

Garima distribution and its application to model behavioral science data

Abstract

In this paper a continuous distribution named “Garima distribution” has been suggested for modeling data from behavioral science. The important properties including its shape, moments, skewness, kurtosis, hazard rate function, mean residual life function, stochastic ordering, mean deviations, order statistics, Bonferroni and Lorenz curves, entropy measure, stress-strength reliability have been discussed. The condition under which Garima distribution is over-dispersed, equi-dispersed, and under-dispersed are presented along with other one parameter continuous distributions. The estimation of its parameter has been discussed using maximum likelihood estimation and method of moments. The application of the proposed distribution has been explained using a numerical example from behavioral science and the fit has been compared with other one parameter continuous distributions.

Keywords: lifetime distribution, moments, hazard rate function, mean residual life function, mean deviations, order statistics, estimation of parameter, goodness of fit

Volume 4 Issue 7 - 2016

Rama Shanker

Department of Statistics, Eritrea Institute of Technology, Eritrea

Correspondence: Rama Shanker, Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea, Email shankerrama2009@gmail.com

Received: September 30, 2016 | **Published:** December 09, 2016

Introduction

The modeling and analyzing lifetime data are crucial in many applied sciences including behavioral science, medicine, engineering, insurance and finance, amongst others. There are a number of continuous distributions for modeling lifetime data such as exponential, Lindley, gamma, lognormal, and Weibull and their generalizations. The exponential, Lindley and the Weibull distributions are more popular than the gamma and the lognormal distributions because the survival functions of the gamma and the lognormal distributions cannot be expressed in closed forms and both require numerical integration. Though each of exponential and Lindley distributions has one parameter, the Lindley distribution has one advantage over the exponential distribution that the exponential distribution has constant hazard rate whereas the Lindley distribution has monotonically decreasing hazard rate.

Recently Shanker^{1,2,3,4} has introduced new lifetime distributions, namely Shanker, Akash, Aradhana, and Sujatha distributions for modeling lifetime data from biomedical sciences, engineering and behavioral sciences and showed their superiority over Lindley⁵ and exponential distributions. The probability density function (p.d.f.) and the cumulative distribution function (c.d.f.) of Sujatha, Aradhana, Akash, Shanker, Lindley and exponential distributions are presented in Table 1.

A New lifetime distribution

The probability density function (p.d.f.) of a new lifetime distribution can be introduced as

$$f_7(x; \theta) = \frac{\theta}{\theta+2} (1+\theta x) e^{-\theta x} ; x>0, \theta>0 \quad (2.1)$$

We would call this distribution, “Garima distribution”. This distribution can be easily expressed as a mixture of exponential (θ) and gamma ($2, \theta$) with mixing proportion $\frac{\theta+1}{\theta+2}$. We have

$$f_7(x, \theta) = p g_1(x) + (1-p) g_2(x) \quad (2.2)$$

where $p = \frac{\theta+1}{\theta+2}$, $g_1(x) = \theta e^{-\theta x}$, and $g_2(x) = \theta^2 x e^{-\theta x}$.

The corresponding cumulative distribution function (c.d.f.) of (2.1) is given by

$$F_7(x; \theta) = 1 - \left[1 + \frac{\theta x}{\theta+2} \right] e^{-\theta x} ; x>0, \theta>0 \quad (2.3)$$

The graphs of the p.d.f. and the c.d.f. of Garima distribution for different values of θ are shown in Figure 1.

Moments and related measures

The r^{th} the moment about origin of Garima distribution (2.1) has been obtained as

$$\mu_r' = \frac{r! (\theta+r+2)}{\theta^r (\theta+2)} ; r=1, 2, 3, \dots$$

and so the first four moments about origin as

$$\mu_1' = \frac{\theta+3}{\theta(\theta+2)}, \quad \mu_2' = \frac{2(\theta+4)}{\theta^2(\theta+2)}, \quad \mu_3' = \frac{6(\theta+5)}{\theta^3(\theta+2)}, \quad \mu_4' = \frac{24(\theta+6)}{\theta^4(\theta+2)}$$

Using the relationship between central moments and the moments about origin, the central moments of Garima distribution are obtained as

$$\begin{aligned} \mu_2 &= \frac{\theta^2 + 6\theta + 7}{\theta^2(\theta+2)^2} \\ \mu_3 &= \frac{2(\theta^3 + 9\theta^2 + 21\theta + 15)}{\theta^3(\theta+2)^3} \\ \mu_4 &= \frac{3(3\theta^4 + 36\theta^3 + 134\theta^2 + 204\theta + 111)}{\theta^4(\theta+2)^4} \end{aligned}$$

