Summer 2012

Properties of Weighted Generalized Beta Distribution of the Second Kind

Yuan Ye

PROPERTIES OF WEIGHTED GENERALIZED BETA DISTRIBUTION OF THE SECOND KIND

by

YUAN YE

(Under the Direction of Broderick O. Oluyede)

ABSTRACT

In this thesis, a new class of weighted generalized beta distribution of the second kind (WGB2) is presented. The construction makes use of the 'conservability approach' which includes the size or length-biased distribution as a special case. The class of WGB2 is used as descriptive models for the distribution of income. The results that are presented generalize the generalized beta distribution of second kind (GB2). The properties of these distributions including behavior of pdf, cdf, hazard functions, moments, mean, variance, coefficient of variation (CV), coefficient of skewness (CS), coefficient of kurtosis (CK) are obtained. The moments of other weighted distributions that are related to WGB2 are obtained. Other important properties including entropy (generalized and beta), which are measures of the uncertainty in this class of distributions are derived and studied. Top-sensitive index, bottom-sensitive index, mean logarithmic deviation (MLD) index and Theil index obtained from generalized entropy (GE) are also applied practically. Dagum distribution is a special case of GB2, properties of Dagum and Weighted Dagum distributions including hazard function, reverse hazard function, moments are presented. Fisher information matrix (FIM) and estimates of model parameters under censoring including progressive Type II for the Dagum distribution are presented. WGB2 proved to be in the generalized beta-F family of distributions, and maximum likelihood estimation (MLE) is used to obtain the parameter estimates. WGB2 is applied as descriptive models for the size distribution of income, and fitted to U.S. family income (2001-2009) data with different values of parameters. The empirical results show the length-biased distribution provides the best relative fit.

Key Words: GB2; WGB2; Moments; Generalized entropy; Beta entropy; Dagum distribution; FIM; Size distribution of income; Generalized beta-F family; MLE

2009 *Mathematics Subject Classification*: 62N05, 62B10

PROPERTIES OF WEIGHTED GENERALIZED BETA DISTRIBUTION OF THE SECOND KIND

by

YUAN YE

B.S. in Applied Mathematics

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment

of the Requirement for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA 2012

PROPERTIES OF WEIGHTED GENERALIZED BETA DISTRIBUTION OF THE SECOND KIND

by

YUAN YE

Major Professor: Broderick O. Oluyede

Committee: Charles Champ
Hani Samawi

Electronic Version Approved:
May 2012

ACKNOWLEDGMENTS

It is my pleasure to express my sincere appreciation to my advisor and committee. Foremost, I feel so very fortunate to have Dr. Broderick O. Oluyede as my advisor. I appreciate very much the time and effort he invested in my work, including twice weekly meetings, clear guidance, encouragement, careful reading of my work and quick response.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Generalized Beta Distribution of the Second Kind

The generalized beta distribution of the second kind (GB2) is a very flexible four-parameter distribution. It captures the characteristics of income distribution including skewness, peakness in low-middle range, and long right hand tail. This distribution, therefore provides good description of income distribution (McDonald,1984). The GB2 also includes several other distributions as special or limiting cases, such as generalized gamma (GG), Dagum, beta of the second kind (B2), Singh-Maddala (SM), gamma, Weibull and exponential distributions.

The probability density function (pdf) of the generalized beta distribution of the second kind (GB2) is given by:

$$f_{GB2}(y; a, b, p, q) = \frac{a y^{ap-1}}{b^{ap} B(p,q)[1 + (\frac{y}{b})^a]^{p+q}} \quad \text{for } y > 0, \text{ and 0 otherwise,} \tag{1.1}$$

where a, p, q are shape parameters and b is scale parameter, $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ is the beta function, and a, b, p, q are positive real values.

The $k^{th}-$ order moments of GB2 are given by (McDonald and Xu, 1995):

$$E_{GB2}\left(Y^k\right) = \frac{b^k \Gamma(p + \frac{k}{a}) \Gamma(q - \frac{k}{a})}{\Gamma(p)\Gamma(q)}. \tag{1.2}$$

The moments exist if $-ap < k < aq$.

1.2 Weighted Distribution

Weighted distribution provides an approach to dealing with model specification and data interpretation problems. It adjusts the probabilities of actual occurrence of events to arrive at a specification of the probabilities when those events are recorded. Fisher (1934) first introduced the concept of weighted distribution, in order to study the effect of ascertainment upon estimation of frequencies. Rao (1965, 1985) unified concept of weighted distribution and use it to identify various sampling situations. Cox (1962) and Zelen (1974) introduced weighted distribution to present length biased sampling. Patil (1978) used weighted distribution as stochastic models in the study of harvesting and predation. The usefulness and applications of weighted distribution to biased samples in various areas including medicine, ecology, reliability, and branching processes can also be seen in Nanda and Jain (1999), Gupta and Keating (1985), Oluyede (1999), Patil (1991), Zelen and Feinleib (1969) are in references therein.

Suppose Y is a non-negative random variable with its natural pdf $f(y;\theta)$, θ is a parameter, then the pdf of the weighted random variable $Y^w$is given by:

$$f^w(y;\theta,\beta) = \frac{w(y,\beta)f(y;\theta)}{\omega}, \tag{1.3}$$

where the weight function $w(y,\beta)$ is a non-negative function, that may depend on the parameter β, and $0 < \omega = E(w(Y,\beta)) < \infty$ is a normalizing constant.

In general, consider the weight function $w(y)$ defined as follows:

$$w(y) = y^k e^{ly} F^i(y) \overline{F}^j(y). \tag{1.4}$$

Setting $k = 0$; $k = j = i = 0$; $l = i = j = 0$; $k = l = 0$; $i \to i-1$; $j = n - i$; $k = l = i = 0$ and $k = l = j = 0$ in this weight function, one at a time, implies probability weighted moments, moment-generating functions, moments, order statistics, proportional hazards and proportional reversed hazards, respectively, where $F(y) = P(Y \leq y)$ and $\overline{F}(y) = 1 - F(y)$. If $w(y) = y$, then $Y^* = Y^w$ is called the size-biased version of Y.

1.3 Outline of Results

The outline of this thesis is as follows: In chapter 2, the weighted generalized beta distribution of the second kind (WGB2) is introduced. Some properties including the cdf, hazard functions, monotonicity, income-share elasticity, and moments (mean, variance, coefficient of skewness and coefficient of kurtosis) are presented. Chapter 3 contains Renyi entropy and generalized entropy of generalize beta distribution and the weighted version. Top-sensitive index, bottom-sensitive index, mean logarithmic deviation (MLD) index and Theil index obtained from generalized entropy (GE) are also presented. Chapter 4 contains properties of the weighted Dagum distribution, Fisher information matrix, and the estimates of the parameters under progressive type II censoring. Chapter 5 contains estimation of parameters in the weighted beta distribution of the second kind with application to U.S.

family income data (2001-2009).

CHAPTER 2

WEIGHTED GENERALIZED BETA DISTRIBUTION OF THE SECOND KIND

In particular, if we set $l = i = j = 0$ in the weight function (1.4),then we have $w(y) = y^k$. With the moments of GB2 in equation (1.2) we can obtain the corresponding pdf of weighted generalized beta distribution of the second kind (WGB2):

$$\begin{aligned} g_{WGB2}(y; a, b, p, q, k) &= \frac{y^k f(y; a, b, p, q)}{E(Y^k)} \\ &= \frac{y^k a y^{ap-1} \Gamma(p)\Gamma(q)}{b^{ap} B(p, q)[1 + (\frac{y}{b})^a]^{p+q} \cdot b^k \Gamma(p + \frac{k}{a})\Gamma(q - \frac{k}{a})} \\ &= \frac{a y^{ap+k-1}}{b^{ap+k} B(p + \frac{k}{a}, q - \frac{k}{a})[1 + (\frac{y}{b})^a]^{p+q}}, \end{aligned} \tag{2.1}$$

where $y > 0$, $a, b, p, q > 0$ and $-ap < k < aq$. WGB2 has one more parameter k compared to GB2.

2.1 Some Properties

The graphs of the pdf are given below:

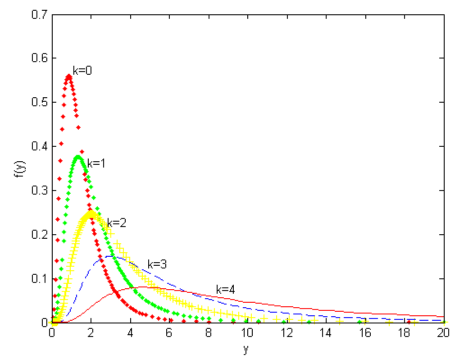

Figure 2.1: pdf of WGB2 (a=1, b=2, p=4, q=6)

Fig.2.1 depicts the pdf of WGB2 as the parameter k changes for representative values of the parameters a, b, p, q : $a = 1, b = 2, p = 4, q = 6$ for $k = 0, 1, 2, 3, 4$. We observe that: as the value of k increases, the "height" of the pdf becomes lower, and the pdf is more skewed right. The parameter k controls the shape and skewness of the density.

In order to further understand WGB2 with weight function $w(y) = y^k$, we discuss some related properties, including the cumulative distribution function (cdf), hazard function, monotonicity properties and elasticity.

The cdf of WGB2 is given by:

$$\begin{aligned} G_{WGB2}(y; a, b, p, q, k) &= \int_0^y \frac{a y^{ap+k-1}}{b^{ap+k} B(p+\frac{k}{a}, q-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+q}} dy \\ &= 1 - I_{[1+(\frac{y}{b})^a]^{-1}}\left(p+\frac{k}{a}, q-\frac{k}{a}\right), \end{aligned} \tag{2.2}$$

where $I_x(\alpha, \beta) = \frac{B_x(\alpha,\beta)}{B(\alpha,\beta)}$ is the incomplete beta function, $y > 0$, $a, b, p, q > 0$ and $-ap < k < aq$.

The graphs of the cdf of WGB2 are given below:

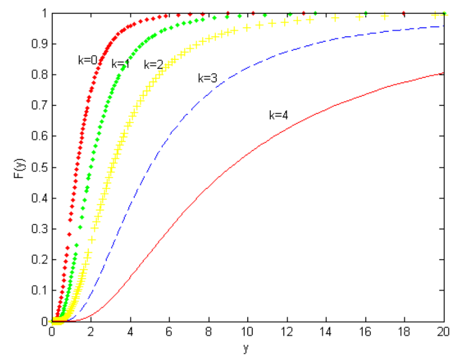

Figure 2.2: cdf of WGB2 (a=1, b=2, p=4, q=6)

Fig.2.2 depicts the cdf of WGB2 as the parameter k changes for representative values of the parameters a, b, p, q : $a = 1, b = 2, p = 4, q = 6$ for $k = 0, 1, 2, 3, 4$. We observe that as the value of k increases, the cdf increases slowly.

In Tables 2.1 and 2.2, some percentiles of WGB2 are presented. In particular, the 50^{th}, 75^{th}, 90^{th} and 95^{th} percentiles of WGB2 are given.

Table 2.1: Percentiles of WGB2 with k=1

a	b	p	q	50th	75th	90th	95th
2.5	2.5	2.5	2.5	2.8986	3.7767	4.8844	5.7719
3				2.7696	3.4467	4.2578	4.8812
3.5				2.6949	3.2474	3.8859	4.3624
4				2.6477	3.1152	3.6411	4.0254
2.5	2.5	2.5	2.5	2.8986	3.7767	4.8844	5.7719
	3			3.4784	4.532	5.8613	6.9262
	3.5			4.0581	5.2873	6.8380	8.0806
	4			4.6378	6.0427	7.8150	9.2348
2.5	2.5	2.5	2.5	2.8986	3.7767	4.8844	5.7719
		3		3.1113	4.0262	5.1868	6.1196
		3.5		3.3043	4.2542	5.4644	6.439
		4		3.4818	4.4649	5.7217	6.7363
2.5	2.5	2.5	2.5	2.8986	3.7767	4.8844	5.7719
			3	2.6265	3.3533	4.2249	4.8932
			3.5	2.4265	3.0555	3.7831	4.3233
			4	2.2713	2.8314	3.4619	3.9189

Table 2.2: Percentiles of WGB2 with k=2

a	b	p	q	50th	75th	90th	95th
2.5	2.5	2.5	2.5	3.3966	4.5133	6.0133	7.281
3				3.0834	3.8815	4.8820	5.6804
3.5				2.9126	3.5359	4.2819	4.855
4				2.8084	3.3212	3.9151	4.3592
2.5	2.5	2.5	2.5	3.3966	4.5133	6.0133	7.281
	3			4.0759	5.416	7.2160	8.7372
	3.5			4.7552	6.3186	8.4187	10.1935
	4			5.4346	7.2213	9.6214	11.6495
2.5	2.5	2.5	2.5	3.3966	4.5133	6.0133	7.281
		3		3.6141	4.7791	6.3507	7.6818
		3.5		3.8136	5.0242	6.6628	8.053
		4		3.9987	5.2524	6.9541	8.3998
2.5	2.5	2.5	2.5	3.3966	4.5133	6.0133	7.281
			3	3.0045	3.8736	4.9623	5.8277
			3.5	2.7349	3.4591	4.3246	4.9849
			4	2.5343	3.1631	3.8892	4.4269

The percentiles increases as b, p increases, and decreases as a, q increases with fixed k.

