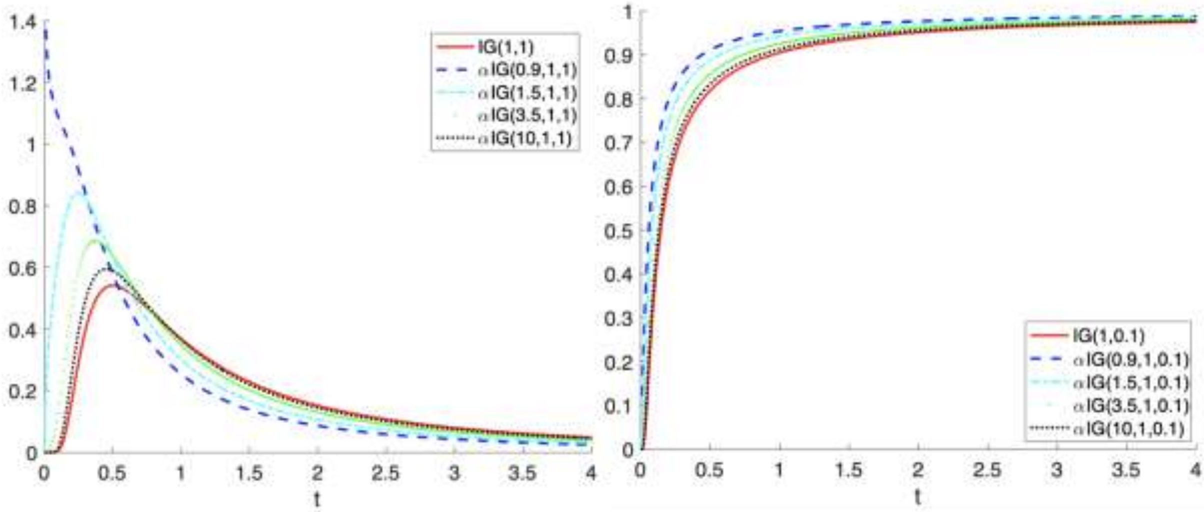
The r-th moment of the αIGam distribution is

$$E[T^r] = E[X^r]\times E\left[Y^{\frac{r}{\alpha}}\right] = \sigma^r\frac{\Gamma(\beta-r)}{\Gamma(\beta)}\frac{\alpha}{\alpha+r}\,. \tag{7}$$

Therefore, random variable T following the αIGam distribution has the square of the coefficient of variation

$$\frac{V[T]}{E[T]^2} = \frac{1}{\alpha(\alpha+2)}\left[\frac{(\alpha+1)^2}{(\beta-2)}+1\right].$$

Fig. 1: Shapes of the pdf and cdf of the αIGam distribution for certain parameter settings



1.2 The properties of the αIGam distribution

The αIGam distribution has some tractable properties given below.

a. Let $T|Y = y \sim \text{IGam}(\beta, \sigma y\hat{}(1/\alpha))$ and $Y \sim \text{Uniform}(0,1)$. Then,

$$\begin{aligned} f_T(t;\alpha,\beta,\sigma) &= \int_0^1 f_X\left(t;\beta,\sigma y^{\frac{1}{\alpha}}\right) f_Y(y)dy \\ &= \alpha\sigma^{-\alpha}\frac{\Gamma(\alpha+\beta)}{\Gamma(\beta)} t^{\alpha-1}\Gamma\left(t^{-1};\alpha+\beta,\sigma\right). \end{aligned} \tag{8}$$

Therefore, the pdf of the αIGam distribution can be expressed as a scale-mixture between the IGam and Uniform(0,1) distributions.

b. The αIGam distribution has an α-monotone density since its pdf satisfies the condition

$$\frac{d}{dt}(\log f_T) \leq \frac{\alpha-1}{t}.$$

c. From the stochastic representation given in (4), the $\alpha\text{IGam}(\alpha,\beta,\sigma)$ converges to the $\text{IGam}(\beta,\sigma)$ when $\alpha \to \infty$; i.e., $\lim_{\alpha\to\infty} f_T(t;\alpha,\beta,\sigma) = f_X(t;\beta,\sigma)$.

The proofs for the corresponding properties can be provided by the author upon request.