Thus the coefficient of variation ($C.V.$), coefficient of skewness ($\sqrt{\beta_1}$), coefficient of kurtosis (β_2) and index of dispersion (γ) of Garima distribution are obtained as

$$C.V = \frac{\sigma}{\mu'_1} = \frac{\sqrt{\theta^2 + 6\theta + 7}}{\theta + 3}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2(\theta^3 + 9\theta^2 + 21\theta + 15)}{(\theta^2 + 6\theta + 7)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3(3\theta^4 + 36\theta^3 + 134\theta^2 + 204\theta + 111)}{(\theta^2 + 6\theta + 7)^2}$$

The condition under which Garima distribution is over-dispersed ($\mu < \sigma^2$), equi-dispersed ($\mu = \sigma^2$) and under-dispersed ($\mu > \sigma^2$) are presented in Table 2 along with other lifetime distributions.

Generating functions

The moment generating function ($M_X(t)$), characteristic function ($\varphi_X(t)$), and cumulant generating function ($K_X(t)$) of Garima distribution (2.1) are given by

$$M_X(t) = \left(1 - \frac{(\theta+1)t}{\theta^2+2\theta}\right) \left(1 - \frac{t}{\theta}\right)^{-2}, \quad |t| \leq 1$$

$$\varphi_X(t) = \left(1 - \frac{(\theta+1)it}{\theta^2+2\theta}\right) \left(1 - \frac{it}{\theta}\right)^{-2}, \quad i = \sqrt{-1}$$

$$K_X(t) = \log\left(1 - \frac{(\theta+1)it}{\theta^2+2\theta}\right) - 2\log\left(1 - \frac{it}{\theta}\right)$$

Using the expansion $\log(1-x) = -\sum_{r=0}^{\infty} \frac{x^r}{r}$, we get

$$K_X(t) = -\sum_{r=0}^{\infty} \left(\frac{\theta+1}{\theta^2+2\theta}\right)^r \frac{(it)^r}{r} - 2\sum_{r=0}^{\infty} \frac{(it/\theta)^r}{r}$$

$$= -2\sum_{r=0}^{\infty} \frac{1}{\theta^r} \frac{(it)^r}{r} - \sum_{r=0}^{\infty} \left(\frac{\theta+1}{\theta^2+2\theta}\right)^r \frac{(it)^r}{r}$$

$$= -2\sum_{r=0}^{\infty} \frac{(r-1)!(it)^r}{\theta^r r!} - \sum_{r=0}^{\infty} \left(\frac{\theta+1}{\theta^2+2\theta}\right)^r \frac{(r-1)!(it)^r}{r!}$$

Thus the r^{th} cumulant of Garima distribution is given by

$$K_r = \text{coefficient of } \frac{(it)^r}{r!} \text{ in } K_X(t)$$

$$= \frac{2(r-1)!}{\theta^r} \frac{(r-1)!(\theta+1)^r}{(\theta^2+2\theta)^r}; r=1, 2, 3, \dots$$

This gives

$$\mu'_1 = K_1 = \frac{\theta+3}{\theta(\theta+2)}$$

$$\mu_2 = K_2 = \frac{\theta^2 + 6\theta + 7}{\theta^2(\theta+2)^2}$$

$$\mu_3 = K_3 = \frac{2(\theta^3 + 9\theta^2 + 21\theta + 15)}{\theta^3(\theta+2)^3}$$

$$\mu_4 = K_4 + 3K_2^2 = \frac{3(3\theta^4 + 36\theta^3 + 134\theta^2 + 204\theta + 111)}{\theta^4(\theta+2)^4}$$

Which the same are as obtained earlier.

Hazard rate function and mean residual life function

Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. The hazard rate function (also known as the failure rate function) and the mean residual life function of X are respectively defined as

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f_7(x; \theta)}{1 - F_7(x; \theta)} \quad (5.1)$$

$$m(x) = E[X - x | X > x] = \frac{1}{1 - F_7(x; \theta)} \int_x^{\infty} [1 - F_7(t; \theta)] dt \quad (5.2)$$

The hazard rate function, $h(x)$ and the mean residual life function, $m(x)$ of Garima distribution are given by

$$h(x) = \frac{\theta(1+\theta+x)}{\theta x + \theta + 2} \quad (5.3)$$

$$\text{and} \quad m(x) = \frac{\theta x + \theta + 3}{\theta(\theta x + \theta + 2)} \quad (5.4)$$

It can be easily verified that $h(0) = \frac{\theta(\theta+1)}{\theta+2} = f(0)$ and $m(0) = \frac{\theta+3}{\theta(\theta+2)} = \mu'_1$. It is also obvious from the graphs of $h(x)$ and $m(x)$ that $h(x)$ is an increasing function of x , and θ , whereas $m(x)$ is a decreasing function of x , and θ . The graph of the hazard rate function and mean residual life function of Garima distribution are shown in Figures 2 & 3.