The hazard function of WGB2 is given by:

$$\begin{aligned} h_{WGB2}(y;a,b,p,q,k) &= \frac{g_{WGB2}(y;a,b,p,q,k)}{\overline{G}_{WGB2}(y;a,b,p,q,k)} \\ &= \frac{g_{WGB2}(y;a,b,p,q,k)}{1-G_{WGB2}(y;a,b,p,q,k)} \\ &= \frac{ay^{ap+k-1}[1+(\frac{y}{b})^a]^{-(p+q)}}{b^{ap+k}B(p+\frac{k}{a},q-\frac{k}{a})I_{[1+(\frac{y}{b})^a]^{-1}}(p+\frac{k}{a},q-\frac{k}{a})} \qquad (2.3) \end{aligned}$$

for $y > 0$, $a, b, p, q > 0$ and $-ap < k < aq$.

The graphs of the hazard functions are given below:

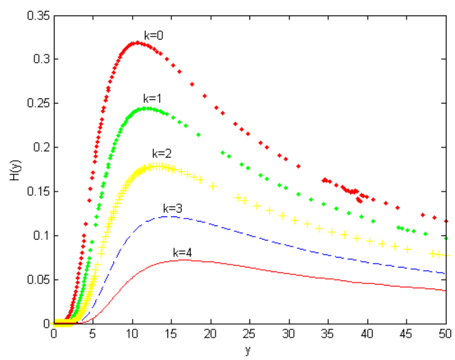

Figure 2.3: Hazard functions of WGB2

Next, we study the monotonicity properties and discuss the income-shared elasticity of WGB2. In order to discuss monotonicity of WGB2, we take the logarithm of its pdf:

$$ln(g_{WGB2}(y;a,b,p,q,k)) = lnC+(ap+k-1)(lny-lnb)-(p+q)ln\left[1+\left(\frac{y}{b}\right)^a\right],$$

where C is a constant. Note that

$$\frac{\partial lng_{_{WGB2}}(y;a,b,p,q,k)}{\partial y}=\frac{(ap+k-1)b^a-(aq-k+1)y^a}{y(b^a+y^a)},$$

where $y>0$, $b,p,q>0$, and$-ap<k<aq$, so $aq-k+1>0$.

It follows therefore that

$$\frac{\partial lng_{_{WGB2}}(y;a,b,p,q,k)}{\partial y}>0\Leftrightarrow y<b\left(\frac{ap+k-1}{aq-k+1}\right)^{\frac{1}{a}},$$

$$\frac{\partial lng_{_{WGB2}}(y;a,b,p,q,k)}{\partial y}=0\Leftrightarrow y=b\left(\frac{ap+k-1}{aq-k+1}\right)^{\frac{1}{a}},$$

$$\frac{\partial lng_{_{WGB2}}(y;a,b,p,q,k)}{\partial y}<0\Leftrightarrow y>b\left(\frac{ap+k-1}{aq-k+1}\right)^{\frac{1}{a}}.$$

The mode of WGB2 is $y_0=b\left(\frac{ap+k-1}{aq-k+1}\right)^{\frac{1}{a}}$.

The income-share elasticity is defined as $\frac{-yg^{'}(y)}{g(y)}$, where $g^{'}(y)=\frac{dg(y)}{dy}$. See Esteban (1986) for additional details. The derivative of $g_{_{WGB2}}(y;a,b,p,q,k)$ with respect to y is given by:

$$\begin{aligned}
g^{'}_{_{WGB2}}(y) &= \left[\frac{ay^{ap+k-1}}{b^{ap+k}B(p+\frac{k}{a},q-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+q}}\right]^{'}\\
&= \frac{a}{bB(p+\frac{k}{a},q-\frac{k}{a})}\left[\left(\frac{y}{b}\right)^{ap+k-1}\left[1+\left(\frac{y}{b}\right)^{a}\right]^{-(p+q)}\right]^{'}\\
&= \frac{ay^{ap+k-2}[ap+k-1-(p+q)\frac{y^a}{b^a+y^a}]}{b^{ap+k}B(p+\frac{k}{a},q-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+q}}.
\end{aligned}$$

The income-share elasticity of WGB2 is given by:

$$\begin{aligned}\delta_{G_{WGB2}}(y;a,b,p,q,k) &= \frac{-yg'_{WGB2}(y;a,b,p,q,k)}{g_{WGB2}(y;a,b,p,q,k)}\\ &= (p+q)\frac{y^a}{b^a+y^a} - ap - k + 1. \qquad (2.4)\end{aligned}$$

2.2 Moments of WGB2 and Related Special Cases

2.2.1 Moments

The non central (j^{th}) moment of WGB2 is given by:

$$\begin{aligned}E_{G_{WGB2}}(Y^j) &= \int_0^\infty y^j g_{WGB2}(y)dy\\ &= \int_0^\infty \frac{y^j a y^{ap+k-1}}{b^{ap+k}B(p+\frac{k}{a}, q-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+q}}dy\\ &= \frac{ab^{j-1}}{B(p+\frac{k}{a}, q-\frac{k}{a})}\int_0^\infty \left(\frac{y}{b}\right)^{ap+k+j-1}\left[1+\left(\frac{y}{b}\right)^a\right]^{-(p+q)}dy\\ &= \frac{ab^{j-1}}{B(p+\frac{k}{a}, q-\frac{k}{a})}\int_0^\infty \left[\left(\frac{y}{b}\right)^a\right]^{p+\frac{a}{k}+\frac{j}{a}-\frac{1}{a}}\left[1+\left(\frac{y}{b}\right)^a\right]^{-(p+q)}dy.\end{aligned}$$

Let $(\frac{y}{b})^a = t$, then $y = bt^{\frac{1}{a}}, dy = \frac{b}{a}t^{\frac{1}{a}-1}dt$, and

$$\begin{aligned}E_{G_{WGB2}}(Y^j) &= \frac{b^j}{B(p+\frac{k}{a}, q-\frac{k}{a})}\int_0^\infty t^{p+\frac{a}{k}+\frac{j}{a}-1}(1+t)^{-(p+q)}dt\\ &= \frac{b^j B(p+\frac{k}{a}+\frac{j}{a}, q-\frac{k}{a}-\frac{j}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})}. \qquad (2.5)\end{aligned}$$

The mean and variance of the WGB2 distribution are given by:

$$\mu_{G_{WGB2}} = E_{G_{WGB2}}(Y) = \frac{bB(p+\frac{k}{a}+\frac{1}{a}, q-\frac{k}{a}-\frac{1}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})}, \qquad (2.6)$$

and

$$\begin{aligned} Var_{G_{WGB2}}(Y) &= E_{G_{WGB2}}(Y^2) - (E_{G_{WGB2}}(Y))^2 \\ &= b^2 \left[\frac{B(p+\frac{k+2}{a}, q-\frac{k+2}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})} - \left(\frac{B(p+\frac{k+1}{a}, q-\frac{k+1}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})} \right)^2 \right] \end{aligned} \tag{2.7}$$

respectively.

The coefficient of variation (CV) is given by:

$$\begin{aligned} CV &= \frac{\sqrt{Var_{G_{WGB2}}(Y)}}{\mu_{G_{WGB2}}} \\ &= \sqrt{\frac{B(p+\frac{k+2}{a}, q-\frac{k+2}{a})B(p+\frac{k}{a}, q-\frac{k}{a})}{B^2(p+\frac{k+1}{a}, q-\frac{k+1}{a})} - 1}. \end{aligned} \tag{2.8}$$

Similarly, the coefficient of skewness (CS) and coefficient of kurtosis (CK) are given by:

$$\begin{aligned} CS &= E\left[\left(\frac{Y-\mu}{\sigma} \right)^3 \right] \\ &= \frac{E[Y^3] - 3\mu E[Y^2] + 2\mu^3}{\sigma^3}, \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} CK &= E\left[\left(\frac{Y-\mu}{\sigma} \right)^4 \right] \\ &= \frac{E[Y^3] - 4\mu E[Y^3] + 6\mu^2 E[Y^2] - 3\mu^4}{\sigma^4}, \end{aligned} \tag{2.10}$$

where

$$\mu = \mu_{G_{WGB2}}, \quad \sigma = \sqrt{Var_{G_{WGB2}}(Y)}, \quad E[Y^2] = \frac{b^2 B(p+\frac{k}{a}+\frac{2}{a}, q-\frac{k}{a}-\frac{2}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})},$$

$$E[Y^3] = \frac{b^3 B(p+\frac{k}{a}+\frac{3}{a}, q-\frac{k}{a}-\frac{3}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})}, \quad E[Y^4] = \frac{b^4 B(p+\frac{k}{a}+\frac{4}{a}, q-\frac{k}{a}-\frac{4}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})}.$$

Since we have obtained the mode, mean, variance, CV, CS and CK of WGB2, we set the values of the parameters a, b, p, q and compute the values of these quantities in Tables 2.3 and 2.4. From the tables we observe the following:

1) When k = 1, mode increases as p increases, decreases as q increases, and does not change as a, b increases; when k = 2, mode increases as b, p increases, decreases as a, q increases.

2) Mean, variance decreases as a,q increases, increases as b, p increases.

3) CV decreases as a, p, q increases, and does not change as b increases.

4) CS decreases as a, b, p, q increases.

5) CK decreases as a, q increases, increases as p increases, and does not change as b increases.

Table 2.3: The mode, mean, variance, CV, CS and CK of WGB2 when k=1

a	b	p	q	mode	mean	variance	CV	CS	CK
2.5	2.5	2.5	2.5	2.5	3.187619	2.072506	0.451629	-13.398078	22.993346
3				2.5	2.954545	1.132674	0.360215	-21.542438	11.61182
3.5				2.5	2.824413	0.723843	0.301227	-32.416597	8.227137
4				2.5	2.743821	0.50716	0.259547	-46.503715	6.672292
2.5	2.5	2.5	2.5	2.5	3.187619	2.072506	0.451629	-13.398078	22.993346
	3			2.5	3.825143	2.984408	0.451629	-14.081444	22.99334
	3.5			2.5	4.462667	4.062111	0.451629	-14.569562	22.993346
	4			2.5	5.100191	5.305615	0.451629	-14.935651	22.993346
2.5	2.5	2.5	2.5	2.5	3.187619	2.072506	0.451629	-13.398078	22.993346
		3		2.6891	3.417945	2.263706	0.440195	-14.444185	23.727663
		3.5		2.8602	3.627291	2.45092	0.431601	-15.321162	24.322348
		4		3.0171	3.820056	2.634649	0.424905	-16.067381	24.813225
2.5	2.5	2.5	2.5	2.5	3.187619	2.072506	0.451629	-13.398078	22.993346
			3	2.3242	2.829373	1.287897	0.401098	-16.618157	11.383826
			3.5	2.1852	2.580454	0.91433	0.370557	-19.08715	7.988561
			4	2.0715	2.394085	0.702145	0.350005	-20.958808	6.438066

Table 2.4: The mode, mean, variance, CV, CS and CK of WGB2 when k=2

a	b	p	q	mode	mean	variance	CV	CS	CK
2.5	2.5	2.5	2.5	2.8445	3.8378	3.9325	0.5167	-10.6613	137.0961
3				2.7339	3.3379	1.7017	0.3908	-18.6434	20.2011
3.5				2.6695	3.0807	0.9626	0.3185	-29.3114	11.2619
4				2.6286	2.9287	0.6268	0.2703	-43.1653	8.2368
2.5	2.5	2.5	2.5	2.8445	3.8378	3.9325	0.5167	-10.6613	137.0961
	3			3.4134	4.6054	5.6628	0.5167	-11.0601	137.0961
	3.5			3.9824	5.3729	7.7077	0.5167	-11.345	137.0961
	4			4.5513	6.1405	10.0672	0.5167	-11.5587	137.0961
2.5	2.5	2.5	2.5	2.8445	3.8378	3.9325	0.5167	-10.6613	137.0961
		3		3.0314	4.0802	4.305	0.5085	-11.1588	140.504
		3.5		3.2024	4.303	4.6694	0.5022	-11.5722	143.2584
		4		3.3607	4.5097	5.0268	0.4972	-11.9217	145.5293
2.5	2.5	2.5	2.5	2.8445	3.8378	3.9325	0.5167	-10.6613	137.0961
			3	2.6116	3.2846	1.9893	0.4294	-15.0613	19.4437
			3.5	2.4342	2.9348	1.2665	0.3835	-18.6267	10.6813
			4	2.2929	2.6874	0.9093	0.3548	-21.4601	7.7605

2.2.2 Special cases

WGB2 includes several other distributions as special or limiting cases, such as weighted generalized gamma (WGG), weighted beta of the second kind (WB2), weighted Singh-Maddala (WSM), weighted Dagum (WD), weighted gamma (WG), weighted Weibull (WW) and weighted exponential (WE) distributions.

