Classical and Bayesian Estimations for the Dagum Distribution

Abstract

The Dagum distribution is a great tool for survival analysis as well as representing the distribution of actuarial, meteorological, and income data. Additionally, it is frequently thought to be the best option when compared to the other three parameter distributions. The inverse Burr distribution is a generalised Beta distribution that is produced from generalised beta-II by taking shape parameter one into consideration. The many characteristics and several techniques for estimating the unknown parameters of parameter Dagum distribution are covered in this article. Although we are able to estimate the parameter of the Dagum distribution using Bayesian methods with both informative and noninformative priors, the methods used to produce the estimators are different. Risk functions are used to compare these estimators.

Key Words:

Dagum Distribution, Bayesian Estimation, MLE Method of Moment Estimators, Jeffreys Prior, Loss Functions, Risk Function.

1. Introduction

Camilo Degum (1977) suggested the Dagum distribution as an alternative to the Pareto and log-normal models for modelling personal income data. This distribution has been widely used in a variety of domains, including survival analysis, reliability, meteorological data, and income and wealth statistics. In the actuarial literature, the inverse Burr XII distribution, often referred as the Dagum distribution, is frequently used. The upside-down bathtub, bathtub, and then upside-down bathtub can all be possible geometries for the hazard function of the Dagum distribution. For further information, see Domma (2002). Several authors have studied the model in various domains as a result of this behaviour. In reality, the Dagum distribution has recently been examined from the perspective of reliability and used to assess survival data. Domma et al. (2011) as well as Domma et al. (2013). The Dagum model's history and uses were thoroughly reviewed by Kleiber and Kotz (2003) and Kleiber (2008). With censored samples, Domma et al. (2011) computed the Dagum distribution's parameters. TL-moments were employed by Shahzad and Asghar (2013) to estimate this distribution's parameter. The class of weighted Dagum distributions and associated distributions were introduced by Oluyede and Ye (2014). This study examines the Bayesian and classical analyses of the Dagum distribution for the entire sample. Under various priors and with various loss functions, the maximum likelihood and Bayes estimators are derived.

The cumulative density function (cdf) is defined as follows for every random variable X that follows the Dagum distribution:

$$F(x) = (1 + \lambda x^{-\delta})^{-\beta} \qquad ; x > 0; \beta, \delta, \lambda > 0 \qquad (1)$$

The probability density function is:

$$f(x) = \beta\delta\lambda\, x^{-\delta-1}(1 + \lambda x^{-\delta})^{-\beta} - 1 \qquad ; ; x > 0; \beta, \delta, \lambda > \qquad 0 \; (2)$$

Where λ is the scale parameter and the shape parameters are β, δ.

2. Some Mathematical and Statistical Properties:

2.1. Hazard Function

The hazard rate (HR) function is one of the fundamental tools for studying a system's ageing and reliability characteristics. The HR provides the system failure rate immediately following time t. The following equation represents the hazard rate function of the Dagum distribution:

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{\beta\lambda\delta\, x^{-(\delta+1)}\left(1 + \lambda x^{-\delta}\right)^{-\beta+1}}{1 - (1 + \lambda x^{-\delta})^\wedge(-\beta)} \tag{3}$$

2.2. Mean Residual Lifetime

$$\mu(x) = E[X - x|T > t] = \frac{1}{\overline{F(t)}} \int_0^t \overline{\overline{F}}(u)du\,, \qquad t \geq 0 \tag{4}$$

2.3. The mean residual lifetime function of X is given by:

$$\mu(x) = \frac{\beta\lambda\delta \int_t^\infty y^{(\delta)}\left(1 + \lambda y^{-\delta}\right)^{-\beta+1} dy}{1 - (1 + \lambda t^{-\delta})^{-\beta}} - t \tag{5}$$

3. Methods of Estimation

3.1. Method of Maximum Likelihood

The most popular technique for estimating parameters is maximum likelihood. The method's popularity is definitely due to its many desirable qualities, such as consistency, asymptotic efficiency, invariance, and just its intuitive appeal. Let ($x_1, \ldots, x_n$) be a random sample of size n from the D(β, λ, δ) distribution with parameters "β, λ and δ".