Stochastic orderings

Stochastic ordering of positive continuous random variables is an important tool for judging their comparative behavior. A random variable Y is said to be smaller than a random variable X in the

- stochastic order ($X \leq_{st} Y$) if $F_X(x) \geq F_Y(x)$ for all x
- hazard rate order ($X \leq_{hr} Y$) if $h_X(x) \geq h_Y(x)$ for all x
- mean residual life order ($X \leq_{mrl} Y$) if $m_X(x) \leq m_Y(x)$ for all x
- likelihood ratio order $f_Y(x)$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x .

The following results due to Shaked & Shanthikumar⁶ are well

known for establishing stochastic ordering of distributions

$$\begin{aligned} X \leq_{lr} Y &\Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y \\ &\Downarrow \\ X &\leq_{st} Y \end{aligned} \quad (6.1)$$

The Garima distribution is ordered with respect to the strongest 'likelihood ratio' ordering as shown in the following theorem:

Theorem: Let $X \sim \text{Garima distribution}(\theta_1)$ and $Y \sim \text{Garima distribution}(\theta_2)$. If $\theta_1 > \theta_2$, then $X \leq_{hr} Y$ and hence $X \leq_{lr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Proof: We have

$$\frac{f_X(x)}{f_Y(x)} = \frac{\theta_1(\theta_2+2)}{\theta_2(\theta_1+2)} \left(\frac{1+\theta_1+\theta_1 x}{1+\theta_2+\theta_2 x} \right) e^{-(\theta_1-\theta_2)x} \quad ; x > 0$$

Now

$$\log \frac{f_X(x)}{f_Y(x)} = \log \left[\frac{\theta_1(\theta_2+2)}{\theta_2(\theta_1+2)} \right] + \log \left(\frac{1+\theta_1+\theta_1 x}{1+\theta_2+\theta_2 x} \right) - (\theta_1-\theta_2)x$$

This gives

$$\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} = \frac{\theta_1-\theta_2}{(1+\theta_1+\theta_1 x)(1+\theta_2+\theta_2 x)} - (\theta_1-\theta_2)$$

Thus for $\theta_1 > \theta_2$, $\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} < 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Mean deviations

The amount of scatter in a population is measured to some extent by the totality of deviations usually from mean and median. These are

known as the mean deviation about the mean and the mean deviation about the median defined by $\delta_1(X) = \int_0^\infty |x-\mu|f(x)dx$ and $\delta_2(X) = \int_0^\infty |x-M|f(x)dx$, respectively, where $\mu = E(X)$ and $M = \text{Median}(X)$. The measures $\delta_1(X)$ and $\delta_2(X)$ can be calculated using the relationships

$$\begin{aligned} \delta_1(X) &= \int_0^\mu (\mu-x)f(x)dx + \int_\mu^\infty (x-\mu)f(x)dx \\ &= \mu F(\mu) - \int_0^\mu x f(x)dx - \mu[1-F(\mu)] + \int_\mu^\infty x f(x)dx \\ &= 2\mu F(\mu) - 2\mu + 2 \int_\mu^\infty x f(x)dx \\ &= 2\mu F(\mu) - 2 \int_0^\mu x f(x)dx \end{aligned} \quad (7.1)$$

$$\begin{aligned} \text{and } \delta_2(X) &= \int_0^M (M-x)f(x)dx + \int_M^\infty (x-M)f(x)dx \\ &= M F(M) - \int_0^M x f(x)dx - M[1-F(M)] + \int_M^\infty x f(x)dx \\ &= -\mu + 2 \int_M^\infty x f(x)dx \\ &= \mu - 2 \int_0^M x f(x)dx \end{aligned} \quad (7.2)$$