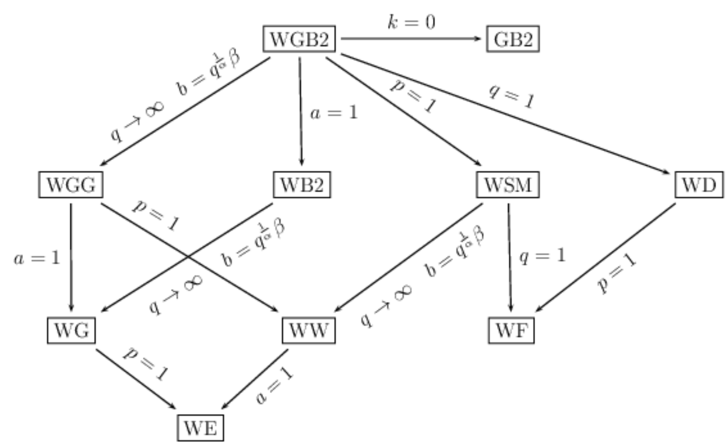

Figure 2.4: Distribution Graph

We can also obtain the j^{th} moments of these distributions with the weight function $w(y) = y^k$.[1]

- Weighted Singh-Maddala (p = 1)

$$E_{G_{WSM}}(Y^j) = \frac{b^j B(1 + \frac{k}{a} + \frac{j}{a}, q - \frac{k}{a} - \frac{j}{a})}{B(1 + \frac{k}{a}, q - \frac{k}{a})}. \tag{2.11}$$

[1]In the special cases, one should consider the restrictions on the values of k and j.

- Weighted Dagum (q = 1)

$$E_{G_{WD}}(Y^j) = \frac{b^j B(p + \frac{k}{a} + \frac{j}{a}, 1 - \frac{k}{a} - \frac{j}{a})}{B(p + \frac{k}{a}, 1 - \frac{k}{a})}. \tag{2.12}$$

- Weighted Beta of the Second Kind (a = 1)

$$E_{G_{WB2}}(Y^j) = \frac{b^j B(p + k + j, q - k - j)}{B(p + k, q - j)}. \tag{2.13}$$

- Weighted Generalized Gamma ($b = q^{\frac{1}{\alpha}}\beta$ as $q \to \infty$)

$$E_{G_{WGG}}(Y^j) = \frac{\beta^j \Gamma(p + \frac{k}{a} + \frac{j}{a})}{\Gamma(p + \frac{k}{a})}. \tag{2.14}$$

- Weighted Fisk (p = 1, q = 1)

$$E_{G_{WF}}(Y^j) = \frac{b^j B(1 + \frac{k}{a} + \frac{j}{a}, 1 - \frac{k}{a} - \frac{j}{a})}{B(1 + \frac{k}{a}, 1 - \frac{k}{a})}. \tag{2.15}$$

- Weighted Gamma (a = 1, $b = q^{\frac{1}{\alpha}}\beta$ as $q \to \infty$)

$$E_{G_{WG}}(Y^j) = \frac{\beta^j \Gamma(p + k + j)}{\Gamma(p + k)}. \tag{2.16}$$

- Weighted Weibull (p = 1, $b = \beta q^{\frac{1}{a}}$ as $q \to \infty$)

$$E_{G_{WW}}(Y^j) = \frac{\beta^j (k + j)\Gamma(\frac{k}{a} + \frac{j}{a})}{k\Gamma(\frac{k}{a})}. \tag{2.17}$$

- Weighted Exponential (a = p = 1, $b = \beta q^{\frac{1}{a}}$ as $q \to \infty$)

$$E_{G_{WE}}(Y^j) = \frac{\beta^j (k + j)!}{k!}. \tag{2.18}$$

2.3 Concluding Remarks

In chapter 2, the weighted generalized beta distribution of the second kind (WGB2) is presented. We showed that WGB2 includes several other distributions as special and limiting cases. The limiting and special cases include weighted generalized gamma (WGG), weighted beta of the second kind (WB2), weighted Singh- Maddala (WSM), weighted Dagum (WD), weighted gamma (WG), weighted Weibull (WW) and weighted exponential (WE) distributions as well as their unweighted or parent versions. Statistical properties of the weighted generalized beta distribution of the second kind (WGB2) including the cdf, hazard functions, monotonicity, and income-share elasticity are also presented. The moments of WGB2 as well as the mean, variance, coefficient of skewness and coefficient of kurtosis are presented.

CHAPTER 3

ENTROPY OF WEIGHTED GENERALIZED BETA DISTRIBUTION OF THE SECOND KIND

3.1 Generalized Entropy

Generalized entropy (GE) is widely used to measure inequality trends and differences. It is primarily used in income distribution. Kleiber and Kotz (2003) derived Theil index for GB2 and Singh-Maddala model.

The generalized entropy (GE) $I(\alpha)$ is defined as:

$$I(\alpha) = \frac{v_\alpha \mu^{-\alpha} - 1}{\alpha(\alpha - 1)}, \alpha \neq 0, \alpha \neq 1, \tag{3.1}$$

where $v_\alpha = \int y^\alpha dF(y)$, $\mu \equiv E(Y)$ is the mean, and $F(y)$ is the cumulative distribution function (cdf) of the random variable Y. The bottom-sensitive index is $I(-1)$, and the top-sensitive index is $I(2)$.

The mean logarithmic deviation (MLD) index is given by:

$$I(0) = lim_{\alpha \to 0} I(\alpha) = log\mu - v_0. \tag{3.2}$$

and Theil index is:

$$I(1) = lim_{\alpha \to 1} I(\alpha) = \frac{\mu}{v_1} - log\mu. \tag{3.3}$$

The generalized entropy of GB2 is given by Jenkins (2007) as:

$$I(\alpha) = \frac{B(p + \frac{\alpha}{a}, q - \frac{\alpha}{a}) B^{-\alpha}(p + \frac{1}{a}, q - \frac{1}{a}) - B^{1-\alpha}(p, q)}{\alpha(\alpha - 1) B^{1-\alpha}(p, q)}, \alpha \neq 0, \alpha \neq 1,$$

with

$$I(0) = \Gamma\left(p + \frac{1}{a}\right) + \Gamma\left(q - \frac{1}{a}\right) - \Gamma(p) - \Gamma(q) - \frac{\psi(p)}{a} - \frac{\psi(q)}{a},$$

and

$$I(1) = \frac{\psi(p+\frac{1}{a})}{a} - \frac{\psi(q-\frac{1}{a})}{a} - \Gamma\left(p+\frac{1}{a}\right) - \Gamma\left(q-\frac{1}{a}\right) + \Gamma(p) + \Gamma(q).$$

From our previous discussions about WGB2, v_α and μ are given by:

$$v_\alpha = \frac{b^\alpha B(p+\frac{k}{a}+\frac{\alpha}{a}, q-\frac{k}{a}-\frac{\alpha}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})}, \quad \text{and} \qquad \mu = \frac{bB(p+\frac{k}{a}+\frac{1}{a}, q-\frac{k}{a}-\frac{1}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})},$$

respectively.

Consequently, the generalized entropy of WGB2 is given by:

$$I(\alpha) = \frac{B(p+\frac{k}{a}+\frac{\alpha}{a}, q-\frac{k}{a}-\frac{\alpha}{a})B^{-\alpha}(p+\frac{k}{a}+\frac{1}{a}, q-\frac{k}{a}-\frac{1}{a}) - B^{1-\alpha}(p+\frac{k}{a}, q-\frac{k}{a})}{\alpha(\alpha-1)B^{1-\alpha}(p+\frac{k}{a}, q-\frac{k}{a})}, \tag{3.4}$$

where $\alpha \neq 0$ and $\alpha \neq 1$. Note that $I(\alpha)$ does not depend on the scale parameter b.

When $\alpha = 0$ or $\alpha = 1$, set $m(\alpha) = v_\alpha \mu^{-\alpha} - 1$, $n(\alpha) = \alpha(\alpha-1)$, then $I(\alpha) = \frac{m(\alpha)}{n(\alpha)}$. By L'Hopital's rule, we have $I(0) = -m'(0)$, $I(1) = m'(1)$,

$$m'(\alpha) = (\mu^{-\alpha})' v_\alpha + (\mu^{-\alpha}) v'_\alpha, \quad (\mu^{-\alpha})' = -\mu^{-\alpha} log\mu,$$

and

$$v'_\alpha = v_\alpha\left[\frac{\psi(p+\frac{k}{a}+\frac{\alpha}{a})}{a} - \frac{\psi(q-\frac{k}{a}-\frac{\alpha}{a})}{a} + logb\right], \quad where \quad \psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}.$$

Consequently the MLD index and Theil index, I(0) and I(1) of WGB2 are

$$I(0) = log\frac{B(p+\frac{k+1}{a}, q-\frac{k+1}{a})}{B(p+\frac{k}{a}, q-\frac{k}{a})} - \frac{\psi(p+\frac{k}{a})}{a} - \frac{\psi(q-\frac{k}{a})}{a}, \tag{3.5}$$

and

$$I(1) = \frac{\psi(p + \frac{k+1}{a})}{a} - \frac{\psi(q - \frac{k+1}{a})}{a} - log\frac{B(p + \frac{k+1}{a}, q - \frac{k+1}{a})}{B(p + \frac{k}{a}, q - \frac{k}{a})} \quad (3.6)$$

respectively. Note that I(0) and I(1) does not depend on the scale parameter b.

We select the values of the parameters a, p, q and compute the bottom-sensitive index (I(-1)), top-sensitive index (I(2)), mean logarithmic deviation (MLD) index (I(0)) and Theil index (I(1)) in Tables 3.1 and 3.2.

Table 3.1: Generalized entropy of WGB2 with k=1

a	p	q	I(-1)	I(2)	I(0)	I(1)
2.5	2.5	2.5	0.0891	0.102	0.0841	0.2544
3			0.0594	0.0649	0.0572	0.2649
3.5			0.0426	0.0454	0.0414	0.2569
4			0.0322	0.0337	0.0315	0.2433
2.5	2.5	2.5	0.0891	0.102	0.0841	0.2544
	3		0.0826	0.0969	0.0791	0.2497
	3.5		0.0779	0.0931	0.0754	0.2464
	4		0.0745	0.0903	0.0726	0.2438
2.5	2.5	2.5	0.0891	0.102	0.0841	0.2544
		3	0.0759	0.0804	0.0712	0.5079
		3.5	0.0678	0.0687	0.0633	0.701
		4	0.0622	0.0613	0.058	0.8568

Table 3.2: Generalized entropy of WGB2 with k=2

a	p	q	I(-1)	I(2)	I(0)	I(1)
2.5	2.5	2.5	0.102	0.1335	0.0981	-0.0294
3			0.0649	0.0764	0.0633	0.0908
3.5			0.0454	0.0507	0.0446	0.1389
4			0.0337	0.0365	0.0333	0.1579
2.5	2.5	2.5	0.102	0.1335	0.0981	-0.0294
	3		0.0969	0.1293	0.0942	-0.033
	3.5		0.0931	0.1261	0.0913	-0.0357
	4		0.0903	0.1236	0.089	-0.0379
2.5	2.5	2.5	0.102	0.1335	0.0981	-0.0294
		3	0.0804	0.0922	0.0767	0.3079
		3.5	0.0687	0.0735	0.0651	0.5465
		4	0.0613	0.063	0.0579	0.7308

From the tables, we observe that:

1) I(-1), I(2) and I(0) decreases as a, p, q increases, and does not change as b increases, since these indices do not depend on the scale parameter b.

2) There is no specific pattern for I(-1), I(1) increases as a increases for the chosen values of the parameters.