1.3 The sub-models of the αIGam distribution

The αIGam distribution includes the distributions given below as a sub-model.

a. If $\beta = 1$, the pdf of the αIGam becomes α-monotone inverse exponential distribution, obtained by Arslan (2021), and has the following pdf

$$f_T(t;\alpha,1,\sigma) = \alpha^2\sigma^{-\alpha}\Gamma(\alpha)t^{\alpha-1}\Gamma(t^{-1};\alpha+1,\sigma);\ \ t>0, \alpha>0, \sigma>0.$$

b. If $\beta = 0.5$ and $\sigma = 0.5c$, the pdf of the αIGam becomes α-monotone Levy distribution and has the following pdf

$$f_T(t;\alpha,0.5,0.5c) = \alpha(0.5c)^{-\alpha}\frac{\Gamma(\alpha+0.5)}{\Gamma(0.5)}t^{\alpha-1}\Gamma(t^{-1};\alpha+0.5,0.5c);\ \ t>0,\alpha>0,c>0.$$

c. If $\sigma = 0.5$, the pdf of the αIGam becomes α-monotone inverse chi-squared distribution and has the following pdf

$$f_T(t;\alpha,\beta,0.5) = \alpha 0.5^{-\alpha}\frac{\Gamma(\alpha+\beta)}{\Gamma(\beta)}t^{\alpha-1}\Gamma(t^{-1};\alpha+\beta,0.5);\ \ t>0,\alpha>0,\beta>0.$$

Note that when $\alpha \to \infty$ α-monotone inverse Exponential, α-monotone Levy, and α-monotone inverse chi-squared distributions converge to the inverse Exponential, Levy, and chi-squared distributions, respectively.

The proofs for the obtaining the pdf of the corresponding sub-models can be provided by the author upon request.

1.4 The ML estimation of the parameters of the αIGam distribution

The ML estimation method, which is based on the maximization of the log-likelihood ($\log L$) function, is used to estimate the parameters of the αIGam distribution. The $\log L$ function of the αIGam distribution is

$$\begin{aligned}\log L\,(\alpha,\beta,\sigma;t) = {} & n\log\alpha - n\alpha\log\sigma + n\log\big(\Gamma(\alpha+\beta)\big) - n\log\big(\Gamma(\beta)\big) \\ & +(\alpha-1)\sum_{i=1}^{n}\log t_i + \sum_{i=1}^{n}\log\Big(\Gamma\big(t_i^{-1};\alpha+\beta,\sigma\big)\Big).\end{aligned} \tag{9}$$

The ML estimates of the parameters α, β, and σ are the points in which the $\log L$ function of the αIGam distribution attains its maximum. Here, optimization tool “fminunc”, which is available in software MATLAB2015b, is used to find the ML estimates of the parameters α, β, and σ, i.e., $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\sigma}$.

Efficiencies of the $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\sigma}$ are shown via a small Monte-Carlo simulation study that can be considered as an example. In the simulations, two different parameter settings are considered for the different sample sizes (n=50, 100, 200) and all simulations are conducted for 1,000 Monte-Carlo runs. For each generated sample, the ML estimates of the parameters were obtained and then the mean, variance and mean squared error (MSE) of the parameter estimates were calculated; see Table 1 for the results. Note that to better understanding the behaviors of the ML estimates of the corresponding parameters, the Monte-Carlo simulation conducted in this study should be extended by considering a wide range of the parameter space.

Tab. 1: The Monte-Carlo simulation results

		$\alpha = 0.5, \beta = 0.3, \sigma = 1.0$			$\alpha = 1.5, \beta = 0.3, \sigma = 1.0$		
n		Mean	Variance	MSE	Mean	Variance	MSE
	$\hat{\alpha}$	0.5933	0.0848	0.0935	1.6036	1.3111	1.3219
50	$\hat{\beta}$	0.3214	0.0069	0.0074	0.3293	0.0050	0.0059
	$\hat{\sigma}$	1.4941	6.1892	6.4333	1.6558	2.0746	2.5047
	$\hat{\alpha}$	0.5511	0.0414	0.0440	1.7528	1.0051	1.0691
100	$\hat{\beta}$	0.3084	0.0024	0.0024	0.3168	0.0023	0.0026
	$\hat{\sigma}$	1.1285	0.5990	0.6155	1.2046	0.4448	0.4867
	$\hat{\alpha}$	0.5184	0.0062	0.0066	1.6404	0.4705	0.4902
200	$\hat{\beta}$	0.3063	0.0012	0.0013	0.3039	0.0011	0.0011
	$\hat{\sigma}$	1.0941	0.2612	0.2701	1.0913	0.1279	0.1362

It can be seen from the Table 1 that the ML estimates for each parameter be closed to the corresponding true parameter value and variances for the parameter estimates decrease when sample size increases. Also, the MSE values for each parameter decrease when sample size increases as expected.