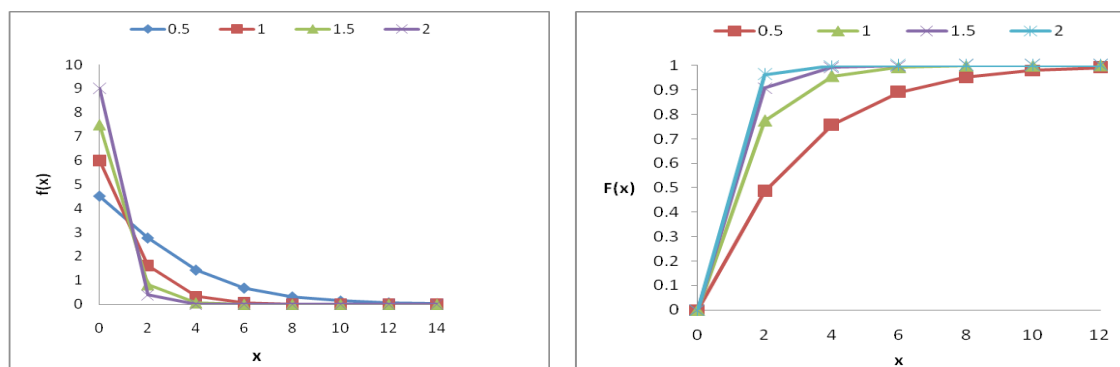


Figure 1 Graphs of the pdf and cdf of Garima distribution for various values of the parameter θ .

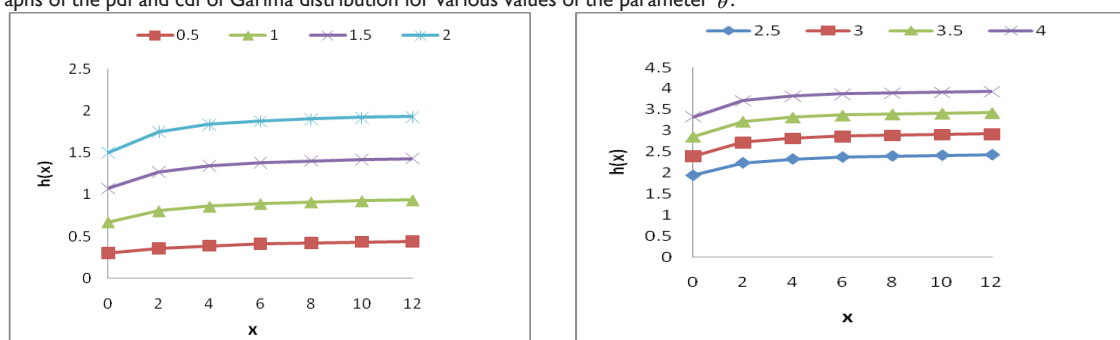


Figure 2 Graph of hazard rate function of Garima distribution for different values of parameter θ .

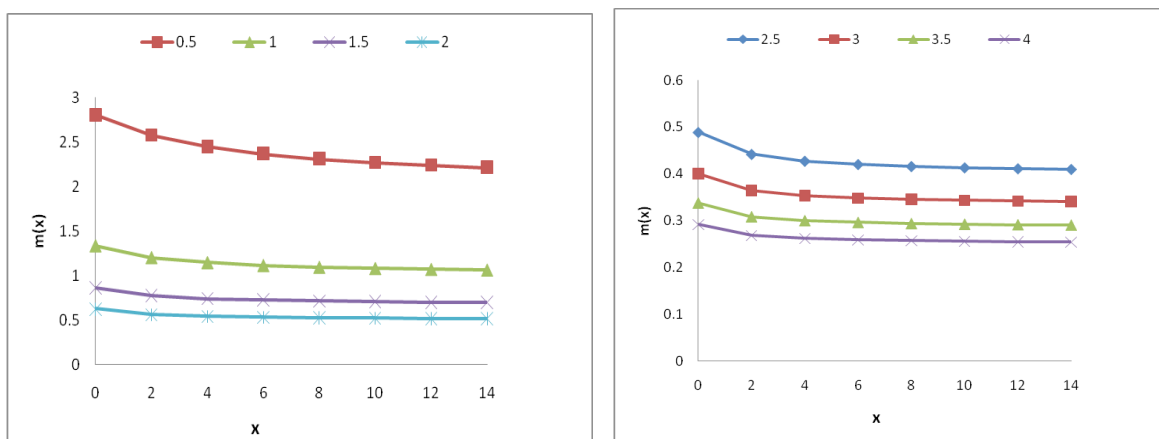


Figure 3 Graph of mean residual life function of Garima distribution for different values of parameter θ .