3.2 Renyi Entropy

Renyi (1961) extended the concept of Shannon's entropy and define Renyi entropy as follows:

$$I_\alpha(\beta) = \frac{1}{1-\beta} log\left(\int_0^\infty f^\beta(y)dy\right), \quad \beta > 0, \beta \neq 1. \tag{3.7}$$

Renyi entropy is important in statistics as a measure of indices of diversity, and tends to Shannon entropy (1948) as $\beta \to 0$.

3.2.1 Renyi Entropy of GB2

Recall the pdf of GB2 distribution is given by:

$$f_{GB2}(y; a, b, p, q) = \frac{ay^{ap-1}}{b^{ap}B(p,q)[1+(\frac{y}{b})^a]^{p+q}} \quad \text{for } y > 0, \text{ and } 0 \text{ otherwise}$$

Note that:

$$f_{GB2}^\beta(y; a, b, p, q) = \left(\frac{a}{bB(p,q)}\right)^\beta \left[\left(\frac{y}{b}\right)^a\right]^{\beta(p-\frac{1}{a})} \left[1+\left(\frac{y}{b}\right)^a\right]^{-\beta(p+q)},$$

where $y > 0$, $a, b, p, q > 0$, $\beta > 0$. Let $(\frac{y}{b})^a = t$, then $dy = \frac{b}{a} t^{\frac{1}{a}-1} dt$, and

$$\begin{aligned} \int_0^\infty f_{GB2}^\beta(y; a, b, p, q)dy &= \left(\frac{a}{b}\right)^{\beta-1} \frac{1}{B^\beta(p,q)} \int_0^\infty t^{\beta p - \frac{\beta}{a} + \frac{1}{a} - 1} \left(\frac{1}{t+1}\right)^{\beta(p+q)} \\ &= \left(\frac{a}{b}\right)^{\beta-1} \frac{B(\beta p - \frac{\beta}{a} + \frac{1}{a}, \beta q + \frac{\beta}{a} - \frac{1}{a} + 2)}{B^\beta(p,q)}. \end{aligned}$$

Consequently, Renyi entropy for GB2 simplifies to:

$$I_R(\beta) = log\left(\frac{b}{a}\right) - \frac{\beta log B(p,q)}{1-\beta} + \frac{B(\beta p - \frac{\beta}{a} + \frac{1}{a}, \beta q + \frac{\beta}{a} - \frac{1}{a} + 2)}{1-\beta}, \tag{3.8}$$

for $a, b, p, q > 0$, $\beta > 0$, $\beta \neq 1$.

3.2.2 Renyi Entropy of WGB2

Recall the pdf of WGB2 distribution is given by:

$$g_{WGB2}(y;a,b,p,q,k) = \frac{ay^{ap+k-1}}{b^{ap+k}B(p+\frac{k}{a},q-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+q}} \tag{3.9}$$

for $y > 0, a, b, p, q, k > 0$, and $-ap < k < aq$. Note that:

$$g_w^{\beta}(y;a,b,p,q,k) = \frac{(\frac{a}{b})^{\beta}[(\frac{y}{b})^a]^{\beta p+\frac{\beta k}{a}-\frac{\beta}{a}}[1+(\frac{y}{b})^a]^{-\beta(p+q)}}{B^{\beta}(p+\frac{k}{a},q-\frac{k}{a})}$$

for $a, b, p, q > 0$, $-ap < k < aq$, $\beta > 0$, $\beta \neq 1$. Let $(\frac{y}{b})^a = t$, then $dy = \frac{b}{a}t^{\frac{1}{a}-1}dt$, and

$$\begin{aligned}\int_0^{\infty} g_w^{\beta}(y)dy &= \frac{(\frac{a}{b})^{\beta-1}}{B^{\beta}(p+\frac{k}{a},q-\frac{k}{a})}\int_0^{\infty}\left(1-\frac{1}{1+t}\right)^{\beta p+\frac{\beta k}{a}-\frac{\beta}{a}+\frac{1}{a}-1}\left(\frac{1}{1+t}\right)^{\beta q-\frac{\beta k}{a}+\frac{\beta}{a}-\frac{1}{a}+2-1}dt \\ &= \frac{(\frac{a}{b})^{\beta-1}B(\beta p+\frac{\beta k}{a}-\frac{\beta}{a}+\frac{1}{a},\beta q-\frac{\beta k}{a}+\frac{\beta}{a}-\frac{1}{a}+2)}{B^{\beta}(p+\frac{k}{a},q-\frac{k}{a})}.\end{aligned}$$

Consequently, Renyi entropy for WGB2 reduces to:

$$I_R(\beta) = log\left(\frac{b}{a}\right) - \frac{\beta logB(p+\frac{k}{a},q-\frac{k}{a})}{1-\beta} + \frac{logB(\beta p+\frac{\beta k}{a}-\frac{\beta}{a}+\frac{1}{a},\beta q-\frac{\beta k}{a}+\frac{\beta}{a}-\frac{1}{a}+2)}{1-\beta}, \tag{3.10}$$

for $a, b, p, q > 0$, $-ap < k < aq$, $\beta > 0$, $\beta \neq 1$.

We select values of the parameters a, b, p, q and compute Renyi entropy for different values of β in Tables 3.3 and 3.4. From the tables we observe the following:

1) When k = 1: for $\beta < 1$, Renyi entropy increases as b, q increases; for $\beta > 1$, Renyi entropy increases as b increases, and decreases as a, p, q increases.

2) When k = 2: for $\beta < 1$, Renyi entropy increases as a, b increases; for $\beta > 1$, Renyi entropy increases as b increases, and decreases as a, p, q increases.

Table 3.3: Renyi entropy of WGB2 with k=1

a	b	p	q	$\beta = 0.25$	$\beta = 0.5$	$\beta = 0.75$	$\beta = 1.25$	$\beta = 1.5$
2.5	2.5	2.5	2.5	0.0043	0.53	4.5354	-17.6808	-13.5768
3				-0.0921	0.421	4.4616	-17.9609	-13.8119
3.5				-0.1795	0.3199	4.3792	-18.1696	-13.9942
4				-0.2599	0.2266	4.2964	-18.3358	-14.1437
2.5	2.5	2.5	2.5	0.0043	0.53	4.5354	-17.6808	-13.5768
	3			0.1866	0.7124	4.7177	-17.4985	-13.3945
	3.5			0.3407	0.8665	4.8719	-17.3444	-13.2404
	4			0.4743	1	5.0054	-17.2108	-13.1068
2.5	2.5	2.5	2.5	0.0043	0.53	4.5354	-17.6808	-13.5768
		3		-0.0843	0.5318	5.027	-19.6824	-15.0812
		3.5		-0.169	0.5261	5.4507	-21.4444	-16.4077
		4		-0.25	0.5158	5.8225	-23.0191	-17.5949
2.5	2.5	2.5	2.5	0.0043	0.53	4.5354	-17.6808	-13.5768
			3	0.1044	0.8711	5.7364	-20.3785	-15.4183
			3.5	0.1848	1.1532	6.75	-22.702	-17.0128
			4	0.251	1.3931	7.6272	-24.7466	-18.4216

Table 3.4: Renyi entropy of WGB2 with k=2

a	b	p	q	$\beta = 0.25$	$\beta = 0.5$	$\beta = 0.75$	$\beta = 1.25$	$\beta = 1.5$
2.5	2.5	2.5	2.5	-0.1758	0.1726	3.6929	-16.5623	-12.9325
3				-0.2436	0.1417	3.846	-17.231	-13.4088
3.5				-0.311	0.091	3.9028	-17.6653	-13.7291
4				-0.3764	0.0329	3.9124	-17.9732	-13.9639
2.5	2.5	2.5	2.5	-0.1758	0.1726	3.6929	-16.5623	-12.9325
	3			0.0065	0.355	3.8752	-16.38	-12.7502
	3.5			0.1607	0.5091	4.0294	-16.2259	-12.596
	4			0.2942	0.6427	4.1629	-16.0923	-12.4625
2.5	2.5	2.5	2.5	-0.1758	0.1726	3.6929	-16.5623	-12.9325
		3		-0.2704	0.1397	4.0278	-18.1282	-14.1235
		3.5		-0.3602	0.1046	4.3179	-19.5189	-15.1831
		4		-0.4454	0.0686	4.5734	-20.7701	-16.1377
2.5	2.5	2.5	2.5	-0.1758	0.1726	3.6929	-16.5623	-12.9325
			3	-0.0429	0.6121	5.2145	-19.9289	-15.2229
			3.5	0.0616	0.9675	6.4713	-22.7763	-17.173
			4	0.1466	1.2651	7.5436	-25.2522	-18.8767

3.3 Concluding Remarks

In chapter 3, the entropy of Weighted Generalized Beta Distribution of the Second Kind (WGB2) are studied. Generalized entropy and beta entropy which are measures of the uncertainty in this class of distributions are derived.Top-sensitive index, bottom-sensitive index, mean logarithmic deviation (MLD) index and Theil index obtained from generalized entropy (GE) are also presented.

CHAPTER 4

RESULTS ON DAGUM AND RELATED DISTRIBUTIONS

4.1 Dagum Distribution

Camilo Dagum (1977) proposed the Dagum distribution for the size distribution of personal income. This distribution is a special case of generalized beta distribution of the second kind(GB2), when q = 1. Kleiber (2007) traced the genesis of Dagum distribution and summeraized several statistical properties of this distribution. Domma (2011) provided the calculation of the Fisher information matrix of Dagum distribution under type I right censoring.

The pdf of generalized beta distribution of the second kind is:

$$f_{GB2}(y;a,b,p,q) = \frac{ay^{ap-1}}{b^{ap}B(p,q)[1+(\frac{y}{b})^a]^{p+q}} \quad \text{for } y>0, \text{ and 0 otherwise.} \tag{4.1}$$

When q = 1, 4.1 becomes the Dagum distribution, and the pdf is given by:

$$\begin{aligned} f_D(y;a,b,p) &= \frac{apy^{ap-1}}{b^{ap}[1+(\frac{y}{b})^a]^{p+1}}. \\ &= \frac{apb^a}{y^{a+1}}\left[\frac{(\frac{y}{b})^a}{1+(\frac{y}{b})^a}\right]^{p+1} \\ &= apb^a y^{-a-1}(b^a y^{-a}+1)^{-p-1}, y>0, a,b,p>0. \end{aligned}$$

Let $b^a = \lambda$, $a = \delta$, $p = \beta$, then the pdf of Dagum distribution can be written as:

$$f_D(y;\beta,\lambda,\delta) = \beta\lambda\delta y^{-\delta-1}(\lambda y^{-\delta}+1)^{-\beta-1}, y>0, \lambda,\beta,\delta>0. \tag{4.2}$$

The corresponding cdf is given by:

$$F_D(y;\beta,\lambda,\delta) = (1+\lambda y^{-\delta})^{-\beta}, y > 0, \lambda, \beta, \delta > 0. \tag{4.3}$$

The hazard function and the reverse hazard function are given by:

$$h_D(y;\lambda,\beta,\delta) = \frac{f_D(y;\lambda,\beta,\delta)}{\overline{F}_D(y;\lambda,\beta,\delta)} = \frac{\beta\lambda\delta y^{-\delta-1}(\lambda y^{-\delta}+1)^{-\beta-1}}{1-(1+\lambda y^{-\delta})^{-\beta}}, \tag{4.4}$$

and

$$\tau_D(y;\lambda,\beta,\delta) = \beta\lambda\delta y^{-\delta-1}(\lambda y^{-\delta}+1)^{-1}, y > 0, \lambda, \beta, \delta > 0. \tag{4.5}$$

respectively.

The k^{th} raw or noncentral moments are:

$$E(Y^k) = \beta\lambda^{\frac{k}{\delta}} B(\beta + \frac{k}{\delta}, 1 - \frac{k}{\delta}), \text{for } \delta > k.$$

4.2 Weighted Dagum Distribution

Recall the pdf of WGB2 is given by:

$$f_{GB2}(y;a,b,p,q) = \frac{a y^{ap+k-1}}{b^{ap+k} B(p+\frac{k}{a}, q-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+q}} \quad \text{for } y > 0, \text{ and } 0 \text{ otherwise.} \tag{4.6}$$

When q = 1, this pdf becomes the weighted Dagum pdf:

$$\begin{aligned} f_{WD}(y;a,b,p,q) &= \frac{ay^{ap+k-1}}{b^{ap+k}B(p+\frac{k}{a},1-\frac{k}{a})[1+(\frac{y}{b})^a]^{p+1}} \\ &= \frac{1}{B(p+\frac{k}{a},1-\frac{k}{a})}\frac{ay^{k-1}}{b^k}\frac{(\frac{y}{b})^{ap}}{[1+(\frac{y}{b})^a]^{p+1}} \\ &= \frac{1}{B(p+\frac{k}{a},1-\frac{k}{a})}\frac{ay^{k-1}}{b^k}\frac{b^a}{y^a}\left[1+\left(\frac{b}{y}\right)^a\right]^{-p-1} \\ &= \frac{1}{B(p+\frac{k}{a},1-\frac{k}{a})}ay^{-a-1+k}(b^a)^{1-\frac{k}{a}}[1+b^ay^{-a}]^{-p-1}, \end{aligned}$$

for $y > 0, a, b, p, k > 0$.