2 Application

In this section two data sets, proposed by Efron (1988), are modeled via the IGam and αIGam distributions. The data sets include the survival time from two groups of patients suffering from head and neck cancer disease. The patients belong to first group were treated only radiotherapy (RT) and the patients in second group were treated combined radiotherapy and chemotherapy (RT+CT); see also Sharma et al. (2015).

Note that Alakus and Erilli (2020), Iranmanesh et al. (2018) and Unal et al. (2018) modeled these data sets by using Weibull-Pareto, inverted gamma and alpha power inverted exponential (APIE) distributions, respectively. The corresponding data are given follows:

Data (RT): 6.53, 7, 10.42, 14.48, 16.10, 22.70, 34, 41.55, 42, 45.28, 49.40, 53.62, 63, 83, 84, 91, 108, 112, 129, 133, 133, 139, 140, 140, 146, 149, 154, 157, 160, 160, 165, 173, 176, 218, 225, 241, 248, 273, 277, 297, 405, 417, 420, 440, 523, 583, 594, 1101, 1146, 1417.

Data (RT+CT): 12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776.

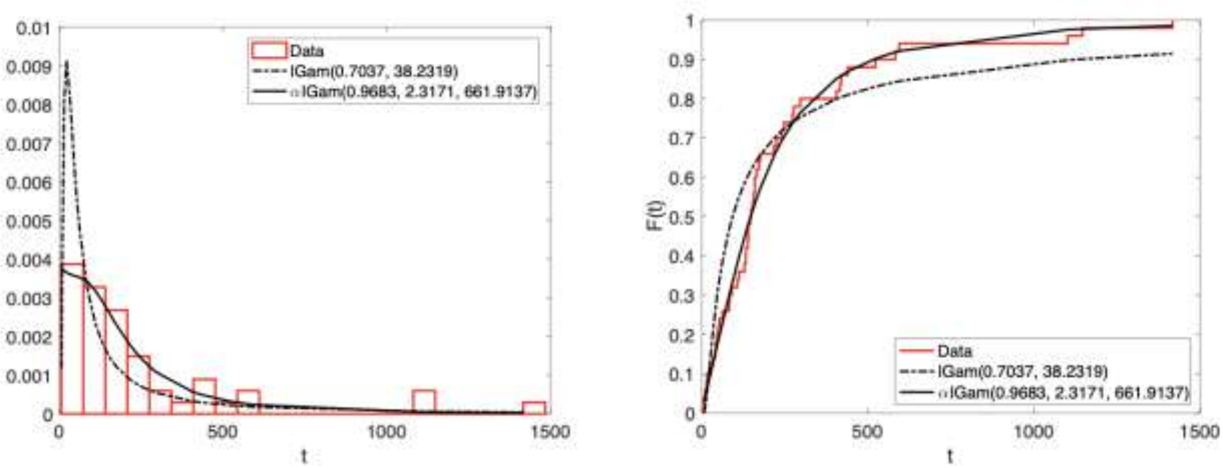
The ML method is used for estimating the parameters of the IGam and αIGam distributions. The parameter estimates, $\log L$, BIC, RMSE and R^2 values of the IGam and αIGam distributions are tabulated in Table 2.

Tab. 2: The modeling results for the corresponding data sets

Data sets	Distributions	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	-log L	BIC	RMSE	R^2
RT	IGam	---	0.7037	38.2319	333.2191	674.2625	0.1033	0.8368
	αIGam	0.9683	2.3171	661.9137	322.3603	656.4566	0.0334	0.9866
RT+CT	IGam	---	1.1234	86.1679	279.3948	566.3580	0.0430	0.9762
	αIGam	71.3642	1.1236	87.4078	279.3925	570.1375	0.0430	0.9762

It is clear from Table 2 that the αIGam distribution has better modeling performance than the IGam distribution when survival times of patients who were treated only RT are considered. Note that the αIGam distribution models the corresponding data better than the APIE distribution proposed by Unal et al. (2018) since it has smaller the BIC and RMSE, and higher the $\log L$ and R^2 values. The fitting performance of the αIGam distribution for RT data is also illustrated graphically by Figure 2.