Table 1 pdf and cdf of Sujatha,⁴ Aradhana,³ Akash,² Shanker,¹ Lindley⁵ and exponential distributions

Distributions	Pdf	Cdf
Sujatha	$f_6(x;\theta) = \frac{\theta^3}{\theta^2 + \theta + 2} (1+x+x^2) e^{-\theta x}$	$F_6(x;\theta) = 1 - \left[1 + \frac{\theta x(\theta x + \theta + 2)}{\theta^2 + \theta + 2} \right] e^{-\theta x}$
Aradhana	$f_5(x;\theta) = \frac{\theta^3}{\theta^2 + 2\theta + 2} (1+x)^2 e^{-\theta x}$	$F_5(x;\theta) = 1 - \left[1 + \frac{\theta x(\theta x + 2\theta + 2)}{\theta^2 + 2\theta + 2} \right] e^{-\theta x}$
Akash	$f_4(x;\theta) = \frac{\theta^3}{\theta^2 + 2} (1+x^2) e^{-\theta x}$	$F_4(x;\theta) = 1 - \left[1 + \frac{\theta x(\theta x + 2)}{\theta^2 + 2} \right] e^{-\theta x}$
Shanker	$f_3(x;\theta) = \frac{\theta^2}{\theta^2 + 1} (\theta + x) e^{-\theta x}$	$F_3(x;\theta) = 1 - \left[1 + \frac{\theta x}{\theta^2 + 1} \right] e^{-\theta x}$
Lindley	$f_2(x;\theta) = \frac{\theta^2}{\theta + 1} (1+x) e^{-\theta x}$	$F_2(x;\theta) = 1 - \left[1 + \frac{\theta x}{\theta + 1} \right] e^{-\theta x}$
Exponential	$f_1(x;\theta) = \theta e^{-\theta x}$	$F_1(x;\theta) = 1 - e^{-\theta x}$

Table 2 Over-dispersion, equi-dispersion and under-dispersion of Garima, Sujatha,⁴ Aradhana,³ Akash,² Shanker,¹ Lindley,⁵ and exponential distributions for varying values of their parameter θ

Lifetime distributions	Over-dispersion ($\mu < \sigma^2$)	Equi-dispersion ($\mu = \sigma^2$)	Under-dispersion ($\mu > \sigma^2$)
Garima	$\theta < 1.164247938$	$\theta = 1.164247938$	$\theta > 1.164247938$
Sujatha	$\theta < 1.364271174$	$\theta = 1.364271174$	$\theta > 1.364271174$
Aradhana	$\theta < 1.283826505$	$\theta = 1.283826505$	$\theta > 1.283826505$
Akash	$\theta < 1.515400063$	$\theta = 1.515400063$	$\theta > 1.515400063$
Shanker	$\theta < 1.171535555$	$\theta = 1.171535555$	$\theta > 1.171535555$
Lindley	$\theta < 1.170086487$	$\theta = 1.170086487$	$\theta > 1.170086487$
Exponential	$\theta < 1$	$\theta = 1$	$\theta > 1$

Using p.d.f. (2.1) and expression for the mean of Garima distribution, we get

$$\int_0^{\mu} x f_7(x; \theta) dx = \mu \frac{\left\{ \theta^2 \mu^2 + (\theta^2 + 3\theta) \mu + (\theta + 3) \right\} e^{-\theta \mu}}{\theta(\theta + 2)} \quad (7.3)$$

$$\int_0^M x f_7(x; \theta) dx = \mu \frac{\left\{ \theta^2 M^2 + (\theta^2 + 3\theta) M + (\theta + 3) \right\} e^{-\theta M}}{\theta(\theta + 2)} \quad (7.4)$$

Using expressions from (7.1), (7.2), (7.3), and (7.4), the mean deviation about mean, $\delta_1(X)$ and the mean deviation about median, $\delta_2(X)$ of Garima distribution are obtained as

$$\delta_1(X) = \frac{2(\theta \mu + \theta + 3)e^{-\theta \mu}}{\theta(\theta + 2)} \quad (7.5)$$

$$\delta_2(X) = \frac{2\left\{ \theta^2 M^2 + (\theta^2 + 3\theta) M + (\theta + 3) \right\} e^{-\theta M}}{\theta(\theta + 2)} - \mu \quad (7.6)$$

Order statistics

Let X_1, X_2, \dots, X_n be a random sample of size n from Garima distribution (2.1). Let $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ denote the corresponding order statistics. The p.d.f. and the c.d.f. of the k^{th} order statistic, say $Y = X_{(k)}$ are given by

$$f_Y(y) = \frac{n!}{(k-1)!(n-k)!} F^{k-1}(y) \{1-F(y)\}^{n-k} f(y)$$

$$= \frac{n!}{(k-1)!(n-k)!} \sum_{l=0}^{n-k} \binom{n-k}{l} (-1)^l F^{k+l-1}(y) f(y)$$

$$\text{and } F_Y(y) = \sum_{j=k}^n \binom{n}{j} F^j(y) \{1-F(y)\}^{n-j}$$

$$= \sum_{j=k}^n \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l F^{j+l}(y),$$

respectively, for $k=1, 2, 3, \dots, n$.