$$L(\beta, \lambda, \delta) = (\beta\lambda\delta)^n \prod_{i=1}^{n} x_i^{-\delta-1}\left(1 + \lambda\, x_i^{-\delta}\right)^{-\beta-1} \tag{6}$$

log-likelihood function is:

$$L(\beta, \lambda, \delta) = nlog\beta + nlog\lambda + nlog\delta - (\delta - 1)\ln \sum \ln x_i - (\beta + 1)\ln \sum \ln\left(1 + \lambda x_i^{-\delta}\right) \tag{7}$$

3.2. Method of Moments

The Method moments of the three-parameter distribution can be obtained by equating the first three theoretical moments of (1) with the sample moments $\frac{1}{n}\sum x_i, \frac{1}{n}\sum_i^n x_i^2$, $\frac{1}{n}\sum_i^n x_i^3$

$$\frac{1}{n}\sum x_i = \beta\lambda^{\left(\frac{1}{\delta}\right)} B - \left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right)$$

$$\frac{1}{n}\sum_i^n x_i^2 = \beta\lambda^{\left(\frac{2}{\delta}\right)} B - \left(1 - \frac{2}{\delta}, \beta + \frac{2}{\delta}\right)$$

$$\frac{1}{n}\sum_i^n x_i^3 = \beta\lambda^{\left(\frac{3}{\delta}\right)} B - \left(1 - \frac{3}{\delta}, \beta + \frac{3}{\delta}\right)$$

3.3. Method of L-Moments

The L-moments estimators are presented in this subsection and can be produced as linear combinations of order statistics. Hosking (1990) first introduced the L-moments estimators, and it has been found that they are more reliable than the traditional moment estimators. Similarly, to how ordinary moment estimators are obtained, the L-moment estimators are similarly obtained by equating the sample L-moments with the population L-moments. Although some higher moments may not exist, L-moment estimation offers a different approach to estimation that is similar to conventional moments. It is also highly resistant to the effects of outliers and has the advantage of existing whenever the distribution's mean exists (Hosking, 1994).

Let the order statistics of a random sample of size n drawn from the Dagum distribution be ($X_{1:n} < \ldots < X_{n:n}$). From Hosking (1990), the first, second, and third sample L-moments are, respectively:

$$l_2 = \frac{2}{n(n-1)} \sum_{i=2}^{n} (i-1) t_{1:n} - t_1$$

$$l_3 = \frac{2}{n(n-1)(n-2)} \sum_{i=3}^{n} (i-1)(i-2) t_{1:n} - \frac{6}{n(n-1)} \sum_{i=2}^{n} (i-1) t_{1:n} + l_1$$

The first, second, and third population Lmoments of (θ=(β,λ,δ)) respectively, are obtained by the quantile function of the Dagum distribution as described in (3).

$$\lambda_1(\theta) = \int_0^1 Q(p|\theta)dp = \beta\, \lambda^{1\backslash\delta}\, B - \left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right)$$

$$\lambda_2(\theta) = \int_0^1 Q(p|\theta)(2p-1)dp = 2\beta\, \lambda^{1\backslash\delta}\, B - \left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right) - \lambda_1(\theta)$$

$$\lambda_2(\theta) = \int_0^1 Q(p|\theta)(6p^2 - 6p + 1)dp = \beta\, \lambda^{1\backslash\delta} \left[6B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right)\right] - 6B - \left(1 - \frac{1}{\delta}, 2\,\beta + \frac{1}{\delta}\right) + B[\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right)]$$

The L-moments estimators $\hat{\beta}_{LME}$, $\hat{\lambda}_{LME}$ and $\hat{\lambda}_{LME}$ of the parameters β, λ and δ can be obtained by solving numerically the equations:

$$\lambda_1\left(\hat{\beta}_{LME}, \hat{\lambda}_{LME}, \hat{\lambda}_{LME}\right) = l_1$$
$$\lambda_2\left(\hat{\beta}_{LME}, \hat{\lambda}_{LME}, \hat{\lambda}_{LME}\right) = l_2$$
$$\lambda_3\left(\hat{\beta}_{LME}, \hat{\lambda}_{LME}, \hat{\lambda}_{LME}\right) = l_3$$

4. Bayesian Estimators using Different Prior and Loss Functions

Types of loss functions: -

If $\hat{\theta}$ represent of estimator for the shape parameter θ , then for:

1. **K-Loss Function (KLF) defined as:(2011)**

$$L(\hat{\theta}, \theta) = \frac{(\hat{\theta} - \theta)^2}{\hat{\theta}\theta} \;\; ; \;\; \hat{\theta}_{KLF} = \sqrt{\frac{E(\theta/x)}{E(p^{-1}/x)}}$$

2. **Entropy loss function (ELF): ELF defined as: (2011)**

$$L(\hat{\theta}, \theta) = \left(\frac{\hat{\theta}}{\theta}\right)^t - t \ln\left(\frac{\hat{\theta}}{\theta}\right)^t - 1; \;\; \hat{\theta}_{elf} = \left(E(\theta^{-t})\right)^{-\frac{1}{t}}$$

4.1. The Posterior distributions with Extension Jeffery's prior

the posterior density function of the shape parameter θ is:

$$P(\theta|\underline{x}) = \frac{L(\theta|\underline{x}).\,p(\theta)}{\int_0^\infty L(\theta|\underline{x}).\,p(\theta)d\theta} \qquad (8)$$

For "Bayesian estimation", we describe two distinct posterior distributions under complete samples, and two distinct prior distributions for the shape parameter.

	Extension Jeffery's priors
$\hat{\theta}_{KLF}$	$\sqrt{\frac{(n-2m+1)(n-2m)}{P}}$

$\hat{\theta}_{ELF}$	$\frac{(n-2m+1)}{P}$

4.2. Bayes Estimators using Loss Functions KLF, ELF

Using KLF, the risk function of $\hat{\theta}_{KLF}$ using K-loss function is given by

$$R\left(\hat{\theta}_{KLF}\right) = E[l(\hat{\theta}_{KLF}, \theta)]$$

$$\theta\left[\frac{\sqrt{\beta_i(\beta_i-1)}}{\theta^2}E\left(\frac{1}{P}\right)-\frac{2}{\theta}+\frac{E(P)}{\sqrt{\beta_i(\beta_i-1)}}\right]$$

$$\left[\frac{\sqrt{\beta_i(\beta_i-1)}}{(n-1)}-2+\frac{n}{\sqrt{\beta_i(\beta_i-1)}}\right]$$

Using ELF, the risk function of $\hat{\theta}_{ELF}$ using entropy loss function is given by

$$R\left(\hat{\theta}_{ELF}\right) = E[l(\hat{\theta}_{ELF}, \theta)]$$

$$\theta\left[\frac{\sqrt{(\beta_i-1)}}{\theta}E\left(\frac{1}{P}\right)-\log(\beta_i-1)-E[\log\theta_P]\right]$$

5. Simulation Study

In this section, we perform numerical calculations to assess the performance of the estimators proposed in previous sections. It is simple to obtain the generation of the Dagum distribution using the inverse transformation, where U is a uniform on (0, 1). We take into account two options for the shape parameter, "β = 1, 2" as well as (n = 20, 60, 80, and 100). We consider the values λ = 1 and δ = 4 in both circumstances. We calculate the bias and the root mean square error (RMSE), respectively, for each estimate as follows:

$$Bias = \frac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i-\theta)$$

$$RMSE(\hat{\theta}) = \frac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i-\theta)^2$$

6. Real Application

For the purpose of comparing the estimators provided in this work, we employed a single data set. The data set, which consists of 100 observations on the breaking stress of carbon fibres, was collected by Nichols and Padgett (Nichols and Padgett, 2006). (in Gba). Data include:

(3.7, 2.74, 2.73, 3.11, 3.27, 2.87, 4.42, 2.41, 3.19, 3.28, 3.09, 1.87, 3.75, 2.43, 2.95, 2.96, 2.3, 2.67, 3.39, 2.81, 4.2, 3.31, 3.31, 2.85, 3.15, 2.35, 2.55, 2.81, 2.77, 2.17, 1.41, 3.68, 2.97, 2.76, 4.91, 3.68, 3.19, 1.57, 0.81, 1.59, 2, 1.22, 2.17, 1.17, 5.08, 3.51, 2.17, 1.69, 1.84, 0.39, 3.68, 1.61, 2.79, 4.7, 1.57, 1.08, 2.03, 1.89, 2.88, 2.82, 2.5, 3.6, 1.47, 3.11, 3.22, 1.69, 3.15, 4.9, 2.97, 3.39, 2.93, 3.22, 3.33, 2.55, 2.56, 3.56, 2.59, 2.38, 2.83, 1.92, 1.36, 0.98, 1.84, 1.59, 5.56, 1.73, 1.12, 1.71, 2.48, 1.18, 1.25, 4.38, 2.48, 0.85, 2.03, 1.8, 1.61, 2.12, 2.05, 3.65).