Fig. 2: The pdf and cdf plots of the αIGam with histogram and empirical cdf of RT data



However, the αIGam and IGam distributions perform more or less the same modeling for the case RT+CT since the αIGam distribution with $\hat{\alpha} = 71.3642$ converges to the IGam distribution. This result is also in agreement with the property of the αIGam distribution given in subsection 1.2 – c. Therefore, a graphical illustration of the fitting performance of the αIGam distribution for RT+CT data is not provided here.

Note that expected value of the αIGam distribution based on the parameter estimates for the RT and RT+CT data sets, i.e., expected lifetime of patients who treated RT and RT+CT, are 247.2404 and 697.2701, respectively. It may conclude that combining the CT with RT changes the distribution of the survival time who suffering from head and neck cancer disease and the effect of this result should be investigated further by the practitioners.

Conclusion

In this study, the αIGam distribution is obtained and some of its properties are shown. The ML is used for estimating the parameters of the αIGam distribution and the Monte-Carlo simulation study is carried out the show the performances of the ML estimates of the parameters.

The αIGam distribution has two shape parameters; therefore, its skewness and kurtosis measures may take values in a broader range than the IGam counters and it makes αIGam preferable over the IGam distribution.

In the application part, two real data sets are modeled by using the IGam and αIGam distributions. The modeling performances of the IGam and αIGam distributions are compared by considering their $\log L$, BIC, RMSE and R^2 values. Modeling result show that the αIGam distribution can be an alternative to the IGam distribution in modeling purpose.

Acknowledgment

This study is supported by “The Scientific and Technological Research Council of Turkey (TUBITAK)” (Grant Number: 1919B022301897) as part of “2224-A Grant Program for Participation in Scientific Meetings Abroad”.

References

Alakus, K., & Erilli, N. (2020). On Weibull-Pareto distribution in censored and uncensored data structures. *Ekoist: Journal of Econometrics and Statistics, 33*, 1-11.

Arslan, T. (2021). An α-monotone generalized log-Moyal distribution with applications to environmental data. *Mathematics, 9*(12), 1400.

Arslan, T. (2023). A new family of unit distributions: Definition, properties and applications. *TWMS J. App. and Eng. Math., 13*(3), 782-791.

Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan– Meier curve. *J. Am. Stat. Assoc., 83*, 414-425.

Iranmanesh, A., Vajargah, K. F., & Hasanzadeh, M. (2018). On the estimation of stress strength reliability parameter of inverted gamma distribution. *Mathematical Sciences, 12*, 71-77.

Jones, M. C. (2020). On univariate slash distributions, continuous and discrete. *Annals of the Institute of Statistical Mathematics, 72*, 645-657.

Lee, C., Famoye, F., & Alzaatreh, A. Y. (2013). Methods for generating families of univariate continuous distributions in the recent decades. *WIREs. Comp. Stats., 5*, 219-238.

Lin, C. T., Duran, B. S., & Lewis, T. O. (1989). Inverted gamma as a life distribution. *Microelectron. Reliab.*, *29*(4), 619-626.

Mead, M. E. (2015). Generalized inverse gamma distribution and its application in reliability. *Comm. Statist. Theory Methods, 44*, 1426-1435.

Sharma, V. K., Singh, S. K., Singh, U., & Agiwal, V. (2015). The inverse Lindley distribution: a stress-strength reliability model with application to head and neck cancer data. *J. Ind. Prod. Eng., 32*(3), 162-173.

Unal, C., Cakmakyapan, S., & Ozel, G. (2018). Alpha power inverted exponential distribution: properties and application. *GU J Sci, 31*(3), 954-965.

Contact

Talha Arslan

Van Yüzüncü Yıl University

Department of Econometrics, Van Yüzüncü Yıl University, 65080 Van, Turkey.

mstalhaarslan@yyu.edu.tr