Let $b^a = \lambda$, $a = \delta$, $p = \beta$, then Weighted Dagum pdf can be written as:

$$f_{WD}(y;\beta,\lambda,\delta,k) = \frac{1}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}\delta\lambda^{1-\frac{k}{\delta}}y^{-\delta-1+k}(\lambda y^{-\delta}+1)^{-\beta-1},$$

for $y > 0, \beta, \lambda, \delta > 0, -\delta\beta < k < \delta$.

The cdf and hazard functions of Weighted Dagum distribution can be obtained from equations (2.2) and (2.3) by setting: $b^a = \lambda$, $a = \delta$, $p = \beta$, $q = 1$.

The cdf of WD is given by:

$$F_{WD}(y;\lambda,\delta,\beta,k) = 1 - I_{(1+\frac{y^\delta}{\lambda})^{-1}}\left(\beta+\frac{k}{\delta},1-\frac{k}{\delta}\right),$$

The hazard function of WD is given by:

$$\begin{aligned} h_{WD}(y;\lambda,\delta,\beta,k) &= \frac{f_{WD}(y;\lambda,\delta,\beta,k)}{\overline{F}_{WD}(y;\lambda,\delta,\beta,k)} \\ &= \frac{\delta\lambda^{1-\frac{k}{\delta}}y^{-\delta-1+k}(\lambda y^{-\delta}+1)^{-\beta-1}}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})I_{(1+\frac{y^\delta}{\lambda})^{-1}}\left(\beta+\frac{k}{\delta},1-\frac{k}{\delta}\right)}, \end{aligned}$$

for $y > 0, \beta, \lambda, \delta > 0, -\delta\beta < k < \delta$.

The j^{th} raw or noncentral moments are:

$$\begin{aligned}
E(Y^j) &= \int_0^\infty \frac{1}{B(\beta-\frac{k}{\delta},1-\frac{k}{\delta})}\delta\lambda^{1-\frac{k}{\delta}}y^{-\delta-1+k+j}(1+\lambda y^{-\delta})^{-\beta-1}dy \\
&= \int_0^\infty \frac{1}{B(\beta-\frac{k}{\delta},1-\frac{k}{\delta})}\lambda^{-\frac{k}{\delta}}(y^{k+j})(\lambda\delta y^{-\delta-1}dy)(1+\lambda y^{-\delta})^{-\beta-1}dy \\
&= \int_0^1 \frac{1}{B(\beta-\frac{k}{\delta},1-\frac{k}{\delta})}\lambda^{-\frac{k}{\delta}}\left(\frac{t^{-1}-1}{\lambda}\right)^{-\frac{k+j}{\delta}}t^{-2}t^{\beta+1}dt \\
&= \int_0^1 \frac{1}{B(\beta-\frac{k}{\delta},1-\frac{k}{\delta})}\lambda^{-\frac{k}{\delta}+\frac{k+j}{\delta}}(1-t)^{-\frac{k+j}{\delta}}t^{\beta-\frac{k+j}{\delta}-1}dt \\
&= \frac{\lambda^{\frac{j}{\delta}}B(\beta-\frac{k+j}{\delta},1-\frac{k+j}{\delta})}{B(\beta-\frac{k}{\delta},1-\frac{k}{\delta})},
\end{aligned}$$

for $\delta > k$.

4.3 Fisher Information Matrix

Let $\theta = (\beta, \lambda, \delta)$, the fisher information is given by:

$$I(\theta) = E[\frac{\partial}{\partial\theta}log(f(Y;\theta))]^2. \tag{4.7}$$

If the second derivative with respect to θ exists for all Y and θ, and the second derivative with respect to θ of $\int f(Y;\theta)dy = 1$ can be obtained by differentiating twice under the integral sign, then the Fisher information becomes:

$$I(\theta) = -E_\theta[\frac{\partial^2}{\partial\theta^2}log(f(Y;\theta))]. \tag{4.8}$$

The logarithm of the weighted Dagum pdf is:

$$\begin{aligned}
& log(f_{WD}(y)) \\
= \; & log(f_{WD}(y;\beta,\lambda,\delta,k)) \\
= \; & log\Gamma(\beta+1) - log\Gamma\left(\beta+\frac{k}{\delta}\right) - log\Gamma\left(1-\frac{k}{\delta}\right) + log(\delta) + \left(1-\frac{k}{\delta}\right)log\lambda \\
& -(\delta+1-k)log(y) - (\beta+1)log(1+\lambda y^{-\delta}).
\end{aligned}$$

The first, second and mixed partial derivatives are:

$$\begin{aligned}
\frac{\partial log f_{WD}(y)}{\partial\beta} &= \frac{\Gamma'(\beta+1)}{\Gamma(\beta+1)} - \frac{\Gamma'(\beta+\frac{k}{\delta})}{\Gamma(\beta+\frac{k}{\delta})} - log(1+\lambda y^{-\delta}) \\
&= \psi(\beta+1) - \psi\left(\beta+\frac{k}{\delta}\right) - log(1+\lambda y^{-\delta}),
\end{aligned}$$

$$\frac{\partial^2 log f_{WD}(y)}{\partial\beta^2} = \psi'(\beta+1) - \psi'\left(\beta+\frac{k}{\delta}\right),$$

$$\frac{\partial^2 log f_{WD}(y)}{\partial\beta\partial\lambda} = \frac{\partial^2 log f(y)}{\partial\lambda\partial\beta} = -\frac{y^{-\delta}}{1+\lambda y^{-\delta}},$$

$$\frac{\partial^2 log f_{WD}(y)}{\partial\beta\partial\delta} = \frac{\partial^2 log f(y)}{\partial\delta\partial\beta} = \frac{k}{\delta^2}\psi'\left(\beta+\frac{k}{\delta}\right) + \frac{\lambda y^{-\delta} log(y)}{1+\lambda y^{-\delta}},$$

$$\frac{\partial log f_{WD}(y)}{\partial\lambda} = \frac{1}{\lambda}\left(1-\frac{k}{\delta}\right) - (\beta+1)\frac{y^{-\delta}}{1+\lambda y^{-\delta}},$$

$$\frac{\partial^2 log f_{WD}(y)}{\partial\lambda^2} = \frac{1}{\lambda^2}\left(\frac{k}{\delta}-1\right) + (\beta+1)\frac{(y^{-\delta})^2}{(1+\lambda y^{-\delta})^2},$$

$$\frac{\partial^2 log f_{WD}(y)}{\partial\lambda\partial\delta} = \frac{\partial^2 log f(y)}{\partial\delta\partial\lambda} = \frac{k}{\lambda\delta^2} + (\beta+1)\frac{y^{-\delta} ln(y)}{(1+\lambda y^{-\delta})^2},$$

$$\begin{aligned}\frac{\partial log f_{WD}(y)}{\partial\delta} &= \frac{k\Gamma'(\beta+\frac{k}{\delta})}{\delta^2\Gamma(\beta+\frac{k}{\delta})} - \frac{k\Gamma'(1-\frac{k}{\delta})}{\delta^2\Gamma(1-\frac{k}{\delta})} + \frac{1}{\delta} + \frac{k}{\delta^2} ln(\lambda) - log(y) + (\beta+1)\frac{\lambda y^{-\delta} log(y)}{1+\lambda y^{-\delta}}\\ &= \frac{k}{\delta^2}\left(\psi\left(\beta+\frac{k}{\delta}\right) - \psi(1-\frac{k}{\delta})\right) + \frac{1}{\delta} + \frac{k}{\delta^2} log(\lambda) - log(y) + (\beta+1)\frac{\lambda y^{-\delta} log(y)}{1+\lambda y^{-\delta}},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 log f_{WD}(y)}{\partial\delta^2} = &-\frac{2k}{\delta^3}\left(\psi\left(\beta+\frac{k}{\delta}\right) - \psi\left(1-\frac{k}{\delta}\right)\right) - \frac{k^2}{\delta^4}\left(\psi'\left(\beta+\frac{k}{\delta}\right) - \psi'\left(1-\frac{k}{\delta}\right)\right) - \frac{1}{\delta^2}\\ &-\frac{2k}{\delta^3} log(\lambda) + \lambda(\beta+1)\left(\frac{y^{-\delta} log(y)}{1+\lambda y^{-\delta}}\right)^2.\end{aligned}$$

We now take the expectations and simplify the results.

$$\begin{aligned}E_1 = E\left(\frac{Y^{-\delta}}{1+\lambda Y^{-\delta}}\right) &= \int_0^\infty \frac{y^{-\delta}}{1+\lambda y^{-\delta}} \frac{1}{B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \delta\lambda^{1-\frac{k}{\delta}} y^{-\delta-1+k} (\lambda y^{-\delta}+1)^{-\beta-1} dy\\ &= \int_0^\infty \frac{1}{B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \delta\lambda^{1-\frac{k}{\delta}} y^{-2\delta-1+k} (\lambda y^{-\delta}+1)^{-\beta-2} dy.\end{aligned}$$

Let $(\lambda y^{-\delta}+1)^{-1} = t$, then $y^{-\delta} = \frac{t^{-1}-1}{\lambda} = \frac{1-t}{\lambda t}$, and $\lambda\delta y^{-\delta-1} dy = t^{-2} dt$,

so that:

$$
\begin{aligned}
E_1 &= \int_0^\infty \frac{1}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}\lambda^{-\frac{k}{\delta}}(\lambda\delta y^{-\delta-1}dy)(y^{-\delta})^{1-\frac{k}{\delta}}(\lambda y^{-\delta}+1)^{-\beta-2} \\
&= \int_0^1 \frac{1}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}\lambda^{-\frac{k}{\delta}}t^{-2}\left(\frac{1-t}{\lambda t}\right)^{1-\frac{k}{\delta}}t^{\beta+2}dt \\
&= \int_0^1 \frac{\lambda^{-\frac{k}{\delta}}}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}(1-t)^{1-\frac{k}{\delta}}\lambda^{\frac{k}{\delta}-1}t^{\beta-1+\frac{k}{\delta}}dt \\
&= \int_0^1 \frac{1}{\lambda B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}(1-t)^{2-\frac{k}{\delta}-1}t^{\beta+\frac{k}{\delta}-1}dt \\
&= \frac{B(\beta+\frac{k}{\delta},2-\frac{k}{\delta})}{\lambda B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}. \qquad (4.9)
\end{aligned}
$$

Similarly, we obtain:

$$E_2 = E\left[\frac{Y^{-\delta}}{(1+\lambda Y^{-\delta})^2}\right] = \frac{B(\beta+\frac{k}{\delta}+1,2-\frac{k}{\delta})}{\lambda B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}, \qquad (4.10)$$

Also,

$$
\begin{aligned}
E_3 = E\left(\frac{Y^{-\delta}logY}{1+\lambda Y^{-\delta}}\right) &= \int_0^\infty \frac{y^{-\delta}logy}{1+\lambda y^{-\delta}}\frac{1}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}\delta\lambda^{1-\frac{k}{\delta}}y^{-\delta-1+k}(\lambda y^{-\delta}+1)^{-\beta-1}dy \\
&= \int_0^\infty \frac{1}{B(\beta+\frac{k}{\delta},1-\frac{k}{\delta})}\delta\lambda^{1-\frac{k}{\delta}}y^{-2\delta-1+k}(\lambda y^{-\delta+1})^{-\beta-2}ln(y)dy,
\end{aligned}
$$