Table 1. Average bias and RMSE of parameters β = 1, λ = 1, δ = 1.5

n		β		λ		δ	
20	**method**	**bias**	**RMSE**	**bias**	**RMSE**	**bias**	**RMSE**
	MLE	0.0762	0.4291	0.1049	0.6225	0.0355	0.3187
	MME	-0.0165	0.0809	-0.0077	0.0925	0.0037	0.1228
	LME	-0.0017	0.0187	-0.0078	0.0502	-0.0005	0.0187
50	**MLE**	0.0076	0.0937	0.0068	0.1186	0.0030	0.0960

	MME	-0.0050	0.0403	-0.0021	0.0432	0.0003	0.0581
	LME	-0.0004	0.0048	-0.0030	0.0237	-0.0007	0.0058
80	**MLE**	0.0076	0.0937	0.0068	0.1186	0.0030	0.0960
	MME	0.0050	0.0403	-0.0021	0.0432	0.0003	0.0581
	LME	-0.0004	0.0048	-0.0030	0.0237	-0.0007	0.0058
100	**MLE**	0.0021	0.0443	0.0020	0.0601	0.0011	0.0555
	MME	-0.0028	0.0284	-0.0008	0.0308	0.0007	0.0449
	LME	-0.0002	0.0025	-0.0018	0.0184	-0.0004	0.0045

Table 2. Average bias and RMSE of parameters $\beta = 1, \lambda = 1,\ \delta = 2.5$

n		$\boldsymbol{\beta}$		$\boldsymbol{\lambda}$		$\boldsymbol{\delta}$	
20	**methods**	**bias**	**RMSE**	**bias**	**RMSE**	**bias**	**RMSE**
	MLE	0.1475	0.4291	0.1047	0.6226	0.0357	0.3186
	MME	-0.0165	0.0806	-0.0077	0.0924	0.0036	0.1227
	LME	-0.0015	0.0188	-0.0077	0.0501	-0.0005	0.0189
50	**MLE**	0.0076	0.0937	0.0068	0.1186	0.0030	0.0960
	MME	-0.0050	0.0403	-0.0021	0.0432	0.0003	0.0581
	LME	-0.0004	0.0048	-0.0030	0.0237	-0.0007	0.0058
80	**MLE**	0.0076	0.0937	0.0068	0.1186	0.0030	0.0960
	MME	0.0050	0.0403	-0.0021	0.0432	0.0003	0.0581
	LME	-0.0004	0.0048	-0.0030	0.0237	-0.0007	0.0058
100	**MLE**	0.0021	0.0443	0.0020	0.0601	0.0011	0.0555
	MME	-0.0028	0.0284	-0.0008	0.0308	0.0007	0.0449
	LME	-0.0002	0.0025	-0.0018	0.0184	-0.0004	0.0045

Performance of Bayes estimates under the Jeffreys prior extension for the scale (θ) parameter of the Dagum distribution. The procedure is repeated 1000 times, and the resulting average is shown in the tables below. The simulation study makes use of the VGAM package. We select n = 20, 50, and 100 to symbolise various sample sizes. The results for various values of b = 3, 2.5, a = 1.4, and 1.5 are summarised in Table 3. The hyper-parameter values chosen are m = (0.5, 1.3, 3.0). The MLE values in R software are used to select these values. In order to determine if the estimators are admissible under various loss functions, their relative risks are computed.