Thus, the p.d.f. and the c.d.f. of k^{th} order statistics of Garima distribution are given by

$$f_Y(y) = \frac{n! \theta (1 + \theta + \theta x) e^{-\theta x}}{(\theta + 2)(k-1)!(n-k)!} \sum_{l=0}^{n-k} \binom{n-k}{l} \times \left[1 - \frac{\theta x + (\theta + 2)}{\theta + 2} e^{-\theta x} \right]^{k+l-1}$$

$$\text{and } F_Y(y) = \sum_{j=k}^n \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l \left[1 - \frac{\theta x + (\theta + 2)}{\theta + 2} e^{-\theta x} \right]^{j+l}$$

Bonferroni and Lorenz curves

The Bonferroni and Lorenz curves⁷ and Bonferroni and Gini indices have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medicine. The Bonferroni and Lorenz curves are defined as

$$B(p) = \frac{1}{p\mu_0} \int_0^q x f(x) dx = \frac{1}{p\mu} \left[\int_0^{\infty} x f(x) dx - \int_q^{\infty} x f(x) dx \right] = \frac{1}{p\mu} \left[\mu - \int_q^{\infty} x f(x) dx \right] \quad (9.1)$$

and

$$L(p) = \frac{1}{\mu_0} \int_0^q x f(x) dx = \frac{1}{\mu} \left[\int_0^{\infty} x f(x) dx - \int_q^{\infty} x f(x) dx \right] = \frac{1}{\mu} \left[\mu - \int_q^{\infty} x f(x) dx \right] \quad (9.2)$$

respectively or equivalently

$$B(p) = \frac{1}{p\mu_0} \int_0^p F^{-1}(x) dx \quad (9.3)$$

$$\text{and } L(p) = \frac{1}{\mu_0} \int_0^p F^{-1}(x) dx \quad (9.4)$$

respectively, where $\mu = E(X)$ and $q = F^{-1}(p)$.

The Bonferroni and Gini indices are thus defined as

$$B = 1 - \int_0^1 B(p) dp \quad (9.5)$$

$$\text{and } G = 1 - 2 \int_0^1 L(p) dp \quad (9.6)$$

respectively.

Using p.d.f. (2.1), we get

$$\int_q^{\infty} x f_7(x; \theta) dx = \frac{\left\{ \theta^2 q^2 + (\theta^2 + 3\theta) q + (\theta + 3) \right\} e^{-\theta q}}{\theta(\theta + 2)} \quad (9.7)$$

Now using equation (9.7) in (9.1) and (9.2), we get

$$B(p) = \frac{1}{p} \left[1 - \frac{\left\{ \theta^2 q^2 + (\theta^2 + 3\theta) q + (\theta + 3) \right\} e^{-\theta q}}{\theta + 3} \right] \quad (9.8)$$

$$\text{and } L(p) = 1 - \frac{\left\{ \theta^2 q^2 + (\theta^2 + 3\theta) q + (\theta + 3) \right\} e^{-\theta q}}{\theta + 3} \quad (9.9)$$

Now using equations (9.8) and (9.9) in (9.5) and (9.6), the Bonferroni and Gini indices of Garima distribution (2.1) are obtained as

$$B = 1 - \frac{\left\{ \theta^2 q^2 + (\theta^2 + 3\theta) q + (\theta + 3) \right\} e^{-\theta q}}{\theta + 3} \quad (9.10)$$

$$G = -1 + \frac{2\left(\theta^2 q^2 + (\theta^2 + 3\theta)q + (\theta + 3)\right)e^{-\theta q}}{\theta + 3} \quad (9.11)$$

Renyi entropy

Entropy of a random variable X is a measure of variation of uncertainty. A popular entropy measure is Renyi entropy.⁸ If X is a continuous random variable having probability density function $f(\cdot)$, then Renyi entropy is defined as

$$T_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \int f^\gamma(x) dx \right\}$$

where $\gamma > 0$ and $\gamma \neq 1$.