Let $(\lambda y^{-\delta}+1)^{-1}=t$, then $logy = -\frac{1}{\delta}[log(1-t)-log\lambda-log(t)]$, and we have:

$$E_3 = \int_0^1 \frac{1}{B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \lambda^{-\frac{k}{\delta}} t^{-2} \left(\frac{1-t}{\lambda t}\right)^{1-\frac{k}{\delta}} t^{\beta+2} \frac{1}{\delta}[log(t) + log\lambda - log(1-t)]dt$$
$$= \int_0^1 \frac{1}{\lambda\delta B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})}(1-t)^{1-\frac{k}{\delta}} t^{\beta+\frac{k}{\delta}-1}[log(t) + log\lambda - log(1-t)]dt$$
$$= \frac{1}{\lambda\delta B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})}\Bigg[log\lambda \int_0^1 (1-t)^{1-\frac{k}{\delta}} t^{\beta+\frac{k}{\delta}-1}dt + \int_0^1 (1-t)^{1-\frac{k}{\delta}} t^{\beta+\frac{k}{\delta}-1} log(t)dt$$
$$- \int_0^1 (1-t)^{1-\frac{k}{\delta}} t^{\beta+\frac{k}{\delta}-1} log(1-t)dt\Bigg]$$
$$= \frac{1}{\lambda\delta B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})}\Bigg[(log\lambda)B\bigg(\beta+\frac{k}{\delta}, 2-\frac{k}{\delta}\bigg) + B\bigg(\beta+\frac{k}{\delta}, 2-\frac{k}{\delta}\bigg)\bigg(\psi\bigg(\beta+\frac{k}{\delta}\bigg)$$
$$- \psi(\beta+2)\bigg) - B\bigg(\beta+\frac{k}{\delta}, 2-\frac{k}{\delta}\bigg)\bigg(\psi\bigg(2-\frac{k}{\delta}\bigg) - \psi\bigg(\beta+2)\bigg)\Bigg]$$
$$= \frac{B(\beta+\frac{k}{\delta}, 2-\frac{k}{\delta})}{\lambda\delta B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})}\Bigg[log\lambda + \psi\bigg(\beta+\frac{k}{\delta}\bigg) - \psi\bigg(2-\frac{k}{\delta}\bigg)\Bigg].$$

Similarly,

$$E_4 = E\bigg(\frac{Y^{-\delta}logY}{(1+\lambda Y^{-\delta})^2}\bigg) \quad = \quad \frac{B(\beta+\frac{k}{\delta}, 2-\frac{k}{\delta})}{\lambda\delta B(\beta-\frac{k}{\delta}, 1-\frac{k}{\delta})}\Bigg[log\lambda + \psi(\beta+\frac{k}{\delta}+1) - \psi(2-\frac{k}{\delta})\Bigg].$$

Also,

$$E_5 = E\Bigg[Y^{-\delta}\bigg(\frac{logY}{1+\lambda Y^{-\delta}}\bigg)^2\Bigg] = \int_0^1 \frac{y^{-\delta}(logy)^2}{(1+\lambda y^{-\delta})^2} \frac{1}{B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \delta\lambda^{1-\frac{k}{\delta}} y^{-\delta-1+k}(\lambda y^{-\delta}+1)^{-\beta-1}dy.$$

Let $(\lambda y^{-\delta}+1)^{-1} = t$, then $logy = -\frac{1}{\delta}[log(1-t) - log\lambda - log(t)]$, and

$$E_5 = \int_0^1 \frac{1}{B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \lambda^{-\frac{k}{\delta}} t^{-2} \left(\frac{1-t}{\lambda t}\right)^{1-\frac{k}{\delta}} t^{\beta+3} \frac{1}{\delta^2} [log(t) + log\lambda - log(1-t)]^2 dt$$

$$= \int_0^1 \frac{1}{\lambda\delta^2 B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} (1-t)^{1-\frac{k}{\delta}} t^{\beta+\frac{k}{\delta}} [log(t) + log\lambda - log(1-t)]^2 dt$$

$$= \frac{1}{\lambda\delta^2 B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \int_0^1 t^{\beta+\frac{k}{\delta}} (1-t)^{1-\frac{k}{\delta}} [(log(1-t) - logt)^2 + (log\lambda)^2$$

$$+ 2log\lambda(logt - log(1-t))]dt$$

$$= \frac{B(\beta+\frac{k}{\delta}+1, 2-\frac{k}{\delta})}{\lambda\delta^2 B(\beta+\frac{k}{\delta}, 1-\frac{k}{\delta})} \left[\left(\psi\left(\beta+\frac{k}{\delta}+1\right) - \psi\left(2-\frac{k}{\delta}\right)\right)^2 + \psi'\left(\beta+\frac{k}{\delta}+1\right)\right.$$

$$\left. + \psi'\left(2-\frac{k}{\delta}\right) + (log\lambda)^2 + 2(log\lambda)\left(\psi\left(\beta+\frac{k}{\delta}+1\right) - \psi\left(2-\frac{k}{\delta}\right)\right)\right]$$

By using E_1 to E_5, we obtain Fisher information matrix (FIM) for the weighted Dagum distribution with the entries:

$$I_{\beta\beta} = -\psi'(\beta+1) + \psi'\left(\beta+\frac{k}{\delta}\right), \tag{4.11}$$

$$I_{\beta\lambda} = E_1, \tag{4.12}$$

$$I_{\beta\delta} = -\frac{k}{\delta^2}\psi'(\beta+\frac{k}{\delta}) - \lambda E_3, \tag{4.13}$$

$$I_{\lambda\lambda} = -\frac{1}{\lambda^2}(\frac{k}{\delta}-1) - (\beta+1)E_2, \tag{4.14}$$

$$I_{\lambda\delta} = -\frac{k}{\lambda\delta^2} - (\beta+1)E_4, \tag{4.15}$$

and

$$I_{\delta\delta} = \tfrac{2k}{\delta^3}\left(\psi(\beta+\tfrac{k}{\delta})-\psi(1-\tfrac{k}{\delta})\right)+\tfrac{k^2}{\delta^4}\left(\psi'(\beta+\tfrac{k}{\delta})-\psi'(1-\tfrac{k}{\delta})\right)+\tfrac{1}{\delta^2} + \tfrac{2k}{\delta^3}log(\lambda)+\lambda(\beta+1)E_5. \quad (4.16)$$

Fisher information matrix (FIM) of weighted Dagum is given by:

$$I(\beta,\lambda,\delta)=\begin{bmatrix} I_{\beta\beta} & I_{\lambda\beta} & I_{\delta\beta} \\ I_{\beta\lambda} & I_{\lambda\lambda} & I_{\delta\lambda} \\ I_{\beta\delta} & I_{\lambda\beta} & I_{\delta\delta} \end{bmatrix}.$$

4.4 Parameter Estimation from Censored Data for Dagum Model

4.4.1 Maximum likelihood estimators

Khedhairi(2007) used the maximum likelihood method to estimate the parameters of the generalized Rayleigh distribution under censoring, we use similar method to obtain the estimates of the parameters of the Dagum distribution in this chapter.

Let $t = (t_1, \cdots, t_k)$, where $t_1 < \cdots < t_k$ denotes the predetermined inspection times and t_k is the completion time of the test, n_i is the number of failures recorded in the time interval $(t_{i-1}, t_i]$, n_{k+1} is the number of units which have not failed by the end of the test, let $t_0 = 0$ and $t_{k+1} = \infty$. Therefore, the likelihood function is given by:

$$L(t;\theta) = C\prod_{i=1}^{k}[P(t_{i-1} < T \leq t_i)]^{n_i}[P(T > t_k)]^{n_{k+1}},$$

where

$$P(t_{i-1} < T \leq t_i) = F(t_i) - F(t_{i-1}),$$

and

$$P(T > t_k) = 1 - F(t_k).$$

Therefore, based on Dagum distribution, we obtain:

$$\begin{aligned} L(t;\theta) &= C[1 - F(t_k)]^{n_{k+1}}\prod_{i=1}^{k}[F(t_i) - F(t_{i-1})]^{n_i} \\ &= C[1 - (1 + \lambda t_k^{-\delta})^{-\beta}]^{n_{k+1}}\prod_{i=1}^{k}[(1 + \lambda t_i^{-\delta})^{-\beta} - (1 + \lambda t_{i-1}^{-\delta})^{-\beta}]^{n_i}. \end{aligned}$$

We let

$$D_i(\lambda, \delta, \beta) = (1 + \lambda t_i^{-\delta})^{-\beta} - (1 + \lambda t_{i-1}^{-\delta})^{-\beta},$$

So that the log-likelihood function can be written as:

$$lnL = lnC + \sum_{i=1}^{k+1} n_i lnD_i(\lambda, \delta, \beta)$$

The partial derivative of $D_i(\lambda, \delta, \beta)$ with respect to λ is given by:

$$\begin{aligned} D_\lambda^{(i)} &= \frac{\partial D_i(\lambda, \delta, \beta)}{\partial \lambda} \\ &= -\beta[(1 + \lambda t_i^{-\delta})^{-\beta-1}t_i^{-\delta} - (1 + \lambda t_{i-1}^{-\delta})^{-\beta-1}t_{i-1}^{-\delta}], \end{aligned}$$

Similarly, we can obtain the partial derivative of $D_i(\lambda, \delta, \beta)$ with respect to β and δ:

$$D_\beta^{(i)} = (1 + \lambda t_{i-1}^{-\delta})^{-\beta} ln(1 + \lambda t_{i-1}^{-\delta}) - (1 + \lambda t_i^{-\delta})^{-\beta} ln(1 + \lambda t_i^{-\delta}),$$

$$D_{\delta}^{(i)} = \beta\lambda[(1+\lambda t_i^{-\delta})^{-\beta-1} ln(t_i) - (1+\lambda t_{i-1}^{-\delta})^{-\beta-1} ln(t_{i-1})],$$

where $i = 1, 2, \cdots, k+1$.

So the partial derivatives of log-likelihood function are given by:

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{k+1} n_i \frac{D_{\lambda}^{(i)}}{D_i(\lambda, \beta, \delta)}. \tag{4.17}$$

$$\frac{\partial lnL}{\partial \beta} = \sum_{i=1}^{k+1} n_i \frac{D_{\beta}^{(i)}}{D_i(\lambda, \beta, \delta)}. \tag{4.18}$$

$$\frac{\partial lnL}{\partial \delta} = \sum_{i=1}^{k+1} n_i \frac{D_{\delta}^{(i)}}{D_i(\lambda, \beta, \delta)}. \tag{4.19}$$

The MLE of λ, β and δ are obtained by solving $\frac{\partial lnL}{\partial \lambda} = \frac{\partial lnL}{\partial \beta} = \frac{\partial lnL}{\partial \delta} = 0$. There is no close form solution, so numerical technique must be applied.

4.4.2 Asymptotic Confidence Interval

Let $\theta = (\lambda, \beta, \delta)$, by using the large sample approximation, the MLE estimators of θ is approximately normal with mean θ and variance-covariance matrix I^{-1}. The elements of the 3×3 matrix I^{-1} can be approximated by the elements of information matrix, and $I_{ij}(\theta) = -E\left[\frac{\partial^2 lnL}{\partial\theta_i \partial\theta_j}\right] \approx -\frac{\partial^2 lnL}{\partial\theta_i \partial\theta_j}$. Also,

$$D_{\lambda\lambda}^{i} = \beta(\beta+1)(1+\lambda t_i^{-\delta})^{-\beta-2} t_i^{-2\delta} - \beta(\beta+1)(1+\lambda t_{i-1}^{-\delta})^{-\beta-2} t_{i-1}^{-2\delta}.$$

$$D_{\lambda\beta}^{i} = (1+\lambda t_i^{-\delta})^{-\beta-1} t_i^{-\delta}[\beta ln(1+\lambda t_i^{-\delta}) - 1] - (1+\lambda t_{i-1}^{-\delta})^{-\beta-1} t_{i-1}^{-\delta}[\beta ln(1+\lambda t_{i-1}^{-\delta}) - 1].$$

$$D^i_{\lambda\delta} = -\beta t_i^{-\delta}(1+\lambda t_i^{-\delta})^{-\beta-2} ln(t_i)(\beta\lambda t_i^{-\delta}-1) - \beta t_{i-1}^{-\delta}(1+\lambda t_{i-1}^{-\delta})^{-\beta-2} ln(t_{i-1})(\beta\lambda t_{i-1}^{-\delta}-1).$$

$$D^i_{\beta\beta} = (1+\lambda t_i^{-\delta})^{-\beta}[ln(1+\lambda t_i^{-\delta})]^2 - (1+\lambda t_{i-1}^{-\delta})^{-\beta}[ln(1+\lambda t_{i-1}^{-\delta})]^2.$$

$$D^i_{\beta\delta} = \lambda t_i^{-\delta}(1+\lambda t_i^{-\delta})^{-\beta-1} ln(t_i)[1-\beta ln(1+t_i^{-\delta})] - \lambda t_{i-1}^{-\delta}(1+\lambda t_{i-1}^{-\delta})^{-\beta-1} ln(t_{i-1})[1-\beta ln(1+t_{i-1}^{-\delta})].$$

$$D^i_{\delta\delta} = \beta\lambda t_i^{-\delta}(1+\lambda t_i^{-\delta})^{-\beta-2}(ln(t_i))^2(\lambda\beta t_i^{-\delta}-1) - \beta\lambda t_{i-1}^{-\delta}(1+\lambda t_{i-1}^{-\delta})^{-\beta-2}(ln(t_{i-1}))^2(\lambda\beta t_{i-1}^{-\delta}-1).$$