Table. 3. Posterior estimates $(\hat{\theta})$ and risks R $(\hat{\theta})$ under extended Jeffreys prior

n		**b**	**a**	**M = 0.5**	**MSE**	**M=1.3**	**MSE**	**M = 3.0**	**MSE**
20	$\hat{\boldsymbol{\theta}}_{KLS}$	3.0	1.4	12.690	0.0519	11.648	0.0594	9.434	0.1428
	$\hat{\boldsymbol{\theta}}_{KLS}$	2.5	1.5	6.031	0.0519	5.536	0.0594	4.483	0.1428
	$\hat{\boldsymbol{\theta}}_{ELS}$	3.0	1.4	12.369	0.0260	11.327	0.0298	9.114	0.0683
	$\hat{\boldsymbol{\theta}}_{ELS}$	2.5	1.5	5.878	0.0260	5.383	0.0298	4.331	0.0683
50	$\hat{\boldsymbol{\theta}}_{KLS}$	3.0	1.4	11.629	0.0203	11.253	0.0214	10.454	0.0317
	$\hat{\boldsymbol{\theta}}_{KLS}$	2.5	1.5	7.147	0.0203	6.916	0.0214	6.425	0.0317
	$\hat{\boldsymbol{\theta}}_{ELS}$	3.0	1.4	11.512	0.0101	11.136	0.0107	11.136	0.0157
	$\hat{\boldsymbol{\theta}}_{ELS}$	2.5	1.5	7.075	0.0101	6.844	0.0107	6.353	0.0157
100	$\hat{\boldsymbol{\theta}}_{KLS}$	3.0	1.4	9.848	0.0100	9.689	0.0103	9.353	0.0127
	$\hat{\boldsymbol{\theta}}_{KLS}$	2.5	1.5	6.651	0.0100	6.544	0.0103	6.316	0.0127
	$\hat{\boldsymbol{\theta}}_{ELS}$	3.0	1.4	9.798	0.0050	9.640	0.0051	9.304	0.0063

	$\hat{\boldsymbol{\theta}}_{ELS}$	2.5	1.5	6.617	0.0050	6.510	0.0051	6.283	0.0063

Conclusion: -

We give explicit representations of various distributional features in this article. In addition to using classical and Bayesian inference, estimation techniques are used to estimate the model parameters. We have performed an extensive simulation study to compare these methods. We have compared estimators with respect to bias and mean-squared error. We have also compared estimators using a real data application.

Tables 2 to 3 exhibit the findings of the simulation study and real data example for various values of n, a, b, and hyper-parameters. It has been noted that

1. As the hyper-parameter values rise, so do the risk values of the under and extension of Jeffreys' previous.
2. The risk functions are constant with respect to the parameter, i.e., Given that there is no influence to increase the parameter's true values under the Dagum distribution, the Bayes estimators are minimax estimators for the parameter. The risk functions never change. In other words, raising the parameter's true values has no effect, which suggests that, in the case of the Dagum distribution, the Bayes estimators are minimax estimators for the parameter.
3. Using ELF, there is minimal risk based on both priors, therefore it is acceptable for all sample sizes.
4. The risks are also seen to decrease as sample size increases.

References

- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain” goodness of fit” criteria based on stochastic processes. The annals of mathematical statistics, 193-212.
- Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. Journal of the American statistical association, 49(268), 765-769.
- Casella, G. & Berger, R. L. (1990), Statistical Inference, Wadsworth & Brooks/Cole. Cheng, R. C. H., & Amin, N. A.
- Dagum, C. (1977). A New Model for Personal Income Distribution: Specification and Estimation. Economic Applique, 30(3), 413–437.
- Domma, F., Giordano, S. & Zenga, M. M. (2011). Maximum likelihood estimation in Dagum distribution with censored sample. Journal of Applied Statistics, 38(12), 2971-2985.
- Domma, F., & Condino, F. (2013). The Beta-Dagum distribution: definition and properties.Communications in StatisticsTheory and Methods, 42(22), 4070-4090.
- Kleiber, C., & Kotz, S. (2003). Statistical size distributions in economics and actuarial sciences. 470. John Wiley & Sons.
- Kleiber, C. (2008). A guide to the Dagum distributions. In Modeling Income Distributions and Lorenz Curves, 97-117. Springer New York.
- Nichols, M. D., & Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. Quality and reliability engineering international, 22(2), 141-151. 91
- Oluyede, B. O., & Ye, Y. (2014). Weighted Dagum and related distributions. Afrika Matematika, 25(4), 1125-1141.
- Shahzad, M. N., & Asghar, Z. (2013). Comparing TL-Moments, L-Moments and Conventional Moments of Dagum Distribution by Simulated data. Revista Colombiana de Estadistica, 36(1), 79-93.