Thus, the Renyi entropy for the Garima distribution (2.1) is obtained as

$$\begin{aligned} T_R(\gamma) &= \frac{1}{1-\gamma} \log \left[\int_0^\infty \frac{\theta^\gamma}{(\theta+2)^\gamma} (1+\theta+x)^\gamma e^{-\theta^\gamma x} dx \right] \\ &= \frac{1}{1-\gamma} \log \left[\int_0^\infty \frac{\theta^\gamma (1+\theta)^\gamma}{(\theta+2)^\gamma} \left(1 + \frac{\theta}{\theta+1}x\right)^\gamma e^{-\theta^\gamma x} dx \right] \\ &= \frac{1}{1-\gamma} \log \left[\int_0^\infty \frac{\theta^\gamma (1+\theta)^\gamma}{(\theta+2)^\gamma} \sum_{j=0}^\infty \binom{\gamma}{j} \left(\frac{\theta}{\theta+1}\right)^j e^{-\theta^\gamma x} dx \right] \\ &= \frac{1}{1-\gamma} \log \left[\sum_{j=0}^\infty \binom{\gamma}{j} \frac{\theta^{\gamma+j} (1+\theta)^{\gamma-j}}{(\theta+2)^\gamma} \int_0^\infty e^{-\theta^\gamma x} x^{j+1-1} dx \right] \\ &= \frac{1}{1-\gamma} \log \left[\sum_{j=0}^\infty \binom{\gamma}{j} \frac{\theta^{\gamma+j} (1+\theta)^{\gamma-j} \Gamma(j+1)}{(\theta+2)^\gamma (\theta^\gamma)^{j+1}} \right] \\ &= \frac{1}{1-\gamma} \log \left[\sum_{j=0}^\infty \binom{\gamma}{j} \frac{\theta^{\gamma-1} (1+\theta)^{\gamma-j} \Gamma(j+1)}{(\theta+2)^\gamma (\gamma)^{j+1}} \right] \end{aligned}$$

Stress-strength reliability

The stress-strength reliability describes the life of a component which has random strength X that is subjected to a random stress Y . When the stress applied to it exceeds the strength, the component fails instantly and the component will function satisfactorily till $X > Y$. Therefore, $R = P(Y < X)$ is a measure of component reliability and in statistical literature it is known as stress-strength parameter. It has wide applications in almost all areas of knowledge especially in engineering such as structures, deterioration of rocket motors, static fatigue of ceramic components, aging of concrete pressure vessels etc.

Let X and Y be independent strength and stress random variables having Garima distribution (2.1) with parameter θ_1 and θ_2 respectively. Then the stress-strength reliability R of Garima distribution can be obtained as

$$R = P(Y < X) = \int_0^\infty P(Y < X | X=x) f_X(x) dx$$

$$= \int_0^\infty f_7(x; \theta_1) F_7(x; \theta_2) dx$$

$$= 1 - \frac{\theta_1 \left[(\theta_1 + \theta_2)^2 (\theta_1 \theta_2 + 2\theta_1 + \theta_2 + 2) + (\theta_1 + \theta_2) (2\theta_1 \theta_2 + \theta_2 + 2\theta_1) + 2\theta_1 \theta_2 \right]}{(\theta_1 + 2)(\theta_2 + 2)(\theta_1 + \theta_2)^3}$$

Estimation of parameter

Maximum likelihood estimates (MLE)

Let $(x_1, x_2, x_3, \dots, x_n)$ be a random sample from Garima distribution (2.1). The likelihood function, L of (2.1) is given by

$$L = \left(\frac{\theta}{\theta+2} \right)^n \prod_{i=1}^n (1+\theta+x_i) e^{-n\theta \bar{x}}$$

The natural log likelihood function is thus obtained as

$$\ln L = n \ln \left(\frac{\theta}{\theta+2} \right) + \sum_{i=1}^n \ln(1+\theta+x_i) - n\theta \bar{x}$$

Now

$$\frac{d \ln L}{d\theta} = \frac{2n}{\theta^2 + 2\theta} + \sum_{i=1}^n \frac{1+x_i}{1+\theta+x_i} - n\bar{x} = 0$$

where \bar{x} is the sample mean.

The maximum likelihood estimate, $\hat{\theta}$ of θ is the solution of the equation $\frac{d \ln L}{d\theta} = 0$ and so it can be obtained by solving the following non-linear equation

$$\sum_{i=1}^n \frac{1+x_i}{1+\theta+x_i} + \frac{2n}{\theta^2 + 2\theta} - n\bar{x} = 0 \quad (12.1.1)$$

Method of moment estimates (MOME)

Equating the population mean of the Garima distribution to the corresponding sample mean, the method of moment estimate (MOME) $\tilde{\theta}$, of θ can be obtained as

$$\tilde{\theta} = \frac{(1-2\bar{x}) + \sqrt{4\bar{x}^2 + 8\bar{x} + 1}}{2\bar{x}}, \bar{x} > 0 \quad (12.2.1)$$

A numerical example

In this section the goodness of fit of the Garima distribution has been discussed with an example from behavioral science. The data is related with behavioral science, collected by Balakrishnan N et al.⁹ The scale "General Rating of Affective Symptoms for Preschoolers (GRASP)" measures behavioral and emotional problems of children, which can be classified with depressive condition or not according to this scale. A study conducted by the authors in a city located at the south part of Chile has allowed collecting real data corresponding to the scores of the GRASP scale of children with frequency in parenthesis, which are:

19(6)	20(15)	21(14)	22(9)	23(12)	24(10)
25(6)	26(9)	27(8)	28(5)	29(6)	30(4)
31(3)	32(4)	33	34	35(4)	36(2)
37(2)	39	42	44		

In order to compare distributions, $-2\ln L$, AIC (Akaike Information Criterion), AICC (Akaike Information Criterion Corrected), BIC (Bayesian Information Criterion), for above data set have been computed and presented in Table 3. The formulae for computing AIC, AICC, and BIC are as follows:

$$AIC = -2\ln L + 2k, \quad AICC = AIC + \frac{2k(k+1)}{(n-k-1)}, \quad BIC = -2\ln L + k \ln n$$

where k = number of parameters and n = sample size

The best distribution is the distribution which corresponds to lower values of $-2\ln L$, AIC, AICC, and BIC.

It can be easily seen from above table that the Garima distribution is better than Aradhana, Sujatha, Akash, Shanker, Lindley and exponential distributions for modeling behavioral science data and thus Garima distribution should be preferred over Aradhana, Sujatha, Akash, Shanker, Lindley and exponential distributions for modeling behavioral science data.

Table 3 MLE's, $-2\ln L$, AIC, AICC, and BIC of Garima, Aradhana,³ Sujatha,⁴ Akash,² Shanker,¹ Lindley⁵ and exponential distributions

Model	ML estimate	$-2\ln L$	AIC	AICC	BIC
Garima	0.05317	188.32	190.32	190.35	193.23
Aradhana	0.11557	989.49	991.49	991.52	994.40
Sujatha	0.11745	985.69	987.69	987.72	990.60
Akash	0.11961	981.28	983.28	983.31	986.18
Shanker	0.07974	1033.10	1035.10	1035.13	1037.99
Lindley	0.07725	1041.64	1043.64	1043.68	1046.54
Exponential	0.04006	1130.26	1132.26	1132.29	1135.16

Conclusion

A one parameter lifetime distribution named, "Garima distribution" has been proposed and studied. Its mathematical properties including shape, moments, skewness, kurtosis, hazard rate function, mean residual life function, stochastic ordering, mean deviations, order statistics, Bonferroni and Lorenz curves, Renyi entropy and stress-strength reliability have been discussed. The condition under which Garima distribution is over-dispersed, equi-dispersed, and under-dispersed are presented along with the conditions under which Sujatha, Aradhana, Akash, Shanker, Lindley and exponential distributions are over-dispersed, equi-dispersed and under-dispersed. The method of moments and the method of maximum likelihood estimation have also been discussed for estimating its parameter. Finally, a numerical example from behavioral science has been considered for the goodness of fit of Garima distribution and the fit has been compared with Sujatha, Aradhana, Akash, Shanker, Lindley and exponential distributions. The goodness of fit of the Garima distribution shows that it is an important model for modeling behavioral science data.

NOTE: The paper is named in loving memory of my niece Garima Satyapriya, daughter of my respected brother Prof. Uma Shanker, Department of Mathematics, K.K College of Engineering & Management, Biharsharif, Nalanda, India.

Acknowledgement

None.

Conflict of interest

None.

References

- Shanker R. Shanker Distribution and Its Applications. *International Journal of Statistics and Applications*. 2015;5(6):338–348.
- Shanker R. Akash Distribution and Its Applications. *International Journal of Probability and Statistics*. 2015b;4(3):65–75.
- Shanker R. Aradhana Distribution and Its Applications. *International Journal of Statistics and Applications*. 2016;6(1):23–34.
- Shanker R. *Sujatha Distribution and Its Applications*. To appear in "Statistics in Transition-New series" 2016;17(3).
- Lindley DV. Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society Series B*. 1958;20(1):102–107.
- Shaked M, Shanthikumar JG. *Stochastic Orders and Their Applications*. Academic Press: New York, USA; 1994.
- Bonferroni CE. *Elementi di Statistica generale*. Seeber: Firenze; 1930.
- Renyi A. *On measures of entropy and information*. Berkeley symposium on Mathematical Statistics and Probability, University of California press: Berkeley, USA; 1961. 547–561 p.
- Balakrishnan N, Ramesh C Gupta, Debasis Kundu, et al. *On mixture model based on Birnbaum-Saunders Distributions and associated inference*. A study conducted by Authors regarding the Scores of the GRASP (General Rating of Affective Symptoms for Preschoolers), in a city located at South Part of the Chile; 2010.