The elements of the matrix can be approximated by:

$$I_{11} = -\frac{\partial^2 lnL}{\partial\lambda^2} = -\sum_{i=1}^{k+1} n_i \frac{D^i_{\lambda\lambda} D^i(\theta) - (D^i_\lambda(\theta))^2}{(D^i(\theta))^2}, \quad (4.20)$$

$$I_{12} = I_{21} = -\frac{\partial^2 lnL}{\partial\lambda\beta} = -\sum_{i=1}^{k+1} n_i \frac{D^i_{\lambda\beta} D^i(\theta) - D^i_\lambda(\theta) D^i_\beta(\theta)}{(D^i(\theta))^2}, \quad (4.21)$$

$$I_{13} = I_{31} = -\frac{\partial^2 lnL}{\partial\lambda\delta} = -\sum_{i=1}^{k+1} n_i \frac{D^i_{\lambda\delta} D^i(\theta) - D^i_\lambda(\theta) D^i_\delta(\theta)}{(D^i(\theta))^2}, \quad (4.22)$$

$$I_{22} = -\frac{\partial^2 lnL}{\partial\beta^2} = -\sum_{i=1}^{k+1} n_i \frac{D^i_{\beta\beta} D^i(\theta) - (D^i_\beta(\theta))^2}{(D^i(\theta))^2}, \quad (4.23)$$

$$I_{23} = I_{32} = -\frac{\partial^2 lnL}{\partial\beta\delta} = -\sum_{i=1}^{k+1} n_i \frac{D^i_{\beta\delta} D^i(\theta) - D^i_{\beta}(\theta) D^i_{\delta}(\theta)}{(D^i(\theta))^2}, \tag{4.24}$$

$$I_{33} = -\frac{\partial^2 lnL}{\partial\delta^2} = -\sum_{i=1}^{k+1} n_i \frac{D^i_{\delta\delta} D^i(\theta) - (D^i_{\delta}(\theta))^2}{(D^i(\theta))^2}. \tag{4.25}$$

Therefore, the approximate $100(1-\gamma)\%$ two-sided confidence intervals for λ, β, δ are given by:

$$\widehat{\lambda} \pm Z_{\frac{\gamma}{2}} \sqrt{I_{11}^{-1}(\widehat{\theta})}, \quad \widehat{\beta} \pm Z_{\frac{\gamma}{2}} \sqrt{I_{22}^{-1}(\widehat{\theta})}, \quad \widehat{\delta} \pm Z_{\frac{\gamma}{2}} \sqrt{I_{33}^{-1}(\widehat{\theta})}$$

respectively.

Where $Z_{\frac{\gamma}{2}}$ is the upper $\frac{\gamma}{2}^{th}$ percentile of a standard normal distribution.

4.4.3 Parameter Estimate Under Progressive Type-II Censoring

Progressive type-II censoring has been wildly used in survival analysis and industrial life testing. A progressive type-II censored sample can be denoted as t_i, r_i, where t_i is the i^{th} failure time and r_i is the number of surviving units which are randomly removed at time t_i. Both type-II censoring and complete sampling are particular cases of progressive type-II censored sample. Domma et al.(2011) provided the Fisher information matrix of Dagum distribution under type I right censored observations. In this section, we present the normal equations for Dagum distribution under Progressive type-II censoring.

Recall that Dagum pdf and cdf are given by:

$$f_D(t;\beta,\lambda,\delta) = \beta\lambda\delta t^{-\delta-1}(\lambda t^{-\delta}+1)^{-\beta-1}, \quad t>0, \beta,\lambda,\delta>0,$$

and

$$F_D(t;\beta,\lambda,\delta) = (1+\lambda t^{-\delta})^{-\beta} \quad t>0, \beta,\lambda,\delta>0,$$

respectively.

Under Progressive type-II censoring, the likelihood function is:

$$L(\beta,\lambda,\delta) \propto \prod_{i=1}^{n} f(t_i)[1-F(t_i)]^{r_i},$$

that is:

$$\begin{aligned} L(\beta,\lambda,\delta) &\propto \prod_{i=1}^{n} \beta\lambda\delta t_i^{\delta-1}(1+\lambda t_i^{-\delta})^{-\beta-1}[1-(1+\lambda t_i^{-\delta})^{-\beta}]^{r_i} \\ &= \beta^n\lambda^n\delta^n \prod_{i=1}^{n} t_i^{-\delta-1} \prod_{i=1}^{n}(1+\lambda t_i^{-\delta})^{-\beta-1} \prod_{i=1}^{n}[1-(1+\lambda t_i^{-\delta})^{-\beta}]^{r_i}. \end{aligned}$$

The log-likelihood function is:

$$lnL = n(ln\beta+ln\lambda+ln\delta)-(\delta+1)\sum_{i=1}^{n} ln(t_i)-(\beta+1)\sum_{i=1}^{n} ln(1+\lambda t_i^{-\delta})+\sum_{i=1}^{n} r_i ln[1-(1+\lambda t_i^{\delta})^{-\beta}]$$

The normal equations are:

$$\frac{\partial lnL}{\partial\beta} = \frac{n}{\beta} - \sum_{i=1}^{n} ln(1+\lambda t_i^{-\delta}) + \sum_{i=1}^{n} \frac{r_i(1+\lambda t_i^{-\delta})^{-\beta} ln(1+\lambda t_i^{-\delta})}{1-(1+\lambda t_i^{-\delta})^{-\beta}},$$

$$\frac{\partial lnL}{\partial\lambda} = \frac{n}{\lambda} - (\beta+1)\sum_{i=1}^{n} \frac{t_i^{-\delta}}{1+\lambda t_i^{-\delta}} + \sum_{i=1}^{n} \frac{r_i\beta(1+\lambda t_i^{-\delta})^{-\beta-1}t_i^{-\delta}}{1-(1+\lambda t_i^{-\delta})^{-\beta}},$$

and

$$\frac{\partial lnL}{\partial \delta} = \frac{n}{\delta} + (\beta+1)\sum_{i=1}^{n}\frac{\lambda t_i^{-\delta} ln(t_i)}{1+\lambda t_i^{-\delta}} - \sum_{i=1}^{n}\frac{r_i\beta(1+\lambda t_i^{-\delta})^{-\beta-1}(\lambda t_i^{-\delta})ln(t_i)}{1-(1+\lambda t_i^{-\delta})^{-\beta}}.$$

The normal equations should be solve numerically.

4.5 Notations

We used the following results in the computations of the results in this chapter.

$$\int_0^1 t^{a-1}(t-1)^{b-1}dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{4.26}$$

Differentiate (4.31) with respect to a, we have:

$$\begin{aligned} \int_0^1 t^{a-1}(t-1)^{b-1}lntdt &= \frac{\Gamma'(a)\Gamma(b)\Gamma(a+b)-\Gamma(a)\Gamma'(b)\Gamma(a+b)}{(\Gamma(a+b))^2} \\ &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}(\psi(a)-\psi(a+b)) \\ &= B(a,b)(\psi(a)-\psi(a+b)). \end{aligned} \tag{4.27}$$

Similarly, differentiate (4.31) with respect to b:

$$\int_0^1 t^{a-1}(t-1)^{b-1}ln(1-t)dt = B(a,b)(\psi(b)-\psi(a+b)). \tag{4.28}$$

Differentiate (4.32) with respect to a:

$$\int_0^1 t^{a-1}(t-1)^{b-1}(lnt)^2dt = B(a,b)[(\psi(a)-\psi(a+b))^2+\psi'(a)-\psi'(a+b)]. \tag{4.29}$$

Differentiate (4.33) with respect to b:

$$\int_0^1 t^{a-1}(t-1)^{b-1}(ln(1-t))^2dt = B(a,b)[(\psi(b)-\psi(a+b))^2+\psi'(b)-\psi'(a+b)]. \tag{4.30}$$

Differentiate (4.32) with respect to b:

$$\int_0^1 t^{a-1}(t-1)^{b-1}(lnt)(ln(1-t))dt = B(a,b)[(\psi(a)-\psi(a+b))(\psi(b)-\psi(a+b))-\psi'(a+b)]. \tag{4.31}$$

From (4.34)-(4.36) we obtain:

$$\int_0^1 t^{a-1}(t-1)^{b-1}(lnt-ln(1-t))^2dt = B(a,b)[(\psi(a)-\psi(b))^2+\psi'(a)+\psi'(b)]. \tag{4.32}$$

4.6 Concluding Remarks

In this chapter, properties of Dagum and Weighted Dagum distributions including hazard function, reverse hazard function, moments are presented. Fisher information matrix (FIM) and estimates of model parameters under censoring including progressive Type II for the Dagum distribution are presented.

CHAPTER 5

ESTIMATION OF PARAMETERS IN THE WEIGHTED GENERALIZED BETA DISTRIBUTION OF THE SECOND KIND

5.1 Estimation of parameters

The WGB2 with weight function $w(y) = y^k$, can be written as:

$$g_w(y; a, b, p, q) = B^{-1}\left(p + \frac{k}{a}, q - \frac{k}{a}\right)\left(\frac{a}{b}\right)\left[\left(\frac{y}{b}\right)^a\right]^{p+\frac{k}{a}}\left[1 + \left(\frac{y}{b}\right)^a\right]^{-(p+q)}.$$

If we set $F(y) = 1 - (1 + (\frac{y}{b})^a)^{-1}$, then $f(y) = \frac{a}{b}(\frac{y}{b})^{a-1}[1 + (\frac{y}{b})^a]^{-2}$ and

$$g_w(y; a, b, p, q) = B^{-1}\left(p + \frac{k}{a}, q - \frac{k}{a}\right) f(y)[F(y)]^{p+\frac{k}{a}-1}[1 - F(y)]^{q-\frac{k}{a}-1}. \quad (5.1)$$

Clearly, this distribution belongs to the $beta - F$ class of distributions with

$$F(y) = 1 - \left[1 + \left(\frac{y}{b}\right)^a\right]^{-1} = \frac{y^a}{b^a + y^a}. \quad (5.2)$$

Let $\theta = (a, b, p, q)^T$ be column vectors of parameters associated with the income distribution. The income distribution is given by:

$$P_i(\theta) = B^{-1}\left(p + \frac{k}{a}, q - \frac{k}{a}\right)\int_{F(y_{i-1})}^{F(y_i)} t^{p+\frac{k}{a}-1}(1 - t)^{q-\frac{k}{a}-1}dt, \quad (5.3)$$

where P_i denotes the estimated proportion of the population in the i^{th} interval of the r income groups defined by the interval $I_i = [y_{i-1}, y_i]$. The multinomial likelihood function for the data is given by:

$$N! \prod_{i=1}^{r} \frac{[P_i(\theta)]^{n_i}}{n_i!},$$

where $n_i, i = 1,...r$ denotes the observed frequency in the i^{th} group and $N = \sum_{i=1}^{r} n_i$. We maximize:

$$L(\theta) = \sum_{i=1}^{r} n_i ln P_i(\theta),$$

where $P_i(\theta) = \int_{F(y_{i-1})}^{F(y_i)} h(t)dt$, and $h(t) = B^{-1}(p+\frac{k}{a}, q-\frac{k}{a})t^{p+\frac{k}{a}-1}(1-t)^{q-\frac{k}{a}-1}$.

Sepanski and Lingji (2007) points out that obtain P_i by computing the cdf of a beta random at $F(y_{i-1})$ and $F(y_i)$ can reduces the complexity of programming required to calculate the integrations.

The first derivative with respect to $\theta = (a, b, p, q)^T$ are:

$$\frac{dL(\theta)}{d\theta} = \sum_{i=1}^{r} \frac{n_i}{P_i(\theta)} \cdot \frac{dP_i(\theta)}{d\theta} \tag{5.4}$$

The partial derivative equations of $P_i(\theta)$ with respect to a,b,p,q are given by:

$$\frac{\partial P_i(\theta)}{\partial a} = h(F(y_i))\frac{b^a y_i^a(lny_i - lnb)}{(b^a + y_i^a)^2} - h(F(y_{i-1}))\frac{b^a y_{i-1}^a(lny_{i-1} - lnb)}{(b^a + y_{i-1}^a)^2}$$
$$+ \int_{F(y_{i-1})}^{F(y_i)} \frac{kh(t)}{a^2}\left[\Psi\left(p + \frac{k}{a}\right) - \Psi\left(q - \frac{k}{a}\right) + ln\frac{1-t}{t}\right]dt, \tag{5.5}$$

$$\frac{\partial P_i(\theta)}{\partial b} = -ab^{a-1}\left[\frac{h(F(y_i))y_i^a}{(b^a + y_i^a)^2} - \frac{h(F(y_{i-1}))y_{i-1}^a}{(b^a + y_{i-1}^a)^2}\right], \tag{5.6}$$

$$\frac{\partial P_i(\theta)}{\partial p} = \int_{F(y_{i-1})}^{F(y_i)} h(t)\left[-\Psi\left(p + \frac{k}{a}\right) + \Psi(p+q) + lnt\right]dt, \tag{5.7}$$

and

$$\frac{\partial P_i(\theta)}{\partial q} = \int_{F(y_{i-1})}^{F(y_i)} h(t)\left[-\Psi\left(q-\frac{k}{a}\right)+\Psi(p+q)+ln(1-t)\right]dt, \qquad (5.8)$$

where $\Psi(x) = \frac{d}{dx}[\Gamma(x)]$. Using the equations (5.5)-(5.8) in equation (5.4) we can obtain the gradient functions of $L(\theta)$ with respect to parameters a, b, p, q.

The partial derivative equation (5.5) exists when $k > 0$. If $k = 0$, the partial derivative equation of P_i with respect to a is given by:

$$\frac{\partial P_i(\theta)}{\partial a} = b^a\left[\frac{h(F(y_i))(y_i)^a(ln(y_i)-ln(b))}{(b^a+y_i^a)^2} - \frac{h(F(y_{i-1}))(y_{i-1})^a(ln(y_{i-1})-ln(b))}{(b^a+y_{i-1}^a)^2}\right] \qquad (5.9)$$

5.2 Applications

In this section we obtain parameter estimates based on our previous discussions and results. WGB2 was fitted to U.S. family nominal income for 2001-2009[1]. The groups consist of families whose income are in the corresponding income interval $I_i = [y_{i-1}, y_i)$, the n_i/N are the observed relative frequencies ($N = \sum n_i$).

[1]The data were taken from the Census Population Report

Table 5.1: U.S. family nominal income for 2001-2009

$[y_{i-1}, y_i)$	observed relative frequencies n_i/N								
(thousand)	2001	2002	2003	2004	2005	2006	2007	2008	2009
[0,15)	13	13.4	12.9	12.6	13	13.3	13.2	12.9	12.4
[15,25)	11.9	12	11.3	11.2	11.5	11.6	11.6	11.4	11.4
[25,35)	11.1	11	10.5	11.1	10.8	11	10.9	10.6	10.5
[35,50)	14.1	14.1	14	14.1	14.2	14.1	14	14.5	14.8
[50,75)	18.1	17.6	18	18.2	18.1	18.1	17.7	18	17.9
[75,100)	11.5	11.9	12	11.6	12.1	12	12.2	12.5	12.6
[100,150)	11.9	11.9	12.7	12.5	12	11.9	12.3	12.3	12.2
[150,200)	4.4	4.3	4.7	4.7	4.3	4.4	4.4	4.2	4.3
[200,∞)	3.8	3.7	4	4	4	3.6	3.7	3.7	3.9

The common way to obtain estimators is to maximize the multinomial likelihood function. Since the likelihood function is nonlinear and complicated, we use MATLAB to search for the maximum value of multinomial likelihood function.[2] The results of this estimation for 2001, 2005, and 2009 are reported in Tables 5.2 - 5.4.

[2]By setting different initial values and use 'fminsearchbnd' to search for the maximum loglikelihood values

Table 5.2: Estimated parameters of WGB2 for income distribution (2001)

	k=0	k=1	k=2	k=3	k=4
a	1.403	1.405	0.669	0.4487	0.3376
b	16.66	16.488	11408.457	5463.111	15213.5
p	0.999	0.288	0.000001	0.001681	0.037333
q	3.963	4.629	470.178	154.668	187.036
sse*10000	1.208234	1.208159	2.509558	8.463283	13.220825

Table 5.3: Estimated parameters of WGB2 for income distribution (2005)

	k=0	k=1	k=2	k=3	k=4
a	1.423	1.383	0.646	0.432	0.339
b	15.502	16.621	12800.159	6595.86	4379.966
p	0.961	0.271989	0.000003	0.004804	0.00976
q	3.582	4.648	446.0185	156.662	127.893
sse*10000	0.520372	0.518495	2.109663	8.049098	13.146924

Table 5.4: Estimated parameters of WGB2 for income distribution (2009)

	k=0	k=1	k=2	k=3	k=4
a	1.090	1.092	0.654	0.440	0.341
b	26.596	26.695	2324.550	3550.805	3964.261
p	1.413	0.493	0.000005	0.001079	0.002585
q	7.209	8.148	157.039	125.837	124.988
sse*10000	0.471333	0.467639	1.078739	6.085883	10.444168

Based on the sse value we can conclude that: the length-biased WGB2 (k=1) provides a better fit than GB2 (k=0) and other WGB2 (k=2,3,4). If we plug the estimated parameters in the partial derivative equations in Section 4, we can obtain the values of these partial derivative equations in Table 5.5 - 5.7. From the tables below we find that these equations are close to zero or very small, this means that the estimated parameters that we obtained are effective.

Table 5.5: Values of partial derivative equations of WGB2 (2009)

	k=0	k=1	k=2	k=3	k=4
$\frac{\partial L(\theta)}{\partial a}$	0.0255	-0.0029	0.0397	-0.1732	-0.1347
$\frac{\partial L(\theta)}{\partial b}$	-0.0017	-0.0001	0	0	0
$\frac{\partial L(\theta)}{\partial p}$	-0.0105	-0.0048	-0.0791	-0.1077	-0.0625
$\frac{\partial L(\theta)}{\partial q}$	0.007	-0.0005	0	0.0004	0

Table 5.6: Values of partial derivative equations of WGB2 for (2005)

	k=0	k=1	k=2	k=3	k=4
$\frac{\partial L(\theta)}{\partial a}$	-0.0069	-0.0341	0.0028	-0.0922	0.0033
$\frac{\partial L(\theta)}{\partial b}$	-0.0021	-0.0017	0	0	0
$\frac{\partial L(\theta)}{\partial p}$	0.0118	0.0422	-0.1075	-0.097	-0.0638
$\frac{\partial L(\theta)}{\partial q}$	-0.0007	-0.0058	0.0002	0.001	-0.0001

Table 5.7: Values of partial derivative equations of WGB2 (2001)

	k=0	k=1	k=2	k=3	k=4
$\frac{\partial L(\theta)}{\partial a}$	-0.0283	-0.0502	-0.2614	-0.6607	3.2358
$\frac{\partial L(\theta)}{\partial b}$	-0.0031	-0.0123	0	0	0
$\frac{\partial L(\theta)}{\partial p}$	0.0173	-0.1533	-0.1243	-0.1381	-0.0055
$\frac{\partial L(\theta)}{\partial q}$	-0.0009	0.0327	0.0003	0.0011	-0.0028

Since we already have some results on WGB2 in Section 3, and we also found out that the length-biased WGB2 provides best fit to income distribution, we can apply the estimated parameters from length-biased WGB2 model to obtain the estimates of mean, variance, coefficient of variation, skewness and kurtosis, bottom sensitive index, top-sensitive index, MLD index and Theil index. The results are presented in Table 5.8.

Table 5.8: Estimated statistics for income distribution (k=1 in WGB2)

Year	Est.mean	Est.Var	Est.CV	Est.CS	Est.CK
2001	6.759468	38.091976	0.913070	-4.242920	24.013806
2005	6.719345	39.169450	0.931423	-4.114463	25.864490
2009	6.629841	37.111483	0.918864	-4.195311	16.203708
Year	Est.I(-1)	Est.I(2)	Est.MLE	Est.Theil	
2001	1.150014	0.416849	0.396929	1.758663	
2005	1.252677	0.433774	0.410256	1.789568	
2009	1.033116	0.422155	0.402408	3.567591	

5.3 Concluding Remarks

In Chapter 5, the weighted generalized beta distribution of the second kind (WGB2) was fitted to U.S. family income data (2001-2009). The maximum likelihood estimation (MLE) is used for estimating the parameters of the income distribution model. The results showed that the length-biased WGB2 provides the best relative fit to income data. Based on previously obtained descriptive measures for WGB2, we estimate the mean, variance, coefficient of variation, coefficient of skewness, coefficient of kurtosis, bottom-sensitive index, top-sensitive index, MLD index and Theil index for the income data.

BIBLIOGRAPHY

[1] Cox, D. R., Renewal Theory, Barnes & Noble, New York, 1962.

[2] Dagum, C., *A New Model of Personal Income Distribution: Specification and Estimation,* Economie Appliquee, 30, 413 - 437, (1977).

[3] Ding-Geng, C. and Yuhlong, L., *A Note on the Maximum Likelihood Estimation for the Generalized Gamma Distribution Parameters under Progressive Type-II Censoring,* International Journal of Intelligent Technology and Applied Statistics, **2**(2), 57 - 64, (2009).

[4] Domma, F., Giordano, S. and Zenga, M., *The Fisher Information Matrix in Right Censored Data from the Dagum Distribution,* Working Paper, (2011).

[5] Esteban, J., *Income-Share Elasticity and the Size Distribution of Income,* International Economic Review, **27**(2), 439 - 444, (1986).

[6] Fisher, R.A., *The Effects of Methods of Ascertainment upon the Estimation of Frequencies,* Annals of Human Genetics, **6**(1), 439 - 444, (1934).

[7] Gupta, C. and Keating, P., *Relations for Reliability Measures Under Length Biased Sampling,* Scan. J. Statist, **13**(1), 49 - 56, (1985).

[8] Jenkins, P., *Inequality and GB2 Income Distribution,* IZA Discussion Paper, **2831**, (2007).

[9] Kleiber, C., *A Guide to the Dagum Distributions*, Springer, New York, (2007).

[10] Kleiber, C. and Kotz, S., *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley, New York, (2003).

[11] Khedhairi, A.A., Sarhan, A. and Tadj, L., *Estimation of the Generalized Rayleigh Distribution Parameters,* King Saud University, (2007).

[12] McDonald, B., *Some Generalized Functions for the Size Distribution of Income,* Econometrica, **52**(3), 647-663, (1984).

[13] McDonald, B. and Xu, J., *A Generalization of the Beta Distribution with Application,* Journal of Econometrics, **69**(2), 133 - 152, (1995).

[14] Nanda, K. and Jain, K., *Some Weighted Distribution Results on Univariate and Bivariate Cases,* Journal of Statistical Planning and Inference, **77**(2), 169 - 180, (1999).

[15] Oluyede, O., *On Inequalities and Selection of Experiments for Length-Biased Distributions,* JProbability in the Engineering and Informational Sciences, **13**(2), 129 - 145, (1999).

[16] Patil, P., *Encountered Data, Statistical Ecology, Environmental Statistics, and Weighted Distribution Methods,* Environmetrics, **2**(4), 377 - 423, (1991).

[17] Patil, P. and Rao, R., *Weighted Distributions and Size-Biased Sampling with Applications to Wildlife and Human Families,* Biometrics, **34**(6), 179 - 189, (1978).

[18] Rao, R., *On Discrete Distributions Arising out of Methods of Ascertainment,* The Indian Journal of Statistics, **27**(2), 320 - 332, (1965).

[19] Renyi, A., *On Measures of Entropy and Information,* Berkeley Symposium on Mathematical Statistics and Probability, **1**(1), 547 - 561, (1960).

[20] Shannon, E., *A Mathematical Theory of Communication,* The Bell System Technical Journal, **27**(10), 379 - 423, (1948).

[21] Shannon, E., *A Mathematical Theory of Communication,* The Bell System Technical Journal, **27**(10), 623 - 656, (1948).

[22] Sepanski, H. and Lingji, *K., A Family of Generalized Beta Distribution For Income,* International Business , **1**(10), 129 - 145, (2007).

[23] Zelen, M. and Feinleib, M., *On the theory of screening for chronic diseases,* Biometrika, **56**(3), 601 - 614, (1969).

[24] Zelen, M., *Problems in Cell Kinetics and Early Detection of Diseae,* Reliability and Biometry, **56**(3), 701 - 726, (1974)